FUNDAMENTALS OF BIOCHEMISTRY III

Henry Jakubowski and Patricia Flatt
LibreTexts

SECTION OVERVIEW

Fundamentals of Biochemistry (Jakubowski and Flatt)

This text is designed for the two-semester introductory biochemistry course but could be customized for a one-semester course as well. Biochemistry is the application of chemistry to the study of biological processes at the cellular and molecular level. It emerged as a distinct discipline around the beginning of the 20th century when scientists combined chemistry, physiology, and biology to investigate the chemistry of living systems. Biochemistry is both a life science and a chemical science - it explores the chemistry of living organisms and the molecular basis for the changes occurring in living cells. It uses the methods of chemistry, physics, molecular biology, and immunology to study the structure and behavior of the complex molecules found in biological material and the ways these molecules interact to form cells, tissues, and whole organisms. The chapter organization generally follows the traditionally used biochemistry texts to allow easier adoption by faculty and use by students.

New! Chapter32 - Biochemistry and Climate Change (1/14/23)

New! Chapter 4.13 - Predicting Structure and Function of Biomolecules Through Natural Language Processing Tools (11/1/23)

Fundamentals of Biochemistry Vol. I - Structure and Catalysis

Licensing
Preface from the Authors
Biochemistry Foundation Concepts and Learning Goals and Objectives
1: The Foundations of Biochemistry
1.1: Cellular Foundations
1.2: Chemical Foundations
1.3: Physical-Chemical Foundations
1.4: Genetic Foundations
1.5: Chapter 1 Questions
2: Water and its Role in Life
2.1: The multiple roles of water
2.2: Weak Acids and Bases, pH and pKa
2.3: Buffering against pH Changes in Biological Systems
2.4: Solubility in an aqueous world - noncovalent interactions in depth
2.5: Solubility in an aqueous world - The Hydrophobic Effect
2.6: Chapter 2 Questions
3: Amino Acids, Peptides, and Proteins
3.1: Amino Acids and Peptides
3.3: Proteins - Analyses and Structural Predictions of Protein Structure
3.4: Protein Purification
3.5: Antibodies in Quantitation and In Vivo Detection
3.2: The Structure of Proteins- An Overview
3.6: Chapter 3 Questions
4: The Three-Dimensional Structure of Proteins
4.1: Main Chain Conformations
4.2: Secondary Structure and Loops
4.3: Tertiary and Quaternary Structures
4.4: Secondary Structural Motifs and Domains
4.5: Protein with Alpha, Alpha-Beta, Beta and Little Secondary Structure
4.6: Intrinsically Disordered Proteins

4.7: Fibrillar Proteins
4.8: Protein Folding and Unfolding (Denaturation) - Dynamics
4.9: Protein Stability - Thermodynamics
4.10: Protein Aggregates - Amyloids, Prions and Intracellular Granules
4.11: Biomolecular Visualization - Conceptions and Misconceptions
4.12: Laboratory Determination of the Thermodynamic Parameters for Protein Denaturation
4.13: Predicting Structure and Function of Biomolecules Through Natural Language Processing Tools
4.14: Chapter 4 Questions
5: Protein Function
5.1: Binding - The First Step Towards Protein Function
5.2: Techniques to Measure Binding
5.3: Oxygen-Binding Proteins and Allosterism
5.4: Complementary Interactions between Proteins and Ligands - The Immune System and Immunoglobulins
5.5: Protein Interactions Modulated by Chemical Energy- Actin, Myosin, and Molecular Motors
5.6: Binding - Conformational Selections and Intrinsically Disordered Proteins
5.7: Binding - Enzyme Linked Immunosorbant Assays (ELISAs)
6: Enzyme Activity
6.1: How Enzymes Work
6.2: Kinetics without Enzymes
6.3: Kinetics with Enzymes
6.4: Enzyme Inhibition
6.5: Enzymatic Reaction Mechanisms
6.6: Enzymes and Protein Regulation
6.7: Ribozymes - RNA Enzymes
6.8: Cofactors and Catalysis - A Little Help From My Friends
7: Carbohydrates and Glycobiology
7.1: Monosaccharides and Disaccharides
7.2: Polysaccharides
7.3: Glycoconjugates - Proteoglycans, Glycoproteins, Glycolipids and Cell Walls
7.4: The Sugar Code and Lectin Decoding
7.5: Working with Carbohydrates
7.6: Chapter 7 Problems - Answer Key
7.7: Chapter 7 Problems
8: Nucleotides and Nucleic Acids
8.1: Nucleic Acids - Structure and Function
8.2: Nucleic Acids - RNA Structure and Function
8.3: Nucleic Acids - Comparison of DNA and RNA
8.4: Chromosomes and Chromatin
8.5: References
8.6: Enzymes for Genetic modifications
9: Investigating DNA
9.1: DNA Isolation, Sequencing, and Synthesis
9.2: Bioinformatics
9.3: Cloning and Recombinant Expression
9.4: DNA Microarrays
9.5: In Situ Hybridization

9.6: References
10: Lipids
10.1: Introduction to lipids
10.2: Lipids Aggregates in Water - Micelles and Liposomes
10.3: Membrane Bilayer and Monolayer Assemblies - Structures and Dynamics
10.4: Working with Lipids
10.5: Problems
11: Biological Membranes and Transport
11.1: Membrane and Membrane Proteins
11.2: Diffusion Across a Membrane - Passive and Facilitated Diffusion
11.3: Diffusion Across a Membrane - Channels
11.4: Diffusion Across a Membrane - Pores
11.5: Active Transport
11.4: Diffusion Across a Membrane - Pores
Index
Analysis of Biomolecules Structure and Function - Problems
1.1: Problem Set 1
1.2: Problem Set 2
Glossary
Detailed Licensing

Fundamentals of Biochemistry Vol. II - Bioenergetics and Metabolism

Licensing
About the Authors
12: Bioenergetics and Biochemical Reaction Types
12.1: Biochemical Reactions and Energy Changes
12.2: Phosphoryl Group Transfers and ATP
12.3: The Chemistry and Biochemistry of Dioxygen
12.4: Biological Oxidation-Reduction Reactions
13: Glycolysis, Gluconeogenesis, and the Pentose Phosphate Pathway
13.1: Glycolysis
13.2: Fates of Pyruvate under Anaerobic Conditions- Fermentation
13.3: Gluconeogenesis
13.4: Pentose Phosphate Pathway of Glucose Oxidation
14: Principles of Metabolic Regulation
14.1: Regulation of Metabolic Pathways
14.2: Basic Principles of Metabolic Control Analysis (MCA)
14.3: The Flux Control Coefficient
14.4: Concentration Control and Elasticity Coefficients
14.5: Metabolism and Signaling: The Steady State, Adaptation and Homeostasis
15: Glucose, Glycogen, and Their Metabolic Regulation
15.1: Insulin Signaling in the Liver
15.2: Glycogenesis
15.3: Glycogenolyis and its Regulation by Glucagon and Epinephrine Signaling
15.4: Regulation of Glycolysis

15.5: Regulation of Gluconeogenesis
16: The Citric Acid Cycle
16.1: Production of Acetyl-CoA (Activated Acetate)
16.2: Reactions of the Citric Acid Cycle
16.3: Regulation of the Citric Acid Cycle
16.4: Variants of the Citric Acid Cycle
17: Fatty Acid Catabolism
17.1: Digestion, Mobilization, and Transport of Fats
17.2: Oxidation of Fatty Acids
17.3: Ketone Bodies
18: Nitrogen - Amino Acid Catabolism
18.1: The Biochemistry of Nitrogen in the Biosphere
18.2: Metabolic Fates of Amino Groups
18.3: Nitrogen Excretion and the Urea Cycle
18.4: An overview of amino acid metabolism and the role of Cofactors
18.5: Pathways of Amino Acid Degradation
19: Oxidative Phosphorylation
19.1: Electron-Transfer Reactions in Mitochondria
19.2: ATP Synthesis
19.3: Regulation of Oxidative Phosphorylation
20: Photosynthesis and Carbohydrate Synthesis in Plants
20.1: Light Absorption in Photosynthesis - An Overview
20.2: The Kok Cycle and Oxygen Evolving Complex of Photosystem II
20.3: Plant Electron Transport and ATP Synthesis
20.4: CO_2 uptake - Calvin Cycle and C3 organisms
20.5: CO2 uptake - C4 and CAM Pathways
20.6: Biosynthesis of Starch, Sucrose and Cellulose
21: Lipid Biosynthesis
21.1: Biosynthesis of Fatty Acids and Eicosanoids
21.2: Biosynthesis of Triacylglycerols
21.3: Biosynthesis of Membrane Glycerolipids
21.4: Biosynthesis of Membrane Sphingolipids
21.5: Biosynthesis of Cholesterol and Steroids
21.6: Biosynthesis of Isoprenoids
22: Biosynthesis of Amino Acids, Nucleotides, and Related Molecules
22.1: Overview of Nitrogen Metabolism
22.2: Biosynthesis of Amino Acids
22.3: Molecules Derived from Amino Acids
22.4: Biosynthesis and Degradation of Nucleotides
Index
Glossary
Detailed Licensing

Fundamentals of Biochemistry Vol. III - Information Pathways

- Licensing
- About the Authors
- 23: Chromosome Structure
  - 23.1: Gene Mapping and Chromosomal Karyotypes
  - 23.2: DNA Transposable Elements
  - 23.3: Chromosome Packaging
- 24: DNA Metabolism
  - 24.1: DNA Replication
  - 24.2: DNA Mutations, Damage, and Repair
  - 24.3: DNA Recombination
- 25: RNA Metabolism
  - 25.1: DNA-Dependent Synthesis of RNA
  - 25.2: RNA Processing
  - 25.3: RNA-Dependent Synthesis of RNA and DNA
  - 25.4: 26.4 References
- 26: Protein Metabolism
  - 26.1: The Genetic Code
  - 26.2: Protein Synthesis
  - 26.3: Translational Regulation and Protein Degradation
- 27: Regulation of Gene Expression
  - 27.1: Regulation of Gene Expression in Bacteria
  - 27.2: Regulation of Gene Expression in Eukaryotes
- Index
- Glossary
- Detailed Licensing

Fundamentals of Biochemistry Vol. V - Problems

- Licensing
- Binding and Noncovalent Interaction Problems
- Chapter 1 - Problems
- Enzyme Kinetics Problems
  - Carbonic Anydrase Inhibition
  - WNovak-Beta Jupyter NB - Ex 1: Under Construction
- Enzyme Mechanisms Questions
  - Carbonic Anhydrase - Mechanism
- Global Challenges with Literature-Based Guided Assessments
  - 1. Global Challenges: An Introduction
  - Carbon Capture Using Carbonic Anhydrase
  - Trauma and Health
- iCn3D for Biomolecular Visualization Learning Themes and Goals
  - Biomolviz- Constructing iCn3D Models to Target Biomolecular Visualization Skills
  - BioMolViz Framework
  - BioMolViz Theme: Alternate Renderings (AR)
  - BioMolViz Theme: Atomic Geometry (AG)

BioMolViz Theme: Construction and Annotation (CA)
BioMolViz Theme: Ligands and Modifications (LM)
BioMolViz Theme: Macromolecular Assemblies (MA)
BioMolViz Theme: Macromolecular Building Blocks (MB)
BioMolViz Theme: Molecular Dynamics (MD)
BioMolViz Theme: Molecular Interactions (MI)
BioMolViz Theme: Structural Model Skepticism (SK)
BioMolViz Theme: Structure-Function Relationship (SF)
BioMolViz Theme: Symmetry/ Asymmetry Recognition (SA)
BioMolViz Theme: Topology and Connectivity (TC)
iCn3D Molecular Modeling Tutorials
iCn3D Basics: File Types
iCn3D Basics: Saving Files
iCn3D Basics: Analysis Menu
iCn3D Basics: Mouse Commands
iCn3D Basics: Selecting and Viewing with a mouse
iCn3D Basics: Style and Color
iCn3D Intro Tutorial A: Modeling a Short Peptide in a Protein
iCn3D Intro Tutorial B: Rendering a Protein
iCn3D Intro Tutorial C: Finding Pockets in Proteins
iCn3D Intro Tutorial D: Modeling Psychoactive Drugs in Target Proteins
iCn3D Skill: Alternative Rendering and Saving Files
iCn3D Skill: Displays surface of a protein - Superoxide Dismutase
iCn3D Skill: Aligning two structures
iCn3D Skill: Analysis of Noncovalent Interactions
iCn3D Skill: Creating and Saving Selections
iCn3D Skill: Mutations
iCn3D Skill: Selection through Sequence and Annotations
iCn3D Skill: Showing a Protein-Protein Interface
iCn3D Tutorial: Binding interactions of SARS-Cov-2 Spike receptor domain with ACE2
iCn3D Tutorial: Overlay many ACE2 receptor binding domain analogous to SARS-Co2 RBD using Blast
iCn3D Tutorial: Protein Kinase B (AKT)
iCn3D Tutorial Question: Intermediate Problem - Cyclooxygenase II
Literature Based Guided Assessments (LGAs)
Voltage-Gated Sodium Channel - Students 082423
Membrane Structure, Diffusion and Transport Problems
Metabolism - Carbohydrate Problems
Metabolism - Lipids: Problems
test pGW
Molecular Modeling Questions
Signal Transduction Problems
NMDA Receptor - Under Construction
Ras Questions
Structure/Function - Carbohydrate Problems
Structure/Function - Lipid Problems
Structure/Function - Nucleic Acid Problems

Structure/Function - Protein Problems
Disulfide Bonds
LGA: Voltage-Gated Sodium Channel - Students 082423
LGA: Protein Stability - Carbonic Anhydase
Index
Glossary
Detailed Licensing

Fundamentals of Biochemistry Vol. IV - Signaling, Integration of Metabolism, and Special Topics

Licensing
About the Authors
28: Biosignaling - Capstone Volume I
28.1: General Features of Signal Transduction
28.2: At the cell membrane- receptors and receptor enzymes
28.3: The Next step - The Kinome and Activation of Kinases at the Cell Membrane
28.4: The next step - Downsteam intracellular signaling
28.5: Small G proteins, GAPs and GEFs
28.6: Phosphatases
28.7: Calcium Signaling
28.8: Receptor Guanylyl Cyclases, cGMP, and Protein Kinase G
28.9: Gated Ion Channels - Neural Signaling
28.10: Integrins- Bidirectional Cell Adhesion Receptors
28.11: Signaling by Steroid Hormones
28.12: mTOR and Nutrient Signaling
28.13: Regulation of the Cell Cycle by Protein Kinases
28.14: Programmed Cell Death
28.15: Signaling in Microorganisms
28.16: Signaling in Plants
28.17: Signal Transduction - Vision and Olfaction
28.18: Signal Transduction - Taste (Gustation)
28.19: Signal Transduction - Temperature
28.20: Signal Transduction - Pressure
29: Integration of Mammalian Metabolism - Capstone Volume II
29.1: Overview of Metabolism
29.2: Regulation of glucose metabolism from a liver-centric perspective
29.3: Emerging role of the brain in the homeostatic regulation of energy and glucose metabolism
29.4: Skeletal Muscle Regulates Metabolism
29.5: Intestinal Fructose and Glucose Metabolism in Health and Disease
29.6: Metabolic consequences of obesity and type 2 diabetes- Balancing genes and environment for personalized care
29.7: Lipid-Induced Mechanisms of Metabolic Syndrome
29.8: Fundamentals of cancer metabolism
29.9: Immunometabolism- Cellular Metabolism Turns Immune Regulator
29.10: Ethanol metabolism- The good, the bad, and the ugly
29.11: Gut microbiota-derived metabolites as central regulators in metabolic disorders
30: Abiotic Origins of Life
30.1: Abiotic Origins of Life

31: Quantum Biochemistry - TBA
31.1: Quantum Biochemistry
32: Biochemistry and Climate Change
32.01A: The Basics of Climate Change
32.01B: Back to the Present and Future of Climate Change
32.2: Use of Isotope Analysis in Measuring Climate Change
32.3: Climate Change - The Carbon Cycle and Carbon Chemistry
32.4: Biofuels A - Corn and Sugar Cane Ethanol
32.5: Biofuels B - Cellulosic Ethanol
32.6: Algae - an Introduction
32.7: Algae - Bioethanol production
32.8: Biodiesel, Syngas and Bioaviation fuels
32.09: Biohydrogen - An Introduction
32.10: Biohydrogen - Hydrogenases
32.11: A Warmer World: Temperature Effects On Chemical Reactions
32.12: A Warmer World: Temperature Effects On Proteins
32.13: Biochemistry, Climate Change and Human Health
32.14: Climate Change, Infectious Disease and Pandemics
32.15: Pandemic Diseases and Drug Discovery - Under Construction
32.16: Fixing Carbon Fixation
32.17: Fixing Nitrogen Fixation
32.18: Turning Trees into Plexiglass - Synthetic Biology For Production of Green Foods and Products
Chapter 33: Your Contribution - Sandbox
Arthur Sikora Lipid rafts
BioMolViz Framework
Chetna IMF Test
Edmund - Structural Basis of Allostery, The Kinase Model
Emily Schmitt Sepiapterin Reductase-Beery Twins Story
Helena-Test
Inserting an iCn3D model into a FOB Chapter section
Inserting an interactive mathematic graph into FOB
Inserting a mathematical simulation (SBML) into a FOB
Insert an Image into FOB
Insert a question with hidden answer and floating hint
KP Procko Test
Make your own Chapter 33.x Section
Make Your Own Customized FOB: Assemble and Remix a Custom Book (Short Version)
Pallavi- Test
Pam Mertz Chapter 33 Test
Rebecca Roberts Chapter 33 testing
Rico Acevedo Testy
Samantha Wilner Ch 33 Test
Subhasish-test
Subhasish-TEST2
Index
Glossary

CHAPTER OVERVIEW

1: The Foundations of Biochemistry

1.1: Cellular Foundations
1.2: Chemical Foundations
1.3: Physical-Chemical Foundations
1.4: Genetic Foundations
1.5: Chapter 1 Questions

1.1: Cellular Foundations

1.1.1: Introduction

You have probably studied the cell many times, either in high school or in college biology classes. There are many websites available that review both prokaryotic (bacterial and archaeal cell types) and eukaryotic cells (protist, fungi, plant, and animal cell types). All cells have some similar structural components, including genetic material in the form of chromosomes, a membrane-bound lipid bilayer that separates the inside of the cell from the outside of the cell, and ribosomes that are responsible for protein synthesis. This tutorial is designed specifically from the viewpoint of chemistry. It explores four classes of biomolecules that are also present in all cell types (lipids, proteins, nucleic acids and carbohydrates) and describes in a simplified pictorial manner where they are found, made, and degraded in a typical eukaryotic, animal cell (i.e. their history). This cell review focuses on the organelle structures common in eukaryotic cells. Subsequent chapters will concentrate on the structure and function of specific biomolecules.

Let's think of a cell as a chemical factory that designs, imports, synthesizes, uses, exports, and degrades a variety of chemicals (in the case of the cell, these include lipids, proteins, nucleic acids, and carbohydrates). It also must determine or sense the amount of raw and finished chemicals it has available and respond to its own and external needs by ramping up or shutting off production. ***Biochemistry*** is the branch of science dedicated to the study of these chemical processes within a cell. Understanding these processes can also lend insight into disease states and the pharmacological effects of toxins, drugs, and other medicines within the body.

The building and breaking down of life-sustaining chemicals within an organism is known as ***Metabolism.*** Overall, the three main purposes of metabolism are: (1) the conversion of food to energy to run cellular processes; (2) the conversion of food/fuel to building blocks for the production of ***primary metabolites***, such as proteins, lipids, nucleic acids, and other ***secondary metabolites***; and (3) the elimination of waste products. These enzyme-catalyzed reactions allow organisms to grow and reproduce, maintain their structures, and respond to their environments.

Metabolic reactions may be categorized as ***catabolic***– the *breaking down* of compounds (for example, the breaking down of proteins into amino acids during digestion); or ***anabolic*** – the *building up* (synthesis) of compounds (such as proteins, carbohydrates, lipids, and nucleic acids). Usually, catabolism releases energy, and anabolism consumes energy.

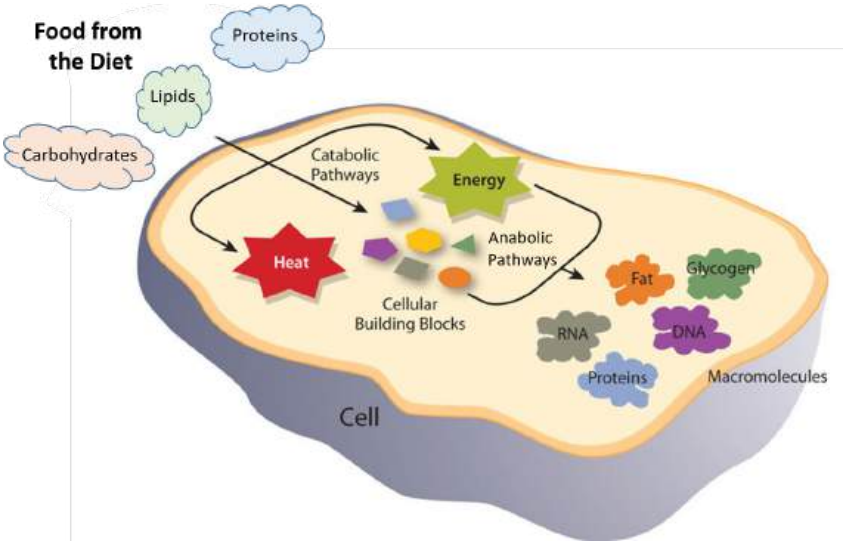


Figure 1.1 Catabolic and Anabolic Reactions. Catabolic reactions involve the breakdown of molecules into smaller components, whereas anabolic reactions build larger molecules from smaller molecules. Catabolic reactions usually release energy whereas anabolic processes usually require energy. (CC BY-SA-NC 3.0; anonymous)

The chemical reactions of metabolism are organized into metabolic pathways, in which one chemical is transformed through a series of steps into another chemical, each step often being facilitated by a specific ***enzyme***. ***Enzymes*** are crucial to metabolism because enzymes act as catalysts – they allow a reaction to proceed more rapidly. In addition, enzymes can provide a mechanism for cells to regulate the rate of a metabolic reaction in response to changes in the cell's environment or to signals from other cells, through the activation or inhibition of the enzyme's activity. Enzymes can also allow organisms to drive desirable reactions that require energy that will not occur by themselves, by coupling them to spontaneous reactions that release energy. Enzyme shape is

critical to the function of the enzyme as it determines the specific binding of a reactant. This can occur by a ***lock and key model*** where the reactant is the exact shape of the enzyme binding site, or by an ***induced fit model***, where the contact of the reactant with the protein causes the shape of the protein to change in order to bind to the reactant. The catalytic mechanisms, kinetics, and regulatory pathways of enzymes will be studied in detail within this text.

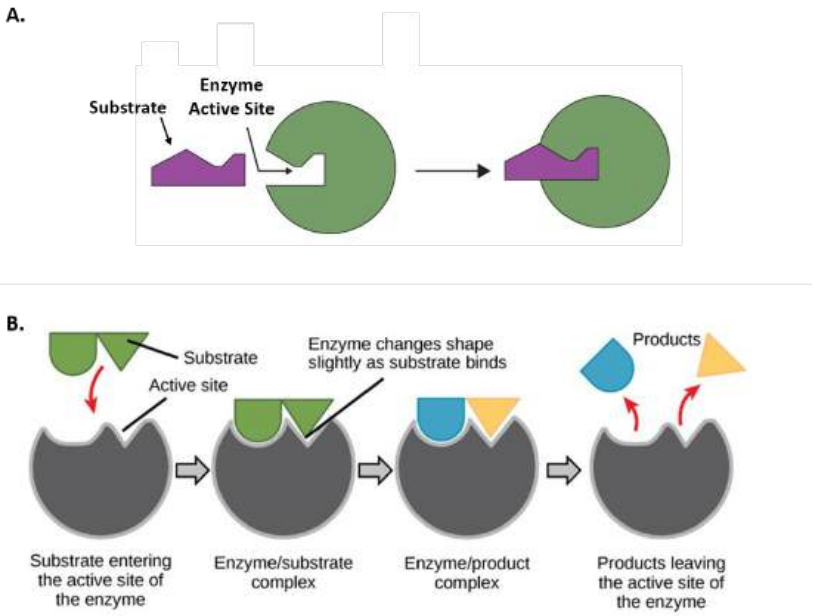


Figure 1.2 Mechanisms of Enzyme-Substrate Binding. (A) In the Lock and Key Model, substrates fit into the active site of the enzyme with no further modifications to the enzyme shape required. (B) In the Induced Fit Model, substrate interaction with the enzyme causes the shape of the enzyme to change to better fit the substrate and mediate the chemical reaction. Figure 1.2A was modified from Socratic and Figure 1.2B was modified from Concepts in Biology

Within eukaryotic cells, the metabolic machinery present allows for the construction of membrane-bound ***organelle*** structures that help to compartmentalize cellular functions. Therefore, ***organelles*** can be thought of as 'little organs' within the cell having discrete cellular functions. The figure of the cell below and in the other linked sites based on it was made available with the kind permission of Liliana Torres. Click on the blue hyperlinks for some of the organelles for more detailed information on them.

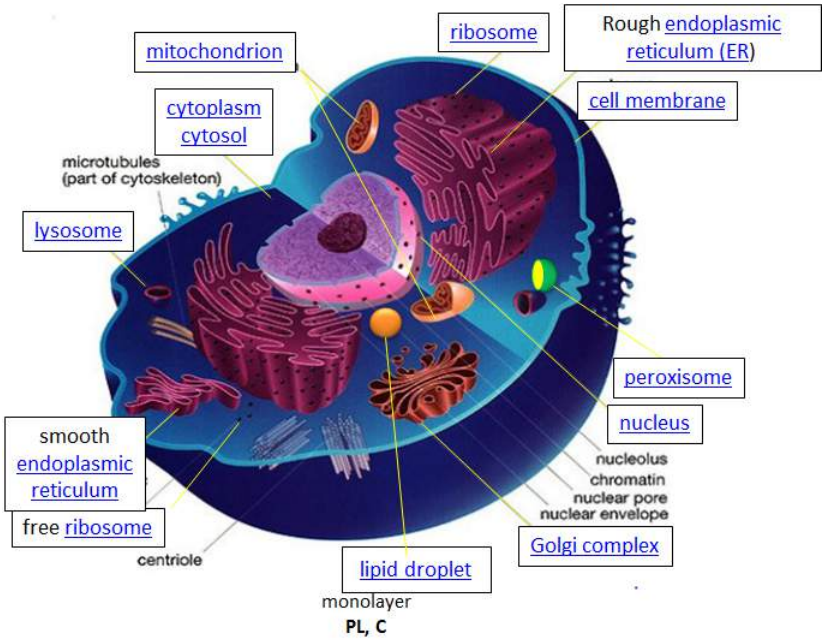


Figure 1.3 Structure of a Typical Eukaryotic Animal Cell. The original figure was acquired from Liliana Torres at: http://torresbioclan.pbworks.com/w/page/22377234/Spikefish%20About%20Cells Use with permission from Liliana Torres. Also at www.animalport.com/animal-cells.html

Design – The design for a cell mostly resides in the blueprint for the cell, the genetic code, which is comprised of the DNA in the cell nucleus and a small amount in the mitochondria. Of course, the DNA blueprint must be read out (transcribed) by ribosomes which themselves were encoded by the DNA and contain a combination of RNA and protein subunits. The genetic code has the master plan that determines the sequence of all cellular proteins, which then catalyze almost all other activities in the cell, including catalysis, motility, architectural structure, etc. In contrast to DNA, RNA, and protein polymers, the length and sequence of polysaccharide polymers and lipids are not driven by such a template but rather by the enzymes that catalyze the synthesis.

Import/Export: Many of the chemical constituents of the cell arise not from direct synthesis but from the import of both small and large molecules. The imported molecules must pass through the cell membrane and in some cases through additional membranes if they need to reside inside membrane-bound organelles. Molecules can move into the cell by passive diffusion across the membrane but usually, their movement is "facilitated" by a membrane transporter protein. Molecules can also move against a concentration gradient in a process called "active transport". Given the amphiphilic nature of the bilayer (polar head group exterior, nonpolar interior), you would expect that polar molecules like glucose would have difficulty in moving across the membrane by passive diffusion. Typically, only small nonpolar molecules move across the membrane via passive transport. Membrane-bound transport proteins are involved in the movement of both nonpolar and polar molecules.

- ***transporters, carrier proteins, and permeases*:** These membrane proteins move specific ligand molecules across a membrane, typically down a concentration gradient. Computer simulations of the facilitated diffusion of lactose across the membrane are shown in the following link. Animation of lactose diffusion through the LacY receptor (The link above and immediately below are from the Theoretical and Computational Biophysics group at the Beckman Institute, the University of Illinois at Urbana-Champaign. These molecular dynamic simulations were made with VMD/NAMD/BioCoRE/JMV/other software support developed by the Group with NIH support.)
- ***ion channels* –** These membrane proteins allow the flow of ions across membranes. Some are permanently open (nongated) while others are gated open or closed depending on the presence of ligands that bind the protein channel and the local environment of the protein in the membrane. The flow of ions through the channel proceeds in a thermodynamically favored direction, which depends on their concentration and voltage gradients across the membrane.
- ***pores*:** Some membranes (nuclear, mitochondria) assemble proteins (such as porins) to form large, but regulated pores. Porins are found in mitochondrial membranes while nucleoporins are found in the nuclear membrane. Small molecules can generally pass through these membrane pores while large ones are selected based on their tendency to form transient intermolecular attractive forces with the pore proteins. The following link shows the diffusion of water through aquaporin. animation of water diffusion through the aquaporin channel,
- ***endocytosis*:** Very large particles [for example, Low Density Lipoproteins (LDL) and viruses] can enter a cell through a process called endocytosis. Initially, the LDL or virus binds to a receptor on the surface of the cell. This triggers a series of events that leads to the invagination of the cell membrane at that point. This eventually pinches off to form an ***endosomal vesicle*** which is surrounded by a protein called clathrin. "Early" endosomes can pick up new proteins and other constituents as well as shed them as they move and mature through the cell. During this maturation process, protein pumps in the endosome lead to a decrease in the endosomal pH which can lead to conformation changes in protein structure and shedding of proteins. Eventually, the "late" endosome reaches and fuses with the lysosome, an internal organelle that contains degradative enzymes. Undegraded components like viral nucleic acids or cholesterol are delivered to the cell. This transport can also go in the reverse direction (called ***exocytosis***) and recycle receptors to the cell membrane. Likewise, vesicles pinched off from the Golgi complex can fuse with endosomes, with some components surviving the process to reenter the Golgi.

Synthesize/Degrade: Cells have to synthesize and degrade small molecules as well as larger polymeric proteins, carbohydrates, lipids, and nucleic acids. The ***anabolic*** (synthetic) and ***catabolic*** (degradative) pathways are often compartmentalized in time and space within a cell. For example, fatty acid synthesis is carried out in the cytoplasm but fatty acid oxidation is carried out in the mitochondria. Proteins are synthesized in the cytoplasm or completed in the endoplasmic reticulum (for membrane and exported proteins) while they are degraded in the lysosome or more importantly in a large multimolecular structure in the cell called the proteasome.

1.1.2: Key Characteristics of a Cell

Let's consider some key characteristics of a cell before we get into the details in later chapters.

1.1.2.1: Cells and their internal compartments have regulated concentrations of ions and hydronium ions.

As expected the pH of the ***cytosol*** (the aqueous substance surrounding all the organelles within the cell) varies from about 7.0-7.4, depending on the metabolic state of the cell. Some organelles have proton transporters that can significantly alter the pH inside an organelle. For example, the pH inside the lysosome, a degradative organelle, is about 4.8. Furthermore, the creation of a pH gradient across the inner mitochondrial membrane is sufficient to drive the thermodynamically unfavored synthesis of ATP.

Compared to the extracellular fluid, the concentration of potassium ions is higher inside the cell, while concentrations of sodium, chloride, and calcium ions are higher on the outside of the cell (see table below). These concentration gradients are maintained by ion transporters and channels and require energy expenditure ultimately in the form of ATP hydrolysis. Changes in these concentrations are integral to the signaling system used by the cell to sense and respond to changes in its external and internal environments. The table below shows approximate ion concentrations in the cell.

Table 1.1 Average Cellular and Extracellular Ion Concentrations

Ion	Inside (mM)	Outside (mM)
Na^+	140	5
K^+	12	140
Cl-	4	15
Ca^{2+}	1 uM	2

1.1.2.2: Cells have an internal framework that provides architectural and internal structural support

The “cytoskeletal” architecture of a (with molecular “cables”- and “girder-like” structures) is not dissimilar from a factory. The internal framework of a cell or ***cytoskeleton,*** is composed of microfilaments, intermediate filaments, and microtubules. These are comprised of monomeric proteins which self-assemble to form the internal architecture. Parts of the cytoskeleton can be seen in Figure 1.4.

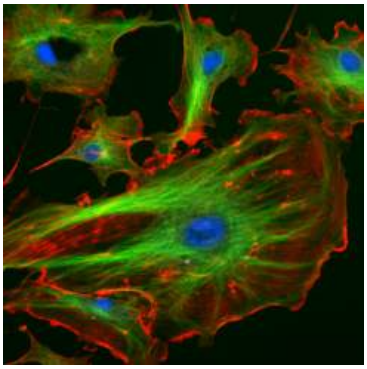


Figure 1.4 Cellular Architecture. The architecture and organization of structural components within a cell *(right picture)* are analogous to the organization seen within a warehouse *(left picture).* In the right image, bovine pulmonary artery endothelial cells have been stained to indicate the nucleus (blue color), tubulin cytoskeletal proteins (fluorescent green color), and F-actin cytoskeletal proteins (fluorescent red color). Source of the factory picture: http://www.cybercom.net/~copters/trips/pictures/factory_inside.jpg Source of the fluorescent cell picture: http://en.Wikipedia.org/wiki/File:FluorescentCells.jpg

Microfilaments of actin monomers (which are stained with a red/orange fluorophore) and microtubules which offer more structural support made of tubulin monomers (stained green) along with the blue-stained nucleus are shown in the image. Organelles are supported and organized by the cytoskeleton (primarily microtubules). Even the cell membrane is supported underneath the inner leaflet by actin (stained orange) and spectrin microfilaments. Motor proteins like myosin (that moves along actin microfilaments) and dynein and kinesin (that move along tubulin microtubules) carry cargo (vesicles, organelles) in a directional fashion. The cell is not a disorganized collection of molecules and organelles. Rather it is highly organized for optimal chemical production, use, and degradation.

Cells have a variety of shapes. Some circulating immune cells must slip through the cells that line capillary walls to migrate to sites of infection. The same process occurs when tumor cells metastasize and escape to other sites in the body. In order to do so, the cell must drastically change shape, a response that requires the dissociation of the cytoskeleton polymers into monomers which are

available later for repolymerization. The following video shows the mobility and flexibility of a Killer T-Cell as it attacks and kills a cancerous cell.

Video 1.1 Killer T Cell Attacking Cancer. Video available on YouTube through creative commons by Cambridge University

1.1.2.3: The cell is an amazingly crowded place

In chemistry labs, we typically work with dilute solutions of solute molecules in a solvent. You have probably heard that the body is comprised of 68% water, but the water concentration is obviously dependent on the cellular environment. Solute molecules like protein and carbohydrates are densely packed. Cells are so crowded that the space between larger molecules like proteins is typically smaller than that of a single protein. Studies have shown that the stability of a protein is increased in such conditions, which would help keep the protein in the correctly folded, native state. Another consequence of high intracellular concentrations is that it limits the diffusion of molecules throughout the cell, as would be expected from an equilibrium perspective in dilute solutions. Thus, cytoplasmic cellular functions can be highly localized within specific regions of the cell creating unique microenvironments and higher differentiation potential within a single cell.

Hence the study of biomolecules in dilute solutions in the lab may not reveal the actual complexities of interactions and activities of the same molecule *in vivo*. Recently investigators have added a neutral copolymer of sucrose and epichlorohydrin to cells *in vitro*. These particles induced the organization of extracellular molecules secreted by the cell, forming an organized extracellular "matrix" which induced the organization of the microfilaments on the inside of the cell as well as inducing changes in cell activity.[1] Furthermore, *in vitro* enzyme activity of a key enzyme in glycolysis dramatically increases under crowded conditions.[2] Another result of crowding may be the spatial and temporal association of key enzymes involved in specific metabolic pathways, allowing for the coordinated passage of substrates and products within the colocalized enzyme system.

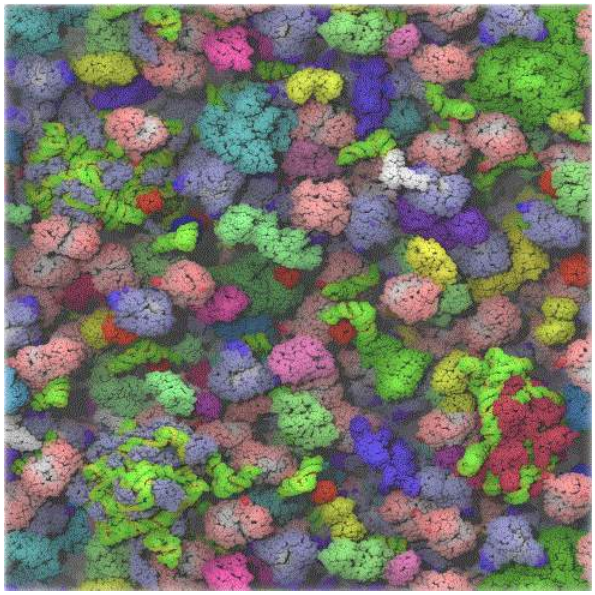


Figure 1.5: The Crowded Cytoplasm of E. Coli. The computer simulation used 50 different types of the most abundant macromolecules of the E. coli cytoplasm and 1008 individual molecules. Rendering of the cytoplasm model at the end of a dynamics simulation. RNA is shown as green and yellow. This figure was prepared with VMD. Figure adapted from: Ufrom McGuffee SR, Elcock AH (2010) PLoS Comput Biol 6(3): e1000694. doi:10.1371/journal.pcbi.1000694 (open source journal)

1.1.2.4: Cell components undergo phase transitions to form substructures within the cell.

A perplexing question is how substructures form within a cell. This includes not only the biogenesis of organelles like mitochondria but also smaller particle such as polysaccharide granules, lipid droplets, protein/RNA particles (including the ribosome) as well as the nucleolus of the cell nucleus. It might be easiest to consider this problem using two examples from the lipid world, lipid droplets and membrane rafts. You are very familiar with phase transitions that occur when a sparing soluble nonpolar liquid is added to water. At a high enough concentration, the solubility of the nonpolar liquid is exceeded and a phase transition occurs as evidenced by the appearance of two separate liquid phases. The same process occurs when triglycerides coalesce into lipid droplets with proteins associated on their outside. Another example occurs within a cell membrane when lipids with saturated alkyl chains self-associate with membrane cholesterol (which contains a rigid planar ring system) to form a membrane microdomain called a ***lipid raft. Lipid rafts*** are characterized by greater packing efficiency, rigidity, and thickness than other parts of the membrane. These lipid rafts often recruit proteins involved in signaling processes within the cell membranes. This process of phase separation is also called ***liquid/liquid demixing*** as two “liquid-like” substances separate.

In a similar manner, it appears that proteins that interact with RNA are composed of less diverse amino acid sequences and have more flexible (“more liquid-like) structures allowing their preferential interaction with RNA to form large RNA-protein particles (like the ribosome and other RNA processing structures) in a fashion that mimics liquid/liquid demixing. All of these interactions are just manifestations of the various intermolecular forces that can exist between molecules. These include ionic interactions, ion-dipole interactions, dipole-dipole interactions, and London dispersion forces (A review of intermolecular forces can be found by Kahn Academy on YouTube).

1.2: Chemical Foundations

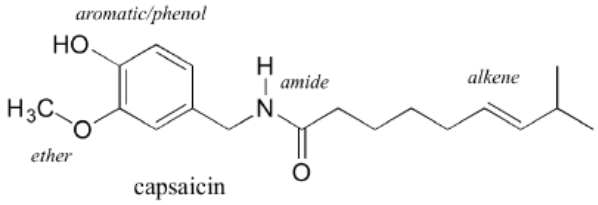
1.2.1: Organic Molecules

On Earth, all carbon-containing molecules have originated from biological, living organisms causing them to be termed **organic compounds**. The number of known organic compounds is quite large. In fact, there are many times more organic compounds known than all the other (inorganic) compounds discovered so far, about 7 million organic compounds in total. Fortunately, organic chemicals consist of relatively few similar parts, combined in different ways. These structural similarities allow us to predict how a compound we have never seen before may react if we know how other molecules containing the same types of parts are known to react.

These parts of organic molecules are called **functional groups** and are made up of specific bonding patterns with the atoms most commonly found in organic molecules (C, H, O, N, S, and P). The identification of functional groups and the ability to predict reactivity based on functional group properties is one of the cornerstones of organic chemistry. Functional groups are specific atoms, ions, or groups of atoms having consistent properties. A functional group makes up part of a larger molecule. For example, -OH, the hydroxyl group that characterizes alcohols, contains oxygen with attached hydrogen. It could be found on any number of different molecules. Just as elements have distinctive properties, functional groups have characteristic chemistries. An -OH functional group on one molecule will tend to react similarly, although perhaps not identically, to an -OH on another molecule.

Organic reactions usually take place at the functional group, so learning about the reactivities of functional groups will prepare you to understand many other aspects about biochemistry.

Functional groups are structural units within organic compounds that are defined by specific bonding arrangements between specific atoms. The structure of capsaicin, the fiery compound found in hot peppers, incorporates several functional groups, labeled in the figure below and explained throughout this section.

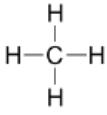


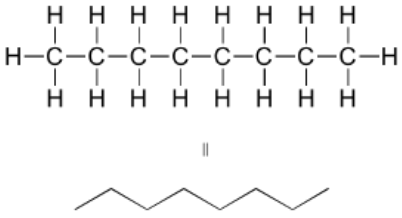
capsaicin

As we progress in our study of biochemistry, it will become extremely important to be able to quickly recognize the most common functional groups, because *they are the key structural elements that define how organic molecules react*. Below is a brief introduction to the major organic functional groups.

1.2.1.1: Alkanes

The ‘default’ in organic chemistry (essentially, the *lack* of any functional groups) is given the term *alkane*, characterized by single bonds between carbon and carbon, or between carbon and hydrogen. Methane, CH_4, is the natural gas you may burn in your furnace. Octane, C_8H_{18}, is a component of gasoline.

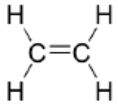

methane


octane

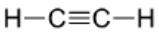
1.2.1.2: Alkenes and Alkynes

Alkenes (sometimes called olefins) have carbon-carbon double bonds, and *alkynes* have carbon-carbon triple bonds. Ethene, the simplest alkene example, is a gas that serves as a cellular signal in fruits to stimulate ripening. (If you want bananas to ripen

quickly, put them in a paper bag along with an apple – the apple emits ethene gas (also called ethylene), setting off the ripening process in the bananas). Ethyne, commonly called acetylene, is used as a fuel in welding blow torches.

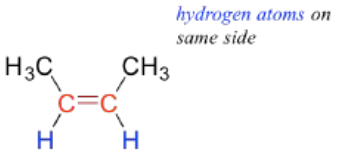


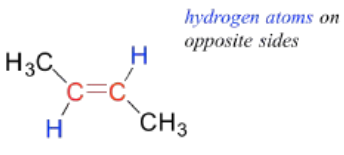
ethene
(an alkene)



ethyne
(an alkyne)

Many alkenes can take two geometric forms: *cis* or *trans*. The *cis* and *trans* forms of a given alkene are different isomers with different physical properties because there is a very high energy barrier to rotation about a double bond. In the example below, the difference between *cis* and *trans* alkenes is readily apparent. Biochemists don't usually use the E (entgegen) and Z (zusammen) labels for groups attached to double bonds (using IUPAC priority numbering).





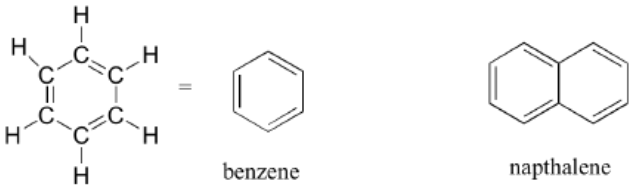
cis-alkene

trans-alkene

Alkanes, alkenes, and alkynes are all classified as *hydrocarbons* because they are composed solely of carbon and hydrogen atoms. Alkanes are said to be *saturated hydrocarbons*, because the carbons are bonded to the maximum possible number of hydrogens – in other words, they are '*saturated*' with hydrogen atoms. The double and triple-bonded carbons in alkenes and alkynes have fewer hydrogen atoms bonded to them – they are thus referred to as *unsaturated hydrocarbons*.

1.2.1.3: Aromatics

The *aromatic group* is exemplified by benzene (which used to be a commonly used solvent on the organic lab, but which was shown to be carcinogenic), and naphthalene, a compound with a distinctive 'mothball' smell. Aromatic groups are planar (flat) ring structures, with conjugated pi bonding with 4n+2 pi electrons. Given the stability of aromatic groups due to delocalization of the pi electrons, these groups are widespread in nature.

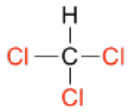


benzene

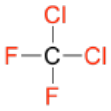
napthalene

1.2.1.4: Alkyl Halides

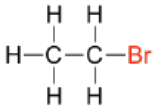
When the carbon of an alkane is bonded to one or more halogens, the group is referred to as an *alkyl halide or haloalkane*. Chloroform is a useful solvent in the laboratory, and was one of the earlier anesthetic drugs used in surgery. Chlorodifluoromethane was used as a refrigerant and in aerosol sprays until the late twentieth century, but its use was discontinued after it was found to have harmful effects on the ozone layer. Bromoethane is a simple alkyl halide often used in organic synthesis. Alkyl halides groups are quite rare in biomolecules.



trichloromethane
(chloroform)



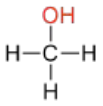
dichlorodifluoromethane
(Freon-12)



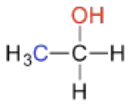
bromoethane

1.2.1.5: Alcohols, Phenols, and Thiols

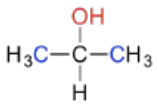
In the ***alcohol*** functional group, a carbon is single-bonded to an OH group (the OH group, when it is part of a larger molecule, is referred to as a ***hydroxyl group***). Except for methanol, all alcohols can be classified as primary, secondary, or tertiary. In a primary alcohol, the carbon bonded to the OH group is also bonded to only one other carbon. In secondary and tertiary alcohols, the carbon is bonded to two or three other carbons, respectively. When the hydroxyl group is *directly* attached to an aromatic ring, the resulting group is called a ***phenol.*** The sulfur analog of an alcohol is called a thiol (from the Greek *thio*, for sulfur).



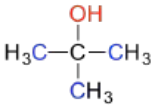
methanol



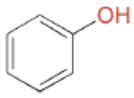
primary alcohol



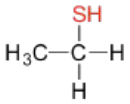
secondary alcohol



tertiary alcohol

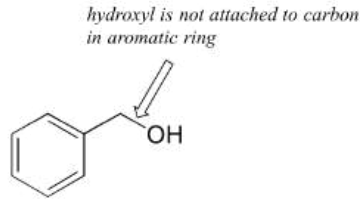


phenol



thiol

Note that the definition of a phenol states that the hydroxyl oxygen must be *directly* attached to one of the carbons of the aromatic ring. The compound below, therefore, is *not* a phenol – it is a primary alcohol.

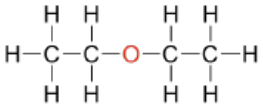


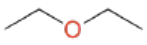
primary alcohol
(*not* a phenol)

The distinction is important because there is a significant difference in the reactivity of alcohols and phenols

1.2.1.6: Ethers and Sulfides

In an *ether* functional group, oxygen is bonded to two carbons. Below is the structure of diethyl ether, a common laboratory solvent and also one of the first compounds to be used as an anesthetic during operations. The sulfur analog of ether is called a *thioether or sulfide*.



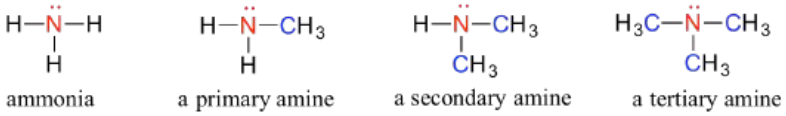


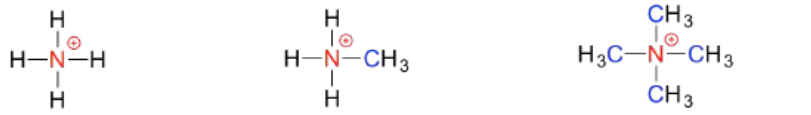
an ether

a sulfide

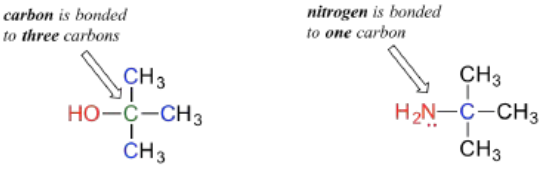
1.2.1.7: Amines

Amines are characterized by nitrogen atoms with single bonds to hydrogen and carbon. Just as there are primary, secondary, and tertiary alcohols, there are primary, secondary, and tertiary amines. Ammonia is a special case with no carbon atoms. One of the most important properties of amines is that they are basic, and are readily protonated to form ammonium cations. In the case where nitrogen has four bonds to carbon (which is somewhat unusual in biomolecules), it is called a quaternary ammonium ion.



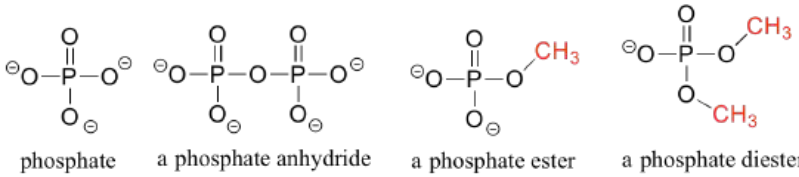


Note: Do not be confused by how the terms ‘primary’, ‘secondary’, and ‘tertiary’ are applied to alcohols and amines – the definitions are different. In alcohols, what matters is how many other carbons the alcohol *carbon* is bonded to, while in amines, what matters is how many carbons the *nitrogen* is bonded to.



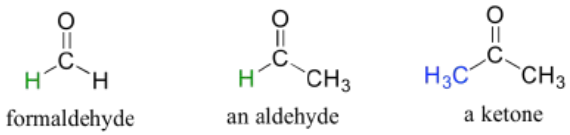
1.2.1.8: Organic Phosphates

Phosphate and its derivative functional groups are ubiquitous in biomolecules. Phosphate linked to a single organic group is called a *phosphate ester*; when it has two links to organic groups it is called a phosphate diester. A linkage between two phosphates creates a phosphate anhydride.



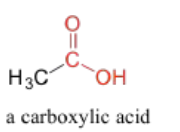
1.2.1.9: Aldehydes and Ketones

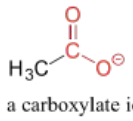
There are a number of functional groups that contain a carbon-oxygen double bond, which is commonly referred to as a *carbonyl*. Ketones and aldehydes are two closely related carbonyl-based functional groups that react in very similar ways. In a *ketone*, the carbon atom of a carbonyl is bonded to two other carbons. In an *aldehyde*, the carbonyl carbon is bonded on one side to hydrogen, and on the other side to carbon. The exception to this definition is formaldehyde, in which the carbonyl carbon has bonds to two hydrogens.

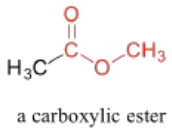


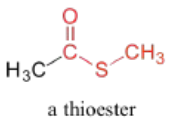
1.2.1.10: Carboxylic Acids and Their Derivatives

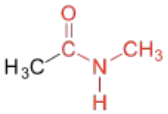
When a carbonyl carbon is bonded on one side to a carbon (or hydrogen) and on the other side to an oxygen, nitrogen, or sulfur, the functional group is considered to be one of the carboxylic acid derivatives, a designation that describes a set of related functional groups. The main member of this family is the *carboxylic acid* functional group, in which the carbonyl is bonded to a hydroxyl group. The *carboxylate ion* form has donated the H^+ to the solution. Other derivatives are *carboxylic esters*(usually just called 'esters'), *thioesters*, *amides*, *acyl phosphates*, *acid chlorides*, and *acid anhydrides*. With the exception of acid chlorides and acid anhydrides, carboxylic acid derivatives are very common in biological molecules and/or metabolic pathways and will be discussed in further detail in a later chapter.

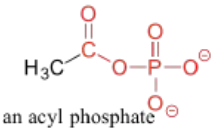

a carboxylic acid

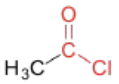

a carboxylate ion

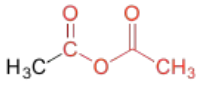

a carboxylic ester


a thioester


an amide

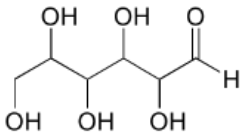

an acyl phosphate

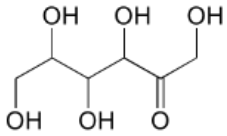

an acid chloride


an acid anhydride

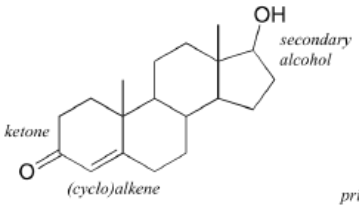
1.2.2: Practice Recognizing Functional Groups in Molecules

A single compound often contains several functional groups, particularly in biological organic chemistry. The six-carbon sugar molecules glucose and fructose, for example, contain aldehyde and ketone groups, respectively, and both contain five alcohol groups. A compound with several alcohol groups is often referred to as a *'polyol'*.

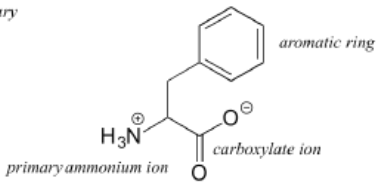

glucose


fructose

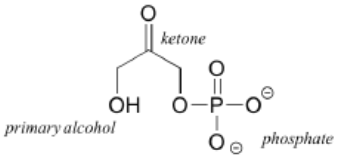
The hormone testosterone, the amino acid phenylalanine, and the glycolysis metabolite dihydroxyacetone phosphate all contain multiple functional groups, as labeled below.



testosterone



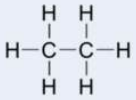
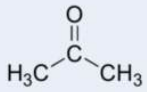
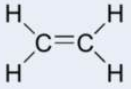
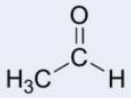
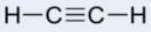
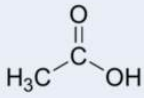
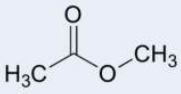
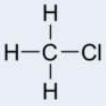
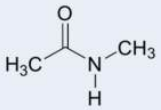
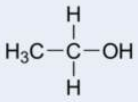
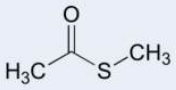
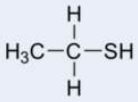
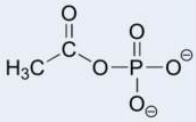
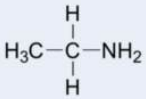
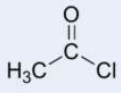
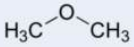
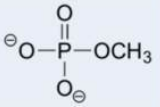
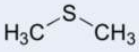
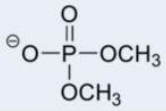
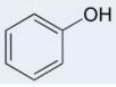
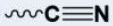
phenylalanine



dihydroxyacetone phosphate (DHAP)

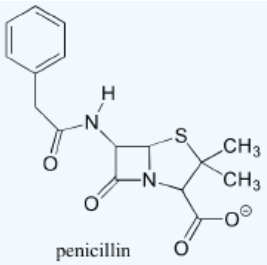
While not in any way a complete list, this section has covered most of the important functional groups that we will encounter in biochemistry. Table 1.7 provides a summary of all of the groups listed in this section.

Table 1.7 Common Organic Functional Groups

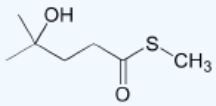
alkane		ketone	
alkene		aldehyde	
alkyne		carboxylic acid	
aromatic		ester	
alkyl halide		amide	
alcohol		thioester	
thiol		acyl phosphate	
amine		acid chloride	
ether		phosphate monoester	
thioether		phosphate diester	
phenol		nitrile	

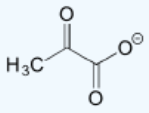
? Exercise 1.2.1

Identify the functional groups (other than alkanes) in the following organic compounds. State whether alcohols and amines are primary, secondary, or tertiary.

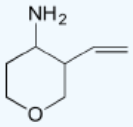
a) 

penicillin

b) 

c) 

pyruvate

d) 

? Exercise 1.2.2

Draw one example of each compound fitting the descriptions below, using line structures. Be sure to designate the location of all non-zero formal charges. All atoms should have complete octets (phosphorus may exceed the octet rule). There are many possible correct answers for these, so be sure to check your structures with your instructor or tutor.

a. a compound with molecular formula $C_6H_{11}NO$ that includes alkene, secondary amine, and primary alcohol functional groups
b. an ion with molecular formula $C_3H_5O_6P^{2-}$ that includes aldehyde, secondary alcohol, and phosphate functional groups.
c. A compound with molecular formula C_6H_9NO that has an amide functional group, and does *not* have an alkene group.

1.2.3: Primary metabolites

Primary metabolites are components of basic metabolic pathways that are required for life. They are associated with essential cellular functions such as nutrient assimilation, energy production, and growth/development. They have a wide species distribution that spans many phyla and frequently more than one kingdom. Primary metabolites include the building blocks required to make the four major macromolecules within the body: carbohydrates, lipids, proteins, and nucleic acids (DNA and RNA).

These are large polymers of the body that are built up from repeating smaller monomer units (Fig. 6.1). The monomer units for building the nucleic acids, DNA and RNA, are the nucleotide bases, whereas the monomers for proteins are amino acids, for carbohydrates are sugar residues, and for lipids are fatty acids or acetyl groups.

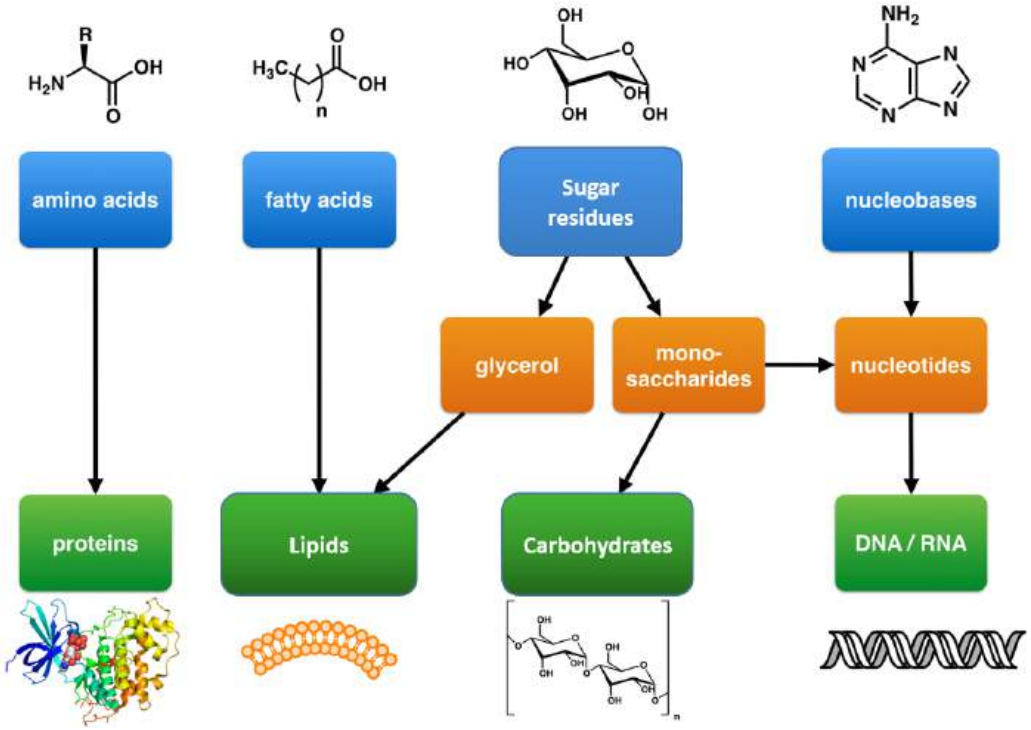


Figure 1.28 The Molecular building blocks of life are made from organic compounds. Modified from: Boghog

1.2.4: Reactions forming the Major Macromolecules

The major macromolecules are built by putting together repeating monomer subunits through the process of dehydration synthesis. Interestingly, the organic functional units used in the dehydration synthesis processes for each of the major types of macromolecules have similarities with one another. Thus, it is useful to look at the reactions together (Figure 1.29)

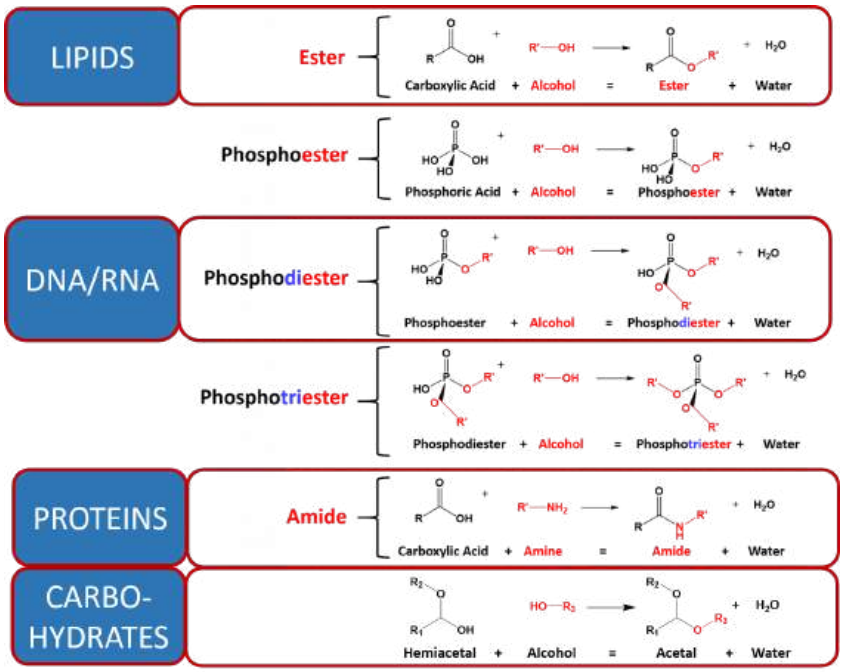


Figure 1.29 Dehydration Synthesis Reactions Involved in Macromolecule Formation. The major organic reactions required for the biosynthesis of lipids, nucleic acids (DNA/RNA), proteins, and carbohydrates are shown. Note that in all of the reactions, there is a functional group that contains two electron withdrawing groups (the carboxylic acid, phosphoric acid and the hemiacetal each have two oxygen atoms attached to a central carbon or phosphorus atom). This forms a reactive partially positive center atom (carbon in the case of the carboxylic acid and hemiacetal, or phosphorus in the case of the phosphoric acid) that can be attacked by the electronegative oxygen or nitrogen from an alcohol or amine functional group. Within biological systems, many functional groups, such as carboxylic acids, require activation before they can be utilized in synthesis reactions and will be detailed in later chapters.

Primary metabolites that are involved with energy production include numerous enzymes that break down food molecules, such as carbohydrates and lipids, and capture the energy released during the hydrolysis of adenosine triphosphate (ATP). *Enzymes* are biological catalysts that speed up the rate of chemical reactions. Typically they are proteins, which are composed of amino acid building blocks. The basic structure of cells and of organisms are also composed of primary metabolites. These include cell membranes (e.g. phospholipids), cell walls (e.g. peptidoglycan, chitin), and cytoskeletons (proteins). DNA and RNA which store and transmit genetic information are composed of nucleic acid primary metabolites. Primary metabolites also include molecules involved in cellular signaling, communication, and transport. The structure and function of primary metabolites are key components of this text. These reactions will be detailed in the following chapters.

1.3: Physical-Chemical Foundations

The types and numbers of chemical reactions that occur in biological cells are staggering. Compared to both physical and chemical reactions that occur in a controlled and closed environment, biological reactions occur in open systems with input and output of both energy and chemical "feedstocks". Yet they are governed by the same physical principles that control all reactions. We can gain insight into biological reactions and how they are controlled by considering the same principles you have used in a myriad of preceding classes, energy changes, equilibria and thermodynamics. Let's review them!

1.3.1: Reactions and Energy Changes

Why do reactions vary in extent from completely irreversible in the forward reaction to reversible reactions favoring the reactants? It might help to understand a simple physical reaction before we try more complicated chemical reactions. Let's start with a simple ball on a hill. Does a ball at the top of a hill roll downhill spontaneously, or does the opposite happen? No one has ever seen a ball roll spontaneously uphill unless a lot of energy was added to the ball. This physical reaction appears to be irreversible and occurs because the ball has lower potential energy at the bottom of the hill than it does at the top. The gap in the potential energy is related to the "extent" and spontaneity of this reaction. As we have undoubtedly observed before, processes in nature tend to go to a lower energy state. By analogy, we will consider the driving force for a chemical reaction to be the free energy difference, ΔG, between reactants and products. ΔG determines the extent and spontaneity of the reaction.

1.3.1.1: Reversible/Irreversible Reactions, Extent of Reactions, Equilibria

Consider a hypothetical reversible reaction in which you start with some reactants, A and B, each at a 1 M concentration (1 mol of each/L solution). but no products, P and Q. For ease assume that the total volume of solution is 1 L, so that we start with 1 mol each of A and B. At time $t = 0$, the concentration of products is 0. The reaction can be written as:

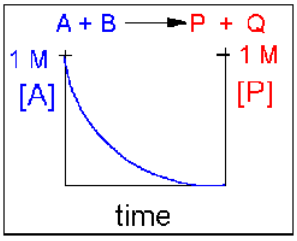
$$\mathrm{A} + \mathrm{B} \rightleftharpoons \mathrm{P} + \mathrm{Q}.$$

As time progresses, the amounts or concentrations of A and B decrease as the amounts or concentrations of products P and Q increase. At some time, no further changes occur in the amount or concentrations of remaining reactants or products. At this point, the reaction is in equilibrium, a term used often in our common vocabulary to denote a system that is undergoing no net change.

Most of the reactions that we will study occur in solution, so we will deal with concentrations (in mol/L or mmol/mL = M). Let's consider how the concentration of reactants and products change as a function of time. Depending on the extent to which a reaction is reversible, 4 different scenarios can be imagined:

Scenario 1: Irreversible reaction in which the reverse reaction occurs to a negligible extent.

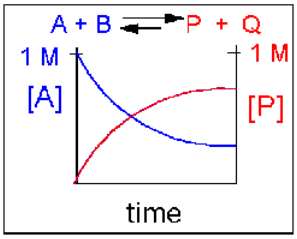
In this reaction, the reverse reaction occurs to such a small extent that we can neglect it. The only reaction that occurs is the conversion of reactants to products. Hence all the reactants are converted to product. At equilibrium [A] = 0. Since 1 mol of A reacted, it must form 1 mol of P and 1 mol Q - i.e. the concentration of products at equilibrium is 1 M. At an earlier time of the reaction, (let's pick a time when [A] = 0.8 M), only part of the reactants have reacted (in this case 0.2 M), producing an equal amount of products, P and Q. Graphs of [A] and [P] as a function of timer are shown below. [A] decreases in a nonlinear fashion to 0 M while [P] increases in a reciprocal fashion to 1 M concentration. This is illustrated in the graph below.



Examples of irreversible reactions are reactions of strong acids (nitric, sulfuric, hydrochloric) with bases (OH- and water), or the much more complicated combustion reactions like the burning of sugars (like trees) and hydrocarbons (like octane) to form CO_2 and H_2O.

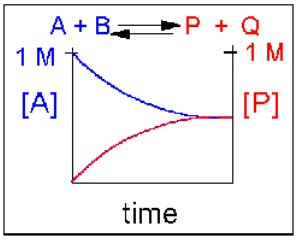
Scenario 2: Reversible reaction in which the forward reaction is favored.

Again [A] decreases and [P] increases, but in this case, some A remains since the reaction is reversible. As [A]and [B] decrease, [P] and [Q] increase, which increases the chance that they will collide and form the product. Since P and Q can react to form reactants, the [A] at equilibrium is not zero as is shown below.



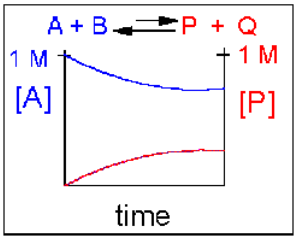
Scenario 3: Reversible Reaction in which forward and reverse reactions are equally favored.

Again [A] decreases and [P] increases, but in this case, some A remains since the reaction is reversible. As [A] and [B] decrease, [P] and [Q] increase, which increases the chance that they will collide and form the product. Since P and Q can react to form reactants, the [A] at equilibrium is not zero as is shown below. Because the reactants and products are equally favored, their concentrations will be equal at equilibrium.



Scenario 4: Reversible Reaction in which the reverse reaction is favored.

Again [A] decreases and [P] increases, but in this case, some A remains since the reaction is reversible. As [A] and [B] decrease, [P] and [Q] increase, which increases the chance that they will collide and form the product. Since P and Q can react to form reactants, the [A] at equilibrium is not zero as is shown below. Because the reaction favors reactants. their concentration will be higher at equilibrium than the products.



An example of this kind of reaction, one that favors reactants, is the reaction of acetic acid (a weak acid) with water.

$$\mathrm{CH_3CO_2H(aq) + H_2O(l) \rightleftharpoons H_3O^+(aq) + CH_3CO_2^-(aq)}$$

1.3.2: Equilibrium Constants

Without a lot of experience in chemistry, it is difficult to just look at the reactants and products and determine whether the reaction is irreversible, or reversible, favoring either reactants or products (with the exception of obvious irreversible reactions described above). However, this data can be found in tables of equilibrium constants. The equilibrium constant, as its name implies, is constant, independent of the concentration of the reactants and products. A $K_{eq} > 1$ implies that the products are favored. A $K_{eq} < 1$ implies that reactants are favored. When $K_{eq} = 1$, both reactants and products are equally favored. For the more general reaction,

$$\mathrm{aA + bB \rightleftharpoons pP + qQ}$$

where a, b, p, and q are the stoichiometric coefficients,

$$K_{eq} = \frac{[P]^p_{eq}[Q]^q_{eq}}{[A]^a_{eq}[B]^b_{eq}} \tag{1.3.1}$$

where all the concentrations are those at equilibrium. For a simple reaction where a, b, p, and q are all 1, then $K_{eq} = ([P][Q])/([A][B])$.

(Note: Equilibrium constants are truly constant only at a given temperature, pressure, and solvent condition. Likewise, they depend on concentration to the extent that their activities change with concentration.)

For an irreversible reaction, such as the reaction of a 0.1 M HCl (aq) in water, $[HCl]_{eq}$ = 0, you cannot easily measure a K_{eq}. However, if we assume the reaction goes in reverse to an almost imperceptible degree, $[HCl]_{eq}$ might equal 10-10 M. Hence K_{eq} >> 1.

In summary, the extent of reactions can vary from completely irreversible (favoring only the products) to reactions that favor the reactants.

Our next goal is to understand what controls the extent of a reaction. That is, of course, the change in the Gibbs free energy. Two different pairs of factors influence the ΔG. One pair is concentration and inherent reactivity of reactants compared to products (as reflected in the Keq). The other pair is enthalpy/entropy changes. We will now consider the first pair.

Contributions of Molecule Stability (K_{eq}) and concentration to ΔG

Consider the reactions of hydrochloric acid and acetic acid with water.

$$HCl(aq) + H_2O(l) \longrightarrow H_3O^+(aq) + Cl^-(aq)$$

$$CH_3CO_2H(aq) + H_2O(l) \longrightarrow H_3O^+(aq) + CH_3CO_2^-(aq)$$

Assume that at t = 0, each acid is placed into water at a concentration of 0.1 M. When equilibrium is reached, there is essentially no HCl left in solution, while 99% of the acetic acid remains. Why are they so different? We rationalized that HCl (aq) is a much stronger acid than H_3O^+(aq) which itself is a much stronger acid than CH_3CO_2H (aq). Why? All we can say is there is something about the structure of these acids (and the bases) that makes HCl much more intrinsically unstable, much higher in energy, and hence much more reactive than the acid it forms, H_3O^+(aq). Likewise, H_3O^+(aq) is much more intrinsically unstable, much higher in energy, and hence more reactive than CH_3CO_2H (aq). This has nothing to do with concentration since the initial concentration of both HCl (aq) and CH_3CO_2H (aq) were identical. This observation is reflected in the K_{eq} for these acids (>>1 for HCl and <<1 for acetic acid). This difference in intrinsic stability of reactants compared to products (which is independent of concentration) is one factor that contributes to ΔG.

The other factor is concentration. A 0.25 M (0.25 mol/L or 0.25 mmol/ml) solution of acetic acid does not conduct electricity, implying that very few ions of H_3O^+(aq) + $CH_3CO_2^-$ (aq) exist in the solution. However, if more concentrated acetic acid is added, a dim light becomes evident. Adding more reactants seemed to drive the reaction to form more products, even though the reverse reaction is favored if one considers only the intrinsic stability of reactants and products. Before the concentrated acid was added, the system was at equilibrium. Adding concentrated acid perturbed the equilibrium, which drove the reaction to form additional products. This is an example of Le Chatelier's Principle, which states that if a reaction at equilibrium is perturbed, the reaction will be driven in the direction that will relieve the perturbation. Hence:

- if more reactant is added, the reaction shifts to form more products
- if more product is added, the reaction shifts to form more reactants
- if products are selectively removed (by distillation, crystallization, or further reaction to produce another species), the reaction shifts to form more product.
- if reactants are removed (as above), the reaction shifts to form more reactants.
- if heat is added to an exothermic reaction, the reaction shifts to get rid of the excess heat by shifting to form more reactants. (opposite for an endothermic reaction).

1.3.3: Change in Free Energy G

Without doing a complicated derivation, these simple examples suggest that the total ΔG can be expressed as the sum of the two contributions showing the effects of the intrinsic stability (K_{eq}) and concentration:

$$\Delta G_{total} = \Delta G_{stability} + \Delta G_{concentration} \tag{1.3.2}$$

which becomes for the simple reaction $\mathrm{A + B \rightleftharpoons P + Q}$ (after a rigorous derivation):

$$\begin{aligned} \Delta G &= \Delta G^0 + RT \ln \frac{[\mathrm{P}][\mathrm{Q}]}{[\mathrm{A}][\mathrm{B}]} \\ &= \Delta G^0 + RT \ln \mathrm{Q_{rx}} \end{aligned} \tag{1.3.3}$$

where ΔG^o reflects the contribution from the relative intrinsic stability of reactants and products and the second term reflects the contribution from the relative concentrations of reactants and products (which has nothing to do with stability). Q_{rx} is the **reaction quotient** which for the reaction A + B ↔ P + Q is given by:

$$Q_{rx} = ([P][Q])/([A][B]) \tag{1.3.4}$$

at any point in the reaction.

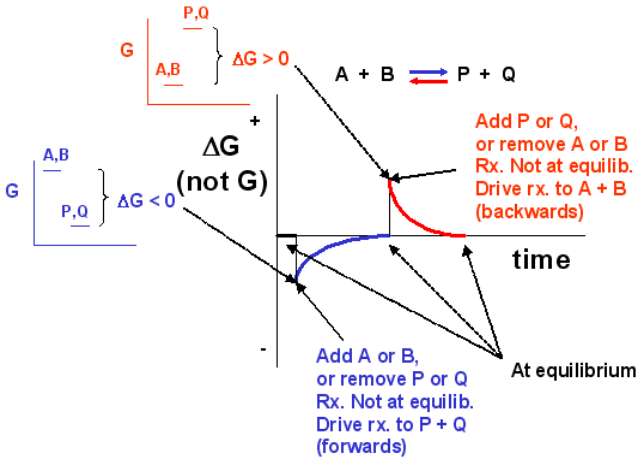
Meaning of ΔG

Remember that ΔG is the "driving" force for a reaction, analogous to the difference in potential energy for a ball on a hill. Go back to that analogy. if the ball starts at the top of the hill, does it roll downhill? Of course. It goes from high potential energy to low potential energy. The reaction can be written as: Ball $_{top}$ → Ball $_{bottom}$ for which the change in potential energy, ΔPE = PE_{bottom} -PE_{top}< 0. If the ball starts at the bottom, will it go to the top? Obviously not. For that reaction, Ball $_{bottom}$ → Ball $_{top}$, ΔPE > 0. If the top of the hill was at the same height at the bottom of the hill (obviously an absurd situation), the ball would not move. It would effectively be at equilibrium, a state of no change. For this reaction, Ball $_{top}$ --> Ball $_{bottom}$, the ΔPE = 0. As the ball starts rolling down the hill, its potential energy gets closer to the potential it would have at the bottom. Hence the ΔPE changes from negative to more and more positive until it gets to the bottom at which case the ΔPE = 0 and movement ceases. If the ΔPE is not 0, the ball will move until the ΔPE = 0.

Likewise, for a chemical reaction that favors products, ΔG < 0. The system is not at equilibrium and the reaction will go in the direction of products. As the reaction proceeds, products buildup, and there is less of a driving force for reactants to go to products (Le Chatelier'sPrinciple), so the ΔG becomes more positive until the ΔG = 0 and the reaction is at equilibrium. A reaction that has a ΔG > 0 is likewise not at equilibrium so it will go in the appropriate direction until equilibrium is reached. Hence for the reaction A + B <==> P + Q,

- if ΔG < 0, the reaction goes toward products P and Q
- if ΔG = 0, the reaction is at equilibrium and no further change occurs in the concentration of reactants and products.
- if ΔG > 0, the reaction goes toward reactants A and B.

We can not measure easily the actual free energy G of reactants or products, but we can measure ΔG readily. These points are illustrated in the graph below of ΔG vs time for the hypothetical reaction A + B ↔ P + Q. (Also notice the two insert graphs - in blue and red - which show, in analogy to the ball on the hill graphs, the values of ΔG at the two points where the perturbation to the equilibrium were made.)



Notice the ΔG is constantly changing until the system reaches equilibrium. Initially, the equilibrium is perturbed so that the system is not in equilibrium (shown in blue). The perturbation was such that the products are favored. After equilibrium was reached, the system was perturbed again, this time in a fashion to favor the reverse reaction. Notice in this case the ΔG for the reaction as written: A + B ↔ P + Q is positive - i.e. it is not in equilibrium. Therefore the reaction (as written) goes backwards to products. It is important to realize that the reported ΔG is for the reaction as written.

Now let's apply $\Delta G = \Delta G^o + RT\ln Q = \Delta G^o + RT\ln([P][Q])/([A][B])$ to two reactions we discussed above:

- $HCl(aq) + H_2O(l) \leftrightarrow H_3O^+(aq) + Cl^-(aq)$
- $CH_3CO_2H(aq) + H_2O(l) \leftrightarrow H_3O^+(aq) + CH_3CO_2^-(aq)$

Assume that at time t=0, 0.1 mole of HCl and CH_3CO_2H were added to two different beakers. At this point the forward reaction is favored, but obviously to different extents. The RTln Q would be identical for both acids since each reactant is present at 0.1 M, but no products yet exist. However, the ΔG^o is negative for HCl and positive for acetic acid since HCl is a strong acid. Hence at t=0, **ΔG** for the HCl reaction is much more negative than for acetic acid. This is summarized in the table below. The direction of the arrow shows if products (-->) or reactants (<---) are favored. The size of the arrow shows very approximately to what extent the ΔG term is favored

Reaction at t=0	ΔG^o	RTln Q	ΔG
$HCl(aq) + H_2O(l)$	---------->	---------->	---------->
$CH_3CO_2H(aq) + H_2O(l)$	<----------	---------->	->

Now when equilibrium is reached, no net change occurs in the concentration of reactants and products, and ΔG = 0. In the case of HCl, there is just an infinitesimal amount of HCl left, and 0.1 M of each product, so concentration favors HCl formation. However, the intrinsic relative stability of reactants and products still favors products. In the case of acetic acid, most of the acetic acid remains (0.099 M) with little product (0.001 M) so concentration favors products. However, the intrinsic relative stability of reactants and products still favors reactants. This is summarized in the table below.

Reaction at equlib.	ΔG^ostab	RTln Q	ΔG
$HCl(aq) + H_2O(l)$	---------->	<----------	favors neither, = 0
$CH_3CO_2H(aq) + H_2O(l)$	<----------	---------->	favors neither, = 0

Compare the two tables above (one at time t= 0 and the other at equilibrium). Notice:

- ΔG^o does not change in a given set of conditions, since it has nothing to do with concentration.
- Only RTln Q changes during the course of a reaction until equilibrium is achieved.

Meaning of ΔG^o

To get a better meaning of the significance of ΔG^o, let's consider the following equations under two different conditions:

$$\Delta G = \Delta G^0 + RT\ln\frac{[P][Q]}{[A][B]} = \Delta G^0 + RT\ln Q_{rx} \tag{1.3.5}$$

Condition I: Reaction at equilibrium, ΔG = 0

The equation reduces to:

$$\Delta G^0 = -RT\ln\frac{[P]_{eq}[Q]_{eq}}{[A]_{eq}[B]_{eq}} = -2.303RT\log K_{eq} \tag{1.3.6}$$

This supports our idea that ΔG^o is independent of concentration since K_{eq} should also be independent of concentration.

Condition II: Concentration of all reactants and products is 1 M (standard state, assuming solution reaction)

The equation reduces to:

$$\Delta G = \Delta G^0 + RT \ln \frac{[1][1]}{[1][1]} = \Delta G^0 + 2.303 RT \log 1 = \Delta G^0$$

$$\begin{aligned} \Delta G &= \Delta G^o + RT \ln\left(\frac{[1][1]}{[1][1]}\right) \\ &= \Delta G^o + 2.303 RT \log 1 \\ &= \Delta G^o \end{aligned} \tag{1.3.7}$$

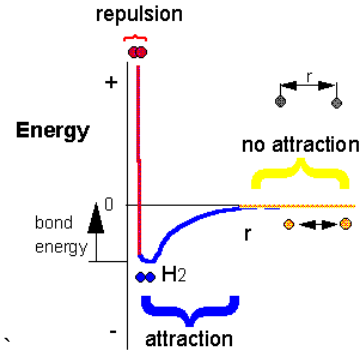
This implies that when all reactants are at this concentration, defined as the standard state (1 M for solutes), the ΔG at that particular moment just happens to be the ΔG^o for the reaction. If one of the reactants or products is H_3O^+, it would make little biological sense to calculate ΔG^o for the reaction using the standard state of $[H_3O^+]$ = 1 M, or a pH of -1. Instead, it is assumed that the pH = 7, $[H_3O^+] = 10^{-7}$ M. A new symbol is used for ΔG^o under these conditions, $\Delta G^{o'}$.

1.3.4: Heat, Enthlapy and Entropy

Consider the association reaction of hydrogen atoms into molecular hydrogen

$$\mathrm{H + H \longrightarrow H_2}.$$

Does this reaction occur spontaneously? It does. You should remember that individual H atoms are unstable since they don't have a completed outer shell of electrons - in this case, a duet. As they approach, they can interact to form a covalent bond and in the process release energy. The bonded state is a lower energy state than two separated H atoms. This should be clear since energy has to be added to a molecule of H_2 to break the bond. We call this the **bond dissociation energy**.



Now consider a more complicated reaction, the burning of octane.

$$\mathrm{2\,C_8H_{18}(l) + 25\,O_2(g) \rightarrow 16\,CO_2(g) + 18\,H_2O(g)}$$

To carry out this reaction, every C-C, C-H and O-O bond in the reactants must be broken (which requires an input of energy) but lots of energy is released on the formation of C-O and H-O covalent bonds in the products. Is more energy needed to break the bonds in the reactants or is more energy released on the formation of bonds in the product? The answer should be clear. The products must be at a lower energy than the reactants since huge amounts of heat and light energy are released on the combustion of gasoline and other hydrocarbons.

These reactions suggest that energy must be released for a reaction to proceed to any extent in a given direction.

Now consider, however, the following reaction:

$$\mathrm{Ba(OH)_2 \cdot 8\,H_2O(s) + 2\,NH_4SCN(s) \longrightarrow 10\,H_2O(l) + 2\,NH_3(g) + Ba(SCN)_2(aq)}$$

When these two solids are added to a beaker and stirred, a reaction clearly takes place, as evidenced by the formation of a liquid (water) and the smell of ammonia. What is surprising is that heat is not released into the surroundings in this reaction. Rather heat was absorbed from the surroundings turning the beaker so cold that it freezes to a piece of wood (with a layer of water added to the wood) on which it was placed. This reaction seems to violate our idea that a reaction proceeds in a direction in which heat is liberated. Reactions, which liberate heat and raise the temperature of the surroundings, are called **exothermic** reactions. Reactions,

which absorb heat from the surroundings and hence lower the temperature of the surroundings, are **endothermic** reactions. To answer the question we need to consider entropy.

1.3.5: A review of thermodynamics

You may remember from General Chemistry that the change in the internal energy of a system, ΔE, is given by:

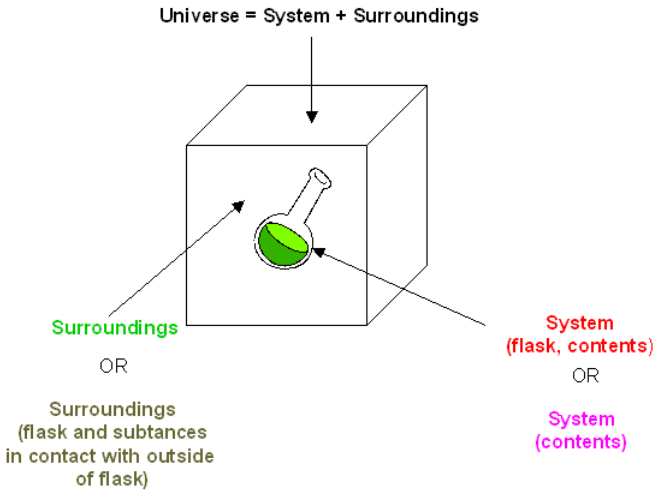
$$\begin{aligned} \Delta E_{sys} &= q + w \\ &= q - P_{ext}\,\Delta V \end{aligned} \tag{1.3.8}$$

where q is the heat (thermal energy) transferred to (+) or from the system (-), w is the work done on (+) or by (-) the system. This is one expression for the 1st Law of Thermodynamics

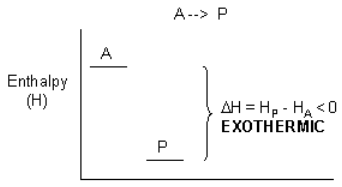
If only pressure/volume (PV) work is done (and not electrical works for example), $w = -P_{ext}\,\Delta V$, where P_{ext} is the external pressure resisting a volume change in the system, ΔV. Under these conditions, the heat transfer at constant P, q_P is given by:

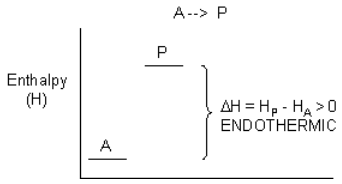
$$\begin{aligned} \Delta \mathrm{E_{sys}} - \mathrm{w} &= \Delta \mathrm{E_{sys}} + \mathrm{P_{ext}}\,\Delta \mathrm{V} \\ &= \mathrm{q_P} \\ &= \Delta \mathrm{H_{sys}} \end{aligned} \tag{1.3.9}$$

q_p, which can easily be measured in a coffee cup calorimeter, is equal to the **change in enthalpy**, ΔH, of the system.



For exothermic reactions, the reactants have more thermal energy than the products, and the heat energy (measured in kilocalories or kilojoules) released is the difference between the energy of the products and reactants. When heat energy is used to measure the difference in energy, we call the energy enthalpy (H) and the heat released as the change in enthalpy (ΔH), as illustrated below.





For exothermic reactions, $\Delta H < 0$. For endothermic reactions, $\Delta H > 0$.

The equation $\Delta E_{sys} = q + w = q - P_{ext}\Delta V$ is one expression of the First Law of Thermodynamics. Another statement of energy conservation, is:

$$\Delta E_{tot} = \Delta E_{universe} = \Delta E_{sys} + \Delta E_{surrounding} = 0 \tag{1.3.10}$$

Clearly, there must be something more that decides whether a reaction goes to a significant extent other tha, if heat is released from the system. That is, the spontaneity of a reaction must depend on more than just ΔH_{sys}. . Another example of a spontaneous natural reaction is the evaporation of water (a physical, not chemical process).

$$H_2O(l) \rightarrow H_2O(g)$$

Heat is absorbed from the surroundings to break the intermolecular forces (H bonds) among the water molecules (the system), allowing the liquid to be turned into a gas. If the surroundings are the skin, evaporation of water in the form of sweat cools the body. What are these reactions spontaneous and essentially irreversible even though they are endothermic? Notice that in both of these endothermic reactions (the reactions of $Ba(OH)_2.8H_2O(s)$ and $2NH_4SCN(s)$ and the evaporation of water), the products are more disorganized (more disordered) than the products. A solid is more ordered than a liquid or gas, and a liquid is more ordered than a gas. In nature, ordered things become more disordered with time. Entropy (S), the other factor (in addition to enthalpy changes) is often considered to be a measure of the disorder of a system. The greater the entropy, the greater the disorder. For reactions that go from order (low S) to disorder (high S), the change in S, $\Delta S > 0$. For the reaction that goes from low order to high order, $\Delta S < 0$.

Caution

However, this common description of entropy is quite misleading. Macroscopic examples describing order/disordered states (such as the cleanliness of your room or the shuffling of a deck of cards) are inappropriate since entropy deals with microscopic states.

The driving force for spontaneous reactions is the dispersion of energy and matter. Increases in entropy for reactions that involve matter occurs when gases or solutes in solution are dispersed, leading to increases in positional entropy. For reactions involving energy changes, entropy increases when energy is dispersed as random, undirected thermal motion, leading to increases in thermal entropy. In this sense, entropy, S (a measure of ("spreadedness") is a measure of the number of different ways (microstates) that particles or energy can be arranged (W), not a measure of disorder! W is an abbreviation for the German word, ***Wahrscheinlichkeit***, which means probability. It can be shown that for a solute dissolving in a solvent, $W_{sys} = W_{solute} \times W_{solvent}$. Note that this is a multiplicative function. Entropy is a logarithmic function of W which allows additivity of solute and solvent W values, a feature found in other thermodynamic state functions like ΔE, ΔH, and ΔS. Hence

$$\ln W \text{ sys} = \ln W_{\text{solute}} + \ln W_{\text{solvent}} \tag{1.3.11}$$

Boltzmann showed that for molecules,

$$S = k \ln W \tag{1.3.12}$$

where k is the Boltzmann constant (1.68×10^{-23} J/K), S units: J/K

or

$$S = kN_A \ln W = R \ln W \tag{1.3.13}$$

Boltzmann realized the connection between the macroscopic entropy of a system and the microscopic order/disorder of a system through the equation $S = k \ln W$, Increasing S (macroscopic property) occurs with increasing numbers of possible microscopic states for the atoms and molecules of a system.

The dissolution of a solute in water and the expansion of a gas into a vacuum, which both proceed spontaneously toward an increase in matter dispersal, are examples of familiar processes characterized by a $\Delta S_{sys} > 0$. We will see in future chapters that entropy changes in the solvent, solutes, and in a protein are critical determinants of protein folding.

The spontaneity of exothermic and endothermic processes will depend on the

$$\Delta S_{tot} = \Delta S_{surr} + \Delta S_{sys} \tag{1.3.14}$$

ΔS_{sys} often depends on matter dispersal (positional entropy). ΔS_{surr} depends on energy changes in the surroundings, $\Delta H_{surr} = -\Delta H_{sys}$ (thermal entropy).

It is more convenient to express thermodynamic properties based on the system which is being studied, not on the surrounding. This can be readily done for the ΔS_{surr} which depends both on ΔH_{sys} and the temperature. First consider the dependency on ΔH_{sys}. Thermal energy flows into or out of the system, and since $\Delta H_{sys} = - \Delta H_{surr}$,

ΔS_{surr} is proportional to $-\Delta H_{sys}$

- For an exothermic reaction, $\Delta S_{surr} > 0$ (since $\Delta H_{sys} < 0$) and the reaction is favored;
- For an endothermic reaction, $\Delta S_{surr} < 0$, (since $\Delta H_{sys} > 0$), and the reaction is disfavored;

ΔS_{surr} also depends on the temperature T of the surroundings:

ΔS_{surr} is proportional to 1/T

If the T_{surr} is high, a given heat transfer to or from the surroundings will have a smaller effect on the ΔS_{surr}. Conversely, if the T_{surr} is low, the effect on ΔS_{surr} will be greater. (Atkins uses the analogy of the effect of a sneeze in library compared to in a crowded street; An American Chemistry General Chemistry text uses the analogy of giving $5 to a friend with $1000 compared to one who has just $10.) Hence,

$$\Delta S_{\text{surr}} = \frac{-\Delta \text{H}_{\text{sys}}}{\text{T}} \tag{1.3.15}$$

(Note: from a more rigorous thermodynamic approach, entropy can be determined from \(dS = dq_{rev}/T\).)

Once again,

$$\Delta S_{\text{tot}} = \Delta S_{\text{surr}} + \Delta S_{\text{sys}} \tag{1.3.16}$$

ΔS_{tot} depends on both enthalpy changes in the system and entropy changes in the surroundings. Hence,

$$\Delta S_{\text{tot}} = \frac{-\Delta \text{H}_{\text{sys}}}{\text{T}} + \Delta \text{S}_{\text{sys}} \tag{1.3.17}$$

Multiplying both sides by $-T$ gives

$$-\text{T}\Delta \text{S}_{\text{tot}} = \Delta \text{H}_{\text{sys}} + \text{T}\Delta \text{S}_{\text{sys}} \tag{1.3.18}$$

The thermodynamic function Gibb's Free Energy, G, can be defined as:

$$G = H - TS \tag{1.3.19}$$

At constant T and P,

$$\Delta G = \Delta H - T\Delta S \tag{1.3.20}$$

Hence

$$\Delta G_{sys} = \Delta H_{sys} - T\Delta S_{sys} = -T\Delta S_{tot} \tag{1.3.21}$$

Spontaneity is determined by ΔS_{tot} OR ΔG_{sys} since $\Delta S_{tot} = -\Delta G_{sys}/T$. ΔG_{sys} is widely used in discussing spontaneity since it is a state function, depends only on the enthalpy and entropy changes in the system, and is negative (as is the potential energy change for a falling object) for all spontaneous processes.

The second law of thermodynamics can be succinctly stated: For any spontaneous process, the $\Delta S_{tot} > 0$. Unlike energy (from the First Law), entropy is not conserved.

1.4: Genetic Foundations

1.4.1: Introduction

The development of complex biological organisms on our planet has arisen through the evolutionary mechanism of ***natural selection***. The British naturalist, Charles Darwin proposed the theory of biological evolution by natural selection in his book, '*On the Origins of Species*' that was published in 1859. Darwin defined ***evolution*** as "descent with modification," the idea that species change over time, give rise to new species, and share a common ancestor. The mechanism that Darwin proposed for evolution is ***natural selection***. Because resources are limited in nature, organisms with heritable traits that favor survival and reproduction will tend to leave more offspring than their peers, causing the traits to increase in frequency within a population over generations. Thus, natural selection causes populations to become ***adapted***, or increasingly well-suited, to their environments over time. Natural selection depends on the environment and requires existing heritable variation in a group.

Natural selection acts on an organism's ***phenotype***, or physical characteristics. ***Phenotype*** is determined by an organism's genetic make-up (***genotype***) and the environment in which the organism lives. When different organisms in a population possess different versions of a gene for a certain trait, each of these versions is known as an ***allele***. It is primarily this genetic variation that underlies differences in phenotype. Some traits are governed by only a single gene, but most traits are influenced by the interactions of many genes. A variation in one of the many genes that contribute to a trait may have only a small effect on the phenotype; together, these genes can produce a continuum of possible phenotypic values.

For example, interactions between different equine coat color genes determine a horse's coat color. Many colors are possible, but all variations are produced by changes in only a few genes. *Extension* and *agouti* are particularly well-known genes with dramatic effects. For example, differences at the *agouti gene* can help determine whether a horse is bay or black in coloration, and a change to the *extension gene* can in turn make a horse chestnut-colored instead (Figure 1.30). Yet other gene variants are responsible for the myriad of other coat color possibilities, including palomino, buckskin, and cremello horses.

Figure 1.30 Genotype Variations as Determinants of Horse Coat Color. Horses that are capable of producing the black pigment, eumelanin, have at least one copy of the dominant *extension gene* (*E/E or E/e*). Interestingly, the *agouti gene* controls the restriction of true black pigment (eumelanin) in the coat. Horses expressing an *extension dominant gene, and* are recessive at the *agouti gene locus (a/a*) will be black in color, as shown in *(a)*. Whereas horses that are dominant for *extension (E/E or E/a)* but are also dominant for the *agouti genotype* (*A/A* or *A/a)*, will never be fully black. Depending on other gene loci, they will instead show coloration patterns such as bay, *as shown in (b). Image (a)* provided by: Serendipityblue; *Image (b)* provided by: CMSporthorses

Thus, the primary molecular mechanism that drives natural selection is controlled by the heritability and mutability of genetic traits housed in the major macromolecule, deoxyribonucleic acid (DNA). In chapter 4, you will learn about the structural characteristics of DNA, whereas chapter 9 focuses on the biochemical mechanisms involved with DNA replication and also details the importance of DNA repair process and molecular mechanisms of evolution at the genetic level.

1.4.2: Genetic Code

Notably, the phenotypic traits determined by the genetic makeup of an organism are not controlled directly by the genetic material, DNA, but by the proteins that are produced from the information housed within the gene. In 1945, geneticist George Beadle proposed the one gene-one enzyme hypothesis suggesting that genes are highly specific when they encode for a protein sequence. However, it would take 16 more years before the biochemical nature of this process was deduced. Efforts to understand how proteins are encoded began after DNA's structure was discovered in 1953. George Gamow postulated that sets of three bases must be employed to encode the 20 standard amino acids used by living cells to build proteins, which would allow a maximum of 4^3 = 64 amino acids.

The Crick, Brenner, Barnett and Watts-Tobin experiment first demonstrated that codons consist of three DNA bases (Figure 1.31). Marshall Nirenberg and Heinrich J. Matthaei were the first to reveal the nature of a codon in 1961.

Figure 1.31 Codons Consist of Sets of Three Bases. A series of codons in part of a messenger RNA (mRNA) molecule. Each codon consists of three nucleotides, usually corresponding to a single amino acid. The nucleotides are abbreviated with the letters A, U, G, C. This is mRNA which uses U (uracil). DNA uses T (thymine) instead. This mRNA molecule will instruct a ribosome to synthesize a protein according to this code. Image by Sverdrup

They used a cell-free system to translate a poly-uracil RNA sequence (i.e., UUUUU…) and discovered that the polypeptide that they had synthesized consisted of only the amino acid phenylalanine. They thereby deduced that the codon UUU specified the amino acid phenylalanine.

This was followed by experiments in Severo Ochoa's laboratory that demonstrated that the poly-adenine RNA sequence (AAAAA…) coded for the polypeptide poly-lysine and that the poly-cytosine RNA sequence (CCCCC…) coded for the polypeptide poly-proline. Therefore, the codon AAA specified the amino acid lysine, and the codon CCC specified the amino acid proline. Using various copolymers most of the remaining codons were then determined.

Subsequent work by Har Gobind Khorana identified the rest of the genetic code. Shortly thereafter, Robert W. Holley determined the structure of transfer RNA (tRNA), the adapter molecule that facilitates the process of translating RNA into protein. This work was based upon Ochoa's earlier studies, yielding the latter the Nobel Prize in Physiology or Medicine in 1959 for work on the enzymology of RNA synthesis.

Extending this work, Nirenberg and Philip Leder revealed the code's triplet nature and deciphered its codons (Figure 1.32). In these experiments, various combinations of mRNA were passed through a filter that contained ribosomes, the components of cells that translate RNA into protein. Unique triplets promoted the binding of specific tRNAs to the ribosome. Leder and Nirenberg were able to determine the sequences of 54 out of 64 codons in their experiments. Khorana, Holley and Nirenberg received the 1968 Nobel for their work.

The three stop codons were named by discoverers Richard Epstein and Charles Steinberg. "Amber" was named after their friend Harris Bernstein, whose last name means "amber" in German. The other two stop codons were named "ochre" and "opal" in order to keep the "color names" theme.

Figure 1.32 The Genetic Code. *Image edited by Seth Miller, Original file designed and produced by: Kosi Gramatikoff courtesy of Abgent*

Each gene contains a reading frame is defined by the initial triplet of nucleotides from which translation starts. It sets the frame for a run of successive, non-overlapping codons, which is known as an ***open reading frame (ORF)***. For example, the string 5′-AAATGAACG-3′, if read from the first position, contains the codons AAA, TGA, and ACG ; if read from the second position, it contains the codons AAT and GAA ; and if read from the third position, it contains the codons ATG and AAC. Every sequence can, thus, be read in its 5′ → 3′ direction in three reading frames, each producing a possibly distinct amino acid sequence: in the given example, Lys (K)-Trp (W)-Thr (T), Asn (N)-Glu (E), or Met (M)-Asn (N), respectively. When DNA is double-stranded, six possible reading frames are defined, three in the forward orientation on one strand and three reverse on the opposite strand. Protein-coding frames are defined by a start codon, usually the first AUG (ATG) codon in the RNA (DNA) sequence.

To terminate the translation process, there are three stop codons: UAG is *amber*, UGA is *opal* (sometimes also called *umber*), and UAA is *ochre*. Stop codons are also called “termination” or “nonsense” codons. They signal the release of the nascent polypeptide from the ribosome.

Mutations

During the process of DNA replication, errors occasionally occur in the polymerization of the second strand. These errors, called mutations, can affect an organism’s phenotype, especially if they occur within the protein-coding sequence of a gene. Error rates are typically 1 error in every 10–100 million bases—due to the “proofreading” ability of DNA polymerases.

Missense mutations and ***nonsense mutations*** are examples of point mutations that can cause genetic diseases such as sickle-cell disease and thalassemia respectively. Clinically important ***missense mutations*** generally change the properties of the coded

amino acid residue among basic, acidic, polar or non-polar states, whereas ***nonsense mutations*** result in a stop codon.

Mutations that disrupt the reading frame sequence by **indels** (insertions or deletions) of a non-multiple of 3 nucleotide bases are known as ***frameshift mutations***. These mutations usually result in a completely different translation than from the original RNA, and likely cause a stop codon to be read, which truncates the protein. These mutations may impair the protein's function and are thus rare in *in vivo* protein-coding sequences. One reason inheritance of frameshift mutations is rare is that, if the protein being translated is essential for growth under the selective pressures the organism faces, the absence of a functional protein may cause death before the organism becomes viable. Frameshift mutations may result in severe genetic diseases such as Tay–Sachs disease.

Although most mutations that change protein sequences are harmful or neutral, some mutations have benefits. These mutations may enable the mutant organism to withstand particular environmental stresses better than wild-type organisms, or reproduce more quickly. In these cases, a mutation will tend to become more common in a population through natural selection. Different sequence variations of the same gene or protein within a single organism, within a population, or between different species are known as ***sequence polymorphisms***. Larger-scale gene duplication events can also lead to evolutionary events.

1.4.3: Similar Proteins

The evolution of proteins is studied by comparing the sequences and structures of proteins from many organisms representing distinct evolutionary clades. If the sequences/structures of two proteins are similar indicating that the proteins diverged from a common origin, these proteins are called ***homologous proteins***. More specifically, homologous proteins that exist in two distinct species are called as ***orthologs***. In contrast, homologous proteins encoded by the genome of a single species are called ***paralogs***. Unrelated genes that have separate evolutionary origins, but that each encode proteins that have similar functions, are termed ***analogs*** (Figure 1.33).

Figure 1.33 Genetic Evolution of Protein Sequences. *(Upper Panel)* An ancestral gene dupliates to produce two paralogs (Gene A and B). A speciation event produces orthologs in the two daughter species. In a separate species, an unrelated gene has a similar function (Gene C) but has a separate evolutionary origin and so is an analog. *(Lower Panel)* 3-D protein models were retrieved or modeled using SWISS-MODEL: Human Histone H1.1 (Q02539), Human Histone H1.2 (P16403), E. coli HNS (P0ACF8). Histone H1.1 from the chimpanzee (*Pan troglodytes* XP_016810512.1) was modeled using Human Histone H1.1 as a template. Note that the *E. coli* HNS protein is typically modeled as a dimer. Only a single monomer is shown here. Upper Image by Thomas Shafee

DNA sequencing techniques have rapidly improved over the last 15 to 20 years making it possible to sequence the entire genomes of organisms and thus, predict the entire proteome of an organism, based on the translation of the sequenced genome followed by the annotation of predicted ORFs using phylogenetic comparison of similar genes/proteins from other known organisms. This has given rise to the field of ***Bioinformatics*** which uses computer science, mathematics, and statistical analysis to analyze the large quantities of biological data created in genome sequencing projects. The phylogenetic relationships, and hence ancestral relationships, of various genes, proteins, and ultimately organisms can be established through the statistical analysis of sequence alignments. Such phylogenetic trees have established that the sequence similarities among proteins reflect closely the evolutionary relationships among organisms.

Protein evolution describes the changes over time in protein shape, function, and composition. Through quantitative analysis and experimentation, scientists have strived to understand the rate and causes of protein evolution. Using the amino acid sequences of hemoglobin and cytochrome c from multiple species, scientists were able to derive estimations of protein evolution rates. What

they found was that the rates were not the same among proteins. Each protein has its own rate, and that rate is constant across phylogenies (i.e., hemoglobin does not evolve at the same rate as cytochrome c, but hemoglobins from humans, mice, etc. do have comparable rates of evolution.). Not all regions within a protein mutate at the same rate; functionally important areas mutate more slowly and amino acid substitutions involving similar amino acids occurs more often than dissimilar substitutions. Overall, the level of polymorphisms in proteins seems to be fairly constant. Several species (including humans, fruit flies, and mice) have similar levels of protein polymorphism.

Gene duplication events followed by mutation can also give rise to paralogs, with unique and different functions within an organism. This can make the annotation of genomes based on sequence alone a difficult task, as homologous protein sequences may not have similar functions *in vivo*. It is estimated that approximately 10-25% of annotations made on sequence homology are incorrect and require experimental validation. For example, human pancreatic ribonuclease is a digestive enzyme utilized to breakdown nucleic acids. The angiogenin protein is a paralog of pancreatic ribonuclease and shares high sequence homology and 3-D shape (Figure 1.34). However, the functions of these proteins are quite different. Angiogenin induces vascularization by activating transcriptional processes in endothelial cells. However, if the function of only one of these homologs was known, it would be easy to mistakenly hypothesize that the homologous protein would be similar in function. Thus, care must be taken when using bioinformatic tools to not overestimate the predictive ability of sequence alignments.

Figure 1.34 Homologous Proteins Do Not Always Have Homologous Functions. In the example above, the digestive enzyme, pancreatic ribonuclease is a paralog of the angiogenin protein and shares an ancestral origin. However, the functions of each of these proteins are quite divergent and have evolved such that they do not share homologous function. 3-D protein models were retrieved using SWISS-MODEL: Human Pancreatic Ribonuclease (P07998) and Human Angiogenin (P03950)

The control of gene expression is critical in all processes of life, allowing for the differentiation of cells to form different body structures and organs, as well as smaller more reversible changes that allow an organism to respond to different environmental situations and stimuli. In chapter 12, you will explore the major biochemical mechanisms used to control gene expression within cells. This will include the discussion of a fairly new and exciting field of study known as ***epigenetics***. In addition to the heritability of traits through the passage of genetic information, it is fast becoming clear that the environmental factors that an organism is exposed to throughout its life can affect gene expression without physically altering the DNA sequence, and that these changes in expression patterns can be long-lasting and can even be inherited in the following generations. The term ***epigenetics*** literally means ‘on top of’ or ‘in addition to’ genetics and focuses on the heritable gene expression patterns that are induced by the exposure or experience of an organism within its environment.

For example in human populations, stressful events such as starvation can have lasting imprints on children that are born under these conditions. These children have higher risks of obesity and metabolic disorders as adults, including the development of type II diabetes. In fact, these predispositions can be carried not only to the children born during starvation but also to their future children indicating that environmental events can affect gene expression patterns through multiple generations. In more controlled laboratory experiments using rats, it has been demonstrated that the more a mother rat licks and nurtures its offspring, the calmer and more relaxed the offspring will be as an adult. Mother rats that are less nurturing and ignore their young, have offspring that will grow up displaying higher levels of anxiety. These changes are not caused by genetic differences between the offspring, but rather by differences in gene expression patterns. In fact, calm and relaxed mice can be altered to show high anxiety by exposing

them to agents that alter gene expression patterns. Mechanisms controlling such heritable alterations in gene expression patterns will be covered in a future chapter.

1.4.4: Central Dogma of Biology

DNA encodes the genetic material. It must be replicated on cell division. Its information is decoded into an RNA in a process called transcription. That information is decoded to form a protein sequence. Collectively these processes are referred to as the Central Dogma of Biology. A variant occurs when RNA is decoded into DNA, a process called reverse transcription. These processes are described briefly below and in great depth in subsequent chapters.

Replication

DNA must be duplicated in a process called **replication** before a cell divides. The replication of DNA allows each daughter cell to contain a full complement of chromosomes.

Animation of Replication

Transcription and Splicing

For a given gene, only one strand of the DNA serves as the template for **transcription**. An example is shown below. The bottom (blue) strand in this example is the template strand, which is also called the minus (-) strand, or the sense strand. It is this strand that serves as a template for mRNA synthesis. The enzyme RNA polymerase synthesizes an mRNA in the 5' to 3' direction complementary to this template strand. The opposite DNA strand (red) is called the coding strand, the nontemplate strand, the plus (+) strand, or the antisense strand.

The easiest way to find the corresponding mRNA sequence (shown in green below) is to read the coding, nontemplate, plus (+), or antisense strand directly in the 5' to 3' direction substituting U for T.

```
5' T G A C C T T C G A A C G G G A T G G A A A G G 3'
3' A C T G G A A G C T T G C C C T A C C T T T C C 5'
```

```
5' U G A C C U U C G A A C G G G A U G G A A A G G 3'
```

As we've learned more about the structure of DNA, RNA, and proteins, it become clear that transcription and translation differ in eukaryotes and prokaryotes. Specifically, eukaryotes have intervening sequences of DNA (introns) within a given gene that separate coding fragments of DNA (exons). A primary transcript is made from the DNA, and the introns are sliced out and exons joined in a contiguous stretch to form messenger RNA which leaves the nucleus. Translation occurs in the cytoplasm. Remember, prokaryotes do not have a nucleus.

- Animation of Transcription
- Animation of mRNA Splicing

Translation

Information in a mRNA sequence is decoded to form a protein. In this process, a triplet of nucleotides (a codon) in the RNA has the information of a single amino acid. Translation occurs on a large RNA-protein complex called the ribosome. An intermediary transfer RNA (tRNA) molecule becomes covalently linked to a single amino acid by the enzyme tRNA-acyl synthetase. This "charged" tRNA binds through a complementary anticodon region to the triplet codon in the tRNA. The ribosome/tRNA complex ratchets down the mRNA allowing a new "charged" tRNA complex to bind at an adjacent site. The two adjacent amino acids form a peptide bond in a process driven by ATP cleavage. This process repeats until a "stop" codon appears in the mRNA sequence. The genetic code shows the relationship between the triplet mRNA codon and the amino acid which corresponds to it in the growing peptide chain.

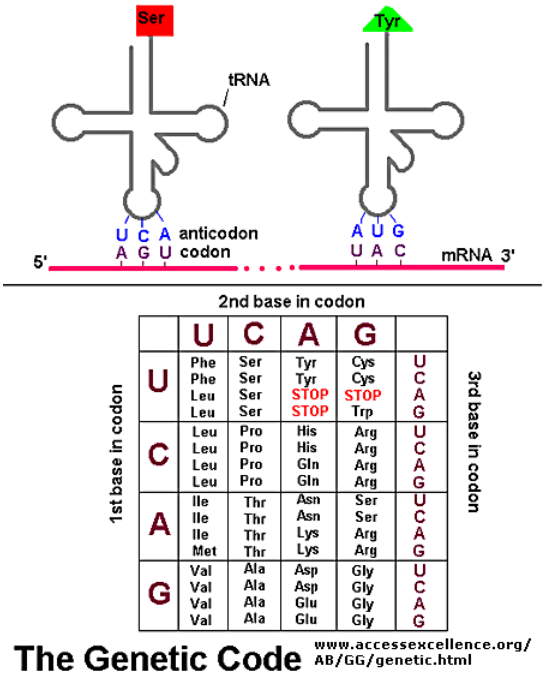


Figure 1.4.1: Codon:Anticodon interactions between mRNA and tRNA. (Copyright; author via source)

As was mentioned in the Protein Chapter (amino acid section) two other amino acids occasionally appear in proteins (excluding amino acids altered through post-translational modification. One is selenocysteine, which is found in Arachea, eubacteria, and animals. The other is just recently found is pyrrolysine, found on Arachea. These new amino acids derive from modifications of Ser-tRNA and probably Lys-tRNA after the tRNA is charged with the normal amino acid, to produce selenocys-tRNA and pyrrolys-tRNA, respectively. The pyrrolysine-tRNA recognizes the mRNA codon UAG which is usually a stop codon, while selenocys-tRNA recognizes UGA, also a stop codon. Clearly, only a small fraction of stop codons in mRNA sequences would be recognized by this usual tRNA complex. What determines that recognition is unclear.

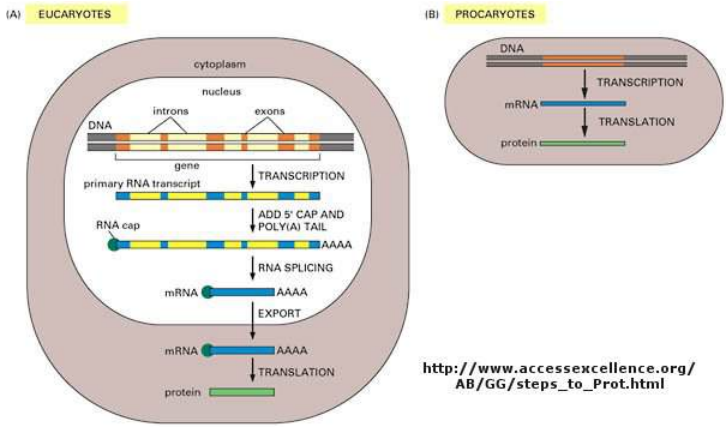


Figure: Central Dogma Differences in Eukaryotes and Prokaryotes

Animation of Translation

1.4.5: What is a gene?

The definition of a gene can differ depending on whom you ask. The world gene has literally become a cultural icon of our day. Can our genes explain what it is to be human? The definition of a gene has changed with time. Eukaryotic genes contain exons (coding regions) and introns (intervening sequences) that are transcribed to produce a primary transcript. In a post-transcriptional process, introns are spliced out by the spliceosome, to produce a messenger RNA, mRNA, which is translated into a protein sequence. (See diagram above).

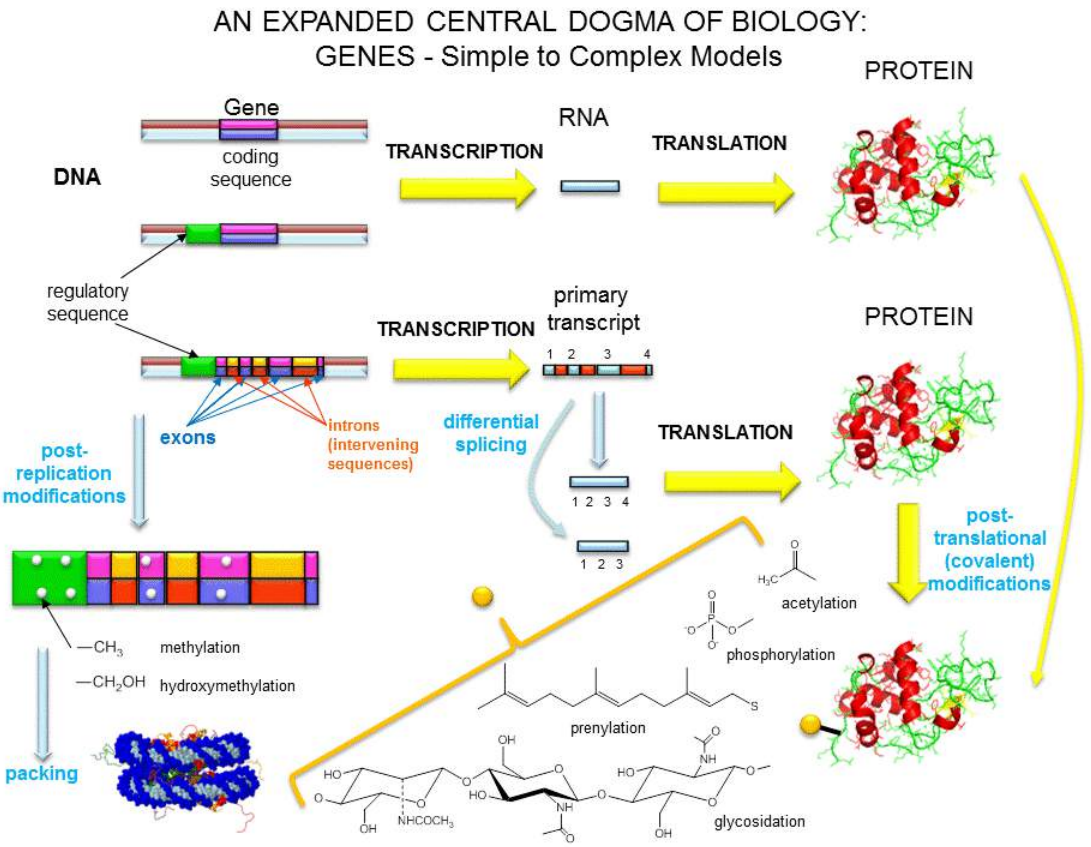


Figure 1.4.1: Figure: A view of genes and their products: Simplicity to Complexity

Over the last 100 years, as our understanding of biochemistry has increased, the definition of a gene has evolved from

- the basis of inheritable traits
- certain regions of chromosomes
- a segment of a chromosome that produces one enzyme
- a segment of a chromosome that produces one protein
- a segment of a chromosome that produces a functional product

The last definition was necessary since some gene products that have functions (structural and catalytic) are RNA molecules. The last definition also includes regulatory regions of the chromosome involved in transcriptional control. Snyder and Gerstein have developed five criteria that can be used in gene identification which is important as the complete genomes of organisms are analyzed for genes.

1. identification of an open reading frame (ORF) - this would include a series of codons bounded by start and stop codons. This gets progressively harder to do if the gene has a large number of exons embedded in long introns.
2. specific DNA features within genes - these would include a bias towards certain codons found in genes or splice sites (to remove intron RNA)
3. comparing putative gene sequences for homology with known genes from different organisms, but avoiding sequences that might be conserved regulatory regions.
4. identification of RNA transcripts or expressed protein (which does not require DNA sequence analysis as the top three steps do) -
5. inactivating (chemically or through specific mutagenesis) a gene product (RNA or protein).

New findings make it even more complicated to define a gene, especially if the transcripts of a "gene region" are studied. Cheng et al studied all transcripts from 10 different human chromosomes and 8 different cell lines. They found a large number of different transcripts, many of which overlapped. Splicing often occurs between nonadjacent introns. Transcripts were found from both strands and were from regions containing introns and exons. Other studies found up to 5% of transcripts continued through the end of "gene" into other genes. 63% of the entire mouse genome, which is comprised of only 2% exons, is transcribed.

1.4.6: The Language of DNA

In this short chapter, you will briefly learn how modern molecular biologists manipulate DNA, the blueprint for all of life. The details will be found in subsequence chapters. The four-letter alphabet (A, G, C, and T) that makes up DNA represents a language that when transcribed and translated leads to the myriad of proteins that make us who we are as a species and as individuals. Let's continue with the metaphor that DNA is a language. To master that language, as with any other language, we need to be able to read, write, copy, and edit that language. If you were using a word processor to find one line in a hundred-page document or one article from one book out of the Library of Congress, you would also need a way to search the large print base available. You might want to compare two different copies of files to see if they differ from each other. From the lab and this online discussion and problem set, you will learn how modern scientists read, write, copy, edit, search, and compare the language of the genome. These abilities, acquired over the last twenty years, have revolutionized our understanding of life and have given us the potential to alter, for good or evil, life itself.

DNA in human chromosomes exists as one long double-stranded molecule. It is too long to physically study and manipulate in the lab. Using a battery of enzymes, the DNA of chromosomes can be chemically cleaved into smaller fragments, which are more readily manipulable. (Similar techniques are used to sequence proteins, which require overlapping polypeptide fragments to be made.) After the fragments have been made, they must be separated from each other in order to study them. DNA fragments can be separated on the basis of some structural feature that differentiates the fragments from each other. Polarity can not be used since all DNA fragments have negatively charged phosphates in the sugar-phosphate backbone of the molecule. Although each fragment would have a unique sequence, it would be hard to separate all the different fragments, by, for instance, attaching some molecule that binds to a unique sequence in the major groove of a given fragment to a big bead and using that bead to separate out that one unique fragment. You would need a different bead for each unique fragment! The best way to separate the fragments from each other is to base the separation on the actual size of the fragment by using electrophoresis on an agarose or polyacrylamide gel.

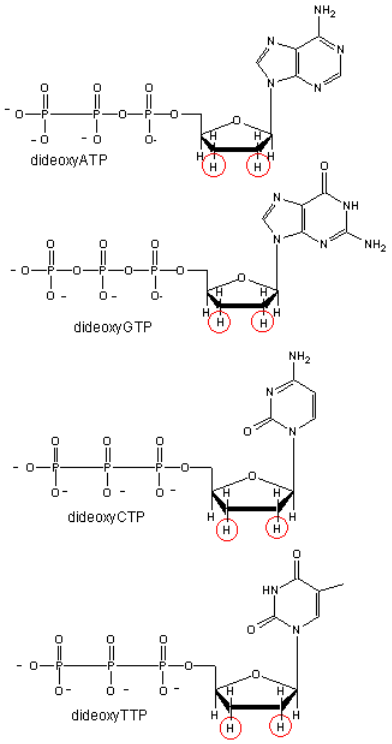
A carbohydrate extract called agarose is made from algae. Water is added to the extract, which is then heated. The carbohydrate extract dissolves in the water to form a viscous solution. The agarose solution is poured into a mold (like warm jello) and is allowed to solidify. A plastic comb with wide teeth was placed in the agarose when it was still liquid. When the agarose is solid, the comb can be removed, leaving in its place little wells. A solution of DNA fragments can be placed in the wells. The agarose slab with the sample is covered with a buffer solution and electrodes are placed at each end of the slab. The negative electrode is placed near the well-end of the agarose slab while the positive electrode is placed at the other end. If a voltage is applied across the agarose slab, the negatively charged DNA fragments will move through the agarose gel toward the positive electrode. This migration of charged molecules in solution toward an oppositely charged electrode is called electrophoresis. Pretend you are one of the fragments. To you, the gel looks like a tangled cobweb. You sneak your way through the openings in the web as you move straight forward to the positive electrode. The larger the fragment, the slower you move because it is hard to get through the tangled web. Conversely, the shorter the fragment, the faster you move. Using this technique and its many modifications, oligonucleotides differing by just one nucleotide can be separated from each other. In the electrophoresis of DNA fragments, a fluorescent, uncharged dye, ethidium bromide, is added to the buffer solution. This dye literally intercalates -between the base pairs of DNA, which imparts a fluorescent yellow-green color to the DNA when UV light is shown on the agarose gel.

Reading DNA

We will discuss one method of reading the sequence of DNA. This method, developed by Sanger won him a second Nobel prize. To sequence a single-stranded piece of DNA, the complementary strand is synthesized. Four different reaction mixtures are set up. Each contains all 4 radioactive deoxynucleotides (dATP, dCTP, dGTP, dTTP) required for the reaction and DNA polymerase. In addition, dideoxyATP (ddATP) is added to one reaction tube The dATP and ddATP attach randomly to the growing 3' end of the complementary stranded. If ddATP is added no further nucleotides can be added after since its 3' end has an H and not a OH. That's why they call it dideoxy. The new chain is terminated. If dATP is added, the chain will continue to grow until another A needs to be added. Hence a whole series of discreet fragments of DNA chains will be made, all terminated when ddATP was added. The same scenario occurs for the other 3 tubes, which contain dCTP and ddCTP, dTTP and ddTTP, and dGTP and ddGTP respectively. All the fragments made in each tube will be placed in separate lanes for electrophoresis, where the fragments will separate by size.

Didexoynucleotides

Figure: Didexoynucleotides



PROBLEM: You will pretend to sequence a single-stranded piece of DNA as shown below. The new nucleotides are added by the enzyme DNA polymerase to the primer, GACT, in the 5' to 3' direction. You will set up 4 reaction tubes, Each tube contains all the dXTP's. In addition, add ddATP to tube 1, ddTTP to tube 2, ddCTP to tube 3, and ddGTP to tube 4. For each separate reaction mixture, determine all the possible sequences made by writing the possible sequences on one of the unfinished complementary sequences below. Cut the completed sequences from the page, determine the size of the polynucleotide sequences made, and place them as they would migrate (based on size) in the appropriate lane of an imaginary gel, which you have drawn on a piece of paper. Lane 1 will contain the nucleotides made in tube 1, etc. Then draw lines under the positions of the cutout nucleotides to represent DNA bands in the gel. Read the sequence of the complementary DNA synthesized. Then write the sequence of the ssDNA that was to be sequenced.

5' T C A A C G A T C T G A 3' (STAND TO SEQUENCE)

3' G A C T 5' (primer)

3' G A C T 5' (primer)

3' G A C T 5' (primer)

3' G A C T 5' (primer)

3' G A C T 5' (primer)

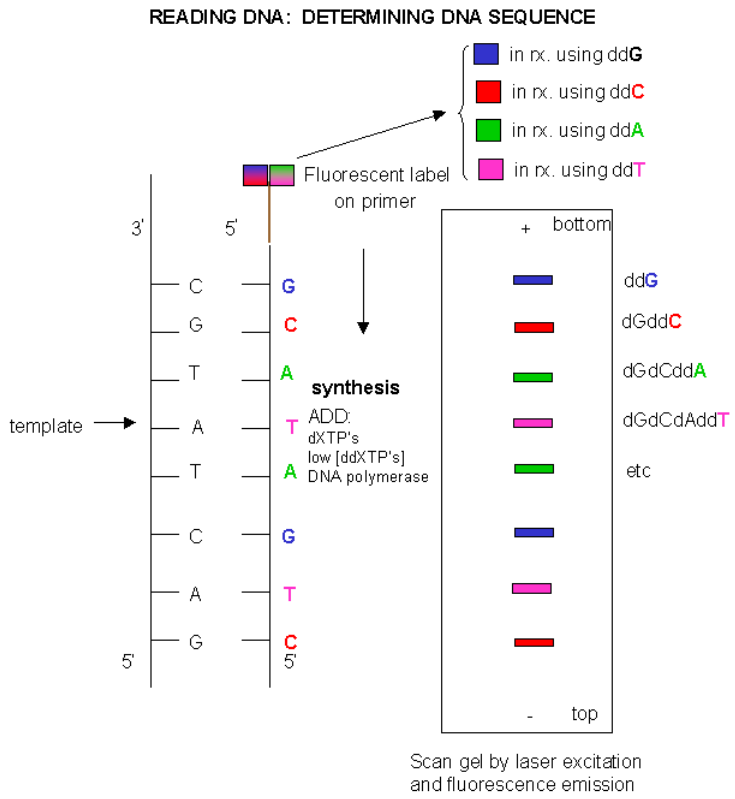
3' G A C T 5' (primer)

3' G A C T 5' (primer)

3' G A C T 5' (primer)

Since the DNA fragments have no detectable color, they can not be directly visualized in the gel. Alternative methods are used. In the one described above, radiolabeled ddXTP's were used. Once the sequencing gel is run, it can be dried and the bands visualized by radioautography (also called autoradiography). A place of x-ray film is placed over the dried gel in a dark environment. The radiolabeled bands will emit radiation which will expose the x-ray film directly over the bands. The film can be developed to detect the bands. In a newer technique, the primer can be labeled with a fluorescent dye. If a different dye is used for each reaction mixture, all the reaction mixtures can be run in one lane of a gel. (Actually, only one reaction mix containing all the ddXTP's together is performed.) The gel can then be scanned by a laser, which detects fluorescence from the dyes, each at a different wavelength.

Figure: DNA sequencing using different fluorescent primers for each ddXTP reaction



One recent advance in sequencing allows for real-time determination of a sequence. The four deoxynucleotides are each labeled with a different fluorophore on the 5' phosphate (not the base as above). A tethered DNA polymerase elongates the DNA on a template, releasing the fluorophore into solution (i.e. the fluorophore is not incorporated into the DNA chain). The reaction takes place in a visualization chamber called a zero mode waveguide which is a cylindrical metallic chamber with a width of 70 nm and a volume of 20 zeptoliters (20 x 10^{-21} L). It sits on a glass support through which laser illumination of the sample is achieved. Given the small volume, non-incorporated fluorescently tagged deoxynucleotides diffuse in and out in the microsecond timescale. When a deoxynucleotide is incorporated into the DNA, its residence time is in the millisecond time scale. This allows for prolonged detection of fluorescence which gives a high signal-to-noise ratio. Newer technology in which sequence is done by moving DNA through pores in membranes could bring sequencing down to $1000/genome or less.

- Animation of Sanger Sequencing
- Nanopore sequencing

Writing DNA

Oligonucleotides can be synthesized on a solid bead. By adding one nucleotide at a time, the sequence and length of the oligonucleotide can be controlled.

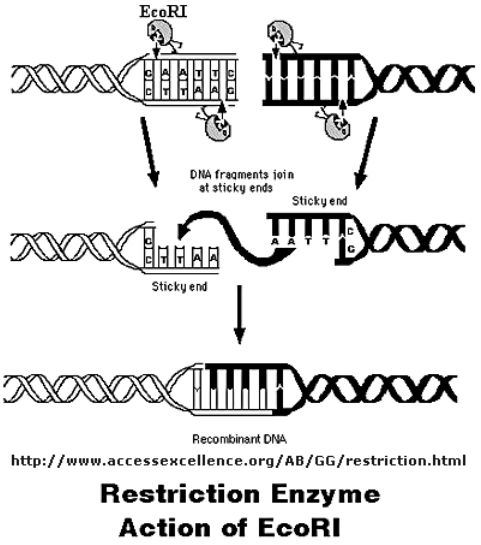
Copying DNA

Several methods exist for copying a sequence of DNA millions of times. Most methods make use of plasmids (which are found in bacteria) and viruses (which can infect any cell). The DNA of the plasmid or virus is engineered to contain a copy of a specific DNA sequence of interest. The plasmid or virus is then reintroduced into the cell where amplification occurs.

Initially, a DNA containing a gene or regulatory sequence of interest is cut at specific places with an enzyme called a restriction endonuclease, or restriction enzyme for short. The enzyme doesn't cleave DNA anywhere, but rather at "restricted" places in the sequence, much as an endoprotease cleaves a protein after a given amino acid within a protein chain. Instead of cleaving one strand, as in proteins, the restriction endonuclease must cleave both strands of dsDNA. It can cut the strands cleanly to leave blunt

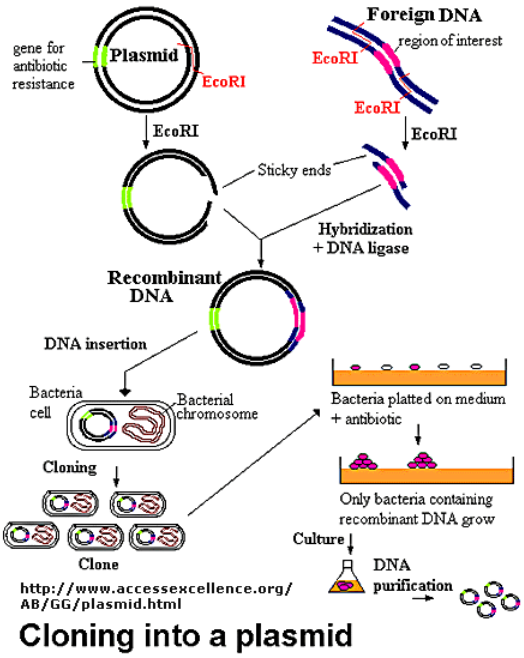
ends, or in a staggered fashion, to leave small tails of ssDNA. Multiple such sites exist at random in the genome. The gene of interest must be flanked on either side by such a sequence. The same enzyme is used to cleave the plasmid or virus DNA.

Figure: Cleaving DNA with the Restriction Enzyme EcoR1



The foreign fragment of DNA can then be added to the plasmid or viral DNA as shown to make a recombinant DNA molecule. This technique of DNA cloning is the basis for the entire field of recombinant DNA technology.

Figure: Cloning a Restriction Fragment into a Plasmid



Animation of Gene Splicing

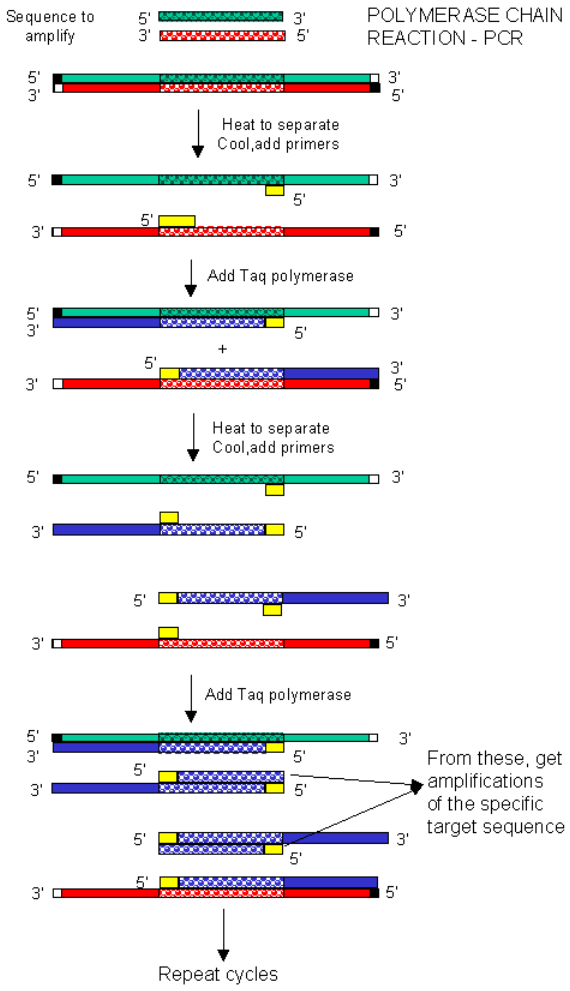
The plasmid can be added to bacteria, which take it up in a process called transformation. The plasmid can be replicated in the bacteria which will copy the DNA fragment of interest. Typically the plasmid carries a gene that can make the bacteria resistant to an antibiotic. Only bacteria that carry the plasmid (and presumably the insert) will grow. To isolate the desired fragment, the plasmids are isolated from bacteria, and cleaved with the same restriction enzyme to remove the desired fragment, after which it can be purified. In addition, the bacteria can be induced to express the protein from the foreign gene. In lab 4, we will transform

bacteria with a plasmid containing the gene for human adipoctye acid phosphatase beta, HAAP-B, and induce expression of the gene.

A similar method can be used to copy DNA in which the foreign fragment is recombined with the DNA of bacteriophage, a virus, which infects bacteria like E. Coli. The recombinant DNA can be packaged into actual viruses, as shown below. When the virus infects the bacteria, it instructs the cells to make millions of new viruses, hence copying the foreign fragment of interest.

Sometimes, "cloning" or copying a fragment of DNA is not what an investigator really wants. If the genomic DNA comes from a human cell, for instance, the gene will contain introns. If you put this DNA into a plasmid or bacteriophage, the introns go with it. Bacteria can replicate this DNA, but often one wants not to just copy (amplify) the DNA but also transcribe it into RNA and then translate it into protein. Bacteria, however, can not splice out the intron RNA, so mature mRNA can not be made. If one could clone into the bacteria's DNA without the introns, this problem would not exist. One such possible method exists in which you start with the actual mRNA for a protein of interest. In this technique, a dsDNA copy is made from a ss-mRNA molecule. Such dsDNA is called cDNA, for complementary or copy DNA. This can then be cloned into a plasmid or bacteriophage vector and amplified as described above.

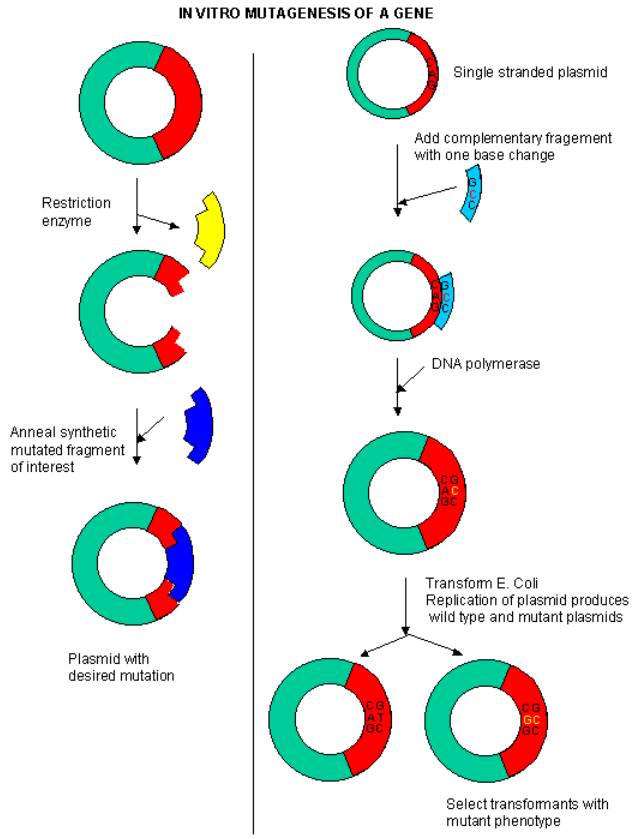
In the mid '80s a new method was developed to copy (amplify) DNA in a test tube. It doesn't require a plasmid or a virus. It just requires a DNA fragment, some primers (small oligonucleotides complementary to sections of DNA on each strand and straddling the section of DNA to be amplified. Just add to this mixture dATP, dCTP, dGTP, dTTP, and a heat-stable DNA polymerase from the organism Thermophilus aquaticus (which lives in hot springs), and off you go. The mixture is first heated to a temperature, which causes the DsDNA strands to separate. The temperature is cooled allowing a large stoichiometric excess of the primers to anneal to the ssDNA. The heat-stable Taq polymerase (from Thermophilus aquaticus) polymerizes DNA from the primers. The temperature is raised again, allowing dsDNA strand separation. On cooling the primers anneal again to the original and newly synthesized DNA from the last cycle and synthesis of DNA occurs again. This cycle is repeated as shown in the diagram. This chain reaction is called the polymerase chain reaction (PCR). The target DNA synthesized is amplified a million times in 20 cycles, or a billion times in 30 cycles, which can be done in a few hours.



- Animation of PCR

Editing DNA

We will spend much time discussing how specific amino acids could be covalently modified to either identify the presence of a specific amino acid or to modify the activity of the protein. Now it is routine to use recombinant DNA technology, to alter one or more nucleotides, to either change the amino acid or add or delete one or more amino acids. This technique, called site-specific mutagenesis, is used extensively by protein chemists to determine the importance of a given amino acid in the folding, structure, and activity of a protein. The techniques are described in the diagram below;



Searching DNA

Where on a chromosome is the gene that codes for a given protein? One way to find the gene is to synthesize a small oligonucleotide "probe" which is complementary to part of the actual DNA sequence of the gene (determined from previous experiments). Attach a fluorescent molecule to the DNA probe. Then take a cell preparation in which the chromosomes can be seen under the microscope. Base is added which unwinds the double-stranded DNA helix. A fluorescent probe is added that will bind to the chromosome at the site of the gene to which the DNA is complementary. Hybridization is the process whereby a single-stranded nucleotide sequence (the target) binds through H-bonds to another complementary nucleotide sequence (the probe).

What if you don't know the nucleotide sequence of the gene, but you know the amino acid sequence of the protein, as in the example shown below? From the genetic code table, you could predict the possible sequence of all possible RNA molecules that are complementary to the DNA in the gene. Since some of the amino acids have more than one codon, there are many possible sequences of DNA that could code for the protein fragment. The link below shows all possible corresponding mRNA sequences that could code for a short amino acid sequence. The 20 mer sequence of minimal degeneracy in the nucleotide sequence should be used as genomic probes.

Comparing DNA

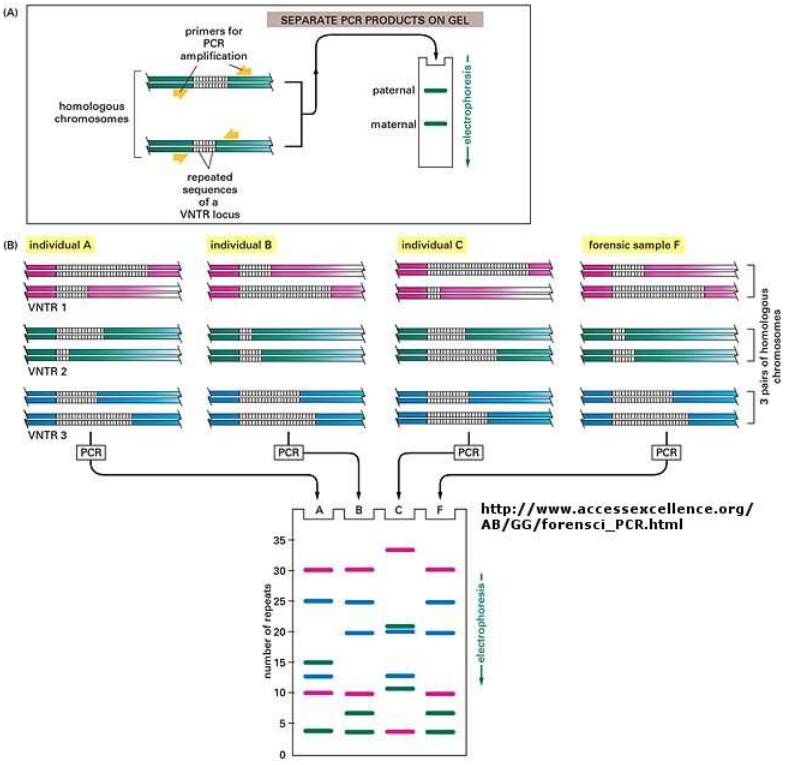
The DNA sequence of each individual must be different from every other individual in the world (with the exception of identical twins). The difference must be less than the differences between a human and a chimp, which are 98.5 % identical. Let us say that each of us have DNA sequences that are 99.9 % identical as compared to some "a normal humans". Given that we have about 4 billion base pairs of DNA, that means we are all different in about 0.001 x 4,000,000,000 which is about 4 million base pairs different. This means that on average we have one nucleotide difference for each 1000 base pairs of DNA. Some of these are in genes, but most are probably in between DNA, and many have been shown to be clustered in areas of highly repetitive DNA at the ends of chromosomes (called the telomeres) and in the middle (called the centromeres).

Now, remember that there are restriction enzyme sites interspersed randomly along the DNA as well. If some of the differences in the DNA among individuals occur within the sequences where the DNA is cleaved by restriction enzymes, then in some individuals a particular enzyme won't cleave at the usual site, but at a more distal site. Hence, the size of the restriction enzyme fragments

should differ for each person. Each person's DNA, when cut by a battery of restriction enzymes, should give rise to a unique set of DNA fragments of sizes unique to that individual. Each person's DNA has a unique **Restriction Fragment Length Polymorphism (RFLP).** How could you detect such polymorphism?

You already know how to cut sample DNA with restriction enzymes, and then separate the fragments on an agarose gel. An additional step is required, however, since thousands of fragments could appear on the gel, which would be observed as one large continuous smear. If however, each fragment could be reacted with a set of small, radioactive DNA probes which are complementary to certain highly polymorphic sections of DNA (like telomeric DNA) and then visualized, only a few sets of discrete bands would be observed in the agarose gel. These discrete bands would be different from the DNA bands seen in another individual's gene treated the same way. This technique is called Southern Blotting and works as shown below. DNA fragments are electrophoresed in an agarose gel. The ds DNA fragments are unwound by heating, and then a piece of nitrocellulose filter paper is placed on top of the gel. The DNA from the gel transfers to the filter paper. Then a small radioactive oligonucleotide probe, complementary to a polymorphic site on the DNA, is added to the paper. It binds only to the fragment containing DNA complementary to the probe. The filter paper is dried, and a piece of x-ray film is placed over the sheet. A set of radioactive fragments (which are not complementary to the probe) are run as well. They serve as a set of markers to ensure that the gel electrophoresis and transfer to the filter paper were correct. This technique is shown on the next page, along with a RFLP analysis from a particular family.

When this technique is used in forensic cases or in paternity cases, it is called DNA fingerprinting. With present techniques, investigators can state unequivocally that the odds of a particular pattern not belonging to a suspect are in the range of one million to one. The x-ray film shown below is a copy of real forensic evidence obtained from a rape case. Shown are the Southern blot results from suspect 1, suspect 2, the victim, and the forensic evidence. Analyze the data.



1.5: Chapter 1 Questions

Section 1 Questions

? Question 1.5.1

In Figure 1.2, two examples of types of enzyme-substrate binding are shown: the Lock-and-Key model and Induced-Fit. What are some situations in which one style of the enzyme would be favored over the other?

Answer

Lock and key enzymes are highly specific for their substrate and therefore do not need a transition state to undergo the catalytic reaction. This could be used for substrate channels like Na+/K+ pumps in which a reaction doesn't need to occur.

Induced fit enzymes utilize a transition state, to convert a substrate into a product. The transition state is able to cause a conformational change in the active site of the enzyme and facilitate high-energy reactions such as breaking or forming chemical bonds

? Question 1.5.2

Label the following type of import/export mechanisms as passive, active, or facilitated and explain why: endocytosis, ion channels, pores, transporters/permeases. *Some may have more than one answer.*

Answer

Endocytosis: Active, facilitated. Endocytosis or "cell eating" is a multi-enzyme mediated process that allows the cell to uptake large particles from its environment. This involves membrane modification, protein receptors, and digestive enzymes and organelles working across gradients.

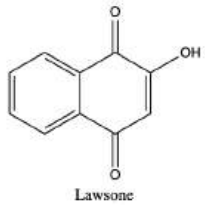
Pores: Passive, facilitated. Once porins establish pores, such as in the nuclear envelope, small molecules like DNA and RNA can passively diffuse in and out of the membrane without the need for carrier proteins.

Ion Channels: Active, facilitated OR passive, facilitated. Active ion channels pump small molecules across a gradient and are typically considered to be "gated," meaning that the enzymes can open and close in a regulated manner to control what is being moved across the membrane. Passive ion channels are permanently open to facilitate transfer and rely on a constantly established concentration gradient to allow for transport to occur.

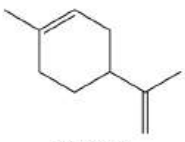
Transporters/Permeases: Active, facilitated. Transporters move larger molecules across a concentration gradient and assist in the movement of soluble proteins and molecules through the hydrophobic membrane

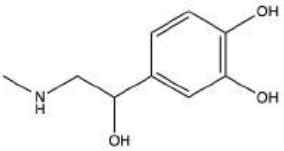
Section 2 Questions

Label the functional groups present in the chemicals shown below:

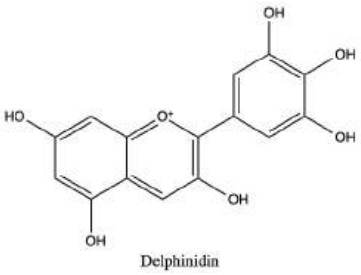


Lawsone

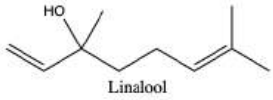

Limonene



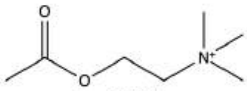
Adrenaline



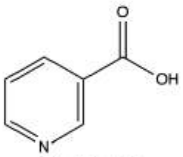
Delphinidin



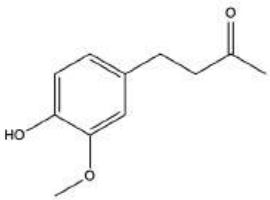
Linalool



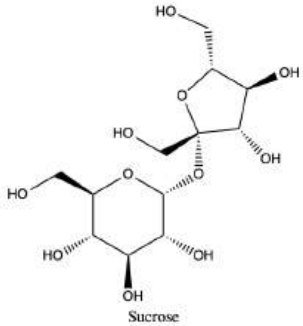
Acetylcholine



Nicotinic Acid



Zingerone



Sucrose

Answers:

Section 3 Questions

1) a. Consider a subset of reactions of glycolysis given below. ΔG'°, substrates, and products are given from colon cancer cells (nmol/g tissue). After examining the conditions of the cell for each enzymatic reaction, predict if the ΔG of the reaction will increase or decrease. (Data from Hirayama A et al. 2009 *Cancer Research*. The ratio of NAD+/NADH is 10:1 and the concentrations of the cofactors are ATP (110) and ADP (300).

Reaction	ΔG'°	[Substrate]	[Product]
#1 Hexokinase Glucose → Glucose-6-phosphate	-16.6	123	75
#2 Phosphoglucose Isomerase Glucose-6-phosphate → Fructose-6-phosphate	1.67	75	50
#3 Phosphofructokinase Fructose-6-phosphate → Fructose-1,6-bisphosphate	-14.2	50	50
#10 Pyruvate Kinase Phosphoenolpyruvate → Pyruvate	-31.4	5	25
Lactate Dehydrogenase	-25.1	25	25,000

Answer:
Reaction #1 - Increase.
Reaction #2 - Decrease.
Reaction #3 - Increase.
Reaction #10 - Increase.
Lactate Dehydrogenase - Increase.

2) Consider the reaction below along with the thermodynamic properties: ΔH° = -760 kJ/mol, ΔS° = -0.185 kJ/mol K, and ΔG = -705 kJ/mol

$Na^+_{(g)} + Cl^-_{(g)} \rightarrow Na^+_{(aq)} + Cl^-_{(aq)}$

At what temperature would this reaction have an equilibrium constant of 1?

Answer: $\Delta G° = RT\ln(K_{eq})$

Because we want know know the temperature at which $K_{eq} = 1$, and we know that the ln(1) = 0, $\Delta G° = 0$ when $K_{eq} = 1$.

$\Delta G° = \Delta H° - T\Delta S°$
0 = -760 kJ/mol - T(-0.185 kJ/mol K); Rearrange and solve for T = 4108.1 °K

CHAPTER OVERVIEW

2: Water and its Role in Life

2.1: The multiple roles of water
2.2: Weak Acids and Bases, pH and pKa
2.3: Buffering against pH Changes in Biological Systems
2.4: Solubility in an aqueous world - noncovalent interactions in depth
2.5: Solubility in an aqueous world - The Hydrophobic Effect
2.6: Chapter 2 Questions

2.1: The multiple roles of water

Search Fundamentals of Biochemistry

"Nothing in the world is as soft and yielding as it,
Yet nothing can better overcome the hard and strong,
For they can neither control nor do away with it.
The soft overcomes the hard,
The yielding overcomes the strong;"

These words come from the Tao Te Ching by Lao Zu. Let's convert this into a chemical riddle and apply it at the nanoscopic level to biochemistry!

"What it loses it gains,
What it donates it accepts,
It is weak yet strong,
It strengthens yet destroys;"

What is it? The answer (one of many possible) is water! It gains and loses protons, donates and accepts electrons, can be both a weaker or stronger acid/base or oxidizing/reducing agent, and can lead to crystal formation or dissolution, depending on circumstances. Water, at least on our planet, appears necessary for life. We know of no biological life form that exists without it. This molecule has a plethora of properties, which make it unique compared to most other liquids and optimal for the type of life on earth. It has contrasting and oppositional properties. Let's investigate a few.

2.1.1: Water as a solvent

Solubility is a property that depends on the nature of both solute and solvent. To a first approximation, We tell students in introductory chemistry and biology courses that for a solute to dissolve in a solvent, and form a solution (an example of a homogenous mixture), the sum of noncovalent interactions (intermolecular forces) between solute and solvent must be greater than those among solute molecules and those among solvent molecules.

As students advance in chemistry classes, nuance is added to that general understanding as entropic contributions to solubility must be considered. Entropy is often described as the degree of apparent disorder in the system. Given that description, changes in entropy would appear to favor the soluble state as a solution of the solute in solvent would be more disordered. That simple description must be adjusted to account for the ordered state of solvent (a clathrate) surrounding a solute and of "holes" in the solvent that accommodate larger solute molecules. Enthalpy considerations also must be considered. The description of entropy as a measure of disorder is not precise. Rather it should be described as a measure of the number of microstates of energy or particles available within a system. An entropy increase would arise from an increase in the number of such available microstates, which could correlate with an increase in the disorder of a system.

Students might often consider a molecule as either soluble or insoluble in a given solvent. This notion can be reinforced by simple liquid/liquid partitioning experiments in organic chemistry experiments in which two immiscible solvents (for example water and an ether) are used. Yet diethyl ether is partially soluble in water (1 g/100 mL). Nonpolar molecules with no or few bond dipoles are generally considered insoluble. Students would know that acetic acid, a two-carbon molecule, is soluble in water, but how many carbons are necessary for the molecule to become essentially insoluble? Molecules with a single polar group (-OH, CO_2H) and a long alkyl/acyl chain are best described as amphiphilic. Amphiphiles like octanol ($C_8H_{17}OH$) and dodecyl sulfate ($CH_3(CH_2)_{10}CO_2H$) can form multimolecular aggregates called **micelles** even as they exist in as a biphasic system, as shown in the following equilibria:

$$C_8H_{17}OH(liq) \leftrightarrow C_8H_{17}OH(aq) \rightleftharpoons C_8H_{17}OH(micelle).$$

Figure 2.1.1 shows an interactive iCn3D model of a micelle below, which consists of 54 self-associated molecules of dodecylphosphocholine fatty acids. It has almost "complete" separation of polar (on the surface) and nonpolar atoms (buried).

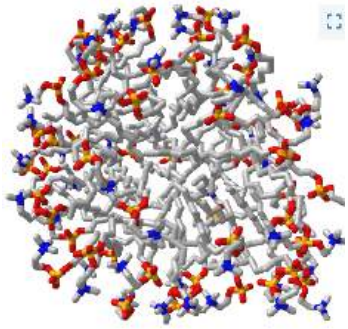


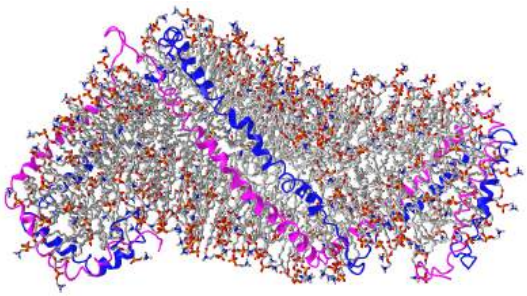
Figure 2.1.1: A Micelle (Copyright; author via source). Click the image for a popup or use this link: https://structure.ncbi.nlm.nih.gov/i...Y7zSYepTQu7yV7.

Note the grey lines representing the nonpolar tails are buried from the surrounding water molecules, which form H bonds with the polar head groups.

Without some limited solubility, the following reaction could not occur:

$$\mathrm{nC_8H_{17}OH(aq) \leftrightarrow 1-C_8H_{17}OH(micelle)}\cdot$$

To solve the general problem of the limited solubility of organic molecules in aqueous-based life, biomolecular structures have evolved to “transport” mostly nonpolar molecules like long-chain carboxylic acids (fatty acids) and cholesterol in circulation. The structure of one such fatty acid and cholesterol-containing particle, nascent high-density lipoprotein (HDL), has been determined by small-angle neutron scattering. Figure 2.1.2 shows an an interactive iCn3D model of it. The gray sticks represent the nonpolar, acyl tails of the long-chain carboxylic (fatty) acids while the polar red (oxygen) and blue (nitrogen) atoms surrounding the surface are polar groups connected to the tails. The long magenta and dark blue "helices" represent a protein that wraps around the particle and stabilizes it.



NCBI iCn3D Figure 2.1.2: Nascent HDL (3k2s) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...fF85h11SpeYJg6

The same ideas can be applied to the solubility of salts. Students will remember general solubility rules (all Gp 1 and Gp 7 salts are soluble) from introductory chemistry. Salts of divalent cations are less soluble as the attractive ion-ion forces within the solid crystal lattice are too strong for the compensatory ion-dipole interactions between the ion and water. Hence salts of Ca^{2+} and Fe^{2+} ions such as $CaCO_3$ and $FeCO_3$ are generally insoluble (K_{sp} values of 1.4 x 10^{-8} and 3.1 x 10^{-11}, respectively). Insoluble calcium salts (carbonates and silicates) are need for shells of Foraminifera and skeletons of vertebrates. Yet free Ca^{2+} and Fe^{2+} ion are found in extracellular and intracellular compartments. Divalent cations like Fe^{2+} can be toxic at a higher concentration so ways to effectively transport and sequester them have evolved. Figure 2.1.3 shows the structure of human heavy-chain ferritin (4zjk), a protein that forms a hollow shell in which is stored Fe^{2+} ions (along with counter ions). The model below shows a ferritin with 120 Fe^{2+} ions (spheres) inside the hollow ferritin sphere.

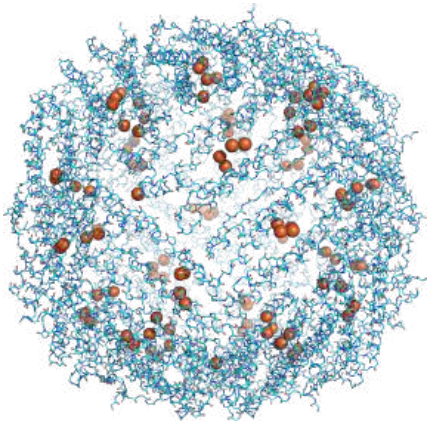


Figure 2.1.3: Human heavy-chain ferritin (4zjk) containing 120 Fe^{2+} ions encapsulated within the hollow sphere formed by the protein.

Finally, let's consider the solubility of gases. The ones that are the most abundant and relevant are O_2 and CO_2 as they are the reactants and products of oxidative respiration. The gases, although they contain oxygen atoms, are nonpolar and have no net dipole. Hence they are quite insoluble in water. However, they must be soluble enough to allow fish to extract it from water. To solve the solubility problem, evolution has produced proteins like vertebrate hemoglobin that bind oxygen through a transition metal complex containing Fe^{2+}-heme complex (hemoglobin in vertebrates). Some invertebrates use the transition metal Cu ions in hemocyanins for the same purpose. Figure 2.1.4 shows an interactive iCn3D model of dioxygen (red spheres), bound to a planar heme (yellow highlights) which contains an Fe^{2+} at its center (not shown) at it center in human hemoglobin (6BB5)

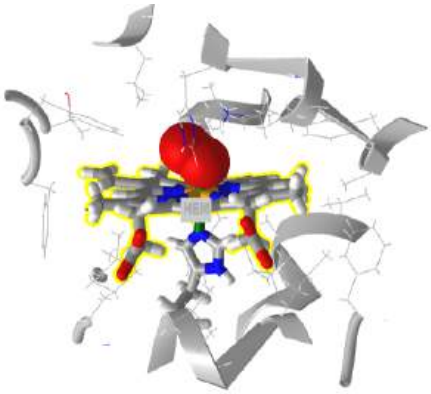


Figure 2.1.4: Oxy-Heme in Human Hemoglobin (6BB5). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...DQ5syZ2XTZ4kR9

Water engages in noncovalent interactions with itself and other molecules. Individual noncovalent interactions are weak but if there are many they can lead to very strong interactions. You've studied noncovalent interactions before, which may have been described as "intermolecular forces". We prefer the term noncovalent interaction. These include ion-ion, ion-dipole, hydrogen bonds, dipole-dipole, induced dipole-induced dipole, and other variants.

All of these interactions originate in the electrostatic *force* between two charged objects. There is only one law that describes the forces of attraction, and that's Coulomb's Law:

$$F = \frac{kQ_1Q_2}{r^2}$$

From this force derives all the electrostatic *interactions* listed above. The magnitude of the attractions for these electrostatic interactions depends on the way charge is distributed in the attracting species. We will explore these in depth in Chapter 2.4.

2.1.2: Water as a reactant: Acids and Bases

H_2O, with its sharable lone pairs and slightly positive Hs is both a Brønsted–Lowry base and acid. Its acid base chemistry hence is among it's most important features.

Water, acting as a base, can react with both strong and weak acids. Examples of reactions of a strong acid (HCl) and weak acids (acetic acids and ammonium) with water as a base are shown in Figure 2.1.5.

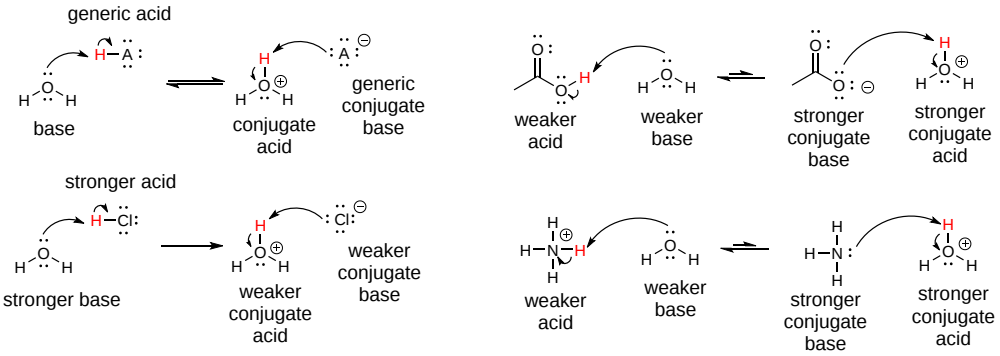


Figure 2.1.5: Reaction of acids with water as a base

Likewise, water can act as an acid as demonstrated in Figure 2.1.6.

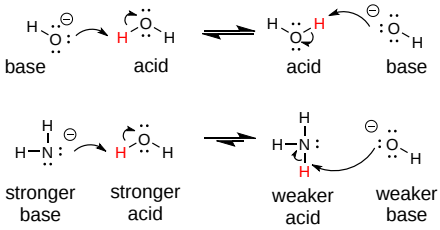


Figure 2.1.6: Reaction of bases with water as an acid

In the first example, no net changes occur. In the second, a negatively charged deprotonated amine (a stronger base than water) can accept a proton from water, which acts as an acid. All acid/base reactions go predominantly in the direction of stronger acid/strong base to weaker acid/weaker base. Whether water reacts with a strong acid, such as HCl, or a weak one like acetic acid, the strongest acid that can actually exist in an aqueous system is $H_3O^+_{(aq)}$. This is an example of the leveling effect.

2.1.3: Water as a reactant: nucleophile/electrophile

In the reactions above, we characterized water as a Brønsted–Lowry acid or base. More generically, we could have said water is a Lewis acid (electron pair acceptor) or Lewis base (electron pair donor). In many reactions, we can also call water a nucleophile (when it shares it lone pair) or an electrophile (when its slightly positive H atoms react with a nucleophile. Here are some examples.

Reaction of water with a transition metal complex.

This reaction below is effectively a nucleophilic substitution reaction in which water displaces ammonia as a ligand as shown in Figure 2.1.7 and in the following chemical equation.

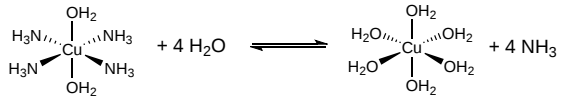


Figure 2.1.7: Reaction of water with transition metal complexes

$$[Cu(NH_3)_4(H_2O)_2]^{2+} + 4\,H_2O \rightleftharpoons [Cu(H_2O)_6]^{2+} + 4\,NH_3$$

Hydration of an alkene

The reaction is catalyzed by the addition of a proton from an acid (like H_2SO_4) which can be called an electrophilic hydration. Once protonated at the carbon which makes the most stable carbocation, water as a nucleophile attacks the positive carbon to produce the alcohol. These steps are illustrated in Figure 2.1.8.

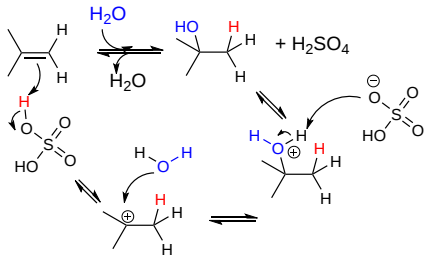


Figure 2.1.8: Mechansim for hydration of an alkene

Nucleophilic substation at an electrophilic carbonyl

This is a very common reaction. When water is the nucleophile, the reaction is also called a hydrolysis reaction. The reactions in Figure 2.1.9 are shown with OH^- as the nucleophile instead of water for simplicity.

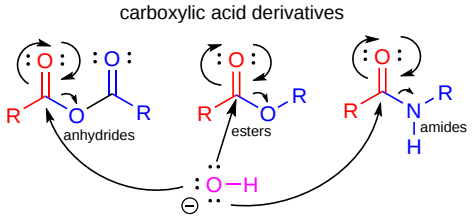


Figure 2.1.9: Reaction of water (OH^-) as a nucleophile

Water as a reactant: Oxidizing/Reducing agent

Everyone knows what happens if you throw a piece of solid Na or K into water. An extremely exothermic reaction occurs which releases H_2 gas which can catch fire and lead to an explosion. The reaction of Na is:

$$2\,Na(s) + H_2O \rightarrow 2\,Na^+(aq) + OH^-(aq) + H_2(g)\cdot$$

The oxidation number of elemental sodium is 0, while Na^+ is +1, indicating that the sodium metal has been oxidized by the water which acts as an oxidizing agent.

This reaction occurs with many pure metals, but some that are less reactive (remember the activity series from introductory chemistry?) required acid, a protonated form of water, as shown in the reaction below:

$$Zn(s) + 2\,H_3O^+(aq) \longrightarrow Zn^{2+}(aq) + 2\,H_2O(l) + H_2(g)$$

As in acid/base reactions, in a redox reaction, an oxidizing agent and a reducing agent react to form a new oxidizing and reducing agent. Other reactants can oxidize water to form oxygen. Consider fluorine gas for example:

$$3\,F_2 + 2\,H_2O \longrightarrow O_2 + 4\,HF$$

F_2 is a strong oxidizing agent (as you would surmise from its electronegativity) than O_2 so the reaction proceeds vigorously to the right.

Of more biological relevance is the oxidation of water to produce O_2 in photosynthesis, a complex series of reactions that is effectively the reverse of combustion:

$$6\,CO_2(g) + 6\,H_2O(l) \rightarrow C_6H_{12}O_6(s) + 6\,O_2(g)\cdot$$

This reaction obviously is endergonic and requires a large input of energy so the reaction proceeds to produce the potent oxidizing agent O_2. The special oxygen-evolving complex in photosynthesis is the powerful oxidant that can oxidize H_2O to form the weaker oxidizing agent, O_2.

2.2: Weak Acids and Bases, pH and pKa

The previous section described the general acid/base properties of water. There are many functional groups in both small and large biomolecules that act as acids and bases. Common weak acids are carboxylic acids and derivatives of phosphoric acid which become negatively charged on donation of a proton. Common weak bases are amines, which become positively charged on protonation. Such charge acquisition changes the properties of the acid or base. A protonated amine is no longer a nucleophile. A deprotonated carboxylic acid can now engage in an ion-ion IMF. The extent of deprotonation depends on the acidity/basicity of the environment. We have to turn to a bit of mathematics to determine that extent.

2.2.1: Reaction of water with self: Autoionization

As shown in the previous section, water can react with itself to produce H_3O^+ and OH^- as illustrated in Figure 2.2.1.

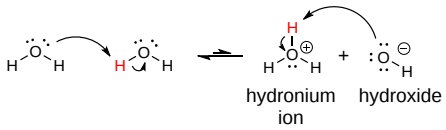


Figure 2.2.1: Reaction of water with self

This autoionization reaction is often represented in a simpler form:

$$\mathrm{H_2O \rightleftharpoons H^+ + OH^-}.$$

The equilibrium constant for this simplified reaction can be written as

$$K_{eq} = \frac{[H^+]\,[OH^-]}{H_2O} \tag{2.2.1}$$

Given the known value of K_{eq} and the concentration of water (55 M), this can be simplistically rewritten as

$$K_a = 55K_{eq} = \left[H^+\right]\left[OH^-\right] = 10^{-14} \tag{2.2.2}$$

(see discussion of the pKa of water below.

Hence pure, neutral water has equal but small concentrations, 10^{-7} M of H_3O^+ and OH^-.

You remember from introductory chemistry and life in general that the pH of pure water is 7. This derives from the general formulas for both pH and a new quantity, pKa.

$$pH = -\log\left[H_3O^+\right] = -\log\left(10^{-7}\right) = 7$$
$$pK_a = -\log K_a = -\log\left(10^{-14}\right) = 14$$

Note

Some texts incorrectly use 15.7 for the pKa of water. Here is a link to an explanation of why 14 is better. The wrong value of 15.7 would make the pKa of water higher than that of methanol (15.3), which simply can't be since the methoxide anion is less stable due to electron release by the methyl group than OH^-.

All acids of the generic formula HA have pKa.

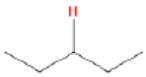
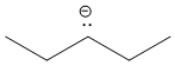
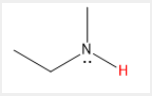
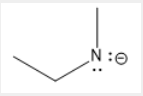
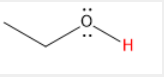
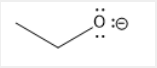
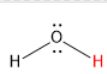
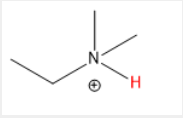
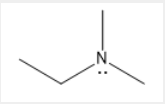
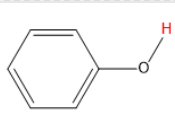
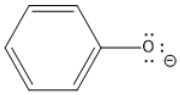
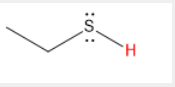
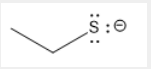
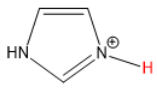
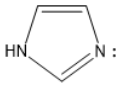
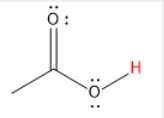
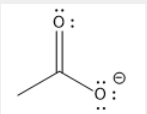
$$\mathrm{HA \rightleftharpoons H^+ + A^-}$$

The equilibrium constant for this simplified reaction can be written as

$$\begin{aligned} K_{eq} &= \frac{[H^+]\,[A^-]}{HA} \\ K_a &= [HA]K_{eq} = \left[H^+\right]\left[A^-\right] \\ pK_a &= -\log K_a \end{aligned} \tag{2.2.3}$$

The pK_a becomes a simple measure of the strength of an acid. The stronger the acid, the larger the K_a and the smaller the pK_a.

Here is a table of pK_a values for common acids and functional groups. The pKa values change with different substituents on the acids differ. The stronger the acid, the weaker the conjugate base. This should make sense as a weak base is unlikely to reabstract a proton and return to its original acidic form. Likewise, the weakest acids produce the strongest conjugate bases which reprotonate to return to the weak acid state.

Group	Example	weaker acid	≈ pKa	Conjugate Base	stronger conj. base
alkane			50		
amine			35		
alkyne			25		
alcohol			16		
water			14		
protonated amine			10		
phenol			10		
thiol			10		
imidazole			7		
carboxylic acid			5		
hydrochloric acid			-8		
		stronger acid			weaker conj. base

2.2.2: The Henderson-Hasselbalch Equation

We can find the pKa for small acids in solution in pKa tables. However, from a biochemical perspective, we often need to know the charge state of the acid. Since the pH is approximately constant in organisms (more on that later), we know the $[H_3O^+]$. Hence we can calculate the ratio of A^-/HA using the **Henderson-Hasselbalch equation** (Equation ???), which is derived below.

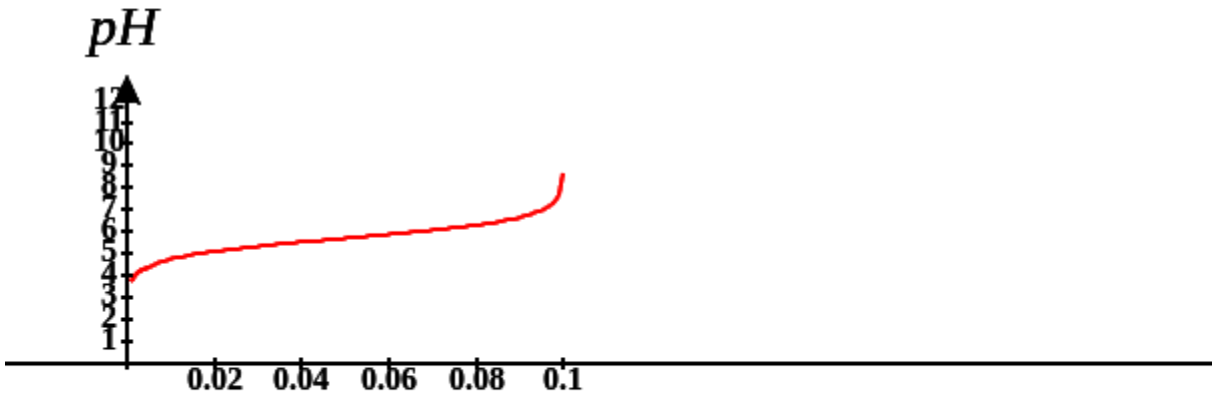
$$K_a = \frac{[H^+][A^-]}{HA}$$
$$-\log K_a = -\log[H^+] - \log([A^-]/[HA]) \tag{2.2.4}$$
$$pK_a = pH - \log([A^-]/[HA])$$

which given the traditional Henderson-Hasselbalch equation below.

$$pH = pK_a + \log \frac{[A^-]}{HA} \tag{2.2.5}$$

In your chemistry class, you certainly would have performed titration curve analyses of acids. What is the chemistry that occurs at each step? Let's assume the pH is low and much lower than the pKa of the acid. From the Henderson-Hasselbalch equation, you would surmise that the ratio of A^-/HA is very small - that is the acid is essentially fully protonated. That should also make intuitive sense. For a weak acid to be coaxed to give up a proton, a reasonably strong base (like OH^-) should be added. So at low pH, the acid exists just as HA. Now consider adding an amount to NaOH to match the concentration of the ionizable proton. At that point in the titration, mass balance would suggest that the acid in its protonated state is gone, and all that remains is A^-. What happens if just enough NaOH is added to react with half of the HA. The mass balance would tell us that A^-=HA and at that point, the pH = pK_a of the acid.

The entire titration curve can be calculated from the Henderson-Hasselbalch equation. A graph of it is shown below.



The graph simply shifts up as the pKa is increased. The pH starts soaring at the end of the graph after the added hydroxide has reacted with the last ionizable proton. After that, the pH is determined by the concentration of the strong base OH^-. The graph is flattest in the middle of the curve at the inflection point of the curve. Note at this pH, **pH = pKa.** In the middle of the curve, the pH changes least on the addition of small amounts of OH^-. This is the basis of buffering which will be covered in the next section.

If you know the pH of a solution and the pKa of the ionizable group, you can very quickly estimate the average charge state of the function group. Let's see what the Henderson-Hasselbach equation (Equation ???) predicts under three specific pH states:

1. If the pH is 2 units below the pKa (i.e under more acidic conditions when you would expect the group to be protonated), the equation becomes, $-2 = logA/HA, or .01 = A/HA$. This means that the functional group will be about 99% protonated (with either 0 or +1 charge, depending of the functional group).
2. If the pH is 2 units above the pKa, the equation becomes $2 = logA/HA, or 100 = A/HA$. Therefore the functional group will be 99% deprotonated.
3. If the pH = pka, the HH equation becomes $0 = logA/HA, or 1 = A/HA$. Therefore the functional group will be 50% deprotonated.

From these simple examples, we have illustrated the **±2 rule** to determine the charge state. This rule is used to quickly determine protonation, and hence charge state, and is extremely important to know (and easy to derive).

2.2.2.1: Polyprotic Oxyacids

Acids that can donate more than one proton are called **polyprotic acids**. They are typically oxyacids, with the ionizable proton on an oxygen atom, which can form a reasonably stable the oxyanion (negative on the oxygen) as the oxygen is electronegative and stabilize the charge. The negative charge on the conjugate base of oxyacids is further stabilized by resonance delocalization involving the doubly bonded oxygen atom. Two of the most biologically relevant oxyacids are shown in Figure 2.2.2.

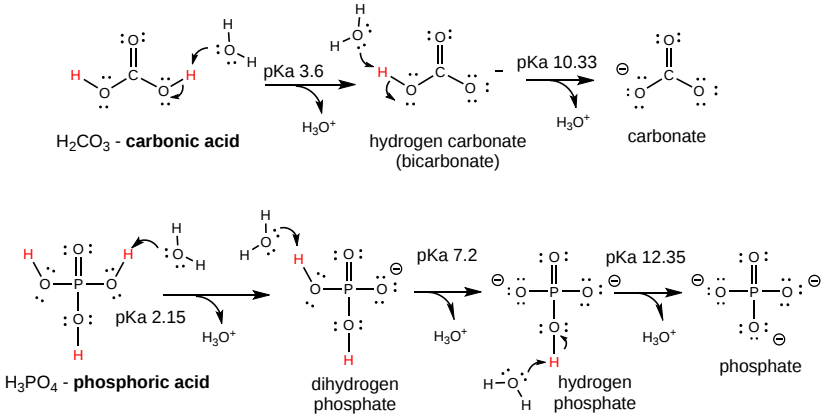


Figure 2.2.2: Reactions of polyprotic acids with water

The pKa for each subsequent ionization is higher since it is more difficult to remove a proton from an increasingly more charged molecular ion. The titration plot of pH vs NaOH is similar to the graph above but has multiple plateaus at pH=pKa,

Derivatives of phosphoric acid are found in all major classes of biomolecules. Nucleic acids contain a sugar-phosphate link in their backbone. Many proteins become phosphorylated after they are synthesized. Membrane lipids usually contain a phosphate group. A whole class of phospholipids are found in biomembranes.

2.2.2.2: Charge State of Biomolecules

The Henderson-Hasselbalch equation can be used to determine the charge state of ionizable functional groups (carboxylic and phosphoric acids, amines, imidazoles, guanidino groups) even on a large macromolecule such as proteins, which contain carboxylic acid (weak acids) and amines (weak bases). Figure 2.2.3 shows how the weakly acidic aspartic and glutamic acids, two common amino acids found in proteins, contribute negative charge to the protein and how the amine of the amino acid lysine, a weak base, contributes to positive charge.

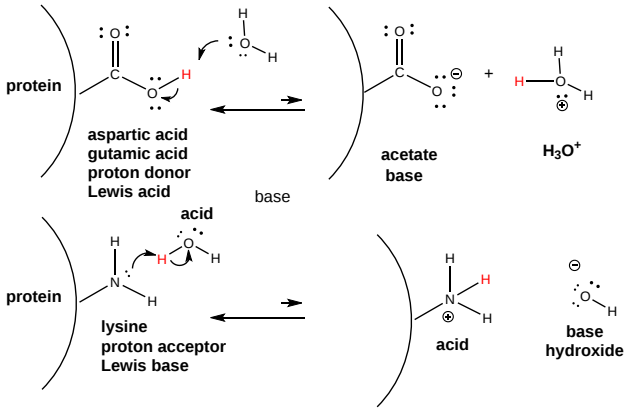


Figure 2.2.3: Deprotonation/Protonation of weak amino acid side chains contributes to protein charge.

Other amino acids that contain an alcoholic function group can also become phosphorylated to produce a phosphoprotein, which converts a neutral ROH group to a phosphoester with a negative two charge as shown in Figure 2.2.4.

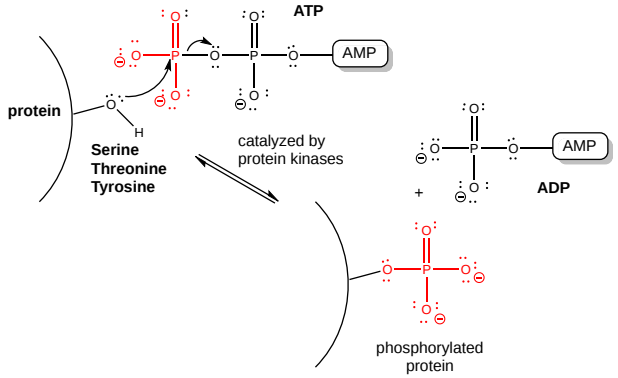


Figure 2.2.4: Phosphorylation of -OH containing side chains of proteins by ATP

2.2.2.3: pKa and Environment

The pKa is really a measure of the equilibrium constant for the reaction. And of course, you remember that ΔG^o = -RT ln Keq. Therefore, pKa is independent of the concentration and depends only on the intrinsic stability of reactants with respect to the products. However, this is true only under a given set of conditions such as temperature, pressure and solvent composition.

Consider, for example, acetic acid, which in aqueous solution has a pKa of about 4.7. It is a weak acid, which dissociates only slightly to form H^+ (in water the hydronium ion, H_3O^+, is formed) and acetate (Ac^-). These ions are moderately stable in water but reassociate readily to form the starting product. The pKa of acetic acid in 80% ethanol is 6.87. This can be accounted for by the decrease in stability of the charged products, which are less shielded from each other by the less polar ethanol. Ethanol has a lower dielectric constant than water. The pKa increases to 10.32 in 100% ethanol, and to a whopping 130 in air!

The pKa values of ionizable groups in proteins vary enormously as they depend on the microenvironment of the group. Consider the amino acid aspartic acid (Asp, D), which has a $-CH_2CO_2H$ R-group or "side chain" similar to acetic acid. In a given protein, a given Asp side chain might be on the surface but another in the same protein might be buried in the protein away from water. You would expect the pKa values for these two different Asp side chains to be different. The average pKa for Asp side chains in 78 different proteins was shown to be 3.5, less than that of acetic acid (4.7) but not dramatically. However, the range of pKa values for Asp in these proteins was huge, with the lowest being 0.5 (a buried Asp in the protein T4 Lysosome) and the highest being 9.2 in the protein thioredoxin from *E. Coli*.

Figure 2.2.5 shows an interactive iCn3D model of the surrounding environment of Asp 70 (D70) in T4 Lysoszyme. Its pKa has been determined experimentally to be 0.5 ,way stronger than acetic acid. The dotted cyan lines show ion-ion interactions between the $-CH_2CO_2^-$ side chain of Asp 70 (D7) and the positively charged imidazolium group of histidine (H31) in the protein. The distance between the two charged groups is about 3.4 A.

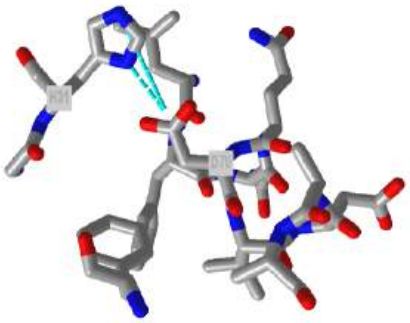


Figure 2.2.5: Surrounding environment of Asp 70 (D70) in T4 Lysoszyme (2b6z). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...RRiVn6saN92tU6

The next model shows the surroundings of Asp 26 (D26) in *E. coli* thioredoxin, It has a pKa of 9.2. The dark blue group is surface exposed positively charge lysine side chain which can stabilize a negative charge on the Asp 26. Note, however, that it is much farther away than the imidazolium group in T4 lysozyme that stabilizes the negative change on D70. The rest of the model is colored based on hydrophobicity, which shows that the Asp 26 side chain is essentially surrounded by nonpolar groups. These would destabilize a negative charge on the D26, enhancing the stability of protonated (neutral) Asp, and elevating its pKa to 9.2.

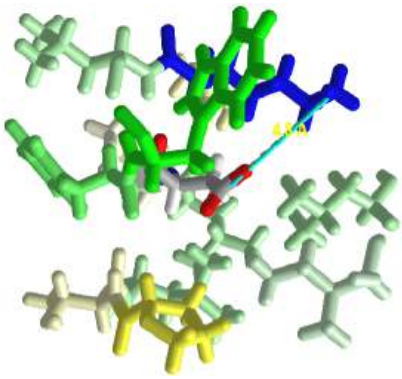


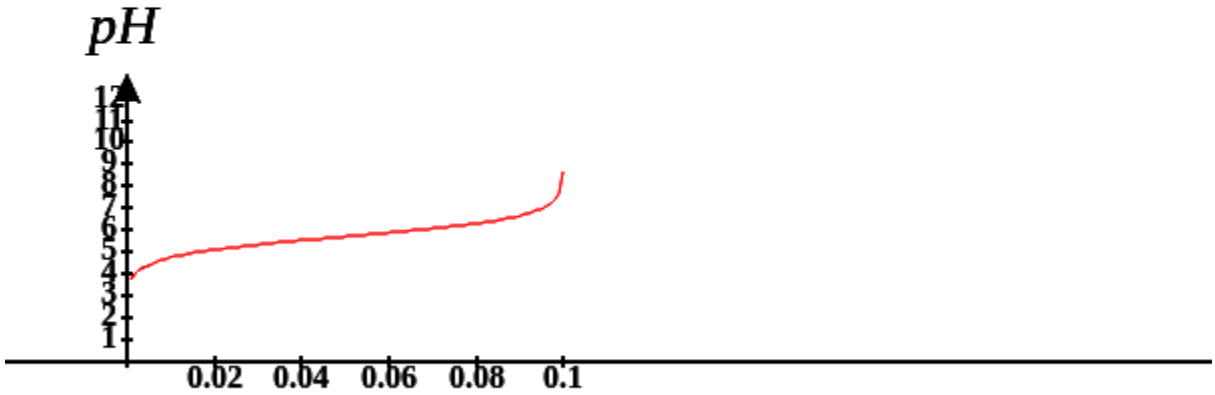
Figure 2.2.6: Surrounding environment of Asp 26 (D26) in E. Coli thioredoxin (5HR2). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...i9BNNdbA2bmP5A

2.3: Buffering against pH Changes in Biological Systems

2.3.1: Introduction

As one way to ensure homeostasis, the pH is maintained between 7.35 and 7.45 in humans. (Much lower pH values, around 4.5, are found in the lysosome). Lower pH values are associated with metabolic and respiratory acidosis while higher pH values are characteristic of metabolic and respiratory alkalosis. pH is maintained by buffering systems that consist of a weak acid and base. If you understand the Henderson-Hasselbalch equation from the previous section, buffer systems become easy to understand.

$$pH = pK_a + \log \frac{[A^-]}{HA} \tag{2.3.1}$$



At the inflection point of the curve, pH = pKa and the system is most resistant to changes in pH on the addition of either acid or base. At this pH, [HA]=[A^-].

If a bit of a strong acid is added, it would react with the strongest base in the solution, which would be the conjugate base of the weak acid:

$$HCl + A^- \rightarrow HA + Cl^-$$

The reaction goes from a strong acid, HCl, to the weak acid, HA. Its concentration would increase a bit but since it's a weak acid, it will only ionize to a small extent. The [HA] in the Henderson-Hasselbalch equations increases a bit but not enough to change the pH significantly. If the same amount of HCl were added to pure water, it would react completely to form an equal amount of H_3O^+ which would significantly alter the pH of pure water (7.0).

If a bit of a strong base is added, it would react with the strongest acid in the solution which would be HA:

$$HA + OH^- \rightarrow H_2O + A^-$$

The reaction goes from a strong base to the weak acid A^-. Its concentration would increase a bit but since it's a weak base, it won't affect the pH significantly. The $[A^-]$ in the Henderson-Hasselbalch equations increases a bit but not enough to change the pH significantly. If the same amount of NaOH were added to pure water, it would react to make the solution basic and significantly alter the pH of pure water (7.0).

To review, buffer solutions contain a weak acid and its conjugate base. They have maximal buffering capacity at a pH = pKa of the weak acid. In general, a buffered solution is best able to withstand a change in pH only with ± 1 pH unit from the pKa.

2.3.2: Biological Buffering Agents

The most relevant systems for biology are the carbonic acid/carbonate buffering system, which controls blood pH and cells and the phosphate buffering system. Proteins, which have many weak acid and base functional groups, can also act as buffering agents.

Carbonic acid/carbonate buffering system: At first glance, the reaction of carbonic acid can be written as follows:

$$H_2CO_3\ (aq) + H_2O(l)\ ßà\ H_3O^+(aq) + HCO_3^-(aq)\ pKa = 3.6$$

However, this system is a bit more complex since we must consider CO_2 (g) solubility and reactivity as well. The overall chemical reactions look like this, where H_2CO_3 is the weak oxyacid, carbonic acid and $HCO_3^-(aq)$ is the weak conjugate based, bicarbonate (or hydrogen carbonate). The $[CO_2(aq)] >> [H_2CO_3\ (aq)]$

$$Rx\ 1:\ CO_2\ (g)\ ßà\ CO_2(aq) + H_2O\ (l)\ ßà\ H_2CO_3\ (aq) + H_2O(l)\ ßà\ H_3O^+(aq) + HCO_3^-(aq)$$

The respiratory system can quickly adjust pH simply by increasing the exhalation of CO_2. The kidneys can respond in a slower fashion to remove H_3O^+ and retain HCO_3^-. The carbonic acid/bicarbonate buffering system can help us understand how shifting equilibria caused by excessive CO_2 released from rapid deep breathing or decreased CO_2 release associated with pulmonary disease or shallow rapid breathing can lead to respiratory alkalosis and acidosis, respectively.

- Respiratory alkalosis can be caused by “hyperventilation” - breathing rapidly. This would lead to breathing out too much CO_2, shifting the above equilibrium to the left, consuming H_3O^+, and increasing pH, making the blood more alkaline. You could breathe into a bag to increase your CO_2 levels.
- Respiratory acidosis is caused by increased CO_2, which can occur when the lungs aren’t working well, and you can’t get rid of CO_2 you produce during respiration Respiratory acidosis can happen with asthma, pneumonia, lung disease or anything that decreases respiration rate.

Inhaling CO_2 can lead to panic. This makes sense as it would mimic suffocation which is lethal to humans. A suffocation response follows. High CO_2 would drive the equilibrium to the right, leading to H_3O^+ production. An acid-sensing ion channel-1a (ASIC1a) in the amygdala has been discovered which appears to mediate the effect. Panic attacks are sometimes associated with hyperventilation which leads to alkalosis, not acidosis. Less noted is that when some people panic, they take short shallow breaths (in a way almost stopping their breath). This would lead to a buildup of CO_2 since it wouldn’t be released in exhalation. The acid channel in the amygdala would be activated and the panic response ensues.

A Dilemma?

How can carbonic acid with a pKa of 3.6 act as a buffer component at pH 7.5?

An astute student might have picked up this conundrum.

The solution to this problem involves looking at the full set of reactions for the components of the buffer system.

Here is the complete set of reactions again:

$$CO_2\ (g)\ ßà\ CO_2(aq) + H_2O\ (l)\ ßà\ H_2CO_3\ (aq) + H_2O(l)\ ßà\ H_3O^+(aq) + HCO_3^-(aq)$$

Let's simplify it since there would be no free "gas bubbles" in blood, so $CO_2\ (g) = CO_2(aq)$:

$$CO_2(aq) + H_2O\ (l)\ ßà\ H_2CO_3\ (aq) + H_2O(l)\ ßà\ H_3O^+(aq) + HCO_3^-(aq)$$

$H_2CO_3\ (aq)$ participates in two different reactions.

Rightwards from $H_2CO_3\ (aq)$:

$$H_2CO_3\ (aq) + H_2O(l)\ ßà\ H_3O^+(aq) + HCO_3^-(aq)$$

Using the simplified equation with H^+ gives

$$K_a = \frac{[H^+]\,[HCO_3^-]}{H_2CO_3} \tag{2.3.2}$$

Hence,

$$[H_2CO_3] = \frac{[H^+]\,[HCO_3^-]}{K_a} \tag{2.3.3}$$

Leftwards from H_2CO_3 (aq) :

H_2CO_3 (aq) ßàCO_2(aq) + H_2O (l)

$$K_2 = \frac{[CO_2]}{H_2CO_3} \tag{2.3.4}$$

so

$$[H_2CO_3] = \frac{[CO_2]}{K_a} \tag{2.3.5}$$

Setting 2.3.3 and 2.3.5 equal to each other gives:

$$[H_2CO_3] = \frac{[H^+]\,[HCO_3^-]}{K_a} = \frac{[CO_2]}{K_2} \tag{2.3.6}$$

Solving for [H^+] gives:

$$[H^+] = \frac{[CO_2]\,(K_a)}{[HCO_3^-]\,(K_2)} \tag{2.3.7}$$

Now take the -log of each side to produce an equation similar to the Henderson-Hasselbalch equation.

$$\begin{aligned} -\log[H^+] &= -\log\left(\frac{[CO_2]}{[HCO_3^-]}\right) - \log\left(\frac{K_a}{K_2}\right) \\ pH &= pK_{aEFFECTIVE} - \log\left(\frac{[CO_2]}{[HCO_3^-]}\right) \end{aligned} \tag{2.3.8}$$

where

$$K_{aEFFECTIVE} = \frac{K_a}{K_2} \tag{2.3.9}$$

This Henderson-Hasselbalch-like equation shows the pH is determined by the ratio K_a/K_2 ratio. $pK_{a\ EFFECTIVE}$ = 6.3. This gives a ratio of CO_2/HCO_3^- of 0.08 = 8/100. There is effectively 12-13 x as much HCO_3^-(aq) as CO_2, making the system primed to react with acid produced metabolically. Yet a second conundrum exists. The pH of the blood (7.4) is outside of the optimal range for a buffer system (in this case ± 1 pH unit from the pKa which is 6.3 in this case). Again, the system is primed to react with acid as it would move the pH close to the optimal buffering pH of 6.3. Other biological systems also must be involved in maintaining pH.

Phosphate buffering system: Phosphates, in the form of dihydrogen ($H_2PO_4^-$) and monohydrogen phosphate (HPO_4^{2-}) are also present in the blood. Given the pKa of HPO_4^{2-}, why is PO_4^{3-} **not** present to any significant degree? Since the concentration of phosphates are low in blood, this system is a minor player in blood.

Proteins: Proteins are found in all cellular and extracellular fluids and they contain weak acids as buffer components. Proteins contain two amino acids, aspartic acid, and glutamic acid) that contain carboxylic acid side chains. Each comprises about 6% of the proteins. In blood, hemoglobin is the most abundant protein by far. Its role in buffering and in O_2 and CO_2 will be discussed in a subsequent chapter.

2.3.3: Making Buffers in the Lab

When studying biomolecules like proteins and nucleic acids in the lab, the pH of the solution is usually maintained under physiological conditions. These macromolecules are either dissolved in or diluted into a buffer solution. Sometimes it's important

to study their properties and activities as a function of pH. A wide variety of buffer systems have been developed for the lab study of these molecules. The dihydrogen ($H_2PO_4^-$)/monohydrogen phosphate (HPO_4^{2-}) pair are commonly used as the pKa of $H_2PO_4^-$ is 7.21, which makes it physiologically relevant. Care must be taken in selecting buffer systems as some of them might bind calcium ions. The pKa of some weak acids vary significantly with temperature as well. Some common biological buffers are listed below.

Buffers	**pKa (at 25°C)**
MES	6.10
Bis-Tris	6.50
ACES	6.78
PIPES	6.76
MOPSO	6.90
MOPS	7.20
HEPES	7.48
Tris	8.06
Tricine	8.05
Gly-Gly	8.20
Bicine	8.26
TAPS	8.40
AMPSO	9.00
CAPS	10.40

There are 3 general ways to make a buffered solution:

1. Make us separate equal concentration solutions of both the weak acid (for example $Na(H_2PO_4)$ and its conjugate base (for example $Na_2(HPO_4)$. Use the Henderson-Hasselbalch equation to calculate how much of each should be added to give the correct $[A^-]/[HA]$ ratio (in the case $[HPO_4^{2-}]/[H_2PO_4^{-1}]$) to give the correct pH.
2. Use a pH meter and monitor the pH when adding both solutions together until the desired pH is reached.
3. Make a solution of one of the components (weak acid or its conjugate base) and bring to near its correct volume for the desired molarity. Monitor the pH as you add a concentrated solution of either HCl or NaOH to get the desired pH. Then bring the solution to the correct volume in a volumetric flask. With this method, you will be adding counter ions (Cl^- with HCl and Na^+ with NaOH) which you may not want in your buffer solution. Often it is not a problem.

2.4: Solubility in an aqueous world - noncovalent interactions in depth

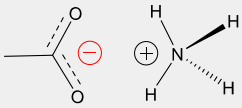
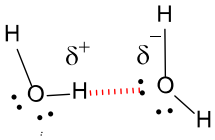
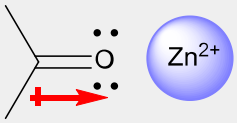
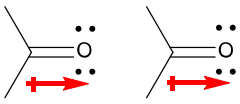
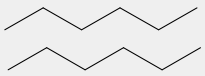
2.4.1: Introduction

In section 2.1, we explored the role of water as a solvent. Using the adage "like dissolves like" that you may have learned in introductory chemistry and biology courses, we can rationalize what substance might dissolve in water. We related this to the types and strengths of attractive interactions that occur between solute and solvent. If in sum they are stronger than self-interactions (solute-solute and solvent-solvent), the solute would dissolve (to a reasonable extent) in the solvent. We also discussed entropic contributions to the dissolution process. For now, we will refocus on the noncovalent interactions.

In introductory science courses, noncovalent interactions are often described as intermolecular forces. This term is ambiguous when applied to biochemistry. Take for example hydrogen bonds. They occur *between* two water molecules, for example, but also *within* larger molecules (like proteins) if hydrogen bond donors and acceptors within the molecule get close enough to each other in space.

The table below summarizes the common noncovalent interactions/"intermolecular forces" that you studied in introductory science classes. It is hard enough for students to recognize and identify these interactions between two small molecules let alone in large molecules like proteins. We will explore these in more detail below, and give examples of noncovalent interactions *between* small molecules and *within* large ones such as proteins. We'll also add a few more specific examples of interactions.

Noncovalent Interactions - "Intermolecular Forces"

Interaction Type	Example	Distance Dependence	Relative Strength Kcal/mol (kJ/mol)	Direction Dependence
Ion-Ion		1/r	60 (250)	nondirectional
H-Bond			3-15 (12-63)	directional
Ion-dipole		$1/r^2$	3-5 (12-21)	directional
Dipole-dipole		$1/r^3$	0.5-1 (2-4)	directional
Induced Dipole-Induced Dipole		$1/r^6$	0.5 (2) (depend on size)	nondirectional

Although there are many different types of noncovalent interactions, one fundamental principle governs all of them. They all originate in the electrostatic *force* between two charged objects. There is one simple law, Coulomb's Law, which you would have discussed in introductory science courses, and one simple equation, that describes the electrostatic force:

$$F = \frac{kQ_1Q_2}{r^2} \tag{2.4.1}$$

where F is the force (attractive or repulsive) between two particles of charge Q_1 and Q_2 with their centers separated by some distance r. Replace the charges with the masses of two objects and you have Newton's Law of Gravitation. Both are inverse squared laws

All the *interactions* described in the table above arise from the electrostatic force. The magnitude of the attractions for the interactions depends on the way charge is distributed in the attracting species. Each has a different dependency on distance.

Different words are used to describe noncovalent interactions. This can be distressing to learners who might hear different terms used by chemists and biologists for the same noncovalent interactions. Some use van der Waals forces to describe induced dipole-induced dipole interactions, while others use London dispersion forces or hydrophobic forces/interactions. Others use van der Waals forces to describe *all* noncovalent interactions except for ion-ion. To avoid ambiguity, let's look at the IUPAC Gold Book Compendium of Chemical Terminology, which offers this definition of van der Waals forces:

Definition: van der Waals Forces

"The attractive or repulsive forces between molecular entities (or between groups within the same molecular entity) other than those due to bond formation or to the electrostatic interaction of ions or of ionic groups with one another or with neutral molecules. The term includes: dipole–dipole, dipole-induced dipole, and London (instantaneous induced dipole-induced dipole) forces. The term is sometimes used loosely for the totality of nonspecific attractive or repulsive intermolecular forces". IUPAC. Compendium of Chemical Terminology, 2nd ed. (the "Gold Book"). Compiled by A. D. McNaught and A. Wilkinson. Blackwell Scientific Publications, Oxford (1997). Online version (2019-) created by S. J. Chalk. ISBN 0-9678550-9-8. https://doi.org/10.1351/goldbook.

Figure 2.4.1 summarizes covalent and noncovalent interactions, using that definition.

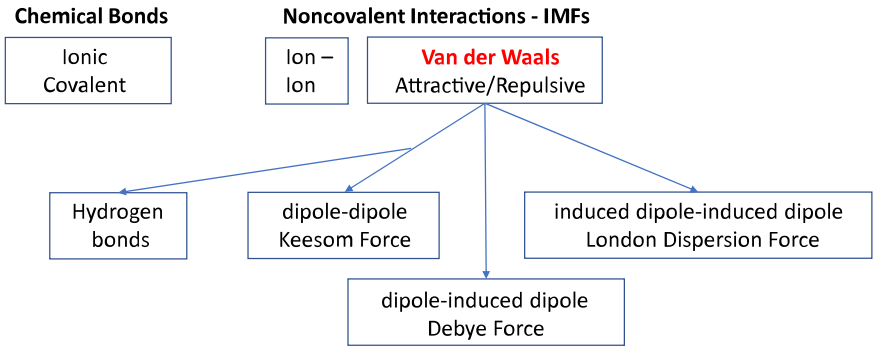


Figure 2.4.1: Noncovalent Interactions

Using this definition, hydrogen bonds are usually considered a type of dipole-dipole interaction. Historically, several of the noncovalent interactions have alternative names based on the person associated with them. Only the names van der Waal and London are commonly used in introductory chemistry courses

Even the word "force" is potentially ambiguous. To a physicist, there are only four known forces:

- gravitational, between two objects with mass
- electromagnetic (between static charges - the electrostatic force, and moving charges - the magnetic force)
- the strong force (holding the nucleus together)
- the weak force (also nuclear and involved in radioactive decay).

We'll try to use the word interaction throughout this book.

Interactions *within* small molecules, such as covalent bonds, and *between* molecules, such as induced dipole-induced dipole, vary as some function of r, the distance between the two interacting particles. Only ion-ion interactions vary as $1/r^2$ ($F= k/r^2$). Attractions lower the overall energy while repulsions raise it. At some optimal distance, the system is in its most favored, lowest energy state.

We just switched from discussing forces to discussing energy, another complex term. The relationship between the potential energy for covalent bond formation and for the noncovalent attraction of two atoms as a function of distance is shown in general form in Figure 2.4.2: below.

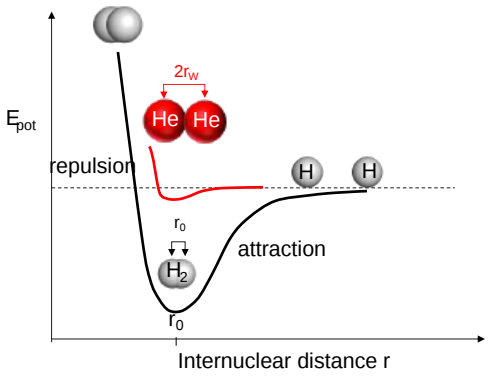


Figure 2.4.2: Potential Energy for Covalent and Noncovlanet interactions

The curve in black shows the shape of E_{pot} vs r for the formation of a covalent bond between H atoms. The Morse potential energy function is used to model energy as a function of r for simple diatomic molecules. The red line shows the shape of E_{pot} vs r for the noncovalent attraction of two He atoms through induced dipole-induced dipole interactions. It is modeled using the Lennard-Jones (6-12) potential function (see below). Each has an optimal r_0 (the bond length for H_2 and two times the van der Waals radius, r_W, of each He in 2He). The energy required to break the induced dipole-induced dipole interactions between He atoms is very small, which accounts for the fact that liquid He, in which many He are interacting, only exists at very cold temperatures (boiling point = -269 Celsius).

Although the graph for H_2 shows the relationship between the potential energy and r_0 for the covalent bond, in reality, the sources of stability of any covalent bond are complex and require, in addition, a term for the kinetic energy of the electron. Fundamentally, the strength of a covalent bond is best described using quantum wave functions for the system. The average single covalent bond strength depends on the atoms bonded and varies between 30-120 kcal/mol (125-500 kJ/mol), a factor of 4.

Another confusing feature when discussing noncovalent interactions is that while we talk about forces (like the electrostatic force), we often draw graphs of energy E vs r, the distance between two interacting particles. Let's briefly examine the relationship between potential energy (E_{pot}) and force for the electrostatic force given by Coulomb's Law by using a more familiar example, the next gravitational force of a stationary ball placed at various points on a hill, as illustrated in Figure 2.4.3.

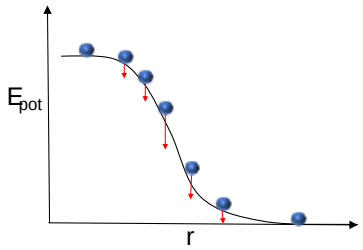


Figure 2.4.3: Potential Energy vs r for a ball on a hill

Assume the ball is motionless at each position in the diagram so only potential energy can be considered. The red arrows (vectors) represent the relative *net* downward force on the ball at each position. The *net* downward force at the top and bottom of the hill is zero. As the *slope* of the hill increases, the net downward force increases. The force is directly proportional to the slope (dE/dr), or simply:

$$F = -\frac{\Delta \mathrm{E}}{\Delta \mathrm{r}} = -\frac{\mathrm{dE}}{\mathrm{dr}} \tag{2.4.2}$$

Now let's apply this same relationship to Coulomb's Law for the force. Rearranging gives

$$dE = -\mathrm{F}\,\mathrm{dr} = -\frac{\mathrm{kq_1q_2}}{\mathrm{r^2}}\mathrm{dr} \tag{2.4.3}$$

Using calculus and integrating both sides of the equations gives this general relationship between E and r for the electrostatic forces:

$$E = \mathrm{kq_1 q_2}\left(\frac{1}{\mathrm{r}}\right) \tag{2.4.4}$$

A graph of E_{pot} vs r for the electrostatic force is shown in Figure 2.4.4. Note that the curves are hyperbolic (1/r) functions of r. There are attractive OR repulsive components.

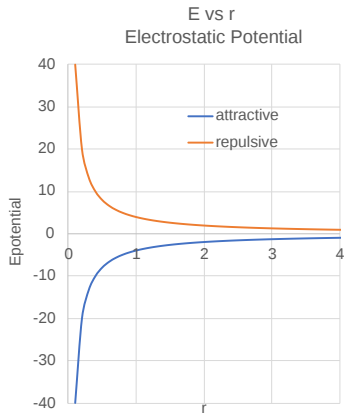


Figure 2.4.4: Electrostatic potential vs distance r

An equation for E_{pot} vs r for the induced dipole-induced dipole interactions can also be derived. For this interaction, E_{pot} has a different dependency on r and has both an attractive (E_{pot} α $-1/r^6$) AND repulsive term (E_{pot} α $+1/r^{12}$) which are added together. This potential is called the Lennard-Jones or 6-12 potential. Figure 2.4.5 shows the total and component attractive and repulsive terms E_{pot} vs r. Note how similar these curves are to the graphs for electrostatic energy.

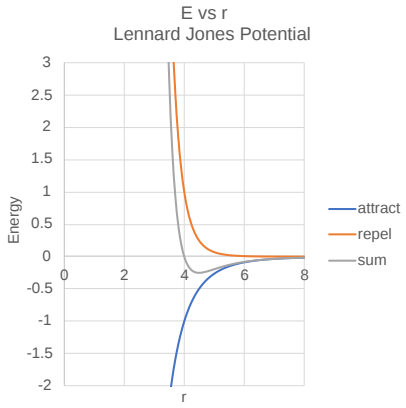


Figure 2.4.5: Leonnard Jones Potential

Now, let's look at the noncovalent interactions more carefully using examples of small and big molecules.

2.4.2: Ion-Ion

All introductory chemistry and biology textbooks differentiate ionic and covalent bonding. Ionic bonding occurs between fully charged species. Some ions are monatomic (like Na^+ or Cl^-), formed from gaining or losing electrons. Others are polyatomic (like ammonium - NH_4^+ or acetate - CH_3COO^-, generally formed from molecules gaining or losing protons in Brønsted acid/base reactions. Polyatomic ions are also called molecular ions. An example of the monatomic salt NaCl and the molecular salt ammonium acetate are shown in 2D Lewis structure and molecular modeling representations (spheres and sticks) in Figure 2.4.6: below.

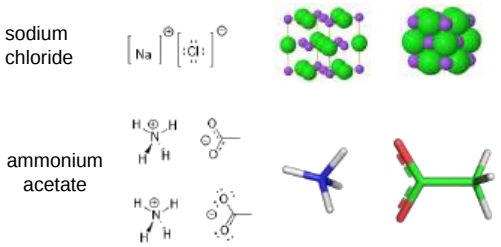


Figure 2.4.6: Representations of monatomic and polyatomic/molecular salts

Now, an intramolecular ionic bond can form within a larger molecule if a negatively charged group in the molecule comes close enough in 3D space to a positively charged group in the same molecule. In contrast to the examples shown above, the ionic bonds within large molecules like proteins do not occur within a large lattice of ions held together by multitudes of similar ionic bonds. Rather a single ionic bond could exist and persist in a larger molecule held together by a multitude of other noncovalent interactions. An ionic bond between a single monatomic or polyatomic cation and an anion would not exist in an aqueous solution long as the species would dissociate into separate ions solvated by water. Hence the ionic bonds that exist between charged groups within a large molecule like a protein exist in such a different environment than a solid crystal lattice that we give it a different name. It is called a ***salt bridge***, as the ionic bond bridges distal parts of the larger molecule. We also categorize it as an **ion-ion** noncovalent attraction.

Figure 2.4.7 shows a salt bridge/ion-ion interaction (represented as a yellow line) between the side chains of two amino acids, aspartic acid (Asp) 67 ($-CH_2COO^-$, similar to acetate) and lysine (Lys) 69 ($-RCH_2NH_3^+$, similar to NH_4^+) in a protein, human lysozyme.

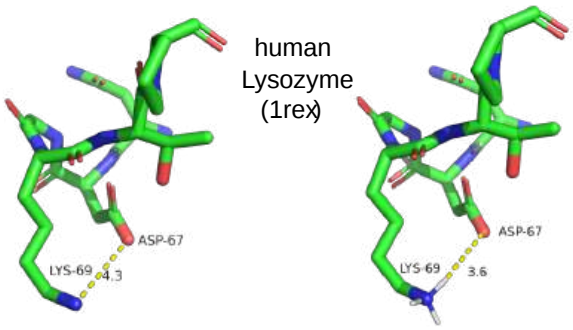


Figure 2.4.7: Salt bridge in human lysozyme

Figure 2.4.8 shows an interactive iCn3D model of a salt bridge/ion-ion interaction between the carboxylate side chain of Asp 67 and the amine side chain of Lys 69 in human lysozyme (1REX).

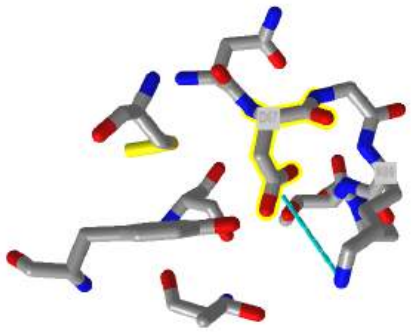


Figure 2.4.8: Salt bridge (represented as a yellow line) between Asp 67 and Lys 69 in human lysozyme (1REX). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...1qpAtSs3CuvVs8

Most of the protein's atoms have been removed to simplify the structure. We haven't studied proteins yet, but to a first approximation, they are polymers consisting of amino acid monomers. The backbone of the polymer contains a repeating amide

group which contains an N-H hydrogen bond donor and a C=O hydrogen bond acceptor. Each amino acid contains an R group side chain oriented away from the backbone. The R groups can be fully charged, polar or nonpolar.

This protein, containing 129 amino acids in a large polymer of over 1000 atoms, has just 10 salt bridges/ion-ion interactions within the most stable structure of the protein. The structure files that contain the x,y, and z coordinates of the atoms in a large biomacromolecule like a protein usually don't give coordinates for hydrogen atoms in the structure since they are too small to detect by techniques such as x-ray crystallography or cryoelectron microscopy, which are used to determine the structure of large biomacromolecules. Computer programs can be used to add them so they can be visualized in modeling programs. The left molecule in Figure 8 shows a stick model of just a small part of the protein containing a single salt bridge/ion-ion interaction. The blue represents nitrogen with a +1 formal charge in the side chain of lysine.

Hydrogen atoms have been added to the right molecule to illustrate the actual distance between adjacent atoms. Quantum calculations of actual electron density in molecular ions such as H_3O^+ and NH_4^+ (and charged amines) show that the electron density in these cations is actually shifted to the electronegative O and N atoms with electron deficiencies over the bonded H atoms (in contrast to the simpler ideal of formal charge), even though we state that the N in a charged amine has a positive formal charge.

As the distance r between interacting groups increases past the optimal interaction distance, the attractions decrease. When modeling most noncovalent interactions in large molecules, programs generally use cutoff values of 5-6 Angstroms, beyond which the interactions do not contribute to stabilization. The ion-ion interaction is the strongest interaction of all, given a fixed distance for comparison.

2.4.3: Hydrogen Bond (H-bond)

The name hydrogen bond is a bit ambiguous, which leads to its misunderstanding by students. It is **not** a covalent bond between two atoms, X and H, such as C-H and O-H in methane and water, respectively. Rather it is a noncovalent interaction between a slightly positive H on a electronegative atom X and a slightly negative electronegative atom Y on another molecule or part of a large molecule. X and Y are electronegative atoms such as F, O or N with lone pairs The H on X-H (for example O-H or N-H) is slightly positive (δ^+) since the X-H bond is polar covalent and electron density in the bond is drawn toward the electronegative atom (for example O or N). Given its small size compared to all other atoms, a slightly positive H, can get very close to a lone pair on a slightly negative (δ^-) electronegative atom Y (for example O or N) on another molecule. Since r, the distance between the δ^+ H and δ^- N or O on two separate molecules is small, Coulomb's Law informs us that the attractive force is significant. This interaction is highly directional and distance-dependent, which accounts for the large range in relative strength (3-15 kcal/mol, 12-63 kJ/mol) for hydrogen bonds within large molecules.

Hydrogen bonds occur between hydrogen bond donors and acceptors. This is determined by looking at the slightly negative electronegative atoms in the two interacting molecules. An alcohol (ROH) can be either a hydrogen bond donor or acceptor, while a ketone (R)C=OR) can only be an acceptor since it has no slightly positive H. In a hydrogen bond between an alcohol and a ketone, the O-H on the alcohol is the hydrogen bond **donor** while a O=C on the ketone is the hydrogen bond **acceptor**. This is illustrated in Figure 2.4.9 below.

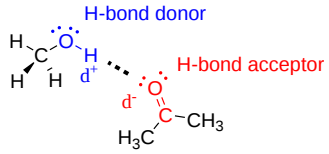


Figure 2.4.9: A hydrogen bond between a donor and acceptor

In a given hydrogen bond, the donor is the X-H with the slightly positive H.

Figure 2.4.9 shows multiple representations of a central water molecule hydrogen-bonded to four other water molecules. The left image shows lone pairs as purple spheres.

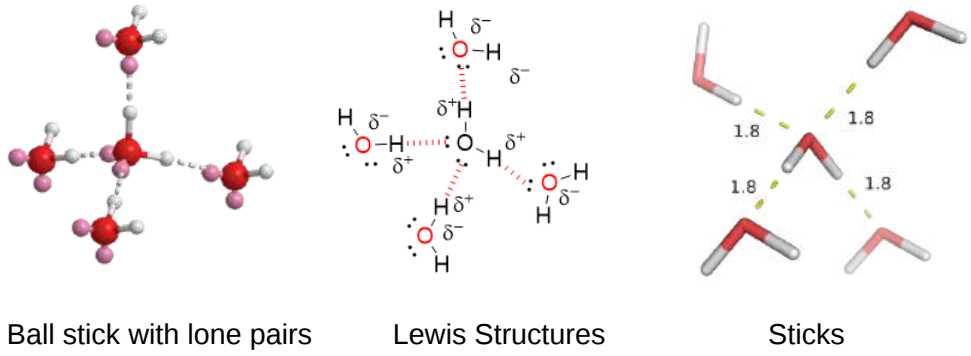


Figure 2.4.9: Multiple representations of hydrogen bonding among 5 water molecules

A common difficulty for students is to identify which of the many hydrogen atoms in any structure can engage in hydrogen bonds. One way is to circle all δ^+ Hs in structures (i.e. those covalently attached to N or O) and see if there are any nearby δ^-: N or :O atoms close enough to form a hydrogen bond. Figure 2.4.10 shows a molecule of methanol forming two hydrogen bonds to two different water molecules. Only 1 of the 4 Hs on methanol is δ^+ (circled in green). The others are bonded to the carbon through nonpolar covalent bonds.

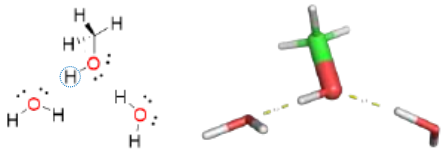


Figure 2.4.10: Hydrogen bond between methanol and waters

Hydrogen bonds are abundant in large molecules like proteins. They occur between backbone atoms, between backbone and side chains atoms, between side-chain atoms, and between protein atoms and water. Their strength depends on the magnitude of δ^+ and δ^- charges on the H bond donor and acceptor atoms, respectively, the distance r between them, and the angle of the bond Three types of H bonds have been categorized based on their relative strengths, based in large part on the distance between the donor and acceptor:

- weak or conventional, 2.4 to 12 kcal/mol (10-50 kJ/mol)
- strong or low barrier, 12 to 24 kcal/mol (50-100 kJ/mol), often called **short hydrogen bonds (SHB)**
- very strong or no barrier >24 kcal/mol (100 kJ/mol), (Frey et al).

In large proteins of known 3D structures, H bonds are calculated by locating all donors and acceptors with 3 +/- x angstroms from each other. Most structural files do not include H atoms, so the 3 Angstrom distance is from the centers of the electronegative atoms, typically N and O, involved in the hydrogen bond, as shown in Figure 2.4.11 (purple bracket).

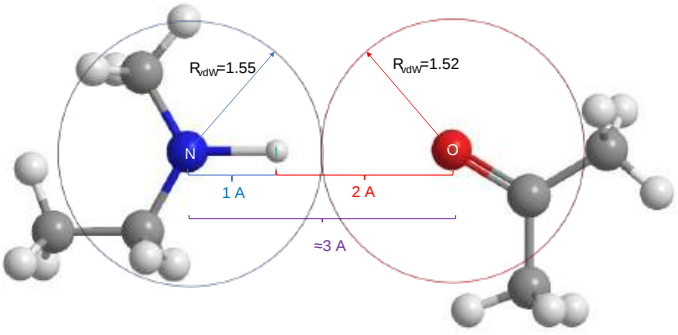


Figure 2.4.11: Hydrogen bonding distances between the center of a participating N with a H donor and an O with a H bond acceptor

Conventional H bonds vary between 2.8-3.2 A, which gives a distance range from the actual δ^+ hydrogen on the donor to the acceptor δ^- N or O (the red line below) of 1.8 to 2.2 A. Short H bonds are < 2.7A which is *smaller than the sums of the van der Waals radii of N and O* (blue and red circle below), suggesting that the bond has a covalent character (see below). Those between 2.5 - 2.7 are characterized as strong, low barrier, or short hydrogen bonds. Analysis of a large number of PDB structures of proteins shows many short hydrogen bonds characterized by these properties:

- the donor and acceptor electronegative atoms A and B are N or O
- r, the separation distance, is 2.3 A to 2.7 °A
- the A–H–B angle is 135^0.

Detailed analyses of high-quality protein structures show one short hydrogen bond for every 16 conventional ones. They are found in proteins, protein-ligand complexes and in DNA. They are involved in many aspects of molecular function.

It would seem likely that the δ^+ H atom, which is covalently attached to a heteroatom like O or N (A), and which is attracted to another heteroatom B, could be exchanged between the two heteroatoms as shown in the chemical equation below, where ---- represents an H bond.

A-H ----B ↔ A ----H-B

A very strong/no barrier H bond occurs if A and B are very close, have similar δ^- charges, and have similar pKa so that the H atom could be equally shared between A and B. It is represented by the representation below.

A ||| H ||| B

An example is FHF^- (F||| H |||F)$^-$ in which there is no barrier for the H to move from one heteroatom to another.

It thus appears that for strong and very strong H bonds, what we call the hydrogen bond has some covalent bond character. Quantum calculations show an overlap between the unoccupied antibonding σ*molecular orbital of X-H (the hydrogen donor) and the non-bonding lone electron pair molecular orbital of the hydrogen bond acceptor molecule.

Even though water is a simple and ubiquitous molecule, scientists still struggle to understand its properties. Lewis structures of water can explain only so much of its physical and chemical properties. However, look at Figure 2.4.12, which shows the electron density around water calculated using quantum theory.

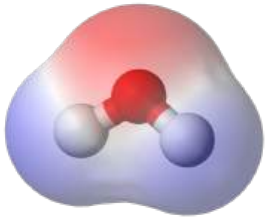


Figure 2.4.12: Electron density around a water molecule. https://commons.wikimedia.org/wiki/F...balls.png#file

Do you see any "rabbit ears" (i.e. lone pairs) emanating from the oxygen atom? Don't think so! Nevertheless, everyone still uses Lewis structures with lone pairs to explain the chemistry of water and other molecules. We present this figure, in advance of a discussion at the end of this section on the halogen bond, which requires an understanding of electron density around bonded atoms.

Now let's look at some hydrogen bonds within a single protein molecule. Figure 2.4.13 hydrogen bonds (yellow dotted line) between serine (Ser) 24 (side chain $-CH_2OH$) and asparagine (Asn) 27 (side chain $-CH_2(C=O)NH_2$ of hen egg white lysozome (1REX). As in the figures above showing salt bridges, two images are shown, one with polar H atoms added. Find the hydrogen bonds between side chains, side chains and backbone, and between backbone hydrogen bond donors and acceptors.

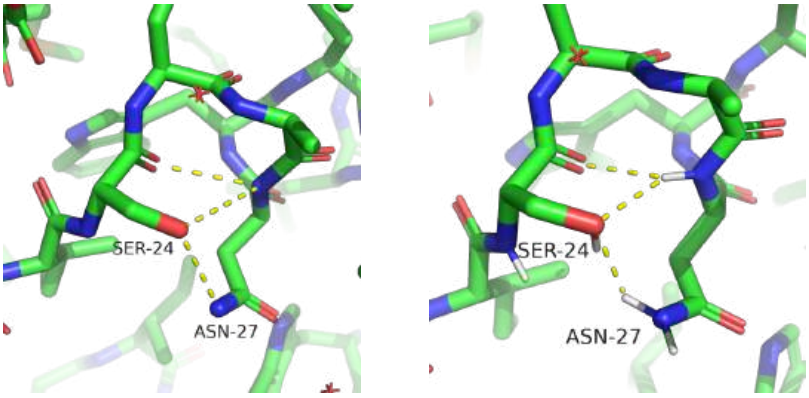


Figure 2.4.13: More hydrogen bonds in hen egg white lysozome (1REX)

Proteopedia has an excellent review of hydrogen bonds.

2.4.4: Dipole-Dipole

This interaction involves the alignment of permanent dipoles in molecules such that the geometric center of the δ^+ of one permanent dipole on one molecule is close to and aligned with the geometric center of δ^- of the permanent dipole on another. Figure 2.4.14 shows two acetone molecules interacting through dipole-dipole interactions. These molecules can't form hydrogen bonds to each other since they both contain just hydrogen bond acceptors.

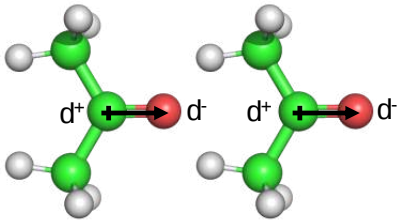


Figure 2.4.14: Attractions of two acetones through dipole-dipole interactions

The arrow represents the *molecule* dipole moment vector (as opposed to individual bond dipole moment for each polar covalent bond in the molecule). Note the difference in Figure 2.4.15. The molecular dipole is the vector sum of the bond dipoles.

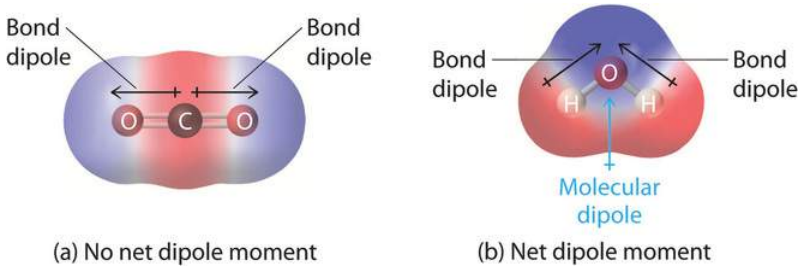


Figure 2.4.15: Difference between bond and molecular dipoles. https://chem.libretexts.org/Bookshel...bond%20dipoles.

None of the H atoms bonded to carbon in acetone are δ^+ so the molecules contain no H bond donors. Although they contain a δ^- oxygen, a hydrogen bond acceptor, two molecules of acetone cannot hydrogen bond to themselves. They can form hydrogen bonds to water. Pure liquid acetone evaporates readily (BP 56^0 C) due to this lack of strong hydrogen bonds.

You can imagine two water molecules forming dipole-dipole interactions as well. However tilting the molecule to align the lone pair on an O with the δ^+H on another water molecule and presto, you have a hydrogen bond. H bonds are often viewed as a special case of a dipole-dipole interaction.

Modeling programs can determine the charge on each atom of a large molecule like a protein and determine the geometric center and magnitude of overall + and - charge. A line drawn between them represents the permanent "dipole" moment of the entire protein. More simply, the molecular dipole is the vector sum of all of the individual bond dipole moments. Entire proteins have a

net dipole moment which probably facilitates the interaction of the protein with other proteins or ligands. Figure 2.4.16 shows the net dipole moment for the protein carboxypeptidase A1 (2v77). This was calculated using the Protein Dipole Moments Server. Proteins, however, do have net charges (not considering any bound counterions) so the molecular dipole for a protein is a bit different conceptually than for a small molecule. Nevertheless, it is a good way to quantitate asymmetric charge distribution in large biomolecules. Asymmetric charge distributions would influence molecular properties.

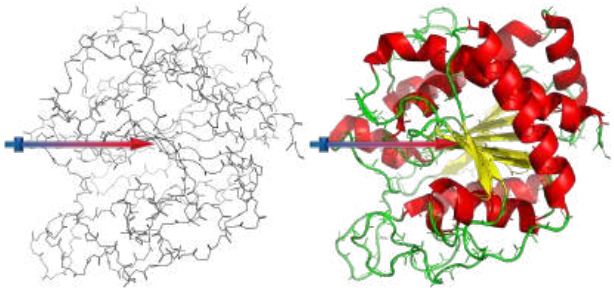


Figure 2.4.16: Net dipole moment of the protein carboxypeptidase A1 (2v77)

2.4.5: Ion-Dipole

Figure 2.4.17 shows interactions between a Na^+ ion and the dipoles of multiple water molecules.

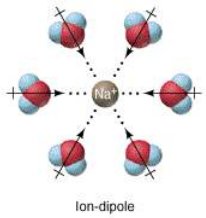


Figure 2.4.17: Ion-dipole interactions between Na^+ and water. https://chem.libretexts.org/Courses/...-Dipole_Forces

Figure 2.4.18 shows an interactive iCn3D model of the molecular ion sulfate SO_4^{2-} bound to a protein through its hydrogen bonding and ion-dipole noncovalent interactions with protein side chain and backbone groups in the sulfate binding protein from Salmonella typhimurium.

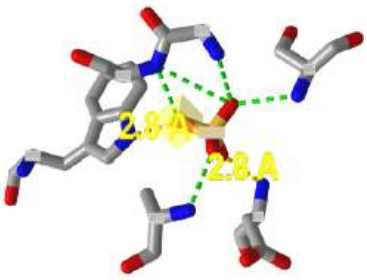


Figure 2.4.18: Sulfate bound to sulfate binding protein from S. typhimurium (1sbp). Hydrogen bonds are shown as green and yellow lines. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...JvqQuZZnvf3AU8

The SO_4^{2-} is buried within the protein. The green and yellow dotted line show hydrogen bonds between the sulfate and amide N-Hs on the protein chain surrounding it and the a side chain of the protein. Modeling programs don't show lines depicting dipole-x interactions. The SO_4^{2-}, through its oxygen, can form hydrogen bonds with nearby donors.

Figure 2.4.19 shows an interactive iCn3D model of another example of a protein backbone and side chains ion-dipole interactions, this time with a Na^+ ion, a simple non-transition state metal ion, which can not form hydrogen bonds. The protein is tryptophan synthase from Salmonella typhimurium (6dz4). The red spheres represent water oxygen atoms (no hydrogen atoms shown).

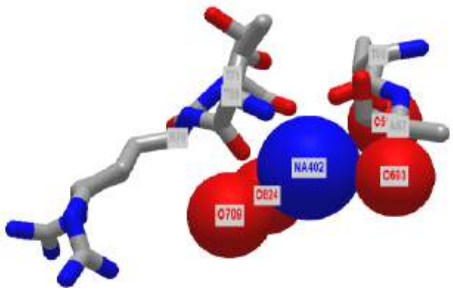


Figure 2.4.19: Sodium ion binding site in tryptophan synthase from Salmonella typhimurium (6dz4). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...pDnU36vXXPeyX7

The ions illustrated in these last two cases are not transition metal ions, whose interactions with ligands can best be considered using ligand field theory and the formation of covalent (dative) bonds between electron pair donors on nucleophilic side chain/main chain atoms and d orbitals on the transition metal.

2.4.6: Induced Dipole - Induced Dipole/Hydrophobic Interactions.

These noncovalent interactions occur when a temporary dipole, created by random fluctuations in electron density in one molecule, induces a temporary dipole in another molecule nearby. These interactions are weak and can easily be broken by raising the temperature. Induced dipole-induced dipole interactions allow nonpolar gases like He, N_2, O_2, and CH_4 to be liquefied, but it takes higher pressures and/or low temperatures to force the molecules close enough and slow them down enough for sufficient interactions to occur to liquefy the molecules. Although individually weak, the larger the molecules, the greater the extent of induced dipole-induced dipole interactions and the stronger the interactions among molecules. This is reflected in the fact that methane, CH_4, is a gas at room temperature, octane, C_8H_{18} is a liquid and $C_{30}H_{62}$ is a solid.

Figure 2.4.20 shows induced dipole interactions between two molecules.

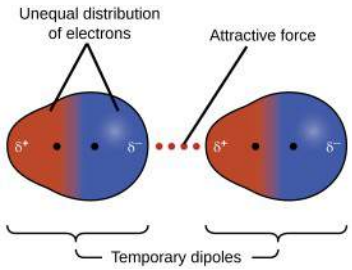


Figure 2.4.20: Induced dipole-induced dipole interactions. https://boisestate.pressbooks.pub/ch...ecular-forces/

Induced dipole-induced dipole interactions are important among large biomolecules as well. Most biologists and probably biochemists prefer to use the words hydrophobic interactions (but not hydrophobic forces) instead of the longer and more formal induced dipole-induced dipole interaction. We will also try to use the more commonly used term within the biochemistry community.

Figure 2.4.21 shows an interactive iCn3D model of a hydrophobic cluster around the side chain of a hydrophobic amino acid, valine 143 in human **carbonic anhydrase II (4ca2)**. Val 143 is highlighted in yellow and shown with normal atom (CPK) colors. White to green indicate nonpolar amino acids while dark blue indicates polar ones.

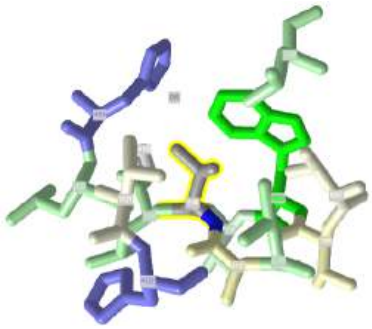


Figure 2.4.21: Hydrophobic cluster around the side chain of a hydrophobic amino acid, valine 143 in human carbonic anhydrase II (4ca2). Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...6rgjajJsg3Xvg8

You can see that the side chain of Val 143 (highlighted in yellow) is completely surrounded by nonpolar amino acids. If the structure was rendered in spacefill instead of sticks, Val 143 would be closely packed to maximize induced dipole-induced dipole (hydrophobic)interactions.

Induced dipole-induced dipole interactions also occur between polar molecules, but they are weaker than the hydrogen bonding and dipole-dipole interactions between them.

2.4.7: Pi stacking

Aromatic rings stacked over each other can interact through induced-induced dipole (hydrophobic) and dipole-induced dipole interactions. These interactions can depend on the presence of heteroatoms in the aromatic ring. Figure 2.4.22 shows an example with benzene in which a staggered arrangement of the rings is more attractive.

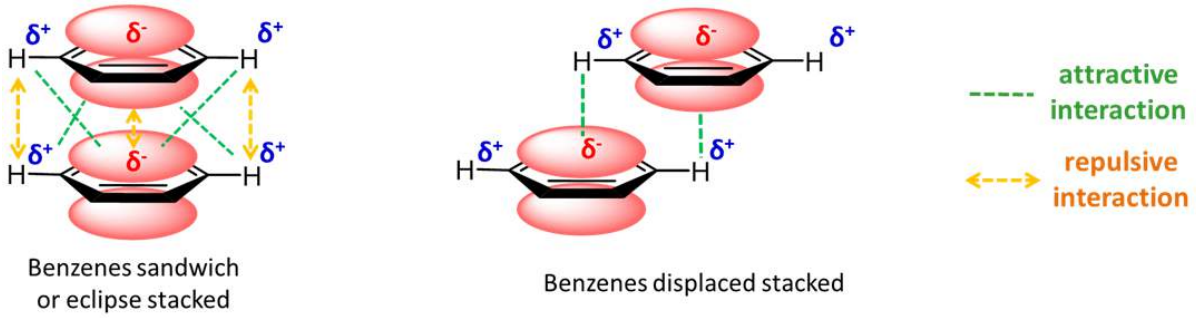


Figure 2.4.22: Pi stacking in benzene. https://chem.libretexts.org/Bookshel...g_interactions

For a biological example, everyone is familiar with the structure of B-DNA in which the bases A, G, C and T point inward perpendicular to the double helix axis and are stacked over each other.

Figure 2.4.23 shows an interactive iCn3D model of a short stretch of DNA with a sugar-phosphate backbone and bases colored in magenta and cyan. Fives bases on one strand are shown in stick and atomic color to show the pi-stacking interactions of the aromatic ring.

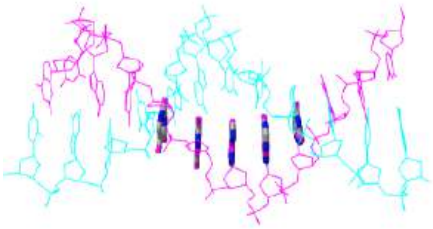


Figure 2.4.23: Pi stacking in B-DNA (5t4w). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...euHaVfXwk18JJ8

Pi stacking also occurs in proteins. Figure 2.4.24 shows an interactive iCn3D model of two sets of pi stacking interactions in the protein arginine kinase (1M15). The aromatic side chains involved in pi stacking are shown in cyan.

Figure 2.4.24: Pi stacking in Arginine Kinase (1M15). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...owSDJ2DUqtwk8A

2.4.8: Cation - Pi

Figure 2.4.25 shows an interactive iCn3D model a specific example of an ion-induced dipole interaction (called a cation-pi interaction) between a sodium ion (blue sphere) and the aromatic ring of the side chain tryptophan (cyan) in hen egg white lysozyme (1lpi).

Figure 2.4.25: Cation-Pi stacking (ion-induced dipole) in hen egg white lysozyme (1lpi). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...13gw1hmVYrBL8A

Example 2.4.1

For another example of a cation-pi interaction, open up iCn3D with 1REX and view the interaction of lysine (K1) side chain with the nonpolar aromatic ring of phenylalanine (F3).

Solution

https://structure.ncbi.nlm.nih.gov/i...4d3wXXsSEYhgv7

Here are some more examples.

Exercise 2.4.1

Select the link below to answer the following questions.

1. What type of noncovalent interaction best describes the red dotted line in the structure?

2. What type of noncovalent interaction best describes the red dotted line in the structure?

Answer

1. cation-pi
2. pi stacking

2.4.9: Halogen Bond

Lastly, we come to the halogen bond. You might ask if there are halogens found in proteins. The answer is no (until one is found!) but halogenated molecules (drugs, xenobiotics, toxins) bind proteins. Consider the C-X bond where X is a halogen. The electronegativity of C is 2.56 while the halogens have these electronegativity values: F (3.98), Cl (3.16), Br (2.96), and I (2.66). Compare these to oxygen (3.44) and N (3.04). Covalent bonds between two bonded atoms whose electronegativity differences are between 0.4 and 1.8 are considered polar covalent, so C-F, C-Cl, and C-Br are considered polar covalent. The C-I bond is the longest and iodine is the most polarizable of these halogens. An alkyl halide with a C-I bond can undergo S_N2 nucleophilic substitution reactions with I^- being an excellent leaving group. Hence the C-I bond behaves somewhat as a polar covalent bond.

Nevertheless, quantum calculations show that the electron density is not uniformly spread around the X halogen in a C-X bond, but rather is pulled more toward the C, leaving the distal *end* of the halogen depleted in electron density and slightly positive. This region of relatively depleted electron density is called the σ-hole. Color-coded renderings of the electron density of the halogen involved in a C-X bond show the halogen atom to have bands (like Jupiter) with the more negative electrostatic potential (represented in blue) closest to C and the more positive potential, the σ-hole (represented in red), at the end farthest from the C atom. Calculations show that this effect is greatest for the heavier halogens (Br, I) which have longer C-X bonds. The halogen's slightly positive σ-hole can act analogously to a hydrogen bond donor in its interactions with nearby δ^- :O and :N atoms/lone pairs. This might take a while to grasp. You have always heard that in general the halogens are more electronegative than carbon and would hence always be δ^- when bonded to it. This case is similar to our chemical intuition about lone pair "rabbit ears" on oxygen, which quantum calculations show not to be an accurate representation of the electron density (see Fig xx).

Figure 2.4.26 shows the electrostatic potential on a halogen X atom covalently attached to a carbon in two different molecules, CF_3-I and :NC-Br. The red distal end is the σ-hole relatively depleted in electron density and with a higher, more positive electrostatic potential. (This is opposite the usual coloration that biochemists use in which oxygen (δ^- or fully -) is colored red and nitrogen (in a protonated amine with a positive charge) is shown in blue.)

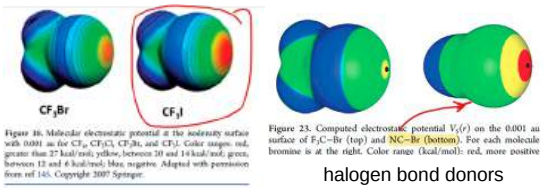


Figure 2.4.26: Electrostatic potential on a halogen X atom on CF_3-I and :NC-Br

Figure 2.4.27 shows a molecule with a carbonyl (a hydrogen bond acceptor with a δ^- :O) interacting with another molecule through either a hydrogen bond or a halogen bond. Again the red distal end of the halogen X is the σ-hole relatively depleted in electron density.

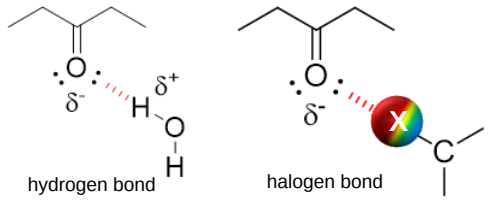


Figure 2.4.27: Comparison of a hydrogen and halogen bond (noncovalent interaction)

Medicinal chemists use halogen substituents on drug molecules to alter drug binding specificity, membrane diffusion, and $t_{1/2}$. Increasingly, they are using halogen bonds in rational drug design to increase drug affinity to target proteins.

Figure 2.4.28 shows an interactive iCn3D model below shows the interaction of a haloaminopyrimidine inhibitor bound to its binding site on the c-Jun N-Terminal Kinase (JNK) protein (2P33).

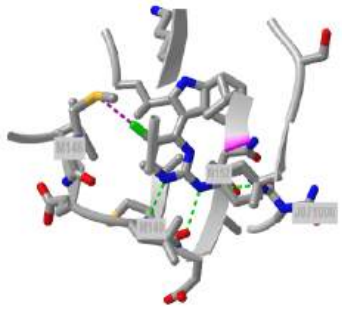


Figure 2.4.28: Haloaminopyrimidine inhibitor bound to its binding site on the c-Jun N-Terminal Kinase (JNK) protein (2P33). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?H4Wabj8my3VsYLSi6

Note that the sulfur of methionine is forming a halogen bond with the Cl atom. Although the electronegativity of sulfur is 2.58, close to that of carbon (2.55), nevertheless, sulfur is larger and more polarizable so it also develops a slightly positive σ-hole distal to the carbon atom. Analysis of PDB files shows that S--O interactions are common in proteins and most likely impact protein stability.

Ultimately all ensembles of molecules/ions reach a low if not the lowest energy state under a given set of conditions. Noncovalent attractions are maximized and repulsions are minimized to achieve this state. Consider for example solid sodium chloride held together by ionic bonds. The ions are closest packed (face-centered cubic) and cannot get closer together (packing density of about 74%) as simple packing considerations and repulsive electrostatic forces and collective van der Waals interactions would prevent it. Each Na^+ is surrounded by 6 Cl^- ions and vice versa.

When large molecules like proteins assume a low energy state, they maximize the attractive noncovalent interactions described in this section while minimizing repulsive ones within a molecule (in given solvent conditions). Packing density reaches similar values as for closest packed spheres (NaCl for example). Figure 2.4.29 shows a slice through a protein and through the crystal lattice of NaCl. The gray circles on the protein show the faces of the sliced atoms. They are superimposed on the surface of the protein shown in colored spheres. If you took a series of cross-sectional slices throughout the protein, you would get a better picture of packing density than a single slice alone. Collective van der Waals interactions are found among all atoms and ions in a protein, which accounts for the closest packing of most atoms, polar and nonpolar, with the packed protein structure.

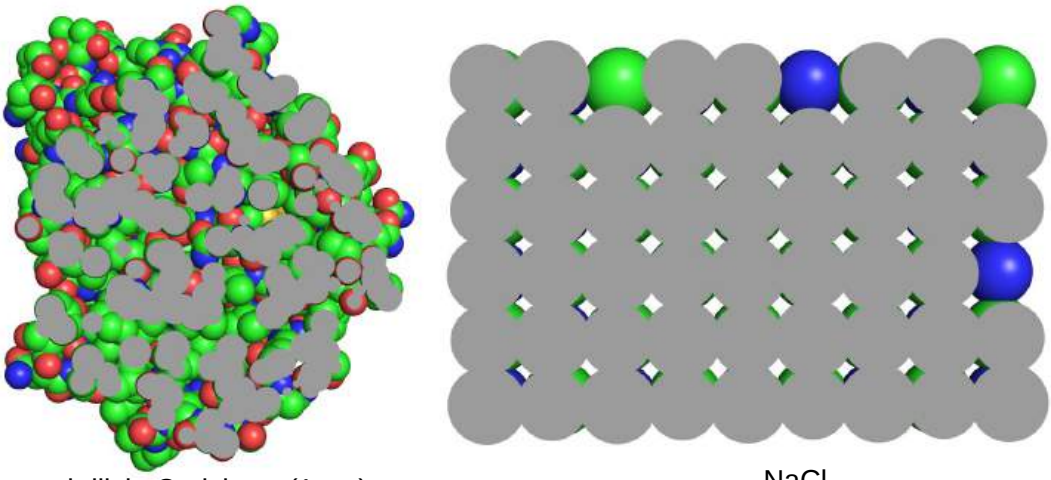


Figure 2.4.29: A slice through a protein and through the crystal lattice of NaCl

Here is a link to a JSmol tutorial by David Marcy et al, An Introduction to Chemical Bonds and Protein Structure

2.4.10: Summary of Noncovalent Interactions in Biomolecules

It is not easy to understand noncovalent interactions among small molecules let alone within solvated and densely packed proteins, for example. To help quantitate strong noncovalent interactions involving amino acid *side chains*, Xie et al have studied amino acids in the gaseous phase using quantum mechanics. Here are some general conclusions (Xie et al. PLoS ONE 10(9): e0137113. https://doi.org/10.1371/journal.pone.0137113. Creative Commons Attribution License)

- Ion-Ion (salt bridge) interactions between acidic amino acids side chains (Glu^- and Asp^-) and alkaline amino acids side chains (Arg^+, Lys^+ and His^+) are the strongest residue-residue interactions. However, this type of interaction may be weakened by solvation effects and broken by lower pH conditions.
- Cation- interactions between protonated amino acid side chains (Arg^+, Lys^+ and His^+) and aromatic amino acid side chains (Phe, Tyr, Trp and His) are 2.5 to 5-fold stronger than common hydrogen bond interactions and are less affected by the solvation environment.
- Amide bridge interactions, shown in Figure 2.4.30 below, which contain two hydrogen bonds, between the two amide-containing amino acid side chains (in the amino acids Asn and Gln) are three times stronger than hydrogen bond interactions, which are less influenced by the pH of the solution.

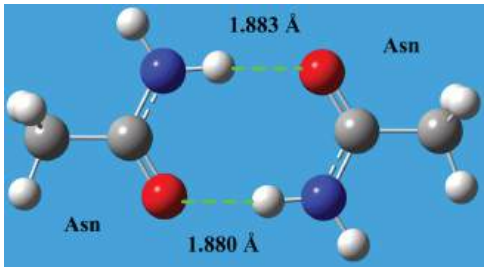


Figure 2.4.30: An "amide" bridge (Xie et al, ibid)

- Ten of the twenty natural amino acids are involved in salt bridge, cation, or amide bridge interactions that often play important roles in protein-protein, protein-peptide, protein-ligand, and protein-DNA interactions.

Another computational study was done to categorize the noncovalent interactions between proteins and small molecules (drugs, inhibitors) that bind to them. These small molecules are generically called **ligands**, a term used in the study of transition metal complexes. They studied 11,016 unique structures found in the Protein Data Bank of small-molecule ligands bound to proteins. A histogram displaying the number of each type of interaction between small ligands and proteins is shown in Figure 2.4.31 below.

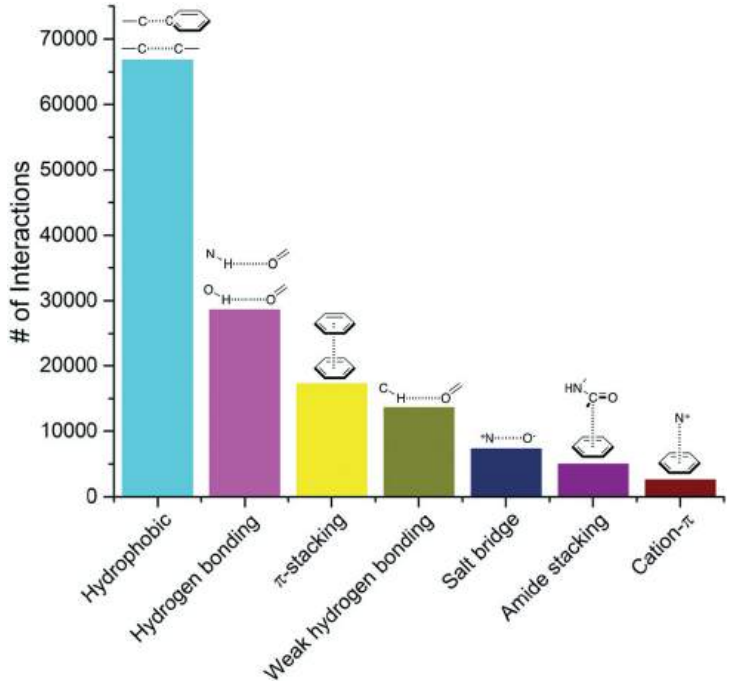


Figure 2.4.31: Frequency distribution of the most common non-covalent interactions observed in protein–ligands extracted from the PDB. de Freitas and Schapira. *Med. Chem. Commun.*, 2017, **8**, 1970-1981. **de Freitas and Schapira.** DOI: 10.1039/C7MD00381A (Research Article)

Two new interactions are shown, the weak hydrogen bond and amide stacking. Amide stacking is readily understandable as an interaction between the slight positive carbonyl carbon of the amide and the electron-dense aromatic ring.

You probably find the weak hydrogen bond more troubling. The hydrogen bond donor is a carbon atom attached to a hydrogen, and a hydrogen bond acceptor, the carbonyl oxygen. We have stated that a C-H bond does not engage in a hydrogen bond. We assume that a C-H bond is sufficiently nonpolar so the carbon atom does not have a slight negative charge which leaves hydrogen without a slight positive. Yet it appears that there are many C–H···O weak hydrogen bonds between ligands and proteins.

C has an electronegativity of 2.5 and H 2.2 with a difference of 0.3, which is a much smaller difference than between N and H (3.04-2.2 =0.8). If the electronic environment around the carbon enhances its slight negative charge, then you could imagine that a C-H could be a hydrogen bond donor. The median distance of the C–H···O hydrogen bonding was 3.4 Å, which is 0.4 Å longer than traditional hydrogen bonds (N–H···O, N–H···N, O–H···O), with an angle of around 130°. The C_{α}–H···O=C interactions are about one-half the strength of an NH···O=C hydrogen bond. Hence they are weak. In the rest of this book, we will **NOT** consider a C-H bond as a candidate for a hydrogen bond. Nevertheless, it is important to mention it given their prevalence.

2.4.11: References

IUPAC. Compendium of Chemical Terminology, 2nd ed. (the "Gold Book"). Compiled by A. D. McNaught and A. Wilkinson. Blackwell Scientific Publications, Oxford (1997). Online version (2019-) created by S. J. Chalk. ISBN 0-9678550-9-8. https://doi.org/10.1351/goldbook.

https://chem.libretexts.org/Bookshel...g_interactions

https://boisestate.pressbooks.pub/ch...ecular-forces/

A low-barrier hydrogen bond in the catalytic triad of serine proteases, PA Frey et al, *Science* 264, 1927-1930 (1994)
DOI: 10.1126/science.7661899

2.5: Solubility in an aqueous world - The Hydrophobic Effect

2.5.1: Introduction

Many biomolecules such as triacylglycerols, cholesterol esters, and waxes are nonpolar. Other biomolecules such as proteins and many lipids have both polar and nonpolar parts. We know from experience that oil floats on the surface of water, showing that it is less dense than water and that it doesn't dissolve in water. You have also probably performed liquid-liquid extractions in chemistry labs in which you utilized the solubility properties of nonpolar molecules to extract them from a mixture in water to a more nonpolar phase such as octanol or chloroform. To understand the stability of biomolecules that contain nonpolar parts in aqueous solutions, we need to understand not only noncovalent interactions of the molecules with water (which we explored in Chapter 2.4) but also the thermodynamics of their molecular interactions in aqueous environments.

We have been taught and internalized the notion that "like-dissolves like". We anthropomorphize molecules to say nonpolar molecules "like" to be in nonpolar environments. We can rationalize solubility properties by examining the noncovalent attractive and repulsive interactions of a molecule in an aqueous solution but when we do so we are usually focusing on enthalpic contributions to stability. What about entropy? We should consider net changes in noncovalent solute:solute, solute:solvent, and solvent:solvent interactions, as well as their thermodynamic contributions to overall stability. When we consider the thermodynamics of the solubility of molecules in water, we need to determine the ΔG, the free energy change, for all processes involved.

2.5.2: The Change in Free Energy (G) and Chemical Potential (μ)

ΔG, the free energy change for a reaction, determines the spontaneity and extent of a chemical or physical reaction. The free energy of a system depends on 3 variables, temperature T, pressure P, and n, the number of moles of each substance. For the latter, think of solute X on two different sides of a permeable membrane. If the concentration of X is the same on each side, as shown in Figure 2.5.1, the system is in equilibrium.

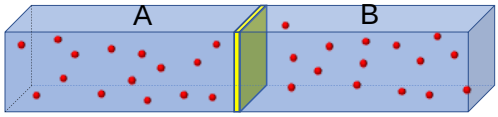


Figure 2.5.1: A system of a molecule in two compartments separated by a membrane

If the system is composed of two different parts, A and B, the system is at equilibrium (ΔG=0) if $T_A = T_B$, $P_A = P_B$, and the change in the absolute free energy per mole of A is $\Delta G_A/\Delta n = \Delta G_B/\Delta n$. More precisely, using simple calculus, we would discuss incremental changes in absolute free energy/mol, dG_A/dn for A, which is the chemical potential of A, μ_A) and dG_B/dn (μ_B) for B. At equilibrium $dG_A/dn = dG_B/dn$. We will use the free energy G here but μ later in this section. G then is the absolute free energy/mol (again chemical potential), where $G=G^o + RTln[A]$. From this the equations you used in introductory chemistry (and that we reviewed in Chapter 1.3), the following equation can be written.

$$\begin{aligned} \Delta G &= \Delta G^0 + RTInQr \\ \Delta G &= \Delta H - T\Delta S \\ \Delta G^0 &= \Delta H^0 - T\Delta S^0 \\ \Delta G^0 &= -RTInKeq \end{aligned} \tag{2.5.1}$$

Now let's apply this to the chemical equation for the solubility of a given solute in water. If you add either a sparing soluble hydrocarbon (HC) or sodium chloride to water, eventually you reach a point of saturation. For the salt, the water is saturated with dissolved NaCl, and no further increase in NaCl (aq) occurs. For a sparing soluble hydrocarbon, the solution reaches saturation after which phase separation occurs.

Let's add a drop of a slightly soluble hydrocarbon liquid (HC_L) into water, as pictured in the diagram below. At t=0, the system is not at equilibrium and some of the HC will transfer from the pure liquid to water, so at time t=0, $\mathbf{\Delta G_{TOT}} < 0$. This is illustrated in Figure 2.5.2.

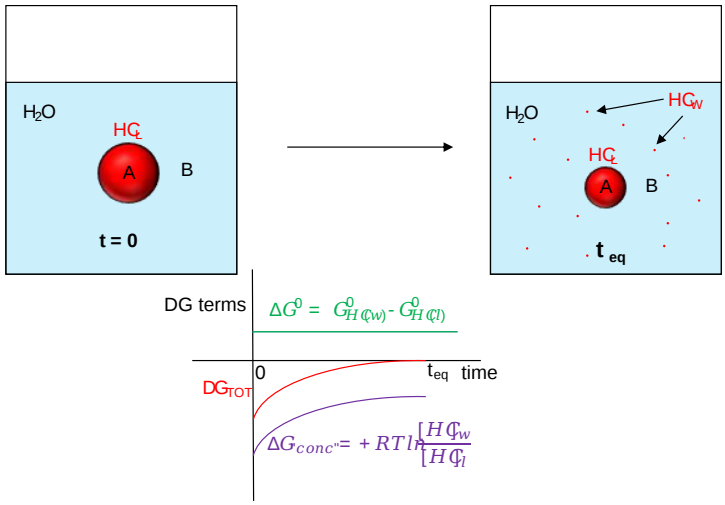


Figure 2.5.2: ΔG vs time for interaction of a hydrocarbon with water

The following equations can be derived.

$$\begin{aligned} \Delta G_{TOT} = (G_{HC-W}) - (G_{HC-L}) = G^0_{HC-W} + RT\ln[HC]_W - \left(G^0_{HC-L} + RT\ln[HC]_L\right) = \\ \Delta G_{TOT} = \left(G^0_{HC-W} - G^0_{HC-L}\right) + RT\ln([HC]_W - \ln[HC]_L) = \\ \Delta G_{TOT} = \Delta G^0 + RT\ln\frac{[HC]_W}{[HC]_L} \end{aligned} \tag{2.5.2}$$

Now add a bit more complexity to the last example. Add a hydrocarbon x, to a biphasic system of water and octanol as shown in Figure 2.5.3. Shake it vigorously. At equilibrium, x would have "partitioned" between the two mostly immiscible phases.

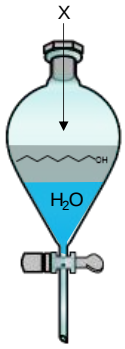


Figure 2.5.3: Use of separatory funnel for separating immiscible liquid phases

A simple reaction can be written for this system: X aq ↔ X oct.

Clearly, if X is a hydrocarbon, $\Delta G < 0$ for the reaction written above. Also, $\Delta G^o < 0$, since this term is independent of concentration and depends only on the intrinsic stability of X in water in comparison to that of octanol. This simple equation holds:

$$\Delta G_{TOT} = \left(G^0_{X-oct} - G^0_{X-w}\right) + RT\ln\frac{[X]_{oct}}{[X]_w} = \Delta G^0 + RT\ln\frac{[X]_{oct}}{[X]_w} \tag{2.5.3}$$

At equilibrium, ΔG^0=0 and the equation can be rewritten as:

$$\Delta G^0 = -RT\ln\frac{[X]_{oct}}{[X]_w} = -RTlnK_{part} \tag{2.5.4}$$

where K_{part} is the equilibrium partition coefficient for X in octanol and water. This can readily be determined in the lab. Just shake a separatory flask with a biphasic system of octanol and water after injecting a bit of X. Then separate the layers and determine the concentration of x in each phase. Plug these numbers into the last equation. You should be able to predict the sign and relative magnitude of ΔG^o since it does not depend on concentration, but only on the intrinsic stability of the molecules in the different

environments. K_{part} values are often determined for drugs since they often must diffuse across cell membranes to move into the cytoplasm where they can act. Drugs hence must have a reasonable K_{part} to pass through the membrane but not so high that they are insoluble.

2.5.3: Introduction to the Hydrophobic Effect

Now let's ask this question: What are the enthalpic and entropic contributions to the **ΔG** for the interaction of a nonpolar molecule HC with water? For this section, we will replace ΔG with Δμ (the change in chemical potential but we will use these terms interchangeably). Likewise, we will use this equation: $\Delta\mu^o = \Delta H^o - T\,\Delta S^o$.

Also instead of framing the reaction as the dissolution of an organic molecule in water, we will frame is as the **transfer of a hydrocarbon X from an aqueous solution to the pure hydrocarbon liquid (HC) or**

$$X(aq) \leftrightarrow X(HC) \tag{2.5.5}$$

Figure 2.5.4 shows the standard free energies of transfer of a hydrocarbon X from an aqueous solution to a pure liquid hydrocarbon (HC), X (aq) ↔ X (HC). where

$$\Delta\mu^\circ = \mu^\circ x(HC) - \mu^\circ x(aq) \tag{2.5.6}$$

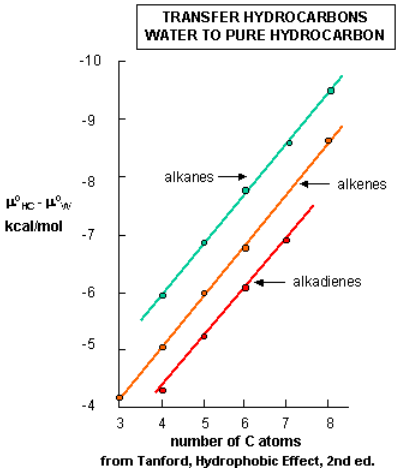


Figure 2.5.4: Standard free energies of transfer (μ^0) of a hydrocarbon X from aqueous solution to a pure liquid hydrocarbon (HC)

$\Delta\mu^o$ is less than 0 since transfer back to the pure HC is favored from a stability perspective. In each graph, $\Delta\mu^o$ is less than 0, and the value of $\Delta\mu^o$ decreases (gets more negative as you go up the y axis which shows increasingly negative values of $\Delta\mu^o$) in a linear fashion with increasing numbers of carbon atoms in the alkyl chain. Notice the lines are unbelievably straight and parallel. Nature is speaking to us in these figures. By determining the surface area of the hydrocarbon molecules and the decrease in $\Delta\mu^o$ with each added CH_2 (methylene group), one can calculate that the $\Delta\mu^o$ decreases by 25 cal/Å^2 (105 J/Å^2), per methylene added.

We expected that $\Delta\mu^o$ for the transfer of X to a pure liquid HC would be negative. We could get more information if we would determine both the entropic and enthalpic contributions. Such data is presented in the table below, which shows the transfer of short, single-chain alcohol X (an amphiphile with a polar head and a longer nonpolar "tail") from the pure liquid alcohol (ROH) to water (the opposite of the previous figures.)

$$X(ROH) \leftrightarrow X(W) \tag{2.5.7}$$

Thermodynamic Parameters for Transfer of Aliphatic Alcohol X from the **Pure Liquid to Water** at 25°C (enthalpy determined by calorimetry)

alcohol X	μ_w^0-μ_{ROH}^0 kcal/mol (kJ/mol)	H_w^0-H_{ROH}^0 kcal/mol (kJ/mol)	S_w^0-S_{ROH}^0 cal/deg mol (J/deg mol)	$(Cp)_w^0$-$(Cp)_{ROH}^0$ **cal**/deg mol (J/deg mol)
ethanol	0.760 (3.18)	-2.43 (-10.2)	-10.7 (-44.8)	39 (163)

n-propanol	1.58 (6.61)	-2.42 (-10.2)	-13.4 (-56.1)	56 (234)
n-butanol	2.4 (10)	-2.25 (-9.41)	-15.6 (-65.3)	72 (301)
n-pentanol (solubility 22g/L H_2O)	3.22 (13.5)	-1.87 (-7.82)	-17.1 (-71.5)	84 (352)

We expect the $\Delta\mu^o$ to be increasingly positive as the chain length gets longer and their solubilities in water become increasingly disfavored. What is perplexing about this data is not that the transfer of these ROHs to water is disfavored, but that transfer is **enthalpically** favored (negative ΔH^0). This seems to be counterintuitive since it goes against the adage that "like dissolves like" as was discussed earlier. From an enthalpic point of view, the amphiphiles prefer (albeit marginally) to be in water. What makes this reaction disfavored is entropy. The data shows that the nonpolar molecule would prefer not to be in the water because it is disfavored entropically.

At first glance, you might guess that the entropy should favor the movement of ROHs into water since they could access a larger volume to access and have greater freedom of motion. Hence here are more possible microstates for the ROH in water. However, this is only part of the process. What we haven't considered is the entropy of the water. To place a hydrocarbon in water, a literal cavity in the water must be created that will accommodate it. Creation of this more ordered cavity must be entropically disfavored (again because the process proceeds to a state with fewer microstates and lower positional entropy).

For the reverse process, transferring the hydrocarbon from water to the pure liquid will dissipate the cavity, and lead to more available microstates for the released solvent, bulk water. It is this entropic contribution that favors the movement of a hydrocarbon from water to the pure hydrocarbon lipid. This "**hydrophobic effect**" is the main thermodynamic drive to move organic molecules out of water.

Image this scenario. When you place a hydrocarbon group into water, water seeks (admittedly an anthropomorphic term) to maintain its hydrogen bonding. Hence it is forced into a more ordered structure around the HC to maintain its H-bonding, characterized by fewer microstates. We will explore the hydrophobic effect in greater detail in a future chapter.

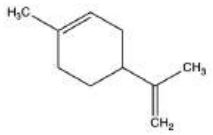
How can we explain the favorable enthalpic contribution on placing a nonpolar molecule into water? Again, this goes against our adage of "like dissolves like". The negative ΔH suggests interactions among all the participants are more favorable when the nonpolar group is in water. One source of such interactions could be the highly structured water in the "cage" that surrounds the nonpolar molecule. If it were more structured than bulk water, hence more "ice-like" in nature, then the formation of these extra H-bonds would contribute to the negative enthalpy change. When the nonpolar molecule is removed from the water, a process which proceeds with a positive ΔH, the cage of "ice-like" water would "melt", which, like the melting of ice, is not favored enthalpically, as heat must be added. Heat energy must be supplied to break the H-bonds as ice changes state to liquid water. This molecular model to understand the thermodynamic data might yet be a simplistic model, but for time being, let's use it.

2.6: Chapter 2 Questions

Section 1 Questions:

Q1) Based on the interactive figure 2.1.1 in the default view, do you hypothesize that water could enter the core of the micelle? Now, let's test your hypothesis! Open the interactive figure, and click Style → Surface Type → Solvent Accessibility. Based on the space-filled solvent accessibility map, do you see any openings for water molecules to enter through? Explain your answer.
A1) Hypothesis can range from no water is accessible, or water could freely occupy the "empty" space in the default view of the micelle. However, once the solvent accessibility filter is applied, it is clear that the micelle will exclude water from entering the hydrophobic core. Some areas of the micelle (the polar head groups) are able to for interactions with the water and therefore show up as red/yellow in this view, but it is key to note there are no pores or solvent-accessible spaces on the surface of the micelle.

Q2) The structure for Limonene, the compound that gives citrus fruits their classic smell is shown below. Examine the structure and answer the following questions.



a) Will Limonene form any associations with water molecules? Explain.
b) For our brain to register the citrus aroma, Limonene needs to bind to a surface protein receptor, what type of interactions could be taking place? Explain.
c) When Limonene binds to its receptor are the interactions from b) stabilized by enthalpy ΔH or entropy ΔS?

A2)
a) No, there are no dipoles or polar functional groups on limonene.
b) The methyl and ethyl functional groups would likely form hydrophobic bonds with nonpolar amino acids in the protein. Because there are no polar/dipole residues, Limonene cannot make hydrogen or ionic bonds.
c) Binding of Limonene to its surface receptor will likely be stabilized by entropy. When the nonpolar Limonene binds to a nonpolar region of the surface receptor, order water will be released, thus creating a +ΔS.

Q3) Consider the following reaction of a ***polar*** substance and water at room temperature (22°C):
$Glucose_{(s)} + H_2O_{(l)} \rightarrow Glucose_{(aq)}$
a) Estimate the enthalpy ΔH and entropy ΔS (+,-,≈0) for the reactants and product. Consider the order/disorder of the over all reaction as well as the net charge of the bond enthalpy.
b) Based on the Gibbs free energy equation ΔG = ΔH - TΔS, will this reaction be spontaneous, non-spontaneous, or at equilibrium?

A3)
a) $Glucose_{(s)} + H_2O_{(aq)} \rightarrow Glucose_{(aq)}$
ΔH: H-bonds H-bonds H-bonds ΔH ≈ 0
ΔS: Solid Liquid Solution ΔS = +

b) ΔG = ΔH - TΔS
= 0 - (+)
= -
This dissolution of polar substances into water is spontaneous.

Q4) Consider the following reaction of a ***nonpolar*** substance and water at room temperature (22°C):
$Methane\ (CH_4)_{(g)} + H_2O_{(l)} \rightarrow Methane\ (CH_4)_{(aq)}$
a) Estimate the enthalpy ΔH and entropy ΔS (+,-,≈0) for the reactants and product. Consider the order/disorder of the overall reaction as well as the net charge of the bond enthalpy.
b) Based on the Gibbs free energy equation ΔG = ΔH - TΔS, will this reaction be spontaneous, non-spontaneous, or at equilibrium?

A4)
a) $Methane\ (CH_4)_{(g)} + H_2O_{(l)} \rightarrow Methane\ (CH_4)_{(aq)}$

ΔH: None H-bonds H-bonds ΔH ≈ 0 or slightly -
ΔS: Gas Liquid Solution ΔS = -

b) ΔG = ΔH - TΔS
= 0/- - (-)
= +
This dissolution of nonpolar substances into water is non-spontaneous. This is primarily due to water forming an ordered cage around the nonpolar methane gas. In terms of enthalpy, the water cage creates a more favorable environment for methane than the pure gas in liquid water, thus the enthalpy can be considered slightly negative. However, the cage structure the water molecules form around the methane gas makes the system ordered. Therefore the overall ΔG for nonpolar substances solubilized with water is positive and therefore non-spontaneous.

Need to add in some reactions? (last chunk of section 1)

Section 2 Questions

Q1) Tyrosine is commonly found in the active sites of enzymes, as the unique structure of its R-group can act as either an acid or base.

a) The enzyme DcpS, an mRNA capping enzyme, utilizes a tyrosine residue as an active site acid, and measurements show that the tyrosine is 75% ionized. What must the local pH need to be for this to occur?
b) Another way for tyrosine to be used as an acid is to lower its pKa by creating weak interactions with the oxygen of its R-group using neighboring amino acids. What type of interaction(s) could achieve this goal?
c) One enzyme utilizing the strategy in b) is Glutathione-S-Transferase. Measurements show that the active site pH is 7.5, and the tyrosine residue is 95% ionized. What must the pKa of tyrosine be for this to occur?

A1)
a)

$$pH = pK_a + log\frac{[A^-]}{[HA]}$$

$$pH = 10.1 + log\frac{[75]}{[25]}$$

pH = 10.1 + (0.477)
pH = 10.58

b) Hydrogen bonding can create a change in the net dipole of the oxygen in tyrosine by making it more acidic and decreasing the pKa.
c)

$$pH = pK_a + log\frac{[A^-]}{[HA]}$$

$$7.5 = pK_a + log\frac{[95]}{[5]}$$

7.5 = pK_a + 1.28
6.22 = pK_a

Section 3 Questions

Q1) Your lab wants to study an enzyme that catalyzes a reaction inside the chloroplast stroma, which has a pH of 8.0 due to the proton gradient that pumps H^+ from the stroma to the thylakoid lumen.
a) Which of the buffers below would be the best choice to study this enzyme, in vitro? Explain your choice. Are there any other buffers that could work?
-Tricine: pKa 8.05
-TAPS: pKa 8.40
-MES: pKa 6.1
-Citrate: pKa 6.40
-HEPES: pKa 7.48
b) How many moles of the conjugate base form of HEPES would there be in 2.5 L of a 175 mM solution at pH 8.0?

A1)
a) Tricine would be the best choice as buffers work best at a pH closest to their pKa. TAPS or HEPES could work in a pinch if Tricine was not available, but remember, pH and pKa are log scales! So, while the pKa values might not seem that far from the intended pH of 8.0, on the log scale that is quite a difference in H^+ concentration.
b)

$$2.5L \times \frac{175mmoles}{L} \times \frac{1mole}{1000mmoles} = 0.4375mmoles$$

Now we want to determine what fraction of those moles are in the conjugate base form:

$$[A^-] + [HA] = 0.4375$$

$$[A^-] = 0.4375 - [HA]$$

$$pH = pK_a + log\frac{[A^-]}{[HA]}$$

$$8.0 = 7.48 + log\frac{[A^-]}{[HA]}$$

$$0.52 = log\frac{[A^-]}{[HA]}$$

$$3.31 = \frac{0.4375 - [HA]}{[HA]}$$

At pH 8.0, [HA] = 0.101 mmoles, which is 23% of the total moles of HEPES in the solution.

Q2) As discussed in Section 2.3, when CO_2 is inhaled, it reacts with water to form the weak acid carbonic acid, acidifying the blood. The reaction is given below for reference.
$CO_2 + H_2O \rightarrow H^+ = HCO_3^-$
Currently, the air you breathe contains about 0.04% CO_2. This number has risen from 0.03% in the 1960s and is projected to increase to 0.08% by 2100 if fossil fuel consumption remains at its current rate. (P_{CO2} =0.0003 atm, Keeling, 1960, P_{CO2} =0.0008 atm MIT System Dynamics Group, 2015).

Using the Ideal Gas Laws, there were 0.44 μM CO_2 in the 1960s, 0.60 μM CO_2 today, and as much as 1.2 μM CO_2 in 2100. The pKa of carbonic acid is 6.35 and the pH of your blood is 7.60.

a) What is the change in blood pH due to the increase in atmospheric CO_2 from the 1960s to today?
b) If you were to do the calculations for what the pH of the blood would rise to in 2100, you'd find it to be pH 7.3. Do you think the body would be able to compensate for this? Use the pKa of histidine to explain your answer.

A2)
a)

$$pH = pK_a + log\frac{[A^-]}{[HA]}$$

$$7.60 = 6.35 + log\frac{[A^-]}{0.6}$$

$$1.25 = log\frac{[A^-]}{0.6}$$

$$17.8 = \frac{[A^-]}{0.6}$$

$[A^-]$ = 10.7 μM

$$pH_1960 = 6.35 + log\frac{10.7}{0.44}$$

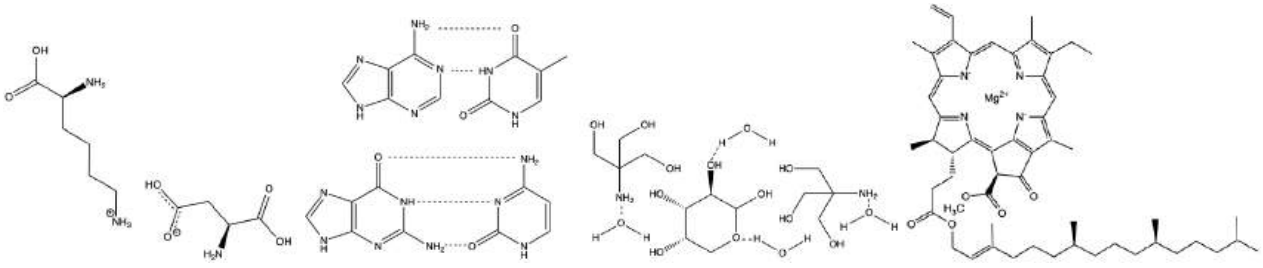
pH_{1960} = 6.35 + 1.39
pH_{1960} = 7.74

Now knowing the pH of the blood from 1960, the change is 7.74 - 7.6 = 0.14 pH units.

b) Yes the body would compensate by increasing the amount of carbonic acid in the blood to buffer the increase in free H^+. If the body did not do this, the protonation state of histidine would change and affect every histidine-containing protein in the blood.

Section 4 Questions

Q1) Categorize the following bonds as ion-ion, ion-dipole, dipole-dipole, H-bond, or hydrophobic, (some maybe be used more than once, or not at all).



A1)

Q2) You discover a new enzyme, "biochemase" and decide to crystalize the protein to determine if you can determine the role from its structure. You hypothesize the protein forms a homodimer when two identical subunits of a protein come together to form one functional structure. You identify two regions that you believe could be the dimer interface. The first region looks to contain several polar amino acids, whose R-groups are less than 5 Å apart, while the second section contains a large cluster of hydrophobic amino acids.

a) You perform some biochemical experiments and discover that when the **polar region** is removed, the enzyme is not able to self-assemble, but maintains two **intact separate structures**. What can you conclude about this region and the type of interactions that are stabilizing the dimer interface?

b) When you remove the **nonpolar** region, you notice the protein aggregates or crashes out of the solution due to an inability to fold correctly. What role can you hypothesize the nonpolar amino acids have in keeping the protein properly folded?

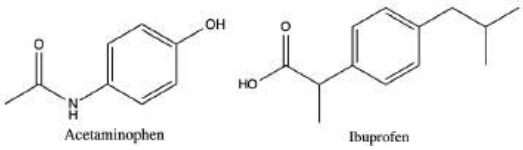
A2)

a) When the polar region is removed, the protein is still able to fold into two stable subunits, but unable to form its homodimer. Therefore we can conclude that the polar region is necessary for the dimer interface, but the protein can still form subunits without it present.

b Now, when the nonpolar region is removed the protein cannot form subunits or dimers. So, we can conclude that the nonpolar region is necessary for protein stability, and without this region present, the protein cannot fold properly. Most likely, the hydrophobic region forces the protein to fold in the correct way by forcing the nonpolar amino acids to the core, and the polar to the surface. This would then allow for a proper dimer interface to form.

Section 5 Questions

Q1) Acetaminophen (Tylenol) and Ibuprofen (Aspirin) are both common pain-relieving/fever-reducing drugs. However, their chemical properties differ, making acetaminophen more suitable for relieving headaches and fever, while ibuprofen can more effectively reduce pain. The pK_{part} for acetaminophen is 0.91, while the pK_{part} is 3.97. (DrugBank) The chemical structures for each are given below:



a) Using equation 2.5.4, determine the ΔG° for each compound, and hypothesize which is more soluble based on your answer and explain. Assume a normal body temperature of T = 37°C

b) With your answer from a) and the information reviews in chapter 1 on functional groups, and types of non-polar interactions in chapter 2, identity the regions of both acetaminophen and ibuprofen that can facilitate interactions.

c) Using your ΔG° calculations and functional group analysis, predict which compound is more likely to stay in the blood and bound to red blood cells, and which can rapidly diffuse across the cell membrane.

Q2) You are tasked with creating a solution of nonpolar and polar solvents to use for thin layer chromatography (a technique we will cover later), to separate lipids. Consider the following enthalpic and entropic values for these nonpolar and polar solvents. Using the information discussed in section 5, which of the following combinations of solvents will yield a homo

---

CHAPTER OVERVIEW

3: Amino Acids, Peptides, and Proteins

3.1: Amino Acids and Peptides
3.3: Proteins - Analyses and Structural Predictions of Protein Structure
3.4: Protein Purification
3.5: Antibodies in Quantitation and In Vivo Detection
3.2: The Structure of Proteins- An Overview
3.6: Chapter 3 Questions

3.1: Amino Acids and Peptides

3.1.1: Introduction

Proteins are one of the most abundant organic molecules in living systems and have the most diverse range of functions of all macromolecules. Proteins may be structural, regulatory, contractile, or protective; they may serve in transport, storage, or membranes; or they may be toxins or enzymes. Each cell in a living system may contain thousands of different proteins, each with a unique function. Their structures, like their functions, vary greatly. They are all, however, polymers of alpha amino acids, arranged in a linear sequence and connected together by covalent bonds.

3.1.2: Alpha Amino Acid Structure

The major building blocks of proteins are called ***alpha (α) amino acids***. As their name implies they contain a carboxylic acid functional group and an amine functional group. The alpha designation is used to indicate that these two functional groups are separated from one another by one carbon group. In addition to the amine and the carboxylic acid, the alpha carbon is also attached to hydrogen and one additional group that can vary in size and length. In the diagram below, this group is designated as an R-group. Within living organisms, there are 20 common amino acids used as protein building blocks. They differ from one another only at the R-group position. The fully protonated structure of an amino acid (at low pH) is shown in Figure 3.1.1.

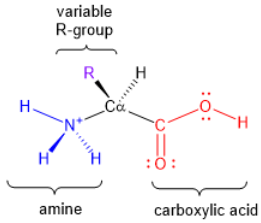


Figure 3.1.1: Generic Structure of an Amino Acid

The twenty common naturally-occurring amino acids each contain an alpha-carbon, an amino, carboxylic acid, and an R group (or side chain). The R group side chains may be either nonpolar, polar and uncharged, or charged, depending on the functional group, the pH, and the pKa of any ionizable group in the side chain.

Two other amino acids occasionally appear in proteins. One is selenocysteine, which is found in Arachea, eubacteria, and animals. Another is pyrrolysine, found in Arachea. Bacteria have been modified to incorporate two new amino acids, O-methyl-tyrosine, and p-aminophenylalanine. The yeast strain *Saccharomyces cerevisiae* has been engineered to incorporate five new unnatural amino acids (using the TAG nonsense codon and new, modified tRNA and tRNA synthetases) with keto groups that allow chemical modifications to the protein. We will concentrate only on the 20 abundant, naturally-occurring amino acids.

Figure 3.1.2 shows the twenty naturally occurring alpha-amino acids as they would appear internally within a protein sequence. The squiggles show that the alpha-amino and carboxyl groups are involved in peptide bonds to adjacent amino acids in the protein sequence. Students often assume that the alpha-amino and carboxyl groups within a protein sequence are free and not part of the peptide bond. This figure should help in resolving that misconception. The three-letter and one-letter abbreviations of each amino acid, as well as their typical pKa values, are also shown. It is important to memorize the three-letter and one-letter codes for the amino acids.

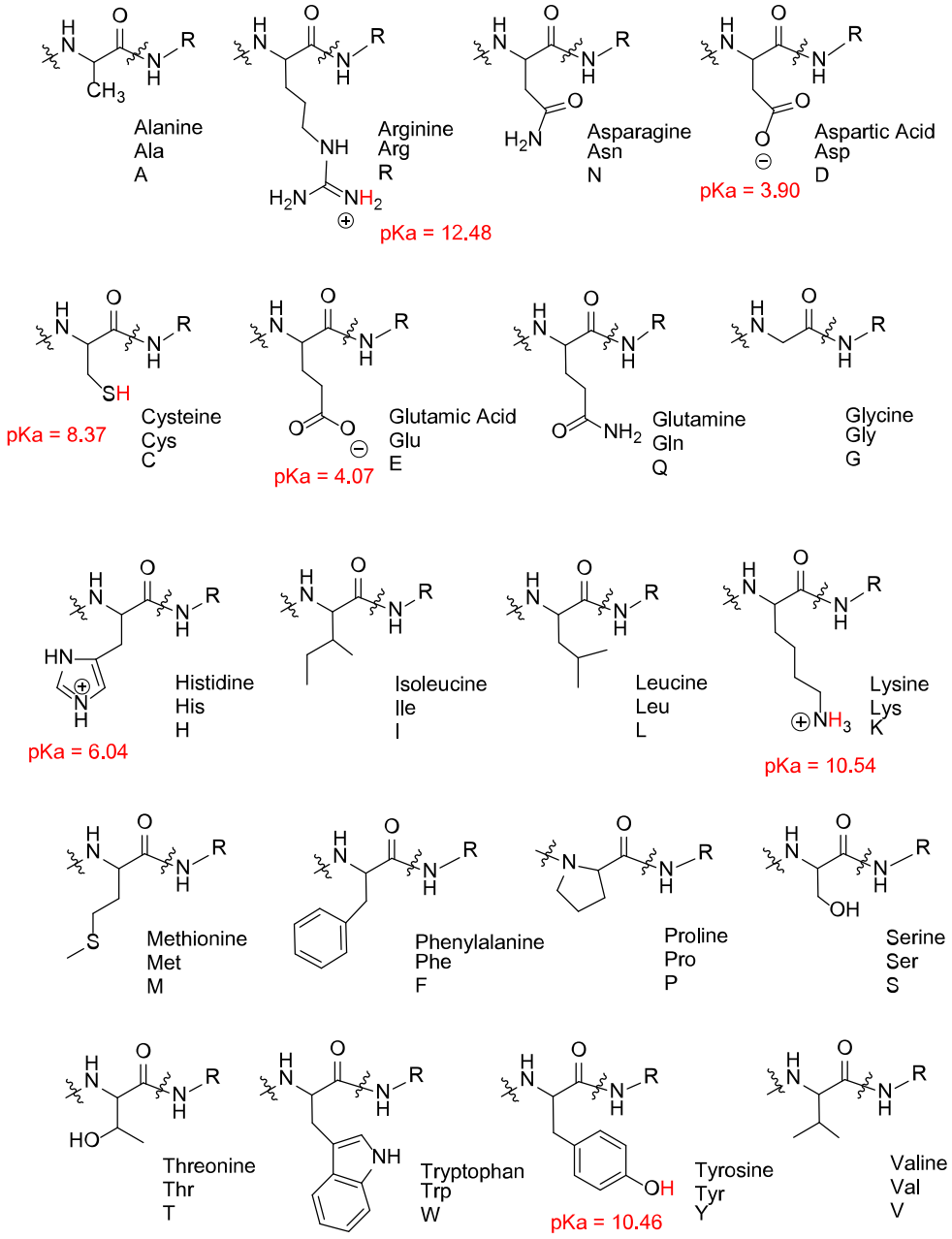


Figure 3.1.2: Side chains of naturally occuring amino acid imbedded in a protein

Amino acids form polymers through a nucleophilic attack by the amino group of an amino acid at the electrophilic carbonyl carbon of the carboxyl group of another amino acid. The carboxyl group of the amino acid must first be activated to provide a better leaving group than OH-. The resulting link between the amino acids is an amide link which biochemists call a peptide bond. In this reaction, water is released. In a reverse reaction, the peptide bond can be cleaved by water (hydrolysis). This is illustrated in Figure 3.1.3.

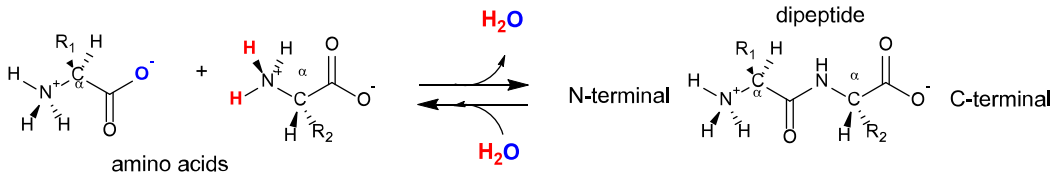


Figure 3.1.3: Amino Acids React to Form a Dipeptide

Proteins are polymers of twenty naturally occurring amino acids. In contrast, nucleic acids are polymers of just 4 different monomeric nucleotides. Both the sequence of a protein and its total length differentiate one protein from another. Just for an octapeptide, there are over 25 billion different possible arrangements of amino acids (820). Compare this to just 65536 different oligonucleotides (4 different monomeric deoxynucleotides) of 8 monomeric units, an 8mer (84). Hence the diversity of possible proteins is enormous.

When two amino acids link together to form an amide link, the resulting structure is called a dipeptide. Likewise, we can have tripeptides, tetrapeptides, and other polypeptides. At some point, when the structure is long enough, it is called a protein. The average molecular weight of proteins in yeast is about 50,000 with about 450 amino acids. The large protein might be titin with a molecular weight of about 3 million (about 30,0000 amino acids). A new class of very small proteins (30 or fewer amino acids and perhaps better named as polypeptides) called smORFs (small open reading frames) have recently been discovered to have significant biological activity. These are encoded directly in the genome and are produced by the same processes that produce regular proteins (DNA transcription and RNA translation). They are not the result of selective cleavage of a larger protein into smaller peptide fragments.

Figure 3.1.4 shows several ways to represent the structure of a polypeptide or protein, each showing differing amounts of information. Note that the atoms in the side chains are denoted alpha, beta, gamma, delta, epsilon ...

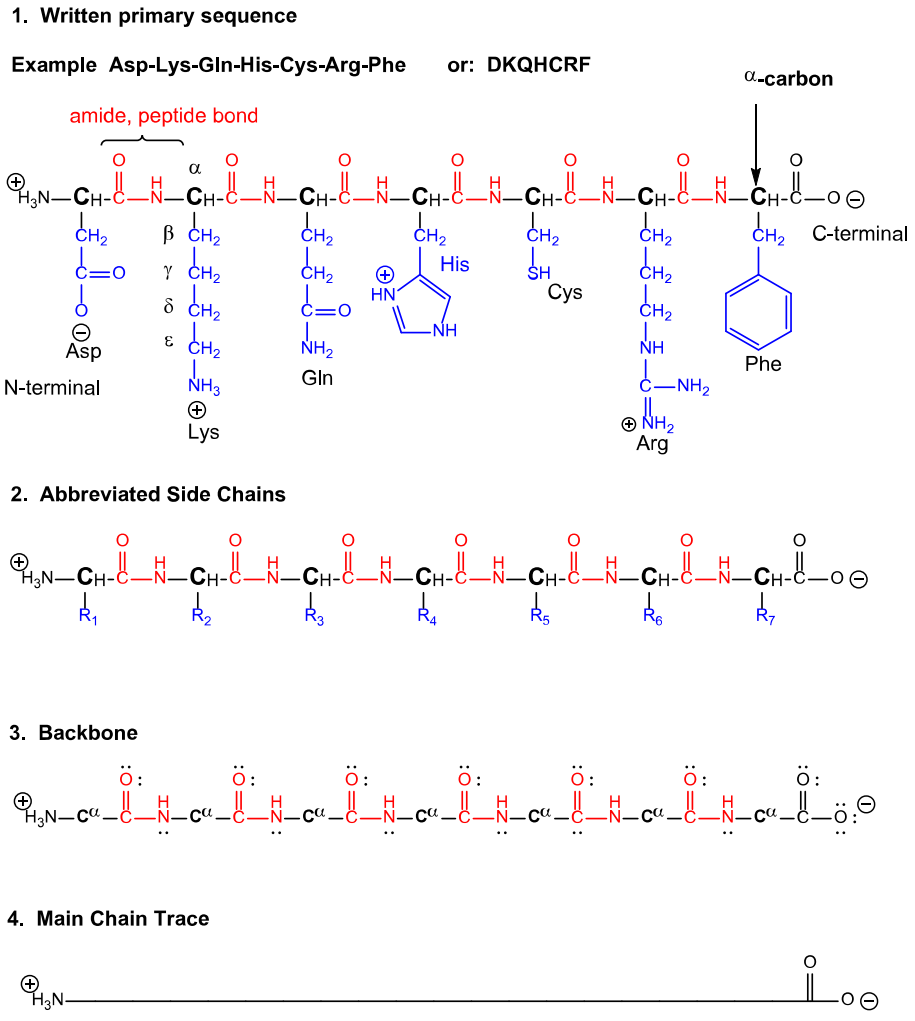
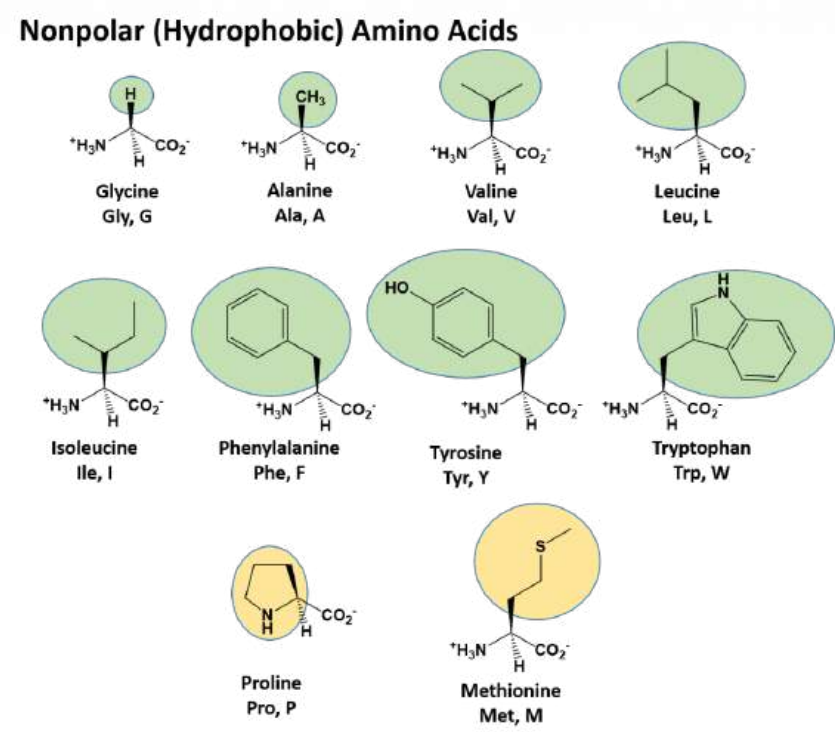


Figure 3.1.4: Different ways to represent the structure of a peptide/protein sequence.

3.1.3: Characteristics of Amino Acids

The different R-groups have different characteristics based on the nature of atoms incorporated into the functional groups. There are R-groups that predominantly contain carbon and hydrogen and are very nonpolar or hydrophobic. Others contain polar uncharged functional groups such as alcohols, amides, and thiols. A few amino acids are basic (containing amine functional

groups) or acidic (containing carboxylic acid functional groups). These amino acids are capable of forming full charges and can have ionic interactions. Each amino acid can be abbreviated using a three-letter and a one-letter code. Figure 3.1.5 shows groupings of the amino acid based on their side chain properties.



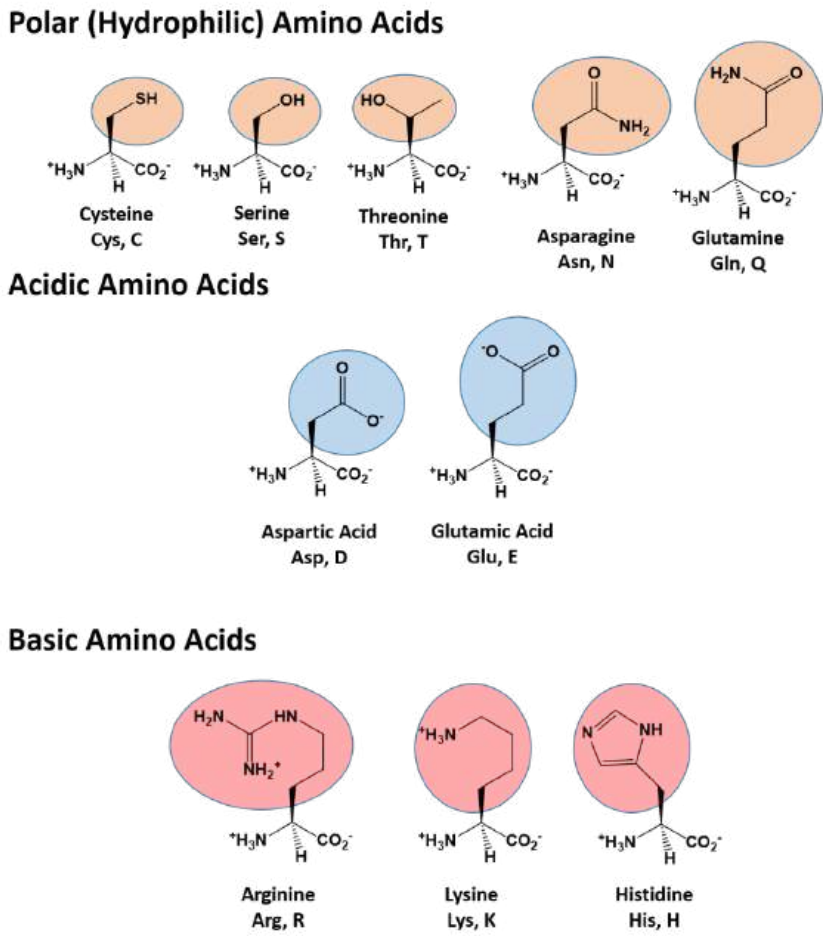


Figure 3.1.5: Structure of the 20 Alpha Amino Acids used in Protein Synthesis. R-groups are indicated by circled/colored portion of each molecule. Colors indicate specific amino acid classes: Hydrophobic - Green and Yellow, Hydrophilic Polar Uncharged - Orange, Hydrophilic Acidic - Blue, Hydrophilic Basic - Rose.

Click Here for a Downloadable Version of the Amino Acid Chart

3.1.4: Nonpolar (Hydrophobic) Amino Acids

The nonpolar amino acids can largely be subdivided into two more specific classes, the ***aliphatic*** amino acids and the ***aromatic*** amino acids. The ***aliphatic amino acids*** ***(glycine, alanine, valine, leucine, isoleucine, and proline)*** typically contain branched hydrocarbon chains with the simplest being glycine to the more complicated structures of leucine and valine. Proline is also classified as an aliphatic amino acid but contains special properties as the hydrocarbon chain has cyclized with the terminal amine creating a unique 5-membered ring structure. As we will see in the next section covering primary structure, proline can significantly alter the 3-dimensional structure of the due to the structural rigidity of the ring structure when it is incorporated into the polypeptide chain and is commonly found in regions of the protein where folds or turns occur.

The ***aromatic amino acids*** ***(phenylalanine, tyrosine, and tryptophan),*** as their name implies, contain an aromatic functional group within their structure making them largely nonpolar and hydrophobic due to the high carbon/hydrogen content. However, it should be noted that hydrophobicity and hydrophilicity represent a sliding scale and each of the different amino acids can have different physical and chemical properties depending on their structure. For example, the hydroxyl group present in tyrosine increase its reactivity and solubility compared to that of phenylalanine.

Methionine, one of the sulfur-containing amino acids is usually classified under the nonpolar, hydrophobic amino acids as the terminal methyl group creates a thioether functional group which generally cannot form a permanent dipole within the molecule and retains low solubility.

3.1.5: Polar (Hydrophilic) Amino Acids

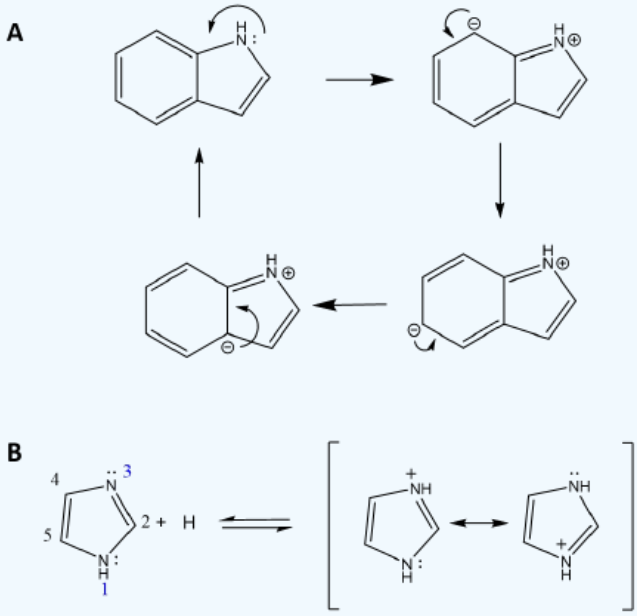
The polar, hydrophilic amino acids can be subdivided into three major classes, the polar uncharged-, the acidic-, and the basic-functional groups. Within the ***polar uncharged class***, the side chains contain heteroatoms (O, S, or N) that are capable of forming permanent dipoles within the R-group. These include **the hydroxyl- and sulfhydryl-containing amino acids, serine, threonine, and cysteine**, and the amide-containing amino acids, **glutamine and asparagine.** Two amino acids, **glutamic acid (glutamate),** and **aspartic acid (aspartate)** constitute **the acidic amino acids** and contain side chains with carboxylic acid functional groups capable of fully ionizing in solution. The **basic amino acids, lysine, arginine, and histidine** contain amine functional groups that can be protonated to carry a full charge.

Many of the amino acids with hydrophilic R-groups can participate within the ***active site*** of enzymes. An ***active site*** is the part of an enzyme that directly binds to a substrate and carries a reaction. Protein-derived enzymes contain ***catalytic groups*** consisting of amino acid R-groups that promote formation and degradation of bonds. The amino acids that play a significant role in the binding specificity of the active site are usually not adjacent to each other in the primary structure but form the active site as a result of folding in creating the tertiary structure, as you will see later in the chapter.

Example 3.1.1

Thought Question: Tryptophan contains an amine functional group, why isn't tryptophan basic?

Answer: Tryptophan contains an indole ring structure that includes the amine functional group. However, due to the proximity of, and electron-withdrawing nature of the aromatic ring structure, the lone pair of electrons on the nitrogen are unavailable to accept a proton. Instead, they are involved in forming *pi*-bonds within several of the different resonance structures possible for the indole ring. Figure 2.3A shows four of the possible resonance structures for indole. Conversely, within the imidazole ring structure found in histidine, there are two nitrogen atoms, one of which is involved in the formation of resonance structures (Nitrogen #1 in Figure 2.3B) and cannot accept a proton, and the other (Nitrogen #3) that has a lone pair of electrons that is available to accept a proton.



Comparison of the Structural Availability of Lone Pair of Electrons on Nitrogen to Accept a Proton in the Indole and Imidazole Ring Structures. (A) Shown are four resonance structures of the indole ring structure demonstrating that the lone pair of electrons on the nitrogen are involved in the formation of pi-bonds. (B) The imidazole ring structure has one nitrogen (1) that is involved in resonance structures (not shown) and is not available to accept a proton, while the second nitrogen (3) has a lone pair of electrons available to accept a proton as shown.

Exercise 3.1.1

Work It Out on Your Own:

Given the example above, describe using a chemical diagram, why the amide nitrogen atoms found in asparagine and glutamine are not basic.

Answer

The lone pair is delocalized into the peptide bond (different resonance structure) so it is unavailable for sharing.

3.1.6: Amino Acid Stereochemistry

The amino acids are all ***chiral***, with the exception of glycine, whose side chain is H. A ***chiral molecule*** is one that is not superimposable with its mirror image. Like left and right hands that have a thumb and fingers in the same order, but are mirror images and not the same, chiral molecules have the same things attached in the same order, but are mirror images and not the same. The mirror image versions of chiral molecules have physical properties that are nearly identical to one another, making it very difficult to tell them apart from one another or to separate them. Because of this nature, they are given a special stereoisomer name called ***enantiomers*** and in fact, the compounds themselves are given the same name! These molecules do differ in the way that they rotate plane-polarized light and the way that they react with and interact with biological molecules. Molecules that rotate light in the right-handed direction are called dextrorotary and are given a small "d" letter designation. Molecules that rotate light in the left-handed direction are called levorotary and are given a small "l" letter designation to distinguish one enantiomer from the other. Biochemists also use the older nomenclature of large "L" and "D" to characterize the 3D stereochemistry of the amino acids. All naturally occurring proteins from all living organisms consist of L amino acids, based on their structural similarities to L-glyceraldehyde.

Again, the d- and l-designations are specific terms used for the way a molecule rotates plane-polarized light. It does **not** denote the absolute stereo configuration of a molecule. An **absolute configuration** refers to the spatial arrangement of the atoms of a chiral molecular entity (or group) and its modern stereochemical description e.g. **R** or **S**, referring to **Rectus**, or **Sinister**, respectively. Absolute configurations for a chiral molecule (in pure form) are most often obtained by X-ray crystallography. Alternative techniques are optical rotatory dispersion, vibrational circular dichroism, the use of chiral shift reagents in proton NMR and Coulomb explosion imaging. When the absolute configuration is known, the assignment of *R* or *S* is based on the **Cahn–Ingold–Prelog priority rules**. The absolute stereochemistry is related to L-glyceraldehyde, as shown in Figure 3.1.6 below.

All naturally occurring amino acids in proteins are L, which corresponds to the S isomer, with the exception of cysteine. As shown in the bottom left of the Figure 6 below, the absolute configuration of the amino acids can be shown with the H pointed to the rear, the COOH groups pointing out to the left, the R group to the right, and the NH_3 group upwards. You can remember this with the mnemonic "CORN".

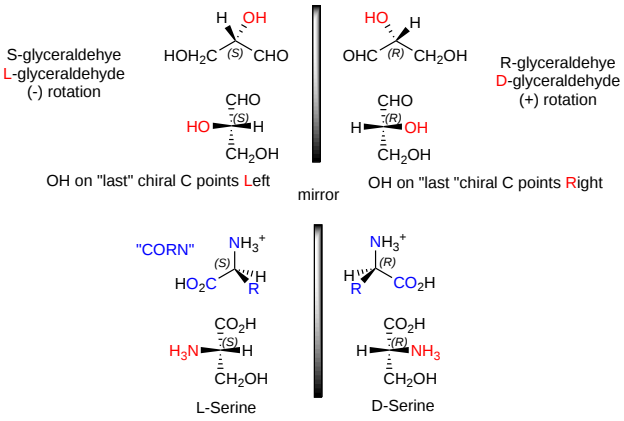


Figure 3.1.6: Stereochemistry of amino acids

Why do Biochemistry still use D and L for sugars and amino acids? This explanation (taken from a website, which may not be available anymore so no reference is available) seems reasonable.

"In addition, however, chemists often need to define a configuration unambiguously in the absence of any reference compound, and for this purpose, the alternative (R,S) system is ideal, as it uses priority rules to specify configurations. These rules sometimes lead to absurd results when they are applied to biochemical molecules. For example, as we have seen, all of the common amino acids are L, because they all have exactly the same structure, including the position of the R group if we just write the R group as R. However, they do not all have the same configuration in the (R,S) system: L-cysteine is also (R)-cysteine, but all the other L-amino

acids are (S), but this just reflects the human decision to give a sulfur atom a higher priority than a carbon atom and does not reflect a real difference in configuration. Worse problems can sometimes arise in substitution reactions: sometimes inversion of configuration can result in no change in the (R) or (S) prefix, and sometimes retention of configuration can result in a change of prefix.

It follows that it is not just conservatism or failure to understand the (R,S) system that causes biochemists to continue with D and L: it is just that the DL system fulfills their needs much better. As mentioned, chemists also use D and L when they are appropriate to their needs. The “explanation” given above of why the (R,S) system is little used in biochemistry is thus almost the exact opposite of reality. This system is actually the only practical way of unambiguously representing the stereochemistry of complicated molecules with several asymmetric centers, but it is inconvenient with regular series of molecules like amino acids and simple sugars."

If you are told to draw the correct stereochemistry of a molecule with 1 chiral C (S isomer for example) and are given the substituents, you could do so easily following the R, S priority rules. However, how would you draw the correct isomer for the L isomer of the amino acid alanine? You couldn't do it without prior knowledge of the absolute configuration of the related molecule, L glyceraldehyde, or unless you remembered the anagram CORN. This disadvantage, however, is more than made up for by the fact that different L amino acids with the same absolute stereochemistry, might be labeled R or S, which makes this nomenclature unappealing to biochemists.

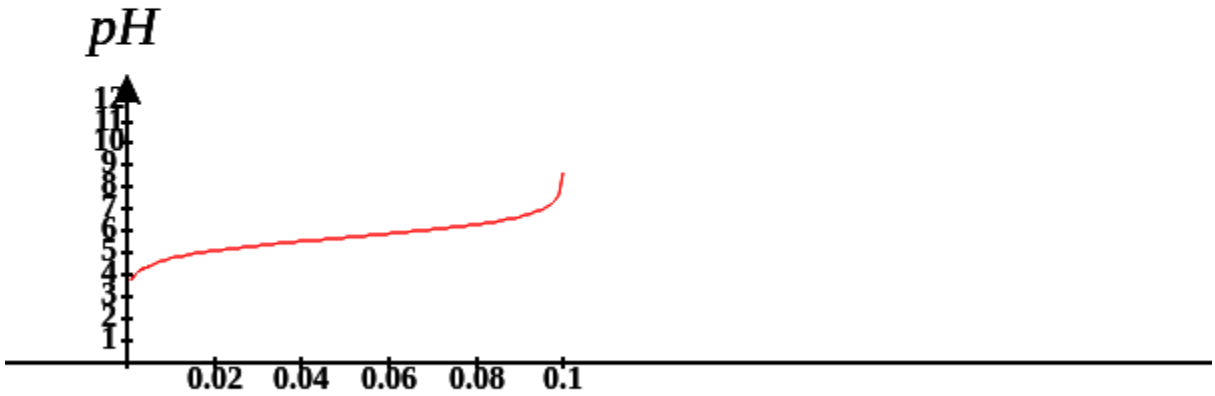
3.1.7: Amino Acid Charges

Monomeric amino acids have an alpha-amino group and a carboxyl group, both of which may be protonated or deprotonated, and a R group, some of which may be protonated or deprotonated. When protonated, the amino group has a +1 charge, and the carboxyl group a zero charge. When deprotonated the amino group has no charge, while the carboxyl group has a -1 charge. The R groups which can be protonated/deprotonated include Lys, Arg, and His, which have a + 1 charge when protonated, and Glu and Asp (carboxylic acids), Tyr and Ser (alcohols) and Cys (thiol), which have 0 charges when protonated. Of course, when the amino acids are linked by peptide bonds (amide link), the alpha N and the carboxyl C are in an amide link, and are not charged.

However, the amino group of the N -terminal amino acid and the carboxyl group of the C-terminal amino acid of a protein may be charged. The Henderson-Hasselbalch equation gives us a way to determine the charge state of any ionizable group knowing the pKa of the group. Write each functional group capable of being deprotonated as an acid, HA, and the deprotonated form as A. The charge of HA and A will be determined by the functional group and the Henderson-Hasselbalch equation from Chapter 2.2.

$$pH = pK_a + \log \frac{[A^-]}{HA}$$

The titration curve for a single ionizable acid with different pKa values is shown below.



At the inflection point of the curve, pH = pKa and the system is most resistant to changes in pH on addition of either acid or base. At this pH, $[HA]=[A^-]$.

The properties of a protein will be determined partly by whether the side chain functional groups, the N terminal, and the C terminal are charged or not. The HH equation tells us that this will depend on the pH and the pKa of the functional group.

- If the pH is 2 units below the pKa, the HH equation becomes -2 = log A/HA, or .01 = A/HA. This means that the functional group will be about 99% protonated (with either 0 or +1 charge, depending of the functional group).
- If the pH is 2 units above the pKa, the HH equation becomes 2 = log A/HA, or 100 = A/HA. Therefore the functional group will be 99% deprotonated.
- If the pH = pKa, the HH equation becomes 0 = log A/HA, or 1 = A/HA. Therefore the functional group will be 50% deprotonated

From these simple examples, we have derived the +2 rule. This rule is used to quickly determine protonation, and hence charge state, and is extremely important to know (and easy to derive). Titration curves for Gly (no ionizable) side chain, Glu (carboxylic acid side chain) and Lys (amine side chain) are shown in Figure 3.1.7. You should be able to associate various sections of these curves with titration of specific ionizable groups in the amino acids.

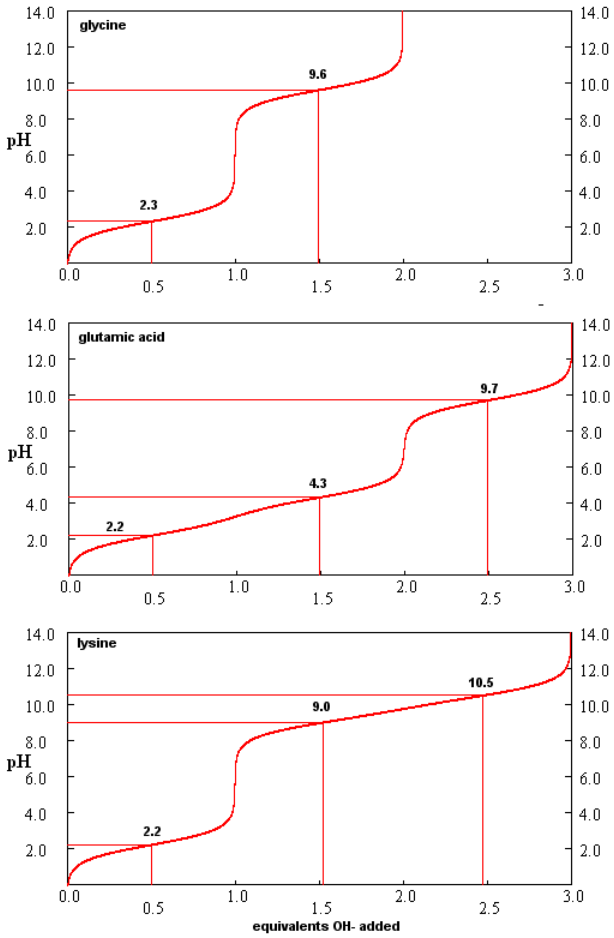


Figure 3.1.7: Titration curves for Gly, Glu, and Lys

New 5/16/23: Download this Excel spreadsheet for Titration Curves for a Triprotic Acid. It has adjustable scroll bars to change pK_a values.

3.1.8: Buffer Review

The Henderson-Hasselbalch equation is also useful in calculating the composition of buffer solutions. Remember that buffer solutions are composed of a weak acid and its conjugate base. Consider the equilibrium for a weak acid, like acetic acid, and its conjugate base, acetate:

$$CH_3CO_2H + H_2O \rightleftharpoons H_3O^+ + CH_3CO_2^-$$

If the buffer solution contains equal concentrations of acetic acid and acetate, the pH of the solution is:

or pH = pKa + log [A]/[HA] = 4.7 + log 1 = 4.7

A look at the titration curve for the carboxyl group of Gly (see above) shows that when the pH = pKa, the slope of the curve (i.e. the change in pH on addition of base or acid) is at a minimum. As a general rule of thumb, buffer solutions can be made for a weak acid/base in the range of +/- 1 pH unit from the pKa of the weak acids. At the pH = pKa, the buffer solution best resists the addition of either acid or base, and hence has its greatest buffering ability. The weak acid can react with the added strong base to form the weak conjugate base, and the conjugate base can react with added strong acid to form the weak acid (as shown below) so pH changes on the addition of strong acid and base are minimized.

- addition of a strong base produces a weak conjugate base: $CH_3CO_2H + OH^- \leftrightarrow CH_3CO_2^- + H_2O$
- addition of strong acid produces weak acid: $H_3O^+ + CH_3CO_2 \rightarrow CH_3CO_2H + H_2O$

There are two simple ways to make a buffered solution. Consider an acetic acid/acetate buffer solution.

- make equal molar solution of acetic acid and sodium acetate, and mix them, monitoring pH with a pH meter, until the desired pH is reached (+/- 1 unit from the pKa).
- take a solution of acetic acid and add NaOH at substoichiometric amounts until the desired pH is reached (+/- 1 unit from the pKa). In this method you are forming the conjugate base,acetate, on the addition of the weak base:

$CH_3CO_2H + OH^- \rightarrow CH_3CO_2^- + H_2O$

- Buffers for pH control: Recipes based on pKas for acids, temperature, and ionic strength

3.1.9: Isoelectric Point

What happens if you have many ionizable groups in a single molecule, as is the case with a polypeptide or protein. Consider a protein. At a pH of 2, all ionizable groups would be protonated, and the overall charge of the protein would be positive. (Remember, when carboxylic acid side chains are protonated, their net charge is 0.) As the pH is increased, the most acidic groups will start to deprotonate and the net charge will become less positive. At high pH, all the ionizable groups will become deprotonated in the strong base, and the overall charge of the protein will be negative. At some pH, then, the net charge will be 0. This pH is called the isoelectric point (pI). The pI can be determined by averaging the pKa values of the two groups which are closest to and straddle the pI. One of the online problems will address this in more detail

- List of pI and MW for proteins derived from 2D gels

Remember that pKa is really a measure of the equilibrium constant for the reaction. And of course, you remember that $\Delta G^o = -RT \ln Keq$. Therefore, pKa is independent of the concentration, and depends only on the intrinsic stability of reactants with respect to the products. This is true only AT A GIVEN SET OF CONDITIONS, SUCH AS T, P, AND SOLVENT CONDITIONS.

Consider, for example, acetic acid, which in aqueous solution has a pKa of about 4.7. It is a weak acid, which dissociates only slightly to form H+ (in water the hydronium ion, H_3O^+, is formed) and acetate (Ac^-). These ions are moderately stable in water but reassociate readily to form the starting product. The pKa of acetic acid in 80% ethanol is 6.87. This can be accounted for by the decrease in stability of the charged products which are less shielded from each other by the less polar ethanol. Ethanol has a lower dielectric constant than water. The pKa increases to 10.32 in 100% ethanol, and to a whopping 130 in air!

Because amino acids are zwitterions, and several also contain the potential for ionization within their R-groups, their charge state *in vivo*, and thus, their reactivity can vary depending on the pH, temperature, and solvation status of the local microenvironment in which they are located. Table 3.1.1 shows the standard pK_a values for the amino acids and can be used to predict the ionization/charge status of amino acids and their resulting peptides/proteins.

pK$_a$ Values for Common Alpha Amino Acids				
Amino Acid Type	Amino Acid	α-COOH	α-NH_3^+	RH or RH^+
Hydrophobic: Aliphatic	Glycine	2.34	9.60	
	Alanine	2.34	9.69	
	Valine	2.32	9.62	
	Leucine	2.36	9.68	
	Isoleucine	2.36	9.68	
	Proline	1.99	10.6	
	Methionine	2.28	9.21	
Hydrophobic: Aromatic	Phenylalanine	1.83	9.13	
	Tyrosine	2.2	9.11	10.07
	Tryptophan	2.38	9.39	
Hydrophilic: Polar Uncharged	Serine	2.21	9.15	
	Threonine	2.63	10.43	
	Cysteine	1.71	10.78	8.33
	Asparagine	2.02	8.8	
	Glutamine	2.17	9.13	
Hydrophilic: Acidic	Aspartic Acid	2.09	9.82	3.86
	Glutamic Acid	2.19	9.67	4.25
Hydrophilic: Basic	Arginine	2.17	9.04	12.48
	Histidine	1.82	9.17	6.00
	Lysine	2.18	8.95	10.53

Table 3.1.1: Summary of pKas of amino acids

However, it should be noted that the solvation status in the microenvironment of an amino acid can alter the relative pK_a values of these functional groups and provide unique reactive properties within the active sites of enzymes. A more in-depth discussion of the effects of desolvation will be given in Chapter 6 discussing enzyme reaction mechanisms.

Printable Version of pKa Values

- A great interactive web site: Amino Acid Acid/Base Titration Curves
- pI calculator for any protein sequence
- Amino Acid Repository: Properties of Amino Acids

3.1.10: Introduction to Amino Acid Reactivity

You should be able to identify which side chains contain H bond donors and acceptors. Likewise, some are acids and bases. You should be familiar with the approximate pKa's of the side chains, and the N and C terminal groups. Three of the amino acid side chains (Trp, Tyr, and Phe) contribute significantly to the UV absorption of a protein at 280 nm. This section will deal predominantly with the chemical reactivity of the side chains, which is important in understanding the properties of the proteins. Many of the side chains are nucleophiles. Nucleophilicity is a measure of how rapidly molecules with lone pairs of electrons can react in nucleophilic substitution reactions. It correlates with basicity, which measures the extent to which a molecule with lone pairs can react with an acid (Bronsted or Lewis). The properties of the atom which holds the lone pair are important in determining both nucleophilicity and basicity. In both cases, the atom must be willing to share its unbonded electron pair. If the atoms holding the nonbonded pair is more electronegative, it will be less likely to share electrons, and that molecule will be a poorer nucleophile (nu:) and weaker base. Using these ideas, it should be clear that RNH_2 is a better nucleophile than ROH, OH^- is better than H_2O and RSH is better than H_2O. In the latter case, S is bigger and its electron cloud is more polarizable - hence it is more reactive. The important side chain nucleophiles (in order from most to least nucleophilic) are Cys (RSH, pKa 8.5-9.5), His (pKa 6-7), Lys (pKa 10.5) and Ser (ROH, pKa 13). The side chain of serine is generally no more reactive than ethanol. It is a potent nucleophile in a certain class of proteins (proteases, for example) when it is deprotonated. The amino group of lysine is a potent nucleophile only when deprotonated.

An understanding of the chemical reactivity of the various R group side chains of the amino acids in a protein is important since chemical reagents that react specifically with a given amino acid side chain can be used to:

- identify the presence of the amino acids in unknown proteins or
- determine if a given amino acid is critical for the structure or function of the protein. For example, if a reagent that covalently interacts with only Lys is found to inhibit the function of the protein, a lysine might be considered to be important in the catalytic activity of the protein.

Figure 3.1.8 shows a summary of nucleophilic addition and substitution at carbonyl carbons.

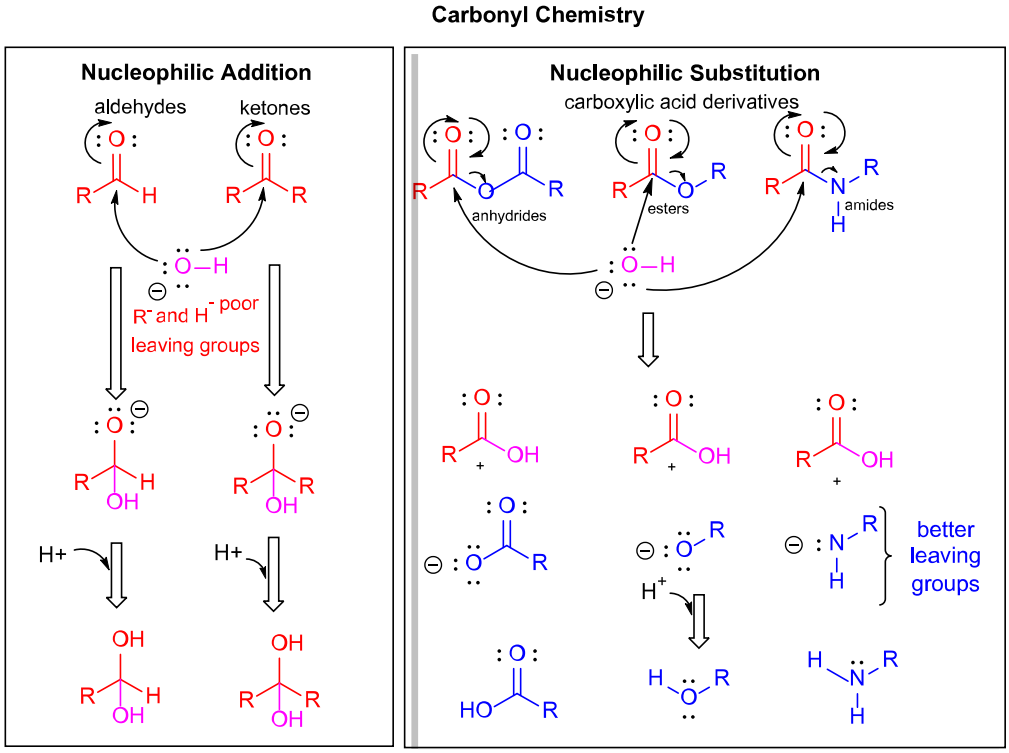


Figure 3.1.8: A review summary of the chemistry of aldehydes, ketones and carboxylic acid derivatives

The rest of the section will summarize the chemistry of the side chains of reactive amino acids. Historically the function of a given amino acid in a protein has been studied by reacting them with side chain-specific chemical modifying agents. In addition, some side chains are covalently modified after they are synthesized in vivo (post-translational modification - see below).

3.1.11: Reactions of Lysine

Figure 3.1.9 the reaction of lysine with anhydrides and ethylacetimidate.

- reacts with anhydrides in a nucleophilic substitution reaction (acylation).
- reacts reversibly with methylmaleic anhydride (also called citraconic anhydride) in a nucleophilic substitution reaction.
- reacts with high specificity and yield toward ethylacetimidate in a nucleophilic substitution reaction (ethylacetimidate is like ethylacetate only with a imido group replacing the carbonyl oxygen). Ethanol leaves as the amidino group forms. (has two N - i.e. din - attached to the C)

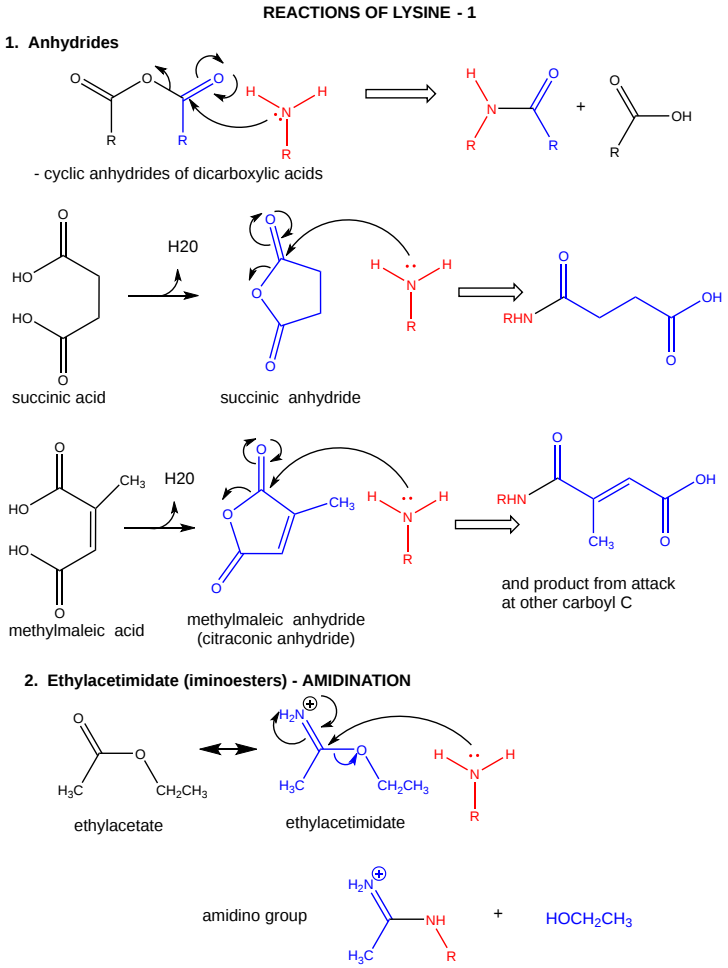


Figure 3.1.9: Reaction of lysine with anhydrides and ethylacetimidate.

Figure 3.1.10 a second set of common reactions of lysine, including those used to attach a chromophore or a fluorescent label to the side chain.

- reacts with O-methylisourea in a nucleophilic substitution reaction. with the expulsion of methanol to form a guanidino group (has 3 N attached to C, nidi)
- reacts with fluorodintirobenzene (FDNB or Sanger's reagent) or trinitrobenzenesulfonate (TNBS, as we saw with the reaction with phosphatidylethanolamine) in a nucleophilic aromatic substitution reaction to form 2,4-DNP-lysine or TNB-lysine.
- reacts with Dimethylaminonapthelenesulfonylchloride (Dansyl Chloride) in a nucleophilic substitution reaction.

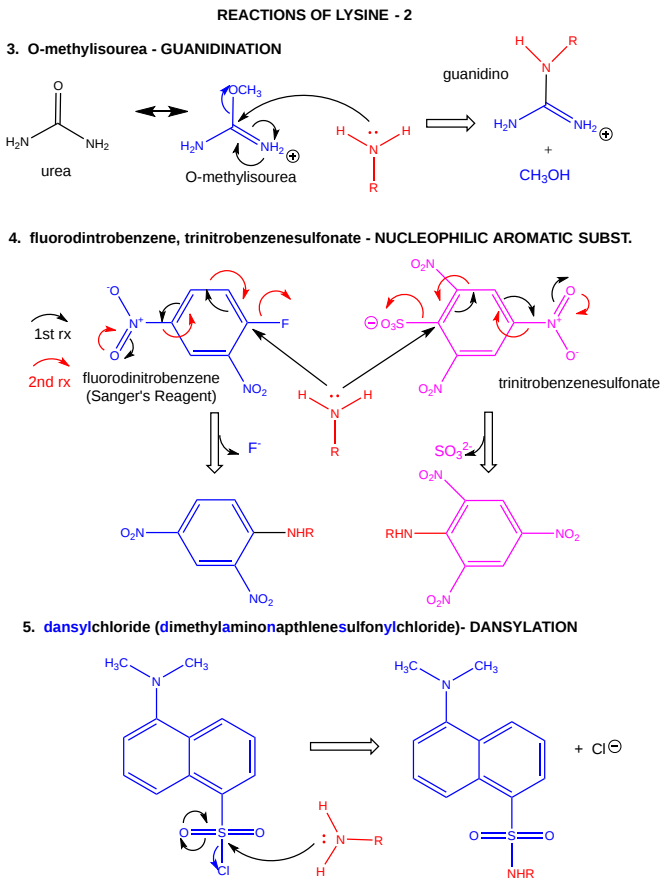


Figure 3.1.10: Reaction of lysine with O-methylisourea, chromophores and fluorophores

Figure 3.1.11 shows a final common reaction we will encounter: the formation of an imine or Schiff base on the reaction of lysine with an aldehyde or ketone.

- reacts with high specificity toward aldehydes to form imines (Schiff bases), which can be reduced with sodium borohydride or cyanoborohydride to form a secondary amine.

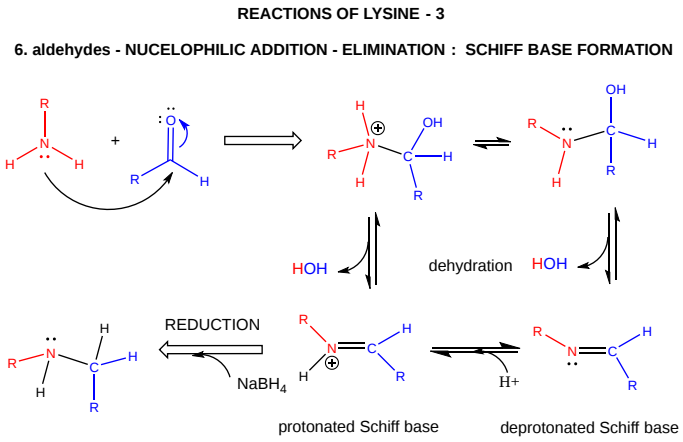


Figure 3.1.11: Reaction of lysine with an aldehyde or ketone to form a Schiff base

3.1.12: Reactions of Cysteine

Cysteine is a potent nucleophile, which is often linked to another Cys to form a covalent disulfide bond.

Figure 3.1.12 shows common reagents used in the lab to label free Cys side chains. These reagents are used to alter Cys side chains to determine if they have functional significance in a protein (such as an active nucleophile in an enzyme-catalyzed reaction.

- reacts with iodoacetic acid in an S_N2 reaction, adding a carboxymethyl group to the S.

- reacts with iodoacetamide in an S_N2 reaction, adding a carboxyamidomethyl group to S.
- reacts with N-ethylmaleimide in an addition reaction to the double bond

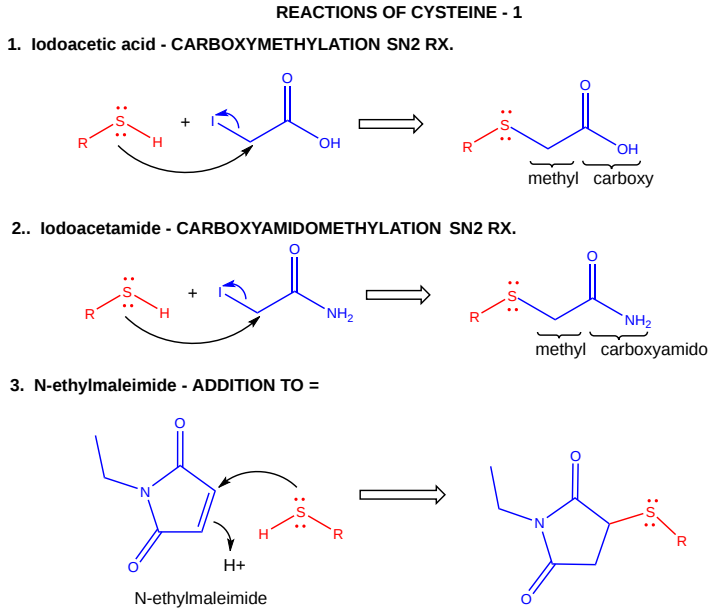


Figure 3.1.12: Common labeling reactions of cysteine

Sulfur is directly below oxygen in the periodic table and, in analogy to water, sulfur-containing amino acids are found in different redox states, as illustrated in Figure 3.1.13

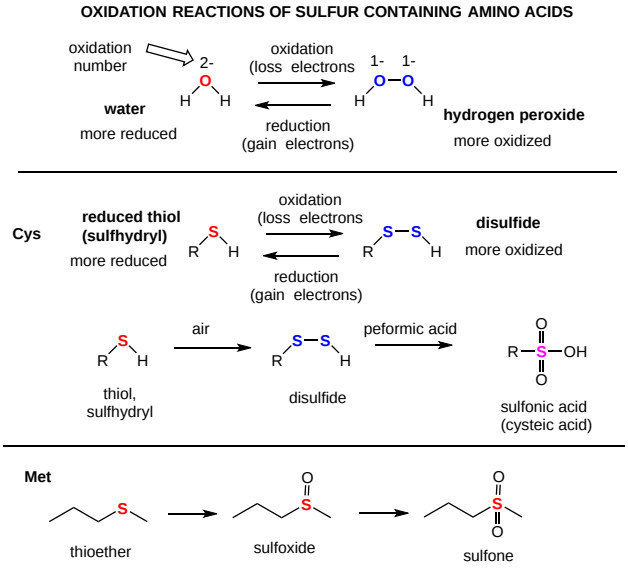


Figure 3.1.13: Oxidation states of sulfur

3.1.13: Cystine Chemistry

Two cysteine side chains can covalently interact in a protein to produce a disulfide (RS-SR) named cystine. Just as HOOH (hydrogen peroxide) is more oxidized than HOH (O in H_2O_2 has an oxidation number of 1- while the O in H_2O has an oxidation number of -2) , RSSR is the oxidized form (S oxidation number -1) and RSH is the reduced form (S oxidation number -2) of thiols. Their oxidation numbers are analogous since O and S are both in Group 6 of the periodic table and both are more electronegative than C.

Cystine can react with a free sulfhydryl (RSH) in a thermodynamically non-challenging disulfide **exchange** reaction, which when conducted with excess free sulfhydryls results in the reduction of cystine in the protein, as shown in Figure 3.1.14

1. Disulfides - INTERCHANGE: REDUCTION OF DISULFIDE

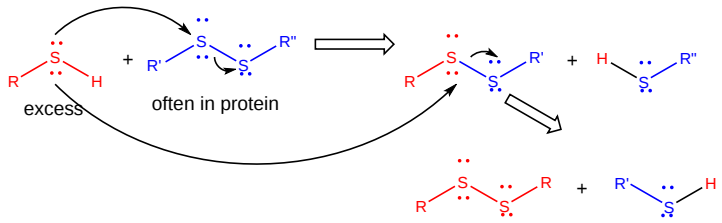


Figure 3.1.14: Disulfide interchange and reduction of protein disulfides

This reaction is often used in the lab to quantitate the amount of free cysteine side chains in a protein using Ellman's reagent, as shown in Figure 3.1.15.

REACTIONS OF CYSTEINE - 2

4. 5,5'-dithiobis(2-nitrobenzoic acid) -DTNB: Ellman's Reagent for quantitation of free sulfhydryls

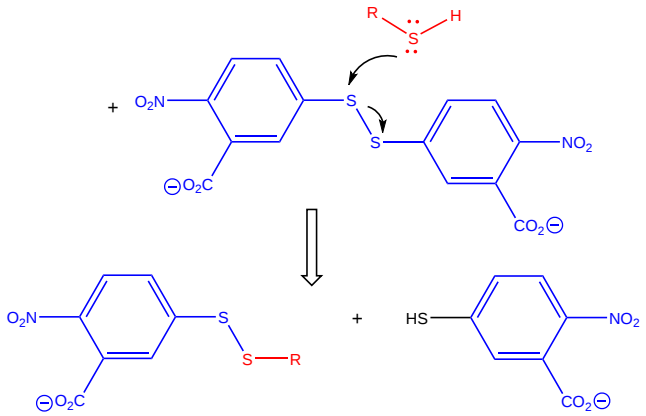
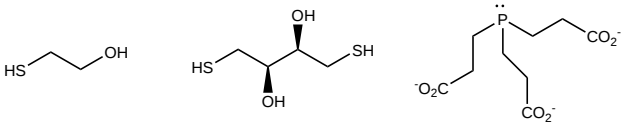


Figure 3.1.15: Reaction of free cysteine with Ellman's reagent.

The 2-nitro-5-thiobennzoic acid anion leaving group absorbs at 412 nm which makes quantitation easy. Only surface and not buried free cysteines will be labeled unless the protein is unfolded to expose all the cysteines.

When a protein folds, two Cys side chains might approach each other, and form an intra-chain disulfide bond. Likewise, two Cys side chains on separate proteins might approach each other and form an inter-chain disulfide. For analysis of the protein structure, disulfides are typically cleaved, and the chains are separated before analysis. The disulfides can be cleaved by reducing agents such as beta-mercaptoethanol, dithiothreitol, tris (2-carboxyethyl) phosphine (TCEP), or by oxidizing agents like performic acid which further oxidizes the disulfide to separate cysteic acids. Three common reagents used in disulfide cleavage reactions in the lab are shown in Figure 3.1.16.



b-mercaptoethanol dithiothreitol tris (2-carboxyethyl) phosphine

Figure 3.1.16: Three common disulfide cleaving (reducing) agents used in the lab

The reaction for beta-mercaptoethanol (BME) and performic acid are shown in Figure 3.1.17 below.

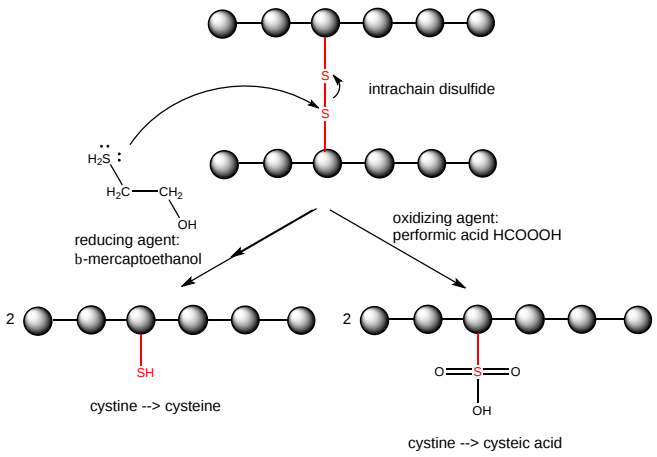


Figure 3.1.17: Cleave of intrachain cystine disulfide bonds in proteins by betamercaptoethanol and performic acid

Figure 3.1.18 shows the reaction for dithiothreitol (DTT). Note that it forms a stable cyclohexane-like ring, which makes this reaction very favored thermodynamically. It does not require as much of an excess of DTT as in the reaction with BME.

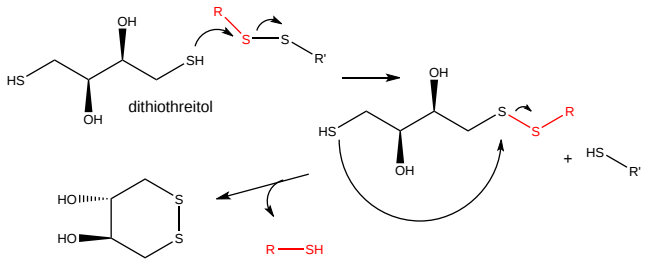


Figure 3.1.18: Cleave of disulfides with dithiothreitol

The reaction with tris (2-carboxyethyl) phosphine (TCEP) is not a disulfide interchange reaction as is shown in Figure 3.1.19.

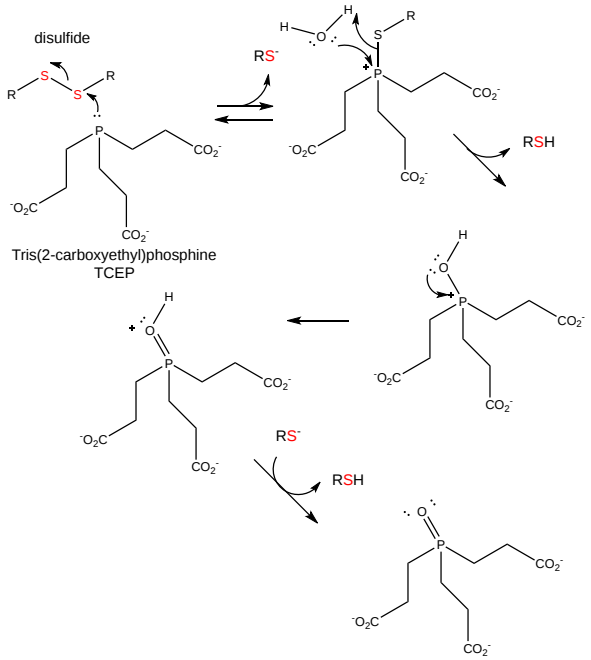


Figure 3.1.19: Reaction of TCEP with disulfides

Cells maintain a reducing environment by the presence of many "reducing" agents, such as the tripeptide gamma-Glu-Cys-Gly (glutathione). Hence intracellular proteins usually do **not** contain disulfides, which are abundant in extracellular proteins (such as those found in blood), or in certain organelles such as the endoplasmic reticulum and mitochondrial intermembrane space where disulfides can be introduced.

Sulfur redox chemistry is very important biologically. As described above, the sulfur in cysteine is redox-active and hence can exist in a wide variety of states, depending on the local redox environment and the presence of oxidizing and reducing agents. A potent oxidizing agent that can be made in cells is hydrogen peroxide, which can lead to more drastic and irreversible chemical modifications to the Cys side chains. If a reactive Cys is important to protein function, then the function of the protein can be modulated (sometimes reversibly, sometimes irreversibly) with various oxidizing agents, as shown in Figure 3.1.20.

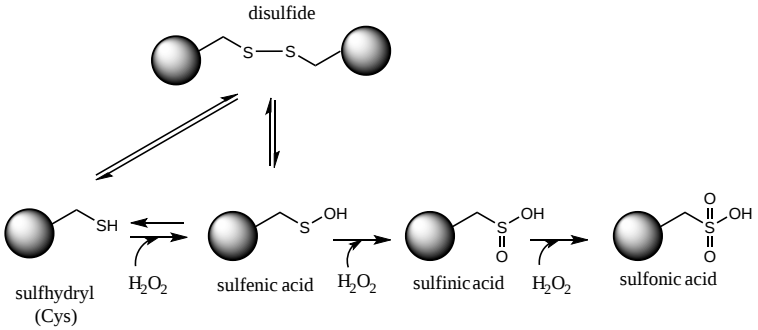


Figure 3.1.20: Reaction of Cysteine with H_2O_2

3.1.14: Reactions of Histidine

Histidine is one of the strongest bases at physiological pH's. His can exist as two tautomers, as shown in Figure 3.1.21. NMR studies show that in model peptides, the proton predominantly is on the ε_2, N_3, or tele N in the imidazole ring, as it has a pK_a 0.6 units higher than δ_1, N_1,or pro N.

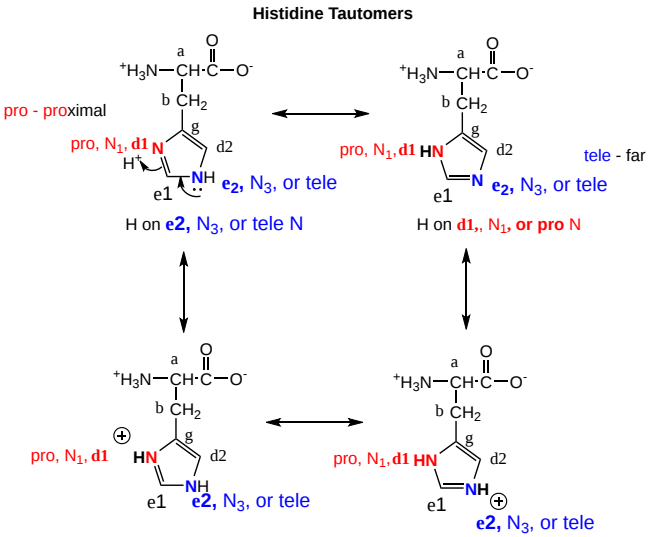


Figure 3.1.21: Histidine tautomers

The nitrogen atom in a secondary amine might be expected to be a stronger nucleophile than a primary amine through electron release to that N in a secondary amine. Opposing this effect is the steric hindrance by the two attached Cs of the N on attach on an electrophile. However, in His, this steric effect is minimized since the 2Cs are restrained by the ring. With a pKa of about 6.5, this amino acid is one of the strongest available bases at physiological pH (7.0). Hence, it can often cross-react with many of the reagents used to modify Lys side chains. His reacts with reasonably high selectivity with diethylpyrocarbonate.

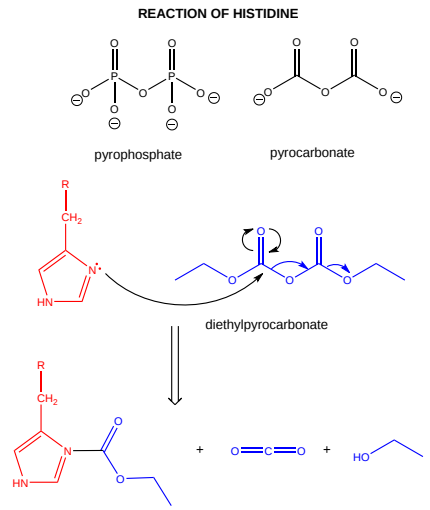


Figure 3.1.22: Reaction of histidine with diethylpyrocarbonate

3.1.15: In vivo Post-Translational Modification of Amino Acids

Amino acids in naturally occurring proteins are also subjected to chemical modifications within cells. These modifications alter the properties of the amino acid that is modified, which can alter the structure and function of the protein. Most chemical modifications made to proteins within cells occur after the protein is synthesized in a process called translation. The resulting chemical changes are termed post-translational modifications. Several are shown in Figure 3.1.23 Note that simple acid/base reactions are included, but these are not considered examples of post-translational modifications.

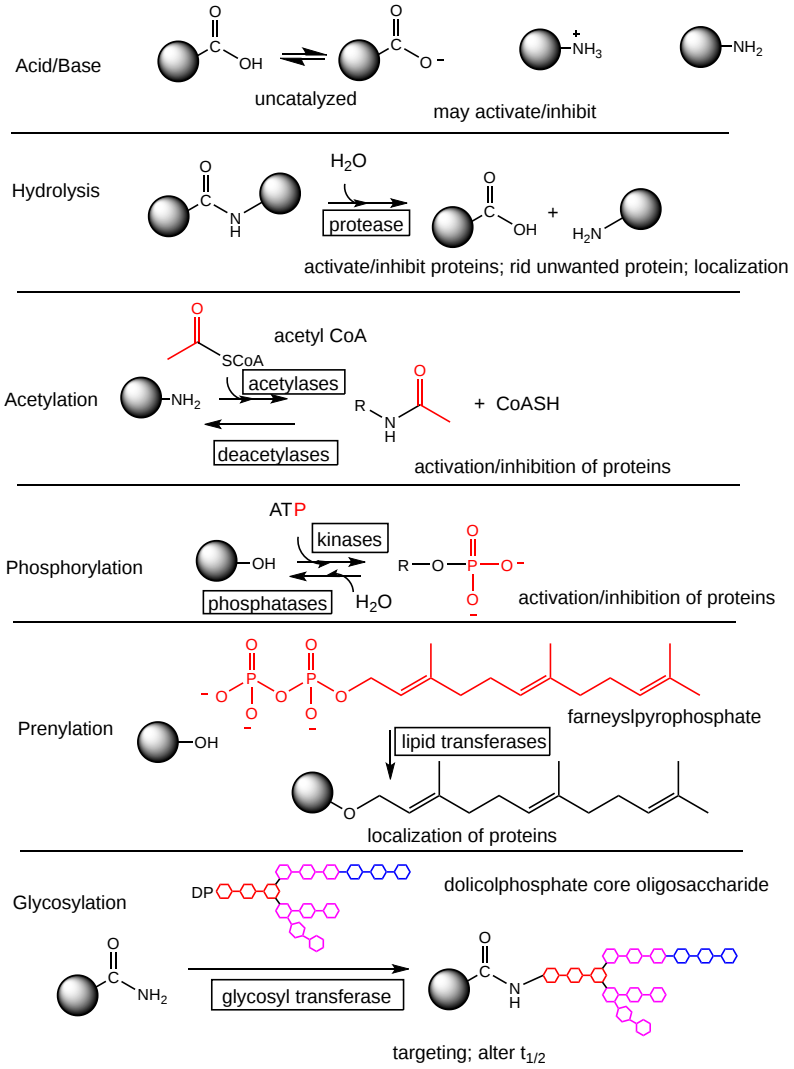


Figure 3.1.23: Common post-translational modifications of protein

There are 100s of PTMs and many are part of an elaborate system within a cell to respond to both external (hormones, neurotransmitters, nutrients, metabolites) and internal chemical signals. The PTMs (like phosphorylation, acetylation, etc.) and their removal by enzymes are part of an elaborate system of cell signaling that we will explore in great detail in Chapter 28. However, not all PTMs are benign. Examples include glycation, oxidation, citrullination, and carbonylation of protein side chains. These are often increased during periods of inflammatory stress (both acute and chronic). These latter modified proteins are degraded within the cell to short peptides that retain the chemical modification. Unfortunately, these can be recognized by the immune system as foreign, which leads to an immune response against self and autoimmune disease. One such potentially deleterious PTM is the carboxyethylation of cysteine, catalyzed by the enzyme cystathionine β-synthase as shown in Figure 3.1.24 below.

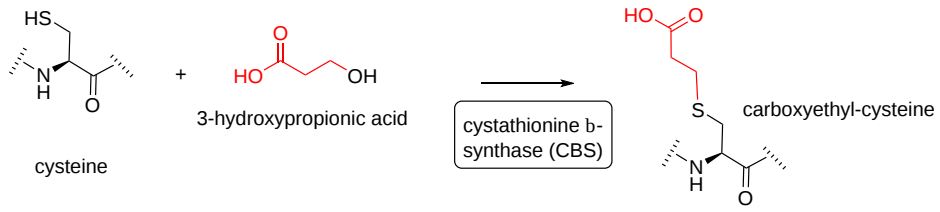


Figure 3.1.24 Carboxyethylation of cysteine

The product is very similar to the carboxymethylation of cysteine shown in Figure 12 above. The modifying reagent, 3-hydroxypropionic acids, is a metabolite released by microbes found in the gut. This modification has been shown to produce an

autoimmune response in the disease ankylosing spondylitis.

3.3: Proteins - Analyses and Structural Predictions of Protein Structure

3.3.1: Introduction

In the last chapter section, we discussed how to purify a protein (mostly through differential salt precipitation and column chromatography) and follow the purity of a protein (mostly through various types of electrophoresis) during the process. Now we want to continue with the analysis of a "pure" protein, so we can understand it structure and the function conferred by that structure. We need many pieces of information to study protein structure/function relationships. Some are "low resolution" characteristics such as knowing the concentration of a protein. At the highest "resolution" end, we would like to know the 3D structure of a protein with a specific interacting ligand bound to a specific site on the protein. Our goal is then to understand proteins at varying levels of complexity or "resolution", some of which are illustrated in Figure 3.3.1.

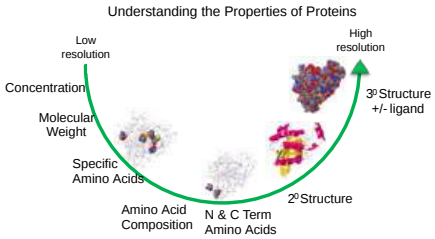


Figure 3.3.1: Understanding proteins from low to high resolution

A variety of chemical and spectra analysis techniques are used to achieve the specified level of structural elucidation. Spectral techniques can give us information on concentration (UV absorbance, fluorescence) and secondary structure (CD spectroscopy). Chemical analyses and mass spectroscopy give information on the amino acid composition and/or sequence. More sophistical techniques (x-ray crystallography, cryoelectron microscopy, NMR spectroscopy) can give us 3D structural information. Each of the analyses shown in the figure above will be summarized below. For some, for which readers might have less experience, more details will be offered.

Sequencing the cDNA can, of course, give you some of the information (amino acid composition, N- and C-terminal amino acids, and the primary structure (omitted from the figure above). Even DNA sequencing won't give information on post-translational modification and other covalent processing (limited proteolysis, disulfide bond formation) that some of the methods below would.

We will explore some commonly used methods for protein analysis in this section. In previous sections, we learned about the charge and chemical reactivity properties of isolated amino acids and amino acids in proteins. The analysis of a whole protein is complicated since each different amino acid might be represented many times in the sequence. Each protein has an N-terminal and C-terminal amino acid and secondary structure. Some proteins exist biologically as multisubunit proteins, which adds to the complexity of the analyses since now the proteins would have multiple N- and C-terminal ends. In addition, isolated proteins might have chemical modifications (post-translational) which add to the functionalities of the proteins but also add to the complexities of the analyses.

Spectral techniques are widely used to give information on concentration (UV absorbance, fluorescence) and secondary structure (CD spectroscopy). Chemical analyses and mass spectroscopy give information on the amino acid composition and/or sequence. More sophistical techniques (x-ray crystallography, cryoelectron microscopy, NMR spectroscopy) can give us 3D structural information. More complete descriptions for two techniques, fluorescence spectroscopy, and mass spectrometry, are presented as their uses in the analyses of biomacromolecules are underrepresented in curriculums as their use in actual laboratories becomes so prevalent.

3.3.2: Low-Resolution Analyzes

3.3.2.1: Protein Concentration

There are multiple ways to determine protein concentrations in samples. Other components of a protein solution may interfere with the assays so the choice of methods has to be carefully determined.

1. Direct mass determination. A known, accurate amount of a dried protein is added to a solution of specific ionic strength and composition. The absorbance at a specific wavelength (usually 280 nm) is determined, which is used to determine an extinction coefficient at that wavelength ($\varepsilon_{1\%}$ = absorbance of a 1% protein solution = 1g protein/100 ml solution). If the molecular weight of the protein is known through sequence analysis, then a molar absorptivity can be determined. The concentration of the same protein in an unknown, pure solution of the protein can then be determined. There are several problems with this technique. It requires relatively large amounts of protein to make accurate measurements. An even more difficult problem is that proteins bind not only water but counter-ions. Even a freeze-dried protein (a protein which has been frozen at -60^{0}C and placed in a vacuum, which causes sublimation of water and volatile salts in the solution) has probably 10% by weight of bound water (water of hydration).

2. ***Quantitative amino acid analysis***. The protein is hydrolyzed completely to amino acids with 6N HCl. The amino acids are then separated by high-performance liquid chromatography. As amino acids elute from the column, they are reacted with a fluorescent reagent such as ninhydrin, fluorescamine, on orthopthaldehdye (OPA) to produce a fluorescent-amino acid conjugate. The fluorescence intensity of the conjugates is proportional to the concentration of the amino acids in the protein. Before hydrolysis, a known quantity of an amino acid not present in proteins (norleucine, beta-alanine) is added and its recovery is determined at the end of the hydrolysis and fluorescence conjugation to normalize the recovery of the other amino acids. Several problems are encountered using this technique. Incomplete peptide bond hydrolysis, and partial or complete destruction of serine, threonine, tryptophan, and tyrosine occur during the acid hydrolysis. The OH-containing amino acids can be determined by quantitating these amino acids after several different time intervals of hydrolysis, and extrapolating the concentrations back to zero time. Incomplete reactions with the detecting reagent can also occur.

The most widely used and perhaps less analytically accurate than the two above are indirect, comparative protein assays based on the chemical properties of amide bonds or the spectrophotometric properties of the side chains Trp, Tyr, Phe. Unknown concentrations can be determined from a standard curve derived from performing the same reactions or spectrophotometric measurements on a series of solutions of known protein concentration. Below is a discussion of each of these techniques.

3. ***Modified Lowry protein determination***. The Lowry Method is actually a modification of the biuret method, whose basis is described below. It is much more sensitive, however. Biuret, as its name implies, derives from the combination of two molecules of urea (***bi-ur-e***t) as shown in the figure (panel A) below. If copper sulfate is added to Biuret in a concentrated hydroxide solution, a violet color results. An illustration of the copper (II) complex with biuret is shown in panel B of the figure. This biuret reaction also arises in the presence of any compound with three or more peptide bonds. Compare in Figure 3.3.2 the structure of Biuret and a polypeptide (Panel A).

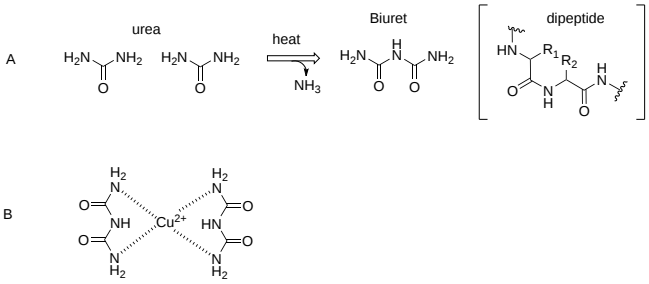


Figure 3.3.2: The biuret reaction for protein quantitation

4. ***Dye binding assay (Bradford method)***. This method is based on the binding of the dye Coomassie Brilliant Blue G-250 to protein, with a resultant change in the absorbance properties of the dye. The structure of the dye is shown in Figure 3.3.3.

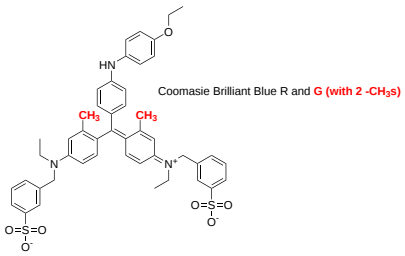


Figure 3.3.3: Structures of Coomassie Brilliant Blue R-250 and G-250 (with 2 -CH_3 groups added)

The magnitude of the difference spectra at 595 nm is directly proportional to the protein concentration. The dye in the unbound, free state, has an absorbance maximum at 465 nm. The method was initially developed by Bradford (Analytic Biochemistry, *72*, 1976), and is available commercially. The dye appears to bind to proteins through both hydrophobic interactions and electrostatic ones through a sulfonic acid group on the dye. The predominant advantage of this method is that its cheap, simple, rapid, 3x more sensitive than the modified Lowry method, and subjected to fewer interferences by other compounds. The color fully develops in about 5 minutes but decreases within 10-15 minutes when the proteins start to precipitate. Precipitation occurs more extensively at higher protein concentrations. Hence, the high-concentration standards will be affected more than the low-concentration standards.

5. ***Bicinchoninic Acid method (BCA)***. This method is based on the reduction of Cu^{++} to Cu^{+} by the peptide bond, and the chelation of the Cu+ by BCA, which is monitored by absorbance at 562 nm. This method, which has been commercialized, is subject to fewer interferences from other compounds than either the modified Lowry or the dye binding assay. Two solutions are required, a BCA solution and a copper sulfate solution. The two are mixed together to form an apple-green working solution. When protein is added, the resulting Cu^{+} chelates with 2 molecules of BCA as shown in Figure 3.3.4.

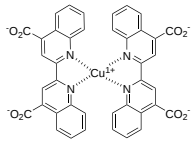


Figure 3.3.4: BCA analysis for protein quantitation

This results in a purple color, which can be monitored spectrophotometrically at 562 nm. An absorbance at 562 nm of 0.012 per microgram of protein added to the working reagent gives this technique high sensitivity.

6. ***Absorbance A_{280} or ratios at different wavelengths***. This method is based on the fact that the three aromatic acids (Tyr, Phe, Trp) have significant absorbances in the UV. The absorption spectra of the three amino acids as a function of wavelength are shown in Figure 3.3.5. Note the log scale on the y-axis.

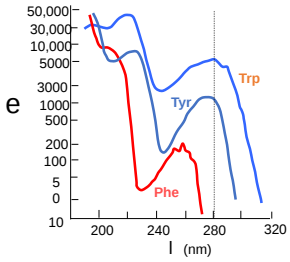


Figure 3.3.5: Log absorbance spectra of tyrosine, phenylalanine, and tryptophan

When separating proteins on chromatography columns, the absorbance at 280 nm of the elute is often measured in isolated fractions or continuously as a measure of the presence and concentrations of the eluting proteins.

The absorbance at 280 nm is often used to estimate the total concentration of proteins (an average protein at a concentration of 1 mg/ml has an A_{280} of about 1). However, only two amino acids absorb significantly at this wavelength, and since proteins have a variable number of these amino acids, this measurement can only be an estimate of the protein concentration of an unknown protein.

The Beer-Lambert law shows that the absorbance of a chromophore in solution is given by

$$A = \epsilon lc \tag{3.3.1}$$

where A is the absorbance at a given wavelength, ε is the molar absorptivity, l is the path length of the cuvette and c is the concentration (mol/L). Pace et al have shown that based on over a hundred measurements on 61 proteins in aqueous solution, the ε(280), the molar absorptivity at 280 nm, is given by this empirical equation:

$$\epsilon(280)\left(\mathrm{M}^{-1}\ \mathrm{cm}^{-1}\right) = (\#\ \mathrm{Trp})(5,500) + (\#\mathrm{Tyr})(1,490) + (\#\ \mathrm{cystine}\ )(125) \tag{3.3.2}$$

Proteins also absorb strongly at wavelengths less than 240 nm. This part of the absorption spectra arises from the above-mentioned amino acids along with contributions from His, Met, Cys, and the peptide bond. At these wavelengths, the absorbance is less dependent on the actual amino acid composition of the protein, but it becomes increasingly susceptible to interfering substances. Contaminating nucleic acids, which absorb maximally at 260 nm, also contribute to the absorbance at 280 nm. Hence the A280/A260 ratio can be determined and through appropriate calculations, the contribution of nucleic acids can be removed. Optimal reliability is obtained by measuring A_{280}/A_{205} values, since at 205 nm, a large fraction of the absorbance is derived from the peptide bond. pH changes have little effect on the absorbance of the peptide bond but has a much larger effect on Tyr. At high pH's, the side change hydroxyl is deprotonated (pKa = 10.5), with concomitant changes in the A_{295}. These changes can be used to follow the titration of the Tyr residues in a protein.

3.3.2.2: Molecular Weight

Molecular weights can be estimated by hydrodynamic techniques including size exclusion chromatography (under denaturing conditions using standards of known molecule weight), ultracentrifugation, and dynamic light scattering. In addition, they can be determined by polyacrylamide gel electrophoresis, again under denaturing conditions using standards. More precisely they could be determined through protein sequence or more easily through cDNA sequence analyses. The most accurate method for smaller proteins is mass spectrometry, as described in below.

3.3.2.3: Specific Amino Acids

Aromatic amino acids can be detected by their characteristic absorbance profiles. Amino acids with specific functional groups can be determined by chemical reactions with specific modifying groups, as shown in a previous section.

3.3.2.4: Amino Acid Composition

At a low level of resolution, we can determine the amino acid composition of the protein by hydrolyzing the protein in 6 N HCl, 100ºC, under vacuum for various time intervals. After removing the HCl, the hydrolysate is applied to an ion exchange or hydrophobic interaction column, and the amino acids are eluted and quantitated with respect to known standards. A non naturally- occurring amino acid, like norleucine, is added in known amounts as an internal standard to monitor quantitative recovery during the reactions. The separated amino acids are often derivatized with ninhydrin or phenylisothiocyantate to facilitate their detection. The reaction is usually allowed to proceed for 24, 36, and 48 hours since amino acids with OH (like ser) are destroyed. A time course allows the concentration of Ser at time t=0 to be extrapolated. Trp is also destroyed during the process. In addition, the amide links in the side chains of Gln and Asn are hydrolyzed to form Glu and Asp, respectively.

3.3.2.5: N- and C-Terminal Amino Acid Analysis

The amino acid composition does not give the sequence of the protein. The N-terminus of the protein can be determined by reacting the protein with fluorodinitrobenzene (FDNB) or dansyl chloride, which reacts with any free amine in the protein, including the epsilon amino group of lysine. The amino group of the protein is linked to the aromatic ring of the dinitrobenzene through an amine and to the dansyl group by a sulfonamide, and are hence stable to hydrolysis. The protein is hydrolyzed in 6 N HCl, and the amino acids are separated by TLC or HPLC. Two spots should result if the protein was a single chain, with some Lys residues. The labeled amino acid other than Lys is the N-terminal amino acid. The C-terminal amino acid can be determined by the addition of carboxypeptidases, enzymes which cleave amino acids from the C-terminal. A time course must be done to see which amino acid is released first. N-terminal analysis can also be done as part of sequencing the entire protein as discussed using Edman degradation.

3.3.3: Primary Sequence

3.3.3.1: Protein Sequencing using Edman Degradation

Edman degradation, developed by Pehr Edman, is a method of sequencing amino acids in a peptide. In this method, the amino-terminal residue is labeled and cleaved from the peptide without disrupting the peptide bonds between other amino acid residues. The reaction is shown in Figure 3.3.6.

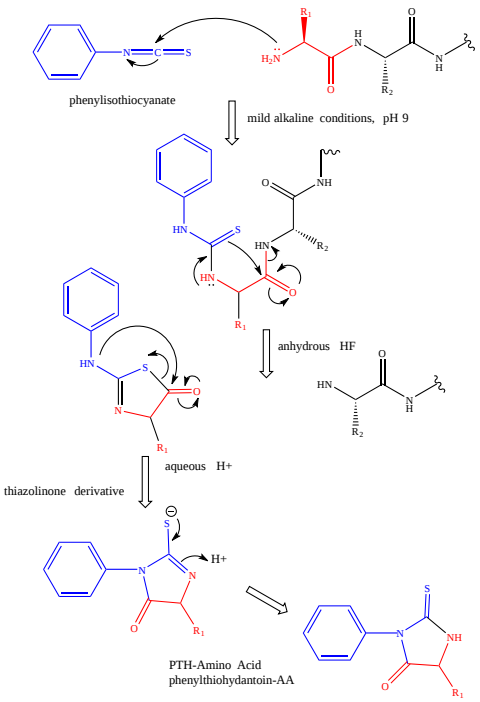


Figure 3.3.6: Mechanism of the Edman Degradation

Phenyl isothiocyanate is reacted with an uncharged N-terminal amino group, under mildly alkaline conditions, to form a cyclical *phenylthiocarbamoyl* derivative. Then, under acidic conditions, this derivative of the terminal amino acid is cleaved as a thiazolinone derivative. The thiazolinone amino acid is then selectively extracted into an organic solvent and treated with acid to form the more stable phenylthiohydantoin (PTH)- amino acid derivative that can be identified by using chromatography or electrophoresis. This procedure can then be repeated again to identify the next amino acid.

A major drawback to Edman degradation is that the peptides being sequenced in this manner cannot have more than 50 to 60 residues (and in practice, under 30). The peptide length is limited due to the cyclical derivatization not always going to completion. The derivatization problem can be resolved by cleaving large peptides into smaller peptides before proceeding with the reaction. It is able to accurately sequence up to 30 amino acids with modern machines capable of over 99% efficiency per amino acid. An advantage of the Edman degradation is that it only uses 10 - 100 pico-moles of peptide for the sequencing process. The Edman degradation reaction was automated in 1967 by Edman and Beggs to speed up the process and 100 automated devices were in use worldwide by 1973.

Because the Edman degradation proceeds from the N-terminus of the protein, it will not work if the N-terminus has been chemically modified (e.g. by acetylation or formation of pyroglutamic acid). Sequencing will stop if a non-α-amino acid is encountered (e.g. isoaspartic acid), since the favored five-membered ring intermediate is unable to be formed. Edman degradation is generally not useful to determine the positions of disulfide bridges. It also requires peptide amounts of 1 picomole or above for discernible results.

3.3.4: Secondary Structure

The percent and type of secondary structure can be determined using circular dichroism (CD) spectroscopy. In this method, right and left circularly polarized light illuminates a protein, which, since it is made of all L-amino acids, is chiral. (The mirror image would be a protein of the same sequence made of D-amino acids.) Differential absorption of the right and left forms give a CD spectrum

Circularly polarized light can be made when plane-polarized light of the same amplitude and wavelength meet out of phase by 90^0. (If they were out of phase by 180^0, they would cancel.)

- To see an animation of how circularly polarized light can be created, go to this page and select: 1. Superposition of plane-polarized waves 2.

If R and L circularly polarized light of the same wavelength and amplitude are passed through an optically inactive medium, the two waves combine (vectorially) to produce **plane-polarized light**.

- To see an animation of how circularly polarized light can be created, go to this page and select: 1. Superposition of circularly polarized waves

Optical activity is observed only when the environment in which a transition occurs is asymmetric.

The peptide (amide) bond absorbs UV light in the range of 180 to 230 nm (far-UV range) so this region of the spectra give information about the protein backbone, and more specifically, the secondary structure of the protein. The main electronic energy transitions are n → π* at 220 nm and π → π * at 190 nm for the peptide bond. There is a contribution from aromatic amino acid side chains but it is small, given the large number of peptide bonds. The lone pair on the nitrogen adjacent to the pi bond can be considered to be rehybridized from sp^3 to sp^2, allowing for conjugation of the p electrons (which lowers the energy of the electrons). The Hückel diagram shown in Figure 3.3.7 below shows 3 molecular (not atomic) orbitals generated from the 3 atomic p orbitals.

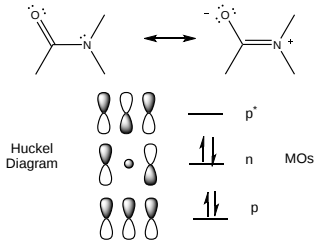


Figure 3.3.7: Hückel molecular orbitals for the peptide bond

The middle one (with 1 node) has energy similar to the separate atomic p orbitals and is considered a nonbonding molecular orbital. This is consistent with the lone nonbonding pair on the nitrogen atom.

The peptide bonds in a protein's asymmetric environment will absorb this wavelength range of light (promoting electrons to higher energy levels). In different secondary structures, the peptide bond electrons will absorb right and left circularly polarized light differently (for example, they have different molar absorptivities). Hence α, β and random coil structures all have distinguishable far UV CD spectra.

To see an animation of circularly polarized light, go to this page and select 1. Circularly polarized Waves

Stated in another way, if plane-polarized light, which is a superposition of right and left circularly polarized light, passes through an asymmetric sample, which absorbs right and left circularly polarized differently (i.e they display circular dichroism), then the light passing through the sample after vector addition of the right and left hand circularly polarized light gives elliptically polarized light.

To see an animation of elliptically polarized light, go to this page and select 2. Plane-polarized waves in a medium with circular dichroism

If the chiral molecules also have a different index of refraction for R and L circularly polarized light, an added net effect is the rotation of the angle of the elliptically of the polarized light. The far-UV CD spectrum of the protein is sensitive to the main chain conformation. The CD spectra of alpha and beta secondary structures are shown in Figure 3.3.8.

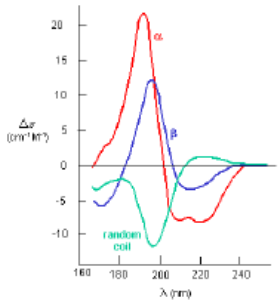


Figure 3.3.8: CD spectra or alpha, beta, and random coil structures in proteins

Protein side chains also find themselves in such an asymmetric environment. If irradiated with circularly polarized UV light in the range of 250-300 nm (near UV), differential absorption of right and circularly polarized light by the aromatic amino acids (Tyr, Phe, Trp) and disulfide bonds occur and a near UV CD spectra result. If the near UV CD spectra of a protein are taken under two different sets of conditions, and the spectra differ, then it can be surmised that the environment of the side chains is different, and hence the proteins have somewhat different conformations. It will not give information about the secondary structure of the backbone since that requires lower wavelengths for absorption of occur. Rather it can show differences in tertiary structure.

3.3.5: Analysis of Proteins Using Fluorescence Spectroscopy

Fluorescence spectroscopy is widely used to study many aspects of protein chemistry. This technique is not often used in lower-level undergraduate classes but it has become so important in the study of biomolecules, that a somewhat detailed explanation is necessary.

When electrons in a molecule absorb energy, they are promoted to higher electronic energy states. This is the basis of absorption spectroscopy. These excited state electrons can return to the ground state in processes that don't emit photons of light (ie. nonradiative processes) or radiative processes that do emit light. In simple absorption spectroscopy, the excited state electrons relax to the ground state through collision interactions. In radiative deexcitation, light is emitted. This process of light emission is called **luminescence**, which can be divided into two categories:

- **fluorescence**: If one electron from a ground state electron pair is excited to a higher energy state, the excited electrons can still be spin paired with its ground state counterpart - i.e. they have opposite spins. The excited electron can return to the ground state without reversing its spin. (The excited state is a **singlet** state with S, the total spin state, given the formula S = 2s +1 where s = 0 (sum of +1/2 and -1/2) and S = 1 for a singlet.) This process, which results in a rapid emission of a photon, is "spin allowed". The rate of photon emission is about 10^8 s^{-1}, which results in a lifetime (the average time between excitation and emission) of the excited state of about 10 ns.
- **phosphorescence**: If, in contrast to the above case, the spin of the excited electron is flipped, then its transition back to the ground state is "spin forbidden" since the excited state electron and its ground state counterpart have the same spin state. (The excited state is a **triplet** state with S, the total spin state, given the formula S = 2s +1 where s = 1 (1/2 + 1/2) and S = 3 for triplet). Hence this transition occurs slowly (in the ms - s range). Toys that glow in the dark display even longer phosphorescence lifetimes. (Note: This guide will concentrate on fluorescence.)

Competing with the two deexcitation process are nonradiative processes (such as through collisions). Given these competing processes, it might be expected that phosphorescence in liquid solutions at room temperature might not be detectable

Molecules which fluoresce are typically aromatic, which absorb readily in the UV and visible light regions. Common fluorophores are quinine, found in tonic water (observe the faint blue glow at the surface when it is placed in direct sunlight), and fluorescein and rhodamine, two fluorophores often added to antifreeze. Atoms are usually nonfluorescent, with the exception of europium and terbium ions from the lanthanide series. These fluoresce when electronic transitions occur between f orbitals, which are shielded from solvent relaxation in these particular ions.

Among biological molecules, some, especially macromolecules with aromatic groups, fluoresce. These groups are called **intrinsic fluorophores**, and include, in the case of proteins, the side chains of tryptophan, tyrosine, and phenylalanine, the aromatic amino acids. The indole side chain of tryptophan is the most fluorescent, and its emission spectra, which is sensitive to solvent conditions, is often blue-shifted when it is buried, and red-shifted when solvent-exposed. Nucleic acids, although they also contain aromatic bases, are poor fluorophores. Many biological molecules can be made fluorescent by covalently modifying them (through nucleophiles on the biological molecule) with exogenously added fluorophores, such as fluorescein isothicyanate, rhodamine isothiocyante, dansyl chloride, etc. These are called **extrinsic fluorophores**. These include molecules that bind noncovalently to structures such as ds-DNA (ethidium bromide) or lipid membranes (diphenylhexatriene). Some biological fluorophores are substrates for enzyme reactions. An example is the oxidized flavins (FAD, FMN) and the reduced form of NAD (i.e. NADH). Another type of useful fluorophore is indicators, whose fluorescent properties change with a change in a parameter like pH or [Ca ion].

The electronic transitions that occur during fluorescence can be represented by a Jablonski diagram as shown in the two-part Figure 3.3.9 (A and B) below.

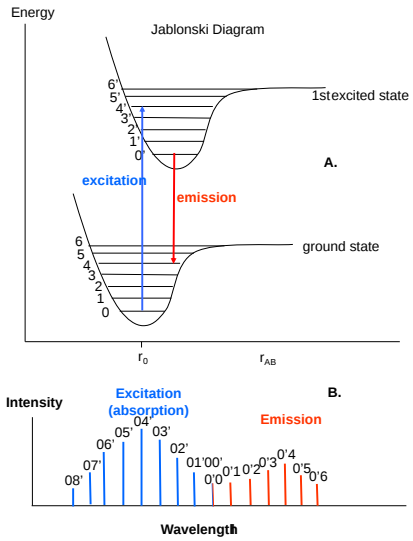


Figure 3.3.9: Jablonski diagran and associated excitation and emission spectra. After Van Holde, et al. *Principles of Physical Biochemistry*, Pearson Prentice Hall (2006)

In panel A, the ground and first excited electronic state are shown. Within each electronic state are multiple vibrational energy levels 0, 1, 2 ..and 0'. 1' 2' .... This simple diagram ignores quenching of fluorescence, resonance energy transfer, etc. The transitions, represented by vertical lines, are considered to be instantaneous. They take about 10^{-15} s so the nuclei don't move in the process. The ground state electron is considered to be in the 0 vibrational level, So, since thermal energy is insufficient to promote it to the next vibrational level. When light is absorbed, the electron is promoted to a higher vibrational level within a higher electronic level. Usually, the excited electrons relaxes quickly (< 1 ps) to the lowest vibrational level of S1 or possibly S2 through a process called **internal conversion**. Fluorescence emission then may occur from the lowest vibrational state of S1 to any of the vibrational states of So. Hence the photon emitted is lower in energy (longer in wavelength) than the absorbed photon. Also since both process involve the movement of the electron to different vibrational levels with absorption or emission of a photon, and nonradiative vibrational relaxation within those levels, the emission spectra is often the mirror image of the absorption spectra. (This assumes that the vibrational levels in So and S1 are similarly spaced. Alternatively, electrons in S1 may flip spin and convert to the T1 state, in a process called intersystem crossing, leading to phosphorescence.

Panel B above shows blue lines corresponding to individual absorbances and red lins corresponding to emissions. Note that the excitation from 0 to 8' has the highest energy of absorbance (lowest wavelength) but gives little intensity as it would occur with low frequency. If you were to draw a line over the tops of the lines in panel B you would get a simple excitation spectra and emission spectra, which would be mirror images of each other. The emission peak is at a longer wavelength since energy was lost on vibrational, nonradiative relaxation of the excited electron. The difference in peak wavelengths of excitation and emission is called the **Stokes Shift.** This shift is greatest for fluorophores in polar environments. Inferences can be made considering the disposition of a side chain (buried or surface) if changes in fluorescence properties (intensity, Stoke's shift) are noted on protein denaturation. Also, many probes are weakly fluorescent in aqueous solution, but fluoresce intensely in nonpolar mediums (bound to a hydrophobic pocket in a protein, in a bilayer or lipoprotein, etc.)

Emission spectra are usually independent of excitation wavelength (Kasha's rule): This occurs because of the rapid relaxation into the lowest vibrational energy level of the excited state. There are also exceptions to the mirror image rule. Deviations arise from a change in the geometry of nuclei in the excited state molecule. This may occur if the lifetime of the S1 state is long, allowing time for motion before emission. An example of this can be seen with p-terphenyl in cyclohexane, in which the rings become more coplanar in the excited state. Since there is an electron shift in the excited state, a complex between the excited fluorophore and another solution component may arise (charge-transfer complex). Alternatively, some fluorophores form complexes with themselves (pyrene) with increasing concentration. At high concentrations, changes in the emission spectra occur, arising from emission from an excited-state dimer or **excimer**. Acridine shows two emission spectra at different pH's, arising from changes in the pKa on excitation (5.45 to 10.7). Finally, exciting a fluorophore at different wavelengths (EX 1, EX 2, EX 3) does not change the emission profile but does produce variations in fluorescence emission intensity (EM 1, EM 2, EM 3) that correspond to the amplitude of the excitation spectrum.

Fluorophores can be used to chemically modify nucleophilic side chains such as lysines and cysteines. Changes in intrinsic fluorescence in proteins can be used to measure the binding of ligands and conformation changes in the protein that occur on binding interactions, change in solution conditions, and protein denaturation. Let's explore a few fluorescence methods widely used to explore protein structure and function.

3.3.5.1: Fluorescence Quenching

Some chemical species (for example iodide and monomeric unpolymerized acrylamide), when added to a protein solution, can decrease the fluorescence from an intrinsic surface accessible fluorophore such as the tryptophan side chain, providing information on the local environment of the intrinsic fluorophore (example tryptophan side chain accessibility). For example, a buried tryptophan or probe will show little change in fluorescence intensity in the presence of a large, polar quencher, while a surface tryptophan or probe will show a significant decrease in fluorescent intensity. It is somewhat amazing that O_2 when added to a solution under increasing pressure, can quench the fluorescence of even buried tryptophan side chain, implying that there are minimal diffusional barriers to O_2 access. This suggests significant conformational flexibility of the protein.

Quenching can be dynamic, occurring on collision of the quench with the intrinsic fluorophore, or static, when the quencher binds to a site near the fluorophore as a prelude to quenching.

Collisional quenching is described by the **Stern-Volmer** equation.

$$\frac{F_0}{F} = 1 + k_q \tau_0 [Q] = 1 + K_D [Q] \tag{3.3.3}$$

where $\mathbf{F_o}$ and **F** are the fluorescent intensities in the absence and presence of the quencher, $\mathbf{k_q}$ is the biomolecular quenching constant, $\boldsymbol{\tau_o}$ is the lifetime of the fluorophore in the absence of the quencher and **[Q]** is the concentration of the quencher. $\mathbf{k_q\tau_o = K_D}$ is the **Stern-Volmer quenching constant.**

A plot of F_o/F vs [Q] is linear, with a slope of K_D. $1/K_D$ is the quencher concentration at which $F_o/F = 2$, or 50% of the fluorescence intensity is quenched. A linear plot indicates a single class of fluorophores, all equally accessible to the quencher. A nonlinear plot would be found for quenching of tryptophan fluorescence in proteins by charged or polar quenchers for proteins with more than one tryptophan and in which some are buried. Static quenching also results in a linear SV plot. Dynamic and static quenching can be distinguished by different dependencies on temperature and viscosity. Since dynamic quenching depends on diffusion, and higher temperatures result in higher diffusion coefficients, k_q should increase with temperature. If static quenching is involved, higher temperatures will probably reduce complex formation.

3.3.5.2: Fluorescence Resonance Energy Transfer (FRET)

If an absorbing species is in close proximity to an excited state fluorophore, and if the emission spectra of the fluorophore overlaps the absorption spectra of the second species, coupling of the two dipoles can occur, and energy can be transferred from the excited state of the fluorophore (donor D) to the second absorbing species (acceptor A). This transfer of energy is through dipole coupling and **not** through the trivial release and absorption of an emitted photon. No photon is produced. This process is called **fluorescence resonance energy transfer (FRET).** Efficiency, E, of FRET for a single donor/acceptor pair at a fixed distance is given by:

$$\mathrm{E} = \frac{\mathrm{R}_0^6}{\left(\mathrm{R}_0^6 + \mathrm{r}^6\right)} = \frac{1}{1 + \left(\frac{\mathrm{r}}{\mathrm{R}_0}\right)^6} \tag{3.3.4}$$

where R_o is the Förster distance (or radius) with a 50% transfer efficiency and r is the distance between the donor and acceptor. R_o is a measure of the spectral overlap of the donor and acceptor (for which most biological macromolecules have a similar value of 30-60 angstroms). This equation shows an efficiency dependent on $1/r^6$, making FRET exquisitely sensitive to distance. FRET and its dependency of distance is illustrated in Figure 3.3.10 below.

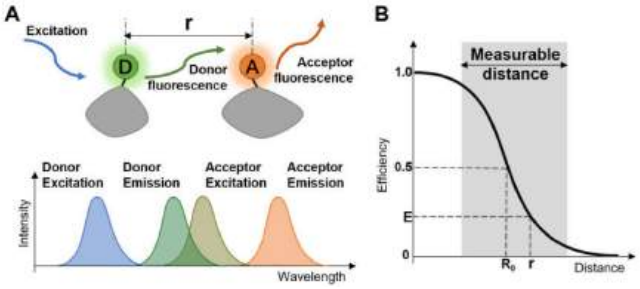


Figure 3.3.10:Fluorescence Resonance Energy Transfer (FRET) Qiao, Y.; Luo, Y.; Long, N.; Xing, Y.; Tu, J. Single-Molecular Förster Resonance Energy Transfer Measurement on Structures and Interactions of Biomolecules. *Micromachines* **2021**, *12*, 492. https://doi.org/10.3390/mi12050492Creative Commons Attribution (CC BY) license (**https://creativecommons.org/licenses/by/4.0/**).

3.3.5.3: Anisotropy or polarization

These measure the extent of rotation of the fluorophore during its fluorescent lifetime. If a small fluorophore binds to a large molecule, its rotation diffusion constant decreases, and its anisotropy increases as illustrated in Figure 3.3.11.

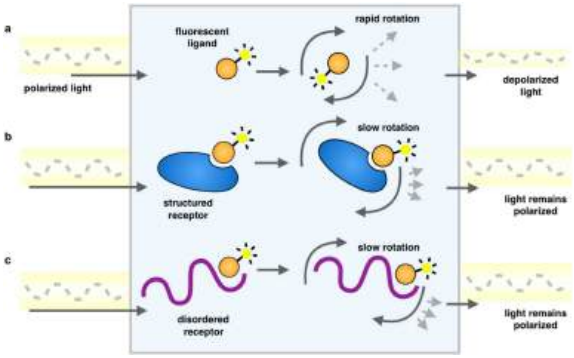


Figure 3.3.11: Fluorescence anisotropy and rotational diffusion of a fluorophore. -2563-4Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/).

Schematic representation of a fluorescence polarization experiment. As a result of rapid tumbling of molecules in solution, when a fluorescently labeled ligand is excited with plane-polarized light, the resulting emitted light is largely depolarized (a). Upon binding another species, a larger proportion of the emitted light remains in the same plane as the excitation energy, because the rotation is slowed as the effective molecular size increases, whether it is an ordered molecular structure (b) or one that is disordered (c)". Since viscosity decreases rates of rotational diffusion, changes in fluorescence (such as inside a bilayer) can be inferred from these measurements. For example, membranes more enriched in saturated fatty acids should show increased anisotropy of a hydrophobic, fluorescent probe, in comparison to the same probe in a bilayer enriched in polyunsaturated fatty acids.

3.3.6: Analysis of Protein Using Mass Spectrometry

Mass spectrometry is supplanting more traditional methods (see above) as the choice to determine the molecular mass and structure of a protein. Its power comes from its exquisite sensitivity and modern computational methods to determine structure through comparisons of ion fragment data with computer databases of known protein structures. In mass spectrometry, a molecule is first ionized in an ion source. The charged particles are then accelerated by an electric field into a mass analyzer where they are subjected to an external magnetic field. The external magnetic field interacts with the magnetic field arising from the movement of the charged particles, causing them to deflect. The deflection is proportional to the mass to charge ratio, m/z. Ions then enter the detector which is usually a photomultiplier. Sample introduction into the ion source occurs though simple diffusion of gases and volatile liquids from a reservoir, by injection of a liquid sample containing the analyte by spraying a fine mist, or for very large proteins by desorbing a protein from a matrix using a laser. Analysis of complex mixtures is done by coupling HPLC with mass spectrometry in a LCMS.

There are many methods to ionize molecules, including atmospheric pressure chemical ionization (APCI), chemical ionization (CI), or electron impact (EI). The most common methods for protein/peptide analyses are electrospray ionization (ESI) and matrix-assisted laser desorption ionization (MALDI).

3.3.6.0.1: Electrospray ionization (ESI)

The analyte, dissolved in a volatile solvent like methanol or acetonitrile, is injected through a fine stainless steel capillary at a slow flow rate into the ion source. A high voltage (3-4 kV) is maintained on the capillary giving it a positive charge with respect to the other oppositely charged electrode. The flowing liquid becomes charged with the same polarity as the polarity of the positively charged capillary. The high field leads to the emergence of the sample as a charged aerosol spray of charged microdrops, which reduces electrostatic repulsions in the liquid. This method essentially uses electrical energy to produce the aerosol instead of mechanical energy to produce a liquid aerosol, as in the case of a perfume atomizer. Surrounding the capillary is a flowing gas (nitrogen) which helps to move the aerosol toward the mass analyzer. The microdrops become smaller in size as the volatile solvent evaporates, increasing the positive charge density on the drops. Eventually, electrostatic repulsions cause the drops to explode in a series of steps, ultimately producing the analyte devoid of solvent. This gentle method of ionization produces analytes that are not cleaved but ready for introduction into the mass analyzer. Proteins emerge from this process with a roughly Gaussian distribution of positive charges on basic side chains. In organic chemistry, you studied mass spectrums of small molecules induced by electron bombardment. This produces ions of +1 charge as an electron is stripped away from the neutral molecule. The highest m/z peak in the spectrum is the parent ion or M+ ion. The highest m/z ratio detectable in the mass spectrum is in the thousands. However, large peptides and proteins with large molecular masses can be detected and resolved since the charge on the ions are great than +1. In 2002, John Fenn was awarded a Noble Prize in Chemistry for the development and use of ESI to study biological molecules.

An example of an ESI spectrum of apo-myoglobin is shown in Figure 3.3.12

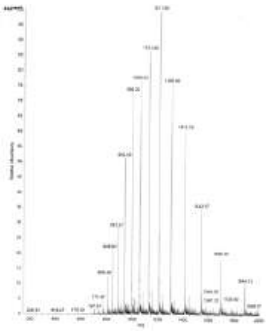


Figure 3.3.12: ESI mass spectrum of apo-myoglobin

Note the roughly Gaussian distribution of the peaks, each of which represents the intact protein with charges differing by +1. Proteins have positive charges by virtue of both protonation of amino acid side chains as well as charges induced during the electrospray process itself. Based on the amino acid sequence of myoglobin and the assumption that the pKa of the side chains are the same in the protein as for isolated amino acids, the calculated average net charges of apomyoglobin (apoMb) would be approximately +30 at pH 3.5, +20 at pH 4.5, +9 at pH 6, and 0 at pH 7.8 (the calculated pI). The mass spectrum below was taken by direct injection into the MS of apoMb in 0.1% formic acid (pH 2.8). Charges on the peptide result from those present on the peptide before the electrospray and changes in charges induced during the process.

The molecular mass of the protein can be determined by analyzing two adjacent peaks, as shown in Figure 3.3.13

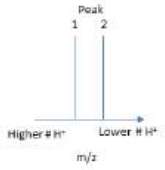


Figure 3.3.13: m/z ratio for two adjacent peaks in the mass spectrum of a protein

If M is the molecular mass of the analyte protein, and n is the number of positive charges on the protein represented in a given m/z peak, then the following equations gives the molecular mass M of the protein for each peak:

$$\begin{aligned} M_{\text{peak 2}} &= n(m/z)_{\text{peak 2}} - n(1.008) \\ M_{\text{peak 1}} &= (n+1)(m/z)_{\text{peak 1}} - (n+1)(1.008) \end{aligned} \tag{3.3.5}$$

where 1.008 is the atomic weight of H. Since there is only one value of M, the two equations can be set equal to each other, giving:

$$n(m/z)_{\text{peak 2}} - n(1.008) = (n+1)(m/z)_{\text{peak 1}} - (n+1)(1.008) \tag{3.3.6}$$

Solving for n gives:

$$n = \frac{(m/z)_{\text{peak1}} - (1.008)}{(m/z)_{\text{peak2}} - (m/z)_{\text{peak1}}} \tag{3.3.7}$$

Knowing n, the molecular mass M the protein can be calculated for each m/z peak. The best value of M can then be determined by averaging the M values determined from each peak (16,956 from the above figure). For peaks from m/z of 893-1542, the calculated values of n ranged from +18 to +10.

3.3.6.0.1: Matrix assisted laser desorption ionization (MALDI)

In this technique, used for larger biomolecules like proteins and polysaccharides, the analyte is mixed with an absorbing matrix material. Laser excitation is used to excite the matrix, leading to energy transfer that results in ionization and the "launching" of the matrix and analyte in ion form from the solid mixture. Parent ion peaks of $(M+H)^+$ and $(M-H)^-$ are formed.

Mass Analyzer takes the created ion and separates them based on m/z ratios. Let's consider two, quadrupole ion tap and time of flight. on mass-to-charge ratios. There are several general types of mass analyzers, including **magnetic sector, time of flight, quadrupole, ion trap**

Quadrupole ion trap (used in ESI) - A complex mixture of ions can be contained (or trapped) in this type of mass analyzer. Two common types are linear and 3D quadrupole. If a dipole has two poles (+ and -) separated by some distance, then a quadrupole has four poles (+, -, +, and -) arranged geometrically such that each + has a - on each side and vice versa. Figure 3.3.14below shows linear and 3D quadrupoles

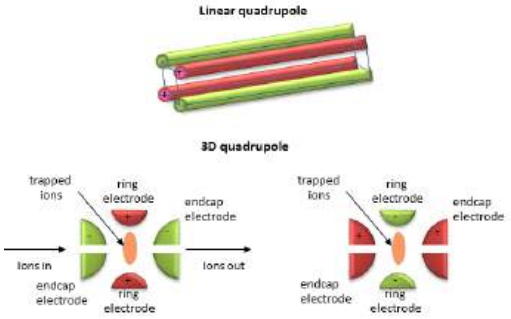


Figure 3.3.14: Linear and 3D quadrupoles

As dipoles display positive and negative charge separation on a linear axis, quadrupoles have either opposite electrical charges or opposite magnetic fields at the opposing ends of a square or cube. In charge separation, the monopole (sum of the charges) and dipoles cancel to zero, but the quadrupole moment does not. The quadrupole traps ions using a combination of fixed and alternating electric fields. The trap contains helium at 1 mTorr. For the 3D trap, The ring electrode has an oscillating RF voltage which keeps the ions trapped. The end caps also have an AC voltage. Ions oscillate in the trap with a "secular" frequency determined by the frequency of the RF voltage, and of course, the m/z ratio. By increasing the amplitude of the RF field across the ring electron, ion motion in the trap becomes destabilized and leads to ion ejection into the detector. When the secular frequency of ion motion matches the applied AC voltage to the end cap electrodes, resonance occurs and the amplitude of motion of the ions increases, also allowing leakage out of the ion trap into the detector.

Tandem Mass Spectrometry (MS/MS): Quadrupole mass analyzers, which can select ions of varying m/z ratios in the ion traps, are commonly used in tandem mass spectrometry (MS/MS). In this technique, the selected ions are further fragmented into smaller ions by a process called collision-induced dissociation (CID). When performed on all of the initial ions present in the ion trap, the sequence of a peptide/protein can be determined. This technique usually requires two mass analyzers with a collision cell in between where selected ions are fragmented by collision with an inert gas. It can also be done in a single mass analyzer using a quadrupole ion trap.

Time of Flight (TOF) tube (used in MALDI) - a long tube is used and the time required for ion detection is determined. The small molecular mass ions take the shortest time to reach the detector.

3.3.6.1: Sequence Determination Using Mass Spectrometry

In a typical MS/MS experiment to determine a protein sequence, a protein is cleaved into protein fragments with an enzyme such as trypsin, which cleaves on the carboxyl side of positively charge Lys and Arg side chains. The average size of proteins in the human proteome is approximately 50,000. If the average molecular mass of an amino acid in a protein is around 110 (18 subtracted since water is released on amide bond formation), the average number of amino acids in the protein would be around 454. If 10% of the amino acids are Arg and Lys, then on average there would be approximately 50 Lys and Arg, and hence 50 tryptic peptides of average molecular mass 1000. The fragments are introduced in the MS where a peptide fragment fingerprint analysis can be performed. The molecular weights of the fragments can be identified and compared to known peptide digestion fragments from known proteins to identify the analyte protein.

Ions with the original N terminus are denoted as a, b, and c, while ions with the original C terminus are denoted as x, y, and z. c and y ions gain an extra proton from the peptide to form positively charged -NH_3^+ groups. Figure 3.3.15below shows peaks for a 4-amino acid peptide fragmentation pattern

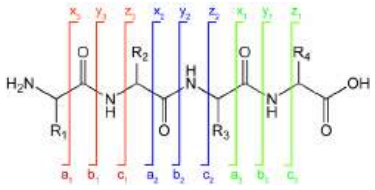


Figure 3.3.15: 4-amino acid peptide fragmentation peaks. https://commons.wikimedia.org/wiki/F...gmentation.gif. Creative Commons Attribution-Share Alike 3.0 Unported

Ions with the original N terminus are denoted as a, b, and c, while ions with the original C terminus are denoted as x, y, and z. c and y ions gain an extra proton from the peptide to form positively charged -NH_3^+ groups. The actual ions observed depend on many factors including the sequence of the peptide, its original charge, the energy of the collision inducing the fragmentation, etc. Low energy fragmentation of peptides in ion traps usually produces a, b, and y ions, along with peaks resulting from loss of NH_3 (a*, b* and y*) or H_2O (ao, bo, and yo). No peaks resulting from the fragmentation of side chains are observed. Fragmentation at two sites in the peptide (usually at b and y sites in the backbone) forms an internal fragment.

The y1 peak represents the C-terminal Lys or Arg (in this example) of the tryptic peptide. Peak y2 has one additional amino acid compared to y1 and the molecular mass difference identifies the extra amino acid. Peak y3 is likewise one amino acid larger than y2. All three y fragments peaks have a common Lys/Arg C-terminal and the charged fragment contains the C-terminal end of the original peptide. All b fragment peaks for a given peptide contain a common N terminal amino acid with b1 the smallest. Note that the subscript represents the number of amino acids in the fragment. By identifying b and y peaks the actual sequence of small peptides can be determined. Usually, spectra are matched to databases to identify the structure of each peptide and ultimately that of the protein. The actual m values for fragments can be calculated as follows, where (N is the molecular mass of the neutral N terminal group, (C) is the molecular mass of the neutral c terminal group, and (M) is the molecule mass of the neutral amino acids. (For N terminal amino acid, add 1 H. for C terminus add OH)

- a: (N)+(M)-CHO
- b: (N)+(M)
- y: (C)+(M)+H (note in the figure above that the amino terminus of the y peptides has an extra proton in the +1 charged peptides.)

m/z values can be calculated from the calculated m values and by adding the one H mass to the overall z if the overall charge is +1, etc.

As an example, from these MW values, the sequence of the human Glu1- fibrinopeptide B can be determined from MS/MS spectra shown in an annotated form in Figure 3.3.16. Note that most of the b peaks are b* resulting from the loss of NH_3 from the N terminus.

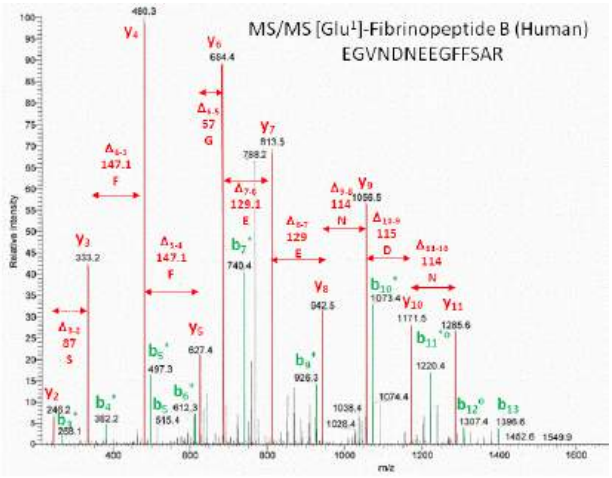


Figure 3.3.16: Annotated MS/MS spectra of human Glu1- fibrinopeptide B

Now let's step back and get a broader picture of structure analysis by mass spectrometry. In general, proteins are analyzed either in a "top-down" approach in which proteins are analyzed intact, or a "bottom-up" approach in which proteins are first digested into fragments. An intermediate "middle-down" approach in which larger peptide fragments are analyzed may also sometimes be used. The top-down approach however is mostly limited to low-throughput single-protein studies due to issues involved in handling whole proteins, their heterogeneity, and the complexity of their analyses.

In the second approach, referred to as the "bottom-up" MS, proteins are enzymatically digested into smaller peptides using a protease such as trypsin, which cleaves peptide chains mainly at the carboxyl side of the amino acids lysine or arginine, except when either is followed by proline. It is used for numerous biotechnological processes. The process is commonly referred to as trypsin proteolysis or trypsinization, and proteins that have been digested/treated with trypsin are said to have been trypsinized.

Subsequently, these peptides are introduced into the mass spectrometer and identified by peptide mass fingerprinting or tandem mass spectrometry. Hence, this approach uses identification at the peptide level to infer the existence of proteins pieced back together with *de novo* repeat detection. The smaller and more uniform fragments are easier to analyze than intact proteins and can be also determined with high accuracy, this "bottom-up" approach is therefore the preferred method of studies in proteomic studies. A further approach that is beginning to be useful is the intermediate "middle-down" approach in which proteolytic peptides larger than the typical tryptic peptides are analyzed.

Proteins of interest are usually part of a complex mixture of multiple proteins and molecules, which co-exist in the biological medium. This presents two significant problems. First, the two ionization techniques used for large molecules only work well when the mixture contains roughly equal amounts of constituents, while in biological samples, different proteins tend to be present in widely differing amounts. If such a mixture is ionized using electrospray or MALDI, the more abundant species have a tendency to "drown" or suppress signals from less abundant ones. Second, the mass spectrum from a complex mixture is very difficult to interpret due to the overwhelming number of mixture components. This is exacerbated by the fact that the enzymatic digestion of a protein gives rise to a large number of peptide products.

In light of these problems, the methods of one- and two-dimensional gel electrophoresis and high-performance liquid chromatography are widely used for the separation of proteins. The first method fractionates whole proteins via two-dimensional gel electrophoresis (Figure 3.31). The first dimension of 2D gel electrophoresis is isoelectric focusing (IEF). In this dimension, the protein is separated by its isoelectric point (pI) and the second dimension is SDS-polyacrylamide gel electrophoresis (SDS-PAGE). This dimension separates the protein according to its molecular weight. Once this step is completed in-gel digestion occurs.

In some situations, it may be necessary to combine both of these techniques. Gel spots identified on a 2D Gel are usually attributable to one protein. If the identity of the protein is desired, usually the method of in-gel digestion is applied, where the protein spot of interest is excised and digested proteolytically. The peptide masses resulting from the digestion can be determined by mass spectrometry using peptide mass fingerprinting. If this information does not allow unequivocal identification of the protein, its peptides can be subject to tandem mass spectrometry for *de novo* sequencing. Small changes in mass and charge can be detected with 2D-PAGE. The disadvantages of this technique are its small dynamic range compared to other methods. Some proteins are still difficult to separate due to their acidity, basicity, hydrophobicity, and size (too large or too small).

The second method, high-performance liquid chromatography is used to fractionate peptides after enzymatic digestion. Characterization of protein mixtures using HPLC/MS is also called shotgun proteomics and MuDPIT (Multi-Dimensional Protein Identification Technology). A peptide mixture that results from the digestion of a protein mixture is fractionated by one or two steps of liquid chromatography. The eluant from the chromatography stage can be either directly introduced to the mass spectrometer through electrospray ionization, or laid down on a series of small spots for later mass analysis using MALDI.

The general schema for analyzing proteins by mass spectrometry is shown in Figure 3.3.17.

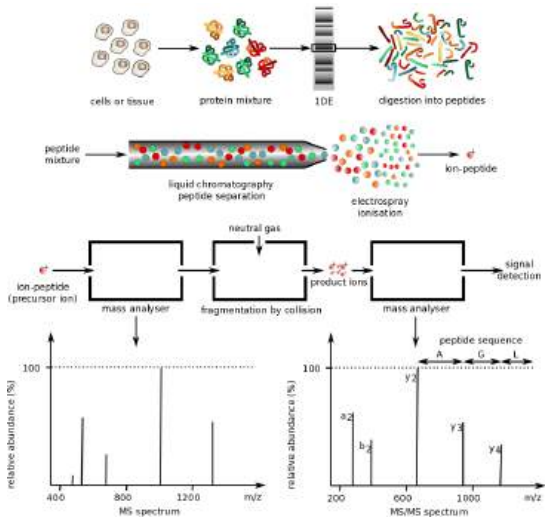


Figure 3.3.17: Schematic of Protein Fingerprinting by Mass Spectrometry. Philippe Hupé - Emmanuel Barillot, Laurence Calzone, Philippe Hupé, Jean-Philippe Vert, Andrei Zinovyev, Computational Systems Biology of Cancer Chapman & Hall/CRC Mathematical & Computational Biology , 2012

Protein mixtures are prepared from cell culture or tissue samples and separated by gel electrophoresis. Single proteins are isolated and digested using trypsin to produce a peptide mixture. Peptides are separated by liquid chromatography and analyzed by mass spectrometry

3.3.7: 3D Structural Determination

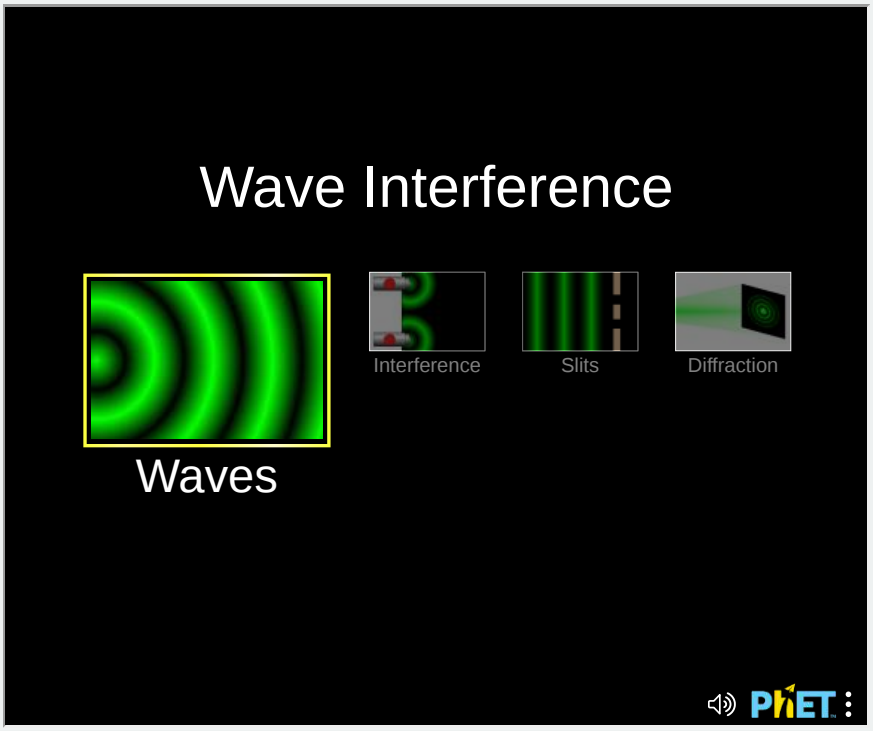
There are four main ways that the 3D structure of a protein can be determined. These include X-ray crystallography, multidimensional NMR, cryoelectron microscopy, and computer modeling using artificial intelligence and machine learning.

3.3.7.1: *X-Ray Crystallography*

In this technique, proteins are induced to form solid crystals in which the individual molecules pack in a well-defined crystal lattice. X-rays are aimed at the crystals. The x-rays are scattered off of the crystal and collected by a detector. The scattered x-rays engage in constructive and destructive interference (much as water waves do) and form a diffraction pattern. The diffraction pattern is determined by the spacing and types of atoms in the crystal. Hence from a given 3D structure, a specific diffraction pattern is formed. Using some sophisticated mathematics (Fourier Transformations), the diffraction pattern can be converted back into the 3D structure of the object, in this case, the atoms with the protein.

PHET Simulation - Diffraction

Constructive and destructive interference and the formation of a "diffraction" pattern can be readily seen performing two PHET simulations. Follow the instructions below.



Simulation 1: Slits

- Select Slits to open the simulation and select these choices in order:
- Type of wave: pick one that looks like a bullet

- Frequency (middle green)
- Amplitude: max
- Check Screen
- Choose 2 Slits
- Slit width 200
- Slit separation 400
- Click the Green button.

You will see light/dark patterns moving toward the screen. The light zones arise from constructive interference and dark from destructive interference of the two waves as they emerge from the slits.

Simulation 2: Diffraction

Now, look at the diffraction pattern arising from light moving through simple and more complex openings and interacting with an object, using the PHET animation and the step below.

- First, refresh the browser window
- Select Diffraction to open the simulation and select these choices in order
- Pick 450 nm wavelength
- Choose in succession the 4 vertical icons (circle, square, circle/square, array of circles, person)
- Observe the diffraction patterns as you change the slit size

X-ray scattering can also be envisioned as light "reflecting" from a series of planes formed by atoms in the crystal in which the planes are separated by specific distances (in the Angstroms range). The x-rays that are "reflected" from innumerable planes recombine constructively and destructively to form a diffraction pattern. X rays are used since the size of the "slits", and the distance between these "reflective" planes, must be comparable to the wavelength of the incident light, which for x-rays is 0.5 – 2.5 Å.

A diffraction pattern is mathematically decoded to form an electron density map since it is the electrons that actually scatter the x-rays. Hydrogen atoms don't appear in x-ray crystal structures since they don't have enough electrons to be effective scattering centers. Computer programs are used to fit the electron density map to a 3D arrangement of atoms separated by characteristic bond distances corresponding to the functions groups and side chains found in protein. The quality/amount of crystals helps determine the quality of the diffraction pattern and the resulting structure. X-ray crystallographers define quality in terms of resolution. A resolution of 5Å - 10Å can reveal the structure of polypeptide chains, 3 Å - 4 Å of groups of atoms, and 1 Å - 1.5 Å of individual atoms.

Figure 3.3.18 shows the process from crystal to model.

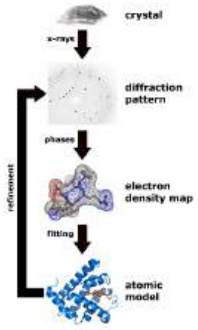


Figure 3.3.18: Overview of structure determination by mass spectrometry. Thomas Splettstoesser (www.scistyle.com), CC BY-SA 3.0 <https://creativecommons.org/licenses/by-sa/3.0>, via Wikimedia Commons

Figure 3.3.19 below shows the electron density map around tyrosine 103 from myoglobin from crystal structures at two different resolutions (left, 1a6m, 1.0 Å resolution and right, 108m, 2.7 Å resolution).

Figure 3.3.19: Electron density map around tyrosine 103 from myoglobin from crystal structures at two different resolutions (left, 1a6m, 1.0 Å resolution and right, 108m, 2.7 Å resolution). Click on the images to see full iCn3D models showing the electron density map around the Y103. (Choice of Y103 from https://pdb101.rcsb.org/learn/guide-...ata/resolution)

This figure shows 2Fo-Fc electron density maps which use the observed diffraction data, Fo, with the diffraction data calculated from the atomic model, Fc. Proteopedia has an excellent description of electron density maps.

Not all proteins can be readily crystallized. The process is in many ways an art as much as it is a science. Membrane proteins fall into this category.

3.3.7.2: Nuclear Magnetic Resonance (NMR)

Many readers have probably performed ^{1}H-NMR on small molecules introductory and organic chemistry labs. Interpreting spectra of molecules with many hydrogen atoms in straight and branched chains, in rings, and in functional groups is not simple. Imagine doing that to determine the structure of a small protein with 1000s of hydrogen atoms! The spectrum

would be essentially indecipherable. Luckily **multi-dimension NMR** techniques have allowed the solution (not crystal) structure of small proteins to be determined. These methods are outside of the scope of this book. For those interested in more detail, read A brief introduction to NMR spectroscopy of proteins by Poulsen.

Let's give a brief introduction to a 2D NMR peak for a simple molecule, ethylacetate. The 1D ^{1}H-NMR spectrum for the molecule is shown in Figure 3.3.20

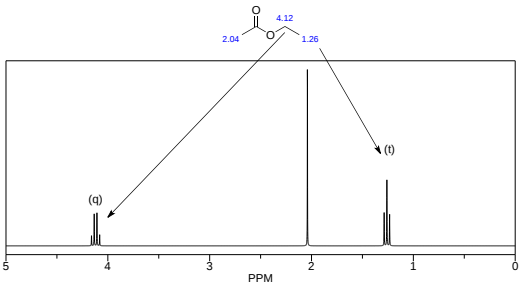


Figure 3.3.20: 1D ^{1}H-NMR spectrum of ethylacetate

Now let's show a simulated 2D COSY spectra of the same molecule. The image and explanation below are adapted from Structure & Reactivity in Organic, Biological and Inorganic Chemistry by Chris Schaller, which is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported License. Structure and Reactivity

In homonuclear correlation spectroscopy (COSY), we can look for hydrogens that are coupled to each other. In ethyl acetate, it's pretty clear where they are. There is a quartet and a triplet; the hydrogens corresponding to those two peaks are probably beside each other in the structure. The COSY spectrum simply takes that ^{1}H spectrum and spreads it out into two dimensions. Instead of being displayed as a row of peaks, the peaks are spread out into a 2D array. Figure 3.3.21 shows an annotated simulated COSY spectrum. The peaks are displayed along one axis and the same peaks are also displayed along the other axis.

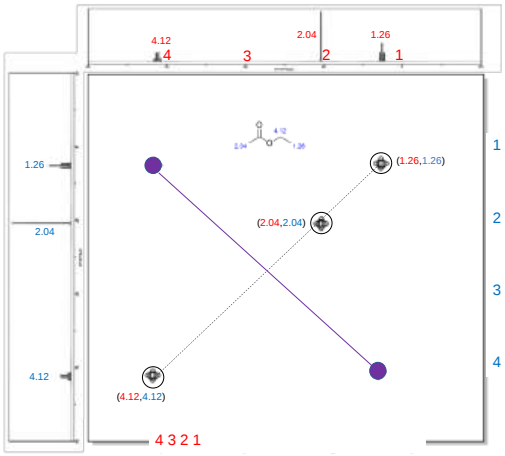


Figure 3.3.21: Annotated simulated COSY spectrum of ethylacetate

What does it mean to be coupled? It means that magnetic information is transmitted between the atoms. How can we tell? Essentially, we can send a pulse of electromagnetic radiation into one set of hydrogens and look for a response somewhere else. Of course, if we send a pulse of radio waves at a frequency that will be absorbed by a particular hydrogen, we will see a response in that hydrogen itself. That's why we see the peaks on the diagonal (dotted line). The peaks along the diagonal in the spectra (1.26, 1.26; 2.04, 2.04; 4.12, 4.12) hence don't give any new information since they are the main peaks in the 1D NMR.

However, we also see responses from other hydrogens that are magnetically linked to the original one. They give the peaks (shown as purple circles) that do **not** appear along the diagonal. Those peaks indicate which hydrogens are coupled to which other hydrogens. The hydrogens at 1.26 ppm are coupled to the ones at 4.12 ppm, and that gives a "**cross-peak**" at (1.26, 4.12). There is also a cross-peak at (4.12, 1.26), because that relationship goes both ways.

This coupling should make sense because the protons that give the signals at 4.12 (-O-C**H**$_2$-CH$_3$) and 1.26 (-CH$_2$-C**H**$_3$) are on adjacent carbons and split each other's signals as seen in the 1D NMR. This examples shows how we can get information on which protons in a 2D NMR spectra are coupled through 3 bonds (H-C-C-H), critical information for determining protein structures by NMR.

There are variants of 2D NMR that are used in protein structure as well. They use NMR-active nuclei in addition to ^{1}H, including ^{13}C (natural abundance 1%) and ^{15}N (natural abundance 0.37%). Given these low abundances, proteins for NMR structure determination are often purified in cells grown in media enriched in ^{13}C and ^{15}N precursors.

HMBC (Heteronuclear Multiple Bond Correlation) and HMQC (Heteronuclear Multiple Quantum Coherence): Just as COSY spectra show which protons are coupled to each other, HMBC (and the related HMQC) give information about the relative relationships between protons and carbons in a structure. In an HMQC spectrum, a ^{13}C spectrum is displayed on one axis and a ^{1}H spectrum is displayed on the other axis. Cross-peaks show which proton is attached to which carbon. COSY spectra show 3-bond coupling (from H-C-C-H), whereas HMQC shows a 1-bond coupling (just C-H).

Nuclear Overhauser Effect Spectroscopy (NOESY): This technique shows through-space interactions within the molecule, rather than the through-bond interactions seen in COSY and HMBC/HMBQ. This method is especially useful for determining stereochemical relationships in a molecule. In two stereoisomers, the atoms are all connected in exactly the same order, through exactly the same bonds. A COSY or an HMBC spectrum wouldn't be able to distinguish between these isomers.

HNCA: This is an example of 3D NMR. It shows a correlation between an amide proton, the amide nitrogen to which it is attached, and the carbons that are attached to the amide nitrogen. HNCA data are viewed in slices, in which you look at one nitrogen at a time. One axis shows the shift of the proton attached to that nitrogen, and the other axis shows the shifts of the carbons attached to the nitrogen. The abbreviation comes from the pathway for transfer of the magnetization (amide H to amide N, and then the attached Cs).

NMR structures found in the Protein Data Bank show an ensemble of slightly different structures. This arises from the dynamic behavior of the proteins in water, compared to when the structure is determined from a crystal. Comparative analyses between crystal and NMR structures show that secondary structures are equally accurate, that loops in NMR structures are probably too flexible, and that loops in protein, often on the surface, are too rigid, which makes sense given the packing restraints with a crystal lattice.

3.3.7.3: Cryo-electron Microscopy

Cryogenic-electron microscopy (cryo-EM) has recently emerged as a powerful technique in structural biology that is capable of delivering high-resolution density maps of macromolecular structures. A cryo-EM and structure determined from it are shown in Figure 3.3.22

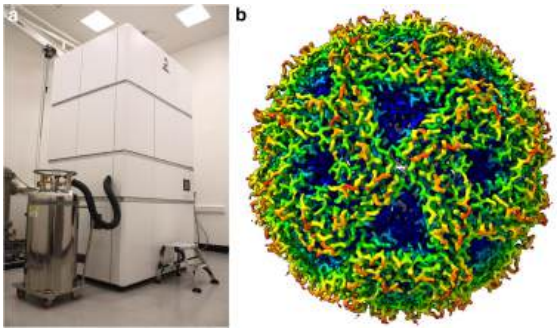


Figure 3.3.22: **Cryo-Electron Microscopy.** (a) the Scottish Centrel for Macromolecular Imaging JEOL CryoARM 300. (b) High-resolution 2.2-Å resolution structure of lumazine synthase. *Bhella, D. (2019) Biophysical Reviews 11:515-519*. Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/)

Resolutions approaching 1.5 Å are now possible and maps in the 1–4-Å range inform the construction of atomistic models with a high degree of confidence. This new capacity for investigators to determine macromolecular structures at high resolution and without the need to form crystals has led to an explosion of interest in adopting cryo-EM.

Protein suspensions are frozen on 3-mm-diameter transmission-electron microscope (TEM) support grids made from a conductive material (e.g. Cu or Au) that are coated with a carbon film with a regular array of perforations 1–2 μm in diameter. A total of 3–5 μl of sample is loaded onto the grid which is then immediately blotted with filter paper with the aim of creating a film of buffer/protein on the grid that, when frozen, will be thin enough for the electron beam to penetrate. Optimizing the ice thickness is a vital step in sample preparation as thicker layers of ice increase the probability that the incident electron will undergo multiple scattering events and thereby reduce the image quality. In the case of extreme ice thickness, the electron beam does not penetrate at all. After blotting, the grid is rapidly plunged into a bath of liquid ethane—a very effective cryogen that freezes water with sufficient rapidity to prevent the formation of ice crystals. The formation of a vitreous layer of ice is the fundamental step in cryo-EM and preserves the target in a near-native state. The resulting vitreous ice layer with suspended protein molecules must then remain close to liquid nitrogen temperature (− 196 °C) during storage and imaging in the TEM to prevent phase changes to other types of ice that are not amenable to high-quality imaging and preservation of protein structure.

Figure 3.3.23shows a summary of the process and single particle structure determination.

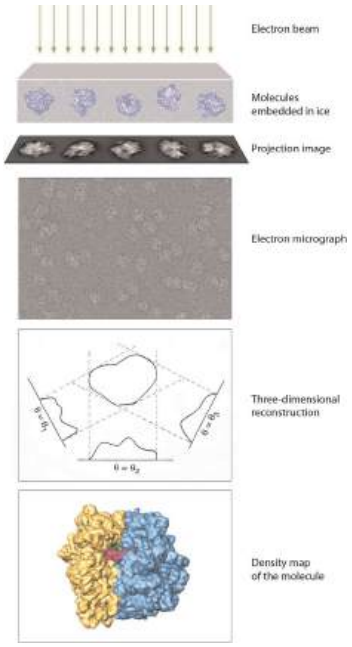


Figure 3.3.23: Principle of cryo-EM and single-particle reconstruction. Agirrezabala, X., Frank, J., 2010. From DNA to proteins via the ribosome: Structural insights into the workings of the translation machinery. Human Genomics 4, 226.. https://doi.org/10.1186/1479-7364-4-4-226. CC BY 4.0

The above figure shows cryoEM structure determination process for ribosomes. When frozen, they are found in random orientations embedded in a thin layer of ice. Exposure to a low-dose electron beam in the transmission electron microscope produces a projection image (i.e the electron micrograph). A typical electron micrograph shows E. coli ribosomes as low-contrast single particles on a noisy background. After the orientations of the particles have been determined, usually by matching them with a reference through computer algorithms, they are used to reconstruct a density map by a back-projection or a similar reconstruction algorithm. This density map is segmented into the different components (subunits, ligands), and the different components are displayed using different colors in a surface representation (bottom panel; small and large subunits are shown in yellow and blue, respectively. A- and P-site tRNAs are colored red pink and green, respectively).

Given the increasing popularity of this technique, we present Figure 3.3.24below which shows another representation with the end-point of a 3D model of higher resolution.

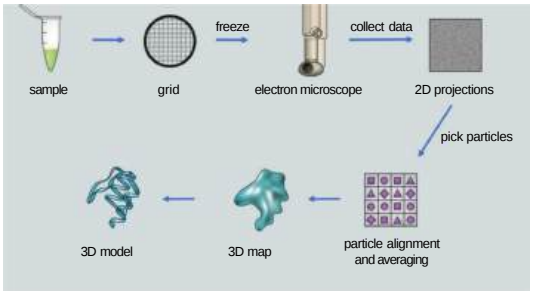


Figure 3.3.24: A schematic of the single-particle reconstruction cryoEM pipeline. Hey Tony et al., 2020Machine learning and big scientific data. *Phil. Trans. R. Soc. A.***378**2019005420190054. http://doi.org/10.1098/rsta.2019.0054. Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/

When the beam strikes the ice lays in which the single structure is found, the ices domes which cause motion. Motion pictures are taken with fast detectors and computers are used to adjust the images for the motion, which would dull and lower the resolution of the structure. In addition, as X-rays damage molecules, so can electron beams. Earlier frames in the movie show less damage. Both the motion and damage effects of the beam can now be corrected to increase the resolution.

Here is a YouTube video from PDB 101 that describes the technique.

Methods for determining atomic structure. PDB-101: Educational resources supporting molecular explorations through biology and medicine. Christine Zardecki, Shuchismita Dutta, David S. Goodsell, Robert Lowe, Maria Voigt, Stephen K. Burley. (2022) *Protein Science* **31**: 129-140 https://doi.org/10.1002/pro.4200. CC By 4.0 license.

3.3.7.4: Homology Modeling and Artificial Intelligence/Machine Learning

None of the above techniques would be required (except for validation) if the 3D structure of a protein could be determined from its sequence. Computationally this is an astoundingly large problem, given the astronomically large number of possible conformations for a given protein sequence. However, using the nearly 200,000 structures in the PDB database, modern computer methods utilizing artificial intelligence and machine learning have perhaps solved the folding problem. New protein structure prediction programs such as AlphaFold (using a neural network-based model) and RoseTTAFold have led to the solving of structures for which crystals or homologous proteins are not available. A comparison of protein structures obtained using the program with known 3D structures obtained through x-ray crystallography or other techniques are almost identical. Different metrics can be used to compared predicted structures to the actual one. The root mean squared deviation (RMSD) is a common one. For example, developers of RoseTTAFold used a **TM-score**, a metric for assessing the topological similarity of protein structures. Compared to RMDS, the TM-score weights smaller distance errors higher than larger distance errors and makes it sensitive to the global fold, not local structural differences. TM values range from 0-100 (100 is a perfect match). Scores below 17 indicate no topology match while those greater than 50 suggest a common fold.

The software is described as a "neural network, meaning it simultaneously considers patterns in protein sequences, how a protein's amino acids interact with one another, and a protein's possible three-dimensional structure. In this architecture, one-, two-, and three-dimensional information flows back and forth, allowing the network to collectively reason about the relationship between a protein's chemical parts and its folded structure". Programs of this type might allow the generation of proteins with new therapeutic or commercial potential just based on sequences. These include vaccines, sensors, specific immune system suppressors or activators, and antivirals.

AlphaFold has now been used to predict the structure of 214 million proteins from more than one million species — essentially all known protein-coding sequences. We have included many AlphaFold iCn3D models throughout this book.

Figure 3.3.25 shows the backbone tube cartoon of the x-ray pdb structure of the small protein (1xww, cyan) and the structure predicted by both RoseTTAFold program and AlphaFold (magenta) just from its primary sequence. Sulfate, a competitive inhibitor, is shown (spacefill) bound in the active site. The alignment is quite spectacular, except for the N-terminal 5 amino acids shown at the very bottom of the figure (6 oclock). This stretch has more disorder even in the x-ray structure as the amino acids have high B-factors, indicating more conformational flexibility.

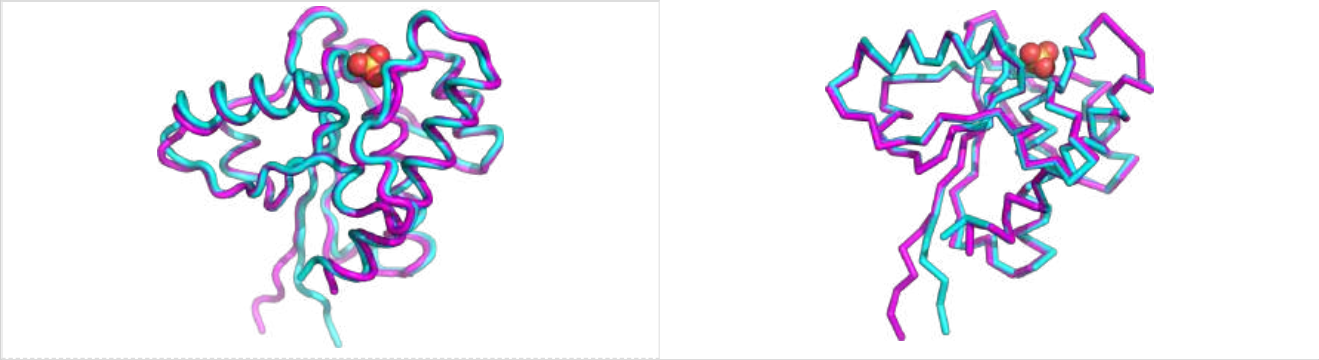


Figure 3.3.25: Comparison of the x-ray and computationally predicted structures of human low molecule weight protein tyrosine phosphatase

Left panel: X-ray structure (cyan) of low molecular weight protein tyrosine phosphatase with bound SO_4^{2-} (1xww) and corresponding structure predicted by the RoseTTAFold (magenta).

Right panel: Same structures using AlphaFold for the structural prediction.

AlphaFold has also been used to predict the structure of protein complexes in which multiples of the same or of a different protein subunit combine to form a larger, quaternary structure.

In yet another expansion of the use of machine learning and artifical intelligence, programs can now start with a desired 3D shape (protein backbone, for example) and determine the amino acid sequence necessary to get it. One such program is called ProteinMPNN. It allows protein structure design, not structure prediction. In addition, predicted symmetric protein oligomers (about 30% of all proteins in the Protein Data Bank) can also be generated using "deep network hallucination" which requires as input the number of monomers in the structure and the length of the monomer. The "hallucinated structures" are quite different from typical homooligomeric structures found in the PDB, but the predicted structures match well with cryoEM structures of the designed oligomers.

We will explore how large language modules in AI/machine learning are used to predict structure and function in macromolecules and even in the analyses of large interconnected metabolic and signaling pathways in Chapter 4.13: Predicting Structure and Function of Biomolecules Through Natural Language Processing Tools.

3.3.8: Comparisons of 3D Structure Determination Methods

Obtaining a pure, highly concentrated (mM) protein sample is a major bottleneck for both x-ray crystallography and NMR. The high concentration is required because both techniques are insensitive to single molecule analysis, and a large population of a particular protein is required to overcome the signal-to-noise barrier. On a similar note, the sample needs to be very homogenous, so protein purification is necessary at some point. Cryo-EM requires considerably less protein than the other two methods, but still 'a lot' by any standard. Typically, Cryo-EM requires preparations at a concentration of 1 mg/ml in a volume of at least 50 µl, whereas crystal formation might require 500 µl of protein at a concentration of 5 -10 mg/ml. Cryo-Em also requires the protein to be prepared in a low-salt buffer, with minimal additives, to ensure good freezing and image contrast.

Recombinant protein production using *E. coli* is the method of choice when large quantities of protein are required. This process involves taking the gene (often cDNA) of the protein of interest, splicing it into a suitable inducible vector, transforming the vector into an *E. coli* host, and growing the culture in a rich medium. The bacterial host will multiply during a growth phase, after which it is induced to express the protein of interest. If all goes well, the protein will express solubly and in high numbers. Unfortunately, this process is easier said than done. Many eukaryotic proteins do not express well in prokaryotic hosts, and oftentimes modifications need to be made to optimize the bacterial host, codon usage, media, etc. to obtain a decent yield of recombinant protein. Additionally, proteins often express insolubly as inclusion bodies and require high concentrations (2M to 8M) of denaturants such as urea or guanidine hydrochloride to solubilize them, and then stepwise dialysis into an appropriate buffer to refold them. Alternatively, eukaryotic organisms such as *S. cerevisiae* (yeast), insect, and mammalian cell lines can be used, especially when post-translation modifications are required, though a decrease in yield and increase in overall cost is common with these organisms.

The difficulty of protein production is compounded for NMR by the fact that all proteins need to be ^{15}N and/or ^{13}C labeled, as only these isotopes have nuclei with + ½ and - ½ spin states which enable the energy transitions required for a radiofrequency NMR signal; note that ^{1}H also has ½ spin states but is highly abundant.

Protein stability is an issue for both crystallography and NMR. Once a protein has been expressed, purified, and concentrated, it must maintain its structural integrity for the duration of the experiments. For crystallography, this involves the crystallization process, where the protein sample is placed in a variety of solutions (most often involving high concentrations of polyethylene glycol) that induce crystallization. Often referred to as a voodoo technique, crystallization conditions are tested in a high throughput method using 96-well screening plates, and any hits are further optimized using a larger volume of the particular solution. While a crystallization condition may eventually be found, the process can take anywhere from a few days to even a year or two to happen, making the crystallization process the rate-limiting step for protein crystallographers. During this time, the protein must stay in solution and maintain its structure so as to produce a high-quality crystal; a condition that is not often the case.

Similarly, a stable, highly concentrated protein sample is required to perform many of the more advanced NMR experiments. This is because many of these experiments require days and even weeks to run, during which the homogeneity of the solution is key to acquiring quality spectra. Should the protein unfold or precipitate out of solution during an experiment, the resulting chemical change would either not produce any signal, or one which could not be used for structure/dynamics determination.

For Cryo-EM, working with frozen-hydrated specimens brings a number of challenges both for manipulations and imaging. When handling cryo-EM grids to load them into the microscope, exposure to atmospheric water vapor rapidly leads to frost buildup on the grid. Under the TEM, these ice crystals on the grid surface appear as huge boulders that completely block the electron beam. Thus, grids are kept under liquid nitrogen as much as possible to minimize frost contamination. Problems with ice conditions are common—insufficient rapid freezing leads to the formation of hexagonal ice, while devitrification occurs when samples warm up, leading to the formation of cubic ice. Various degrees of contamination may occur, and frosting at atmospheric pressure causes the above-mentioned ice crystal deposition, while contamination within the column or under low-vacuum conditions gives rise to a more subtle artifact.

One of the hallmarks of protein crystallography is that size does not matter. Whether one is working with a 25 kDa monomeric protein, or a 900 kDa multimeric complex, if it can be crystallized and produce a high-resolution diffraction pattern its structure can be determined. This is due to the fact that once in crystal form, a protein is in a more-or-less static conformation which, after passing it through the x-ray beam at different angles, can produce a single structural model. Cryo-EM is similar in this regard. Very large structures, including massive nucleoprotein complexes, such as the ribosome, can be elucidated using Cryo-EM. The same cannot be said for NMR.

In NMR, the protein is in a soluble state and therefore in constant movement. The most important movement that governs the spectral quality is that of the molecular tumbling rate. For proteins larger than about 40 kDa, the tumbling rate decreases significantly, in turn increasing the transverse relaxation rate (T_2). Essentially, this results in a weaker and rapidly decaying NMR signal, which manifests itself in peak broadening and spectral overlap.

One of the major advantages of NMR is its ability to record small and large-scale protein dynamics, a phenomenon that is generally suppressed when a protein is crystallized. Although a crystallized protein may exhibit a certain amount of motion within the lattice, the motions manifest themselves as static or dynamic disorder, the former of which may result in two different conformations of a particular region, and the latter in averaged electron density. In general, crystallization may restrict a protein's natural flexibility and motions. Cryo-EM suffers from this same limitation as samples are frozen and immobile. However, cryo-EM is capable of capturing a snapshot of the native structure as freezing is instantaneous and does not require the formation of a crystal lattice.

Crystallography, however, is not left in the cold when it comes to dynamic structural analysis. Time-resolved crystallography can be used to monitor changes in the protein structure upon the addition of some ligand, or change in the environment. Because all protein crystals are highly hydrated, they are able to serve as crucibles for some biochemical reactions. The crystal is typically soaked in a solution containing the ligand of interest to initiate the biochemical reaction, after which the crystal is quickly placed into the beam line and the diffraction pattern is obtained. This can be performed multiple times if necessary to obtain a variety of structural intermediates. The process though requires many things to go right: the protein cannot become disordered nor should the crystal become cracked during the soaking process, and a high-powered synchrotron is required to collect high-quality diffraction data over short exposure times.

In the end, protein X-ray crystallography, cryo-EM, and NMR spectroscopy are not mutually exclusive techniques; one can easily pick up where the other falls short. In analyzing NMR dynamics experiments, for example, one can greatly benefit from existing crystal structure data, or cryo-EM data onto which the NMR structural data can be superimposed. Similarly, NMR structure data can be used to supplement a cryo-EM or crystal structure with more information on the protein's dynamics, binding information, and conformational changes in solution.

Recent Updates 8/4/23

3.3.9: Molecular Dynamics Simulations

Introduction to Molecular Mechanics and Molecular Dynamics

Molecular modeling and computational chemistry are important parts of modern biochemistry. Modeling is important to display in a meaningful and instructive fashion the large amounts of data produced when X-ray crystallography and NMR are used to determine the structure of large biological molecules and complexes. Remember, however, that primary X-ray crystal data (in the form of electron density maps) are just that, and the data must be interpreted like any other type of data. Structures need to be refined and energy minimized to produce more realistic structures (without van der Waals overlap or missing atoms, for example). In addition, atoms within any molecule are not static, but move as bonds vibrate, angles bend, etc. This implies that large biomolecules could adopt many possible conformations of different energies. For proteins, some of these conformations might center around an average conformation situated at a local or global energy minimum separated from each other by activation energy barriers.

In contrast to small molecules whose structure can be minimized using ab initio or semi-empirical quantum mechanics (using programs such as Spartan), large molecular structures (like DNA, RNA, proteins, and their complexes) must be minimized using molecular mechanics, based on Newton's laws. Atoms are treated as masses, and bonds as springs with appropriate force constants. A force field, containing all the relevant parameters for a given atom (for example sp^3, sp^2, sp^2 aromatic, and sp C) and bond types is used to solve energy equations that sum all energies over all atoms and bonds in the molecule. These energies include interactions among bonded atoms (stretching, bending, torsion, wagging) and those among nonbonded atoms (electrostatic and van der Waals). For minimization calculations, the positions of the atoms within a molecule must be systematically or randomly moved and the energy recalculated with the goal of finding a lower energy and hence more stable molecule. Minimization calculations can not probe all conformational space and can not easily move a structure from a local minimum to a global minimum if two are separated by a large energy barrier. Energy minimizations are usually done in the absence of solvent. Common force fields used for macromolecules are CHARMM, AMBER, and GROMOS. Parameters for specific atom type in a given bond include atomic mass, van der Waals radius, partial charge for atoms (from quantum mechanics) and bond length (from electron diffraction data), angles, and force constants for bonds (modeled as springs, obtained from IR). These parameters are derived from experiments and theoretical (usually quantum mechanical) calculations on small organic molecules. A potential energy equation comprised of terms from bond stretching, angle bending, and torsion angle changes (bonded interactions) as well as electrostatic and van der Waals interactions (nonbonded) is then solved (described below).

The goal of molecular dynamics is to simulate the actual changes in a molecule as a function of time after an energy input (heat application at a higher temperature) is added to a molecule at equilibrium. To make the simulation realistic, the structure is placed in a "bath" of thousands of water molecules. As is described below, if the energies of atoms in a large molecule are known, the forces acting on those atoms can be deduced. From Newton's Second Law (F=ma), the velocity or change of position of an atom in the structure with time can be determined. If the dynamic simulation can be run for a long enough period of time, alternate conformations (perhaps those centered around a global minimum as well as those nearby in energy space - a local minimum) may be sampled. By determining what fraction of the simulated conformations resemble the two alternative conformations, the ΔG for the interconversion of the two states can be calculated. As you can imagine, these calculations require large amounts of computer time. They give very important information, however, since protein conformational changes are often, if not always associated with the binding of a biological molecule to a binding partner. *In silico* experiments offer important clues and support to results obtained using other methods of study.

Molecular mechanics (MM) and molecular dynamics (MD) have become powerful tools in analyzing and predicting the properties of complex biological structures. The Noble Prize in Chemistry in 2013 was awarded to Martin Karplus, Michael Levitt, and Arieh Warshel "for the development of multiscale models for complex chemical systems". Karplus in particular developed much of the present basis for MD simulations.

Energy (E), Force (F), and Motion

To make the energy equations for the individual components more understandable, it is useful to consider the relationship between force and energy. You have studied two general force equations in introductory chemistry and physics courses. One is Coulomb's Law, which describes the electrostatic force of attraction, F_C, between two charges, q_1 and q_2, separated by a distance r.

$$F_C = k\frac{q_1 q_2}{r^2} \tag{3.3.8}$$

The other is Hooke's Law, which describes the restorative force on a mass connected to a spring on stretching or compression of the spring.

$$F = -kx \tag{3.3.9}$$

where x is the displacement of the spring from an equilibrium (at rest) position.

Our first interest is to understand how these equations might lead us to equations that describe the potential energy of a two-charge system or of a compressed or stretched spring. We can best understand this by studying the simple example of a ball placed at various locations on a hill. If placed on a flat surface at the top and bottom of the hill, there is no net force on the ball (F_{net} = 0), so it will not move. If placed at various locations on the down slope, it will experience a net downward force, shown in a qualitative fashion in Figure 3.3.26 below.

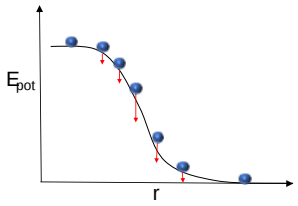


Figure 3.3.26: Potential Energy vs r for a ball on a hill

Astute observers will note that the magnitude of the force vector is proportional to the slope. From this simplistic approach, we come to the following equation relating F to E:

$$F = -\frac{dE}{dr} \tag{3.3.10}$$

The minus sign is required since the force is downward but the energy increases upward.

This simplified approach can be extended into three dimensions, to give the following equation (which will have meaning to those with advance calculus background) where F is the negative gradient of the potential energy:

$$F = -\left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y} + \frac{\partial}{\partial z}\right) E = -\nabla E \tag{3.3.11}$$

Applying the 1D equation to Hooke's Law gives

$$\begin{aligned} dE &= -F dr = -kx dx \\ \int dE &= -k \int x dx \end{aligned} \tag{3.3.12}$$

which gives

$$E = \frac{kx^2}{2} \tag{3.3.13}$$

This gives a parabolic graph of E vs displacement. Figure 3.3.27 below shows an interactive PHET simulation of Hooke's Law. Click Energy and then select energy to see the parabolic plot. Change the force constant to alter the "steepness" of the resulting parabolic curve.

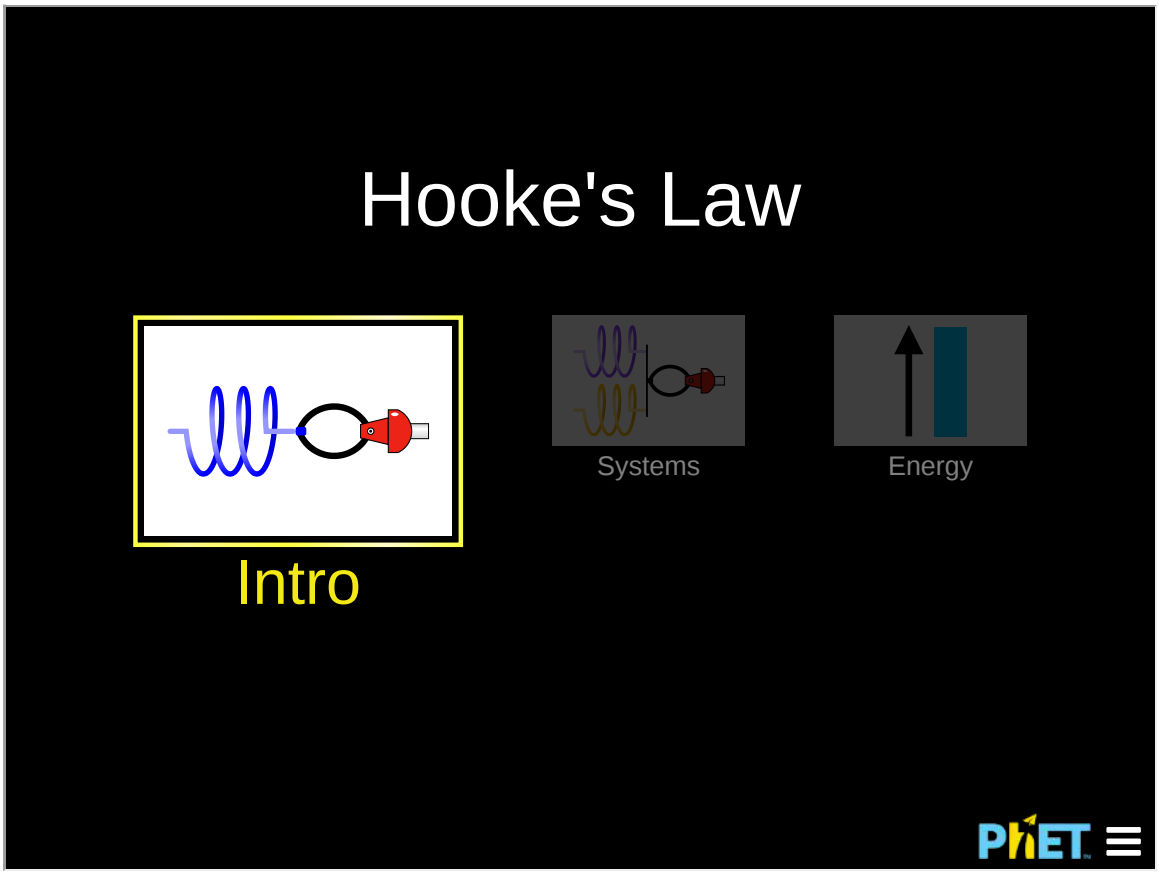


Figure 3.3.27: PHET simulation of Hooke's Law. https://phet.colorado.edu/en/simulations/hookes-law

The same approach can be applied to Coulomb's Law. Notice that the result equation for E results in increasingly negative values as r get smaller only when q1 and q2 have opposite charges.

$$dE = -F\,dr = -k\frac{q_1 q_2}{r^2}dr$$
$$\int dE = -kq_1q_2\int r^{-2}dr \tag{3.3.14}$$

gives

$$E = \frac{kq_1q_2}{r} \tag{3.3.15}$$

A graph of E vs r for both attractive and repulsive interactions is shown in Figure 3.3.28 below.

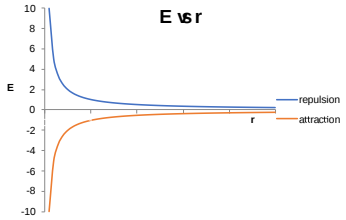


Figure 3.3.28: E vs r for both attractive and repulsive interactions

Molecular Mechanics

Note: The following review is based on an NIH Guide to Molecular Modeling (1996), which to the best of our knowledge was removed from the web.

Molecular mechanics uses Newtonian mechanics to calculate energy of atoms in large molecules like proteins. It assumes that nuclei and electrons are one particle with radii and calculated charges. Bonds are treated as springs connecting atoms. Energies are calculated classically (not with quantum mechanics). Parameters, many based on quantum mechanics calculations on small molecules, are assigned to all bonds, angles, dihedrals, etc. Interactions are bonded (local) and nonbonded.

Bonded interactions involved atoms connected by one bond (bond stretch), two bonds (angle bending) and 3 bonds (dihedral angle change). These three types of bonded interactions are shown with black arrows on the right-hand side of Figure 3.3.29 below.

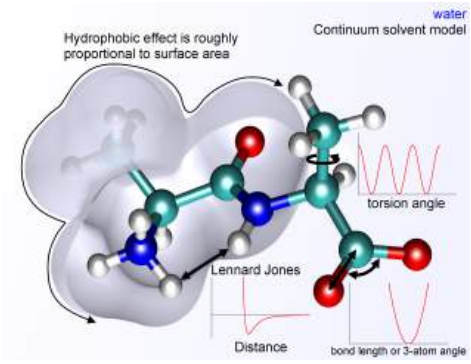


Figure 3.3.29: Bonded and Nonbonded Interactions in Proteins. Force field (chemistry). (2023, July 15). In *Wikipedia*. https://en.wikipedia.org/wiki/Force_field_(chemistry). CC BY-SA 3.0

Non-bonded atoms (greater than two bonds apart) interact through induced dipole-induced dipole interactions (one example of which is steric repulsions) and electrostatic attraction/repulsion. And example of a nonbonded interaction is shown in the double black arrow labeled Lennard Jones in the above figure.

All energy terms from these interactions are summed to give the energy of a given conformation. The energy should be considered relative to those of other conformations. Here is the basic energy equation for all of these energy terms:

$$\text{Energy (E)} = \text{E}_{\text{Stretch}} + \text{E}_{\text{Bending}} + \text{E}_{\text{Torsion}} + \text{E}_{\text{Non-bonded Interactions}} \tag{3.3.16}$$

The "force field" consists of the energy equations and the parameters for each of the energy terms. There are many different commercially available force fields.

Bonded Interaction Energies

The mathematical form of the energy terms varies from force-field to force-field. The more common forms will be described.

Stretching (Vibrational) Energy

$$\text{E}_{\text{stretch}} = \Sigma_{\text{bonds}}\,\text{k}_\text{b}\left(r - \text{r}_\text{o}\right)^2 \tag{3.3.17}$$

Figure 3.3.30 illustrates bond stretching or vibration

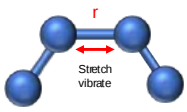


Figure 3.3.30: Bond stretching or vibration

The stretching energy equation is based on Hook's law. The kb parameter defines the stiffness of the bond spring. R_0 is the equilibrium distance between the two atoms. It should make sense that deviations from the equilibrium length would be associated with higher energy. The E vs r curve is hence a parabola as shown in Figure 3.3.3 below for a system when the lowest energy is at r=6.

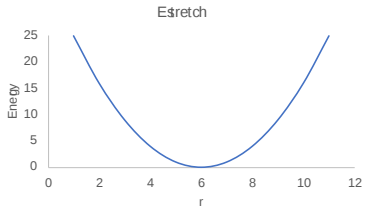


Figure 3.3.31: Energy vs r (r_0=6) for bond stretching (vibration)

Obviously, only small changes in r are allowed as too large an r value would lead to bond breaking.

Bending Energy

$$\mathrm{E_{bending}} = \Sigma_{\mathrm{angles}}\, k_\Theta\,(\Theta - \Theta_0)^2 \tag{3.3.18}$$

Figure 3.3.32 illustrates bond bending

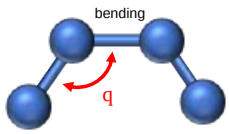


Figure 3.3.32: Bond bending

The bending energy equation is also based on Hooke's law. The kΘ parameter controls the stiffness of the angle spring, while the Θ_0 is the equilibrium angle. As above, the graph of E vs theta is expected to be a parabola as shown in Figure 3.3.33 below for a system when the lowest energy is at θ=45

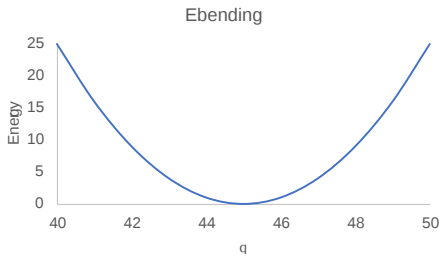


Figure 3.3.33: $E_{bending}$ vs r when lowest energy is at θ=45.

Torsion Energy

$$E_{\mathrm{torsion}} = \Sigma_{\mathrm{torsions}} A[1 + \cos(\,\mathrm{ntau}\, - \Theta)] \tag{3.3.19}$$

Figure 3.3.34 illustrates torsion angle rotation

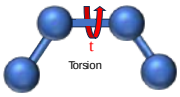


Figure 3.3.34: Torsion angle rotation

The torsion energy is modeled by a periodic function, much as you have seen with energy plots associated with Newman projections sighting down the central C-C bond in butane, for example. Two different torsion energy functions are shown in Figure 3.3.35 below.

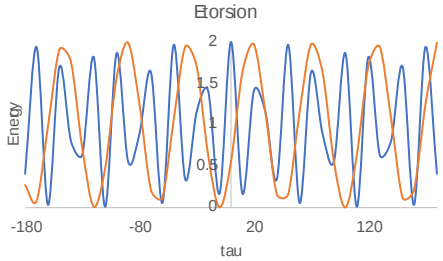


Figure 3.3.35: Torsion energy vs tau for A=1, n=1, phi=0 (blue) and A=1, n=2,phi=90 (orange).

Non-Bonded Interaction Energy

The non-bonded energy is calculated for all possible pairs of nonbonded atoms, i and j:

$E_{nonbonding} = \Sigma_i \Sigma_j [-B_{ij}/r_{ij}^6 + A_{ij}/r_{ij}^{12}] + \Sigma_i \Sigma_j (q_i q_j) / r_{ij}$

$$E_{\text{nonbonding}} = \Sigma_i \Sigma_j \left[-B_{ij}/r_{ij}^6 + A_{ij}/r_{ij}^{12} \right] + \Sigma_i \Sigma_j \left(q_i q_j \right) / r_{ij} | \tag{3.3.20}$$

The first term represents van der Waals interactions while the second term represents Coloumbic electrostatic interactions. Figure 3.3.36 illustrated nonbonded interactions.

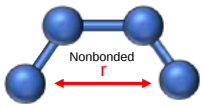


Figure 3.3.36: Nonbonded interactions

You should remember that induced dipole-induced dipole interactions are short-range and occur among all atoms. The 6-12 energy equation based on the Lennard-Jones' potential, shows a negative (attractive) term proportional to $-1/r^6$ and a repulsive term proportional to $+1/r^{12}$. Figure 3.3.37 below shows a graph of the attractive, repulsive, and summation of the energy terms in the Lennard Jones potential.

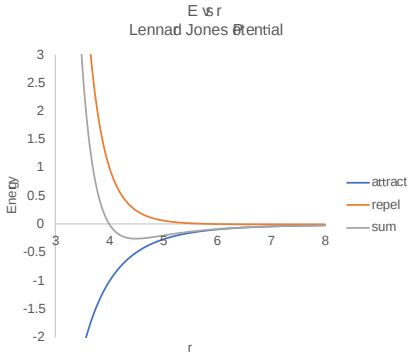


Figure 3.3.37: Lennard Jones potential vs r

The A and B parameters control the depth and position (interatomic distance) of the potential energy well for a given pair of non-bonded interacting atoms (e.g. C:C, O:C, etc.). In effect, A determins the degree of stickiness of the van der Waals attraction, and B determines the degree of hardness of the atoms (e.g. marshmallow-like, billiard ball-like, etc.).

The B parameter is related to the "stickiness" of the interactions and is related to the polarization of the atoms. B can be obtained from atomic polarizability measurements, or it can be calculated quantum mechanically. The A parameter is empirically derived to fit nonbonded contacts between atoms in crystal structures.

Summary Interactions

Some programs assign charges using rules or templates, especially for macromolecules. In some force fields, the torsional potential is calibrated to a particular charge calculation method (rarely made known to the user). The use of a different method can invalidate the force-field consistency. Sometimes, an additional bonded interaction term, improper dihedrals, is added as illustrated below. The potential for that is given by the following equation:

$E_{improper} = \Sigma_{angles}\, k\omega\, (\omega - \omega_0)^2$

$$E_{\text{improper}} = \Sigma \text{ angles } k\omega(\omega - \omega_0)^2 \tag{3.3.21}$$

Molecular Dynamics

This section comes directly from the NIH tutorial:

"In the broadest sense, MD is concerned with molecular motion. Motion is inherent to all chemical processes. Simple vibrations, like bond stretching, and angle bending, give rise to IR spectra. Chemical reactions, hormone-receptor binding, and other complex processes are associated with many kinds of intra- and intermolecular motions.

The driving force for chemical processes is described by thermodynamics. The mechanism by which chemical processes occurs is described by kinetics. Thermodynamics dictates the energetic relationships between different chemical states, whereas the sequence or rate of events that occur as molecules transform between their various possible states is described by kinetics.

Conformational transitions and local vibrations are the usual subjects of molecular dynamics studies. MD alters the intermolecular degrees of freedom in a step-wise fashion, analogous to energy minimization. The individual steps in energy minimization are merely directed at establishing a downhill direction to a minimum. The steps in MD, on the other hand, meaningfully represent the changes in atomic position, ri, over time (i.e. velocity).

Newton's equation ($F_i = m_i a^i$) is used in the MD formalism to simulate atomic motion. The rate and direction of motion (velocity) are governed by the forces that the atoms of the system exert on each other as described by Newton's equation. In practice, the atoms are assigned initial velocities that conform to the total kinetic energy of the system, which in turn, is dictated by the desired simulation temperature. This is carried out by slowly heating the system (initially at absolute zero) and then allowing the energy to equilibrate among the constituent atoms. The basic ingredients of MD are the calculation of the force on each atom, and from that information, the position of each atom through a s specified period of time (typically on the order of picoseconds = 10-12 seconds).

The force on an atom can be calculated from the change in energy between its current position and its position a small distance away. This can be recognized as the derivative of the energy with respect to the change in the atom's position: $-dE/dr_i = F_i$.

Energies can be calculated using either MM or quantum mechanics methods. MM energies are limited to applications that do not involve drastic changes in electronic structure such as bond-making/breaking. Quantum mechanical energies can be used to study dynamic processes involving chemical changes. The latter technique is extremely novel and of limited availability.

Knowledge of the atomic forces and masses can then be used to solve for the positions of each atom along a series of extremely small time steps (on the order of femtoseconds). The resulting series of snapshots of structural changes over time is called a trajectory. The use of this method to compute trajectories can be more easily seen when Newton's equation is expressed in the following form

$-dE/dr_i = m_i a = md^2r_i/dt^2$

$$-dE/dr_i = m_i a = md^2r_i/dt^2 \tag{3.3.22}$$

In practice, trajectories are not directly obtained from Newton's equation due to the lack of an analytical solution. First, the atomic accelerations are computed from the forces and masses. The velocities are next calculated from the accelerations based on the following relationship:
$a_i = dv_i/dt$. Lastly, the positions are calculated from the velocities: $v_i = dr_i/dt$. A trajectory between two states can be subdivided into a series of sub-states separated by a small time step, delta t (e.g. 1 fs).

The initial atomic positions at time t are used to predict the atomic positions at time t = delta t. The positions at t = delta t are used to predict the positions at t = 2Δt, and so on.

The leapfrog method is a common numerical approach to calculating trajectories based on Newton's equation. The method derives its name from the fact that the velocity and position information successively alternate at ½ time step intervals. MD has no defined point of termination other than the amount of time that can be practically covered.

Here is a molecular dynamics simulation of a small protein module (NTL9) folding.

MD simulations can be used to obtain theoretical values for ΔG and Keq values for conformational changes, binding of small ligands, and changes in protonation states for side chains. This process is based on the idea that the conformations sampled in silico MD simulations reflect those found in vitro (i.e. they are part of the thermodynamically expected and available conformations available to the molecules during normal conformational shifts). This is called the Ergodic Hypothesis. Given the short time spans for MD simulations (limited by computer power) this hypothesis can't apply to the dynamic results unless the sample conformations are close in energy without a large activation energy barrier between them. If it is then the following equation could apply:

$$\begin{aligned}\Delta G^0 &= -RT\ln K_{\text{eq}} \\ \Delta G^0 &= -RT\ln P_2/P_1 = -RT\ln f_2/f_1\end{aligned} \tag{3.3.23}$$

where

P_n is the probability of being in a given state and f_n is the fraction in a given state.

You will note that none of the potential energy functions use quantum mechanical parameters. This is due, in part, to the complexity of the systems studied. This is beginning to change as more effort is devoted to understanding quantum mechanical aspects of complex bonded systems. New advances in even simple systems can illustrate this point. Take for example ethane. The conformational analysis of this simple molecule is discussed in all organic chemistry books. Plots of energy vs dihedral angle (viewing the molecule down the C-C bond and measuring the angle between C-H bonds on adjacent C atoms) oscillates every 120°. The E is at a maximum when the dihedral angle is 0, 120, 240 and 360°, each representing the eclipsed conformation. It reaches minimums at the staggered (gauche) conformations at 60, 180, and 270°. Why is the eclipsed form higher in energy than the staggered form? All organic books would state that there is greater steric repulsion (of the electron clouds) in the eclipsed forms, which raises their energy compared to the staggered forms? However, Pophristic shows that to be incorrect. For the correct answer, you must turn to quantum mechanics and the phenomena of hyperconjugation. The staggered conformation is energetically favored not since it is less sterically restricted, but since its is a lower energy form due to resonance like stabilization of the σ CH molecular orbitals. There is greater correct phase overlap of σ CH and s* CH molecular orbitals on the adjacent Cs when they are in the staggered conformation than in the eclipsed form.

3.3.10: Proteome Analysis

The proteome is the entire set of proteins that are produced or modified by an organism or system. Proteomics has enabled the identification of ever-increasing numbers of protein. This varies with time and distinct requirements, or stresses, that a cell or organism undergoes. Proteomics is an interdisciplinary domain that has benefited greatly from the genetic information of various genome projects, including the Human Genome Project. It covers the exploration of proteomes from the overall level of protein composition, structure, and activity. It is an important component of functional genomics.

After genomics and transcriptomics, proteomics is the next step in the study of biological systems. It is more complicated than genomics because an organism's genome is more or less constant, whereas proteomes differ from cell to cell and from time to time. Distinct genes are expressed in different cell types, which means that even the basic set of proteins that are produced in a cell needs to be identified.

In the past, this phenomenon was assessed by RNA analysis, but it was found to lack correlation with protein content. Now it is known that mRNA is not always translated into protein, and the amount of protein produced for a given amount of mRNA depends on the gene it is transcribed from and on the current physiological state of the cell. Proteomics confirms the presence of the protein and provides a direct measure of the quantity present.

A cell may make different sets of proteins at different times or under different conditions, for example during development, cellular differentiation, cell cycle, or carcinogenesis. Further increasing proteome complexity, as mentioned, most proteins are able to undergo a wide range of post-translational modifications.

Therefore, a proteomics study may become complex very quickly, even if the topic of study is restricted. In more ambitious settings, such as when a biomarker for a specific cancer subtype is sought, the proteomics scientist might elect to study multiple blood serum samples from multiple cancer patients to minimize confounding factors and account for experimental noise. Furthermore, many proteins undergo post-translational modifications such as phosphorylation. Many of these post-translational modifications are critical to the protein's function. Thus, complicated experimental designs are sometimes necessary to account for the dynamic complexity of the proteome.

3.3.11: References

Molnar, C. and Gair, J. (2013) Antibodies. Chapter in Concepts in Biology, Published by B.C. Open Textbook Project. Available at: https://opentextbc.ca/biology/chapter/23-3-antibodies/

The Human Atlas Project. (2019) Methods. Available at: https://www.proteinatlas.org/learn/method

Uhlén M et al, 2015. **Tissue-based map of the human proteome.** *Science*
PubMed: 25613900 DOI: 10.1126/science.1260419

Thul PJ et al, 2017. **A subcellular map of the human proteome.** *Science.*
PubMed: 28495876 DOI: 10.1126/science.aal3321

Uhlen M et al, 2017. **A pathology atlas of the human cancer transcriptome.** *Science.*
PubMed: 28818916 DOI: 10.1126/science.aan2507

Ahern, K. and Rajagopal, I. (2019) Biochemistry Free and Easy. Published by Libretexts. Available at: https://bio.libretexts.org/Bookshelves/Biochemistry/Book%3A_Biochemistry_Free_and_Easy_(Ahern_and_Rajagopal)/09%3A_Techniques/9.04%3A_Gel_Exclusion_Chromatography .

Magdeldin, S. (2012) Gel Electrophoresis - Principles and Basics. Published by InTech under Creative Commons Attribution 3.0. Available at: https://pdfs.semanticscholar.org/4b93/70ac3946cec6e12c369679c4178a5ef38e61.pdf

Structural Biochemistry/Proteins/X-ray Crystallography. (2018, November 19). *Wikibooks, The Free Textbook Project*. Retrieved 15:40, August 17, 2019 from en.wikibooks.org/w/index.php?title=Structural_Biochemistry/Proteins/X-ray_Crystallography&oldid=3488057.

UCD: Biophysics 200A (2019) "NMR Spectroscopy vs X-ray Crystallography", Chapter published in Current Techniques in Biophysics. Published by Libretexts and available at: https://phys.libretexts.org/Courses/University_of_California_Davis/UCD%3A_Biophysics_200A_-_Current_Techniques_in_Biophysics/NMR_Spectroscopy_vs._X-ray_Crystallography

Wikipedia contributors. (2019, June 27). Protein purification. In *Wikipedia, The Free Encyclopedia*. Retrieved 23:32, July 28, 2019, from en.Wikipedia.org/w/index.php?title=Protein_purification&oldid=903657925

Wikipedia contributors. (2019, February 15). Fast protein liquid chromatography. In *Wikipedia, The Free Encyclopedia*. Retrieved 17:14, August 15, 2019, from en.Wikipedia.org/w/index.php?title=Fast_protein_liquid_chromatography&oldid=883530035

Wikipedia contributors. (2019, July 9). Protein mass spectrometry. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:27, August 16, 2019, from en.Wikipedia.org/w/index.php?title=Protein_mass_spectrometry&oldid=905547289

Wikipedia contributors. (2019, July 8). Peptide synthesis. In *Wikipedia, The Free Encyclopedia*. Retrieved 06:13, August 17, 2019, from en.Wikipedia.org/w/index.php?title=Peptide_synthesis&oldid=905401648

Mass Spectral Analysis

3.4: Protein Purification

3.4.1: Introduction

Before a protein or other biological macromolecule can be rigorously studied from a structural and functional basis, it must be purified. The problems that can arise during protein purification become clear when one considers that a single protein has to be purified from a mixture of as many 10,000 other cellular or tissue proteins, each of which is made up of the same constituent amino acids. Proteins differ in size (how many amino acids), charge (how many positively and negatively charged amino acids), sequence, and presence of specific binding sites on the proteins. Any technique that could be used to purify protein must be based on these inherent differences. Once the protein is purified, it must be analyzed, typically by a spectral or electrophoretic technique.

Protein purification is a series of processes intended to isolate and purify a single protein or complex from cells, tissues, or whole organisms. Protein purification is vital for the characterization of the function, structure, and interactions of the protein of interest. Separation steps usually exploit differences in protein size, physical-chemical properties, binding affinity, and biological activity.

Protein purification is either *preparative* or *analytical*. **Preparative purifications** aim to produce a relatively large quantity of purified proteins for subsequent use. Examples include the preparation of commercial products such as enzymes (e.g. lactase), nutritional proteins (e.g. soy protein isolate), and certain biopharmaceuticals (e.g. insulin). Many steps and much quality control is required to remove other host proteins and other biomolecules, which pose a potential threat to the patient's health. **Analytical purification** produces a relatively small amount of a protein for a variety of research or analytical purposes, including identification, structural characterization, and studies of the protein's structure, post-translational modifications, and function.

The choice of a starting material is key to the design of a purification process. In plants or animals, a particular protein usually isn't distributed homogeneously throughout the body; different organs or tissues have higher or lower concentrations of the protein. The use of tissues or organs with the highest concentration decreases the volumes needed to produce a given amount of purified protein. If the protein is present in low abundance, or if it has a high value, scientists may use recombinant DNA technology to develop cells that will produce large quantities of the desired protein. These techniques will be discussed in greater detail in Chapter 5.

3.4.2: Sample Processing

If the protein of interest is not secreted by the organism into the surrounding solution, the first step of each purification process is the disruption of the cells containing the protein. Depending on how fragile the protein is, one of several techniques could be used including repeated freezing and thawing, sonication, homogenization by high pressure (French press), homogenization by grinding (bead mill), and permeabilization by detergents (e.g. Triton X-100) and/or enzymes (e.g. lysozyme). Finally, the cell debris can be removed by centrifugation so that the proteins and other soluble compounds remain in the supernatant.

Also proteases are released during cell lysis, which will start digesting the proteins in the solution. As the protein of interest may be sensitive to proteolysis, it is important to proceed quickly and conduct many steps at low temperatures to reduce unwanted proteolysis. Alternatively, one or more protease inhibitors can be added to the lysis buffer immediately before cell disruption. Sometimes it is also necessary to add DNase in order to reduce the viscosity of the cell lysate caused by a high DNA content.

3.4.3: Centrifugation

Centrifugation is a process that uses centrifugal force to separate mixtures of particles of varying masses or densities suspended in a liquid. When a vessel (typically a tube or bottle) containing a mixture of proteins or other particulate matter, such as bacterial cells, is rotated at high speeds, the inertia of each particle yields a force in the direction of the particle's velocity that is proportional to its mass. The tendency of a given particle to move through the liquid because of this force is offset by the resistance the liquid exerts on the particle. The net effect of "spinning" the sample in a centrifuge is that massive, small, and dense particles move outward faster than less massive particles or particles with more "drag" in the liquid. When suspensions of particles are "spun" in a centrifuge, a "pellet" may form at the bottom of the vessel that is enriched for the most massive particles with low drag in the liquid.

Non-compacted particles remain mostly in the liquid called "supernatant" and can be removed from the vessel thereby separating the supernatant from the pellet. The rate of centrifugation is determined by the angular acceleration applied to the sample, typically measured in comparison to the *g*. If samples are centrifuged long enough, the particles in the vessel will reach equilibrium wherein the particles accumulate specifically at a point in the vessel where their buoyant density is balanced with centrifugal force. Such an "equilibrium" centrifugation can allow extensive purification of a given particle.

In ***sucrose gradient centrifugation***, a linear concentration gradient of sugar (typically sucrose, glycerol, or a silica-based density gradient media, like Percoll) is generated in a tube such that the highest concentration is on the bottom and the lowest on top. Percoll is a trademark owned by GE Healthcare companies. A protein sample is then layered on top of the gradient and spun at high speeds in an ultracentrifuge. This causes heavy macromolecules to migrate toward the bottom of the tube faster than lighter material. During centrifugation in the absence of sucrose, as particles move farther and farther from the center of rotation, they experience greater centrifugal forces (the further they move, the faster they move). However, the useful separation range within the vessel is restricted to a small observable window. A properly designed sucrose gradient will counteract the increasing centrifugal force so the particles move in close proportion to the time they have been in the centrifugal field. After separating the protein/particles, the gradient is then fractionated and collected. These are described in Figure 3.4.1.

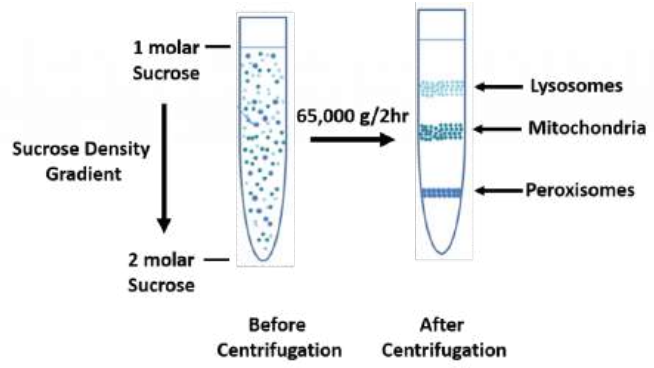


Figure 3.4.1: Sucrose Density Gradient. *Image derived from Michel Awkal*

3.4.4: Precipitation and Differential Solubilization

In bulk protein purification, a common first step to isolate proteins is precipitation using a salt such as ammonium sulfate $(NH_4)_2SO_4$. Ammonium sulfate is often used as it is highly soluble in water, has relative freedom from temperature effects, and typically is not harmful to most proteins. Proteins are precipitated by $(NH_4)_2SO_4$ in their native state, which is important if you need the protein for structure/function studies. Furthermore, ammonium sulfate can be removed by dialysis as described in Figure 3.4.2.

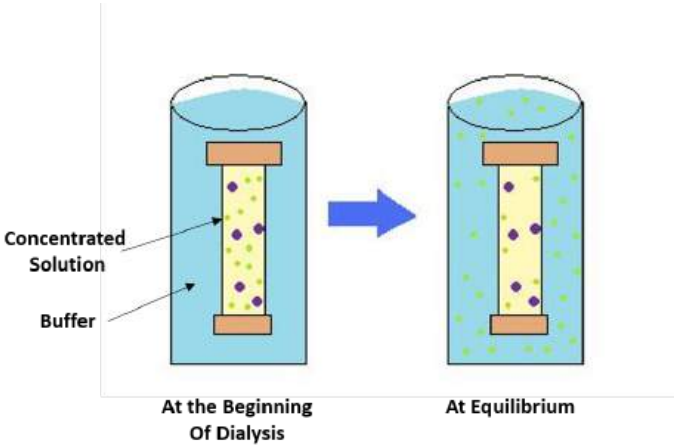


Figure 3.4.2: Dialysis of a macromolecule (blue dots). *Image adapted from Gjk003*

The process of dialysis separates dissolved molecules by their size. The biological sample is placed inside a closed membrane, where the protein of interest is too large to pass through the pores of the membrane, but through which smaller ions can easily pass. As the solution comes to equilibrium, the ions become evenly distributed throughout the entire solution, while the protein remains concentrated in the membrane. This reduces the overall salt concentration of the suspension.

The mechanism underlying salt precipitation is actually quite complicated. High concentrations of sodium chloride don't precipitate protein. Other salts like guanidinidum chloride unfold proteins and do not lead to precipitation. Salt ions interact with both the protein and solvent water in somewhat complicated ways (which we will explore later). For now, we will simply be satisfied with the empirical observation that ammonium sulfate is the salt of choice to precipitate and also concentrate proteins from a solution.

One advantage of $(NH_4)_2SO_4$ precipitation of protein from solution is that it can be performed inexpensively with very large volumes, so it is used early in many purification proteins. Different proteins precipitate at different $(NH_4)_2SO_4$ concentrations, so differential precipitation is often used. $(NH_4)_2SO_4$ concentrations are increased in a step-wise fashion until the protein of interest is precipitated.

Some proteins are not soluble in water. These include transmembrane proteins that span cell membranes and large fibrous proteins. Membrane proteins can be solubilized by the addition of detergents like sodium dodecyl sulfate (SDS), which unfolds the proteins, and octylglucoside or Triton X-100, which keeps the protein structure intact.

3.4.5: Chromatography

Chromatography is used in almost all protein purification methods and is the key that allows the separation of a given protein from the 1000s of different proteins in cells and tissues. The separation of proteins on a chromatography column depends on the type of column and chemical/physical properties of the molecule. There are four main types used of chromatographies used to separate proteins:

- **size exclusion** chromatography in which proteins can be separated according to their size/shape or molecular weight
- **ion exchange** chromatography in which proteins are separated by their charge/isoelectric point;
- **hydrophobic interaction** chromatography (similar to reverse phase columns for purifying organic molecules) in which they are separated based on their relative hydrophobicity
- **affinity** chromatography in which proteins are separated based on binding to a ligand covalently attached to a column bead.

For preparative protein purification, the purification protocol generally contains one or more chromatographic steps. The basic procedure in chromatography is to flow the solution containing the protein through a column packed with a chromatography resin selected to separate proteins based on a specific property of the protein. Different proteins interact differently with the column material, and can thus be separated by the time required to pass the column, or the conditions required to elute the protein from the column. Usually, proteins are detected as they are eluting from the column by measuring the absorbance at 280 nm, at which the aromatic amino acids absorb.

3.4.5.1: Size Exclusion Chromatography (also known as Gel Filtration Chromatography)

This method is used to separate proteins based on size and shape. The chromatography beads have tiny openings/pores into which proteins of a size less than the pore diameter, can enter. Large proteins that can't enter the pore flow around the beads and elute faster than small ones that enter the pores. They diffuse out of the pores and enter the rest of the moving solvent before getting "trapped" again for a short time in more pores. Eventually, they work their way through the column and elute at a volume much greater than proteins, which can't enter the pores. Thus, proteins will be separated based on their size as illustrated in Figure 3.4.3. The eluate is collected in sequential test tubes (or fractions). Note that the figure below shows the pores as actual channels that go through the bead. In actuality, the openings in resin beads should be considered to be pores, not channels.

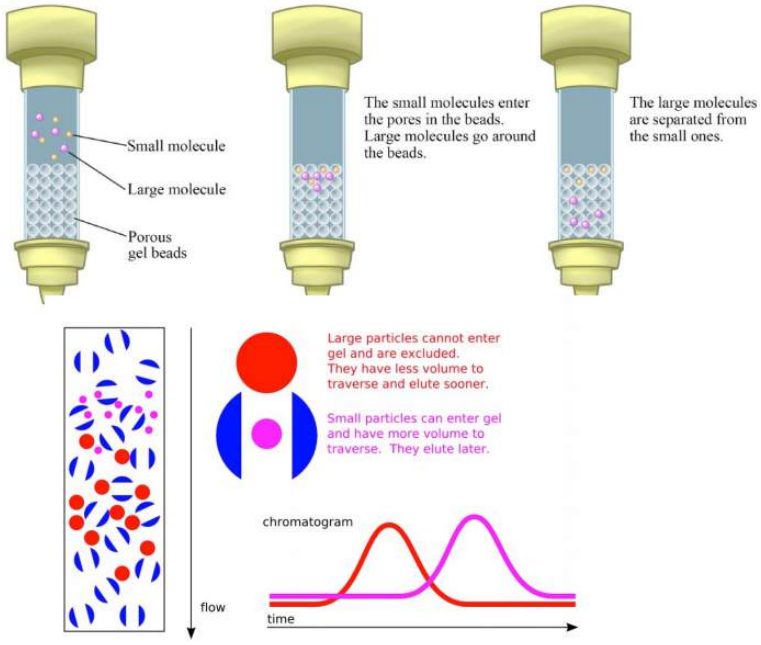


Figure 3.4.3: **Size Exclusion Chromatography.** *Image from Dr. Kevin Ahern and Indira Rajagopal*

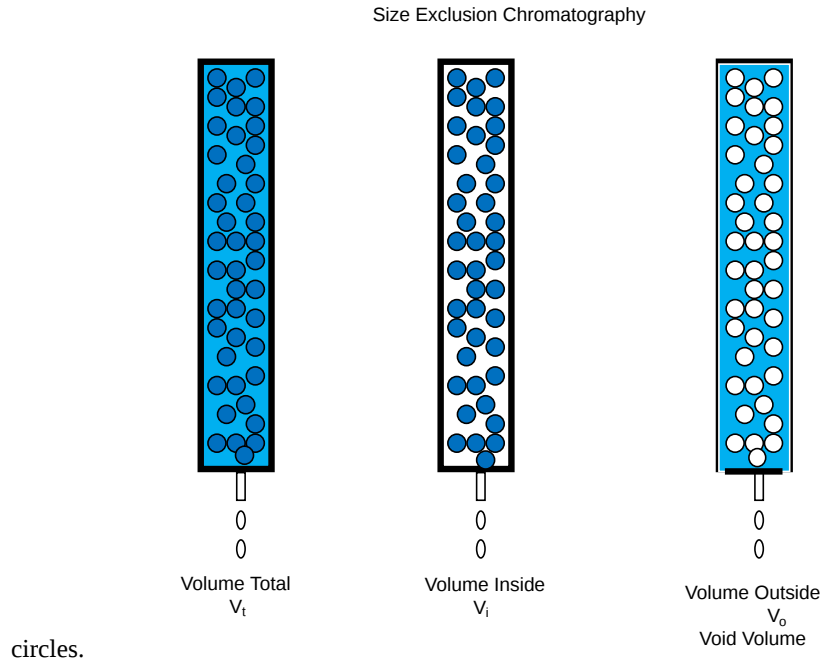
Also known as gel filtration chromatography, is a low-resolution isolation method that involves the use of beads that have tiny “tunnels" in them that each have a precise size. The size is referred to as an “exclusion limit," which means that molecules above a certain molecular weight will not fit into the tunnels. Molecules with sizes larger than the exclusion limit do not enter the tunnels and pass through the column relatively quickly by making their way between the beads. Smaller molecules, which can enter the tunnels, do so, and thus, have a longer path that they take in passing through the column. Because of this, molecules larger than the exclusion limit will leave the column earlier, while smaller molecules that pass through the beads will elute from the column later. This method allows the separation of molecules by their size.

In any chromatography system, there is a mobile and stationary phase. For size exclusion chromatography, the stationary phase is usually a polymerized agarose or acrylamide bead, which contains pores of various sizes filled with the solvent. Let's pretend that the solvent (typically aqueous buffered solution) inside of the bead is trapped there and doesn't exchange with the solvent moving around the bead, so it would be part of the stationary phase. The mobile phase is the solvent used to elute the column which flows around the bead. The chromatography beads are often supplied in dried form, which must be swollen in the solvent before they are packed in the column. The actual volume of the agarose or acrylamide bead is very small compared to the volume of solution within their hydrated forms.

3.4.5.2: Size and shape effects in size exclusion chromatography

Size-size exclusion chromatography is so common, so we will explore it in greater detail

Several different column volumes can be defined as shown in Figure 3.4.4, where the packed chromatography beads are shown as



circles.

Figure 3.4.4: Define volumes in size exclusion chromatography

If we consider the mass of the beads to offer a *negligible amount to the volume of the bead*, the actual volume in the bead is mostly from the trapped solution, which can be considered to be the "stationary" phase. The volume around the bead is called the **void volume**, V_o. It should be apparent the volume inside the bead is given by

$$V_i = V_t - V_o \tag{3.4.1}$$

A solute elutes from the column in a broad peak. If the sample volume applied to the column is very small compared to V_t, the volume at which a solute elutes, V_e, is considered to be the center of the elution peak. This is true when $V_{sample} \gg V_e$.

If we view this chromatography as a partitioning of solute between the mobile and stationary phases (the basis of all chromatography), we might be interested in what fraction of the stationary phase, V_i, a solute might partition into. Such a ratio would be given by:

$$K = \frac{V_e - V_o}{V_t - V_o} \tag{3.4.2}$$

where V_t-V_o (= V_{inside}) represents 100% of the stationary phase, where K is a **distribution coefficient**. Consider two cases:

1. A very large solute compared to the pore size of the bead: In this case, V_e-V_o = 0 since V_e would be equal to V_o. (The solute wouldn't "see" any of the V_i.) In this case, K = 0. The solute would elute in the void volume of the column since it is too large to partition into the volume within the beads. All solutes of molecular weight greater than or equal to the smallest solute that can't enter the gel beads will all elute in the void volume. Hence solutes greater than this minimal size will co-elute from the column and not be separated. V_o is usually about 30-40% of the V_t.
2. A very small solute compared to the pore size. In this case V_e-V_o = V_t - V_o, since V_e would be equal to V_t. The solute would "see" all of the solvent within the bead. In this case, K = 1. Similar to above, all solutes of MW equal to or less than the largest solute that can partition into the entire volume within a bead will co-elute at a volume near V_t.

Hence K is a **partition coefficient**, which varies from 0 - 1, and represents that fraction of V_i into which a solute could partition. This K is not exactly a partition coefficient, however, since the actual volume of the gel matrix is assumed to be zero above. The graph in Figure 3.4.5 shows typical V_e as a fraction of V_t for solutes of different sizes (x-axis is V_e/V_t).

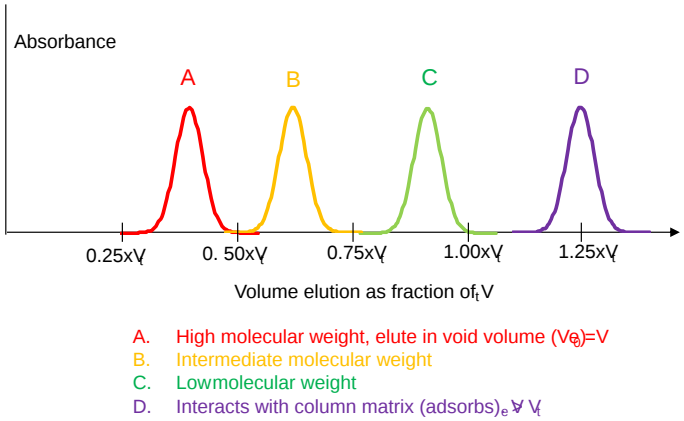


Figure 3.4.5: Elution volumes (V_e for different size macromolecules)

Large species that cannot enter the pores in the beads flow around it and elute in the void volume (V_0) which is about 35-40% of V_t (red bell-shaped curve). Very small species can partition into both V_0 and V_i so the elute near V_t (green bell-shaped curve). If a species adsorbs to the column bead through noncovalent interactions (such as hydrogen bonds or ion-ion interactions), it may elute after V_t (purple bell-shaped curve).

K depends on the size and shape of the solute. The size and shape of an object determine its flow properties in a fluid. Frictional resistance (itself a force, which acts in the opposite direction to the velocity, another vector quantity), can be shown to be proportional to the velocity.

$$F_f \propto v \tag{3.4.3}$$

or

$$F_f = -fv \tag{3.4.4}$$

where f is the frictional coefficient, which depends on the shape. Clearly, the bigger the object, the more frictional resistance to movement. For a sphere it can be shown that:

$$f = 6\pi\eta R_s \tag{3.4.5}$$

where η is the viscosity (a measure of the resistance to flow of a liquid - water has a low viscosity, real maple syrup a high viscosity), and R_s (Stokes radius) is the radius of the hydrated sphere (the larger R_s, the larger the frictional coefficient, the larger the F_f which resists motion). For an irregularly shaped object, the Stokes radius is the radius of a sphere that would have the same frictional coefficient as the object. Hence the R_s for a protein molecule that was not spherical in shape would be much larger than the R_s for another protein molecule of identical molecular weight that was spherical. Hence the V_e and the K value for a solute on a gel filtration column would best be related to the Stokes radius, since R_s values take into account both size and shape.

If you separate two proteins of equal mass but one is highly elongated and the other is spherical, the elongated one, with a large R_S, would elute first (assuming that both don't elute together in the void volume, V_0.

Gel filtration can be used to determine the molecular weight of an unknown, spherical (globular) protein when compared to a standard curve generated from other globular proteins of known molecule weight. To ensure the protein have the same "effective" shape, the proteins are eluted under denaturing conditions to remove shape contributions to the elution order.

3.4.5.3: Separation on the basis of charge - Ion Exchange Chromatography

The chromatography resin in this type of chromatography consists of an agarose, acrylamide, or cellulose resin or bead which is derivatized to contain covalently linked positively or negatively charged groups. Proteins in the mobile phase will bind through electrostatic interactions to the charged group on the column. In a mixture of proteins, positively charged proteins will bind to a resin containing negatively charged groups, like the carboxymethyl group, CM ($-OCH_2COO^-$) or sulfopropyl, SP, ($-OCH_2CH_2CH_2SO_3^-$) while the negatively charged proteins will pass through the column. The positively charged proteins can be eluted from the column with a mobile phase containing either a gradient of increasing salt concentration or a single higher salt

concentration (isocratic elution). The most positively charged protein will be eluted last, at the highest salt concentration. Likewise, negatively charged proteins will bind to a resin containing positively charged groups, like the diethylaminoethyl group, DEAE (-$OCH_2CH_2NH(C_2H_5)_2^+$) or a quaternary ethyl amino group, QAE, and can be separated in an analogous fashion.

Ion exchange chromatography separates compounds according to the nature and degree of their ionic charge. The column to be used is selected according to its type and strength of charge. Anion exchange resins have a positive charge and are used to retain and separate negatively charged compounds (anions), while cation exchange resins have a negative charge and are used to separate positively charged molecules (cations).

Before the separation begins a buffer is pumped through the column to equilibrate the opposing charged ions. Upon injection of the sample, solute molecules will exchange with the buffer ions as each competes for the binding sites on the resin. The length of retention for each solute depends upon the strength of its charge. The most weakly charged compounds will elute first, followed by those with successively stronger charges. Because of the nature of the separating mechanism, pH, buffer type, buffer concentration, and temperature all play important roles in controlling the separation.

Figure 3.4.6 shows a ***cation exchange column***. The beads (brown) contain negatively charged functional groups which can bind positive protein (blue) or concentrated regions of positive charge on a protein.

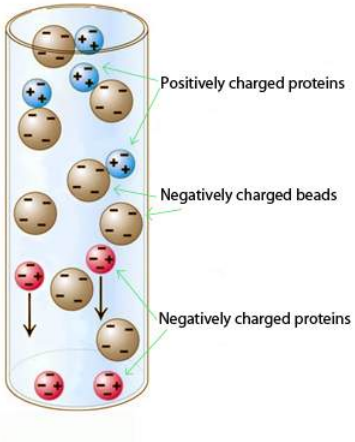


Figure 3.4.6: Ion Exchange Chromatography https://upload.wikimedia.org/wikiped...n_exchange.jpg. Jspiteri at English Wikibooks, Public domain, via Wikimedia Commons

Before loading the column with protein, the negatively-charged beads would interact with positively charged counter cations (often Na^+) from the column equilibration buffer. When the protein solution is introduced to the column, the positively charged protein will exchange with the bound Na^+ ions (hence the name cation exchanger). Conversely, an anion exchanger consists of positively charged beads, which exchange bound anions. Proteins bounds through ion-ion interactions can be eluted by increasing the Na^+ concentration in the eluting solution either in a stepwise or gradient fashion. Ion exchange chromatography is a very powerful tool for use in protein purification and is frequently used in both analytical and preparative separations.

3.4.5.4: Affinity Chromatography

In this technique, the chromatography resin is derivatized with a group that binds to a specific site on a given protein of interest. It may be a group that binds to the active site of an enzyme (such as benzamidine-agarose which is used for the purification of trypsin) or an antibody, which recognizes a specific amino acid sequence (an epitope) on a protein. For example, an antibody can be made to a specific peptide from albumin, the antibody covalently linked to agarose, and the antibody-agarose column then used to purify albumin specifically. This is a powerful technique since antibodies can be made that will bind selectively to a single protein. Knowing only the DNA sequence of a protein which has never been previously isolated, the amino acid sequence of the unknown protein can be derived from the DNA sequence. A 10-12 amino acid peptide from that protein can be synthesized in the lab (seethe last section below), and an antibody raised against the peptide. The antibody will most likely bind to the unknown protein as well as to the peptide, and hence could be used to purify the protein.

These features of affinity chromatography are illustrated in Figure 3.4.7.

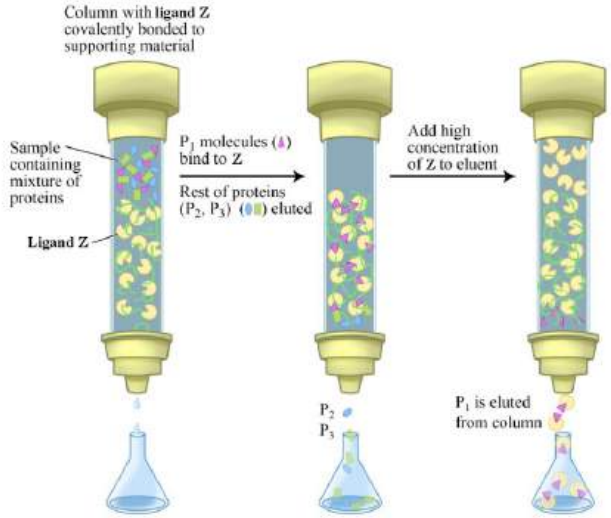


Figure 3.4.7: **Example of Affinity Chromatography.**

In this example in Figure 3.4.7, protein P1 has affinity for ligand Z and will bind to the column while proteins P2 and P3 will pass through the column. Protein P1 can then be eluted from the column using high concentrations of free ligand Z.

In vitro peptide synthesis for antibody production

When making anti-peptide antibodies that recognized target proteins, or to study an isolated peptide by itself, it is more difficult to isolate and purify a peptide from its original protein than to synthesize it in the lab using solid-phase synthesis. We describe this technique below.

Peptides are chemically synthesized by the condensation reaction of the carboxyl group of one amino acid to the amino group of another. Two chemical challenges must be addressed. The formation of an amide bond between the carboxylic acid of one amino acid and the amine of the other is thermodynamically unfavorable, so the carboxyl end must be activated typically by the reaction of the incoming amino acid with a reagent such as dicyclocarbodiimide. Secondly, reactive functional groups on the side chains and the amine of the carboxyl group-activated amino acid must be protected from unwanted reactions. Chemical peptide synthesis most commonly starts at the carboxyl end of the peptide (C-terminus), and proceeds toward the amino-terminus (N-terminus). Protein biosynthesis in living organisms occurs in the opposite direction. Chemical synthesis facilitates the production of peptides, which incorporate unnatural amino acids, peptide/protein backbone modification, and the synthesis of D-amino acids.

The established method for the production of synthetic peptides in the lab is known as solid-phase peptide synthesis (SPPS). Pioneered by Robert Bruce Merrifield, SPPS allows the rapid assembly of a peptide chain through successive reactions of amino acid derivatives on an insoluble porous support. The solid support consists of small, polymeric resin beads functionalized with reactive groups (such as amine or hydroxyl groups) that link to the nascent peptide chain. Since the peptide remains covalently attached to the support throughout the synthesis, excess reagents and side products can be removed by washing and filtration. This approach circumvents the comparatively time-consuming isolation of the product peptide from the solution after each reaction step, which would be required when using conventional solution-phase synthesis.

Each amino acid to be coupled to the peptide chain N-terminus must be protected on its N-terminus and side chain using appropriate protecting groups such as t-Boc (t-butyloxycarbonyl-, acid-labile) or flourenylmethyloxycarbonyl (Fmoc, base-labile), depending on the side chain and the protection strategy used (see below).

The general SPPS procedure involves repeated cycles of alternate N-terminal deprotections and coupling reactions. The resin can be washed between each step to remove side products. The mechanism for the solid phase synthesis of a dipeptide is shown in Figure 3.4.8.

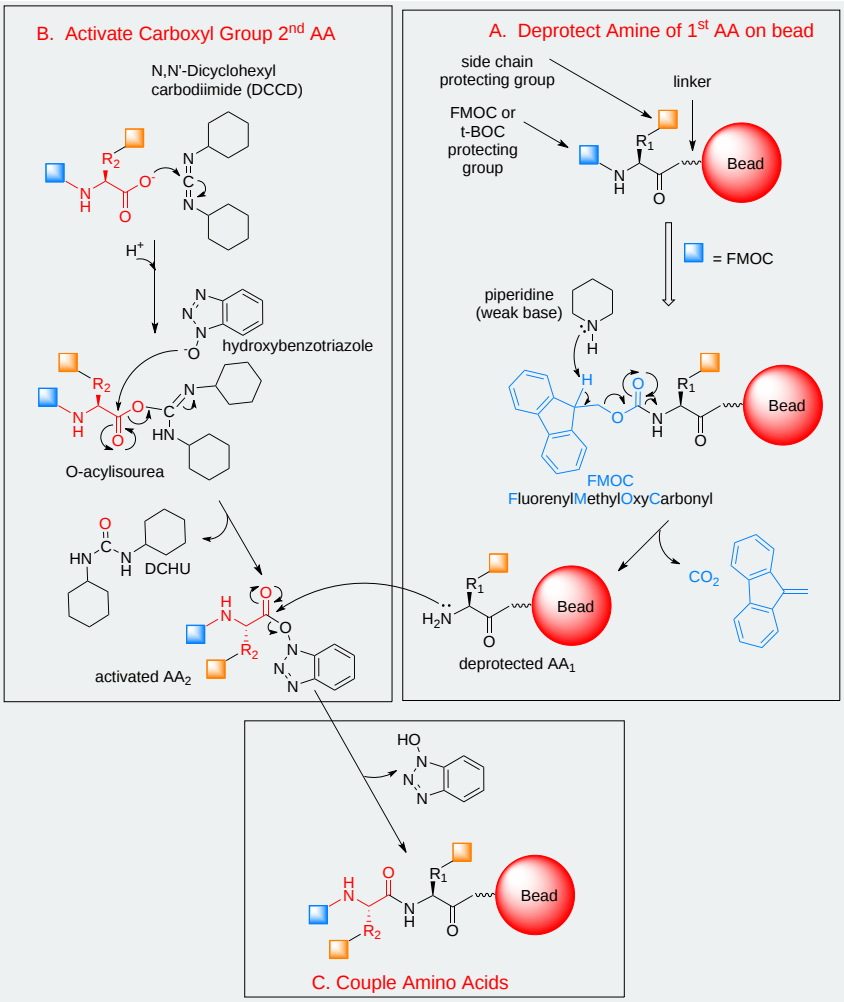


Figure 3.4.8: Solid phase peptide/protein synthesis

A. Deprotection of AA_1: The first amino acid is coupled to the resin or purchased pre-coupled. The amine terminus contacting an FMOC group is deprotected with piperidine. The hydrogen abstracted from the FMOC is acidic as its negatively charged conjugated base is aromatic, since the negative charge on that C becomes sp^2 hybridized to create the aromatic anion. The weak base piperidine is used to avoid side reactions.

B. Activation of AA_2: The carboxyl group of AA_2 reacts with a carbodiimide, which is attacked by the carboxylate of AA_2 leading to the formation of an isourea derivative. This can react with a second nucleophilic catalyst (which is regenerated in step C), hydrobenzotriazole (HBT), to form the activated HBT ester and the very stable urea derivative.

C. Coupling Reaction: The activated AA_2 now reacts with the amine of solid phase N-terminal deprotected AA1 to form the peptide bond.

This cycle repeats until the desired sequence has been synthesized. At the end of the synthesis, the crude peptide is cleaved from the solid support while simultaneously removing all protecting groups using a reagent strong acid like trifluoroacetic acid. The crude peptide can be precipitated from a non-polar solvent like diethyl ether in order to remove organic soluble by-products and then purified using reversed-phase HPLC. The purification process, especially of longer peptides can be challenging, because small amounts of several byproducts, which are very similar to the product, have to be removed. For this reason, so-called continuous chromatography processes such as MCSGP are increasingly being used in commercial settings to maximize the yield without sacrificing purity levels.

SPPS is limited by reaction yields, and typically peptides and proteins in the range of 70 amino acids are pushing the limits of synthetic accessibility. Synthetic difficulty also is sequence-dependent; typically aggregation-prone sequences such as amyloids are difficult to make. Longer lengths can be accessed by using ligation approaches such as native chemical ligation, where two shorter fully deprotected synthetic peptides can be joined together in solution.

Increasingly, proteins made in cells can be engineered through manipulation of the protein's gene to contain a molecular tag, which can either be a small peptide or a protein, for which antibodies are commercially available. The tag is expressed at either the N- or C-terminal end of the target so as to not interfere with the folding of the expressed target protein. Examples of peptide tags include the His (sequence HHHHHH), FLAG (sequence DYKDDDDK) and HA (YPYDVPDYA) tags. The HA tags derives from the influenza hemagglutinin protein. A small protein such as the green Fluorescent Protein - GFP) can also be used as a tag. The resulting fusion protein of GFP connected to the target protein can also allow the target protein to be localized and followed by confocal fluorescence microscopy within cell. Chromatography resins with covalently attached antibodies to the His, FLAG, HA peptide tags, and GFP are commercially available as affinity chromatography resins as shown in the right-hand side of Figure 3.4.9: below.

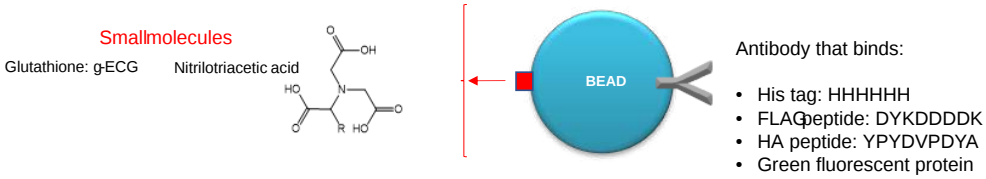


Figure 3.4.9: Protein Tags and small molecule bead ligands for affinity chromatography

Affinity reagents other than antibodies can be attached to the beads, as shown in the left-hand side of Figure 3.4.9. Two, in particular, are Ni-Nitrilotriacetic acid (Ni-NTA) and the short peptide glutathione (γ-gluatmylcysteinylglycine). They also bind tagged proteins. The Ni-Nitrilotriacetic binds the His tag by chelation of the nickel ion with the 6 histidine imidazole groups on the His-tagged protein. (Note that His tags can also be bound to anti-His tag antibody beads.) Glutathione binds a protein tag, Glutathione-S-Transferase (GST) linked in a fusion protein to the target.

The His tag, which is probably the most widely used, binds strongly to divalent metal ions such as nickel and cobalt. The protein can be passed through a column containing Ni-nitrilotriacetic. All untagged proteins pass through the column. The protein can be eluted with imidazole, which competes with the imidazole side chain on the His tag for binding to the column, or by a decrease in pH (typically to 4.5), which decreases the affinity of the tag for the resin. While this procedure is generally used for the purification of recombinant proteins with an engineered affinity tag (such as a 6xHis tag), it can also be used for natural proteins with an inherent affinity for divalent cations.

3.4.5.5: Hydrophobic Interaction Chromatography (HIC)

HIC media is similar to reverse phase chromatography in which a matrix like silica (very polar with exposed OH groups) is derivatized with ester or ether links from the silica surface hydroxyl OHs to nonpolar molecules, usually containing 8 or 18 carbons in the acyl or alkyl chain. Proteins with exposed hydrophobic groups would preferentially bind to the bead. The interactions of the protein with the derivatized beads are increased by adding high concentrations of salt to the aqueous solution, making water effectively more polar. This would shift the equilibrium towards binding of the surface-exposed nonpolar region on the protein to the nonpolar C8 or C18 chains. The ionic strength of the buffer is then reduced to elute proteins in order of increasing hydrophobicity, as shown in Figure 3.4.10.

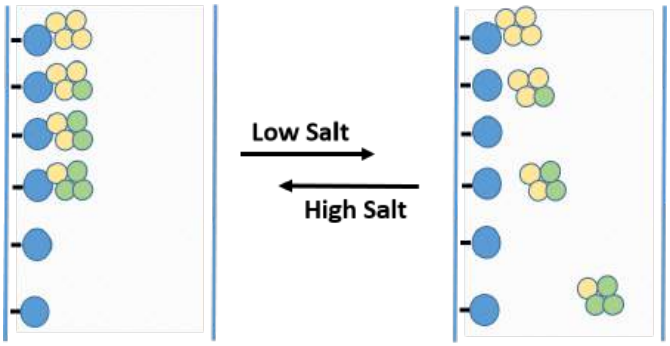


Figure 3.4.10: Hydrophobic Interaction Chromatography.

The column matrix, shown in blue has a hydrophobic ligand covalently attached. In high salt conditions, proteins will bind to the matrix with differing affinity, with more hydrophobic proteins (shown in yellow) binding more tightly than more hydrophilic

proteins (shown in green) When the salt concentration is decreased, proteins that are more hydrophilic will be released first, followed more hydrophobic proteins.

3.4.5.6: High Performance Liquid Chromatography (HPLC) and Fast Protein Liquid Chromatography (FPLC)

High performance liquid chromatography or high pressure liquid chromatography (HPLC) is a form of chromatography applying high pressure to drive the solutes through the column faster than using gravity-forced flow of solvent through the column. The packing beads are small and very closely packed which allows less diffusion and greatly increased resolution. Because of the close packing of the small beads, no flow would occur with an external pump. The most common form of HPLC is "reversed phase" HPLC, where the column packing material is hydrophobic. The proteins are eluted by a gradient of water and increasing amounts of an organic solvent, such as acetonitrile. The proteins elute according to their hydrophobicity. After purification by HPLC, the protein is in a solution that only contains volatile compounds, and can easily be ***lyophilized (freeze-dried)***. HPLC purification frequently results in the denaturation of the purified proteins and is thus not applicable to proteins that do not spontaneously refold.

Due to the drawbacks of HPLC, an alternative technique using lower pressure was developed and is called ***Fast protein liquid chromatography (FPLC).*** In FPLC, the mobile phase is an aqueous solution, or "buffer". The buffer flow rate is controlled by a positive-displacement pump and is normally kept constant, while the composition of the buffer can be varied by drawing fluids in different proportions from two or more external reservoirs. The stationary phase is a resin composed of beads, usually of cross-linked agarose, packed into a cylindrical glass or plastic column. FPLC resins are available in a wide range of bead sizes and surface ligands depending on the application.

In the most common FPLC purification systems as shown in Figure 3.4.11, an ion exchange resin is typically chosen.

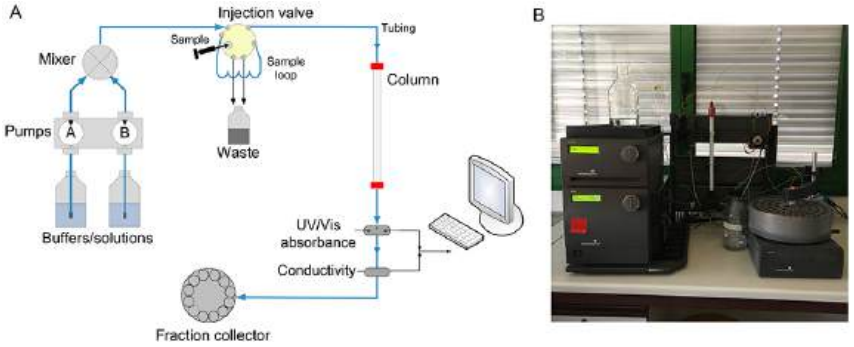


Figure 3.4.11: **Typical FPLC System.** A. Scheme of basic components and typical flow path for a chromatography system. B. Picrue of GE Healthcare AKTA FPLC apparatus. *Image provided by LaVerde, V., Dominici, P. and Astegno, A. (2017)* Bio-protocol *7(8): e2230.*

A mixture containing one or more proteins of interest is dissolved in 100% buffer A and pumped into the column. The proteins of interest bind to the resin while other components are carried out in the buffer. The total flow rate of the buffer is kept constant; however, the proportion of Buffer B (the "elution" buffer) is gradually increased from 0% to 100% according to a programmed change in concentration (the "gradient"). Buffer B contains high concentrations of the exchanger ion. Thus as the concentration of Buffer B gradually increases, bound proteins will dissociate depending on their ionic interactions with the column matrix and appear in the eluant. The eluant passes through two detectors which measures salt concentration (by conductivity) and protein concentration (by absorption of ultraviolet light at a wavelength of 280 nm). As each protein is eluted it appears in the eluant as a "peak" in protein concentration and can be collected for further use.

3.4.6: Purification Scheme

During the protein purification process it is necessary to have a quantitative system to determine, the total amount and concentration of total and target protein at each step, the biological activity of the target protein, and its overall purity. This will help guide and optimize the purification method being developed. Ineffective separation techniques can be disregarded and other techniques that give higher yield or that retain biological activity of the protein can be adopted.

Thus, each step in the purification scheme is quantitatively evaluated for the following parameters: total protein, total activity, specific activity, yield, and purification level. Each of these parameters will be defined within the protocol given below.

Pretend you are a researcher that wants to isolate a novel, unknown protein from a bacterial culture. You grow 500 ml of the bacteria overnight at 37ºC and harvest the bacteria by centrifugation. You remove the culture broth and retain the bacterial pellet.

You then lyse the bacteria using freeze/thaw in 10 mL of reaction buffer. You then centrifuge the lysed bacteria to remove the insoluble materials and retain the supernatant that contains the soluble proteins. Your protein of interest has a biological activity that you can measure using a simple assay that causes a color change in the reaction mixture, as illustrated in Figure 3.4.12 You also note that this reaction rate increases with increasing concentrations of your protein supernatant.

Figure 3.4.12: **Example of a Chemical Reaction that causes a color change from orange to brown depending on increasing concentration.** *Image from: Ludwig, N., et. al. (2015) on* Research Gate

At this point, you can measure your baseline concentrations for the first purification level (bacterial lysis and removal of insoluble proteins and other cellular debris by centrifugation).

Total Protein is calculated by measuring the concentration in a fraction of the sample, and then multiplying that by the total volume of your sample. In this case, you are starting with 10 mL of supernatant. In a typical assay to measure protein concentration, you will use 50 - 200 µL of sample to determine the protein concentration. For example, if you calculate that there is 7.5 µg/µL in your initial assay, you would need to convert that value into mg/mL and then multiply it by 10 mL for a total of 75 mg of protein in 10 mL of supernatant (Table 3.1)

Total Activity is measured as the enzyme activity within the assay, multiplied by the total volume of the sample. For example, in the initial sample, you might use 5 to 50 µL of sample in your biological reaction (Figure 3.10). If you calculated the activity in your assay to be 2.5 units/µL, this would be equivalent to 2,500 units/mL or 25,000 units/10 mL of supernatant. Note that, the ***enzyme unit***, or international unit for the enzyme (symbol U, sometimes also IU) described the ***enzyme's catalytic activity***. 1 U (µmol/min) is defined as the amount of the enzyme that catalyzes the conversion of one micromole of substrate per minute under the specified conditions of the assay method.

Specific Activity is measured by dividing the Total Activity by the Total Protein. In our example, 25,000 units divided by 75 mg of protein = 333.3 units/mg.

Yield is a measure of the biological activity retained in the sample after each purification step. The amount in the first step is set to be 100%. All subsequent yield steps will be evaluated using the first purification step. It is calculated by dividing the total activity of the current step, by the total activity of the first step and then multiplying by 100.

Purification level evaluates the purity of the protein of interest by dividing the specific activity calculated after each purification step by the specific activity of the first purification step. Thus, the first step always has a value of 1. A typical purification analysis scheme is shown in Table 3.4.1 below.

Purification Step	Purification Method	Total Protein (mg)	Total Activity (units)	Specific Activity (units/mg)	Yield (%)	Purification Level
1	Cell Lysis - Centrifugation	75	25,000	333.3	100	1
2	Salting Out	54	23,416	433.6	93.7	1.3
3	Ion Exchange Chromatography	12.3	21,227	1725.8	84.9	5.2
4	Size Exclusion Chromatography	0.93	18,633	20,035	74.5	60.1

Table 3.4.1: A typical purification analysis scheme.

Note that after each purification step, the ***Total Protein*** goes down, as you are separating the target protein away from other proteins in the mixture. ***Total Activity*** also goes down with each purification step, as some of your protein of interest is also lost at each purification step, because (1) some protein will stick to the test tubes and glassware, (2) some protein won't bind with 100% efficiency to your column matrix, (3) some protein may bind too tightly to be removed from the column matrix during elution, and (4) some protein may be denatured or degraded during the purification process.

The amount of your protein of interest that is retained is represented within the overall ***percent yield*** for each purification step. If the ***percent yield*** is too low alternative purification methods should be explored.

Note that in a good protein purification scheme that the ***specific activity*** should go up substantially with each level of purification as the amount of your protein of interest makes up a greater percentage of the total protein within that fraction. If the specific activity only increases modestly within a purification step, or if it decreases during a purification step, this could indicate that (1) your protein of interest is being substantially lost at that step, (2) your protein of interest is being denatured or degraded and is no longer biologically active, or (3) that a required cofactor or binding protein is being reduced at that purification step. Additional experiments may need to be conducted to determine which of the causes predominate, so that steps can be taken to reduce protein inactivation. For example, many proteins are temperature sensitive and will degrade or denature at room temperature. Completing purification steps on ice can often reduce degradation.

Overall, the fold increase in ***purification level*** should increase exponentially during the purification process. Note that in our example, if after 4 steps of purification our proteins is close to 95% pure, this would indicate that our protein of interest makes up approximately 1.24% of the total protein within the sample.

3.4.7: Electrophoresis: Separation and Analysis

In column chromatography, flow through the column is driven by hydrostatic pressure causing flow from higher regions of higher pressure at the top of the column reservoir to lower pressure (drops eluting from the bottom of the column). Ultimately the hydrostatic pressure (in columns not driven by mechanical pumps) derives from the gravitational force. However, proteins are also charged particles and can be moved by an external electric field instead of a gravitational field. Electrophoresis is the movement of charged particles in an electric field. As we will show below, the movement of a charged protein within a static matrix in the presence of an external electric field depends on both size and shape. Electrophoresis can be used for both analytical and preparative separations of proteins. The most common uses are for analytical separations.

3.4.7.1: Theory

What determines how a protein moves in an electric field? Consider the simple case of a charged particle (+Q) moving in an electric field (E) in a nonconducting medium, such as water. If the particle is moving at a constant velocity toward the cathode (-electrode where cations go), the net force F_{tot} on the particle is 0 (since F=ma, and the acceleration (a) of the particle is 0 at constant velocity). Two forces are exerted on the particle, one F_E, the force exerted on the charged particle by the field, which is in the direction of the motion (toward the cathode), and the other, F_f, the frictional force on the charged particle, which retards its motion toward the cathode, and hence is in the direction opposite to the motion (toward the anode (+) electrode). This is shown in the Figure 3.4.13:

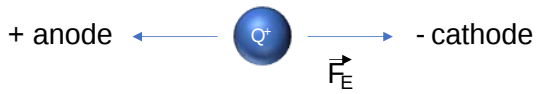


Figure 3.4.13: Movement of charge Q^+ in an electric field

Therefore:

$$F_{tot} = F_E + F_f \tag{3.4.6}$$

where F_E, the electrical force, is

$$F_E = QE \tag{3.4.7}$$

and

F_f, the frictional force, is

$$F_f = -fv \tag{3.4.8}$$

In the last equation, v is the velocity of the particle, and f is a constant called the frictional coefficient. This equation shows that the force F_f hindering motion toward the cathode is proportional to the velocity of the particle. This is intuitive since one would expect the higher the velocity, the greater the F_f which would hinder the motion. The frictional coefficient depends on the size and shape

of the molecule. The larger the molecule, the larger the frictional coefficient (i.e. more resistance to the motion of the molecule). It can be shown that the frictional coefficient for a spherical particle is given by

$$f = 6\pi\eta R_S \tag{3.4.9}$$

where η is the viscosity (a measure of the resistance to flow of a liquid - water has a low viscosity, real maple syrup a high viscosity), and R_s (Stokes radius) is the radius of the hydrated sphere (the larger R_s, the larger the frictional coefficient, the larger the F_f which resists motion toward the cathode). This equation should be intuitive from your experiences. From (1), (2), and (3), $F_e = F_f$, or

$$QE = fv \tag{3.4.10}$$

Hence v/E = Q/f = U = the **electrophoretic mobility**, or

$$U = \frac{V}{E} = \frac{Q}{6\pi\eta R_S} \tag{3.4.11}$$

Therefore, the electrophoretic mobility U is proportional to the **charge density (charge/size, Q/R_s)** of the particle, not just the size as is the case for spherical proteins in size exclusion chromatography. Macromolecules of different **charge density** can thus be separated by electrophoresis. This discussion deals with the simplest case, since in reality there are counter ions in the solution (from salts) which would form a cloud around the charged macromolecule, and partially shield the charged particle from the electric field E.

Modern day electrophoresis is conducted in solid gels (such as polyacrylamide), which are formed from liquid acrylamide solutions after the addition of a polymerizing agent. The solid gel is porous to solute and solvent molecules and serves as a medium for electrophoresis while helping to eliminate convection forces in the liquid which interfere with the separation. Electrophoretic experiments have been conducted on the space station in weightless conditions in order to prevent such perturbations.

One complication that affects this idealized description of electrophoresis in polyacrylamide gels is that the gels have pores through which the macromolecules move. Think of the protein moving under an electric force through a "spider web-like" matrix. As in gel chromatography, the smaller molecules can pass through the pores more readily than larger molecules, so there is an additional sieving mechanism that contributes to the *effective* mobility (Also, the gel could alter the local effective electric field). The sieving effect of the gel actually increases the resolving power of this technique.

It has been determined that the actual electrophoretic mobility of the protein, U, is a function of the mobility of the protein in a concentrated sucrose solution (Uo) and T, the total concentration of the acrylamide in the polymerized gel. The higher the concentration of acrylamide in the unpolymerized gel solution, the smaller the size of the pores in the polymerized gel. An equation showing the relationship between U, Uo, and T is shown below:

$$\log U = \log U_0 - K_r T \tag{3.4.12}$$

where K_r is the slope of a plot of log U vs T for a given protein. Since K_r is a function of the radius of the molecule, it is possible to determine the molecular weight of a protein molecule by performing several electrophoretic separations in gels of different acrylamide concentrations (T), and extrapolating results to T = 0, hence eliminating pore size effects. Problems arise, however, if the proteins are **not** spheroid in shape

Is there any way to obtain molecular weight information, in addition to purity determination, on a single gel? What would result if two different proteins, each with the same molecular weight and total net charge, but different shapes, were run on a single acrylamide gel? The one having the more elongated shape (large Stokes radius) would have lower electrophoretic mobility (U =

$Q/6\pi\eta R_s$). A larger Rs would also cause the protein to enter the pores at a slower rate. Hence both electrophoretic mobility and sieving effects would cause this protein to run anomalously slow and have a higher apparent molecular weight. Also ,imagine two globular proteins of different sizes but with compensatory charge differences which might allow the two proteins to migrate at the same speed in the gel.

An astute reader might quickly recognize a problem with the separation of proteins by electrophoresis in a gel. Some proteins are negatively charged (pH > pI), some would be neutral (pH=pI) and the rest would be positive (pH < pI). Only some proteins would enter the gel and move to the electrode at the bottom of the gel. Luckily there is a way to eliminate both charge and shape effects in the electrophoresis of proteins and that is to run the gel under denaturing conditions when all proteins have the same charge density. The denaturant of choice for electrophoresis is usually sodium dodecyl sulfate (SDS), which is an ionic detergent with the structure $CH_3(CH_2)_{10}CH_2OSO_3^-$ (a single chain amphiphile). This detergent binds to and denatures most proteins, with about 1.4 g SDS binding/g of protein (about 1 SDS/2 amino acids). Since there is 1 negative charge/SDS, the binding of SDS masks any of the charges on the protein, and gives all proteins an overall large negative charge. *Additionally*, SDS-proteins complexes have been shown to generally have an elongated cylindrical-like shape. Since the amount of SDS bound per unit mass of protein is constant, the overall charge density on all proteins is similar, so the electrophoretic mobility is only determined by sieving effects.

SDS also eliminates shape differences in the proteins as a variable, since all proteins have the same general rod-like shape. (The use of SDS is analogous to the use of 8M urea in the gel chromatographic separation of proteins to determine molecular weights). Mobility becomes only a function of the molecular weight of the protein, and not its shape. The molecular weight of an unknown protein can be determined by comparing the protein's position on an SDS polyacrylamide gel with a series of known molecular weight standards from which a linear plot of the ln Mr vs R_f can be used to calculate unknown molecular weights. This is similar to the analysis in gel chromatography, where ln Mr is a linear function of K_{avg}, the distribution coefficient, when the gel is run under denaturing conditions. However, some proteins run anomalously on such gels (due to incomplete or excess binding SDS), so alternative techniques of molecular weight determination should be used in conjunction with this technique.

Proteins are usually heated in SDS to 100^oC for 3 minutes, in the presence of a reducing agent such as β-mercaptoethanol (βME), to completely denature the protein to a rod-shaped protein. Apparent molecular weight can be obtained under non-reducing conditions (without βME), but these should be considered just estimates. Running proteins both in the presence and absence of the reducing agent can provide important information on the subunit structure of a protein. A multimeric protein whose subunits are held together by disulfide bonds can be resolved into its individual components when the reducing agent is added. If the subunits are held together by noncovalent intermolecular attractions, the proteins will run identically under the denaturing conditions (SDS), which will eliminate subunit interactions, in the presence or absence of b-ME. To determine the subunit composition of a protein held together by noncovalent interactions, the electrophoresis should be performed in the absence of denaturing agents.

Electrolytic vs Galvanic Cells

Electrode nomenclature might be confusing to some of you. As mentioned above, cations move towards the cathode (where reduction occurs), so the cathode must be negative. Likewise, anion move towards the anode (where oxidation occurs), so the anode must be positive. This is the opposite of what you might remember from introductory science courses when you discussed primarily galvanic cells. In galvanic cells, an electrical current is generated from a spontaneous set of redox half-reactions. In electrophoresis, electrolytic cells are used, in which reactions such as the electrolysis of water ($2H_2O(l) \rightarrow 2H_2(g) + O_2(g)$) or the productions of $Cl_2(g)$ and Mg(s) from the aqueous electrolyte $MgCl_2(aq)$ occur. Those who have done electrophoresis will have seen robust bubble production from the electrodes arising from the electrolysis (a redox reaction) of water ($2H_2O \rightarrow 2H_2(g) + O_2(g)$). In electrolytic cells, a power supply must supply the current to drive the nonspontaneous (unfavored thermodynamically) reactions, such as outlined above. These differences are illustrated in Figure 3.4.14

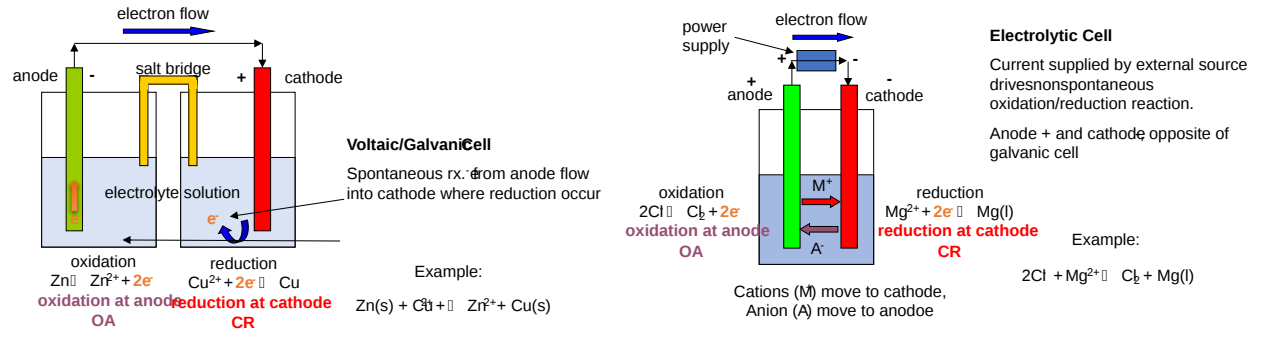


Figure 3.4.14: Galvanic vs electrolytic cells

3.4.7.2: Polyacrylamide Gel Electrophoresis - PAGE

Electrophoresis is performed in a porous, yet solid medium, to eliminate any problems associated with convection currents. Such media are formed from the solidification of a liquid solution of agarose (used mostly for electrophoresis of DNA fragments and very large proteins) or the polymerization of a solution of acrylamide. Polymerization of acrylamide is initiated by the additions of ammonium persulfate in the presence of tetramethylenediamine (TEMED), along with a dimer of acrylamide (N,N'-methylene-bis(acrylamide) connected covalently between the amide nitrogens of the acrylamides by a methylene group. The structures of these compounds are shown in Figure 3.4.15.

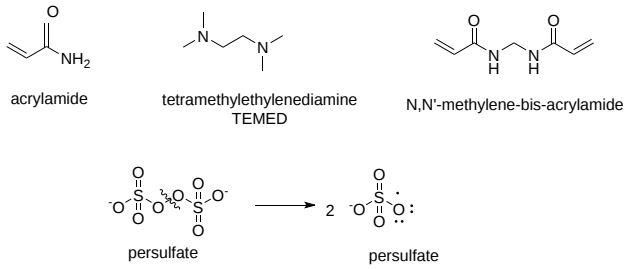


Figure 3.4.15: Structures of reagents use to make PAGE gels

The free radical polymerization of the acrylamide is initiated by the addition of ammonium persulfate, which on dissolving in water, forms free radicals, as shown above

The radical initiates polymerization of the acrylamide, as shown below. The TEMED, through its ability to exist as a free radical, acts as an additional catalyst for polymerization. A rigid gel is only formed, however, when N,N'-methylene-bis(acrylamide is added to the mixture during the polymerization, which cross-links adjacent acrylamide polymers as shown in Figure 3.4.16.

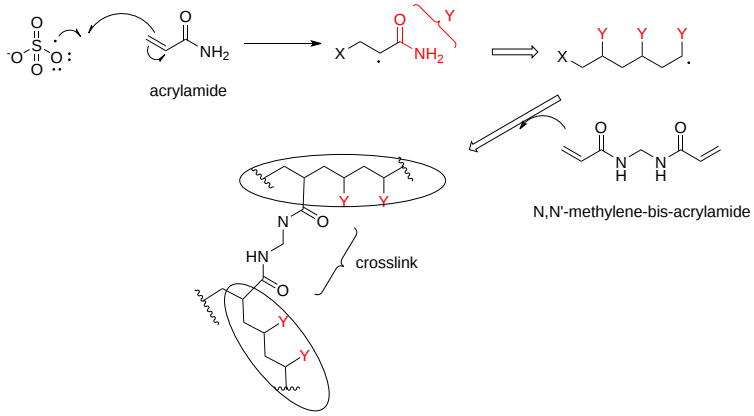


Figure 3.4.16: Polymerization of acrylamide

The amount of bisacrylamide added during polymerization controls the degree of cross-linking, and hence the pore size of the polymerized gel. The effect of pore size is **OPPOSITE** to that in gel chromatography. In both cases, large proteins have a difficult time entering the pore. In gel chromatography, large proteins partition preferentially into the mobile liquid phase (the void volume) and are eluted most **QUICKLY** from the column. In electrophoresis, large proteins, which can not readily enter the pores in the gel,

are not as easily transported by the electric field through the gel, and elute most **SLOWLY**. Pore size can not be controlled as accurately as in the manufacture of gel chromatography resins.

How do proteins migrate through the gel? A viscous protein solution is layered on the top of the gel in a small well molded into the gel during the polymerization process. The bottom and top parts of the gel are inserted into reservoirs containing a buffered solution and the appropriate electrode. The electric field is applied and the proteins migrate through the hydrated gel. The nature of the buffer solution in the reservoir and in the polymerized gel is important. The components of the buffer must not bind to the proteins to be separated. Additionally, for native (non-denatured gels), the pH of the medium must be such that the proteins have the appropriate charge, so they will migrate in the expected direction.

There are many variations of electrophoresis commonly used. Gels can be polymerized in tubes, or slabs, and in the presence or absence of denaturing agents. Additionally, a given slab might consist of two separate slabs polymerized one on top one other, each with a different acrylamide concentration and pH values. The top part is the stacking gel, te bottom is called the running gel. Other gels have a continuous gradient of acrylamide concentrations (from low at the top to high at the bottom). Most commercially available precast gels use continuous acrylamide concentration gradients. Figure 3.4.17 shows a gel placed in an electrophoresis chamber.

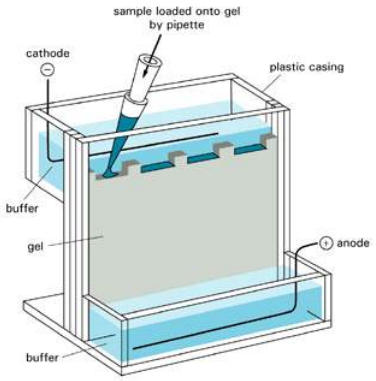


Figure 3.4.17: Electrophoresis gel and chamber https://openwetware.org/mediawiki/in...&oldid=1096467

Whether the gel has a continuous gradient or is discontinuous, the top part of the gel is a low concentration acrylamide (2-4%), often in a Tris HCl buffer solution (pH 6.5) usually 2 pH units below that used in the running gel. The lower part of the gel 8-15% acrylamide, depending on the choice of gel, which is selected based on the molecular weight of the proteins to be separated. The upper buffer reservoir contains Tris-buffered with a weak acid such as glycine (pKa2 = 9.6) to the same pH as the running gel.

Protein electrophorese quickly through the local concentration stacking gel and at the top of continuous gradient gels, and effectively "stack" as they hit the interface between the stacking and running gels, or before they enter too far into the continuous gradient gel. This increases the compactness of the proteins before they enter the "running" section of the gel and increases resolution.

For discontinuous gels, how does this stacking process work? When the electrophoresis is started, glycine ions from the upper reservoir (at pH 8.7) enter the stacking gel since at that pH they have an average partial negative charge. The stacking gel buffer ions continue moving in the stacking gel, but when the glycine ions enter the pH 6.5 of the stacking gel, they become zwitterions with a net charge of zero, and hence stop their motion toward the anode. The electrical resistance in the stacking gel then increases since the number of ions moving through the stacking gel decreases. To maintain constant current throughout the circuit, there will be a *localized* increase in the voltage *in the stacking gel* (from Ohms Law, V=iR). This will cause the proteins to migrate quickly and all stack in a single, very thin disc right behind the Cl- ions in the stacking gel (which are in front because they have the highest charge density and electrophoretic mobility of any ion in the stacking gel). The proteins will not pass the Cl- ions since if they did, they would immediately slow down since they would no longer be in an area of diminished charged carriers and higher voltage. At the stacking gel/running gel interface, the proteins can not all migrate at the same speed, due to sieving effects of the more concentrated gel, and hence will be separated in the running gel. The glycine eventually enters the running gel, assumes its fully charged state at that pH (8.7), will pass the proteins, and restore the deficiency in charge that occurred in the stacking gel.

3.4.7.3: Detection of proteins in the gel:

Most proteins do not absorb at visible wavelengths of light, and hence will not be visible during the course of electrophoresis. To ensure that the proteins are not eluted from the gel into the lower buffer reservoir, a small molecular weight, anionic dye,

bromophenol blue is added to the protein before it is placed on the gel. The electrophoresis is halted when the dye reaches the bottom of the gel. The gel assembly is removed from the electrophoresis chamber, the glass plates separated, and the gel washed into a series of solutions with the goal of rendering the banded proteins visible to the eye.

- Coomassie Brililant Blue dye: This is the most common stain used in labs. It is dissolved in a methanol/acetic acid solution so it generates significant waste. Proteins bind this dye, with a concomitant spectral shift in the absorbance properties of the bound dye. The methanol and acetic acid in the dye solution also help to "fix" the proteins in the gel, and prevent their diffusion into the solution. After the gel is stained, the background stain in the gel is removed with acetic acid/methanol, leaving the blue-colored protein bands. Some proteins will not be stained with Coomassie blue.
- Silver staining: This involves the reduction of Ag(I) to elemental silver and its deposition by protein in the appropriate reaction solutions, much as in a photographic process. (Remember in the BCA assay, peptide bonds reduce Cu(II) to Cu(I), which is chelated to BCA.) A developer and fixer solution is required. This technique is 10-50 X more sensitive than Coomassie Blue staining. Figure 3.4.18 shows gels stained with Coomasie Blue (A)and silver staining (B).

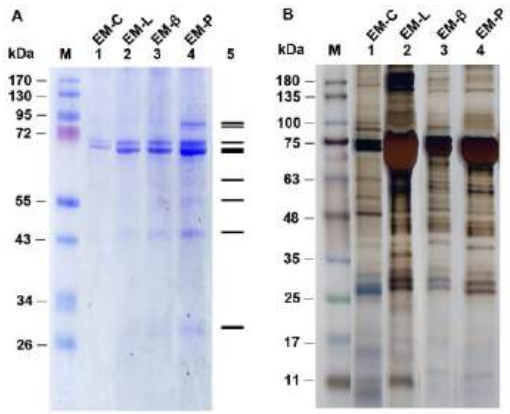


Figure 3.4.18: PAGE gels stained with Coomasie Blue (A) and silver stain (B) Chen Y-Y, Chen J-C, Lin Y-C, Kitikiew S, Li H-F, Bai J-C, et al. (2014) Endogenous Molecules Induced by a Pathogen-Associated Molecular Pattern (PAMP) Elicit Innate Immunity in Shrimp. PLoS ONE 9(12): e115232. https://doi.org/10.1371/journal.pone.0115232. Creative Commons Attribution License

- Pre-electrophoresis fluorescent or radioactive modification of the proteins. These allow even greater sensitivity. After the electrophoresis of a radiolabeled protein, the gel can be dried and overlaid with X-ray film for periods as long as months, if necessary, to allow sufficient exposure of the film by a low concentration protein. This visualization technique is called autoradiography.

3.4.7.4: Variations on polyacrylamide gel electrophoresis:

Isoelectric focusing: In this technique, a pH gradient is set up within the polyacrylamide gel or strip. This is accomplished by pre-electrophoresing a series of low molecular weight molecules containing amino and carboxyl groups called **ampholytes,** each with a different isoelectric point. When subjected to an electric field, the most negative of the species will concentrate at the anode, while the most positive will concentrate toward the cathode. The remaining ampholytes will migrate in-between, with the net effect being that the ampholytes migrate to their isoelectric point and set up a linear pH gradient in the gel.

Proteins initially in regions with a pH below its isoelectric point are positively charged and migrate toward the cathode, while those that are in media with pH lower than its pI will be negatively charged and migrate towards the anode as shown below in Figure 3.4.19. The migration will lead to a region where the pH coincides with its pI. There the protein will have a zero net charge and stop. Thus amphoteric molecules are located in narrow bands where the pI coincides with the pH. In this technique the point of application is not critical as molecules will always move to their pI region. The stable pH gradient between the electrodes is achieved using a mixture of low molecular weight ampholytes whose pIs covers a preset range of pH.

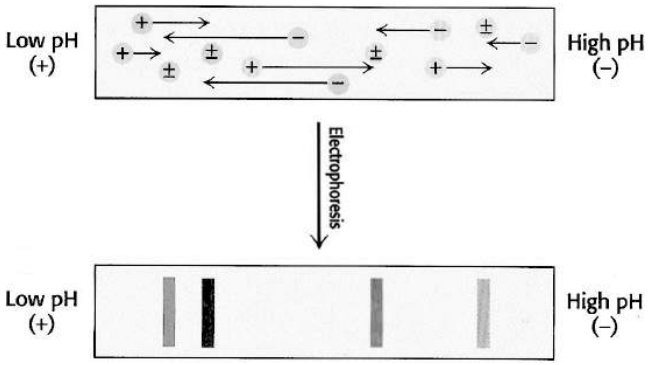


Figure 3.4.19: Isoelectric Focusing. A pH gradient is established in a gel before loading the sample. After the sample is loaded a voltage is applied. The protein will migrate to their isoelectric pH, which they have no net charge. *Image from Magdeldin, S.*

2D electrophoresis: Two-dimensional gel electrophoresis (2-DE) is based on separating a mixture of proteins according to two molecular properties, one in each dimension. The most used is based on a first dimension separation by isoelectric focusing (IEF) and a second dimension according to molecular weight by SDS-PAGE. A conditioning step is applied to proteins separated by IEF prior to the second-dimension run. This process reduces disulfide bonds and alkylates the resultant sulfhydryl groups of the cysteine residues. Concurrently, proteins are coated with SDS for separation on the basis of molecular weight. After the IEF, the tube or strip is placed across the top of a slab gel and subjected to SDS-polyacrylamide gel electrophoresis in a direction 90^{o} from the initial isoelectric focusing experiment. If the proteins were derived from cells labeled with ^{35}Met, representing unique proteins can be obtained from a given cell population. Figure 3.4.20 shows a 2D electrophoresis gel.

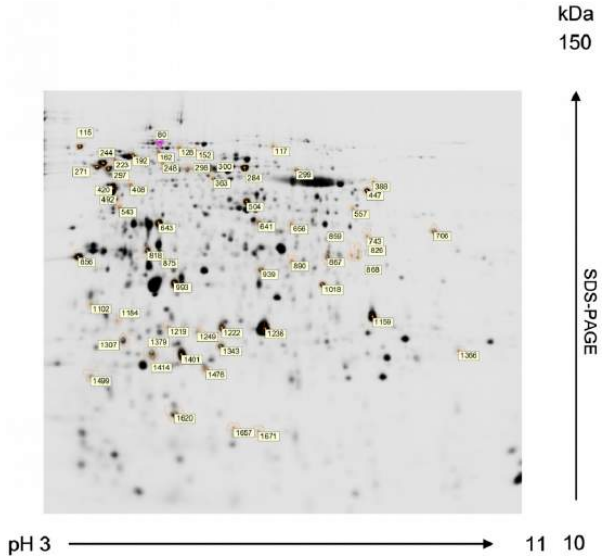


Figure 3.4.20: **Two-Dimentional Gel Electrophoresis.** *Image from Magdeldin, S*

In Figure 3.4.20, proteins of *Chlamydomonas reinhardtii are* resolved by 2-DE from preparative gels stained with MALDI-MS compatible silver reagent for peptide mass fingerprinting analysis. First dimension: isoelectric focusing in a 3-11 pH gradient. Second dimension: SDS-PAGE in a 12% acrylamide (2.6% crosslinking) gel (1.0 mm thick). Numbered spots marked with a circle correspond to proteins compared to be subsequently identified by MALDI-TOF MS. The MALDI-TOF MS analysis of protein sequences is discussed in more detail in section 3.3 below.

One of the biggest problems in 2-DE is the analysis and comparison of complex mixtures of proteins. Currently there are databases capable of comparing two-dimensional gel patterns. These systems allow automatic comparison of spots for the precise identification of those needed in the quantitative analysis. Once interesting proteins are identified, the proteins can be excised from gels, destained, and digested to prepare for their identification by mass spectrometry. This technique is known as peptide mass fingerprinting. The ability to precisely determine molecular weight by matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS) and to search databases for peptide mass matches has made high-throughput protein identification possible. Proteins not identified by MALDI- TOF can be identified by sequence tagging or de novo sequencing using the Q-TOF electrospray LC-MS-MS.

Western blotting: After a standard SDS-slab electrophoresis experiment is run, the gel is overlaid with a piece of nitrocellulose membrane. The sandwich of gel and filter paper is placed back into an electrophoresis chamber, such that the proteins migrate from the gel into the nitrocellulose, where they irreversibly bind. This is illustrated in the figure below. Note however that in the absence of staining, the protein bands in either the PAGE gel or Western blot would not be visible. Standards (lane 5) would be visible if they were labeled with chromophores, as shown in Figure 3.4.21.

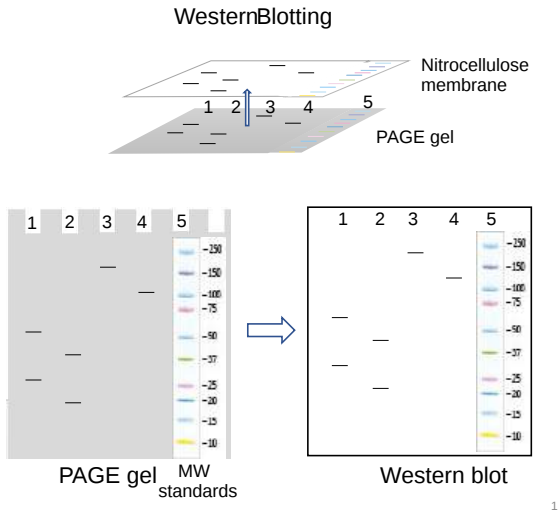


Figure 3.4.21: Western blotting of proteins

If a cell lysate was applied to a lane of a PAGE gel, after staining with any technique, the bands would appear as overlapping smears on the stained gel. What makes Western blots so useful is that specific bands can be specifically visualized (stained) on the nitrocellulose membrane by using a detection system linked to an antibody that recognizes just a specific target protein. This is illustrated in Figure 3.4.22

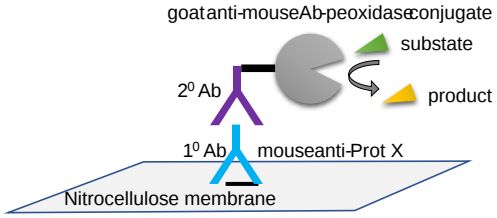


Figure 3.4.22: Developing a Western blot

3D electrophoresis: To detect specific proteins in a 2D electrophoresis experiment, a 3rd dimension of separation, a Western blot, could be performed on the PAGE gel and the nitrocellulose stained with an antibody specific to a target protein. That is illustrated in Figure 3.4.23

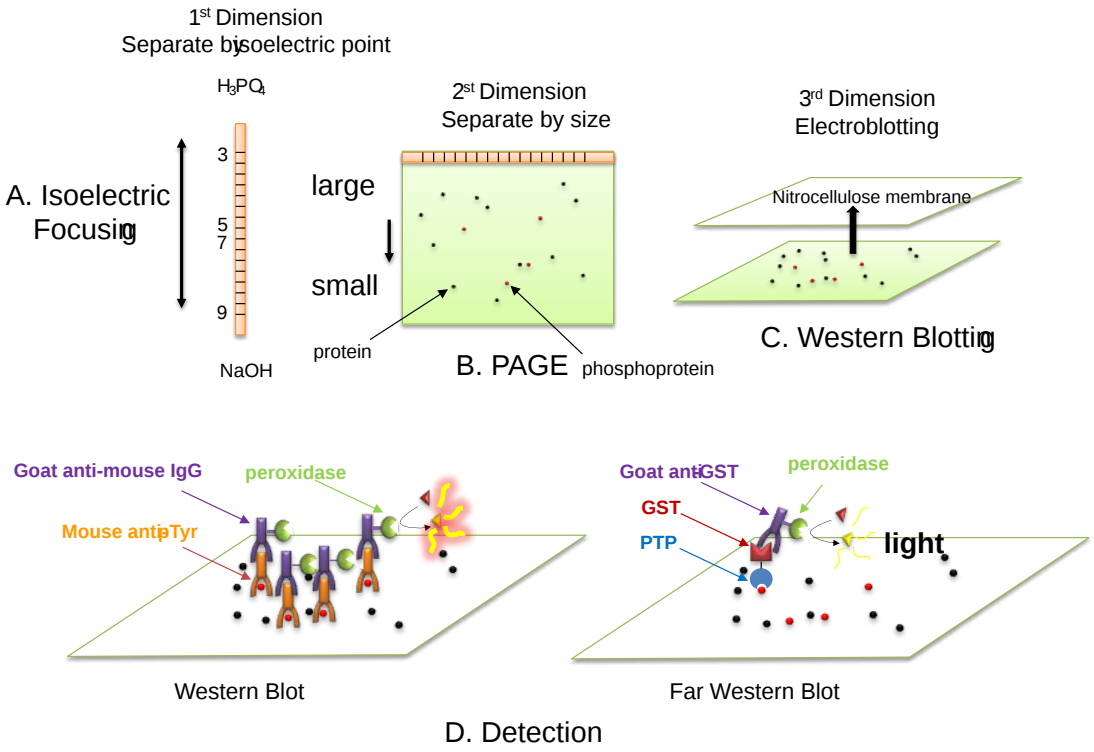


Figure 3.4.23: "3D" electrophoresis - Isoelectric focusing, PAGE and Western blotting

Part A, isoelectric focusing, is followed by a PAGE gel (B). The red dots represent proteins that have undergone a post-translational modification in which a phosphate group has been added to tyrosine side chains (for example). Western blotting is performed in panel C and staining in panel D. The left blot in D uses an antibody that recognizes phosphorylated tyrosine side chains on protein. The right blot is D is sometimes called a Far Western blot. If the protein on the nitrocelluose membrane retains some 3D native structure or can be induced to refold, it can be probed on the blot by a protein that binds to the native form of the protein on the blot. In the example shown in panel D above, the p-Tyr-protein target on the nitrocellulose membrane recognizes a fusion protein of PTP-GST. GST is a protein tag for detection. PTP is a protein tyrosine phosphatase, an enzyme that hydrolyzes p-Tyr on specific phosphorylated target proteins.

3.4.8: References

Molnar, C. and Gair, J. (2013) Antibodies. Chapter in Concepts in Biology, Published by B.C. Open Textbook Project. Available at: https://opentextbc.ca/biology/chapter/23-3-antibodies/

The Human Atlas Project. (2019) Methods. Available at: https://www.proteinatlas.org/learn/method

Uhlén M et al, 2015. **Tissue-based map of the human proteome.** *Science*
PubMed: 25613900 DOI: 10.1126/science.1260419

Thul PJ et al, 2017. **A subcellular map of the human proteome.** *Science.*
PubMed: 28495876 DOI: 10.1126/science.aal3321

Uhlen M et al, 2017. **A pathology atlas of the human cancer transcriptome.** *Science.*
PubMed: 28818916 DOI: 10.1126/science.aan2507

Ahern, K. and Rajagopal, I. (2019) Biochemistry Free and Easy. Published by Libretexts. Available at: https://bio.libretexts.org/Bookshelves/Biochemistry/Book%3A_Biochemistry_Free_and_Easy_(Ahern_and_Rajagopal)/09%3A_Techniques/9.04%3A_Gel_Exclusion_Chromatography.

Magdeldin, S. (2012) Gel Electrophoresis - Principles and Basics. Published by InTech under Creative Commons Attribution 3.0. Available at: https://pdfs.semanticscholar.org/4b93/70ac3946cec6e12c369679c4178a5ef38e61.pdf

Structural Biochemistry/Proteins/X-ray Crystallography. (2018, November 19). *Wikibooks, The Free Textbook Project*. Retrieved 15:40, August 17, 2019 from en.wikibooks.org/w/index.php?title=Structural_Biochemistry/Proteins/X-ray_Crystallography&oldid=3488057.

UCD: Biophysics 200A (2019) "NMR Spectroscopy vs X-ray Crystallography", Chapter published in Current Techniques in Biophysics. Published by Libretexts and available at: https://phys.libretexts.org/Courses/University_of_California_Davis/UCD%3A_Biophysics_200A_-_Current_Techniques_in_Biophysics/NMR_Spectroscopy_vs._X-ray_Crystallography

Wikipedia contributors. (2019, June 27). Protein purification. In *Wikipedia, The Free Encyclopedia*. Retrieved 23:32, July 28, 2019, from en.Wikipedia.org/w/index.php?title=Protein_purification&oldid=903657925

Wikipedia contributors. (2019, February 15). Fast protein liquid chromatography. In *Wikipedia, The Free Encyclopedia*. Retrieved 17:14, August 15, 2019, from en.Wikipedia.org/w/index.php?title=Fast_protein_liquid_chromatography&oldid=883530035

Wikipedia contributors. (2019, July 9). Protein mass spectrometry. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:27, August 16, 2019, from en.Wikipedia.org/w/index.php?title=Protein_mass_spectrometry&oldid=905547289

Wikipedia contributors. (2019, July 8). Peptide synthesis. In *Wikipedia, The Free Encyclopedia*. Retrieved 06:13, August 17, 2019, from en.Wikipedia.org/w/index.php?title=Peptide_synthesis&oldid=905401648

3.5: Antibodies in Quantitation and In Vivo Detection

Written by ?? Needs work

We will describe three different uses of antibodies for the detection and quantitation of analytes.

3.5.1: Enzyme-Linked Immunosorbent Assay (ELISA)

(work derived from the Human Atlas Project)

Since the very first use of antibodies for the detection of antigens, many different technologies have been developed that make use of the antibodies' capability to bind to other molecules. During the 1950s, Yalow and Berson developed a method where radioactivity is used to determine the amount of an analyte in a solution. This 'radioimmunoassay' (RIA), for which Yarlow received the Nobel prize in 1977, was a very sensitive method for the detection of hormones but using radioactivity for antigen detection is not safe and suitable for general use. Hence, an alternative procedure was developed by linking enzymes to antibodies instead of a radioactive molecule, and by adhering molecules to surfaces. This is the basis of the widely-used "enzyme-linked immunosorbent assay" (ELISA). Many variants of experimental procedures have been developed, and it is common to build assays using more than one antibody to detect a target of interest (see Figure 1). To further enhance the possibilities offered by the immunoassay format, applications based on microarrays have been developed which allow the measurement of more than one analyte in a single reaction chamber (see below). Different detection methods are described in Figure 3.5.1.

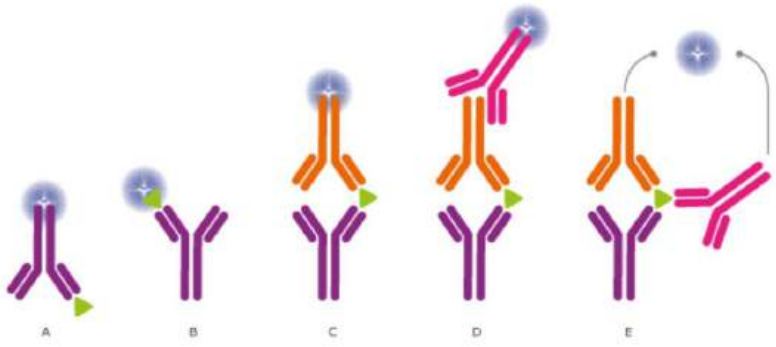


Figure 3.5.1: **Different detection methods for ELISA and Other Immunossays**. (CC BY-SA 3.0; *The Human Atlas Project*)

In ELISA assays, the antibodies may (A) detect an immobilized antigen, (B) capture a labeled antigen, (C) capture an unlabeled antigen and use a second, labeled antibody to detect the captured antigen, or (D) use a third antibody for detection, or even use two antibodies for detection (E). Direct labeling of the antibody or antigen as in (A), (B), and (C) is the simplest and fastest method for detection. Using a secondary antibody as a detection method, as shown in (D) and (E), will further increase the sensitivity and selectivity of the analysis. The method used in (D) also allows greater flexibility, whereas method (E) further increases the specificity, as three antibodies must bind the antigen in order to produce a reporter molecule. Out of the presented assays, the most commonly used concepts are shown in (C) and (D). We will describe how ELISA data is used to determine concentrations of analytes in Chapter 5.

A new era in immunoassays started with the development of a technology called microarrays. The term microarray most commonly describes the ordered organization of small-volume droplets that have dried on a small surface area. The reaction dimensions are miniaturized so that many assays can be performed in multiple samples in parallel. Glass slides can be used and robotic pipettors can deposit very small drops of liquid (1 nL = 10^{-9} L on the glass surface in an ordered fashion, with the spots having sizes of around 0.15 mm. Another common technique for multiplexing is to use even smaller and color-coded particles (diameter of 0.005 mm). These particles can be coated with antibodies to fish out the analyte from the solution.

Microarray assays are used for parallel analysis of DNA and RNA molecules. In addition, multiplexed techniques are used to determine many proteins simultaneously, and to study post-translation modifications such as phosphorylation. Another example is the analysis of antibodies circulating in the blood of patients. Protein microarrays can reveal the interactions of ligands with the whole protein or larger protein fragments, while peptide microarrays are used to detect small peptides (epitopes) of the proteins that bind antibodies. A typical epitope mapping result is shown in Figure 3.5.2 (Edfors et al., 2014).

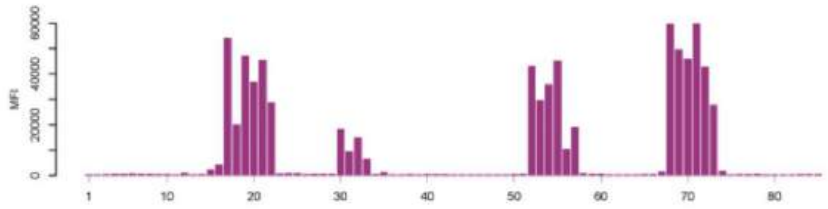


Figure 3.5.2: **Epitope Mapping of Polyclonal Antibodies.** Polyclonal antibodies binding to a peptide array where the result displays four distinct linear epitopes and the consecutive overlapping peptides which are bound. X-axis: peptides, Y-axis: mean fluorescence intensity (MFI). (Edfors et al., 2014) *Image from The Human Atlas Project*

Synthesizing millions of overlapping peptides with only one amino acid residue shift on such arrays enables the mapping of antibody binding regions at high resolution. This gives a very detailed analysis of the linear (continuous) epitopes recognized by an antibody.

3.5.2: Immunohistochemistry - Detecting Proteins in Vivo

(work derived from the Human Atlas Project)

Immunohistochemistry (IHC) is a powerful microscopy-based technique for visualizing cellular components, for instance, proteins or other macromolecules in tissue samples. The strength of IHC is the intuitive visual output that reveals the existence and localization of the target protein in the context of different cell types, biological states, and/or subcellular localization within complex tissues.

The IHC technique was invented during the 1940s (Coons, Creech, & Jones, 1941) and is routinely used as an important tool in health care and pathology for e.g. diagnostic purposes or to stratify patients for optimized treatment regimes. IHC is also widely used in research where molecules of interest are analyzed to study their roles in both healthy and diseased cells and tissues on the molecular, cellular or tissue level. There are many different ways to perform visualization of targets in tissues using IHC or IHC-based methods, and numerous protocols exist for different applications and assays. Even though IHC is generally a robust and established method, new assays often need careful optimization depending on the tissue or on the properties of the target protein, binder-molecule and/or reporter system.

The classical IHC assay is illustrated in Figure 3.5.3 and involves the detection of epitopes expressed by a single protein target within a tissue sample using a "primary antibody" capable of binding those epitopes with high specificity. After the epitope-antibody binding event, a "secondary antibody" capable of binding the primary antibody with high specificity is added. The secondary antibody is coupled to a reporter molecule and after the antibody-antibody binding event, a chemical substrate is added which reacts with the reporter molecule to produce a colored precipitate at the site of the whole epitope-antibody complex.

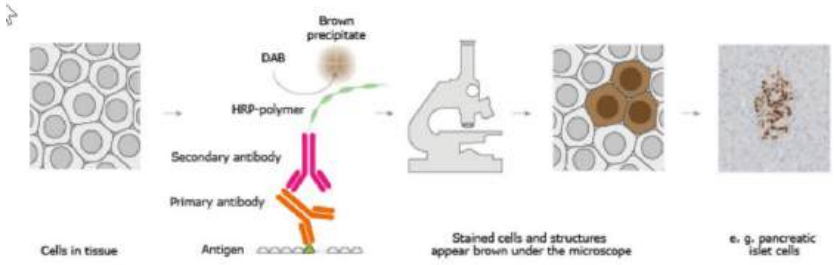


Figure 3.5.3: **The Basic Principle of Immunohistochemistry.** *mage from The Human Atlas Project*

In the schematic illustration, a formalin-fixed paraffin-embedded tissue section is stained using a primary antibody directed toward a specific protein target. A solution containing the primary antibody is added to the tissue section and the antibodies are allowed some time to find and bind to their target. After this step, unbound and surplus antibodies are washed away and the secondary antibody is added. The secondary antibody, which carries a linker molecule with horseradish peroxidase (HRP) enzymes, is also allowed some time to bind to the primary antibody, followed by another washing step. After this, 3,3' Diaminobenzidine (DAB) is added. The HRP enzyme transforms the DAB substrate into a brownish precipitate that is deposited in the tissue at the site of the reaction, thus producing a visual representation of where the primary antibody first bound its target.

Tissue preparation

The tissue plays a central role in the experiment and it is important that it is processed so that epitopes and proper morphology is preserved. The most common processing for IHC is to prepare formalin-fixed paraffin-embedded (FFPE) tissue blocks. The purpose of formalin fixation is to produce chemical cross-linking of proteins within the tissue. This terminates all cellular processes and freezes the cellular components at the place and in the conformation at the time of fixation and also to prevent degradation. After adequate fixation, the tissue is further processed and ultimately embedded in paraffin blocks, which are then sectioned into thin slices (usually 4-10µm) using a microtome. The sections are transferred to glass slides and allowed to adhere prior to further processing.

Other methods for fixation besides formalin are sometimes used. These include other types of aldehydes or using different alcohol solutions. The best choice of fixative is very much dependent on the assay. A common alternative to FFPE is to prepare frozen tissue samples. In this case, the tissue is embedded in a cryoprotective medium and frozen, and fixation is performed post-sectioning. Frozen tissues are sectioned in cryostats and have the advantage of short processing times and of better preservation of sensitive epitopes, but can often be inferior to FFPE tissues in terms of preserving histological morphology.

Antigen (epitope) retrieval

A concern associated with cross-linking fixatives like formalin or the length of time spent in the fixative medium is the masking of epitopes, which can obstruct the primary antibody from binding to its target. Especially with FFPE samples, there is often a need to revert some of the chemical crosslinking and "retrieve" the epitopes before proceeding to the actual IHC. There are several antigen retrieval protocols available and the main strategies include treating the tissue slide with heat, digestive enzymes, detergents, or combinations thereof. The most common method for antigen retrieval in FFPE samples is to pressure-boil the tissue slides in an acidic citrate buffer for around 15-20 minutes.

Antibody binding

The quality and specificity of the binding molecule are crucial for any IHC based technique, and the choice of binder can directly affect the outcome, reliability, and possibly also the interpretation of the assay. Antibodies are by far the most common type of binding-molecule used for IHC, and although most antibodies are able to adequately detect the correct molecule of interest, they may also vary greatly in their specificity for their intended target. Antibodies with high specificity are therefore more reliable when interpreting "on-target" binding, since they produce little or no "off-target" binding or "background". Antibodies that are less specific can produce more off-target binding, and the resulting background will possibly interfere with the correct interpretation of the true on-target signals. There are two main types of antibodies; polyclonal antibodies which are a heterogeneous mix of antibodies that bind different epitopes on the target and monoclonal antibodies which bind the same epitope. Polyclonal antibodies are often very potent due to their ability to detect and bind multiple epitopes on the same target. However, the epitopes they bind are often poorly defined, and with multiple and varying epitope-specificities comes the increased likelihood of off-target binding events and background noise. However, the potency of polyclonal antibodies can be advantageous since the concentration of binding events around the on-target molecule usually outweighs potential background noise. A drawback is that polyclonal antibodies are usually limited resources since they are derived from animal sera. Monoclonal antibodies, by contrast, have more continuity since they can be produced in hybridoma cell lines. Monoclonal antibodies are also often well-defined in terms of epitope binding, but can still generate results that are hard to interpret if the specificity is low or if the target epitope is present in low abundance.

Careful optimization and titration of antibody concentration for each assay are needed, since the result is dependent not only on the antibody's specificity and affinity for the target, but also on the concentration and availability of on-target and potential off-target epitopes present in the sample. Adding too many antibodies to the sample will increase the number of possible low-affinity off-target binding events once the on-target epitope(s) are saturated with binders. By lowering the antibody concentration, off-target binding events become rarer as they usually have lower affinity than on-target binding events. The risk when attempting to reduce background while using a low-affinity antibody is that the on-target signals are concomitantly weakened to the point of providing a false negative result.

Other types of binder molecules sometimes used in IHC-based techniques include affibodies, peptides, antibody fragments or other small molecules.

Detection systems

The whole purpose of performing IHC is to obtain a visual representation of where the target can be found within the experimental tissue and preferably and also gain information about the target's expression pattern among heterogeneous cell populations and/or

subcellular sites. This is exemplified in Figure 3.5.4, which illustrates how different antibodies are used to visualize different cellular or tissue compartments within a complex tissue. To visualize the target-antibody interaction, some kind of detection system that produces an observable stain or signal is needed. The most common method for introducing a detection system to the experiment is to use a secondary antibody that carries a pre-bound reporter molecule, i.e. enzyme or fluorophore. Secondary antibodies are usually targeted specifically towards antibody molecules from different animal species. For example, if the primary antibody is raised in a rabbit, then the secondary antibody must be raised in another animal and targeted specifically towards rabbit antibodies.

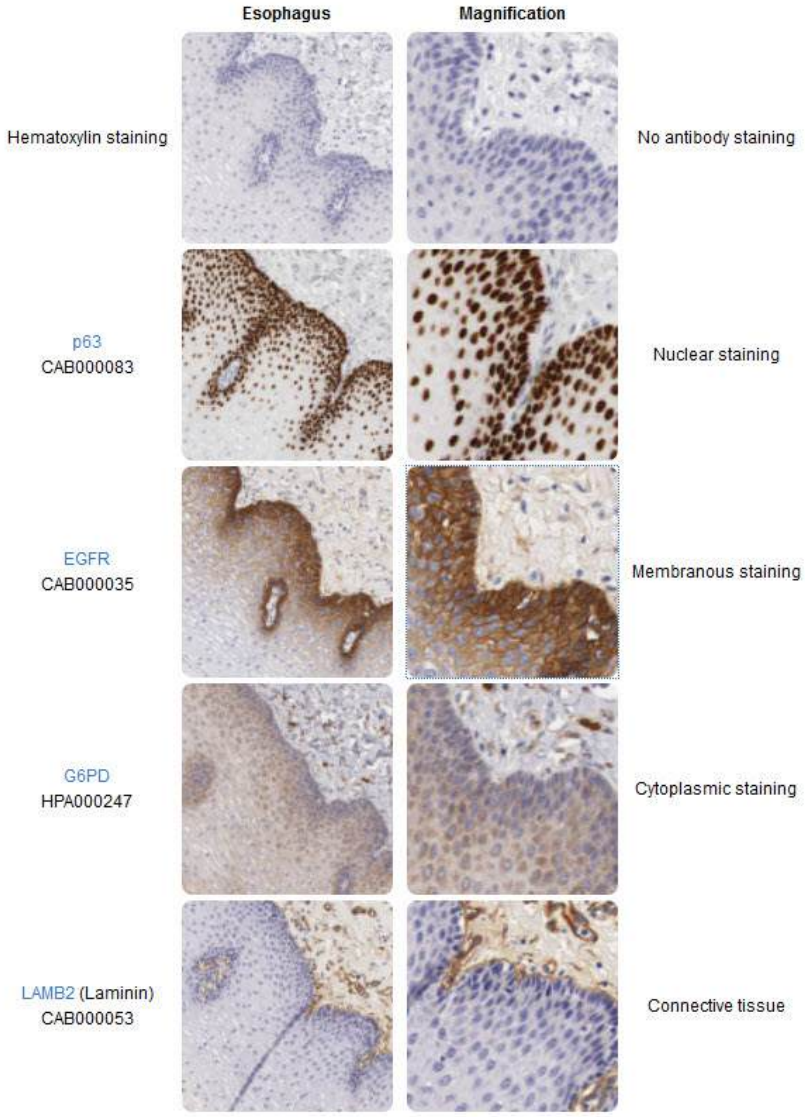


Figure 3.5.4: Visualizing different protein targets in complex tissues. *Image from The Human Atlas Project*

The right column shows a magnification of the corresponding images in the left column. In the IHC image, consecutive sections of a human esophagus stained using four different antibodies allow for direct comparison of different protein expression patterns within the tissue and within subcellular compartments. The top images are only counterstained for hematoxylin for comparison. The p63 antibody stains cell nuclei in a population of cells that reside in the basal part of the esophageal epithelium. The EGFR (Epidermal growth factor receptor) antibody appears to stain the same cell population as p63, but stains cellular membranes instead of nuclei. The G6PD (Glucose-6-phosphate dehydrogenase) antibody stains the cytoplasm of a wider repertoire of esophageal epithelial cells and also cells residing in the connective tissue. The Laminin (LAMB2) antibody stains only cells and structures in the connective tissue underlying the esophagus.

Image from The Human Atlas Project

For FFPE tissue samples the most common detection method is to use enzymatic reactions to generate a colored precipitate at the site of antibody binding. Secondary antibodies with an attached enzyme, e.g. horseradish peroxidase (HRP) or alkaline phosphatase (AP), are capable of converting chromogens like 3,3' Diaminobenzidine (DAB) or 5-bromo-4-chloro-3-indolyl phosphate/ p-nitroblue tetrazolium chloride (BCIP/NBT) into brown or bluish precipitates that are deposited in the tissue at the site of the reaction. Chromogenic stains are observable using light microscopy and are usually very stable over long periods of time, which is beneficial if the experiment needs to be archived or reviewed at a later time point.

For frozen tissue sections, it is more common to use fluorophore-linked secondary antibodies that emit a specific color (usually green, red, or blue) when excited by the correct wavelengths of light. Moreover, fluorophores are usually not stable for long periods of time. However, the benefit of using fluorophores is that they provide an easy method for performing double-labeling experiments where several antibodies directed toward multiple targets are assayed in the same sample. The secondary antibodies need to be targeted towards different primary antibodies and also to be coupled to different fluorophores. The different secondary antibodies are then observed separately by exciting them sequentially with different wavelengths of light. These different excitation results are saved as separate images (or color channels) and may later be overlaid to infer protein co-localizations etc.

Using reporter-carrying secondary antibodies for detection is in itself an amplification step since several secondary antibodies are able to bind a single primary antibody, but sometimes further amplification steps are desired to increase the signal and sensitivity of the experiment. In such cases, the secondary antibody may instead carry "linker molecules", for instance biotin polymers, which are able to recruit a larger number of reporter molecules in subsequent steps. This strategy for amplifying signals is useful for both enzymatic and fluorescent detection methods.

Counterstaining

Immunohistochemical staining using chromogens offers benefits from having a counterstain applied that enhances the contrast and facilitates the observation of histological features. The most common type of counterstain used for FFPE samples is hematoxylin, which stains cellular cytoplasm with a pale bluish color, and stain cell nuclei in a darker bluish nuance. Fluorescent stainings are usually not counterstained with hematoxylin, since the detection method is not based on light microscopy. Instead, the most common way to obtain counterstaining for fluorescence is to label cell nuclei by adding fluorescent dyes that bind nucleic acids.. After the actual immunohistochemical reaction, the only remaining steps are to use a coverslip to seal and protect the sample and for long-term storage. The most common way is to "glue" the coverslip to the sample using commercially available purpose-made resins.

Add figure

Figure 3.27 **Endothelial cells under the microscope.** Nuclei are stained blue with DAPI, microtubules are marked green by an antibody bound to FITC and actin filaments are labeled red with phalloidin bound to TRITC. Bovine pulmonary artery endothelial cells

Image from NIH ImageJ-Programmpaket

Specific examples

IHC is widely used in both research and clinical practice. The Human Protein Atlas (HPA) project is a prime example of how high-throughput IHC is used to achieve large-scale mapping of the human proteome in a multitude of tissues, cancers, and cells. In the HPA project, a streamlined in-house large-scale antibody production chain facilitates the generation of specific antibodies, which after passing basic characterization and validation regimes, are used to systematically stain tissue microarrays containing hundreds of tissue cores within a single experiment. The system for IHC employed by HPA relies heavily on the standardization of protocols and automatization using machines, but the evaluation of the optimal titration for each antibody is performed manually before the antibody is approved for staining on the full set of tissues. Each stained tissue core is annotated with respect to immunohistochemical staining in tissues and cell types, and thereafter published as a high-resolution image on the web portal to be freely viewed by anyone.

In clinical practice, IHC is mainly used within pathology to aid physicians to evaluate tissue specimens with respect to healthy and or diseased states, to set diagnoses, and to define the molecular subtype of different types of cancer. A specific example where IHC is used diagnostically is when pathologists are presented with a metastatic tumor sample and the tissue origin of the primary tumor is unknown. In these cases, pathologists use a panel of different antibodies that target tissue-specific proteins, such as prostate-specific antigen for prostate cancer, and estrogen receptor for gynecological cancers, or cytokeratin 20 for gastrointestinal cancers (Gremel et al., 2014). Once a broad classification is made, additional tissue-specific antibodies are used to further pinpoint the origin of the primary tumor. This information is useful for choosing the best or most appropriate strategy for drug therapy and/or to locate the primary tumor for radiation therapy and/or surgery.

3.5.3: References

Molnar, C. and Gair, J. (2013) Antibodies. Chapter in Concepts in Biology, Published by B.C. Open Textbook Project. Available at: https://opentextbc.ca/biology/chapter/23-3-antibodies/

The Human Atlas Project. (2019) Methods. Available at: https://www.proteinatlas.org/learn/method

Uhlén M et al, 2015. **Tissue-based map of the human proteome.** *Science*
PubMed: 25613900 DOI: 10.1126/science.1260419

Thul PJ et al, 2017. **A subcellular map of the human proteome.** *Science.*
PubMed: 28495876 DOI: 10.1126/science.aal3321

Uhlen M et al, 2017. **A pathology atlas of the human cancer transcriptome.** *Science.*
PubMed: 28818916 DOI: 10.1126/science.aan2507

Ahern, K. and Rajagopal, I. (2019) Biochemistry Free and Easy. Published by Libretexts. Available at: https://bio.libretexts.org/Bookshelves/Biochemistry/Book%3A_Biochemistry_Free_and_Easy_(Ahern_and_Rajagopal)/09%3A_Techniques/9.04%3A_Gel_Exclusion_Chromatography.

Magdeldin, S. (2012) Gel Electrophoresis - Principles and Basics. Published by InTech under Creative Commons Attribution 3.0. Available at: https://pdfs.semanticscholar.org/4b93/70ac3946cec6e12c369679c4178a5ef38e61.pdf

Structural Biochemistry/Proteins/X-ray Crystallography. (2018, November 19). *Wikibooks, The Free Textbook Project*. Retrieved 15:40, August 17, 2019 from en.wikibooks.org/w/index.php?title=Structural_Biochemistry/Proteins/X-ray_Crystallography&oldid=3488057.

UCD: Biophysics 200A (2019) "NMR Spectroscopy vs X-ray Crystallography", Chapter published in Current Techniques in Biophysics. Published by Libretexts and available at: https://phys.libretexts.org/Courses/University_of_California_Davis/UCD%3A_Biophysics_200A_-_Current_Techniques_in_Biophysics/NMR_Spectroscopy_vs._X-ray_Crystallography

Wikipedia contributors. (2019, June 27). Protein purification. In *Wikipedia, The Free Encyclopedia*. Retrieved 23:32, July 28, 2019, from en.Wikipedia.org/w/index.php?title=Protein_purification&oldid=903657925

Wikipedia contributors. (2019, February 15). Fast protein liquid chromatography. In *Wikipedia, The Free Encyclopedia*. Retrieved 17:14, August 15, 2019, from en.Wikipedia.org/w/index.php?title=Fast_protein_liquid_chromatography&oldid=883530035

Wikipedia contributors. (2019, July 9). Protein mass spectrometry. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:27, August 16, 2019, from en.Wikipedia.org/w/index.php?title=Protein_mass_spectrometry&oldid=905547289

Wikipedia contributors. (2019, July 8). Peptide synthesis. In *Wikipedia, The Free Encyclopedia*. Retrieved 06:13, August 17, 2019, from en.Wikipedia.org/w/index.php?title=Peptide_synthesis&oldid=905401648

3.2: The Structure of Proteins- An Overview

3.2.1: Peptide Bond Formation and Primary Protein Structure

Proteins are polymers of amino acids that fold into shapes that confer function on the proteins. In biological systems, the amino acids are linked together by a large ribonucleic acid/protein nanoparticle called the ***ribosome***. Thus, as the amino acids are linked together to form a specific protein, they are placed within a very specific order that is dictated by the genetic information contained within a specific type of RNA called messenger RNA (mRNA). The mRNA sequences are encoded in the genomic DNA sequence. The specific ordering of amino acids is known as the protein's ***primary sequence***. The translation mechanism used by the ribosome to synthesize proteins will be discussed in detail in Chapter 26.

The amino acids are linked together using ***dehydration synthesis*** (loss of water) reaction that connects the carboxylic acid of the upstream amino acid with the amine functional group of the downstream amino acid to form an amide linkage (Figure 2.10). You will remember from other chemistry courses that the formation of an amide from a carboxylic acid (thermodynamically stable) and an amine requires activation of the carboxylic acid end to form a derivative with a better leaving group. This carbonyl of the modified end serves as an electrophile in the attack of the amine nitrogen, a nucleophile) in a nucleophilic substitution reaction. The activation reaction, which we will discuss in subsequent chapters involves the transfer of a phosphate from a phosphoanhydride, ATP, to the carboxylic acid group to form a mixed anhydride with the phosphate serving as a leaving group. Note that the reverse reaction is ***hydrolysis*** and requires the incorporation of a water molecule to separate two amino acids and break the amide bond. Notably, the ***ribosome*** serves as the enzyme that mediates the dehydration synthesis reactions required to build protein molecules, whereas a class of enzymes called ***proteases*** is required for protein hydrolysis.

Within protein structures, the amide linkage between amino acids is known as the ***peptide bond***. Subsequent amino acids will be added onto the carboxylic acid terminal of the growing structure. Proteins are always synthesized in a directional manner starting with the amine and ending with the carboxylic acid tail. New amino acids are always added onto the carboxylic acid tail, never onto the amine of the first amino acid in the chain. The directionality of protein synthesis is dictated by the ribosome. Figure 3.2.1 below shows an overly simplistic version of the reaction that produces the amide bond.

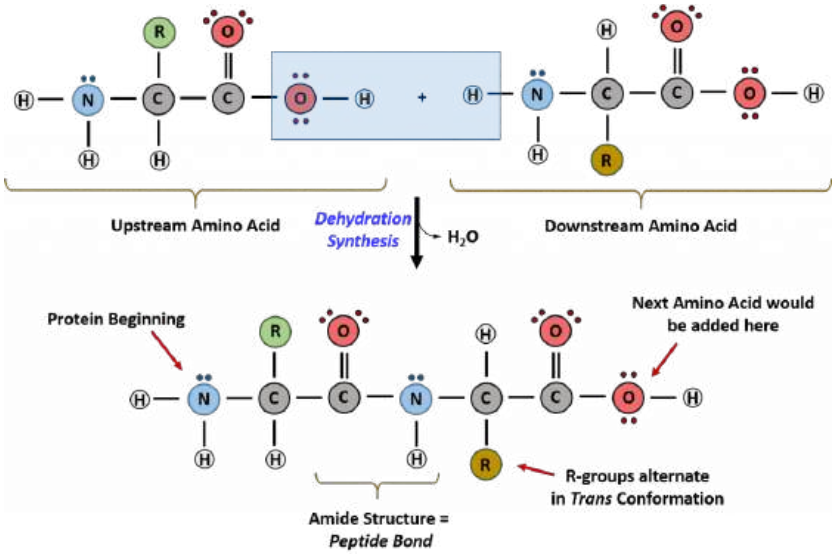


Figure 3.2.1: Formation of the Peptide Bond. The addition of two amino acids to form a peptide requires dehydration synthesis. (Copyright; author via source)

Please note two features of the reaction as shown in the diagram:

1. The activation step (phosphorylation of the carboxylic acid end of the amino acid by ATP) is not shown.
2. The reaction is shown with an unlikely protonation state. If the carboxyl group is protonated, which would occur at a low pH, the amine would also be protonated and should correctly be shown as RNH_3^+. The protonation state in the above figure was chosen to emphasize the loss of H_2O (dehydration) in the reaction. Many textbooks that aren't rigorously based in chemistry show unlikely protonation states for this reaction. By discussing this now, we hope to highlight common mistakes and misconceptions found in many resources.

Figure 3.2.2 shows a generic structure from a longer peptide or protein.

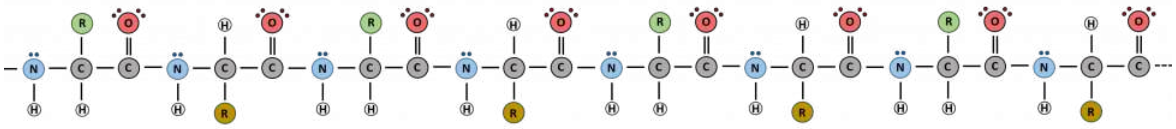


Figure 3.2.2: Peptides and Proteins are macromolecules built from long chains of amino acids joined together through amide linkages. The order and nature of amino acids in the primary sequence of a protein determine the folding pattern of the protein based on the surrounding environment of the protein

Proteins range in size from around 50 amino acids in length to the largest known protein, titin (aka called connectin), a muscle protein. The human version has over 34,000 amino acids and a molecular weight > 3.9 million! Some consider structures with fewer than 50 amino acids to be **peptides** (Figure 2.13). Others suggest that structures with 40-50 amino acids should be considered **small proteins**. One way to differentiate them is by how they are synthesized in vivo. Storz et al consider polypeptides to be small proteins if they are encoded in the genome by a continuous stretch of DNA base in an "open reading frame" (doi: 10.1146/annurev-biochem-070611-102400). Peptides, on the other hand, could be structures that are:

- "intrinsically disordered" with no definite fold,
- derived from proteins by proteolysis, and/or
- not synthesized by ribosomes

As genomes were sequences and annotated, some arbitrary parameters were set. For the yeast genome, annotated proteins were defined as those made from an open reading frame leading to a polypeptide of 100 amino acids (which on average has a molecular weight of 11,000). If no cutoff was used, the number of proteins encoded by the genome would be huge. Submissions of DNA sequences to the NIH GenBank must encode proteins no smaller than about 66 amino acids (MW about 7250). Even this ignores small proteins that have been isolated and characterized from cells. So the cutoff of 50 amino acids (MW about 5500) derived from open reading frames seems like the best arbitrary cutoff going from peptides to proteins.

Due to the large pool of amino acids that can be incorporated at each position within the protein, there are billions of different possible protein combinations that can be used to create novel protein structures! For example, think about a tripeptide made from this amino acid pool. At each position, there are 20 different options that can be incorporated. Thus, the total number of resulting tripeptides possible would be 20 X 20 X 20 or 20^3, or 8,000 different tripeptide sequences! Now think about how many options there would be for a small peptide containing 40 amino acids. There would be 20^{40} options, or a mind-boggling 1.09 X 10^{52} potential sequence options! Each of these options would vary in the overall protein shape, as the nature of the amino acid side chains helps to determine the interaction of the protein with the other residues in the protein itself and with its surrounding environment.

Nearly 200,000 3D structures of biomacromolecules are known and over a million have been determined using artificial intelligence computer programs. How can we simplify our understanding of the diversity of protein structures? Is each structure completely unique? What do they have in common?

To simplify and inform our understanding of the diversity of biological organisms, we place them into groups (from domains and kingdoms to genuses and species), based on common characteristics. Likewise, proteins are divided into a hierarchy of structures with increasing information content. This overview describes the four basic levels of protein structure, primary (1^0), secondary (2^0), tertiary (3^0) and quaternary (4^0). Each will be probed in greater deal in the next chapter. These classes of structures will be illustrated below with a protein named **hydroxynitrile lyase** (5Y02). (This protein has been simplified to illustrate key features of structure as will be described at the end).

3.2.2: Primary (1^0) Structure: the amino acid sequence of a protein.

The primary (1^0) structure of a protein is simply the amino acid sequence of a protein written from N- to C-terminal. It does not require visualization to describe it. Consider two different short continuous sequences from the hydroxynitrile lyase:

- Gln-Lys-Gln-Ile-Asp-Gln-Ile or in single letter code QKQIDQI. This is the sequence for amino acids 20-26 in the protein. This stretch of 1^0 structure has multiple repeated amino acids.
- Asp-Leu-Gly-Pro-Ala-Val or in single letter code DLGPAV. This is the sequence for amino acids 48-53 in the protein. This stretch of 1^0 structure does not contain repetitive amino acids.

A 2-D line drawing of the sequence offers more information about the sequence but does not provide information about the actual conformation of these sections of 1^0 structure within a given protein. These can be shown in Figure 3.2.3 in which the overall structure of the protein is shown in grey sticks with short stretches of primary structure shown in colored spacefill and 2D line drawings.

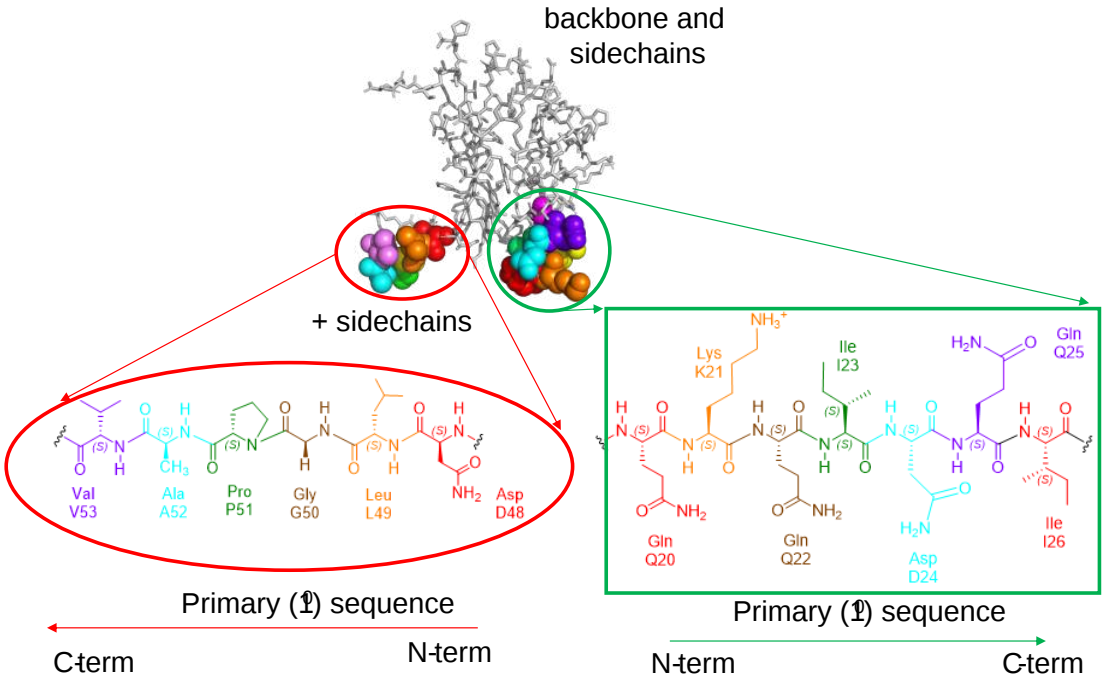


Figure 3.2.3: Alternative renderings of a "primary" sequence within a protein

3.2.3: Secondary (2^0) and Tertiary (3^0) Structures

Secondary (2^0) structures are repetitive structures within a protein held together by hydrogen bonds between amide Hs and carbonyl Os in the backbone main chain atoms. It's most easily examined through the specific rendering of the overall **tertiary (3^0)** or 3-D structure of the protein. Five different renderings showing the 3D (the **tertiary** structure of the protein are shown in Figure 3.2.4.

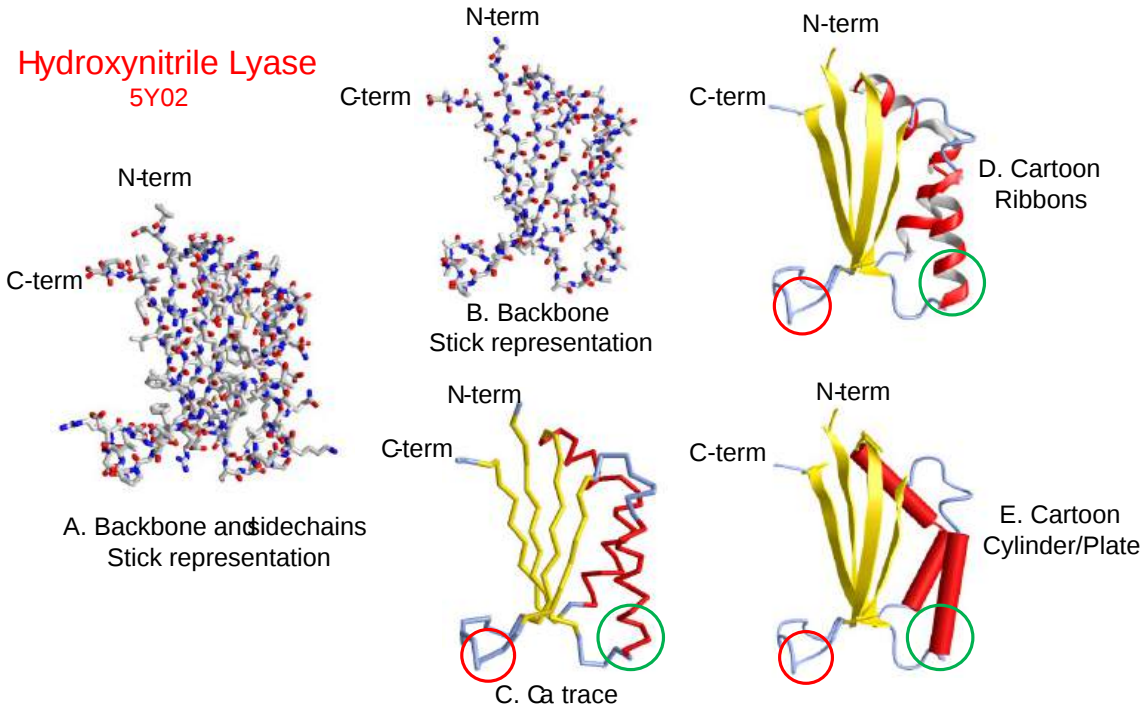


Figure 3.2.4: Alternative renderings of a protein

Representation **A** shows a stick drawing of the protein with red indicating bonds to oxygen and blue bonds to nitrogen. No bonds to hydrogen are shown as these are too small to be detected using common techniques used to determine structures of such large molecules. It looks like a complicated mess of bonds so understanding unique features with the 3^0 structure of the protein are difficult. Representation **B** shows just the backbone of the protein. The outline of how the protein twists and turns in space becomes more evident. The N- and C-terminal ends are more clearly seen.

Representation **C** shows just the bonds connecting the alpha C atoms of each amino acid. The protein's overall topology is now clearly evident. If you follow the chain from the N- to C-terminal ends, it should be evident that there are regularities in the conformations of the protein chain. The individual yellow zig-zags are called **beta** strands. These strands appear elongated and aligned with other beta strands to form a larger **beta-sheet**. The sheet is held together through hydrogen bonds between backbone amide Hs and carbonyl Os on adjacent strands. Beta strands are a type of **secondary structure**.

The red zig-zag lines represent another type of secondary structure called the **alpha-helix**. The helix is again held together by hydrogen bonds between amide Hs and carbonyl Os within a single continuous strand. The backbone of the alpha helix appears less elongated than in a beta-strand as it is wound into a coil (the alpha helix) along a central axis. If you took tweezers (using atomic force microscopy) and pulled on the helix, it could stretch and become more elongated like the beta strands.

The rest of the protein alpha carbon chain shown in blue is less regular. However, it is still ordered as it propagates through space in what is termed a **random coil**. In the protein, it adopts mostly a fixed conformation but it has more conformational flexibility than alpha helices and beta sheets. The alpha helices and beta strands (sheets) are examples of secondary structures.

Representations **D** and **E** are cartoon drawings clearly showing the alpha helices (red) and beta stands and sheets (yellow). It would be extremely difficult to discern alpha or beta secondary structures with stick representations showing all the bonds in a protein. Some of the atoms must be visually (not literally) removed to see the repetitive propagation of the protein backbone through the overall structure. If your goal was to understand the disposition of side chains in a small part of a protein, a cartoon view by itself would not be useful. Modeling programs allow mixed rending of a protein to include both cartoon and stick representations together.

Secondary structures, held together by hydrogen bonds between backbone atoms are characterized by repetitive changes in the angle of propagation of the chain between connected amino acids in an alpha helix and strand. In a given beta-strand, the relative change in the angle of propagation is nearly 0^0 compared to a much large angular change required to bend the amino acid backbone

into an alpha helix. Here is the IUPAC definition of secondary structure. We added the word "repetitive" to clearly show that random coils are not an example of secondary structures.

Definition: Secondary Structure (from the IUPAC Gold Book)

The [repetitive] conformational arrangement (α-helix, β-pleated sheet, etc.) of the backbone segments of a macromolecule such as a polypeptide chain of a protein without regard to the conformation of the side chains or the relationship to other segments.

3.2.4: Quaternary Structure

Separate protein chains often interact through noncovalent interactions and sometimes through disulfide bond formation between free cysteine side chains on different chains to form dimers, trimers, tetramers, octamers, etc. Dimers can be homodimers (if the two chains are identical) or heterodimers (if they are different). The example we used in the section, hydroxynitrile lyase, forms a homodimer, as shown in Figure 3.2.5. The left images shows a cartoon version, with one monomer in orange and the other identical monomer in green. The images to the right shows a translucent surface representation of the dimer, with the car

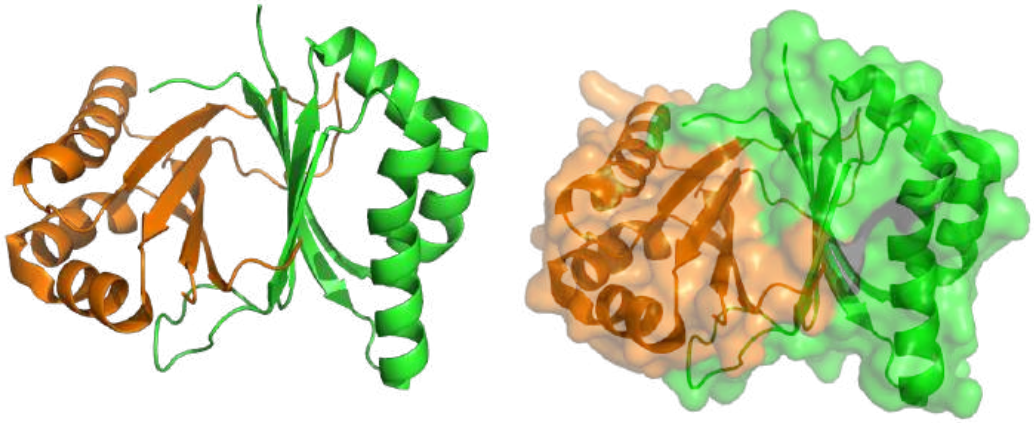


Figure 3.2.5: Dimer of hydroxynitrile lyase

The mixed-rendered image on the right shows a translucent surface image of each monomer and underneath the cartoon image

- primary structure: the linear amino acid sequence of a protein
- secondary structure: regular repeating structures arising when hydrogen bonds between the peptide backbone amide hydrogens and carbonyl oxygens occur at regular intervals within a given linear sequence (strand) of a protein or between two adjacent strands

Disulfide bonds within individual chains and between them stabilize both tertiary and quaternary structures of both peptides and proteins. These are illustrated in Figure 3.2.6.

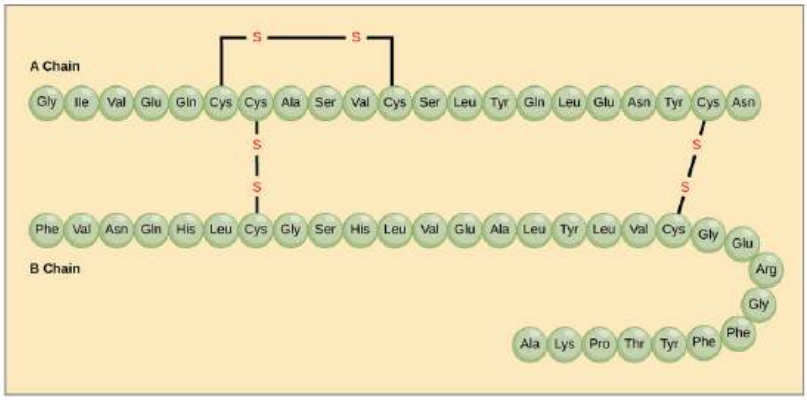


Figure 3.2.6: **Disulfide Bonds.** Disulfide bonds are formed between two cysteine residues within a peptide or protein sequence or between different peptide or protein chains. In the example above the two peptide chains that form the hormone insulin are depicted. Disulfide bridges between the two chains are required for the proper function of this hormone to regulate blood glucose levels. Image by: CNX OpenStax via Wikimedia Commons

3.2.5: Protein Shape and Function

The primary structure of each protein leads to the unique folding pattern that is characteristic of that specific protein. In summary, the primary sequence is the linear order of the amino acids as they are linked together in the protein chain as shown in Figure 3.2.7. In the next section, we will discuss protein folding that gives rise to secondary, tertiary, and sometimes quaternary protein structures.

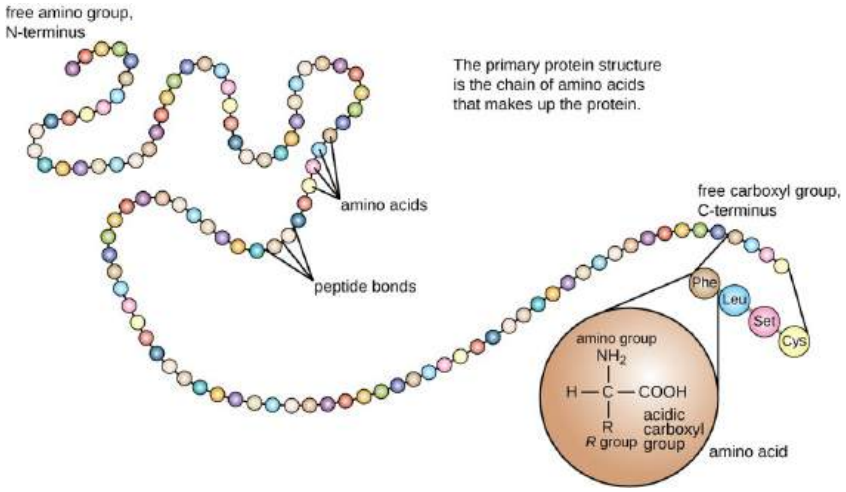


Figure 3.2.7: Primary protein structure is the linear sequence of amino acids. (credit: modification of work by National Human Genome Research Institute)

3.2.6: References

OpenStax, Proteins. OpenStax CNX. Sep 30, 2016 http://cnx.org/contents/bf17f4df-605c-4388-88c2-25b0f000b0ed@2.

File:Chirality with hands.jpg. (2017, September 16). *Wikimedia Commons, the free media repository*. Retrieved 17:34, July 10, 2019 from commons.wikimedia.org/w/index.php?title=File:Chirality_with_hands.jpg&oldid=258750003.

Wikipedia contributors. (2019, July 6). Zwitterion. In *Wikipedia, The Free Encyclopedia*. Retrieved 21:48, July 10, 2019, from en.Wikipedia.org/w/index.php?title=Zwitterion&oldid=905089721

Wikipedia contributors. (2019, July 8). Absolute configuration. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:28, July 14, 2019, from en.Wikipedia.org/w/index.php?title=Absolute_configuration&oldid=905412423

Structural Biochemistry/Enzyme/Active Site. (2019, July 1). *Wikibooks, The Free Textbook Project*. Retrieved 16:55, July 16, 2019 from en.wikibooks.org/w/index.php?title=Structural_Biochemistry/Enzyme/Active_Site&oldid=3555410.

Structural Biochemistry/Proteins. (2019, March 24). *Wikibooks, The Free Textbook Project*. Retrieved 19:16, July 18, 2019 from en.wikibooks.org/w/index.php?title=Structural_Biochemistry/Proteins&oldid=3529061.

Fujiwara, K., Toda, H., and Ikeguchi, M. (2012) Dependence of a α-helical and β-sheet amino acid propensities on teh overall protein fold type. BMC Structural Biology 12:18. Available at: https://bmcstructbiol.biomedcentral.com/track/pdf/10.1186/1472-6807-12-18

Wikipedia contributors. (2019, July 16). Keratin. In *Wikipedia, The Free Encyclopedia*. Retrieved 17:50, July 19, 2019, from en.Wikipedia.org/w/index.php?title=Keratin&oldid=906578340

Wikipedia contributors. (2019, July 13). Alpha-keratin. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:17, July 19, 2019, from en.Wikipedia.org/w/index.php?title=Alpha-keratin&oldid=906117410

Open Learning Initiative. (2019) Integumentary Levels of Organization. Carnegie Mellon University. In Anatomy & Physiology. Available at: https://oli.cmu.edu/jcourse/webui/syllabus/module.do?context=4348901580020ca6010f804da8baf7ba.

Wikipedia contributors. (2019, July 16). Collagen. In *Wikipedia, The Free Encyclopedia*. Retrieved 03:42, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Collagen&oldid=906509954

Wikipedia contributors. (2019, July 2). Rossmann fold. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:01, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Rossmann_fold&oldid=904468788

Wikipedia contributors. (2019, May 30). TIM barrel. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:46, July 20, 2019, from en.Wikipedia.org/w/index.php?title=TIM_barrel&oldid=899459569

Wikipedia contributors. (2019, July 16). Protein folding. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:30, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Protein_folding&oldid=906604145

Wikipedia contributors. (2019, June 11). Globular protein. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:49, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Globular_protein&oldid=901360467

Wikipedia contributors. (2019, July 11). Intrinsically disordered proteins. In *Wikipedia, The Free Encyclopedia*. Retrieved 19:52, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Intrinsically_disordered_proteins&oldid=905782287

3.6: Chapter 3 Questions

Section 1 Questions:

Q1) A small protein has the amino acid sequence below:
$C_1NVC_2KYAPITALYC_3AEEC_4QQH$
There are four cysteine residues in the sequence and are designated by the subscripts. The protein is digested with chymotrypsin, and the resulting slurry is followed by an anionic exchange column.
a) Where are the disulfide bonds in this structure? Are they disrupted by any of the above treatments?
b) Identify the peptide fragments created by the chymotrypsin digest. Which fragment will elute first from the anion exchange column?

A1)
a) The two disulfides are occurring between C_1/C_2 and C_3/C_4. Due to the proline in the center of this structure, we know there is likely not a folding pattern that would let another combination of disulfides to occur. No, the disulfides are not disrupted by the chymotrypsin treatment.
b) Since chymotrypsin cleaves aromatic amino acids at the C-terminus, the fragment following digestion would be:
1) C_1NVC_2KY
2) APITALY
3) C_3AEEC_4QQH
When determining which will elute off an anionic column (a positively charged bead that binds to negative residues), we need to determine the overall charge of the peptide fragments. The lysine (K) in fragment 1 will give an overall positive charge. Fragment 2 contains no charged amino acids, and the two glutamic acid (E) residues in fragment 3 will yield an overall negative charge. Since we know the column will bind most tightly to negative residues and repel positives, Fragment 1 should be eluted first, followed by fragment 2, and lastly fragment 3.

Q2) In most proteases, there is a Ser or Cys residue in the active site. Site-directed mutagenesis experiments have shown that active site Ser can be replaced with a Cys, and vice versa, with the protease, still remaining catalytically active.
a) Based on the structure of Cys and Ser, suggest an explanation as to why this could be.
b) There are other amino acids with R-groups that have similar to Ser and Cys. Hypothesize if a site-directed mutagenesis experiment changing a Ser/Cys to Thr, Tyr or Met would still retain catalytic function. Explain your reasoning.

A2) a)For both amino acids, their R-groups contain elements with unbonded lone pairs of electrons. These lone pairs on Ser (:OH) and Cys (:SH) can each act as nucleophiles and allow the protease to engage in SN-2 type reactions for proteolytic cleavage.
b) When considering the likelihood of Thr, Tyr or Met to maintain activity of the active site, the most likely candidate would be threonine. Threonine is structurally the most similar to serine and cysteine and therefore has the highest chance of not causing any steric hindrances in the active site. Threonine is however larger, so one could potentially expect a decrease in activity. Tyrosine while still containing a hydroxyl group, has a large aromatic group that would likely disrupt the interactions of the binding pocket and substrate recognition. Similarly, methionine does contain a sulfur group, but the length of the side chain most likely will cause an inhibition of substrate binding/recognition.

Q3) What pH would you use for an ion-exchange chromatography column to separate the following small peptides? Explain your answer.
Peptide 1) RGAG
Peptide 2) RGAE
Peptide 3) HGAE
Peptide 4) EGAE

A3) First, assign the formal charge on each peptide:
P1) +1
P2) Neutral
P3) -1* Note the histidine here!
P4) -2

Next to ensure that each peptide is able to maintain a unique charge, we need to assess if any pH changes could affect the net charge of the peptides and the pKas of the relevant amino acids that can contribute to the charge. R=12.48, E=4.25, H=6.00. Now, to ensure arginine is positive, glutamic acid is negative, and histidine is neutral, the pH of the column would need to be higher than 6 but less than 12.48.

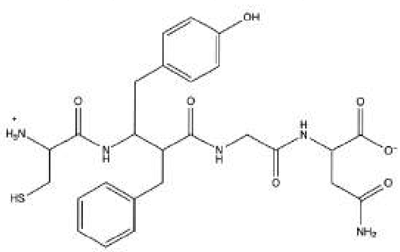
Q4) Match the three letter code to the one letter code for the following amino acids.
Tyr W
Ala P
Asp A
Asn Y
Pro T
Trp D
Thr N

A4) Tyr = Y; Ala = A; Asp = D; Asn = N; Pro = P; Trp = W; Thr = T

Q5) Vasopressin is a small peptide hormone with its C-terminus converted to an amide. It is produced in the hypothamus, but ultimately becomes functionally active in the pituitary gland. One in the bioactive form, vasopressin aids in kidney function by increasing water retention, as well as increases blood pressure in the arteries. (Möller and Mari, *Biochem J* 2007; doi:10.1042/BJ20061480)
The peptide sequence for the mammalian protein is given below:
$C_1YFGNC_2PRG\text{-}NH_2$

a) Write the sequence of vasopressin in the 3 letter code
b) Draw out the structure of the first 5 amino acids of vasopressin
c) What is the overall charge of vasopressin at pH 5? Assume the carboxylate group has a pKa of 4.0 and the amino group has a pKa of 10.
d) What is the isoelectric point of vasopressin?

A5) a) Cys-Tyr-Phe-Gly-Asn-Cys-Pro-Arg-Gly

b) 

c) When we are looking at the amino acids whose R-groups can affect the overall charge of the peptide, for vasopressin that would be Tyr, and Arg. At a pH of 6 there is not enough hydroxide present to deprotonate the -OH of Tyr, so there is no charge as its pKa is 10.46. However, the pKa of Arginine is 12.48, and therefor at a pH of 6, the amine of the R-group remains protonated. Additionally, the N-terminal amino group at pH 6 contains a +1 charge and without a carboxylate group on the C-terminus to balance the charges, an additional +1 is added to the peptide. Therefore, the overall charge of the peptide is +2.
Reminder: Cys in disulfide!
d) When considering which amino acid's pKa to choose for pI calculation, you first need to determine which in the peptide can contribute to a charge on the molecule. For vasopressin, that would be Cys (pKa =8.37), Tyr (pKa = 10.46), Arg (pKa = 12.48), and the NH_3^+ (pKa = 10) at the N-terminus. Remember! The C-terminus of vasopressin is not a free carboxyl! Rather, it is aminated resulting in no net charge. So, we want to determine which to pKa values are the closest and straddle the pH at which the molecule has a net neutral charge.
With the pH < 8.37, the peptide has a net charge of +2. However when the pH is 10 < pH > 8.37, the cysteine residues are now deprotonated resulting in an S^-. When we factor in both cysteine residue's negative charge and the +1 from the N-terminus and the +1 from arginine, the peptide is now at a net neutral!
So, to calculate the pI, we need to add the two pKa values that straddle the net neutral pH (8.37 and 10) and find the average between them. This results in a pI of 9.18!|

Section 2 Questions:

Q1) When a growing peptide chain is being synthesized by the ribosome, what terminus (amino or carboxy) is added on to? Also, which amino acid is always the "starting" amino acid for a polypeptide?
A1) Much like how DNA is always synthesized from 5' → 3', proteins are always synthesized from the amino terminus to the carboxy terminus, beginning with the start codon Methionine.

Q2) Match the type of protein structure to its definition:
a) Primary Structure (1°) 1) This structure can also be called a homodimer or heterodimer, results because of two proteins forming interactions
b) Secondary Structure (2°) 2) A growing polypeptide synthesized directly out of the ribosome
c) Tertiary Structure (3°) 3) The 3D structure of a protein, fully synthesized and correctly folded
d) Quaternary Structure (4°) 4) The protein structure formed using the R-groups of the polypeptide to create α-helices and β-sheets.
A2)
a = 2
b = 4
c = 3
d = 1

Section 3 Questions

Q1) You are considering choosing between traditional centrifugation and density centrifugation for the following scenarios. Explain which method would result in the best result for your experiment and why.
a) You want a crude cell pellet free of all supernatant.
b) From a plant cell, you want to separate the nucleus from the chloroplast.
c) You want to separate the nuclear envelope from the nucleus (hint: think about the structure of the nucleus!)

A2)
a) Traditional centrifugation. All that is needed for this experiment is the cell material to be pelleted away from the supernatant, so traditional centrifugation is sufficient.
b) Density centrifugation. To separate on an organellar level, density centrifugation should be used to increase efficiency. Bonus fun fact! Choosing an osmotically inert material such as Percoll can improve separation by not inducing hyper- or hypotonic lysis!
c) Density centrifugation. To fractionate on a suborganellar level, density centrifugation is a must. Proceeded by the appropriate experiments, an osmotic material such as sucrose can be used to fraction the nuclear envelope away from the nucleus.

Q2) You are creating a cell-free extract of Arabidopsis proteins that you want to keep for extended storage at -80°C. However, when thawed you still want the protein to remain functional for future assays and you know just adding glycerol will cause the protein to denature.
a) What is a method you can use to add glycerol to your protein extract while keeping it stable in solution? Explain why.

A2)
To increase the concentration of glycerol in the buffer without adding directly, dialysis should be used. This allows for an exchange of buffer components, ultimately bringing both solutions to equilibrium. So, if you have a higher concentration of glycerol in the dialysis solution than in the cell-free extract, dialysis will cause the concentration of glycerol in the cell-free extract to increase slowly over time, preventing protein precipiation.

Q3) The components of a cell-free extract contain: 25 mM HEPES, 100mM KCl, 5 mM $MgCl_2$ 250 mM sucrose 10% glycerol and 1 mM dithiothreitol. You dialyze with a buffer containing 25 mM HEPES, 100 mM KCl, 12 mM $MgCl_2$, 17% glycerol, and 2 mM dithiothreitol. Hypothesize if the components of the cell-free extract will increase, decrease, or stay the same following overnight dialysis. (Li et al., 2002 doi: 10.1105/tpc.010258)

A3) The HEPES and KCl concentrations will stay the same, as both the cell extract and dialysis solution contain equal concentrations. The concentration of sucrose will decrease. The concentrations of the $MgCl_2$, glycerol, and dithiothreitol will all increase.

Q4) Match the type of chromatography to its definition.

a) Cation-exchange 1) Antibodies are bound to beads and bind to tagged proteins
b) Affinity 2) A bead or gel matrix is created resulting in low molecular weight proteins exiting the column first, and larger last
c) Anion-exchange 3) A negatively charged bead binds to net positively charged proteins causing net negative and neutral proteins to elute faster
d) Size-exclusion 4) A positively charged bead binds to net negatively charged proteins causing net positive and neutral proteins to elute faster
e) Hydrophobic Interacting 5) A high salt concentration causes all proteins to bind to the column matrix. Decreasing the salt solution causes hydrophilic proteins to elute first, followed by hydrophobic

A4) a = 3; b = 1; c = 4; d = 2; e = 5

Q5) You want to separate 4 proteins with the following molecular weights: 120 kDa, 100 kDa, 150kDa, and 70 kDa.
a) What percentage of acrylamide gel would you use to resolve these proteins, 7%, 12%, or 15%? Explain.
b) After seeing your data from a), you decide a 25 kDa protein is also of interest. Can you use the same acrylamide % as in a)? Why or why not? And would you be able to
resolve all 5 proteins on your gel?

A5) a) A 7% gel would be the best option here. The lower percentage of acrylamide used for the gel, the larger the pores allowing for easier movement of high weight molecular weights.
b) No, a 7% acrylamide gel would not be able to resolve a 25 kDa protein. You would need a higher percentage gel, at least 12% to resolve a protein with a 25 kDa MW. With a 12% gel, it might be hard to distinguish between 100 and 120 kDa. You might consider running a gradient gel (4-15%) to see all 5 proteins.

Q6) You are studying a protein that undergoes the posttranslational modification of phosphorylation to become active, the addition of negatively charged phosphates to the outer surface of the protein. While you know the protein is phosphorylated to become active, you want to determine how many phosphates are added. When the protein is inactive (not-phosphorylated) the pI is 8.0.
What experiment could you plan to determine the number of phosphates added to the active form of your protein?

A6) You could plan to do isoelectric focusing of your protein in the inactive and active form to visualize the shift in pI. Knowing how much negative charge each phosphate adds to the pI, you can determine the number of phosphates added by how your protein moves on the isoelectric gel.

Section 5 Questions

Q1) Given the data below from a Bradford Absorbance experiment, determine the concentration of a protein extract:

Q2) The following data is from a circular dichroism experiment, based on the absorbance pattern, what is the predominant secondary structure of the protein ?

Q3) FRET Question

Q4) Question about MS with poor sequence coverage that doesn't contain a lot of basic AAs; use a new enzyme to digest.

Q5) NMR Question (use data?)

Q6) Question about Cryo-EM?

CHAPTER OVERVIEW

4: The Three-Dimensional Structure of Proteins

4.1: Main Chain Conformations
4.2: Secondary Structure and Loops
4.3: Tertiary and Quaternary Structures
4.4: Secondary Structural Motifs and Domains
4.5: Protein with Alpha, Alpha-Beta, Beta and Little Secondary Structure
4.6: Intrinsically Disordered Proteins
4.7: Fibrillar Proteins
4.8: Protein Folding and Unfolding (Denaturation) - Dynamics
4.9: Protein Stability - Thermodynamics
4.10: Protein Aggregates - Amyloids, Prions and Intracellular Granules
4.11: Biomolecular Visualization - Conceptions and Misconceptions
4.12: Laboratory Determination of the Thermodynamic Parameters for Protein Denaturation
4.13: Predicting Structure and Function of Biomolecules Through Natural Language Processing Tools
4.14: Chapter 4 Questions

4.1: Main Chain Conformations

Search Fundamentals of Biochemistry

4.1.1: Introduction

In Chapter section 3.2, we discussed primary, secondary, tertiary, and quaternary structures of proteins. We used as an example the protein **hydroxynitrile lyase** (5Y02). An interactive iCn3D model of hydroxynitrile lyase is shown in Figure 4.1.1.

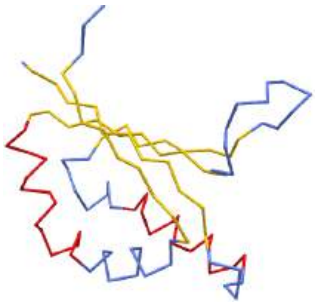


Figure 4.1.1: hydroxynitrile lyase (5Y02). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?gGY8QBkCP8dPmfoH6 (Copyright; author via source)

Blue indicates random coil, red indicates alpha helices and yellow shows beta strands. As we discussed in Chapter 3.2, the beta strand backbones are quite extended compared to the alpha helices which are more compressed. Atomic force microscopy can be used to pull the helix into alternative extended shapes. Let's explore main chain conformations in greater detail.

If we wish to understand protein structure, it is important to first understand the possible conformations of the backbone chain before we start considering the effect of side chains and the overall structure of the protein. Where do we start? It's a very complicated topic but we can simplify it by considering the structures of short hydrocarbons and other organic groups and use ideas you learned in earlier chemistry courses. So let's start with butane, a 4C hydrocarbon.

You will remember from previous chemistry courses that different rotational conformations (conformers) of butane have different energies. These conformers are often visualized using Newman projections sighting down the internal C2-C3 carbon bond in butane. The figure below shows ball and stick representations as well as Newman projects for three different conformers formed by rotation around the C2-C3 bond. You are familiar with bond angles, which is the angle between two bonds connecting three connected atoms. For this discussion, we will consider the dihedral angle which describes the angle of rotation around the middle bond of a 4-atom, 3-bond connected system. It's easiest to use Newman projections to understand dihedrals, which represents the angle between groups bonded to the front and back (circle) atoms sighted down the front-to-back atoms. Figure 4.1.2 shows multiple conformers of butane, represented in ball and stick form or as Newman projections.

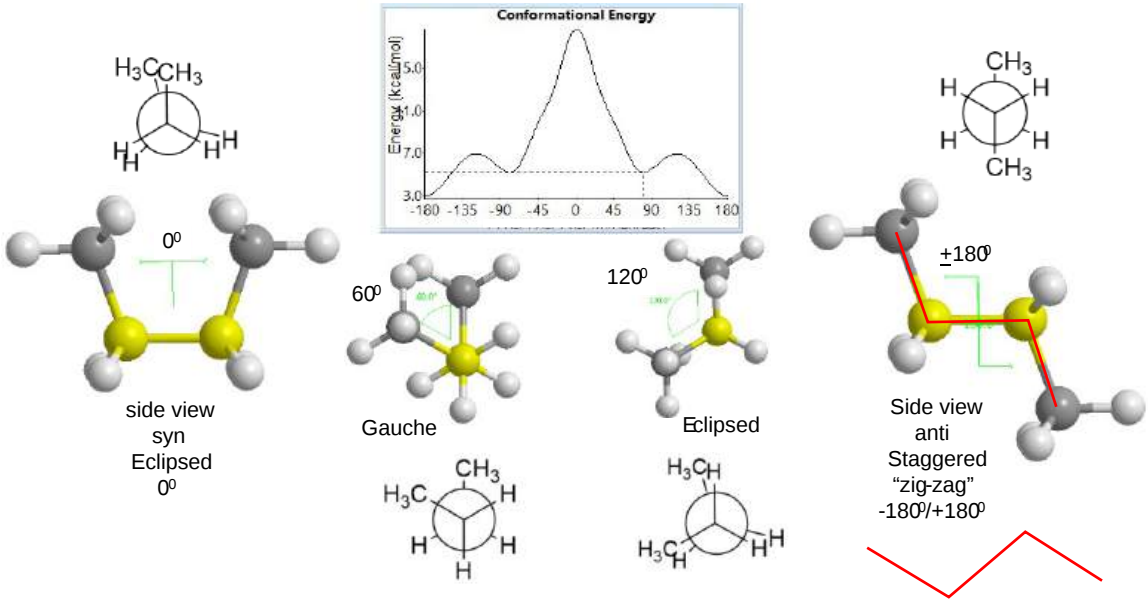


Figure 4.1.2: Newman projections for butane

The syn and eclipsed conformer with a dihedral angle of 0° has the highest energy since the two eclipsed bonds have high torsional energy and the two closest approaching H atoms on the methyl C atoms have steric (1,6) interactions. Rotating the front C 60° gives a gauche conformation of lower energy. Rotating to 120° leads to higher energy given the torsional strain between the eclipsed bonds. Rotating fully to 180° gives the conformer of lowest energy and torsional (bond-bond) strain. Steric strain between the two closest H atoms on the methyl groups also raises the energy.

A red line has been drawn between the 4 carbon atoms connected by three bonds in the anti or trans-conformer. Notice that the red line is a zig-zag as shown in the rotated red line drawing beneath the ball and stick conformer shown. All the conformers are available but the anti or trans-confomer is most stable and abundant.

Now, let's try this approach on a more complicated molecule, a 12-carbon atom carboxylic acid, dodecanoic acid. Figure 4.1.3 shows the molecule in the all- trans conformation (where all the carbon atoms in the chain are arranged in an anti, zig-zag manner), and a second gauche conformation for the bond highlighted by a red line. Note that rotation around that bond to produce the gauche conformer, a higher energy form, produces a kink in the chain. We will see this again when we discuss lipid structure.

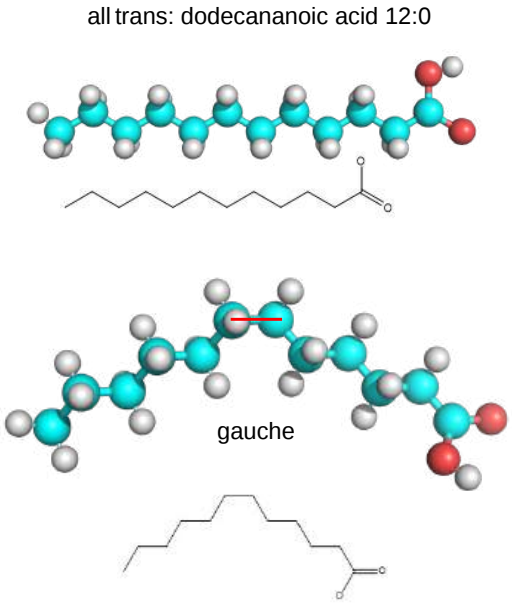


Figure 4.1.3: Conformations of dodecanoic acid

Now let's compare the all-trans dodecanoic acid structure to one largely extended conformer of a tetrapeptide, Gly-Gly-Ala-Gly (GGAG) as shown in Figure 4.1.4.

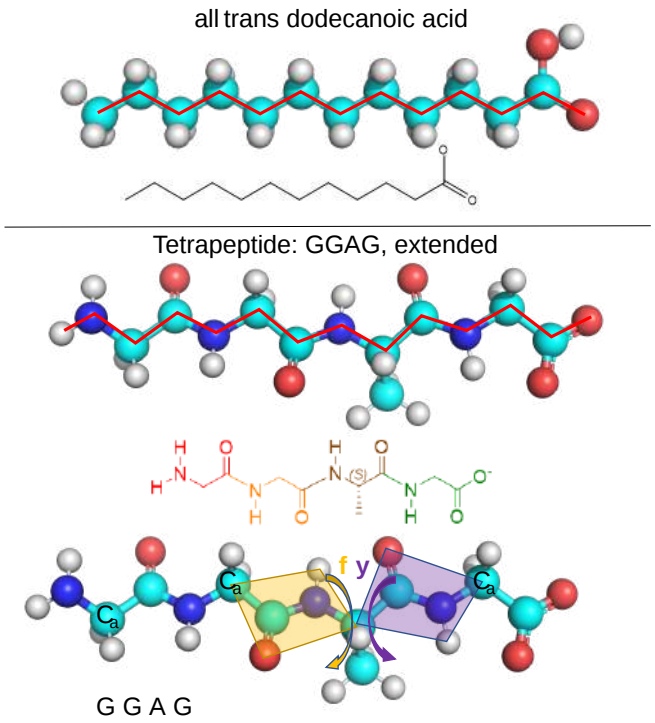


Figure 4.1.4: Extended conformations of dodecanoic acids and a tetrapeptide

Notice the similarities in these two structures. The atoms in the main chain of the peptide, N, C_a, C (the carbonyl C) are arranged in the familiar zig-zag fashion, characteristic of the lowest energy conformer. The dihedral angles along the backbone would be, to a first approximation, around $\pm$ 180° depending on whether the rotation is clockwise or counterclockwise from the syn conformer. However, the actual conformation of the GGAG peptide, if found as part of the primary structure of a protein, could adopt a whole

range of other conformers with potentially different dihedral angles for each bond in the main chain. Look back at the iCn3D model for **hydroxynitrile lyase** above and it should be clear that if the stretch of amino acids Gly-Gly-Ala-Gly is in a beta sheet, its dihedral angles for the main chain would be closer to $\pm 180^0$ than if they were in an alpha helix, which would dihedral angles much less than 180°, perhaps closer to $\pm 60^0$

It might seem like all possible dihedral angles are possible for the main chain atoms of a protein, which would make understanding protein structure too complicated for our minds to comprehend. However, that are two major things that simplify conformational analyzes of proteins.

4.1.2: Trans and Cis Peptide Bonds

Note in the extended peptide example in the above figure, the alpha Cs of adjacent amino acids are on opposite sides of the peptide bond between the carbonyl carbon atom and the amide nitrogen atom. They are **trans** to that bond. A rectangle was drawn on the above figure centered on the C-N peptide bond for two adjacent amino acids to clearly show the trans orientation of the alpha Cs.

Rotation around single bonds can occur, so an isomerization of the trans peptide bond to the cis isomer can occur. Both isomeric forms of the dipeptide Ala-Ala are shown in the top part of Figure 4.1.5.

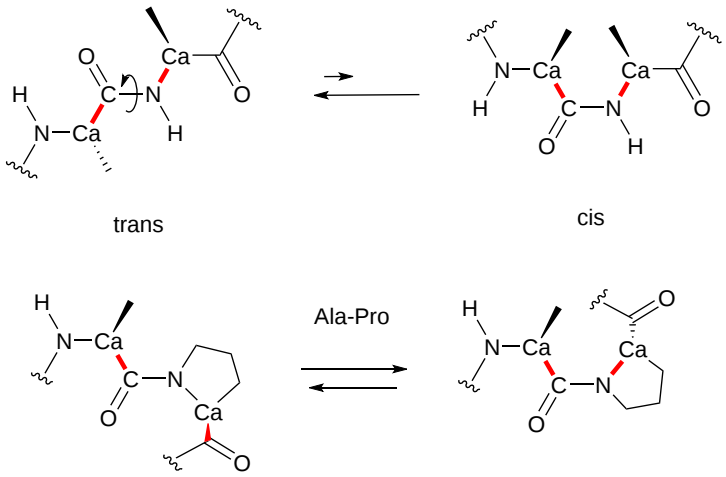


Figure 4.1.5: Cis/Trans peptide bonds

This trans arrangement of the alpha Cs in Ala-Ala and most internal X-Y amino acids pairs in a protein is sterically favored by a factor of 1000/1 over the cis form, which clearly shows torsional between the aligned (red) bonds and steric strain between amino acid side chain atoms.

The case is different for X-Y amino acid pairs when Y is proline, which is shown in the bottom part of the above figure. Proline is a cyclic amino acid and as such would be expected to be more sterically restricted in a protein sequence. In X-Pro peptide bonds in proteins, the trans/cis ratio is about 4/1. Clearly the torsional (between bonds) and steric strain (between atoms) are similar in both isomers.. You would expect to find both trans and cis X-Pro peptide bonds in proteins given the constraints placed on protein conformation in the tertiary structure of the protein.

Figure 4.1.6 is an interactive iCn3D model showing two pairs of X-Y peptide bonds in the protein carboxypeptidase. Answer the questions below about the two pairs of peptide bonds, which are shown in stick form and CPK colors.

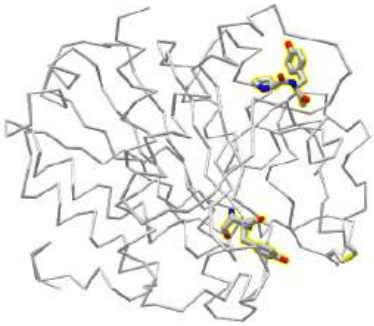


Figure 4.1.6: 2 peptide bonds in carboxypeptidase. Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...3ARF97MgN4427A (Copyright; author via source)

Exercise 4.1.1

Two X-Y amino acids are shown in stick form in the protein carboxypeptidase A

1. What is the sequence of each pair shown from N to C terminus

2. Are the X-Y peptides bonds cis or pro?

Answer

Add texts here. Do not delete this text first.

4.1.3: Backbone dihedral angles and Ramachandran Plots

Now let's focus on possible dihedral angles for the peptide bond in protein chains. Just as saturated fatty acid chains have preferred conformations (all anti or trans), peptide chains also have preferred conformations. The complexity is much greater, however. For our example using dodecanoic acid, we dealt only with torsion or dihedral angles around the methylene ($-CH_2-$) carbons. For proteins, we must consider the covalent links which attach the amino acids together, as well as the rotations possible in 20 different amino acids. The peptide bond connects the carbonyl C of the i^{th} amino acid to the alpha amine N of the i th+1 amino acid. The resulting bond is an amide link. X-ray analysis shows that the C-N bond has double bond character. This can be accounted for by delocalizing the nonbonding electron pair of the N to the carbonyl C forming a double bond, with the pi-bonded electrons of the carbonyl C-O bond moving to the O. These resonance structures lead to a planar arrangement of the peptide carbonyl C and amide N and the two other atoms connected to each, since the hybridization of the C and N has sp^2 character, with 120° bond angles, as illustrated in Figure 4.1.7.

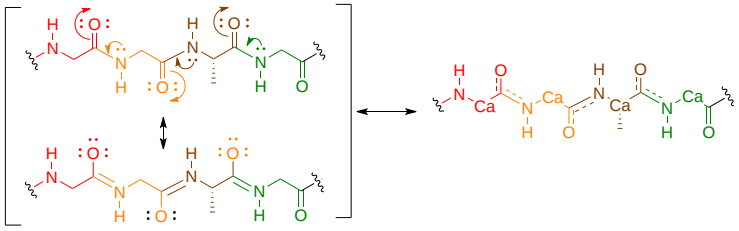


Figure 4.1.7: Resonance forms of peptide bond

This greatly simplifies the number of conformations that a protein can adopt since these 6 atoms can be considered to reside and move in a plane. The alpha C serves as the corner attachment point of two different planes, each of which can rotate independently of the other plane. The two planes can twist around the alpha carbon. The rotation angles for the two planes are called phi (φ) and psi (ψ) are analogous to the torsion angles in the acyl chain of dodecanoic acid. They can vary from -180 to +180°. The R group substituent attached to the alpha C can also rotate around the alpha C and the beta C of the side chain. This angle is defined as chi. Other rotations also occur within the side chain. We will concentrate on phi (φ) and psi(ψ) angles.

Figure 4.1.8 shows the peptide Gly-Gly-Ala-Gly (at high pH to give the protonation state shown) with the six atoms around the C-N peptide bonds shown in rectangles. Phi (φ) and psi (ψ) angles are also shown.

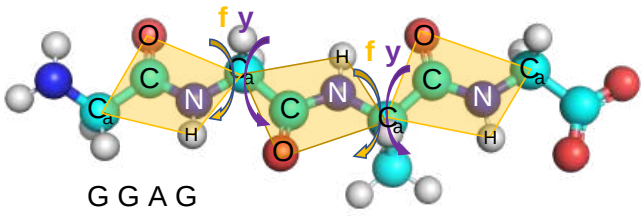


Figure 4.1.8: Phi/Psi angles for Gly-Gly-Ala-Gly

Using these ideas, we can consider a protein backbone as a series of linked sequences of rigid, planar peptide units which can rotate around phi/psi angles. When the chain is fully extended (as shown in the links above), φ/ψ are ± 180°, with the syn eclipse form

defined as having 0^0 φ/ψ angles. When φ and ψ equal 0^o, the two peptide bonds flanking the alpha Cs are in the same plane. This conformation is prohibited since the O of the C=O on one plane and the H of the H-N on the other are overlapping - i.e. they approach closer than their van der Waals radii.

The Proteopedia site below gives a phenomenal explanation and visual representation of φ/ψ angles in proteins.

Proteopedia

:

Ramachandran principle and phi psi angles

Figure 4.1.9 is an animated gif showing rotation of phi and psi angles (made from Proteopedia link above). You can see that when either φ or ψ approach 0^0, steric clashes are most significant.

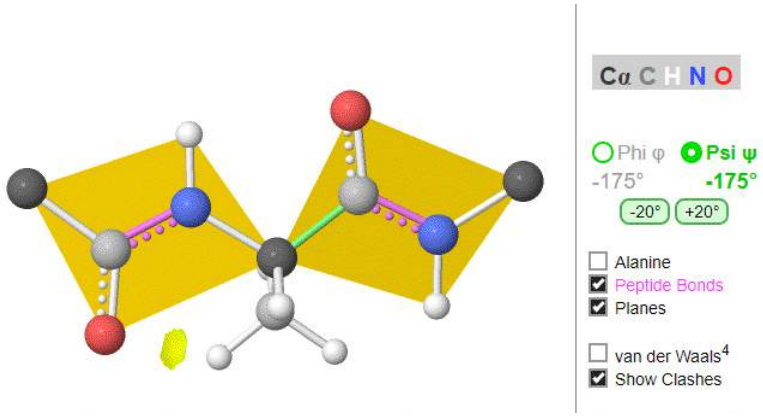


Figure 4.1.9: Animation of rotation around Phi/Psi angles f

Ramachandran was the first to calculate which φ/ψi angles are allowed. He modeled the angles permitted to a tripeptide, assuming the atoms were hard spheres. The angles allowed depended in part on the limiting distance chosen for interatomic contacts. (i.e. the usual H -- H distance is 2.0 angstroms, and 3.0 for C--C bonds.) The plot below shows the allowed regions in red. Only 3 small regions of conformational space are available. If you allow a closer approach by 0.1 angstroms, more conformation space is available, but only one new area is available, shown in yellow in Figure 4.1.10.

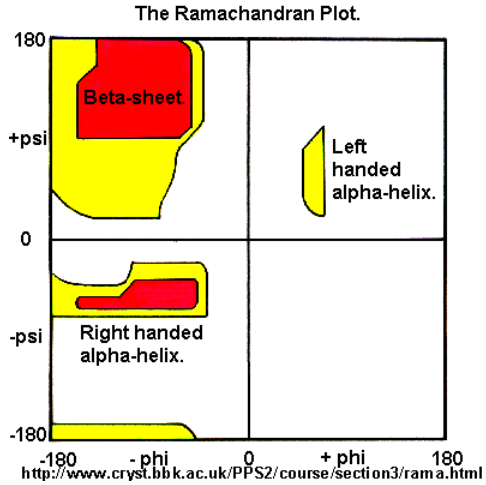


Figure 4.1.10: Ramachandran plot. http://www.cryst.bbk.ac.uk/PPS95/cou...etry/rama.html

Right-hand alpha helices fall at -57^0,-47^0 while left hand alpha helices fall at $+57^0$,$+47^0$. (Notice these are not mirror images of each other. The mirror image of a right-handed alpha helix would be a left-handed helix made of D-amino acids.) Parallel beta sheets are at -119^0, -113^0, while antiparallel sheets falls at -139^0, $+135^0$. Note that beta sheets have φ/ψ angles closer to the fully extended + 180^0 of the all anti (trans) conformation of dodecanoic acid than does the alpha helix which is more compact. Other types of helices also are found. The 310 helix, a sharper helix with 3 amino acids/turn, falls at -49^0,-26^0. All of these examples of secondary structure fall in allowed regions.

Modern Ramachandran plots do not model the atoms as hard spheres but instead consider the potential energy of the atoms using the Lennard-Jones potential (6-12 potential) for induced dipole-induced dipole interactions. We discussed this potential function Chapter 2.4.

A Ramachandran plot of Ala-Ala-Ala is nearly identical to the plot for Phe-Phe-Phe (which is unbranched at the beta carbon, the first methylene C in the side chain). The plot for Thr-Thr-Thr, which has a branch at the beta C (with OH and CH_3 attached) shows somewhat less room than the other plots. Pro-Pro-Pro is most restricted for obvious reasons.

For a longer chain than a tripeptide, there are more restrictions than for $(Ala)_3$, since the chain can't assume a conformation when it passes through itself. The plots for actual proteins have many points which do fall in forbidden regions. However, these points would be allowed if the peptide bonds twist a few degrees. Gly bonds also fall outside the allowed regions. This is understandable, since the side chain of Gly is H, and it is used in proteins where sharp turns of the chain are necessary.

Figure 4.1.11 shows Ramachandran plots for Gly, Ala, Tyr, and Pro in actual proteins (made years OK but can't find reference)

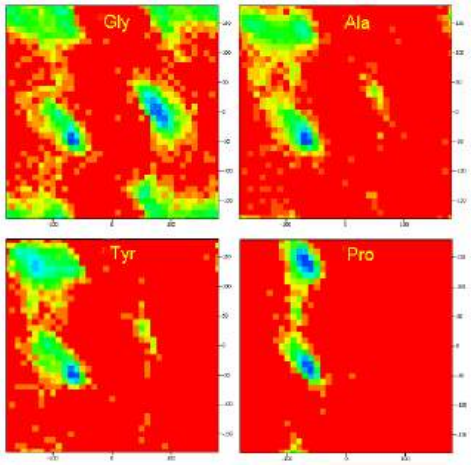


Figure 4.1.11: Ramachandran plots for Gly, Ala, Tyr and Pro from actual proteins

Figure 4.1.12, taken from the Proteopedia page above, shows the Ramachandran plots for 100,000 proteins from the Protein Data Bank.

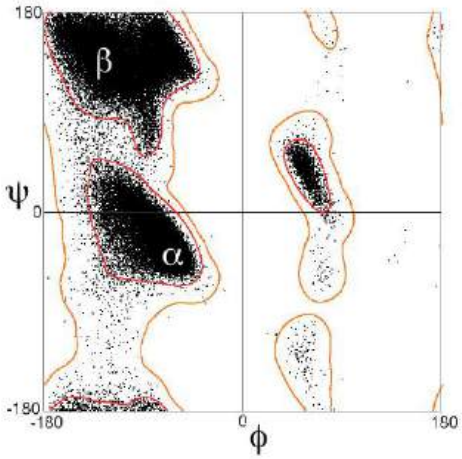


Figure 4.1.12: Ramachandran plots for 100,000 proteins. https://proteopedia.org/wiki/index.p...neral_100K.jpg

Now it's time to explore in more detail secondary, tertiary and quaternary structures in more detail. There are many ways to organize the structural organization of the protein world. For the rest of this chapter, we will adopt the approach used by Carl Branden and John Tooze in their seminal book, "Introduction to Protein Structure".

4.2: Secondary Structure and Loops

Secondary structures are those repetitive structures involving H bond between amide Hs and carbonyl Os in the protein backbone. These include

- helices (alpha - α , 3_{10} and pi - π), in which the hydrogen bonds are within a short continuous stretch of amino acids (a strand),
- beta strands (sheets) in which the hydrogen bonds are between backbone atoms (again amide Hs and carbonyl Os) on noncontinuous stretches of the protein, and
- reverse turns, which occur within a very short continuous stretch of amino acids.

4.2.1: Helices

A schematic showing idealized geometries of helices, with amino acids shown as dots for simplicity, is shown in Figure 4.2.1.

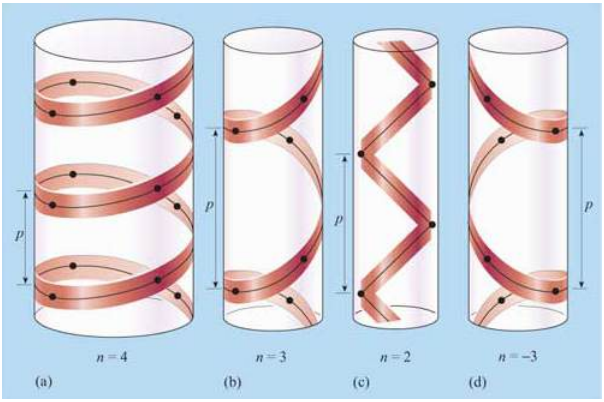


Figure 4.2.1: Helix geometries Open University. https://www.open.edu/openlearn/scien...-section-1.3.1.

The pitch (*p*) represents the spacing between the chain on one side of the helix, the number of amino acids per turn (*n*), and handedness (+ = right-handed or – l= left- handed) are shown in the figure. There are three major types of helices in proteins, the alpha helix (n = 3.6), 3_{10} helix (n = 3) and the pi helix (n = 4.4). Note that n is most commonly not an integer.

4.2.1.1: Alpha

The alpha helix is the most common type of helix. They are formed when the carbonyl O of the i^{th} amino acid forms hydrogen bonds to the amide H of the i^{th+4} aa (4 amino acids away). Figure 4.2.2 show a short section of an alpha helix running from N-terminal (bottom) to C-terminal (top) with the sequence DTASDAA. The amino acids i, i+1, ... i+4 are labeled at their alpha carbons. The red oval highlights the intrastrand H bond between the C=O of the i^{th} amino acid (Asp) and the amide H of the i^{th+4} aa (Ser).

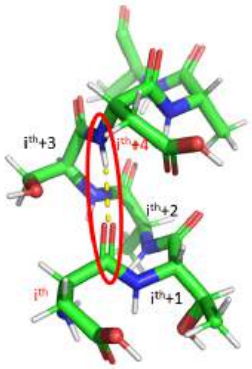


Figure 4.2.2: A short alpha helix

Figure 4.2.3 shows a cartoon image showing a more extended helix and a schematic showing hydrogen bonding partners.

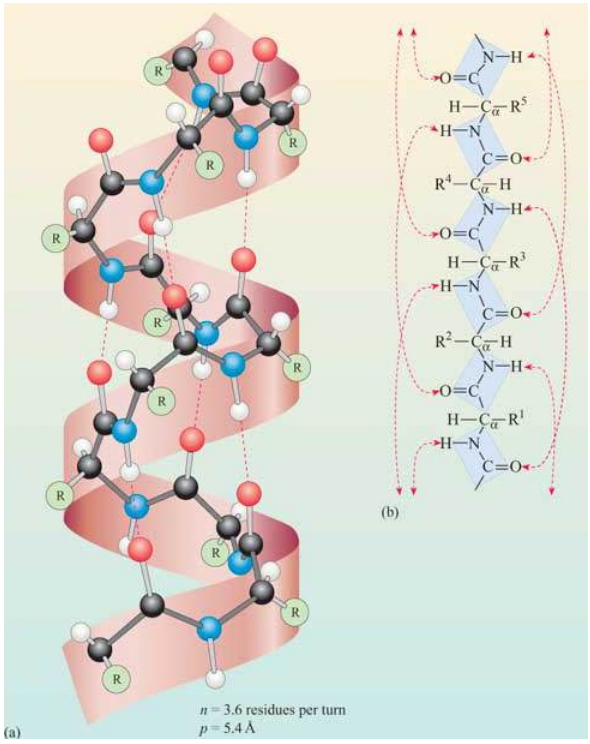


Figure 4.2.3: An extended alpha helix Open University. https://www.open.edu/openlearn/scien...-section-1.3.1.

The phi/psi angles for amino acids in the alpha helix are approximately - 57,-47, which emphasizes the regular repeating nature of the structure. It can also be characterized by n (the number of amino acid units/turn = 3.6) and the pitch (the helix rise/turn = 5.4 angstroms = 0.54 nm). Since there are 3.6 amino acids per turn, and a full circle or turn is 360^0, each amino acid is staggered at 100^0 increments looking down on the helix axis. To refresh your mind, the phi/psi diagram for a fully extended polypeptide chain (phi 180^0, psi 180^0) is shown below in Figure 4.2.4.

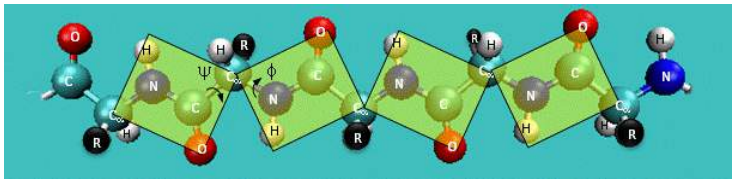


Figure 4.2.4: Extended alpha helix with phi/psi planes

Figure 4.2.5 shows side and end-on view of a helix from the antifreeze protein (1wfa) from winter flounder. The green coil (often shown in red when displaying alpha helices in full proteins) shows the repetitive nature of the backbone. Note that side chains are pointing away from the helix axis. H-bonds are shown as yellow dotted lines within the backbone (one is also shown between two side chains on the top). The spacefill rendering is shown in colors optimal for those who are colorblind. The end-on view shows that the center of the helix is fully packed with atoms from the helix and is **NOT** open (a common misconception among students).

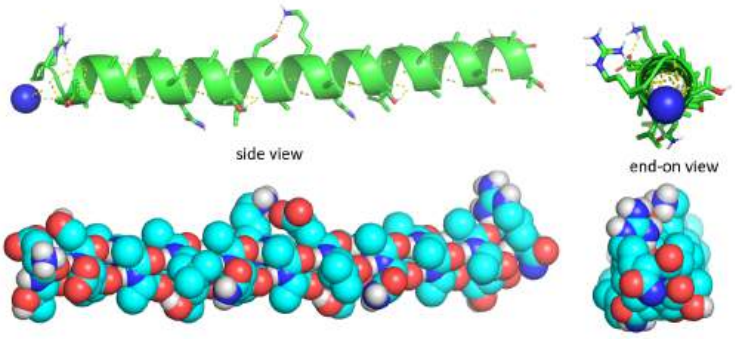


Figure 4.2.5: Multiple views and renders of an alpha helix

Some facts:

- the alpha helix is more compact than the fully extended polypeptide chain with phi/psi angles of 180°
- in proteins, the average number of amino acids in a helix is 11, which gives 3 turns.
- the left-handed alpha helix, although allowed from inspections of a Ramachandran plot, is never observed, since the side chains are too close to the backbone.
- the core of the helix is packed tightly. No central cavities or or pores are present in the helix.
- All the R-groups extend backward and away from the helix axis.
- Some amino acids are more commonly found in alpha helices. Amino acids can be divided into two kinds, those with branches at the beta C and those with none. Consider first those that aren't branched. Gly is too conformationally flexible to be found with high frequency in alpha helices, while Pro is too rigid. The amino acids with side chains that can H-bond (Ser, Asp, and Asn) and aren't too long appear to act as competitors of main chain H bond donors and acceptors, and destabilize alpha helices. The rest with no branches at the beta C can form helices. Those with branches at the beta carbon (Val, Ile) destabilize the alpha helix due to steric interactions of the bulky side chains with the helix backbone. (Remember left-handed alpha helices are not found in nature for similar reasons.)
- alpha keratins, the major component of hair, skin, fur, beaks, and fingernails, are almost all alpha helix.

Figure 4.2.6 shows an interactive iCn3D model of an alpha helix from bacteriophage T4 lysozyme (1DYG). Side chains, which are not involved in helix-stabilizing hydrogen bonds are shown in cyan. H-bonds are shown as green dotted lines.

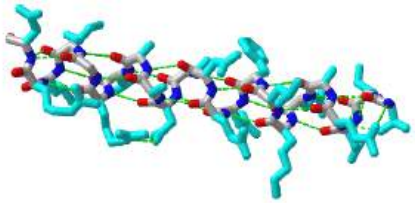


Figure 4.2.6: An alpha helix from bacteriophage T4 lysozyme (1DYG). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...qhkPT6cMqxvp88 (Copyright; author via source)

The amino acid side chain R-groups can be hydrophilic or hydrophobic, and can be localized in specific positions on the helix forming amphipathic regions on the protein or fully hydrophobic helices may also extend through the plasma membrane as shown in Figure 4.2.7.

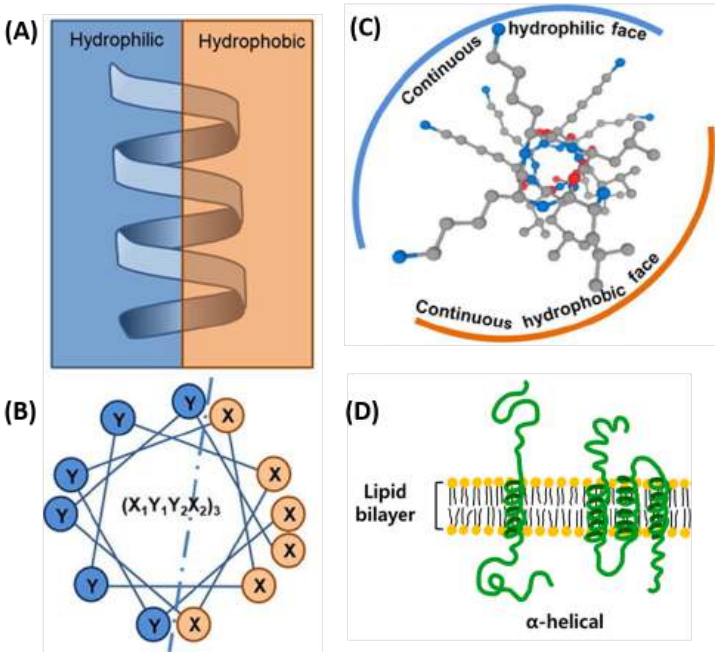


Figure modified from: Khara, J.S., et al. (2017) *Acta Biomat* 57:103-114 and Ryu, H., et al. (2019) *Int J Mol Sci* 20 (6) 1437

In amphipathic helices, hydrophilic residues are positioned on one side of the helix and hydrophobic on the other as shown in the side view (A) or top-down views (B & C). R-groups may also be fully hydrophobic within alpha helices that span the plasma

membrane as shown in (D).

Helical wheel projections can be made showing the polarity of the faces of the helix looking down the axis. Here are two such helical wheel projections:

For the sequence MLQSMVSLLQSLVSLIIQ, Figure 4.2.8 shows that the helix is amphiphilic.

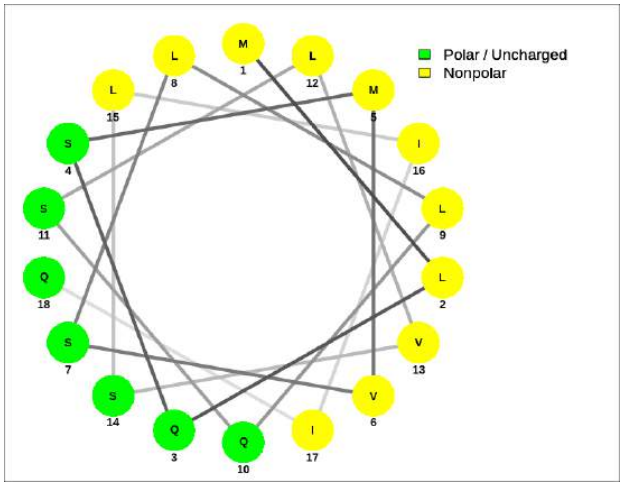


Figure 4.2.8: Helical wheel for MLQSMVSLLQSLVSLIIQ http://lbqp.unb.br/NetWheels/

A helical wheel for the membrane-crossing section of the human receptor-type tyrosine-protein phosphatase C protein, ALIAFLAFLIIVTSIALLVVL, is shown in Figure 4.2.9.

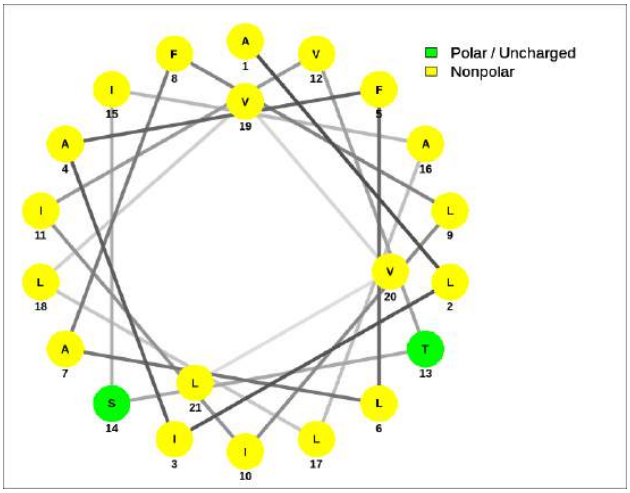


Figure 4.2.9: Helical wheel for a membrane helix

The amide bond in the peptide has a significant permanent dipole moment. Since the dipoles of individual amino acids are oriented in the same direction in an alpha helix, the whole helix has a significant dipole moment. This is illustrated in Figure 4.2.10

recreate this image:

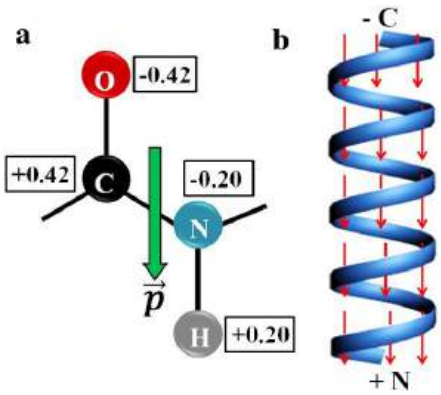


Figure 4.2.10: Helix Dipole (https://www.sciencedirect.com/scienc...05273612003860; reproduce image)

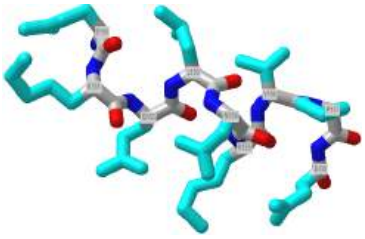
In the alpha helix all the amides in the backbone point in one direction from the C-terminus to N-terminus. This leads effectively to one long dipole of magnitude n(3.5) Debyes, where n is the number of amino acids in the helix. A really long helix can then produce a significant electric field and affect protein binding properties.

4.2.1.1.1: 3_{10} helices

The 3_{10} helix is stabilized by hydrogen bonds between the carbonyl O of the i^{th} amino acid and the amide H of the i^{th+3} aa (3 amino acids away). It has 3 residues/turn, and a pitch (rise per turn) of 6 angstroms, with a rise of 1.3-2 angstroms/residue. Typical phi/psi angles are -50^0,-26°. As with the alternative description of the alpha helix, the $\mathbf{3_{10}}$ helix has **3** amino acids per turn and **10** atoms in the main chain/turn (counting Cα-N-C-Cα atoms). It is longer (for the same number of amino acids) and thinner. The amino acid side chains are staggered at 120^0 increments as you look around the helix axis. Although not very prevalent (about 3 percent of protein amino acids are in 3_{10} helix with an average 3.3 amino acids in the helix), they presumably serve some function. 3_{10} helices as long as 11 residues have been found.

Within a protein, a helix will be stable if the packing around it will allow it. Many more alpha-helix side chain interactions with surrounding protein are likely, given a 100^0 staggering of side chains compared to a 120^0 staggering in a 3_{10} helix, which lineup in 3 ridges looking down the helical axis. Molecular dynamics studies suggest that parts of a 3_{10} helix might reversibly interconvert to an alpha helix, allowing conformational and binding flexibility. The S4 helix in some voltage-sensitive potassium ion channels with a canonical $R_1xxR_2xxR_3xxR_4xxK_5xxR_6$ (where R and K are Arg and Lys) have been shown to adopt a 3_{10} helical conformation.

Figure 4.2.11 shows an interactive iCn3D model of a 9 amino acids (150-158) 3_{10}helix from dienelactone hydrolase (1DIN)



NCBI iCn3D Figure 4.2.11: A 3_{10} helix from from dienelactone hydrolase (1DIN). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...Kufi7gmpSBdjYA (Copyright; author via source)

4.2.1.1.2: π (pi) helices

This helix has 4.4 residues/turn, and a helix pitch (rise) of about 4.1 nm angstroms. It has hydrogen bonds between the carbonyl O of the i^{th} amino acid and the amide H of the i^{th+5} amino acids (5 amino acids away). It rise is about 1.2 angstroms/residue and has approximate phi/psi angles of -55^0, -70^0. An alternative designation for the pi helix is 4.4_{16} with 16 main chain atoms in 1 full turn (see above for the alpha helix). Some reasons for its low abundance include minimal contact between the main chain atoms given the larger radius, and phi-psi angles close to disallowed values. It might also not form kinetically as fast as the other helices since nucleation of it would be more difficult. At the same time, molecular dynamic simulation simulations show that alpha helices can interconvert reversibly with pi helices. They are often found between two alpha helices, again suggesting dynamic interconversions between the two forms are likely.

About 55% of characterized pi helices contain 5 amino acids. Each side chain is staggered by 85^0, with a rise of about 1.3 Angstroms. Figure 12 below shows an interactive iCn3D model of a short pi helix (aa 265-276) from barley beta-D-glucan glucohydrolase (1x38).

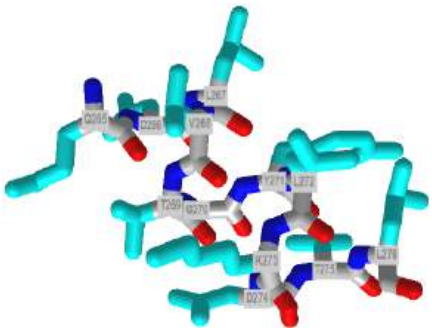


Figure 4.2.12: A short pi helix (aa 265-276) from barley beta-D-glucan glucohydrolase (1x38). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...U7H6xjzHd95nbA (Copyright; author via source)

Helices in Proteins: Comparison of alpha, 3_{10} and pi helices

4.2.2: Beta Structure

Beta Structure: **Parallel** and **antiparallel** beta strands are much more extended than alpha helices (phi/psi of -57,-47) but not as extended as a fully extended polypeptide chain (with phi/psi angles of +/- 180) as shown in the figure below. Parallel beta strands have phi/psi angles of -119, +113, while the antiparallel angles are -139, +135. Figure 4.2.13

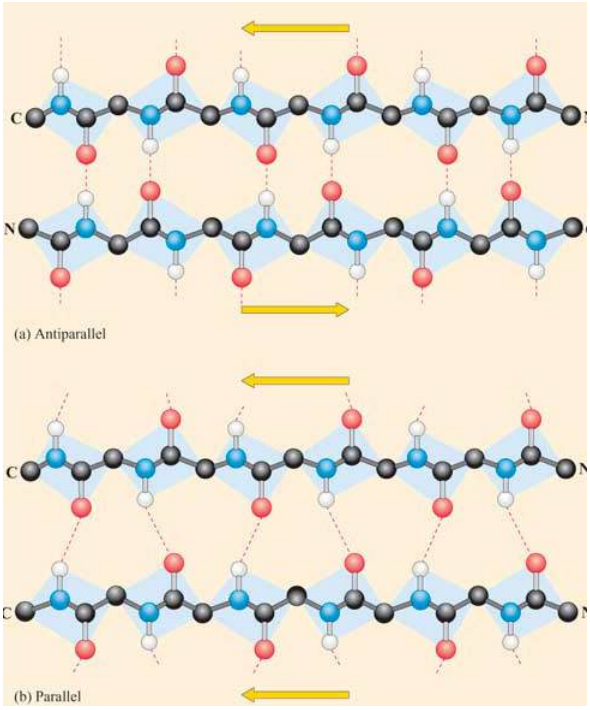


Figure 4.2.13Open University. https://www.open.edu/openlearn/scien...-section-1.3.2

Each single strand of the beta-sheet can be pictured as a twofold helix, i.e. a helix with 2 residues/turn. The arrangement of each successive peptide plane is pleated due to the tetrahedral nature of the alpha C. Hydrogen bonds are inter-strand, not intra-strand as in the alpha helix.

The figure below shows how the "pleats" in a sheet containing parallel beta strands can be envisioned as rippled sheets. Figure 4.2.14

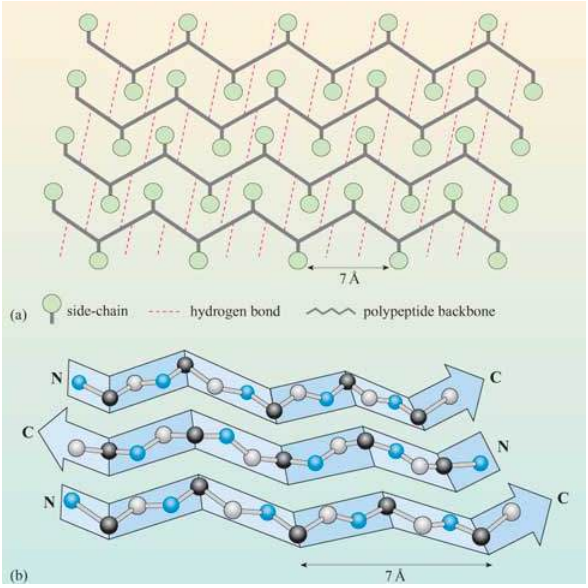


Figure 4.2.14

They can be visualized by folding a sheet of paper into narrow folds or pleated strips side by side to make a "pleated sheet" of paper. Each strip of paper can be pictured as a single peptide strand in which the peptide backbone makes a zig-zag along the strip, with the alpha carbons lying at the folds of the pleats. The R groups are attached to the carbons and extend above and below the folds of the pleat in the *trans* conformation.

Consider a strand as a continuous and contiguous polypeptide backbone propagating in one direction. Hence, using this definition, a helix consists of a single strand, and all the H-bonds are within the strand (or intrastrand). A beta sheet would then consist of multiple strands, since each "strand" is separated from other "strands" by an intervening contiguous stretch of amino acid which bends within the protein in a way that allows the next section of the peptide backbone, the next "strand", to H-bond with the first "strand". But remember, even in this case, all the H-bonds holding the alpha and beta structures together are intramolecular.

In a parallel beta sheet structure, the optimal H bond pattern leads to a less extended structure (phi/psi of -119, +113) than the optimal arrangement of the H bonds in the antiparallel structure (phi/psi of -139, +135). Also, the H bonds in the parallel sheet are bent significantly. (i.e. the carbonyl O on one strand is not exactly opposite the amide H on the adjacent strand, as it is in the antiparallel sheet.) Hence antiparallel beta strands are presumably more stable, even though both are abundantly found in nature. Short parallel beta sheets of 4 strands or less are not common, which might reflect their lower stability.

The side chains in the beta sheet are normal to the plane of the sheet, extending out from the plane on alternating sides. Parallel sheets characteristically distribute hydrophobic side chains on both sides of the sheet, while antiparallel sheets are usually arranged with all the hydrophobic residues on one side. This requires an alternation of hydrophilic and hydrophobic side chains in the primary sequence. Antiparallel sheets are found in silk with the sheets running parallel to the silk fibers. The following repeat is found in the primary sequence: (Ser-Gly-Ala-Gly)n), with Gly pointing out from one face, and Ser or Ala from the other.

Unfortunately, there is no PDB structure of the silk "amyloid" protein showing this repetitive structure. The monomer and aggregates of this protein are quite insoluble so few x-ray structures for proteins like this are available. Figure 4.2.15 shows an interactive iCn3D model of the N-terminal part (domain) of the *Bombyx mori* fibroin silk protein (pdb = **3UA0),** which does give an excellent example of antiparallel beta sheets. Notice that two chains align to form a face of the curved antiparallel beta sheet.

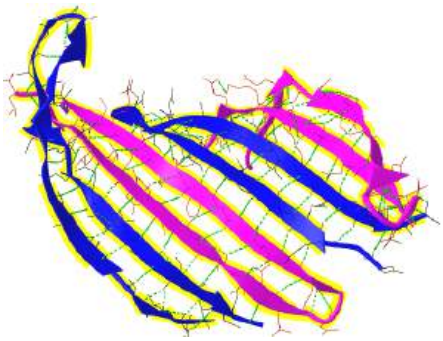


Figure 4.2.15: N-terminal Domain of the *Bombyx mori* fibroin silk protein (3UA0**).** Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...TuDiMoi64TaAx9 (Copyright; author via source)

Figure 4.2.16 shows a more detailed static image of the antiparallel beta sheets in 3UA0. Note the yellow sticks between the strands representing the H-bonds.

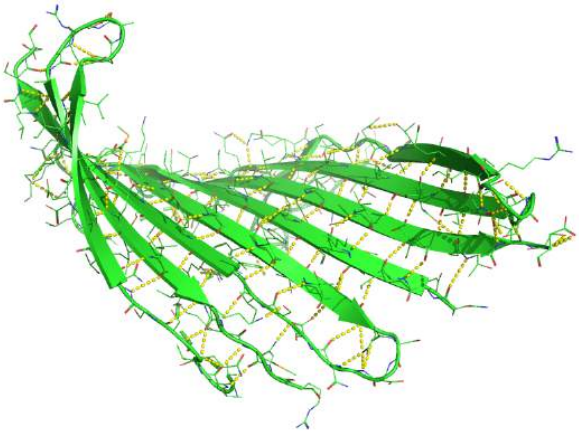


Figure 4.2.16:

Beta strands have a tendency to twist in the righthand direction. This leads to important consequences in how the beta strands are connected. Parallel strands can form twisted sheets or saddles as well as beta barrels.

Figure 4.2.17shows an interactive iCn3D model of the parallel beta sheet structure from the arabinose binding protein (1ABE).

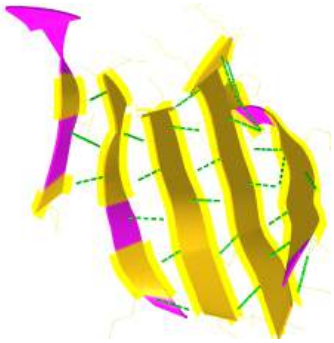


Figure 4.2.17: Parallel beta sheet structure from the arabinose binding protein (1ABE). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...7A2akqvuCpavu8 (Copyright; author via source)

Figure 4.2.18 shows a static image of the parallel beta sheets in 1ABE. Note the yellow sticks between the strands representing the H-bonds.

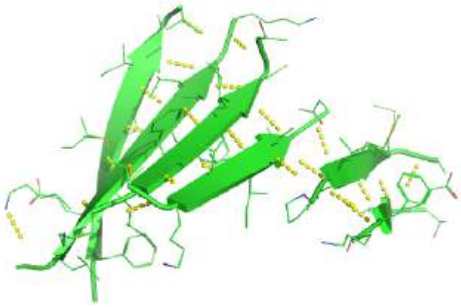


Figure 4.2.18: Parallel beta sheet structure from the arabinose binding protein (1ABE)

Figure 4.2.19 shows an interactive iCn3D model of a parallel beta barrel from the triose phosphate isomerase (1WYI).

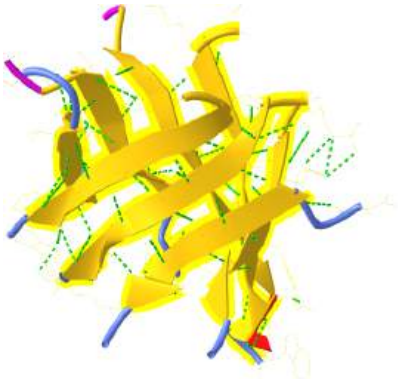


Figure 4.2.19: Parallel beta barrel from the triose phosphate isomerase (1WYI) (Copyright; author via source)). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...7dki3WjRDtiaRA

Figure 4.2.20 shows a static image of the parallel beta sheets in triose phosphate isomerase. Note the yellow sticks between the strands representing the H-bonds. Also, note that the barrel is not hollow but is filled with side chains.

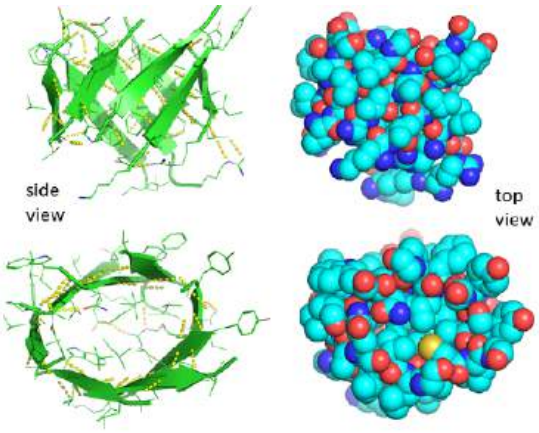


Figure 4.2.20: Parallel beta sheets in triose phosphate isomerase

Some facts on parallel beta structure:

- in parallel strands, right-handed connectivity is common.
- in a protein with parallel strands in register, and given the inherent twist in the stands, the strands arrange in a way to have the H bonds stretched equally at the ends of the chains, giving rise to a twisted saddle shape (top structure above).
- in a protein with parallel strands out of register, and given the inherent twist in the stands, the strands arrange in a way to have the H bonds stretched equally at the ends of the chains, giving rise to a beta barrel (bottom structure above).

4.2.3: Connectors, Loops, Linkers and Bends

About 50% of the amino acids in a globular (spherical) protein are in regular secondary structure (alpha or beta). The amino acids in helices and beta strands are connected by stretches of amino acids which still have order, but that order is less regular than those found in helices and beta strands, which are characterized by stretches of amino acids with the same phi/psi angles. Some bear the hallmarks of secondary structure - they are held together by intrachain hydrogen bonds. We will consider a few here.

Turns, Reverse Turns and Hairpins

One example of a connector involving secondary structure (i.e. hydrogen bonds between amide Hs and carbonyl Os of the backbone), is a reverse turn called the **beta bend or beta turn**. These turns often connect successive antiparallel beta strands and are then called beta hairpins. A ***hairpin*** is a special case of a turn, in which the direction of the protein backbone reverses and the flanking secondary structure elements interact. For example, the ***beta hairpin*** connects two hydrogen-bonded, antiparallel β-strands. The word *beta* can be confusing. It does not mean that the structure has hydrogen-bonded amino acids that have the same phi-psi angles as beta strands. It's easy to remember the name beta as the beta bend connects two antiparallel beta strands. The term beta really comes from the fact that it is a member of a class of turns named with Greek letters, including alpha-, gamma-, delta-, pi- and beta- turns.

They are almost always on the surface, and usually consist of 4 amino acids. However, there are several types of beta turns and different ways to classify them. One involves the number of residues (n) *between the two residues that are hydrogen bonded*.

n=2: These contain four amino acids. Amino acids 1 and 4 form hydrogen bonds with n=2 amino acids in between. Another way to describe them is that the hydrogen bond between residues 1 and 4 is between the backbone carbonyl O of the i^{th} amino acid and the amide H of the i^{th}+3 aa (three amino acids away) so the structure contains 4 amino acids (i^{th}, i^{th}+1, i^{th}+2, and i^{th}+3). There are two common types:

- Type I: phi 2 = -60, psi 2 =-30; phi 3 = -90, psi 3 = 0; The first amino acid in the actual turn (i^{th} + 1) is actually in a left-handed alpha helix conformation. Glycine, asparagine or aspartate are stable at this position since glycine is small and the side chains of Asp and Asn can form hydrogen bonds to the main chain. Glycine is usually found in the second position of the actual turn (i^{th} + 2).
- Type II: phi 2 = -60, psi 2=120; phi 3 = 90, psi 3 = 0; The first residue of the actual turn is typically Gly while the second often is a polar amino acid such as Ser and Thr

Here are 2D line drawings showing Type I and Type II beta turns. Figure 4.2.21

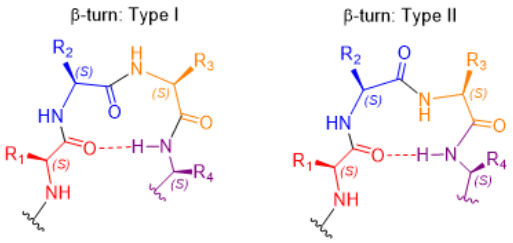


Figure 4.2.21

The figure below shows Type I (left) and Type 2 (right) from human egg white lysozyme. Figure 4.2.22

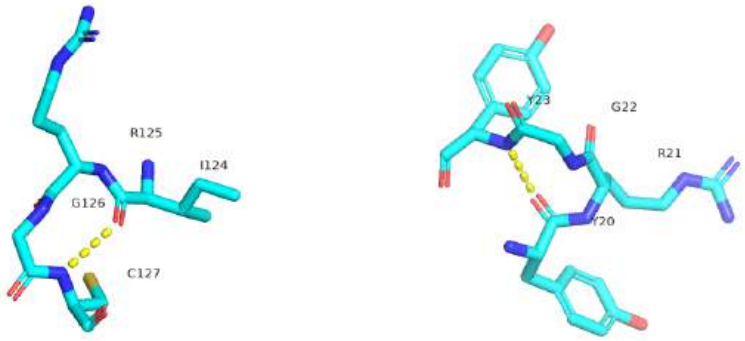


Figure 4.2.22

Figure 4.2.23 shows an interactive iCn3D model showing two reverse turns in hen egg white lysozyme (1dpx). Notice the tightness of the reverse turn and the presence of proline and glycine.

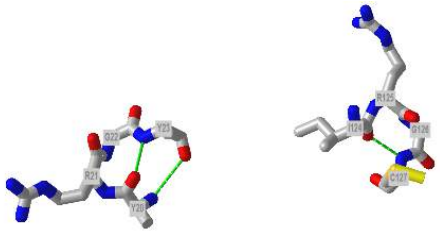


Figure 4.2.23: Two reverse turns in hen egg white lysozyme (1dpx) (Copyright; author via source). Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...U2BzPt98cPYHFA

n=3: These contain five amino acids. Amino acids 1 and 5 form hydrogen bonds with n=3 amino acids in between. Another way to describe them is that the hydrogen bond between residues 1 and 4 is between the backbone carbonyl O of the i^{th} amino acid and the amide H of the i^{th}+4 aa (four amino acids away) .

Shifting back to the Greek letter naming system, the **gamma turn**, the second most common turn, has just three total residues, (i^{th}, i^{th}+1, and i^{th}+2) with the hydrogen bond between the backbone carbonyl of the i^{th} amino acid and the backbone amide H of the i^{th}+2 amino acid.

An ***ω-loop*** is a catch-all term for a longer, extended or irregular loop **without** fixed internal hydrogen bonding. Turns are sometimes found within flexible linkers or loops connecting protein domains. Linker sequences vary in length and are typically rich in polar uncharged amino acids. ***Flexible linkers*** allow connecting domains to freely twist and rotate to recruit their binding partners via protein domain dynamics.

Mostly modeling programs display the linear sequence of a protein and in addition a linear cartoon rendering showing alpha structure as squiggles or helices, beta structure as yellow arrows. and connecting amino acids between secondary structures as lines. The figure below shows a 1D connectivity diagram for part of the protein alpha-lactalbumin (1a4v). Figure 4.2.24

Figure 4.2.24:

4.2.4: Amino Acid Propensities for Secondary Structures

Finally, what types of amino acids are most likely to be found in different types of secondary structures? Some rationales for the "propensity" for secondary structure are shown in the figure below. Figure 4.2.25

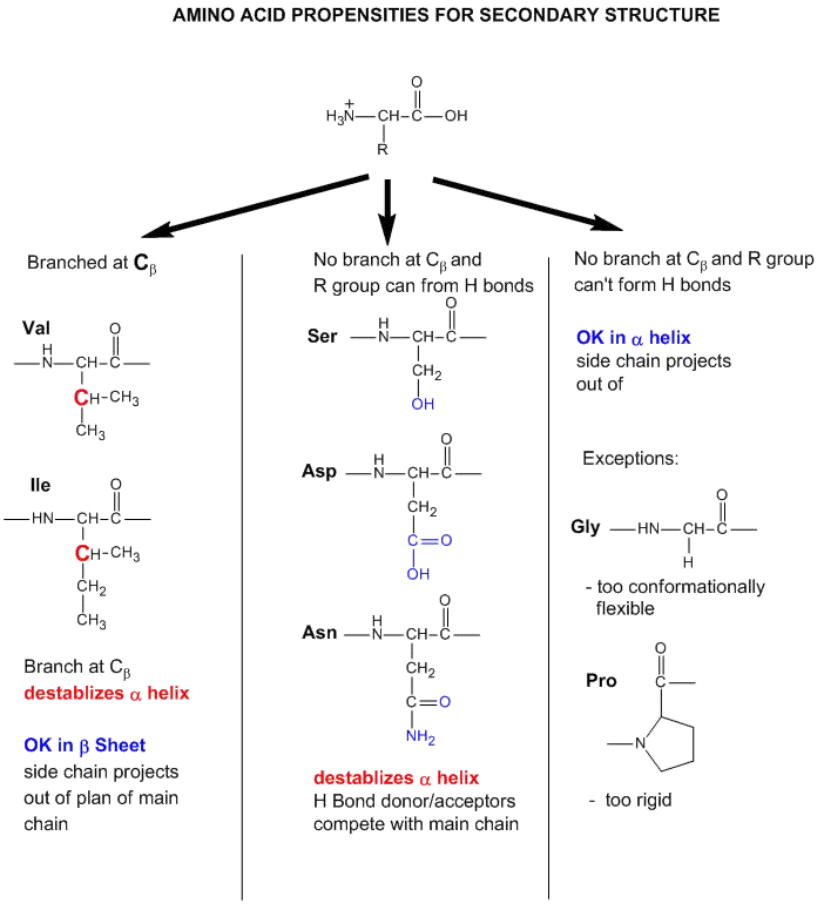


Figure 4.2.25:Amino Acid Propensities for Secondary Structure

OpenStax, Proteins. OpenStax CNX. Sep 30, 2016 http://cnx.org/contents/bf17f4df-605c-4388-88c2-25b0f000b0ed@2.

File:Chirality with hands.jpg. (2017, September 16). *Wikimedia Commons, the free media repository*. Retrieved 17:34, July 10, 2019 from commons.wikimedia.org/w/index.php?title=File:Chirality_with_hands.jpg&oldid=258750003.

Wikipedia contributors. (2019, July 6). Zwitterion. In *Wikipedia, The Free Encyclopedia*. Retrieved 21:48, July 10, 2019, from en.Wikipedia.org/w/index.php?title=Zwitterion&oldid=905089721

Wikipedia contributors. (2019, July 8). Absolute configuration. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:28, July 14, 2019, from en.Wikipedia.org/w/index.php?title=Absolute_configuration&oldid=905412423

Structural Biochemistry/Enzyme/Active Site. (2019, July 1). *Wikibooks, The Free Textbook Project*. Retrieved 16:55, July 16, 2019 from en.wikibooks.org/w/index.php?title=Structural_Biochemistry/Enzyme/Active_Site&oldid=3555410.

Structural Biochemistry/Proteins. (2019, March 24). *Wikibooks, The Free Textbook Project*. Retrieved 19:16, July 18, 2019 from en.wikibooks.org/w/index.php?title=Structural_Biochemistry/Proteins&oldid=3529061.

Fujiwara, K., Toda, H., and Ikeguchi, M. (2012) Dependence of a α-helical and β-sheet amino acid propensities on teh overall protein fold type. BMC Structural Biology 12:18. Available at: https://bmcstructbiol.biomedcentral.com/track/pdf/10.1186/1472-6807-12-18

Wikipedia contributors. (2019, July 16). Keratin. In *Wikipedia, The Free Encyclopedia*. Retrieved 17:50, July 19, 2019, from en.Wikipedia.org/w/index.php?title=Keratin&oldid=906578340

Wikipedia contributors. (2019, July 13). Alpha-keratin. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:17, July 19, 2019, from en.Wikipedia.org/w/index.php?title=Alpha-keratin&oldid=906117410

Open Learning Initiative. (2019) Integumentary Levels of Organization. Carnegie Mellon University. In Anatomy & Physiology. Available at: https://oli.cmu.edu/jcourse/webui/syllabus/module.do?context=4348901580020ca6010f804da8baf7ba.

Wikipedia contributors. (2019, July 16). Collagen. In *Wikipedia, The Free Encyclopedia*. Retrieved 03:42, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Collagen&oldid=906509954

Wikipedia contributors. (2019, July 2). Rossmann fold. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:01, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Rossmann_fold&oldid=904468788

Wikipedia contributors. (2019, May 30). TIM barrel. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:46, July 20, 2019, from en.Wikipedia.org/w/index.php?title=TIM_barrel&oldid=899459569

Wikipedia contributors. (2019, July 16). Protein folding. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:30, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Protein_folding&oldid=906604145

Wikipedia contributors. (2019, June 11). Globular protein. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:49, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Globular_protein&oldid=901360467

Wikipedia contributors. (2019, July 11). Intrinsically disordered proteins. In *Wikipedia, The Free Encyclopedia*. Retrieved 19:52, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Intrinsically_disordered_proteins&oldid=905782287

4.3: Tertiary and Quaternary Structures

4.3.1: Tertiary Structure

The tertiary structure of a single chain protein is the overall 3D structure of the protein. A protein of a given primary structure folds to form a 3D structure with embedded secondary structures, super secondary structures and domains. Folded proteins can have a variety of shapes from a roughly spherical or "globular" to a more extended "fibrillar" form. Let's consider the more globular one first. How a protein folds will be discussed in greater detail later, but a more descriptive and simpler view will help us understand the structural features of the folded protein in its tertiary structure.

Start with an unfolded protein. It has a polar backbone with dangling polar charged, polar uncharged, and nonpolar sides attached to the alpha carbon of each amino acid in the primary sequence. On folding, where do these side chains of varying polarity end up? To a first approximation, you may think that a globular protein might fold such that all the hydrophobic amino acid side chains are buried in the interior of the protein, surrounded by other hydrophobic side chains. In a similar fashion, you might expect the polar and charged side chains could be on the surface, exposed to water.

Such a model would be analogous to a **micelle,** which has an almost "perfect" separation of polar (on the surface) and nonpolar atoms (buried). Figure 4.3.1 shows an interactive iCn3D model of a micelle below, which consists of 54 self-associated molecules of dodecylphosphocholine fatty acids.

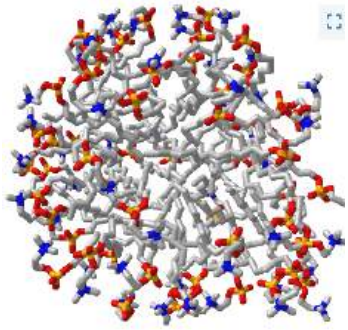


Figure 4.3.1: A Micelle (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...GkJxUQJKdHNfJ7

If protein folding was only that simple! Topologically, it is impossible for a protein to fold in an intramolecular fashion in a strict analogy to the intermolecular aggregation of single chain amphiphiles into a micelle. Consider also that the entire backbone is polar! To a first approximation we would expect the bulk of nonpolar groups would be buried, surrounded by other nonpolar groups. Likewise, we would expect the bulk of polar and charged groups would be on the surface. Figure 4.3.2 shows an interactive iCn3D model of part of human low molecular weight protein tyrosyl phosphatase (1xww). It shows one buried nonpolar side chain (Phe 10) surrounded by essentially all nonpolar side chains of other amino acids.

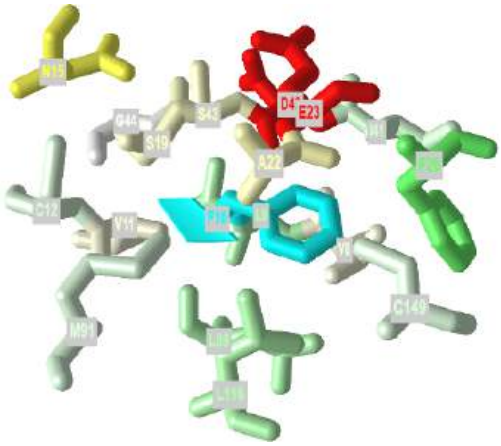


Figure 4.3.2: A buried phenylalanine in low molecular weight protein tyrosyl phosphatase (1xww) (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...LC53qmf4FyszT7

The completely nonpolar Phe 10 side chain is shown in cyan. The atoms with 5 Angstroms are color-coded using the Wimley–White whole residue hydrophobicity scale, which uses not only side chain but peptide bond contributions that are determined experimentally by determining free energy of transfers of groups to nonpolar environments. The color scale runs from cyan (nonpolar) to red (polar/charged. Like a micelle, the protein is roughly spherical.

Noncovalent interactions between atoms within a protein chain help drive protein folding. The noncovalent interactions (also termed intermolecular forces in traditional introductory chemistry classes, include ion-ion, ion-dipole, hydrogen bonds, dipole-dipole, induced dipole - induced dipole (often called London Dispersion Forces) and variants including ion-induced dipole, etc. We generally use the same terms when describing these interactions within and between proteins. Ion-Ion interactions are usually called salt bridges, and induced dipole-induced dipole are often called hydrophobic interactions.

Let's look at some specific interactions with a given protein chain (the light chain of the mouse immunoglobulin G, PDB 1D = 4hdi). Manipulate the iCn3D model in the exercises below and answer the following questions. Use your mouse to hover over amino acids to help in identification.

Exercise 4.3.1

Name the types of interactions between the following side chains in the iCn3D image above. A large images can see by using this link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?DrdnVubujRaihTcr9

a. Leu 52 and Ile 53, Val 63, Phe 67 and Leu 78

b. Arg 24 and Asp 75

c. Asp 170 and Lys 108

d. Tyr 178 and Lys 147

e. Glu 190 and Arg 160

f. Tyr 191 and Phe 214

Answer

Add texts here. Do not delete this text first. a

a. hydrophobic interactions, induced dipole-induced dipole

b. salt bridge, ion-ion

c. salt bridge, ion-ion

d. pi-cation

e. salt bridge, ion-ion

f. Aromatic-aromatic, induced dipole-induced dipole

You can analyze the noncovalent interactions within and between a protein using PIC- protein interactions calculator

In the next exercise, identify the likely hydrogen donors and acceptors in the pairs shown

Exercise 4.3.1

Are these hydrogen bonds

- side chain to side chain?
- main chain to main chain?
- side chain to main chain?

a. Ile 111: Gln 171

b. Gln 6: Thr 107

c. Ile 53: Trp 40

d. Try 37:Thr 97

Answer

a. Ile 111: Gln 171 - side chain to side chain

b. Gln 6: Thr 107 - side chain to side chain

c Ile 53: Trp 40 - main chain to main chain

d. Try 37:Thr 97 - main chain to main chain

4.3.2: A more realistic understanding of noncovalent interactions

But are all the nonpolar side chains buried? How about the polar uncharged and polar charged side chains? What are the preferred dispositions of side chains in proteins as derived from the crystal structure of thousands of proteins? Here are some conclusions from a paper by Pace (Biochemistry. 40, pg 310 (2001).

- On average, about 50% of the amino acids are in secondary structures. On average, there is about 27% alpha helix, and 23% beta structure. Of course, some proteins are almost all alpha-helical, and some are almost all beta structure, but most are a mixture.
- The side chain location varies with polarity. Nonpolar side chains, such as Val, Leu, Ile, Met, and Phe are predominately (83%) in the interior of the protein.
- Charged polar side chains are almost equally partitioned between being buried or exposed on the surface. (54% - Asp, Glu, His, Arg, Lys are buried away from water, a bit startling!)
- Uncharged polar groups such as Asn, Gln, Ser, Thr, Tyr are mostly (63%) buried, and not on the surface (a bit startling).
- Globular proteins are quite compact, with water excluded. The packing density (Vol_{vdw}/Vol_{tot}) is about 0.75, which is like the NaCl crystal and equals the closest packing density of 0.74. This compares to organic liquids, whose density is about 0.6-0.7.

4.3.3: Tertiary structure and pKa Values

If a charged side chain is buried in a protein, you would expect that it would be surrounded, in general, by either oppositely charged side chains, to which it could form an internal salt bridge (ion-ion interaction), or a polar uncharged group with which it could interact through dipole-dipole or, more specifically, H bond interactions. You would also expect that if it were not near an oppositely charged side chain, that it would exist, if buried, in an uncharged state.

Hence the pKa of side chains would be dramatically affected by the nature of its microenvironment (as we have already seen with the pKa of acetic acid in solvents of different polarity). NMR spectroscopy has been used to determine the pKa values of specific side chains in proteins whose crystal structure is known. Pace et al (2009) summarize data on the properties of ionizable side chains in a series of proteins whose structure has been determined. The intrinsic pKa, pKa_{int} or prototypical pKa value for a side chain exposed to water can be determined using a pentapeptide containing the target amino acid X surrounded by 2 Ala on either each side with both the N and C termini of the peptide blocked so they are uncharged. Table 4.3.1 below shows the pKa values of ionizable side chains in a series of proteins compared to that in the control pentapeptide.

Group	Content %	Buried %	pKa int in AAXAA	pKa avg	low pKa	high pKa	# measurements
Asp	5.2	56	3.9	3.5 + 1.2	0.5	9.2	139
Glu	6.5	48	4.3	4.2 + 0.9	2.1	8.8	153
His	2.2	72	6.5	6.6 + 1.0	2.4	9.2	131
Cys	1.2	90	8.6	6.8 + 2.7	2.5	11.1	25
Tyr	3.2	67	9.8	10.3 + 1.2	6.1	12.1	20
Lys	5.9	34	10.4	10.5 + 1.1	5.7	12.1	35

Group	Content %	Buried %	pKa int in AAXAA	pKa avg	low pKa	high pKa	# measurements
Arg	5.1	56	12.3				
C term			3.7	3.3 + 0.8	2.4	5.9	22
N term			8.0	7.7 + 0.5	6.8	9.1	16

Table 4.3.1: pKa values of side chains in actual proteins

A quick glance a the table shows a huge variation in the pKas of ionizable side chains in proteins with the pKa of Asp varying over a range of 8.7 pH units, showing that it can act at physiological pH as either a strong acid or a moderate base. Three majors effects can perturb the pKa of ionizable side chains:

1. Dehydration of the side chain as it is buried in a protein (Born Effect): The stability of a charged group depends on the polarity of the medium in which it exists. Ions are more stable in water than in nonpolar solvents as the water molecules can reorient and interact with the ion through ion-dipole or ion-H bond interactions, which effectively shields the ion from other counter ions. The shielding effect of water is related to the dielectric constant, ε, of the solvent. Coulomb's law can be written as:

$$F = \frac{kQ_1Q_2}{r^2} = \frac{Q_1Q_2}{4\pi\varepsilon r^2}$$

Epsilon is the dielectric constant of the solvent. Water has a higher dielectric constant (80) than nonpolar solvents (4-10) and shields opposing charges more, stabilizing them. Hence the pKa of side chains of those amino acids whose deprotonated states are charged will have their pKa values raised (so they are less acidic) in nonpolar environments. The reverse holds for side chains whose protonated form is charged. Pace cites as an example two mutants of staphylococcal nuclease in which a buried Val 66 is changed either to Asp or Lys. The buried Asp has a pKa of 8.9 compared to 5.5 for the buried Lys. These changes were not compensated for with new charge-charge interactions, so the change can be attributed to the dehydration (or Born) effect.

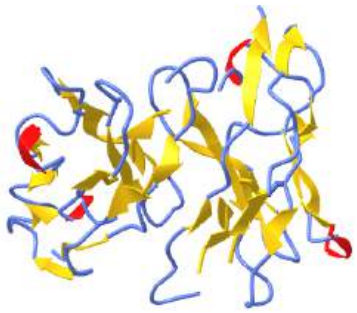
2. Ion-Ion interactions with another charged side chain through Coulombic forces: This effect can be most readily observed at the surface of the protein. Pace cites a study of RNase S that is devoid of Lys and has a pI of 3.5. Five Asp and Glu were replaced on the surface using site-specific mutagenesis with Lys, which changed the pI of the protein to 10.2. At pH 7, the protein without Lys had a charge of -7 while the protein with 5 Lys had a charge of +3. The crystal structures were similar so Coulombic interactions would determine the differences in the pKa of the 11 common side chains. On average the mutant pKas were higher by 0.75 pH units, which makes sense as the mutant had a high pI. Calculated pKa values were similar to those determined by NMR. These data are consistent with the idea that Coulombic interactions are the chief cause of pKa changes in surface side chains.

3. Charge-dipole interactions and H bonds: It should be obvious that charge states of ionizable side chains would be adjusted to optimize H bond (and more generally charge-dipole) interactions in proteins. If the interactions are optimal in the charged state, pKa values for His and Lys would be increased and for Asp, Glu, Cys, and Tyr they would be decreased. Pace cites the buried Asp 76 in RNase T1 in which the Asp is charged but does not form an internal salt bridge. It has a depressed pKa of 0.6 and has 3 H bonds to the side chains of Asn 9, Tyr 11 and Thr 91. Mutants were made to remove the H bonds to see the effect on the pKa of Asp 76. Removing 1, 2, or 3 H bonds changed the pKa to 3.3, 5.1, and 6.4 respectively. The 6.4 value is much higher than the pKint, which can be attributed to the Born effect.

4.3.4: Quaternary Structure

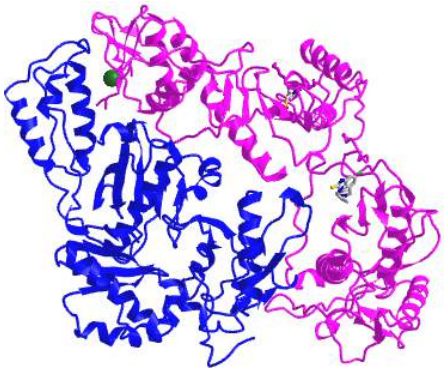
Primary structure is the linear sequence of the protein. Secondary structure is the repetitive structure formed from H-bonds among backbone amide H and carbonyl O atoms. Tertiary structure is the overall 3D structure of the protein. Quaternary structure is the overall structure that arises when separate protein chains aggregate with self to form homodimers, homotrimers, or homopolymers OR aggregate with different proteins to form heteropolymers. Most protein subunits in a larger protein displaying quaternary structure are held together by noncovalent interactions (intermolecular forces), although in some, they are also held together by disulfide bonds (an example includes immunoglobulins).

Figure 4.3.3 shows an interactive iCn3D model of a homodimer, the variable domain of the T cell receptor delta chain (1tvd). Carefully rotate the model to see the two identical chains held together by noncovalent interactions



NCBI iCn3D Figure 4.3.3: variable domain of the T cell receptor delta chain (1tvd) (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...yN6B43P7tvHcR7

Figure 4.3.4 shows an interactive iCn3D model of a heterodimer, reverse transcriptase (1rev). Carefully rotate the model to see the two identical chains held together by noncovalent interactions



NCBI iCn3D Figure 4.3.4: variable domain of the T cell receptor delta chain (1tvd) (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...2mmkgdR9Z2dTQ6

4.3.5: Macromolecule Oligomer Formation and Symmetry

Many proteins are found in aggregated states and hence have quaternary structure. Hemoglobin consists of two alpha and two beta monomers (or protomers) which assemble to produce the biologically relevant heterotetrameric protein. A given monomer can self-aggregate to form homooligomers such as dimers (M_2), trimers (M_3), tetramers(M_4), or or higher oligomers (M_n). The oliogomers often display symmetry with respect to the geometric arrangement of the subunits. Symmetry is an important component of the many kinetic models for catalysis.

Most oligomeric proteins contain protomers that are symmetrically arranged. What mechanism determines whether a monomeric protein forms a homooligomer? Why do they stop at a certain n value? Can proteins be engineered to do so? If mutation can induce oligomer formation, then fewer mutations would be required to produce a symmetric oligomer from subunits since fewer mutations would be required. Why? A single mutation in a single monomer would be represented n times in a single oligomer of n monomers. This fact probably underlies the reason that oligomers display exquisite symmetry. Hence a basic knowledge of the symmetry of protein oligomers is necessary.

In the study of small molecules, chemists describe symmetry through the use of mathematical symmetry operations and elements, which find great use in analysis of structure and in molecular spectroscopy. These concepts are usually first encountered in physical and inorganic chemistry classes. They are a bit complicated so we will offer a limited introduction.

A symmetry operation is a movement of an object like a molecule that leads to an identical, superimposable molecule. Each operation has a symmetry element (point, line, or plane) about which the motion occurs. Some examples are shown in Table 4.3.2 below:

Element (with Jmol link)	Operation
inversion center (i)	projection through center (point) of symmetry of point x,y,z to point -x,-y,-z

proper rotation axis (Cn)	rotation around a Cn axis by 360^o/n where C denotes Cyclic
horizontal (σ_h) and vertical (σ_v) symmetry plane (reflection)	reflection across a horizontal (h) or vertical (v)plane
improper rotation axis (Sn)	rotation around a Sn axis by 360^o/n followed by reflection in plane perpendicular to the axis.

Table 4.3.2: Symmetry Elements and Operations

Luckily for students trying to apply these rules to protein oligomers, biomolecules made up of chiral monomers containing L-amino acids, can not be converted to identical structures using inversion or reflection since the chirality of the monomer would change. For proteins, this would entail an L to D amino acid change. That excludes all but proper rotation axes (Cn) from the list above.

A **point group** is a collection of symmetry operations that define the symmetry about a point. The 4 types of symmetries around a point are those described above: rotational symmetry, inversion symmetry, mirror symmetry, and improper rotation. We'll just consider two versions of rotational symmetry.

4.3.5.1: Cyclic (Cn) - Single Cn rotation axis.

These are very common in proteins that form dimers, trimers, tetramers, etc of identical monomers. These are called **homo n-mers**. In this point group note that the n in Cn is equal to the number of monomers and the angle of rotation is 360^o/n. Figure 4.3.5, adapted from Voet and Voet, shows a cartoon model for C_2 symmetry.

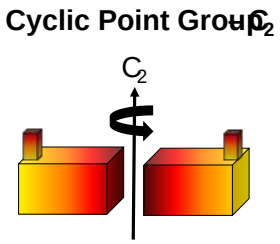
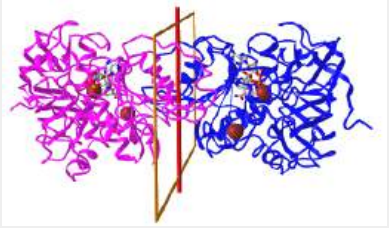
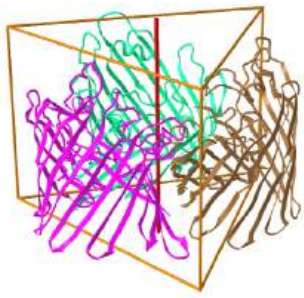
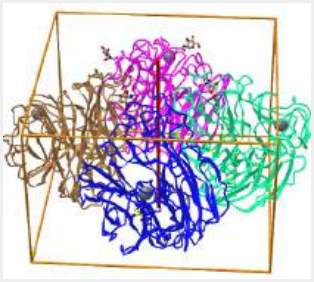


Figure 4.3.5: Cartoon model for C_2 symmetry (adapted from Voet and Voet)

Here are some protein examples with C_n symmetry with the symmetry axis shown as a red vertical line.

Symmetry/ homo n-mer **Protein (**click link for new window)	NCBI iCn3D **Symmetry** click for popup model
C2: homo 2-mer alcohol dehydrogenase (1HSO)	
C3: homo 3-mer porin (2POR)	

C4: homo 4-mer neuramindase (2HTY)	

Symmetric complexes can also have empty interior volumes designed for specific function. Take for example, the human mitochondrial Hsp60-Hsp10 chaperonin complex (6MRD), which assists in protein folding in the mitochondria. It is a hetero 14-mer (A_7B_7) and displays C_7 symmetry, as shown in Figure 4.3.6, an interactive iCn3D model below. The cyan surface is from 7 A monomers while the red is from 7 B monomers

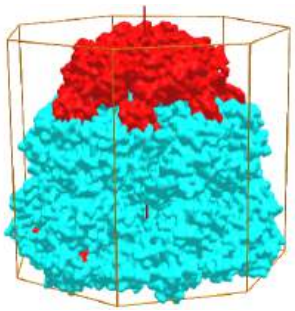


Figure 4.3.6: Hsp60-Hsp10 chaperonin complex (6MRD) with C_7 symmetry (Copyright; author via source). Click the image for a popup (long load) or use this external link:(long load) https://structure.ncbi.nlm.nih.gov/i...jEEY5tu19nXcp6

4.3.5.2: Dihedral (Dn) - Mutually perpendicular rotation axes.

These display higher symmetry as they contain (a) C2 axis(es) perpendicular to a single Cn axis. The minimal number of subunits is n. Most protein oligomers fall into this category. The packing (or asymmetric) unit does not have to be a single monomer but could be a heterodimer. D_n symmetries are more difficult to see but structures in the PDB conveniently provide the type of global symmetry and stoichiometry when symmetry is present.

A D2 point group has 1 C2 axis and 2 perpendicular C2 axes, and 4 monomers (like Hb). These proteins can dissociate into two dimers (such as two α/β dimers for Hb). Note that a different arrangement of 4 monomers could produce an oligomer with C4 symmetry instead of D2. Also note the hemoglobin, a tetramer ($\alpha_2\beta_2$) displays **pseudo** D2 symmetry since the α and β subunits are slightly different in sequence but their folds are almost identical. It also displays C_2 symmetry.

Figure 4.3.7, adapted from Voet and Voet, shows a cartoon model for D_2 symmetry.

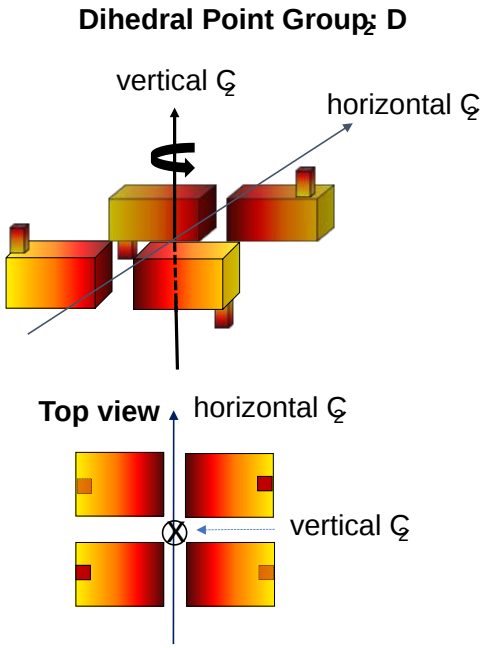
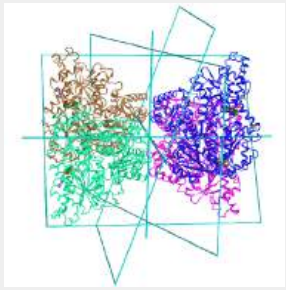
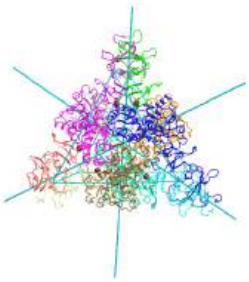
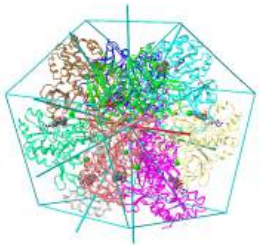


Figure 4.3.7: D_2 symmetry (adapted from Voet and Voet)

A D4 point group has 1 C4 axis and 4 C2 axes, along with 2n=8 subunits. An example of a D4 point group is ribulose bisphosphate carboxylase/oxygenase (RuBisCO) which has 8 subunits (where a subunit, or more technically the asymmetric subunit, is a dimer of a small and large molecular weight protein). This point group could arise from quaternary structure of two C4 tetramers or four C2 dimers.

Here are some homo n-mer protein examples with D_n symmetry.

Symmetry/ homo n-mer **Protein (**click link for new window)	NCBI iCn3D **Symmetry** click for popup model
D2: homo 4-mer phosphofructokinase (4XYJ)	
D3: hetero 12-mer A_6B_6 asparatate carbamoyltransferase **1Q95**	

D5: homo 10-mer glutamine synthetase (2OJW)	

4.3.5.3: Cubic Groups

Now let's consider some cases in which a large number of rotation axes exist that fit the symmetry of simple geometric shapes called the Platonic solids, in which all faces, edges and angles are congruent. The five Platonic shapes, described by Plato, are shown inscribed in spheres in Figure 4.3.8:

Figure 4.3.8:The five Platonic solids (from left to right): tetrahedron (4 sides), octahedron (8 sides), cube (6 sides), icosahedron (20 sides) and the dodecahedron (12 sides). Quantum 4:293 (2020) DOI: 10.22331/q-2020-07-09-293. Creative Commons License: CC BY 4.0

These can also be inscribed in cubes. A tetrahedron and octahedron (actually the overlap of 2 tetrahedron) are shown inscribed in a cube in Figure 4.3.9.

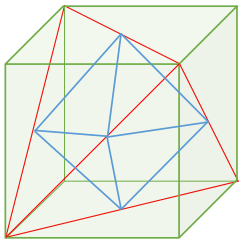


Figure 4.3.9: Tetrahedron (red) and Octahedron (blue) inscribed in a Cube (green).

Since all of the Platonic solids can be inscribed in a cube, they all have basic **cubic symmetry**. Cubes have a total of 13 symmetry axes: three C4 axes passing through the centers of opposite faces, four C3 axes passing through opposite vertices (diagonals), and six C2 axes passing through the centers of opposite edges. The other Platonic solids have related C3 axes (diagonals connecting opposite corners for cubes, diagonals from a vertex to the opposite face for tetrahedrons, lines connecting two opposite faces for octahedron, etc) but they all can be considered to be part of the cubic point group. The rotation axes for each of the Platonic solids are shown in Table 4.3.3below.

# and type rotation axes	# monomers/asymmetric units	Shape
3 C2, 4 C3	12	tetrahedron
6 C2, 4 C3, 3 C4	24	cube/octahedron
15 C2, 10 C3, 6 C5	60	dodecahedron/icosahedron

Table 4.3.3: Cn axes and number of monomers in Platonic solids symmetries

We'll consider examples of tetrahedral, octahedral and icosahedral symmetries, which depend on the overall shape and number of monomers in the functional structure. In some cases, the symmetry of the packed monomers is not perfect. This applies to monomers in clathrin since the monomers can form different structures) and hemoglobin (in which the four subunits are very similar (alpha and beta) but not identical, as described above.

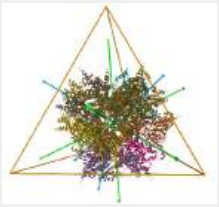
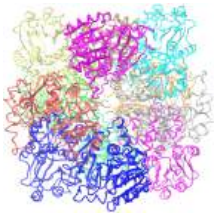
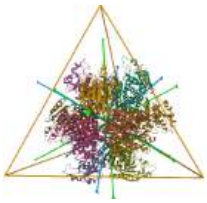
When many monomers form oligomers, it seems that homomers are favored evolutionarily over heteromers. One explanation for this is that in symmetrical arrangements of monomers, there are fewer unique subunit-subunit interfaces that have evolved for complementary of fitness of shape and noncovalent interactions than for heteromers. The same argument applies to the formation of symmetric vs asymmetric arrangements of homo n-mers.

A final note: The words tetrahedron, cube, octahedron, dodecahedron and isocahedron refer to the shape of structures. The words tetrahedral, cubic, octahedral, dodecahedral and isocahedral should be reserved for the type of symmetry. This can be a great source of confusion. Take the case of the E2p, dihydrolipoyl acyltransferase. The homo 24-mer version (from *Azotobacter vinelandii*, 1DPB) has the shape of a cube with octahedral symmetry, while the homo 60-mer (from *Bacillus stearothermophilus*, 1B5S) has the shape of a dodecahedron with the icosahedral symmetry.

It is sometimes difficult to determine the actual biological structure and its symmetry from crystal structure given the artificial packing of the protein in the crystal state. Also, other than for icosohedral virus assemblies, there aren't that many examples of proteins that show tetrahedral, octahedral and dodecahedral symmetries. No structure in the Protein Data Bank is listed with a global symmetry of cubic. We'll describe a few structures with tetrahedral, octahedral and icosohedral global symmetries, knowing that they all fall in the cubic point group.

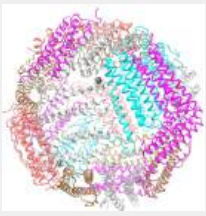
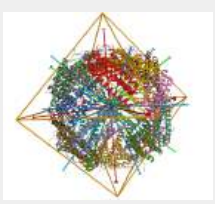
Tetrahedral - 4 sided

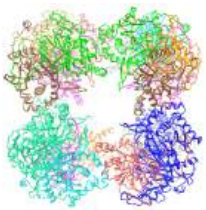
A tetrahedron has four C_3 axes - diagonals from a four corners/vertices to the opposite faces as well as three C2 axes, which are the same as for the cube (since a tetrahedron can be inscribed inside a cube).

Protein (pdb)	**Symmetry/Homo X-mer** (click link)	**Structure - subunits**	**Symmetry**
L-aspartate beta-decarboxylase (2zy2)	tetrahedral/homo 12-mer		
ornithine carbamoyltransferase (1A1S)	tetrahedral, homo 12-mer		

Octahedral - 8 sided

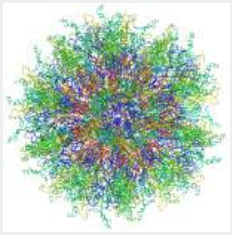
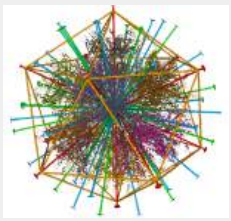
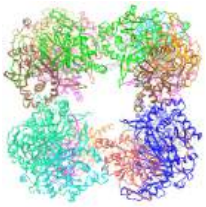
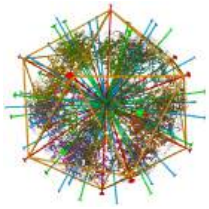
C3 axes - line connecting two opposite faces for an octahedron. Since an octahedron can be aligned with a cube, it also has the same symmetry axes. Examples include human ferritin, octahedral, 24 asymmetric subunits. octahedral, homo 24-mer (4V6B) and dihydrolipoyl acyltransferase from *Azotobacter vinelandii*, octahedral, homo 24-mer (1DPB)

Protein (pdb)	**Symmetry/Homo X-mer** (click link)	**Structure - subunits**	**Symmetry**
human ferritin (4V6B)	octahedral/homo 24-mer		

dihydrolipoyl acyltransferase from *Azotobacter vinelandii* (1DPB)	octahedral/homo 24-mer		

Icosohedral - 20 sided, 60 mers

Examples include adenovirus Ad3 virus, Icosahedral , homo 60-mer 4AQQ and dihydrolipoyl transacetylase from *Bacillus stearothermophilus*, iscosohedral,homo 60-mer (1B5S)

Protein (pdb)	**Symmetry/Homo X-mer** (click link)	**Structure - subunits**	**Symmetry**
Adenovirus Ad3 virus (4AQQ)	Icosahedral/homo 60-mer		
dihydrolipoyl transacetylase from *Bacillus stearothermophilus* (1B5S)	iscosohedral/homo 60-mer		

In summary, our study of symmetry is not in vain since almost half of all proteins appear to form complexes of identical or similar monomers. This probably stabilizes them and adds functional attributes such as kinetic regulation of their activities. Homo n-mers display symmetry, with C_n and D_n being the most common. The cubic symmetries are far less common. What is interesting about cubic symmetries is that structures displaying cubic symmetries have significant empty interior volumes, or "holes" which can be used to encapsulate chemical species. They include the icosahedral viruses, whose interiors contain proteins and viral genomes, and cellular ferritin, which houses up to 4500 Fe^{3+} (oxidized) ions in the form of hydroxide and phosphate complexes. Ferritin delivers Fe^{2+} into cells. Storing iron ions in ferritin prevents its spurious and harmful oxidation and precipitation in nonregulated environments. The encapsulated available volumes have diameters ranging from about 2-4 nm for tetrahedral, 4-8 nm for octahedral and 8-18 nm for icosahedral symmetries.

4.3.6: Filaments

Proteins, especially those involved in cytoskeletal filaments, can form fibers with helical symmetry which differs from those described above since the monomers at the ends of helical fibers, although they have the same tertiary structures as those in the middle of the helical fibers, do not contact the same number of monomers as monomers internal in the oligomer. Hence they have different microenvironments.

Grueninger et al. address the question of whether the process of oligomerization can be programmed into the genome. Can simple amino acid substitutions lead to oligomerization? Oligomerization can be beneficial (formation of cytoskeleton filaments) or detrimental (formation of fibers in sickle cell anemia and prion disease). Oligomers with long half-lives (for example cytoskeletal filaments such as actin and tubulin) and short half-lives (for example proteins that are associated with transient biological activities) are regulated by oligomer formation.

It has long been noted that if a protein chain forms oligomers, then a single amino acid change in the chain would be found n times in an oligomer of n chains. Mutations could either promote chain contact and oligomer formation or dissociation into monomeric or other asymmetric subunit composition if the mutation were in a region involved in subunit association (a contact region).

Experimental work in this field of study is hampered by the fact that mutants made by site-specific mutagenesis to prefer the monomeric state often fail to fold (due to hydrophobic exposure and aggregation). Studies have shown that most contact areas between monomers or other asymmetric units are hydrophobic in nature and the contact regions must be complementary in shape. Obviously, mutations that replace hydrophobic side chains involved in subunit contact with polar, polar charged, or bulkier hydrophobic side chains would inhibit oligomer formation.

Grueninger et al were able to successfully engineer dimer formation and oligomer formation as well. First, consider the simplest case of a mutation in a monomer that can produce a dimer with C2 symmetry. This is illustrated below in Figure 4.3.10. The figure illustrates how a mutation that produces a weak interaction in a monomer could also produce a long helical aggregate (which can't be crystallized) without symmetry (as described above). A mutation at 2 could promote either oligomer helix formation or dimerization.

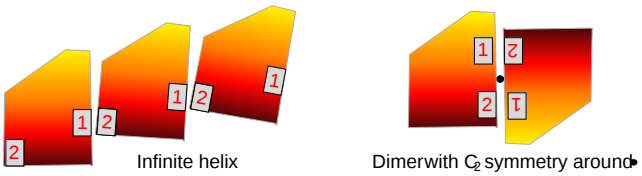


Figure 4.3.10: Mutations causing Dimer with C2 symmetry or Infinite Helix (adapted from Grueninger et al. Science, 319, 206-209 (2008)

It should be noted that mutation could lead to dimer or oligomer formation by producing a more global conformational change in the monomer (not indicated in the example above) which leads to aggregate formation, as we have seen previously in the formation of dimers and aggregates of proteins associated with neurodegenerative diseases (like mad cow disease).

Grueninger produces mutants of two different proteins that showed dimer formation as analyzed by gel filtration chromatography (but did not crystallize so no 3D structures were determined). In addition, the group modified urocanase, a C2 dimer, at 3 side chains to form a tetramer with D2 symmetry. Also, they modified L-rhamnulose-1-phosphate aldolase, a C4 tetramer, at a single position to form an octamer with D4 symmetry. The latter two were analyzed through x-ray crystallography. Their work suggests ways that complex symmetric protein structures arose in nature from simple mutation and evolutionary selection.

4.3.7: References

OpenStax, Proteins. OpenStax CNX. Sep 30, 2016 http://cnx.org/contents/bf17f4df-605c-4388-88c2-25b0f000b0ed@2.

File:Chirality with hands.jpg. (2017, September 16). *Wikimedia Commons, the free media repository*. Retrieved 17:34, July 10, 2019 from commons.wikimedia.org/w/index.php?title=File:Chirality_with_hands.jpg&oldid=258750003.

Wikipedia contributors. (2019, July 6). Zwitterion. In *Wikipedia, The Free Encyclopedia*. Retrieved 21:48, July 10, 2019, from en.Wikipedia.org/w/index.php?title=Zwitterion&oldid=905089721

Wikipedia contributors. (2019, July 8). Absolute configuration. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:28, July 14, 2019, from en.Wikipedia.org/w/index.php?title=Absolute_configuration&oldid=905412423

Structural Biochemistry/Enzyme/Active Site. (2019, July 1). *Wikibooks, The Free Textbook Project*. Retrieved 16:55, July 16, 2019 from en.wikibooks.org/w/index.php?title=Structural_Biochemistry/Enzyme/Active_Site&oldid=3555410.

Structural Biochemistry/Proteins. (2019, March 24). *Wikibooks, The Free Textbook Project*. Retrieved 19:16, July 18, 2019 from en.wikibooks.org/w/index.php?title=Structural_Biochemistry/Proteins&oldid=3529061.

Fujiwara, K., Toda, H., and Ikeguchi, M. (2012) Dependence of a α-helical and β-sheet amino acid propensities on teh overall protein fold type. BMC Structural Biology 12:18. Available at: https://bmcstructbiol.biomedcentral.com/track/pdf/10.1186/1472-6807-12-18

Wikipedia contributors. (2019, July 16). Keratin. In *Wikipedia, The Free Encyclopedia*. Retrieved 17:50, July 19, 2019, from en.Wikipedia.org/w/index.php?title=Keratin&oldid=906578340

Wikipedia contributors. (2019, July 13). Alpha-keratin. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:17, July 19, 2019, from en.Wikipedia.org/w/index.php?title=Alpha-keratin&oldid=906117410

Open Learning Initiative. (2019) Integumentary Levels of Organization. Carnegie Mellon University. In Anatomy & Physiology. Available at: https://oli.cmu.edu/jcourse/webui/syllabus/module.do?context=4348901580020ca6010f804da8baf7ba.

Wikipedia contributors. (2019, July 16). Collagen. In *Wikipedia, The Free Encyclopedia*. Retrieved 03:42, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Collagen&oldid=906509954

Wikipedia contributors. (2019, July 2). Rossmann fold. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:01, July 20, 2019, from https://en.Wikipedia.org/w/index.php?title=Rossmann_fold&oldid=904468788

Wikipedia contributors. (2019, May 30). TIM barrel. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:46, July 20, 2019, from en.Wikipedia.org/w/index.php?title=TIM_barrel&oldid=899459569

Wikipedia contributors. (2019, July 16). Protein folding. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:30, July 20, 2019, from https://en.Wikipedia.org/w/index.php?title=Protein_folding&oldid=906604145

Wikipedia contributors. (2019, June 11). Globular protein. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:49, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Globular_protein&oldid=901360467

Wikipedia contributors. (2019, July 11). Intrinsically disordered proteins. In *Wikipedia, The Free Encyclopedia*. Retrieved 19:52, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Intrinsically_disordered_proteins&oldid=905782287

4.4: Secondary Structural Motifs and Domains

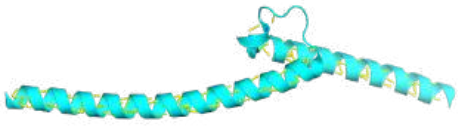
4.4.1: Common Structural Motifs

Given the number of possible combinations of 1^o, 2^o, and 3^o structures, one might guess that the 3D structure of each protein is quite distinctive. This is in general true. However, it has been found that similar substructures are found in proteins. For instance, common secondary structures are often grouped together to form common structural **motifs,** often called **super-secondary structures**. Often the same motif is found in proteins with similar functions (such as proteins that bind DNA, Ca^{2+}, etc). Let's explore some of the common motifs.

4.4.1.1: Alpha-loop-Alpha

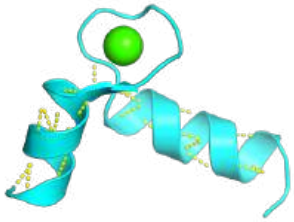
These are found in DNA-binding proteins that regulate transcription and also in calcium-binding proteins, in which the motif is often called the EF hand. The loop region in calcium-binding proteins is enriched in Asp, Glu, Ser, and Thr. Why? The EF hand shown below is from calmodulin.

Figure 4.4.1 shows an interactive iCn3D model of a basic helix-turn-helix from the c-Myc protein (1NKP). The iCn3D model shows the helices interacting with the major grove of DNA, which is shown in spacefill.



NCBI iCn3D Figure 4.4.1: Basic helix-turn-helix from the c-Myc protein (1NKP). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...kDv9DGzWWWoMZ8

Figure 4.4.2 shows an interactive iCn3D model of the "EF hand" from the calcium-binding protein calmodulin (1cll)



NCBI iCn3D Figure 4.4.2: EF hand from Calmodulin (1cll): Secondary Structure Motif. (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...J9jefZYfcpkRu8

The EF Hand can be envisioned as a hand gripping a ball (calcium ion) with the index finger and thumb representing alpha helices, as shown in Figure 4.4.3.

Figure 4.4.3: "EF hand" gripping a ball

The EF hand motif of calmodulin is used in a variety of Ca^{2+} binding proteins. Figure 4.4.4 shows the alignment of the first 50 residues of human calmodulin with four other human calcium-binding proteins. The EF hand (F12-L29) of calmodulin consists of the second half of the first helix (F12-L18), an intervening loop (F19-T28), and the second helix (T29-L29). Sometimes, it is annotated to encompass a larger stretch (8-43)

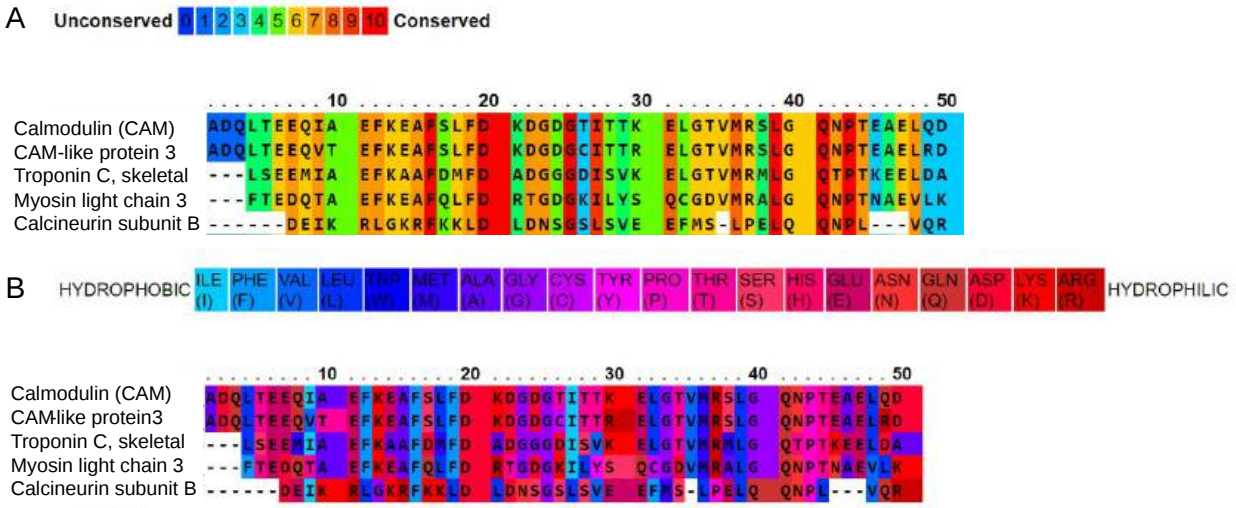
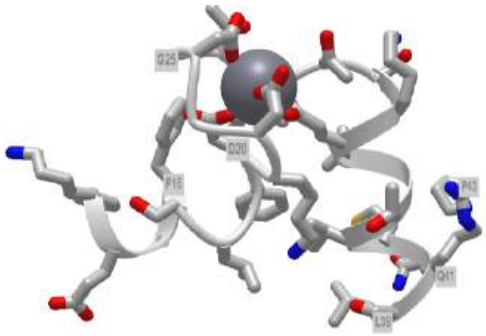


Figure 4.4.4: Amino acid sequence homology of Ca^{2+} binding protein

Part A shows the degree of conservation of amino acids in this first Ca^{2+}-binding EF hand. Part B shows the general conservation of key hydrophobic (F12, F19, I27) as well additionally, those of similar polarity (36 and 39)

Figure 4.4.5 shows an interactive iCn3D model of the first EF hand in human calmodulin with key amino acids labeled.



NCBI iCn3D Figure 4.4.5: First EF hand of Calmodulin (1cll) (Copyright; author via source).

Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...c6YcMn9dj77Wr9

Hover over the amino acid side chains that are coordinating the Ca^{2+} ion. Are they what you would expect?

A linear connectivity "wiring" diagram showing secondary structure connected by connecting regions is shown in Figure 4.4.6. This particular wiring diagram shows a 2-residue beta strand, which is insignificant in length to be considered an actual strand.

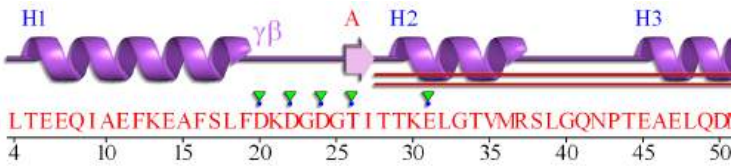


Figure 4.4.6: Wiring Diagram for 1st 50 amino acids in calmodulin

A more complicated 2D topology map is shown in Figure 4.4.7. In this case, it is linear given the small section of amino acids depicted. We will see more complicated 2D topology maps with more complicated structures below.

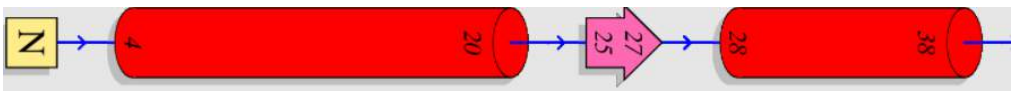


Figure 4.4.7: 2D Topology Map EF Hand Calmodulin

It is presented on its side to save space on this page.

4.4.1.2: Beta-hairpin or beta-turn

This motif is present in most antiparallel beta structures, both as an isolated ribbon and as part of beta sheets.

Figure 4.4.8 shows an interactive iCn3D model of the beta hairpin from bovine pancreatic trypsin inhibitor (1k6u)

Beta hairpin from bovine pancreatic trypsin inhibitor (1k6u)

NCBI iCn3D Figure 4.4.8: Beta hairpin from bovine pancreatic trypsin inhibitor (1k6u) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...eMFdHkGogJHCCA

Figure 4.4.9 shows the 2D homology map for the beta-hairpin.

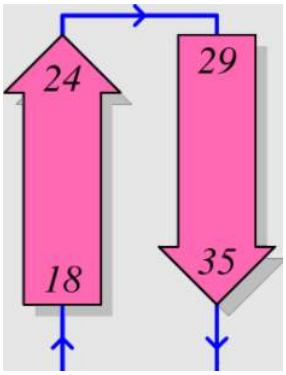


Figure 4.4.8: 2D homology map of beta-hairpin from bovine pancreatic trypsin inhibitor (1k6u)

4.4.1.3: Greek Key

The "Greek Key" symbol represents infinity and the eternal flow of things and resembles in part primitive keys. The Greek Key motif in proteins can be seen in the structure of antiparallel beta sheets in the ordering of four adjacent antiparallel beta strands as shown in Figure 4.4.9. The figure also shows the repetitive Greek key, which you will see many times if you visit Greece and tour its antiquities.

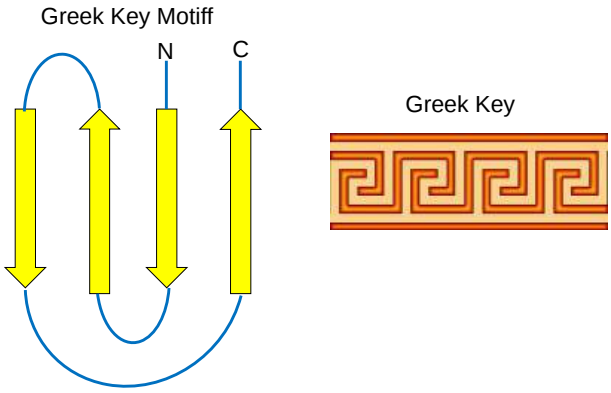


Figure 4.4.9: Greek Key Motif

Figure 4.4.10 shows a partial 2D topology map of Staphylococcus nuclease (2SNS).

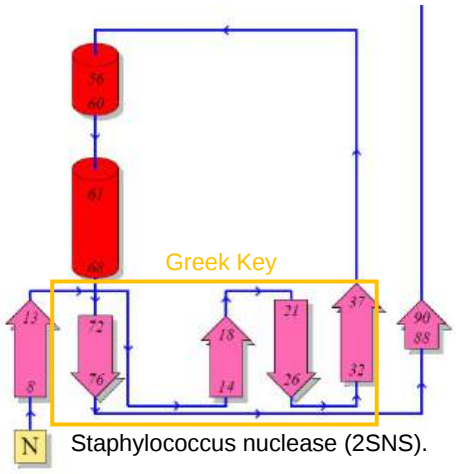
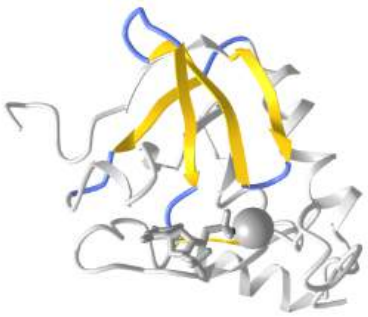


Figure 4.4.10: 2D topology map of Staphylococcus nuclease (2SNS)

NCBI iCn3D Figure 4.4.11 shows an interactive iCn3D model of the Greek Key motif from Staphylococcus nuclease (2SNS). The involved beta strands are shown in yellow.



NCBI iCn3D Figure 4.4.11: Greek Key motif from Staphylococcus nuclease (2SNS) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...x2ef4xpttXrFb9

4.4.1.4: Beta-Alpha-Beta

The motif is a common way to connect two **parallel** beta strands as compared to beta hairpins, which are used to connect antiparallel beta strands.

Figure 4.4.12 shows an interactive iCn3D model of the beta-alpha-beta structure from triose phosphate isomerase (1amk).

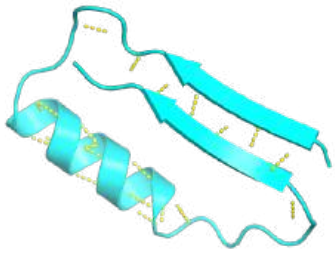


Figure 4.4.12: Beta alpha beta motif from triose phosphate isomerase (1amk) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...8L3VTKtYjVPy97

Figure 4.4.13 shows the 1D wiring diagram for the first beta-alpha-beta motif in triose phosphate isomerase.

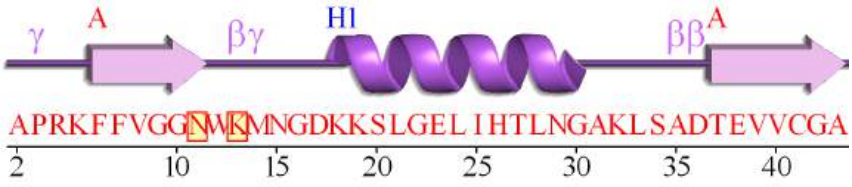


Figure 4.4.13: 1D wiring diagram for the first beta-alpha-beta motif in triose phosphate isomerase

Figure 4.4.14shows the 2D topology diagrams showing this motif.

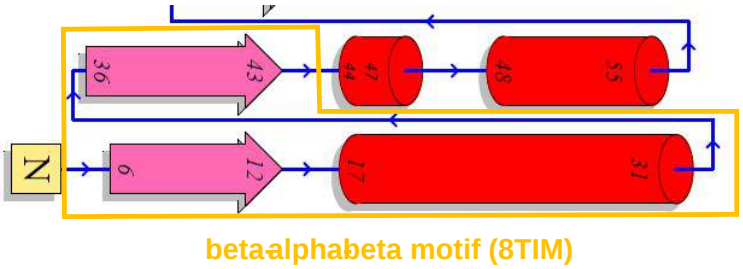


Figure 4.4.14: 2D topology diagrams of beta-alpha-beta motif (8tim).

4.4.2: Larger Structural Motifs - Protein Architecture

Some proteins combine larger secondary and supersecondary structural components, often in a repeated fashion to produce more complex structures. We've seen this with larger twisted sheets and beta barrels, such as the TIM barrel. Let's consider three of these, which can be considered examples of protein architectures without considering connectivity within the protein.

4.4.2.1: The Rossman Fold

Structural motifs can serve particular functions within proteins such as enabling the binding of substrates or cofactors. For example, the Rossmann fold is responsible for binding to nucleotide cofactors such as nicotinamide adenine dinucleotide (NAD^+) as shown in Figure 4.4.15. The Rossmann fold is composed of six parallel beta strands that form an extended beta sheet. The first three strands are connected by α-helices resulting in a beta-alpha-beta-alpha-beta structure. This pattern is duplicated once to produce an inverted tandem repeat which contains six strands. Overall, the strands are arranged in the order of 321456 (1 = N-terminal, 6 = C-terminal). Five stranded Rossmann-like folds are arranged in the sequential order 32145. The overall tertiary structure of the fold resembles a three-layered sandwich wherein the filling is composed of an extended beta sheet and the two slices of bread are formed by the connecting parallel alpha helices.

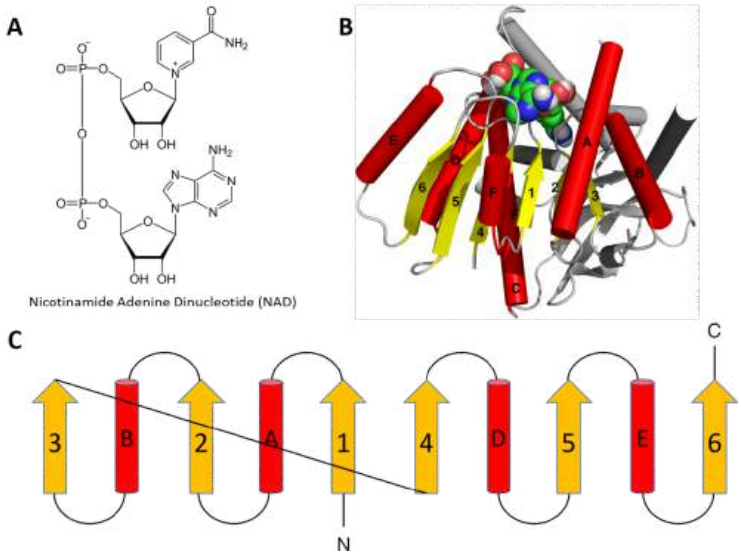


Figure 4.4.15 The Rossman Fold. (A) Structure of Nicotinamide Adenine Dinucleotide (NAD^+) (B) Cartoon diagram of the Rossmann Fold (helices A-F red and strands 1-6 yellow) from *E. coli* malate dehydrogenase enzyme. The NAD^+ cofactor is shown binding as the space filling molecule. (C) Schematic diagram of the six stranded Rossmann fold.

Image modified from: Boghog

One of the features of the Rossmann fold is its co-factor binding specificity. The most conserved segment of Rossmann folds is the first beta-alpha-beta segment. Since this segment is in contact with the ADP portion of dinucleotides such as FAD, NAD and NADP it is also called as an "ADP-binding beta-beta fold".

Figure 4.4.16 shows an interactive iCn3D model of the Rossman fold of malate dehydrogenase (5KKA) from E. Coli. The beta strands (yellow) and connecting alpha helices (red), and coil (blue) of the Rossman fold are shown in the context of the rest of the monomeric version of the protein, which is shown in gray.

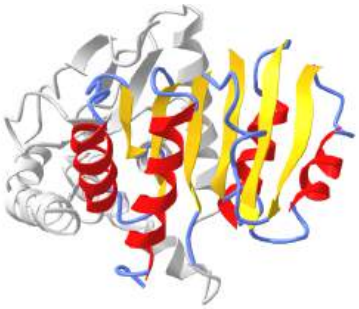


Figure 4.4.16: Rossman fold of malate dehydrogenase (5KKA) from E. Coli (Copyright; author via source). Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/icn3d/share.html?igcXKbc8aU96eG896

4.4.2.2: The TIM barrel revisited

Interestingly, similar structural motifs do not always have a common evolutionary ancestor and can arise by convergent evolution. This is the case with the TIM Barrel, a conserved protein fold consisting of eight α-helices and eight parallel β-strands that alternate along the peptide backbone. It is illustrated in Figure 4.4.17. The structure is named after triosephosphate isomerase, a conserved metabolic enzyme. TIM barrels are one of the most common protein folds. One of the most intriguing features among members of this class of proteins is although they all exhibit the same tertiary fold there is very little sequence similarity between them. At least 15 distinct enzyme families use this framework to generate the appropriate active site geometry, always at the C-terminal end of the eight parallel beta-strands of the barrel.

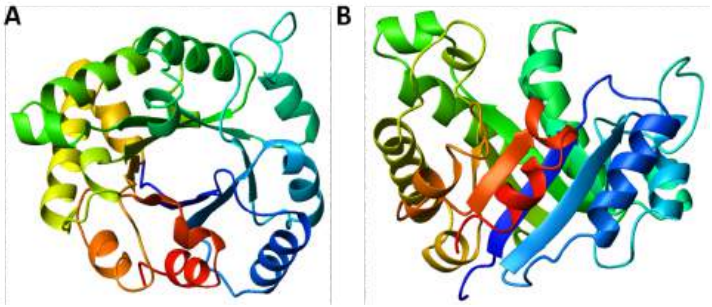


Figure 4.4.17 The TIM Barrel. TIM barrels are considered α/β protein folds because they include an alternating pattern of α-helices and β-strands in a single domain. In a TIM barrel, the helices and strands (usually 8 of each) form a solenoid that curves around to close on itself in a doughnut shape, topologically known as a toroid. The parallel β-strands form the inner wall of the doughnut (hence, a β-barrel), whereas the α-helices form the outer wall of the doughnut. Each β-strand connects to the next adjacent strand in the barrel through a long right-handed loop that includes one of the helices, so that the ribbon N-to-C coloring in the top view (A) proceeds in rainbow order around the barrel. The TIM barrel can also be thought of, then, as made up of 8 overlapping, right-handed β-α-β super-secondary structures, as shown in the side view (B).

Image modified from: WillowW

Although the ribbon diagram of the TIM Barrel shows a hole in the protein's central core, the amino acid side chains are not shown in this representation (Figure 2.26). The protein's core is actually tightly packed, mostly with bulky hydrophobic amino acid residues although a few glycines are needed to allow wiggle room for the highly constrained center of the 8 approximate repeats to fit together. The packing interactions between the strands and helices are also dominated by hydrophobicity and the branched aliphatic residues valine, leucine, and isoleucine comprise about 40% of the total residues in the β-strands.

The figure 4.4.18 below shows an interactive iCn3D model of the TIM barrel (1WYI) from Chapter 4.2).

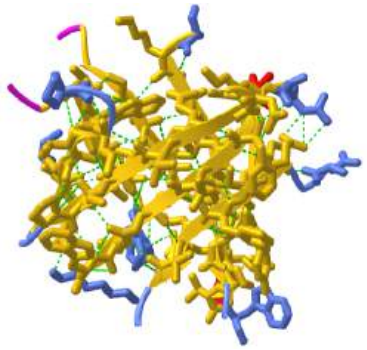


Figure 4.4.18: TIM barrel (1WYI) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...7dki3WjRDtiaRA

As our knowledge continues to increase about the myriad of structural motifs found in nature's treasure trove of protein structures, we continue to gain insight into how protein structure is related to function and are better enabled to characterize newly acquired protein sequences using *in silico* technologies.

4.4.2.3: Beta Helices

These right-handed parallel helical structures consist of a contiguous polypeptide chain with three parallel beta strands separated by three turns forming a single rung of a larger helical structure which in total might contain as many as nine rungs. The intrastrand H-bonds are between parallel beta strands in separate rungs. These seem to prevalent in pathogens (bacteria, viruses, toxins) proteins that facilitate the binding of the pathogen to a host cell.

Figure 4.4.19 shows an interactive iCn3D model of the C-terminal fragment of the phage T4 GP5 beta helix (4osd).

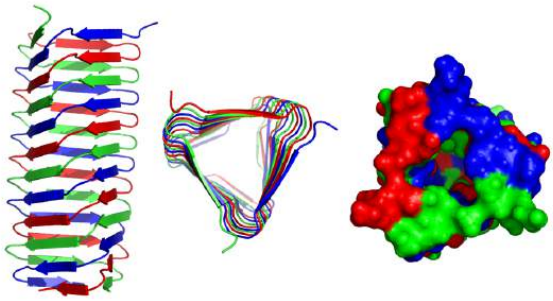


Figure 4.4.19: Beta helix from the C-terminal fragment of the phage T4 GP5 (4osd) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...jmRABoThf5dXJ9

Beta helices and found in the following organisms (with the diseases they cause in humans): *Vibrio cholerae* (cholera), *Helicobacter pylori* (ulcers), *Plasmodium falciparum* (malaria), *Chlamyidia trachomatis* (VD), *Chlamydophilia pneumoniae* (respiratory infection), *Trypanosoma brucei* (sleeping sickness), *Borrelia burgdorferi* (Lyme disease), *Bordetella parapertussis* (whooping cough), *Bacillus anthracis* (anthrax), *Neisseria meningitides* (menigitis) and *Legionaella pneumophilia* (Legionaire's disease).

4.4.2.4: Beta Propellors

Protein with this structure has 4-8 blade-shaped beta sheets arranged around a central axis, forming an active site shaped like a funnel.

Figure 4.4.20 shows an interactive iCn3D model of the C-terminal domain of Tup1 (1ERJ), a yeast transcription factor, which has a seven-bladed beta propeller. Each blade contains a **WD40 repeat** sequence (around 40 amino acids) that often ends in tryptophan-aspartic acid (W-D). The particular protein has four WD dipeptides sequences, shown in sticks colored with CPK colors.

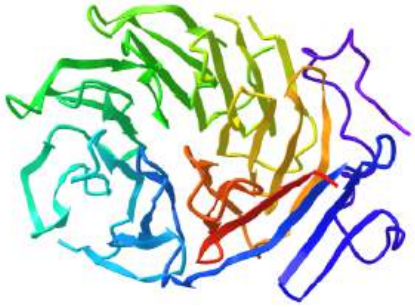


Figure 4.4.20: Seven-bladed beta propeller from the C-terminal domain of Tup1 (1ERJ) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?ydS1vZedDUPupfN28

The funnel provides binding sites for proteins and other molecules, with the ones with more blades usually acting as enzymes.

4.4.3: Domains

Domains are the fundamental unit of 3^{o} structure. Domains can be considered a chain or part of a chain that can independently fold into a stable tertiary structure. Domains are units of structure but can also be units of function. Some proteins can be cleaved at a single peptide bond to form two separate domains. Often, these can fold independently of each other, and sometimes each unit retains an activity that was present in the uncleaved protein. Sometimes binding sites on the proteins are found in the interface between the structural domains. Many proteins seem to share functional and structural domains, suggesting that the DNA of each shared domain might have arisen from the duplication of a primordial gene with a particular structure and function.

Evolution has led toward increasing complexity which has required proteins of new structure and function. Increased and different functionalities in proteins have been obtained with addition of domains to base proteins. Chothia (2003) has defined domain in an evolutionary and genetic sense as "an evolutionary unit whose coding sequence can be duplicated and/or undergo recombination". Proteins range from small with a single domain (typically from 100-250 amino acids) to large with many domains. From recent analyzes of genomes, new protein functionalities appear to arise from the addition or exchange of other domains which, according to Chothia, result from

- duplication of sequences that code for one or more domains
- divergence of duplicated sequences by mutations, deletions, and insertions that produce modified structures that may have useful new properties to be selected
- recombination of genes that result in novel arrangement of domains.

Structural analyzes show that about half of all protein-coding sequences in genomes are homologous to other known protein structures. There appear to be about 750 different families of domains (i.e. small proteins derived from a common ancestor) in vertebrates, each with about 50 homologous structures. About 430 of these domain families are found in all the genomes that have been solved.

Proteins with multiple domains also are more likely **not** to misfold if each domain can fold somewhat autonomously. In addition, they provide a myriad of binding sites which increase the number of biological functions expressed in a single protein. Multidomain proteins can also express multiple catalytic activities, allowing for a reaction product from one domain to diffuse to another catalytic domain (or interface between domains). This would reduce the dimensionality of the search for a substrate from 3D to more of a 1D or 2D search, enormously speeding up the net reaction. The process is often called **substrate channeling**.

Figure 4.4.21 shows an interactive iCn3D model of the three domains of the enzyme pyruvate kinase (1pkn). These include a nucleotide (ADP/ATP) binding domain (blue) made of beta strands, a substrate binding domain (green) in the middle composed of alpha/beta structure, and a regulatory domain (red) composed of alpha/beta structure. These domains were analyzed by a web program called CATH-Gene3D.

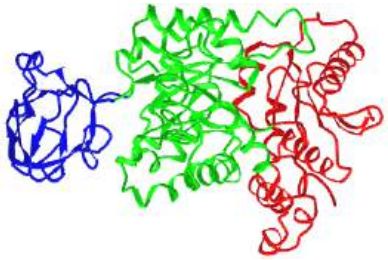


Figure 4.4.21: three domains of the enzyme pyruvate kinase (1pkn) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...3eMGhKSto7ceW7

The **CATH** programs offer a complete classification of protein structure based on the following hierarchy of organization: Class, Architecture, Topology, and Homologous Superfamilies - CATH.

- **C**lass: the highest level of organization which consists of four classes - mainly alpha, mainly beta, alpha-beta, and few secondary structures
- **A**rchitecture (40 types): describes the shape of domain based on secondary structures but doesn't describe how they are connected. Ex: beta barrel, beta propeller
- **T**opology (or fold group, 1233 types): members in topology groups have a common fold or topology in the "core" of the domain structure.
- **H**omologous Superfamilies (2386 types): These groups are homologous in sequence or structure and derive from a common precursor gene/protein.

An alternative computer program, Pfam, shows this enzyme as having 2 major domains, a pyruvate kinase beta barrel domain and a pyruvate kinase alpha/beta domain.

Pfam domains are determined by sequence analysis while CATH are determined by structural comparisons. Domains determined by both programs show about a 75% overlap.

At a simpler level, domains are built from the kinds of motif structures we discussed above. Since proteins are very packed structures, the organizational structure of proteins can be thought of as closely packed motifs, but not all possible combinations are found. For example, if you have one beta hairpin next to another to form a 2-unit Greek key, there are 24 likely ways to connect them but only eight are common. The two below appear to account for more than the sum of the other 22. These are shown in Figure 4.4.22

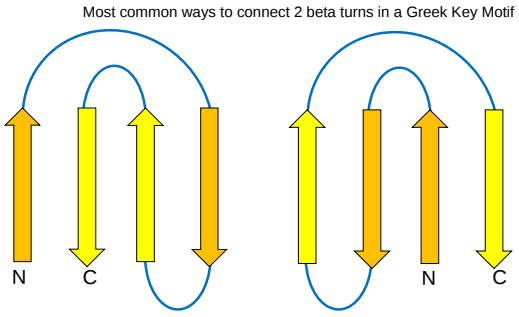
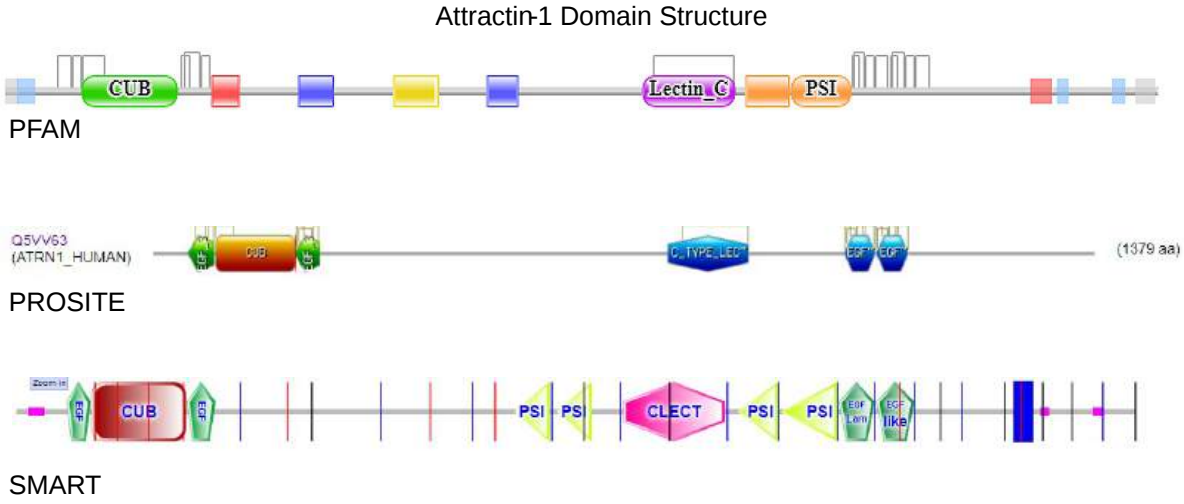


Figure 4.4.22: Most common ways to connect 2 beta turns in Greek Key Motif

Figure 4.4.23 shows an example of the architecture of the multi-domain protein, human Attractin-like protein 1. This protein is an example of a lectin, a carbohydrate-binding protein, which we will explore in a subsequent chapter. It binds Ca^{2+}, so it is considered a C-Lectin. Three different programs were used to analyze the domain structure.



4.4.3.1:

Figure 4.4.23: Architecture of the multi-domain protein, human Attractin-like protein 1

4.5: Protein with Alpha, Alpha-Beta, Beta and Little Secondary Structure

Proteins can also be classified as to the type and extent of secondary structure found in the protein. A detailed description of protein classes can be found at CATH. Here we will describe the basiv types with a few examples offered for each.

4.5.1: Alpha proteins

In these proteins, the core of the protein is composed of alpha helices. The CATH classification shows two major types

a. *Orthogonal bundles*. Example: the Z[beta] Domain of the RNA-editing Enzyme ADAR1 (1xmk), shown in the interactive iCn3D model in Figure 4.5.1.

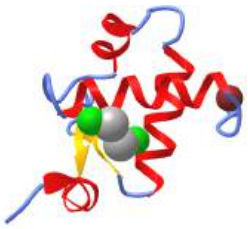


Figure 4.5.1: Alpha protein, Orthogonal bundles: Z[beta] Domain of the RNA-editing Enzyme ADAR1 (1xmk) (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...fxHFM9ErUptjY7

b. *Updown bundles*. Example: Phospholipase A2 from Agkistrodon acutus venom (1mc2), shown in the interactive iCn3D model in Figure 4.5.2.

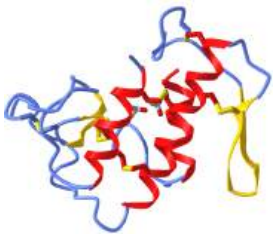


Figure 4.5.2: Alpha protein, Updown bundles: Phospholipase A2 from Agkistrodon acutus venom (1mc2) (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...NGSeK9JAMe3Qh9

4.5.2: Beta proteins

In these proteins, the core is typically antiparallel beta sheets. There are many types of these, including single sheets, rolls, beta barrels, clams, sandwich, propellers, etc (some of which we have already discussed). Here are a few interesting examples.

a. *Roll*. Example: The second SH3 domain from ponsin (2O9S), shown in the interactive iCn3D model in Figure 4.5.3.

Figure 4.5.3: Beta protein Roll: Second SH3 domain from ponsin (2O9S), (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...xYZBXHWpt3njf7

b. *Sandwich*. Example: Mcg immunoglobulin light chain variable domain (4unu), shown in the interactive iCn3D model in Figure 4.5.4.

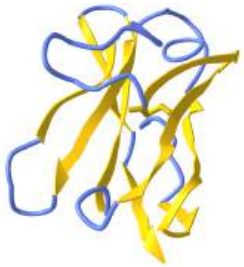


Figure 4.5.4: Beta protein Sandwich - Mcg immunoglobulin light chain variable domain (4unu) (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...4vcY5svrBveJp7

4.5.3: Alpha/Beta proteins

These are the most common class and contain both many beta-alpha-beta motifs with mostly parallel beta strands surrounded by alpha helices. Again, there are many variants for these proteins, including the alpha-beta barrel. We will show two other common ones.

a. *Two layer sandwich*. Example: HIV-1 Nef-SF2 Core Domain (4U5W), shown in the interactive iCn3D model in Figure 4.5.5.

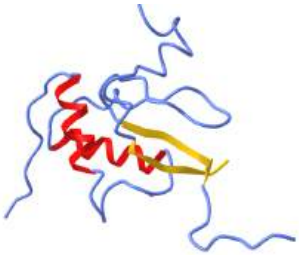


Figure 4.5.5: Alpha-Beta *Two layer sandwich* - HIV-1 Nef-SF2 Core Domain (4U5W) (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...6Fx6J3ecPJbF3A

b. *Three layer sandwich (aba)*. Example: Human biliverdin IX beta reductase (1hdo), shown in the interactive iCn3D model in Figure 4.5.6.

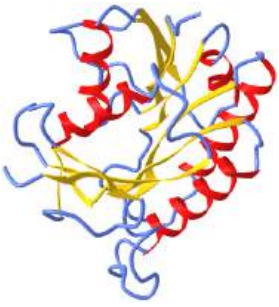


Figure 4.5.6: Alpha-Beta Three layer sandwich (aba) - Human biliverdin IX beta reductase (1hdo) (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...s8GrdasiecmZC8

4.5.4: Little Secondary Structures

These proteins most likely can adopt different conformationdon binding to target molecules. Here are two examples

a. HIV-1 TAT (Transactivating) Protein (1JFW), shown in the interactive iCn3D model in Figure 4.5.7.

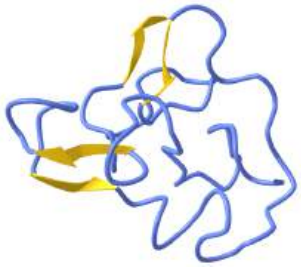


Figure 4.5.7: Few Secondary Structures - HIV-1 TAT (Transactivating) Protein (1JFW) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...zVYofj3HFasWBA

b. Rat rat metallothionein-2 (4MT2), shown in the interactive iCn3D model in Figure 4.5.8.

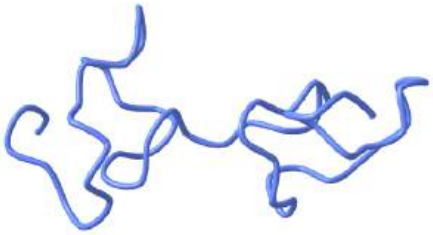


Figure 4.5.8: Few Secondary Structures - Rat rat metallothionein-2 (4MT2) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...fYY4xF2RWd7pD7

4.6: Intrinsically Disordered Proteins

4.6.1: Intrinsically Disordered Proteins (IDPs) and Metamorphic Proteins

Many examples of proteins that are partially or completely disordered but still retain biological function have been found. At first glance this might appear to be unexpected, since how could such a protein bind its natural ligand with specificity and selectivity to express its function? Of course, one could postulate ligand binding would induce conformational changes necessary for function (such as catalysis) in an extreme example of an induced fit of a ligand compared to a "lock-and-key" fit. Decades ago, Linus Pauling predicted that antibodies, proteins that recognize foreign molecules (antigens), would bind loosely to the antigen, followed by a conformational change to form a more complementary and tighter fit. This was the easiest way to allow for a finite number of possible protein antibodies to bind a seemingly endless number of possible foreign molecules. This is indeed one method in which antibodies can recognize foreign antigens. Antibodies that bind to antigens with high affinity and hence high specificity are more likely to bind through a lock and key fit. (Pauling, however, didn't know that the genes that encode the proteins chains in antibodies are differentially spliced and subjected to enhanced mutational rates which allow the generation of incredible antibody diversity from a limited set of genes.)

4.6.1.1: Intrinsically Disordered Proteins (IDPs)

It's been estimated that over half of all native proteins have regions (greater than 30 amino acids) that are disordered, and upwards of 20% of proteins are completely disordered. Regions of disorder are enriched in polar and charged side chains which follow since these might be expected to assume many available conformations in aqueous solutions compared to sequences enriched in hydrophobic side chains, which would probably collapse into a compact core stabilized by the hydrophobic effect. Mutations in the disordered regions tend to preserve the disordered region, suggesting that the disordered region is advantageous for "future" function. In addition, mutations that cause a noncoding sequence to produce a coding one invariably produce disordered protein sequences. Disordered proteins tend to have regulatory properties and bind multiple ligands, in comparison to ordered ones, which are involved in highly specific ligand binding necessary for catalysis and transport. The intracellular concentration of disordered proteins has also been shown to be lower than ordered proteins, possibly to prevent occurrences of inappropriate binding interactions mediated through hydrophobic interactions, for example. Processes to accomplish this include more rapid mRNA and protein degradation and slower translation of mRNA for disordered proteins. For a similar reason, misfolded proteins are targeted for degradation as well. Figure $4.6.1$ shows characteristics of intrinsically disordered proteins.

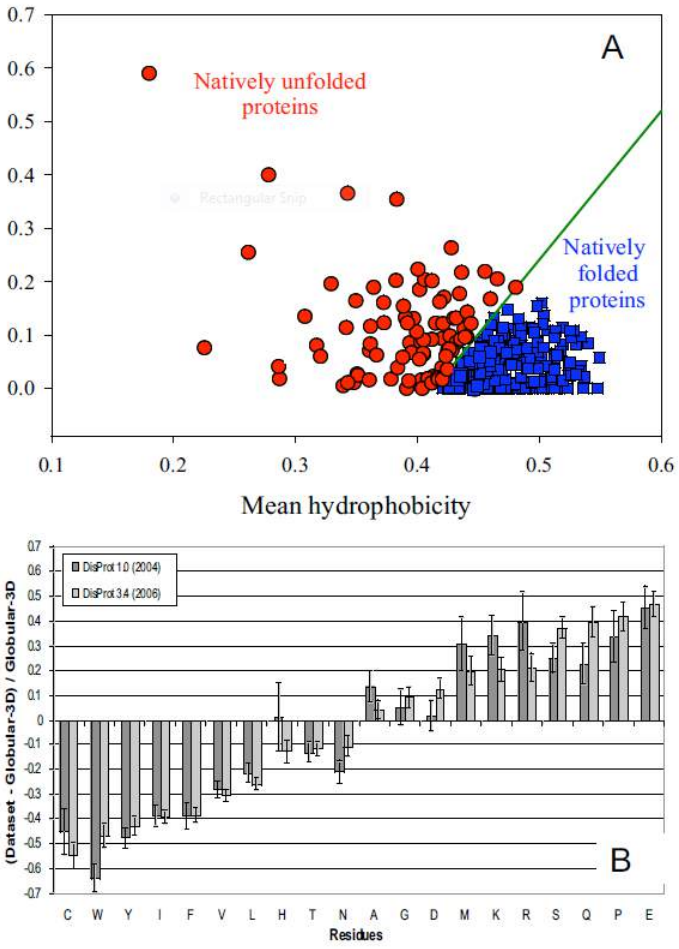


Figure 4.6.1: Characteristics of Intrinsically Disordered Proteins. From open access journal: Dunker, A. et al. BMC Genomics 2008, 9(Suppl 2):S1 doi:10.1186/1471-2164-9-S2-S1. Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0)

Panel A shows the mean net charge vs the mean hydrophobicity for 275 folded and 91 natively unfolded proteins. Panel B shows the relative amino acid composition of globular (ordered) proteins compared to regions of disorder greater than 10 amino acids in disordered proteins. The two different grey bars were obtained with two different versions of the software used to analyze the proteins. Again the graph shows the enrichment of hydrophilic amino acids in disordered proteins.

Many experimental methods can be used to detect disordered regions in proteins. Such regions are not resolved well in X-Ray crystal structures (have high B factors). NMR solution structures would show multiple, and differing conformations. CD spectroscopy likewise would show ill defined secondary structure. In addition solution measurements of size (light scattering, centrifugation) would show larger size distributions for a given protein.

What types of proteins contain disorder? The above experimental and new computational methods have been developed to classify proteins as to their degree of disorder. There appear to be more IDPs in eukaryotes than in archea and prokaryotes. Many IDPS are involved in cell signaling processes (when external molecules signal cells to respond by proliferating, differentiating, dying, etc). Most appear to reside in the nucleus. The largest percentage of known IDPs bind to other proteins and also to DNA. These results suggest that IEPs are essential to protein function and probably confer significant advantages to eukaryotic cells as multiple functions can be elicited from the interaction of a single IEP (derived from a single gene) with different protein binding partners. This would greatly extend the effective genome size in humans, for example, from around 20,000 protein-encoding genes with specified functions, to many more. This doesn't even take into account the increase in functionalities derived from post-translational chemical modifications.

Protein structure is fluid and complex and our simple notions and words to denote proteins as either native or denatured are misguided and constrain our ideas about how protein structure elicits biological function. For example, what does the word "native" mean, if proteins exist in multiple states in vivo and in vitro simultaneously? Dunker et al (2001) have coined the concept "Protein

Trinity" to move past the notion that a single protein folds to a single state which elicits a single function. Rather each of the states in the "trinity", the ordered, collapsed (or molten globule) and extended (random coil) coexist in the cell, as shown in Figure 4.6.2.: Characteristics of Intrinsically Disordered Proteins. Hence all can be considered "native" and all contribute to the function of the cell. A single IDP could bind to many different protein partners, each producing different final structures and functions. IDPs would also be more accessible and hence susceptible to proteolysis, which would lead to a simple mechanism to control their concentrations, an important way to regulate their biological activity. Their propensity to post-translational chemical modification would likewise lead to new types of biological regulation.

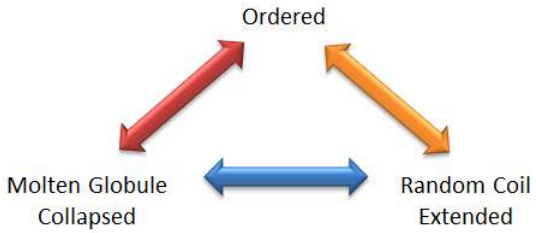


Figure 4.6.2: The Protein Trinity: Ordered, Collapsed and Extended States

These ideas have profound ramifications for our understanding of the expression of cellular phenotype. In addition, a whole new world of drug targets is available by finding drugs that modulate the transitions between ordered, collapsed and extended protein states. Likewise, side effects of drugs might be understood by investigating their effects on these transitions in IDPs that were not initially targeted for anal. Several web database, including PONDR - Predictor of Naturally Occurring Disorder and Database of Protein Disorder are available.

IDPs cover a spectrum of states from fully unstructured to partially structured and include random coils, (pre-)molten globules, and large multi-domain proteins connected by flexible linkers. They constitute one of the main types of protein (alongside globular, fibrous and membrane proteins). Figure 4.6.3 shows the conformational flexibility in SUMO-1 protein (PDB:1a5r), which is a composite of 10 NMR structures. The central part shows relatively ordered structure. Conversely, the N- and C-terminal regions (left and right, respectively) show ‘intrinsic disorder’.

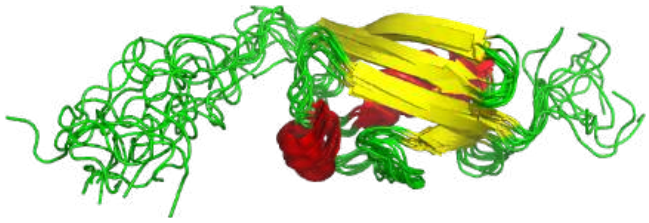


Figure 4.6.3: Conformational flexibility in SUMO-1 protein (1a5r) showing intrinsically disordered regions

History of IDPs

It's interesting to explore the history of our understanding of IDPs. In the 1930s -1950s, the first protein structures were solved by protein crystallography. These early structures suggested that a fixed three-dimensional structure might be generally required to mediate biological functions of proteins. When stating that proteins have just one uniquely defined configuration, Mirsky and Pauling did not recognize that Fisher's work would have supported their thesis with his 'Lock and Key' model (1894). These publications solidified the central dogma of molecular biology in that the sequence determines the structure which, in turn, determines the function of proteins. In 1950, Karush wrote about 'Configurational Adaptability' contradicting all the assumptions and research in the 19th century. He was convinced that proteins have more than one configuration at the same energy level and can choose one when binding to other substrates. In the 1960s, Levinthal's paradox suggested that the systematic conformational search of a long polypeptide is unlikely to yield a single folded protein structure on biologically relevant timescales (i.e. seconds to minutes). Curiously, for many (small) proteins or protein domains, relatively rapid and efficient refolding can be observed *in vitro*. As stated in Anfinsen's Dogma from 1973, the fixed 3D structure of these proteins is uniquely encoded in its primary structure (the amino acid sequence), is kinetically accessible and stable under a range of (near) physiological conditions, and can therefore be considered as the native state of such "ordered" proteins.

During the subsequent decades, however, many large protein regions could not be assigned in x-ray datasets, indicating that they occupy multiple positions, which average out in electron density maps. The lack of fixed, unique positions relative to the

> crystal lattice suggested that these regions were "disordered". Nuclear magnetic resonance spectroscopy of proteins also demonstrated the presence of large flexible linkers and termini in many solved structural ensembles. It is now generally accepted that proteins exist as an ensemble of similar structures with some regions more constrained than others.

Some people differentiate a particular type of IDP called **Intrinsically Unstructured Proteins (IUPs),** which occupy the extreme end of this spectrum of flexibility, whereas IDPs also include proteins of considerable local structure tendency or flexible multidomain assemblies. These highly dynamic disordered regions of proteins have subsequently been linked to functionally important phenomena such as allosteric regulation and enzyme catalysis.

Many disordered proteins have their binding affinity with their receptors regulated by post-translational modification. Hence it has been proposed that the flexibility of disordered proteins facilitates the conformational requirements for binding their modifying enzymes as well as their receptors. Intrinsic disorder is particularly found in proteins implicated in cell signaling, transcription and chromatin remodeling functions. Here are some types or characteristics of IDPs.

Flexible linkers

Disordered regions are often found as flexible linkers or loops connecting domains. Linker sequences vary greatly in length but are typically rich in polar uncharged amino acids. Flexible linkers allow the connecting domains to freely twist and rotate to recruit their binding partners via protein domain dynamics. They also allow their binding partners to induce larger-scale conformational changes by long-range allostery.

Linear motifs

Linear motifs are short disordered segments of proteins that mediate functional interactions with other proteins or other biomolecules (RNA, DNA, sugars etc.). Many roles of linear motifs are associated with cell regulation, for instance in control of cell shape, subcellular localization of individual proteins and regulated protein turnover. Often, post-translational modifications such as phosphorylation tune the affinity (not rarely by several orders of magnitude) of individual linear motifs for specific interactions. Unlike globular proteins, IDPs do not have premade active pockets. Nevertheless, in 80% of IDPs (~3 dozen) subjected to detailed structural characterization by NMR, there are linear motifs termed PreSMos (pre-structured motifs) that are transient secondary structural elements primed for target recognition. In several cases, it has been demonstrated that these transient structures become full and stable secondary structures, e.g., helices, upon target binding. Hence, PreSMos are the putative active sites in IDPs.

Coupled folding and binding

Many unstructured proteins undergo transitions to more ordered states upon binding to their targets. The coupled folding and binding may be local, involving only a few interacting residues, or it might involve an entire protein domain. It was recently shown that the coupled folding and binding allow the burial of a large surface area that would be possible only for fully structured proteins if they were much larger. Moreover, certain disordered regions might serve as "molecular switches" in regulating certain biological functions by switching to ordered conformations upon binding small molecules, nucleic acids or ions.

Disorder in the bound state (fuzzy complexes)

Intrinsically disordered proteins can retain their conformational freedom even when they bind specifically to other proteins. The structural disorder in the bound state can be static or dynamic. In ***fuzzy complexes*** structural multiplicity is required for function and the manipulation of the bound disordered region changes activity. The conformational ensemble of the complex is modulated via post-translational modifications or protein interactions. The specificity of DNA binding proteins often depends on the length of fuzzy regions, which is varied by alternative splicing. Intrinsically disordered proteins adapt many different structures *in vivo* according to the cell's conditions, creating a structural or conformational ensemble.

Therefore, their structures are strongly function-related. However, only few proteins are fully disordered in their native state. Disorder is mostly found in ***intrinsically disordered regions (IDRs)*** within an otherwise well-structured protein. The term intrinsically disordered protein (IDP) therefore includes proteins that contain IDRs as well as fully disordered proteins.

The existence and kind of protein disorder is encoded in its amino acid sequence. As described above, IDPs are characterized by a low content of bulky hydrophobic amino acids and a high proportion of polar and charged amino acids, usually referred to as low hydrophobicity. This property leads to good interactions with water. Furthermore, high net charges promote disorder because of electrostatic repulsion resulting from equally charged residues. Thus disordered sequences cannot sufficiently bury a hydrophobic

core to fold into stable globular proteins. In some cases, hydrophobic clusters in disordered sequences provide clues for identifying the regions that undergo coupled folding and binding (refer to biological roles).

Many disordered proteins reveal regions without any regular secondary structure These regions can be termed as flexible, compared to structured loops. While the latter are rigid and contain only one set of Ramachandran angles, IDPs involve multiple sets of angles. The term flexibility is also used for well-structured proteins, but describes a different phenomenon in the context of disordered proteins. Flexibility in structured proteins is bound to an equilibrium state, while it is not so in IDPs. Many disordered proteins also reveal low complexity sequences, i.e. sequences with over-representation of a few residues. While low complexity sequences are a strong indication of disorder, the reverse is not necessarily true, that is, not all disordered proteins have low complexity sequences. Disordered proteins have a low content of predicted secondary structure.

4.6.2: Silent Single nucleotide polymorphisms (SNPs)

For some amino acids, multiple triplet nucleotide sequences (codons) in the coding regions of a gene for a protein lead to the incorporation of the same amino acid in the protein sequence. Hence two proteins identical in amino acid sequence might have slightly different nucleotide sequences in the gene that encodes them. Such single nucleotide polymorphisms (SNPs) in coding regions were thought to have no effect on the tertiary structure and biological function of a protein if the single nucleotide variation did not lead to the insertion of a different amino acid into the growing peptide chain (i.e the codons were synonymous and the mutations presumably silent with no effect). Recently single nucleotide polymorphisms (SNPs) in the gene for the product of the MDR1 (multidrug resistance 1) gene, P-glycoprotein, was shown to result in a protein with different substrate specificity and inhibitor interactions, and hence a different 3D structure. One possible explanation for this observation is a difference in the rate of translation of the mRNA for this membrane protein. Different rates might lead to different intra- and intermolecular associations, which could lead to different final 3D structures as the protein cotranslationally folds and inserts into the membrane. This would especially be true if two possible structures were close enough in free energy but separated by a significant activation energy barrier, precluding simple conformational rearrangement of one conformation to another.

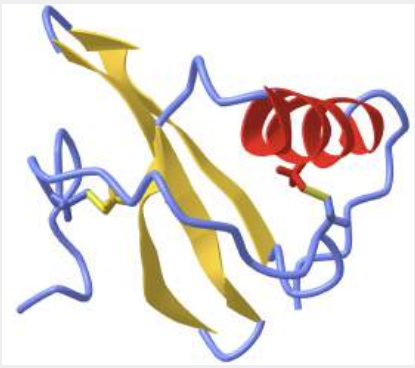
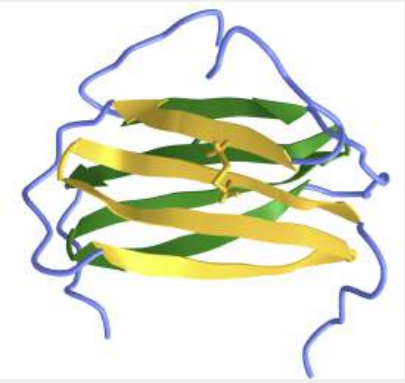
It has been shown in yeast that synonymous mutations (those that don't change the amino acid on mutation of the DNA encoding the particular amino acid) generally have the same effect on the "health" of yeast as do non-synonymous mutations (those that change the amino acid). This rather startling result upends much dogma. Some possible expected effects of synonymous mutation include alteration in gene expression of the mutated gene and possible effects on the stability of the transcribed RNA from the mutated RNA. mRNA levels are lowered from both types of mutations as well as fitness levels of the yeast, as defined by speed of growth.

4.6.3: Metamorphic Proteins

In addition to prion proteins, it appears that many proteins can adopt more than one conformation under the same set of conditions. In contrast to prion proteins, however, in which the formation of the beta-structure variant is irreversible since the conformational change is associated with aggregation, many proteins can change conformations reversibly. Often, these changes do not appear to be associated only with binding interactions that trigger the change. Murzin has described proteins that change conformations on change of the pH (viral glycoproteins), redox state (chloride channel), disulfide isomerization (lysozyme), and bound ligand (RNA polymerase as it initiates and then elongates the growing RNA polymer). He cites two proteins that appear to change state without external signals. These include Mad2, in which the two conformers share an extensive similarity, and Ltn10 (lymphotactin), in which they don't. One form of lymphotactin (Ltn 10) binds to similar lymphokine receptors, while the other (Ltn 40) binds to heparin. Folding kinetics may play a part in these examples as well, as proteins capable of folding to two conformers independently and quickly might prevent misfolding and aggregation that might occur if they had to completely unfold first before a conformational transition. Both Mad2 and Ltn10 alter conformation through transient formations of dimers, which facilitate conformational changes without widespread unfolding. Mutations in Ltn10 can cause the protein to adopt the Ltn40 conformation, Hence primordial "metamorphic" proteins could, by simple mutation, produce new protein functionalities.

Metamorphic proteins, which display large structural changes, usually involving large changes in hydrogen bonding and hence secondary structure, are different from simpler allosteric proteins chose conformation changes are smaller. Few metamorphic proteins have been found but some speculate they could account for as much as 5% of proteins. A wonderful example of such a protein is the human chemokine XCL1 (lymphotactin) protein, an immune regulatory protein. It undergoes a huge transition from a form that has a typical chemokine fold to a dimer that has an extensive beta structure. Ltn lacks one of the two disulfide bonds found in all other chemokines, which allows greater conformation flexibility.

Figure 4.6.3 shows interactive iCn3D model of the solution structures of monomeric and XCL1 (lymphotactin, PDB **2HDM**)

Monomeric XCL1 (lymphotactin PDB **2HDM**)	Dimeric XCL1 (lymphotactin PDB **2JP1)**
	
(Copyright; author via source). Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...b8BwBXkuTAJvx6	(Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...1fuU5jg76RAbVA

4.6.4: References

f17f4df-605c-4388-88c2-25b0f000b0ed@2.

File:Chirality with hands.jpg. (2017, September 16). *Wikimedia Commons, the free media repository*. Retrieved 17:34, July 10, 2019 from commons.wikimedia.org/w/index.php?title=File:Chirality_with_hands.jpg&oldid=258750003.

Wikipedia contributors. (2019, July 6). Zwitterion. In *Wikipedia, The Free Encyclopedia*. Retrieved 21:48, July 10, 2019, from en.Wikipedia.org/w/index.php?title=Zwitterion&oldid=905089721

Wikipedia contributors. (2019, July 8). Absolute configuration. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:28, July 14, 2019, from en.Wikipedia.org/w/index.php?title=Absolute_configuration&oldid=905412423

Structural Biochemistry/Enzyme/Active Site. (2019, July 1). *Wikibooks, The Free Textbook Project*. Retrieved 16:55, July 16, 2019 from en.wikibooks.org/w/index.php?title=Structural_Biochemistry/Enzyme/Active_Site&oldid=3555410.

Structural Biochemistry/Proteins. (2019, March 24). *Wikibooks, The Free Textbook Project*. Retrieved 19:16, July 18, 2019 from en.wikibooks.org/w/index.php?title=Structural_Biochemistry/Proteins&oldid=3529061.

Fujiwara, K., Toda, H., and Ikeguchi, M. (2012) Dependence of a α-helical and β-sheet amino acid propensities on teh overall protein fold type. BMC Structural Biology 12:18. Available at: https://bmcstructbiol.biomedcentral.com/track/pdf/10.1186/1472-6807-12-18

Wikipedia contributors. (2019, July 16). Keratin. In *Wikipedia, The Free Encyclopedia*. Retrieved 17:50, July 19, 2019, from en.Wikipedia.org/w/index.php?title=Keratin&oldid=906578340

Wikipedia contributors. (2019, July 13). Alpha-keratin. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:17, July 19, 2019, from en.Wikipedia.org/w/index.php?title=Alpha-keratin&oldid=906117410

Open Learning Initiative. (2019) Integumentary Levels of Organization. Carnegie Mellon University. In Anatomy & Physiology. Available at: https://oli.cmu.edu/jcourse/webui/syllabus/module.do?context=4348901580020ca6010f804da8baf7ba.

Wikipedia contributors. (2019, July 16). Collagen. In *Wikipedia, The Free Encyclopedia*. Retrieved 03:42, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Collagen&oldid=906509954

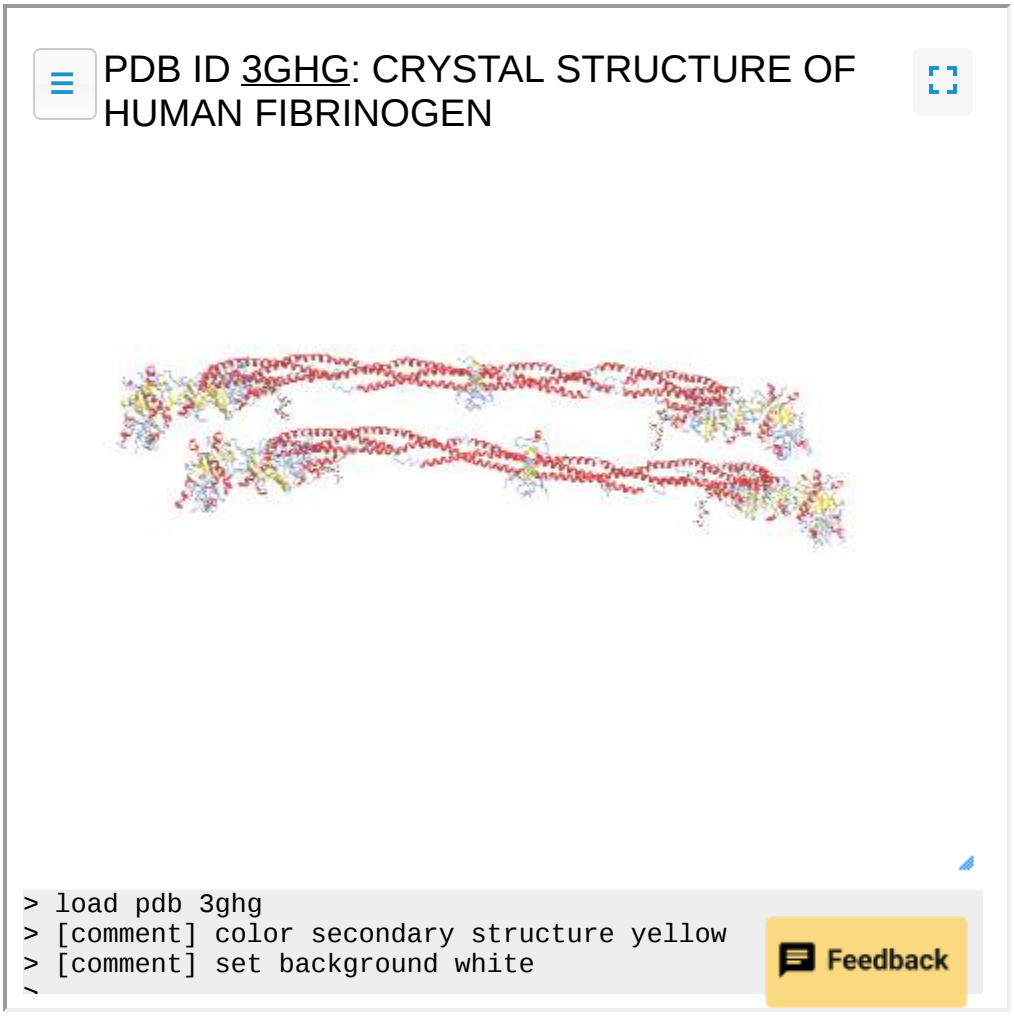
Wikipedia contributors. (2019, July 2). Rossmann fold. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:01, July 20, 2019, from https://en.Wikipedia.org/w/index.php?title=Rossmann_fold&oldid=904468788

Wikipedia contributors. (2019, May 30). TIM barrel. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:46, July 20, 2019, from en.Wikipedia.org/w/index.php?title=TIM_barrel&oldid=899459569

Wikipedia contributors. (2019, July 16). Protein folding. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:30, July 20, 2019, from https://en.Wikipedia.org/w/index.php?title=Protein_folding&oldid=906604145

Wikipedia contributors. (2019, June 11). Globular protein. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:49, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Globular_protein&oldid=901360467

Wikipedia contributors. (2019, July 11). Intrinsically disordered proteins. In *Wikipedia, The Free Encyclopedia*. Retrieved 19:52, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Intrinsically_disordered_proteins&oldid=905782287



- Comprehensive Database for Protein Analysis - Biozon
- SCOP: Structural Characterization of Proteins - Database showing folds, superfamiles, families, and domains

4.7: Fibrillar Proteins

4.7.1: Introduction

Most proteins have a roughly spherical or "globular" tertiary structure. However, there are many proteins that form elongated fibrils with properties like elasticity, which allows deformation on the application of a force and subsequent return to the original state. Elastic molecules must store energy (go to a higher energy state) when the elongating force is applied, and the energy must be released on return to the equilibrium resting structure. Structures that can store energy and release it when subjected to a force have resiliency. Proteins that stretch with an applied force include elastin (in blood vessels, lungs and skins where elasticity is required), resilin in insects (which stretches on wing beating), silk (found in spider web and whose structure we showed in 4.2) and fibrillin (found in most connective tissues and cartilage). Some proteins have high resiliency (90% in elastin and resilin), while others are only partially resilient (35% in silk, which has a tensile strength approaching that of stainless steel).

In contrast to rubber, which has an amorphous structure, which imparts elasticity, these proteins, although they have a dissimilar amino acid sequence, seem to have a common structure inferred from their DNA sequences. In some (like fibrillin), the protein has a folded beta sheet domain, which unfolds like a stretched accordion. Others, like elastin and spider silk, have a beta sheet domain and other secondary structures (alpha-helices and beta turns) along with Pro and Ala repetitions. Scientists are studying these structures to help in the synthesis of new elastic and resilient products.

Fibrous Proteins are characterized by elongated protein structures. These types of proteins often aggregate into filaments or bundles forming structural scaffolds in biological systems. Within animals, the two most abundant fibrous protein families are collagen and α-keratin. Let's start our exploration of fibrillar proteins with these.

4.7.2: Collagen

Collagen is the most abundant protein in mammals, making 25% to 35% of the whole-body protein content. It is found predominantly in the extracellular space within various connective tissues in the body. Collagen contains a unique quaternary structure of three protein strands wound together to form a triple helix. It is mostly found in fibrous tissues such as tendons, ligaments, and skin.

Depending upon the degree of mineralization, collagen tissues may be rigid (bone), compliant (tendon), or have a gradient from rigid to compliant (cartilage). It is also abundant in corneas, blood vessels, the gut, intervertebral discs, and dentin in teeth. In muscle tissue, it serves as a major component of the endomysium. Collagen constitutes one to two percent of muscle tissue and accounts for 6% of the weight of strong, tendinous, muscles. The fibroblast is the most common cell that creates collagen. Gelatin, which is used in food and industry, is collagen that has been irreversibly hydrolyzed. In addition, partially and fully hydrolyzed collagen powders are used as dietary supplements. Collagen also participates in many binding interactions with target proteins in addition to its role in the structure.

The name *collagen* comes from the Greek (*kólla*), meaning "glue", and the suffix *-gen*, denoting "producing". This refers to the compound's early use in the process of boiling the skin and tendons of horses and other animals to obtain glue.

Over 90% of the collagen in the human body is type I. However, as of 2011, 28 types of collagen have been identified, described, and divided into several groups according to the structure they form. The five most common types are:

- Type I: skin, tendon, vasculature, organs, bone (the main component of the organic part of bone)
- Type II: cartilage (the main collagenous component of cartilage)
- Type III: reticulate (the main component of reticular fibers), commonly found alongside type I
- Type IV: forms basal lamina, the epithelium-secreted layer of the basement membrane
- Type V: cell surfaces, hair, and placenta

Let's focus on Type I collagen, which has unusual amino acid composition and sequence:

- Glycine is found at almost every third residue.
- Proline makes up about 17% of collagen.
- Collagen contains many hydroxyproline and hydroxylysine which are formed on post-translational modifications by different enzymes, both of which require vitamin C as a cofactor.
- Some hydroxylysines are glycosylated, mostly with disaccharides.

Figure 4.7.1 shows the post-translational hydroxylations of lysine and proline.

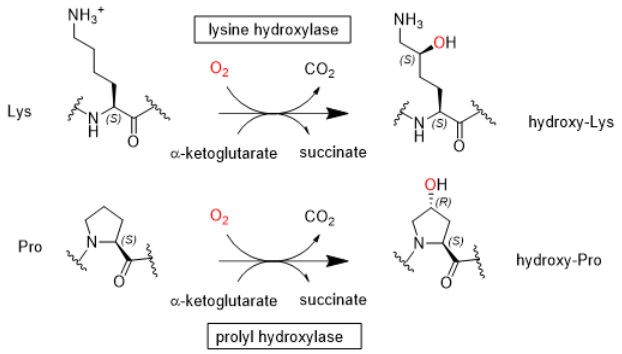
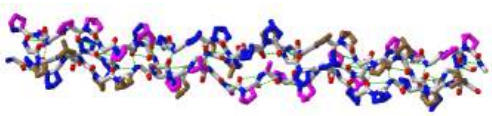


Figure 4.7.1: Post-translational hydroxylations of lysine and proline

Vitamin C deficiency causes scurvy, a serious and painful disease in which defective collagen prevents the formation of strong connective tissue. Gums deteriorate and bleed, with loss of teeth; skin discolors, and wounds do not heal. Prior to the 18th century, this condition was notorious among long-duration military, particularly naval, expeditions during which participants were deprived of foods containing vitamin C. Many bacteria and viruses secrete virulence factors, such as the enzyme collagenase, which destroys collagen or interferes with its production.

Collagen has many $(GXY)_n$ repeats, where G is glycine (Gly), and X and Y are frequently proline (Pro) and hydroxyproline (Hyp). Three strands of collagen self-associate to for a triple-stranded helix with 10 GXY triplets in 3 complete turns of the helix. Others suggest that there are seven triplet units in 2 turns of the stands. Note that the helix of each strand in the triple helix is **not** an alpha helix and has different phi/psi angles. Each strand is "frameshifted" by one amino acid, resulting in a staggered arrangements of the individual stands and helices. The glycines are buried along the central axis so there is no essential hydrophobic core. The X, Y amino acids are solvent-exposed. All the other side chains, both hydrophobic and hydrophilic, are likewise exposed to solvent. Hydrogen bonding occurs between the amide hydrogen of the peptide bond of Gly and the carbonyl O of an X amino acid in another chain.

Figure 4.7.2 shows an interactive iCn3D model of a **triple helical collagen-like peptide** (4Z1R). The main chain atoms, shown in CPK colors, are shown forming hydrogren bonds with neighboring chains. The side chains are color based on the three chains (blue, brown and magenta). Two sets of Pro-HPro-Gly repeats are labeled.



NCBI iCn3D Figure 4.7.2: Triple helical collagen-like peptide (4Z1R). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?TCrji1wPhekypJtc6

4.7.3: α-Keratin

α-keratin is the key structural element making up hair, nails, horns, claws, hooves, and the outer layer of skin. Due to its tightly wound structure, it can function as one of the strongest biological materials and has various uses in mammals, from predatory claws to hair for warmth.

The first sequences of α-keratins were determined by Hanukoglu and Fuchs. These sequences revealed that there are two distinct but homologous keratin families which were named Type I keratin and Type II keratins. There are 54 keratin genes in humans, 28 of which code for type I, and 26 for type II. Type I proteins are enriched in Asp and Glu amino acids, while type II proteins contain more basic amino acids, such as lysine. This differentiation is especially important in α-keratins because in the formation of a keratin dimer, the ***coiled coil***, one protein coil must be type I, while the other must be type II. Even within type I and II, there are acidic and basic keratins that are particularly complementary within each organism. For example, in human skin, K5, a type II α-keratin, pairs primarily with K14, a type I α-keratin, to form the α-keratin complex of the epidermis layer of cells in the skin.

Coiled-coil dimers then assemble into a tetramer of two staggered coiled-coil dimers. Two tetramers can then pack together to form an elongated protofilament, a very stable, left-handed superhelical structure as shown in the figure below. The keratin filaments

stay associated through hydrophobic interactions between apolar residues along the keratin's helical segments. This is illustrated in Figure 4.7.3.

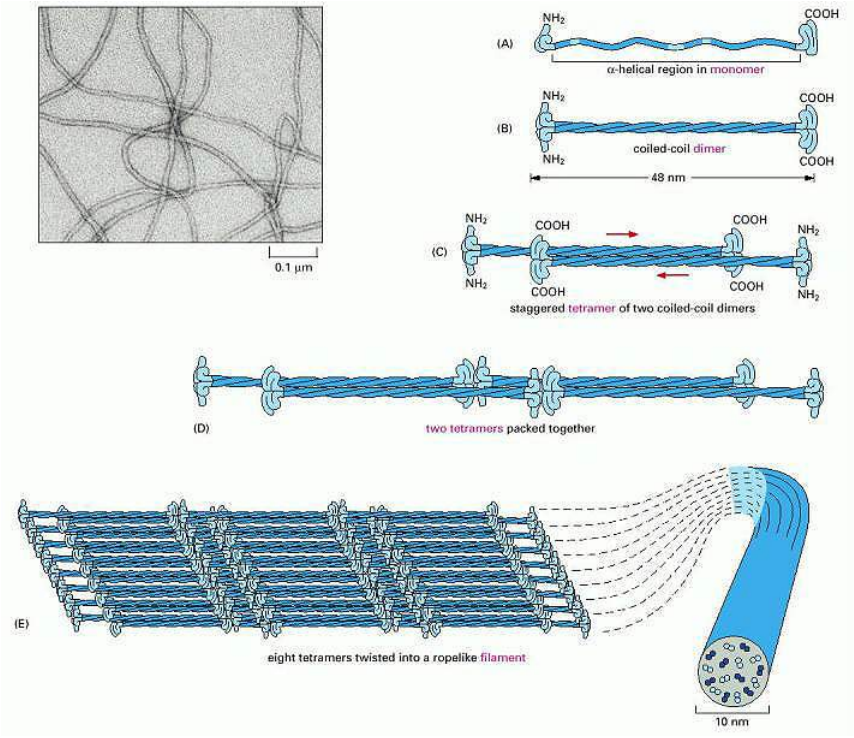
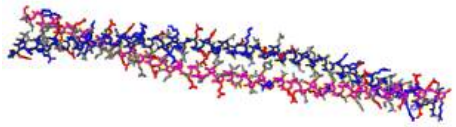


Figure 4.7.3: Assembly of Keratin Fibers. Wiki lectures. https://www.wikilectures.eu/w/Indivi...e_and_function

Initially, two keratin monomers (A) form a coiled coil dimer structure (B) Two coiled coil dimers join to form a staggered tetramer (C), next, the tetramers start to join together (D), ultimately forming a sheet of eight tetramers (E). The sheet of eight tetramers is then twisted into a lefthanded helix forming the final intermediate filament (E) An electron micrograph of the intermediate filament is shown in the upper left hand corner.

Figure 4.7.4 shows an interactive iCn3D model of a dimer of Type I alpha-keratin (magenta backbone) and Type II (blue backbone) (6JFV). Acidic (red) side chains (Asp and Glu) and basic (blue) side chains (Lys) can be seen projecting away from the dimer. Both A and B chains have negative and positive side chains. The A (more acidic) chain in this structure has 5 Lys , 7 Arg, 1 Asp and 16 Glu side chains for a net charge of +12 -17 = -5. The B (more basic) chain in this structure has 8 Lys , 9 Arg, 5 Asp and 12 Glu side chains for a net charge of +17 -17 = 0. It is clearly more basic with 17 Lys and Arg side chains, compared to the A chain with 12. Depending on their 3D orientation, they could present a positive face to the more negative monomer in the dimer.



NCBI iCn3D Figure 4.7.4: Dimer of a Type I and Type II alpha-keratin backbones (6JFV). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...g3L3sQh9dj5pw8

Acidic (red) side chains (Asp and Glu) and basic (blue) side chains (Lys, Arg) can be seen projecting away from the dimer. Both A and B chains have negative and positive side chains. The A (more acidic) chain in this structure has 5 Lys , 7 Arg, 1 Asp and 16 Glu side chains for a net charge of +12 -17 = -5. The B (more basic) chain in this structure has 8 Lys , 9 Arg, 5 Asp and 12 Glu side chains for a net charge of +17 -17 = 0. It is clearly more basic with 17 Lys and Arg side chains, compared to the A chain with 12. Depending on their 3D orientation, they could present a positive face to the more negative monomer in the dimer.

Figure 4.7.5 shows an interactive iCn3D model of a spacefill model of the dimer. Note that a significant fraction of the nonpolar side chains is pointed inward between the two monomers and are much less exposed to solvent.

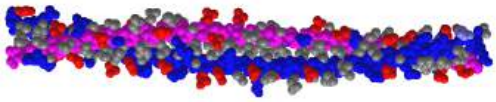


Figure 4.7.5: Dimer of a Type I and Type II alpha-keratin backbones in spacefill (6JFV). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...uyJrw28mZ6x5h7

4.7.4: Elastin

As its name implies, the protein confers elasticity in target structures such as connective tissue and blood vessels. It has low-complexity hydrophobic domains and the protein is cross-linked to form larger structures. It contains repeating hydrophobic amino acid sequences mostly of valine, proline, glycine and alanine, and mimetics of the repeating structure $(LGGVG)_6$ have been studied. This protein also displays significant disorder.

4.7.5: Resilin

The following description of resilin is taken directly from an article under Creative Commons Attribution 4.0 International License available at http://creativecommons.org/licenses/by/4.0/. Balu, R., Dutta, N.K., Dutta, A.K. *et al.* Resilin-mimetics as a smart biomaterial platform for biomedical applications. *Nat Commun* **12,** 149 (2021). https://doi.org/10.1038/s41467-020-20375-x

Resilin-Mimetics

Native elastomeric proteins are biomaterials that have been perfected over billions of years by natural selection to act as molecular springs in a wide range of biological systems to drive unique functions. Among native proteins, resilin is purported to be one of the most efficient elastic proteins known. It is essentially a structural protein, which exists mainly in insect exoskeleton structures and exhibits outstanding resilience and fatigue life. The first description of resilin was made in 1960s as a rubber-like protein observed in locust-wing hinge and dragonfly tendon. Early studies on the composition and structure of resilin revealed the protein to contain about 66% hydrophobic residues (much lower than elastin) with about 45% proline and glycine residues combined. In native state, resilins exist as di- and trityrosine crosslinked hydrogels, and exhibit highly amorphous structures when examined using X-ray diffraction and electron microscopy. During biosynthesis, pro-resilins (uncrosslinked) are secreted from the apical surface of the epidermal cells into the subcuticular space, where they are crosslinked by an enzyme-mediated process to form hydrogels. Over the course of next three decades, resilin was also identified in many other insects and arthropods, including copepods, reduviidae and moth. In arthropods, resilin is largely involved in a number of different functions, including the flexibility and deformability of membranous cuticle and joint systems, the storage of elastic energy in locomotion (jumping, flying, etc.) and catapulting systems, the adaptability to surface topography by multiple contact attachment, and prey catching systems and the reduction of fatigue and damage in feeding and traumatic reproductive system.

The amino acid sequence of resilin was first identified in early 2000s from the CG15920 gene segment of the fruit fly *Drosophila melanogaster*, which opened up new opportunities for synthesis and development of biomimetic resilins. The CG15920 gene comprises N-terminal (exon-1), C-terminal (exon-3) and the middle chitin-binding (exon-2) domains, where exon-1 and exon-3 consist of 18 and 11 copies of consensus amino acid sequences: GGRPSDSYGAPGGGN and GYSGGRPGGQDLG, respectively. The first recombinant pro-resilin or resilin-like polypeptide (RLP), namely Rec1-resilin (encoded from the exon-1 of CG15920 gene) was synthesized in mid-2000s as a water soluble polypeptide expressed in the bacteria *Escherichia coli*. The synthesized pro-resilin was photo-crosslinked (dityrosine) using a ruthenium-persulfate crosslinking system to form hydrogels, which exhibited 97% resilience, outperforming native resilin dissected from dragonfly tendon (92%), natural elastin (90%) and synthetic polybutadiene rubber (80%)

The synthesized RLPs have several advantages over other elastomeric polypeptides, such as elastin-like polypeptides (ELPs), silk-like polypeptides (SLPs) and collagen-like polypeptides (CLPs). These include:

- Unique sequences rich in uncharged, polar amino acids and devoid of canonical hydrophobic residues, and contain high proportions of glycine- and proline-rich segments.
- Average negative hydropathy index.
- Intrinsically disordered protein structure with rapidly-interchangeable conformational ensemble in physiological conditions.

- Multi-stimuli (pH, temperature, ions, mechanical stress, other molecules, etc.) responsiveness, including dual-phase transition behavior (existence of both upper critical solution temperature, UCST and lower critical solution temperature, LCST).
- Low stiffness, high extensibility, outstanding resilience and excellent fatigue life.
- No inflammatory response.

Figure 4.7.6 shows possible transitions in the resilin protein which demonstrate such resiliency.

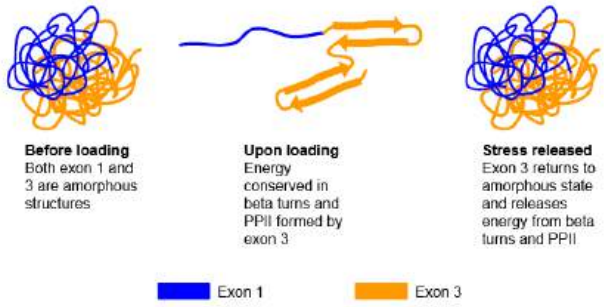


Figure 4.7.6: Transitions in the resilin protein By MTLE4470 grp10 ss - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/inde...curid=58499811

Given its disordered nature, there are no available PDB structures for resilin.

4.7.6: Fibrinogen

This very large molecule is a hexamer of three monomers (Aα, Bβ and γ) each present in two copies ($\alpha_2\beta_2\gamma_2$). Disulfide bond connect one structural unit (αβγ) with another. When two fibrinopeptides (FpA and FpB) are cleaved from the amino ends of the α and β chains by the clotting enzyme thrombin, small structural "knobs" form that bind to "holes" on another fibrinogen, causing the formation of large fibrils of fibrin clots. Figure 4.7.7 shows an interactive iCn3D model of human fibrinogen (3ghg). The "floppy" parts of the alpha chain (*α*C region) and FpA and FpB peptides are not shown as they were not resolved (due to their disorder) in the crystal structure.

NCBI iCn3D Figure 4.7.7: Human fibrinogen (3ghg). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...S3dXzzsrkujKZ8

It is a very long, flexible molecule. The two alpha chains are shown in magenta and light green, the beta chains as dark blue and gray, and the gamma chains as brown and orange. Note the helical chains are actually *alpha*-helical for this molecule.

Figure 4.7.8 shows the domain structure and hints at the flexibility of this long molecule, which is required to form a fibrous mesh clot as well as to be accessible to an enzyme, plasmin, which cleaves fibrin clots, facilitating their removal. The A*α*, B*β* and *γ* are shown in blue, red and green, respectively. Carbohydrates are shown in orange in the spacefilling model (b).

Figure 4.7.8: Domain Structue of FibrinogenKöhler S, Schmid F, Settanni G (2015) The Internal Dynamics of Fibrinogen and Its Implications for Coagulation and Adsorption. PLoS Comput Biol 11(9): e1004346. https://doi.org/10.1371/journal.pcbi.1004346. Creative Commons Attribution License.

The coiled coils contain mostly alpha helical structures. The central E region is where the N-terminal ends of all of the fibrinogen chains are located and where the fibrinopeptides A (FpA, 16 residues) and B (FpB, 14 residues) are located, where they will be cleaved by thrombin in clot formation. The C-terminals of the chains are located at the distal D region which houses two C domains. After cleavage of the fibrinopeptides, conformational changes occur in the central E region to produce the "knobs" A (with starting sequence Gly-Pro-Arg) and B (Gly-His-Arg). These knobs bind through noncovalent interactions corresponding "holes" a and b at the distal D regions of another fibrinogen to form a dimer and subsequently a fibrin clot.

4.7.7: Myosin Heavy Chain

Myosin is a chief component of muscles and in complex with actin and other proteins allows muscle contraction. It has 2 clear domains and an elongated rod-like tail sequence, which has 28-residue repeats of 4 heptapeptides, characteristic of alpha-helical proteins that form the coiled-coils quaternary structure. Figure 4.7.9 shows the domain structure of human myosin heavy chain 1. The orange represents the motor domain of the protein, which binds and hydrolyzes ATP, providing the free energy that powers muscle contraction.

Figure 4.7.9: Domain structure of human myosin

Figure 4.7.10 shows a cartoon of myosin heavy chains (blue) associated with myosin light chains, and how they interact with actin in the actinomyosin complex in striated muscle cells. The motor domain also binds actin filaments. In a simplistic way, the myosin thick filaments can be considered to slide back and forth in muscle contraction and extension.

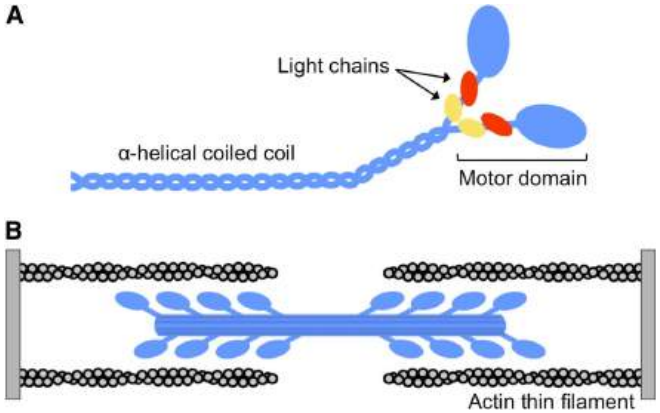
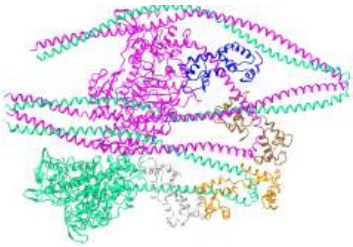


Figure 4.7.10: Myosin-Actin Complex. Lee, L.A., Karabina, A., Broadwell, L.J. *et al.* The ancient sarcomeric myosins found in specialized muscles. *Skeletal Muscle* **9,** 7 (2019). https://doi.org/10.1186/s13395-019-0192-3Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/

Myosin can exist in two major conformations. One is the "6S" (extended tail) form that assembles into myosin filaments, which interacts with actin as shown in Figure 4.7.10 to transduce the chemical energy from ATP hydrolysis into mechanical forces and filament sliding. The other is the "10S" conformation which is folded on itself. The heads interact with each other and the tail. In this compact form, all necessary steps (ATP cleavage, filament assembly, actin activation) required for actin/myosin- mediated contraction are inhibited.

Figure 4.7.11 shows an interactive iCn3D model of an *inactive* (nonextended) conformation of myosin heavy chain (6xe9) from smooth muscle. The long cyan and green chains are the myosin (II) heavy chains from smooth muscle.



NCBI iCn3D Figure 4.7.11: Inactive (nonextended) conformation of myosin heavy chain (6xe9). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...thwKbyQQ5ezB96

References

OpenStax, Proteins. OpenStax CNX. Sep 30, 2016 http://cnx.org/contents/bf17f4df-605c-4388-88c2-25b0f000b0ed@2.

File:Chirality with hands.jpg. (2017, September 16). *Wikimedia Commons, the free media repository*. Retrieved 17:34, July 10, 2019 from commons.wikimedia.org/w/index.php?title=File:Chirality_with_hands.jpg&oldid=258750003.

Wikipedia contributors. (2019, July 6). Zwitterion. In *Wikipedia, The Free Encyclopedia*. Retrieved 21:48, July 10, 2019, from en.Wikipedia.org/w/index.php?title=Zwitterion&oldid=905089721

Wikipedia contributors. (2019, July 8). Absolute configuration. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:28, July 14, 2019, from en.Wikipedia.org/w/index.php?title=Absolute_configuration&oldid=905412423

Structural Biochemistry/Enzyme/Active Site. (2019, July 1). *Wikibooks, The Free Textbook Project*. Retrieved 16:55, July 16, 2019 from en.wikibooks.org/w/index.php?title=Structural_Biochemistry/Enzyme/Active_Site&oldid=3555410.

Structural Biochemistry/Proteins. (2019, March 24). *Wikibooks, The Free Textbook Project*. Retrieved 19:16, July 18, 2019 from en.wikibooks.org/w/index.php?title=Structural_Biochemistry/Proteins&oldid=3529061.

Fujiwara, K., Toda, H., and Ikeguchi, M. (2012) Dependence of a α-helical and β-sheet amino acid propensities on teh overall protein fold type. BMC Structural Biology 12:18. Available at: https://bmcstructbiol.biomedcentral.com/track/pdf/10.1186/1472-6807-12-18

Wikipedia contributors. (2019, July 16). Keratin. In *Wikipedia, The Free Encyclopedia*. Retrieved 17:50, July 19, 2019, from en.Wikipedia.org/w/index.php?title=Keratin&oldid=906578340

Wikipedia contributors. (2019, July 13). Alpha-keratin. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:17, July 19, 2019, from en.Wikipedia.org/w/index.php?title=Alpha-keratin&oldid=906117410

Open Learning Initiative. (2019) Integumentary Levels of Organization. Carnegie Mellon University. In Anatomy & Physiology. Available at: https://oli.cmu.edu/jcourse/webui/syllabus/module.do?context=4348901580020ca6010f804da8baf7ba.

Wikipedia contributors. (2019, July 16). Collagen. In *Wikipedia, The Free Encyclopedia*. Retrieved 03:42, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Collagen&oldid=906509954

Wikipedia contributors. (2019, July 2). Rossmann fold. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:01, July 20, 2019, from https://en.Wikipedia.org/w/index.php?title=Rossmann_fold&oldid=904468788

Wikipedia contributors. (2019, May 30). TIM barrel. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:46, July 20, 2019, from en.Wikipedia.org/w/index.php?title=TIM_barrel&oldid=899459569

Wikipedia contributors. (2019, July 16). Protein folding. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:30, July 20, 2019, from https://en.Wikipedia.org/w/index.php?title=Protein_folding&oldid=906604145

Wikipedia contributors. (2019, June 11). Globular protein. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:49, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Globular_protein&oldid=901360467

Wikipedia contributors. (2019, July 11). Intrinsically disordered proteins. In *Wikipedia, The Free Encyclopedia*. Retrieved 19:52, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Intrinsically_disordered_proteins&oldid=905782287

Comprehensive Database for Protein Analysis - Biozon

SCOP: Structural Characterization of Proteins - Database showing folds, superfamiles, families, and domains

---

4.8: Protein Folding and Unfolding (Denaturation) - Dynamics

4.8.1: Introduction

We've seen many static and rotatable images of lipid aggregates (the micelle) as well as proteins. We have learned some rules about the disposition of amino acid side chains in a folded proteins. However, when we think about how proteins fold, we have to think dynamically as well as thermodynamically. Figure 4.8.1 shows a fun but clearly unrealistic animation of how a protein might fold from an unfolded state with exposed hydrophobic side chains (orange) to a folded state when they are mostly, but not fully, buried.

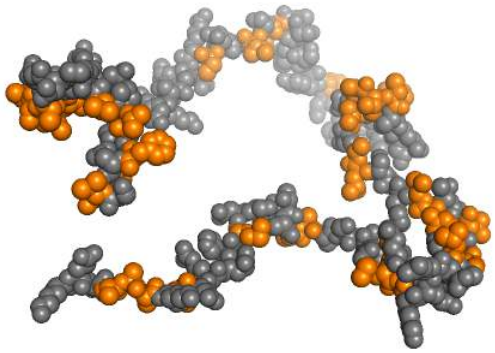


Figure 4.8.1: A morph showing an unfolded protein collapsing to the folded state. (CC BY-NC; Henry Jakubowski via LibreTexts)

Luckily we have the tools of molecular dynamics (MD) at our fingertips which helps us imagine how these processes take place and concomitantly how to probe protein folding experimentally.

As protein folding occurs in 3D, let's explore a free energy (G) landscape for folding from an extraordinarily large number of unfolded states of higher free energy to a single low energy folded state. Two such landscapes, created by Ken Dill's group, are shown in Figure 4.8.2.

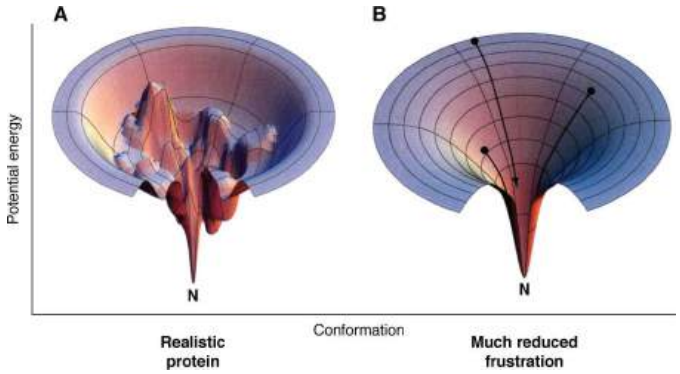


Figure 4.8.2: Free energy protein folding landscapes. Gershenson et al. JBC Reviews 295, 15-33 (2020). Open Access. https://doi.org/10.1074/jbc.REV119.006794. Creative Commons BY 4.0 license

The right image B shows a simplistic ("much-reduced frustration") funnel view with a simple one-step path for any unfolded protein to reach a single global free energy minimum state, a process that occurs without any intermediates. A more realistic view is shown in A in which there are a series of local minimums and one global minimum. Like in any activation energy curve you have encountered in chemistry courses, traversing from a local minimum to other local minimums or the global minimum is possible if enough energy is provided to overcome the activation energy for that step. Panel A implies that are many intermediates on the road to the final global free energy minimum native state but that local minimum could be populated and either stable or metastable depending on their activation energies. Intermediates might be "trapped" in these local minima. The view also conforms to our view that proteins are conformationally flexible and can adopt a variety of conformations. This is especially true for intrinsically disordered proteins.

Now let's add some additional complexity. A protein on the path to a folded state has more hydrophobic exposure than the native state, so you would expect that it could aggregate with other self proteins and form intermediates and end products off of the

normal folding pathway. These are illustrated the Figure 4.8.3 which shows the "normal" protein folding (blue) and a misfolding (orange) landscape. The misfolded landscape is populated with amorphous aggregates, oligomers, protofibrils and fibrils.

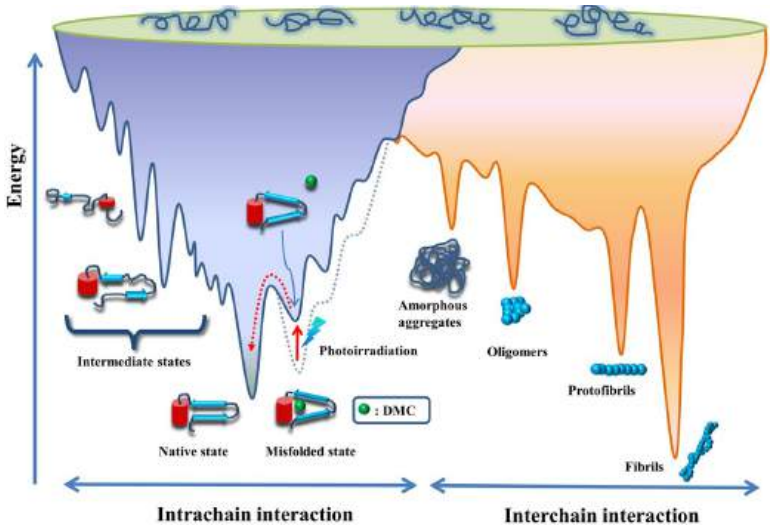


Figure 4.8.3: Protein Folding and Misfolding Landscape Eric H.-L. Chen et al. Scientific Report 7: 8691 DOI:10.1038/s41598-017-08385-0 . Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

We will discuss protein folding done in the lab (in vitro) as well as protein folding in the cell (in vivo). In vitro folding is done in very defined conditions, typically using low concentrations of small proteins to minimize misfolding opportunities. Folding in vivo occurs as a protein is being made on a ribosome. It also occurs when a fully-folded protein misfolds (such as during fevers in disease states) and has been prevented from folding by association with other molecules. Folding in vivo is often assisted by other proteins called folding chaperones.

In either case, given the number of possibly nonnative states, it is amazing that proteins fold to the native state at all, let alone in a reasonable time frame. Consider this greatly simplified view of protein folding for a protein containing 100 amino acids. If each amino acid can adopt only 3 possible conformations, the total number of conformations could be $3^{100} = 5 \times 10^{47}$. Assuming that it would take 10^{-13} s to change each conformation, the time required to "test" all conformations would be 5×10^{34} s or 10^{27} years, longer than the age of the universe (14×10^{9} yr). Yet the protein can fold within seconds. This paradox is called the **Levinthal paradox**, after Cyrus Levinthal.

Lubert Stryer, (in his classic Biochemistry text), shows a way out of this dilemma by using an analogy of a monkey sitting at a typewriter, and typing this line out of Hamlet: "Me thinks it is like a weasel." Random typing would produce that line after 10^{40} keystrokes on average, but if the correct letters were maintained, the number of keystrokes would be in the realm of a few thousand. Proteins could fold more quickly if they retain native-like intermediates along the way. Also, remember that much of conformational space is already restricted by allowed phi/psi angles (remember the empty areas in the Ramachandran plots).

We will explore the classic study of the folding of RNase done by Anfinsen, for which he won the Nobel Prize. RNase is a small protein with 4 disulfide bonds and can refold from denaturing and reducing conditions in vitro. If he had chosen a larger protein, he might not have met with success as they are more prone to aggregation. We'll discuss mostly small proteins in this section.

Before we start, let's anticipate what might happen to fully reduced and denatured RNase with 8 free Cys side chains. If the denatured protein is suddenly placed in a refolding solution without denaturant and in an oxidizing environment (such as oxidized glutathione), the reduced cysteine side chains could start forming disulfide pairs, but only one combination of such pairs would be native. What would be the probability that a purely random process of forming disulfides would produce 4 correct ones and a fully native state?

To think about that, try this thought experiment. You have 4 pairs of socks, with each pair having a different color as shown in Figure 4.8.4. You threw them into a drawer unpaired after washing. Now without looking, take one sock out and then a second and tie them together to form a pair. Continue doing this without looking until all four are paired. What is the probability that all the socks will be correctly matched when you are done? Once you have the answer, you're ready to understand Anfinsen's classic experiment. To think about that, try this thought experiment.

Exercise 4.8.1 Socks and Disulfides

You have 4 pairs of socks, with each pair having a different color as shown in Figure 4.8.4. You threw them into a drawer unpaired after washing. Now without looking, take one sock out and then a second and tie them together to form a pair. Continue doing this without looking until all four are paired. What is the probability that all the socks will be correctly matched when you are done? Once you have the answer, you're ready to understand Anfinsen's classic experiment.

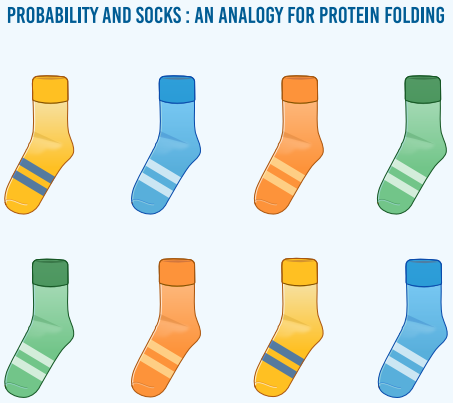


Figure 4.8.4: Socks and protein folding. (CC BY-NC; Ümit Kaya via LibreTexts)

Answer

We'll calculate the probability that you get 4 perfect matches. You pick one sock initially. You have a 1/7 chance to get the match from the remaining 7 unpaired socks. You have 6 socks left. Pick one again. From the remaining 5 you have a 1/5 chance of getting a match. Now pick another. From the remaining 3 you have a 1/3 chance of getting a match. Now pick another. You have a 1/1 changes of getting the match. Each is independent so the total probability of getting 4 matched pairs of socks is given by the products of the probability at each step, or

(1/7) x (1/5) x (1/3) x (1/1) = 1/105 or about 1%

The classic experiment of Anfinsen showed that, at least for some proteins, all the necessary and sufficient information required to direct the folding of a protein into the native state is present in the primary sequence of a protein. Anfinsen studied the in vitro folding of a single chain protein, RNase, which has four intrachain disulfide bonds, as shown in the interactive iCn3D model in Figure 4.8.5.

Figure 4.8.5: RNase with four intrachain disulfide bonds (yellow sticks) (1KF5). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...Risnvb1SfNrww6

We have previously discussed how chemical agents (such as beta-mercaptoethanol, a disulfide reducing agent) can covalently interact with specific protein functional groups. Some can bind through complementary intermolecular forces to the active site or other cavities on the surface. Other reagents, like urea, acting through generalized solvent changes or nonspecific interactions with the protein, can alter protein folding. Anfinsen used two different reagents, 8 M urea and beta-mercaptoethanol (βME), in combination to unfold, or denature, RNase to the nonnative or denatured state. He then removed the βME using dialysis, allowing the disulfides to reform. Next, he removed the denaturing reagent, urea. To monitor if the protein was correctly refolded or

renatured, he tested the activity of the protein compared to the native protein. He found that the "refolded" protein retained only 1% of its initial activity (compare this to the matching socks activity). If, however, he added catalytic amounts of βME, the protein soon retained 100% of its initial activity. For his work, he was awarded the Nobel Prize in Chemistry in 1972. A general outline of his experiment is shown in Figure 4.8.6.

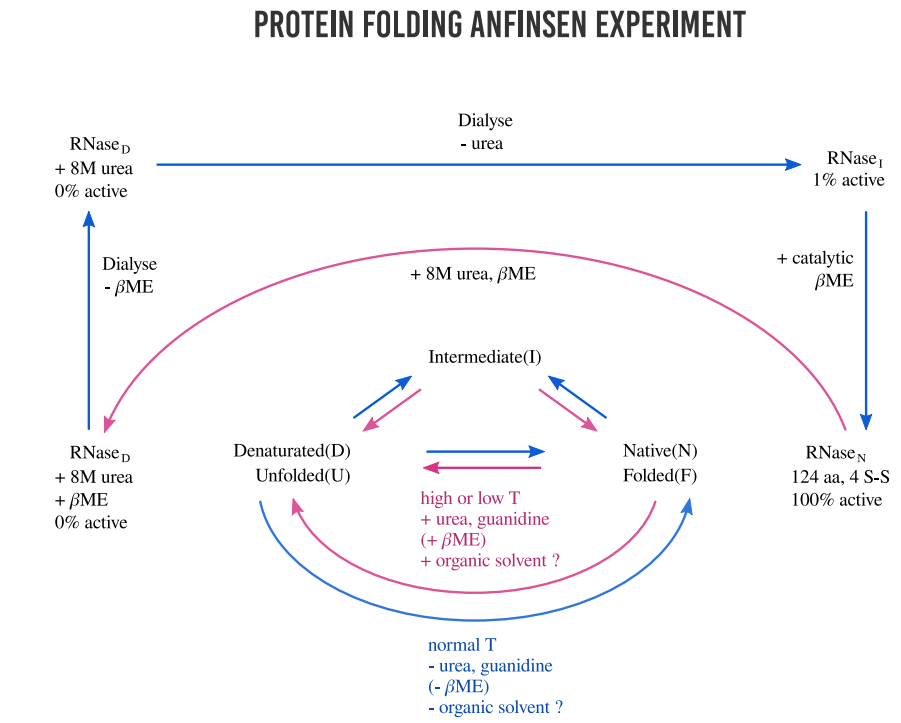


Figure 4.8.6: Anfinsen Experiment: Folding of RNase. (CC BY-NC; Ümit Kaya via LibreTexts)

Figure 4.8.7 presents a chemical mechanism to show how catalytic (non-stoichiometric) amounts of beta-mercaptoethanol can lead to full recovery of the most stable set of a protein with two disulfide bonds (right hand side).

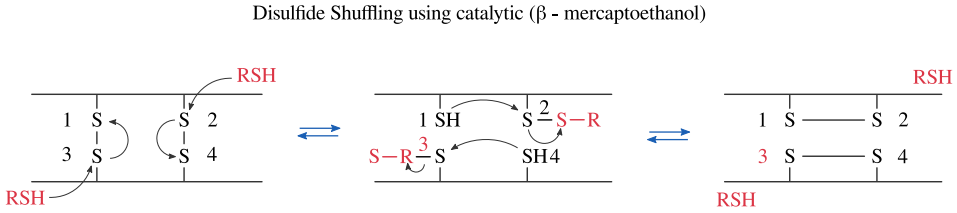


Figure 4.8.7: Catalytic Shuffling of Disulfides With Beta-mercaptoethanol. (CC BY-NC; Ümit Kaya via LibreTexts)

Scientists have investigated the folding of proteins both in vitro and in vivo. In vitro experiments involve denaturing the protein with urea, guanidine hydrochloride, or heat, then refolding the protein by removing the perturbant (denaturing agent), using spectral techniques to follow the process. In vivo experiments involve the study of intracellular proteins that assist folding. An understanding of protein folding can not be separated from an understanding of protein stability, and an understanding of the nature of the native and denatured state as illustrated in the protein folding landscapes shown in Figure 4.8.1 and Figure 4.8.2.

In studying protein folding and stability/structure of the native and denatured states, both equilibrium (thermodynamic) and timed (kinetic) measurements are made. Folding occurs in the ms to second range, which limits the ability to study the presence of intermediates in the process. Some clever methods have been developed to study intermediates in protein folding by trapping specific intermediate structures and investigating their structure and stability in a "leisurely" fashion. Alternatively, intermediates can be studied as they occur using **stop-flow kinetics**. In this technique, a protein under denaturing conditions is rapidly mixed

with a solution containing no denaturant or protein by injecting both solutions into a mixer/cuvette using syringes. The denaturant in the protein solution is now diluted such that renaturation can occur. Spectral measurements can begin at once.

A diagram summarizing these methods is shown in Figure 4.8.8. Study it in conjunction with the text which follows.

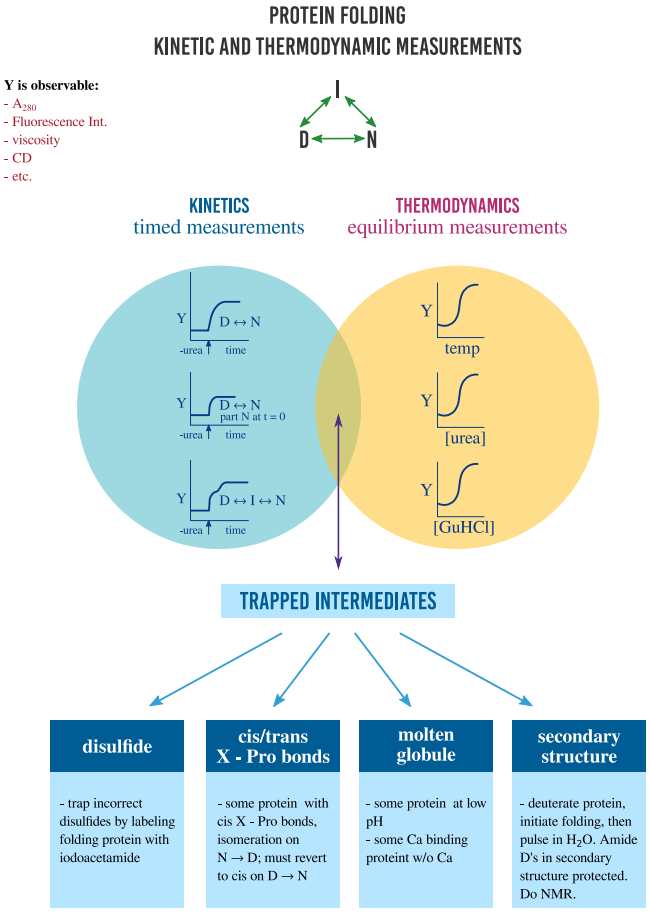


Figure 4.8.8: Kinetic and thermodynamic measurements of proteins stability and folding. (CC BY-NC; Ümit Kaya via LibreTexts)

Folding appears to proceed not by an obligatory pathway but a probabilistic or stochastic search of possible conformations. Evolution has surely selected for sequences that can make it to the folded state. Localized secondary structure motifs (like a short alpha helix and beta turns) can form quickly (about 1 ms). The folding of small proteins occurs, depending on their structure, over a wide time frame (ms to minutes). Most likely, a small number of amino acids coalesce into a core, which serves as a nucleus for folding into structures that are similar to the native state. Finally packing interactions collapse the structure into the native state.

In general, the more complex the fold of the backbone, the longer it takes the protein to fold. If complexity requires more interactions among distal regions of the polypeptide change, then the more complex the fold, the less probable that random interactions would lead to quick protein folding. The mechanisms of folding for larger proteins (greater than 100 amino acids) appear to proceed through intermediates, suggesting that different domains of the protein can fold independently.

4.8.2: Protein Folding In Vitro

Early studies of protein folding involved small proteins which could be denatured and refolded in a reversible fashion. A two state model, D ↔ N, was assumed. The denaturants were heat, urea, or guanidine HCl. Since the denatured states are less compact than the native state, the viscosity of the solution can be used as a measure of denaturation/renaturation. Likewise, the amino acid side chains in the differing states, would be in different environments. The aromatic amino acid Trp, Phe, and Tyr absorb UV light. After excitation, the electrons decay to the ground state through several processes. Some vibrational relaxation occurs, bringing the electrons to lower vibrational energy levels. Some of the electrons can then fall to various vibrational levels at lower principle energy states through a radiative process. The photons emitted are lower in energy and hence longer in wavelength. The emitted

light is termed fluorescence. The wavelength of maximum fluorescent intensity and the lifetime of the fluorescence decay is very sensitive to the environment of the amino acids. Hence fluorescence can also be used to measure changes in protein conformation. Other spectral techniques like CD spectroscopy as well as simple absorbance measurements, are used. For small, single domain proteins (such as RNase) undergoing reversible denaturation, graphs showing the extent of denaturation using each technique above, as shown in Figure 4.8.9, are superimposable, giving strong validity to the two state model.

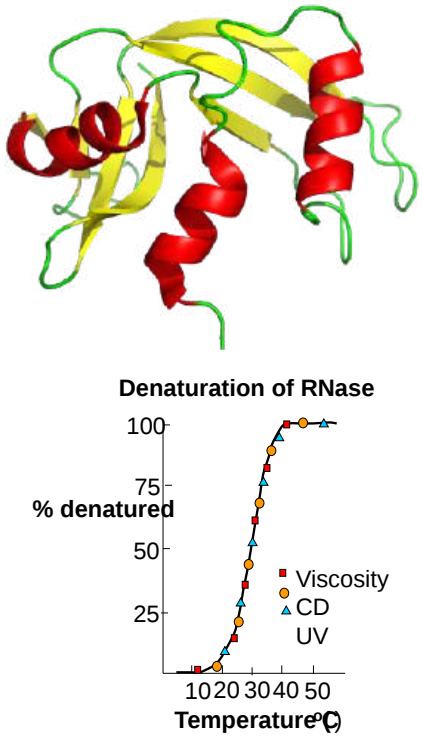


Figure 4.8.9: Denaturation of RNase - % Denaturation vs Temperature (after after Ginsburg and CArroll, Biochemistry 4, pg 2159 (1965)

Proteins that fold without easily discernable, long-lived intermediates and follow a simple two-state model, D ↔ N are said to undergo **cooperative folding**. This simple model needed to be expanded as more proteins were studied. Some intermediates in the process were detected.

Some proteins show two steps, one slow one, and one quick one, in refolding studies, suggesting an intermediate. The longer a protein is kept in the denatured state, the more likely it is to display an intermediate. One accepted explanation for this observation is that during an extended time in the D state, some X-Pro bonds might isomerize from trans to the cis state, to form an intermediate. Alternatively, as in the case of RNase, which has a cis X-Pro bond in the native state, denaturation causes an isomerization to the trans-state. In the case of RNase, to refold, the accumulating intermediate I must reisomerize in a slow step to the cis state, followed by a quick return to the N state.

Some proteins, which contain multiple disulfide bonds that must reform correctly after reductive denaturation could refold into intermediates with the wrong S-S partner. Such intermediates can be trapped by stopping further S-S formation during refolding with the addition of iodoacetamide as shown in Figure 4.8.10.

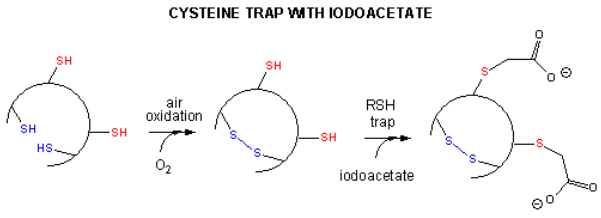


Figure 4.8.10: Cysteine trapping with iodoacetate.

By adding iodoacetamide at varying times along the folding pathway, and separating starting (unfolded), trapped intermediates on the pathway and the final native state by HPLC, the entire folding pathway can in principle be determined. This has been done for

many proteins including bovine pancratic trypsin inhibitor (BPTI), small protein with 58 amino acids, a molecular weight of 6512, and three sets of disulfide bonds (C5-C55, C14-C38 and C30-C51). The structure of BPTI with native disulfides is shown in Figure 4.8.11.

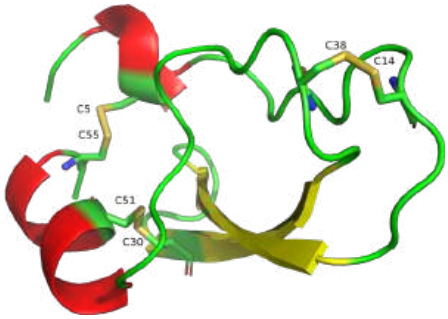


Figure 4.8.11: BPTI (2hex) showing 3 pairs of disulfide bonds

Given its characteristics, BPTI has been used as a model protein to study protein folding. These studies have shown that no non-native disulfides form as intermediates in the pathway. Two major intermediates form quickly, each with two of the three native disulfide bonds. These intermediates are named N',with disulfide pairs 14-38 and 30-51, and N*, with pairs 5-55 and 14-38. These two intermediates then slowly form a common intermediate N^{SH}_{SH}, with disulfide pairs 5-15 and 30-51, which then converts very quickly to the native N state with three correct pairs (5-55,14-38 and 30-51). This folding pathway is shown in part a and b of Figure 4.8.12.

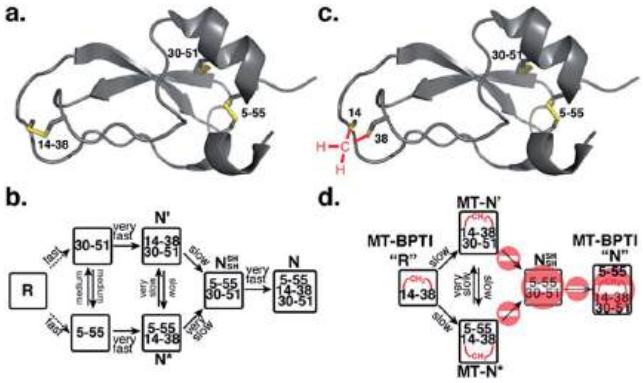


Figure 4.8.12: Folding Pathways for BPTI. Reem Mousa et al. *Chem. Sci.*, 2018, **9**, 4814-4820. DOI: 10.1039/C8SC01110A (Edge Article) *Chem. Sci.*, 2018, **9**, 4814-4820dkjfkdjkf. Creative Commons Attribution-Non Commercial 3.0 Unported Licence

In going from N' or N* to N^{SH}_{SH}, the **14-38** disulfide, which is very solvent-exposed (see Figure 4.8.11), must break before proceeding to the N (native state). Mousa et al replaced that disulfide with a stable methylene thioacetal bridge (MT, alternative name methylenedithioether, S-CH_2-S link) with the idea that once this stable (i.e irreversible) 14-S-CH_2-S38 bond formed in N' and N*, it would not break again. The formation of the methylenedithioether bond and its properties compared to disulfide bonds are shown in Figure 4.8.13.

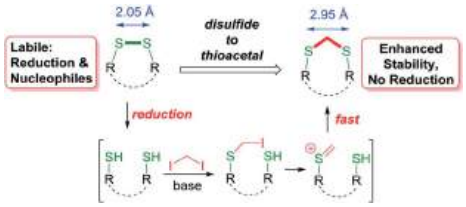


Figure 4.8.13: Formation and properties of the methylenedithioether bond. C. M. B. K. Kourra and N. Cramer, ***Chem. Sci.***, 2016,**7**, 7007-7012. **https://doi.org/10.1039/C6SC02285E.** C. M. B. K. Kourra and N. Cramer. Creative Commons. **Attribution-NonCommercial 3.0 Unported (CC BY-NC 3.0)**

Hence the final last intermediate, N^{SH}_{SH}, would not form, leaving just N' and N*. This is illustrated in Part c an d in Figure 4.8.12.

In actuality, MT-BPTI **does** fold to the native form, whose crystal structure is virtually identical to native BPTI. Figure 4.8.12 shows "2D" folding plots that show HPLC retention time of reactants, intermediates on one axis and folding time on the other for

wild-type BPTI (part a) and MT-BPTI (part b). Hence the folding scheme is more complicated than the one shown in Figure 4.8.14

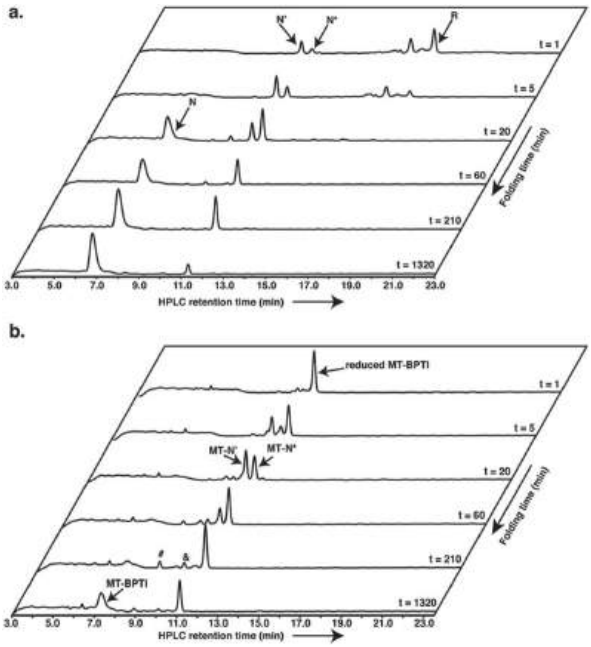


Figure 4.8.14: BPTI and MT-BPTI Folding Pathway In Vitro -

Some proteins form partially folded but stable intermediates when folded under partially denaturing conditions. A good example is lactalbumin, which under mildly acidic conditions (pH 4), low levels of guanidine HCl, or neutral pH and low ionic strength in the absence of calcium (which normally binds to the protein), forms a stable, isolatable intermediate (I) called the **molten globule (MG)**. Figure 4.8.15 shows the folded state with a bound calcium ion and circular dichroism spectra of the protein in various states.

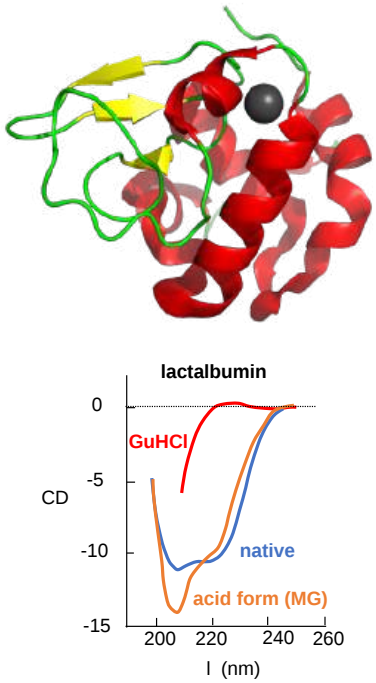


Figure 4.8.15: CD spectrum of Native, denatured (GuHCl) and molten globule (MG) forms of Lactalbumin (after Kuwajima. Biochemistry 24, pg 874,1985)

Data show that the MG is about 50% larger in volume than the N state. This compares to the denatured state, which can be 300% larger than the native state. Hence, it is more like the native state as studied by hydrodynamic techniques, but with more solvent accessibility of hydrophobic side chains. The MG has a similar CD spectrum as the native state, but the aromatic side chains display the same UV absorption and fluorescent characteristics as the protein in 6 M guanidine HCl, suggesting that the final tertiary state has not yet completely formed. The secondary structure in the MG may not be the same as in the native state

NMR can also be used to detect folding intermediates. Using this technique, proteins are unfolded in D_2O, which will cause the exchange of all Cs with ionizable protons, including, the amide Hs. An amine is a weak base (pK_B around 3.5) so its conjugate acid, the protonated amine, has a pK_a of around 9.5. An amide or peptide bond would be a weaker base than an amine since it's lone pair is less available (due to delocalization through resonance) for sharing with a proton. The pKa for the conjugate acid of the amide (in which the amide N is protonated and has a plus charge) is much lower, around -0.5, than the pKa for the conjugate acid of an amine. At 2 pH units greater than its pKa, the charged amide N is close to 100% deprotonated The pka of the protonated group is important since the rate of H exchange is related to the pKa, holding other variables constant. The pka of an unprotonated amine ($RNH_2 \rightarrow RNH^-$ is very high (30s) and hence deprotonation of the RNH_2 amine to form RNH- is not likely under normal conditions. Figure 4.8.16 shows NMR D/H exchange/protection experiments for the formation of an alpha helix in a protein folding experiment.

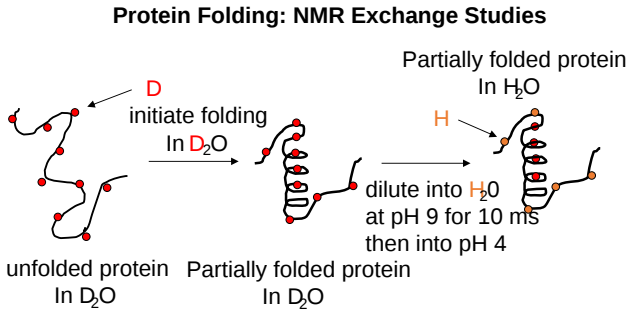
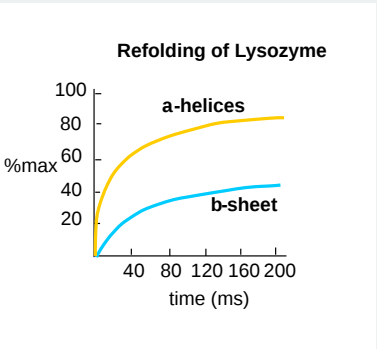


Figure 4.8.16: NMR D/H exchange/protection in protein folding

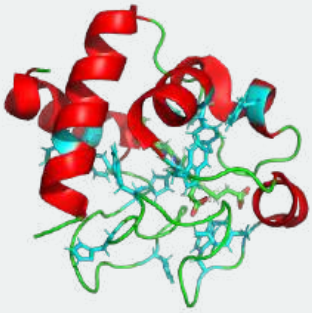
Refolding is initiated by diluting the protein into a solution without the denaturatant, but still in D_2O. As the protein folds and becomes more compact, the buried atoms are now sequestered from the solvent, and no longer readily exchange Ds. Then the protein is placed in H_2O at pH 9.0 for 10 ms, after which the pH is changed to pH 4.0. D → H exchange is promoted at high pH, and quenched for the amide Ds and Hs at low pH. Amide Hs that continue to exchange must be accessible to water. Those that aren't are usually buried in secondary structure.

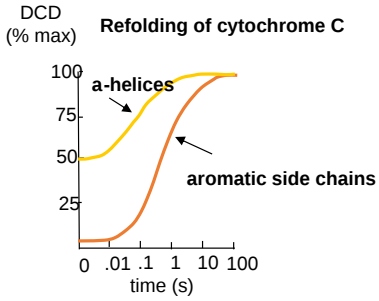
How would you interpret these structure and data?

Hen Egg white lysozyme: Radford et al, Nature 358, pg 303 (1992) 2YVB



Cytochrome C: Elove et al, Biochemistry, 31, pg 6879 (1992). Cyan aromatic amino acids; Heme not shown. 5TY3





When the same techniques are applied to large, multidomain or oligomeric proteins, only a few percent refold in vitro. Incorrect intermolecular interactions and heterogeneous aggregation appear to be the main problems, which prevent correct protein folding in vitro

Here is a movie of a 6 us molecular dynamics simulation of the small protein villin.

The movie starts with the final crystal structure of villin and show how it folds into its final structure.

With permission from the Beckman Institute for Advanced Science and Technology National Institutes of Health //National Science Foundation, Physics, Computer Science, and Biophysics at University of Illinois at Urbana-Champaign

4.8.3: The Denatured State

Although the structure of native and native-like states can be determined using x-ray crystallography and in solution using NMR, little detailed information exists on the actual structure of denatured and intermediate states. Intermediate states are difficult to trap in a way that allows detailed structural analysis. In contrast to the "native" state which consists of an ensemble of closely related states, intermediates and denatured states would consist of an ensemble of many different states, making structural analysis more difficult. Religa and others from Fersht's lab have engineered a mutant of the engrailed homeodomain (En-HD) from Drosophila melanogaster that allows such structural analyses to be performed. The mutation, Leu16Ala (L16A), destabilizes the protein such that it can be denatured simply by changing ionic strength. It is stable at high ionic strength and folds quickly under those conditions. However, at physiological ionic strength, it is "denatured" but contains significant alpha-helical structure but has nonnative contacts. It behaves like an early folding intermediate in that if placed in solutions of higher ionic strength it rearranges to form the native state. If placed in lower ionic strength, it progressively "unfolds" to yet other states. Given the ambiguities in how to define denatured and early folding intermediates states, Ferscht's group suggests an "explicit" definition of the denatured state. They define the unfolded state (U) as the "maximally unfolded state of a protein, in which the backbone NH groups have little protection against 1H/2H exchange". They define the denatured state, D, as the "lowest energy non-native state under a defined set of conditions". In this scenario, the denatured state could also be a folding intermediate if placed in conditions that promote folding. Previous work from the group showed that the denatured state of En-HD has three helices protected from 2H exchange and was 1kcal/mol l(4.1 kJ/mol) lower in energy than the unfolded state.

In section 4.6, we discussed the added complexities to the notion of a simple 2-state D ↔ N model for protein folding. These include Silent Single nucleotide polymorphisms, Metamorphic Proteins and Intrinsically Disordered Proteins (IDPs).

4.8.4: Protein Folding In Vivo

There are many differences between how a protein might fold or unfold in a cell compared to a test tube.

- The total concentration of all the proteins and nucleic acids in cells is estimated to be about 350 g/L, or 350 mg/ml. Most measurements in the lab are conducted in the range of 0.1 to 10 mg/ml
- Proteins are synthesized in cells from an N to C terminal direction. Hence the nascent protein, as it emerges from its site of synthesis (the ribosome), might fold into intermediate structures since not all of the protein sequence is yet available for direct folding.
- Proteins are synthesized in the cytoplasm, but they have to find their final place in the cell. Some end up in membranes, some must translocate across one or even two different membranes to end up in specific organelles like the Golgi, mitochondria, chloroplasts (in plant cells), nuclei, lysosomes, peroxisomes, etc. Do they translocate in their native state?

Additional evidence suggests that protein folding/translocation requires assistance (i.e. catalysis) in the cell.

- Mutant cells defective in certain proteins can lead to the accumulation in the cells of misfolded and aggregated proteins.
- eukaryotic genes (taken from higher cells, which contain nuclei and internal organelles), when transferred into prokaryotes (bacteria, like E. Coli), can be expressed to form protein, but they often misfold and aggregate in the bacterial cells and form structures called inclusion bodies.

Hence recombinant proteins expressed in vivo have the same problems in folding as larger proteins in vitro. In both cases, conditions favor the accumulation of nonnative proteins with exposed hydrophobic groups leading to aggregation. Aggregation also occurs in vivo when a protein is over-expressed or expressed at a higher temperature than normal. Why? Mutant cells have been selected that actually suppress inclusion bodies in vivo. This effect was mediated by a class of proteins, which are expressed by the bacteria and other cells when their temperature is raised. The function of these proteins, called **heat shock proteins (Hsp)**, was unknown until it was realized that they facilitate correct protein folding, in part, by binding to denatured proteins in the cells before they aggregate into inclusion bodies. Further studies discovered a large number of proteins that seem to facilitate protein folding and prevent aggregation in vivo. These proteins are now called molecular chaperones. Many are still named **Hsp#**, where # is the approximate molecular weight of the protein in kD on a PAGE gel.

In a broader sense, a protein homeostatic environment exists within cells to maintain the proteome. This system includes chaperones proteins which facilitate folding and refolding, the proteasome, which cleaves undesired proteins marked by ubiquination, and the lysosome, which facilitates the process of autophagy of structures such as damaged mitochondria. Here we will concentrate on chaperones, which also are involved in protein transport in the cell and preventing aggregation. Some have used novel names to describe the various activities of chaperones. These include holdases/translocases, unfoldases/foldases and disaggregases, which are used to process the various potential intermediates and end products that occur. A simplified 2D free energy folding diagram showing the involvement of chaperones is shown in Figure 4.8.17.

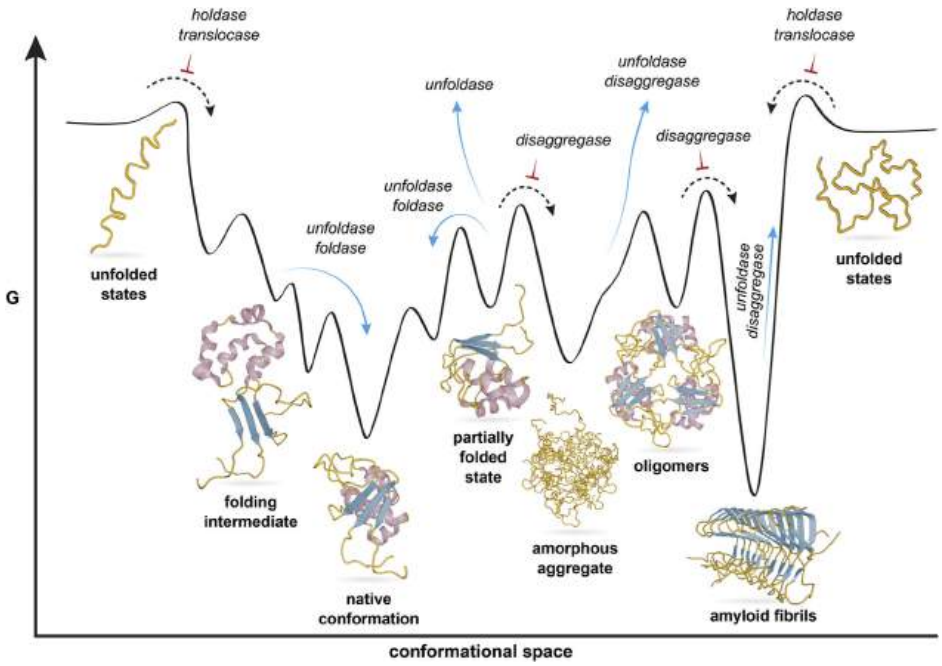


Figure 4.8.17: 2D Free Energy Folding Diagram with Chaperone Assist. Moseck et al. Front. Mol. Biosci., 14 June 2021 | https://doi.org/10.3389/fmolb.2021.683132 Frontiers in Molecular Biosciences 8, pg 514, 2021 https://www.frontiersin.org/article/...lb.2021.683132. **Creative Commons Attribution License (CC BY).**

The nomenclature used to categorize chaperones is confusing. In general, chaperones, found in all cells, interact with unfolded or misfolded proteins and essentially catalyze their folding. They can do so by removing them from environments that would inhibit folding. Some can prevent aggregate formation or remove them. Some interact with proteins exiting the ribosome as they are being synthesized while some escort damaged proteins to sites of proteolysis.

Many chaperons are used to maintain proteostasis. They can be classified as to molecular size or groups based on overall structure and mechanism. We'll use the second approach. Here are some of the key player

4.8.4.1: Cage Chaperonins: Oliogomeric High Molecular Nanoparticles

Chaperonins, whose subunits can still be called chaperones, assemble into oligomeric, high molecular weight nanoparticle cages. These complexes are involved in the folding of larger proteins. The complexes consist of two stacked rings. **In**side is an inner cavity **in** which larger proteins can fold in isolation (an easy way to remember the name chaperon**in**). A somewhat silly analogy would be a person dressing alone in a closet. These complexes are ATPase as the cleavage of ATP drives protein folding.

There are two main classes of chaperonins, Class I and Class II

Group I: These are found in bacteria, in some archaea and in mitochondria, which derived in evolution through an endosymbiotic relationship with bacteria. Proteins in this group include **Hsp60**, which forms a homo 14-mer, and the co-chaperone **Hsp10**, which forms a homo 7-mer. In E. Coli, hsp60 is also called **GroEL**, while the small Hsp10 is called **GroES**. GroEL and GroES together form the GroEL/GroES Complex, which is shown in the interactive iCn3D model in Figure 4.8.18 (note: it loads slowly given its large size).

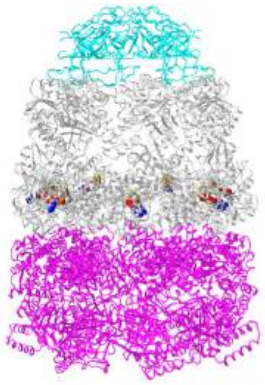


Figure 4.8.18: GroEL/GroES Complex (1pcq). (Copyright; author via source).
Click the image for a popup (which loads slowly) or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?JNESMy2hS9ZBTST66

The 14 GroEL subunits form separate 7-mer (magenta) and 7-mer (gray) hollow rings, which stack on each other. The gray ring monomers each have an ADP/AlF_3^- (spacefill) bound. Seven GroES monomers, shown in cyan, form a "lid" over the end of the complex. This hetero 21-mer displays a C7 global symmetry. Lid binding requires ATP.

Figure 4.8.19 shows a mechanism for the GroEL/GroES protein folding cycle (shown in a linear arrangement).

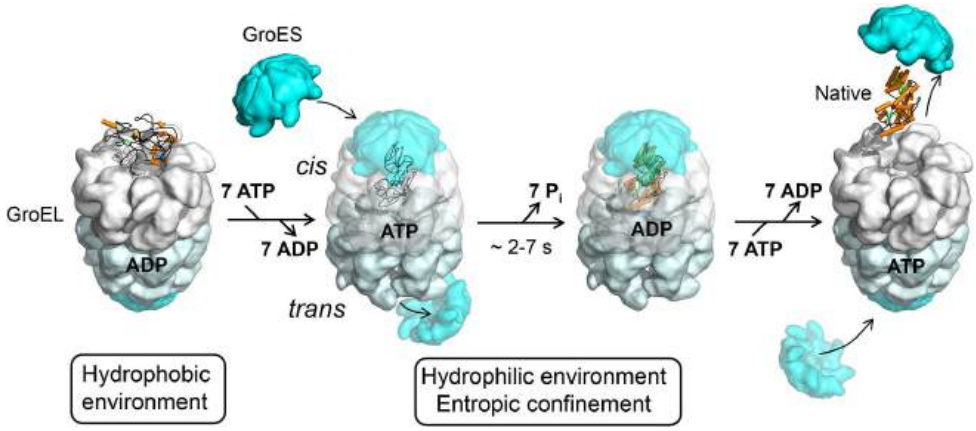


Figure 4.8.19: GroEL/GroES protein folding cycle in linear form. F. Ulrich Hartl Mol Biol Cell. 2017 Nov 1; 28(22): 2919–2923. doi: 10.1091/mbc.E17-07-0480. Attribution–Noncommercial–Share Alike 3.0 Unported Creative Commons License (http://creativecommons.org/licenses/by-nc-sa/3.0).

In the left-most structure, a substrate protein (colored line drawing) binds in the cavity formed by the GroEL ring that does not have the GroES 7-mer cap (cyan). Rearrangement of the substrate protein and interactions with the interior wall trigger binding of ATP, binding of the top GroES cap on the cis end (same end as bound protein substrate), and release of the previously bound ADP, the distal trans GroES cap and a folded protein (not shown) from the trans end. Transient interactions and substrate protein conformation changes lead to the folding of the colored protein substrate within the time frame needed for the cleavage of the 7 bound ATPs. Binding of ATP and a new GroES cap at the bottom trigger release of the folded protein at the top along with the top GroES cap. The cycle then repeats with the binding of a new target protein substrate, ATP and a GroES cap leading to the dissociation of the ADP and the other GroES cap.

The sequestration of the bound protein substrate clearly prevents aggregation of the folding protein. The binding of any polymer into a confined small volume must be entropically disfavored since the polymer's conformational flexibility is reduced. Somehow an entropic activation energy barrier is reduced inside of the cage, maybe in a fashion similar to the restraining effects of disulfides on protein folding, which allows less conformational space to be explored on folding. GroEL has also been shown to bind in its hydrophobic cavity a fluorescent CdS semiconductor nanoparticle which can be released on addition and cleavage of ATP

Group 2: These are found in Archaea and in eukaryotes. The eukaryotic chaperonin in this group is **CCT** (cytosolic chaperonin containing TCP1) also known as **TRic** (tailless complex polypeptide 1 ring complex or TCP1-ring complex). These also contain

two stacked rings but this time they are 8-mers of molecular weight 50-60K. This is similar to the Hsp60 monomers in the GroEL/GroES complex, but the monomer in Group 2 are not named with the beginning letters Hp. In addition, the complex does NOT have a cap like the GroES co-chaperone 7-mer in the GroEL/GroES complex. Rather, they have a built-in cap that closes on folding. The **CCT/TRic** chaperonin complex interacts with other "co-chaperones" including a **prefoldin 6-mer** complex, which inhibits aggregation and which "delivers" the protein substrate to **CCT/TRic.** It also binds to phosducin-like proteins. Well know substrate proteins for CCT/TRic include actin and tubulin (cytoskeletal proteins) and proteins involved in cell cycle control, but many other proteins (up to 10% of cytosolic proteins) might interact with it.

The genes encoding the 8 monomers in CCT/TRic arose from gene duplication and subsequent mutation so they are different from each other (technically they are called paralogs). The subunits are named α, β, ξ, δ, ε, ζ, η and θ. Figure 4.8.20 shows generalized structures of **CCT/TRic.**

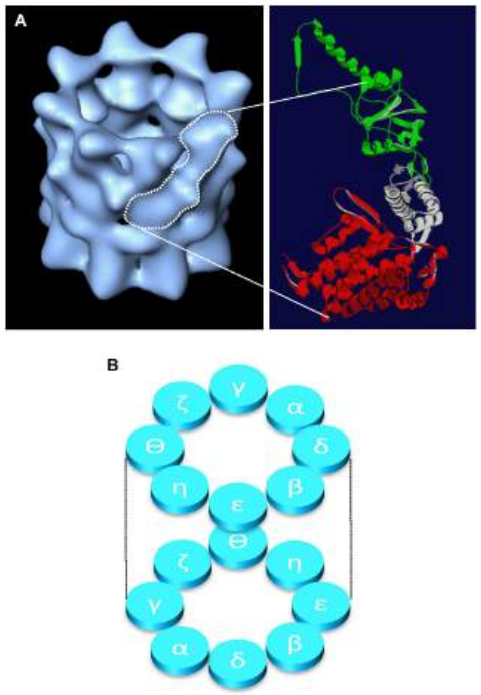


Figure 4.8.20: Low resolution models of the **CCT/TRic** chaperonin complex. Julie Grantham. Front. Genet., 19 March 2020 | https://doi.org/10.3389/fgene.2020.00172. **Creative Commons Attribution License (CC BY)**.

Panel A shows a cryoEM of the complex with one subunit outlined in white and an adjacent PDB structure of the alpha subunit (1A6D).The red domain binds ATP. The relative arrangements of the monomers is shown in B.

The Human TRiC/CCT complex (7LUP) with a protein substrate, reovirus outer capsid protein sigma-3, bound in the internal cavity is shown in the interactive iCn3D model in Figure 4.8.21 (note: it loads slowly given its large size).

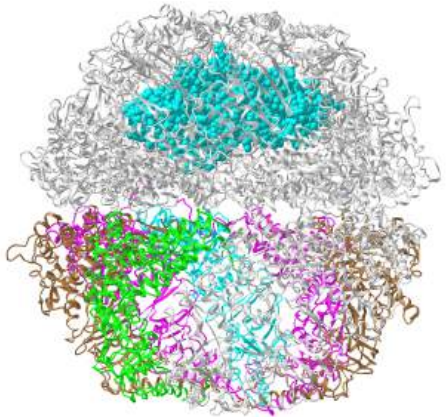


Figure 4.8.21: Human TRiC/CCT complex with reovirus outer capsid protein sigma-3 (7LUP) (Copyright; author via source).

Click the image for a popup (which loads slowly) or use this external link: https://structure.ncbi.nlm.nih.gov/i...Dxr3n6oy6zLsZ7

The cyan spacefill protein substrate is the encapsulated reovirus outer capsid protein sigma-3. The top ring 8-mers are shown in gray ribbon while the bottom ones are shown in different colors. It appears that each ring binds 4 ATPs for a total of 8 when the "attached lids" are closed.

4.8.4.2: Non-cage chaperones

A variety of chaperones that are not part of the nanoparticle chaperonin complex are also involved in protein folding in vivo. Some classify chaperones on the basis of molecular weight into 5 classes, Hsp60 (which we discussed above as monomers in the chaperonin cage complexes), **Hsp70, Hsp90, Hsp104, and the small Hsps.** We'll focus mostly in Hsp70 and its "co-chaperone", Hsp 40, and their bacterial analogs, DnaK and DnaJ, respectively.

Hsp70/Hsp 40: Hsp70 binds to hydrophobic regions of proteins which are more prevalent in unfolded and partially folded proteins and helps refold them through repeated cycles of binding and release, which is dependent on ATP cleavage. It also helps unfolded proteins through membranes and helps form/dissociate protein complexes. It's found in the cytoplasm and in a variety of organelles including chloroplasts, mitochondria, nuclei and the ER. There is a family of Hsp70 proteins (1, 1A, 1B, 2, 3, 4, 4L, 5–10, 12–18). Hsp70 proteins are made up of two regions. The **nucleotide binding domain (NBD**), which has intrinsic low **ATPase** activity, is located in the N-terminal region. The polypeptide **substrate binding domain (SBD)** is located in the carboxyl terminal end. It transiently recognizes short peptides in the target protein. A flexible stretch of conserved amino acids connects the two domains. The substrate binding domain has a beta sheet where peptide substrates bind and an alpha-helical region that acts like a lid.

From an evolutionary sense, they are one of the most conserved proteins. This shows the critical importance of protein folding to life. Some members of the Hsp70 family include **DnaK** (a bacterial Hsp70 that has been widely studied along with its Hsp40 **co-chaperone partner, DnaJ**) and BiP (HSPA5) that assists protein folding in the ER. and alpha crystalline in eukaryotes. Alpha crystalline comprises 30% of the lens proteins in the eye, where it functions, in part, to prevent nonspecific, irreversible aggregates. Some Hsp70 proteins are constitutively expressed, while others are induced by increased temperatures or other stress perturbants that affect protein structure, including radiation, inflammation and exposure to heavy metals. They are key to the maintenance of proteostasis.

Hsp70 proteins:

- bind to growing polypeptide chains as they are synthesized on ribosomes.
- express activity as monomers.
- have ATPase activity - i.e. they cleave the phosphoanhydride ATP (which can drive reactions).
- bind short, extended peptides, which stimulates the ATPase activity
- release bound peptides after ATP cleavage

Hsp70s are highly flexible so it is difficult to get detailed structural information comparing bound (to target peptides) and free states. Hsp70 appears to clamp onto hydrophobic regions of target proteins, making transient interactions characterized by multiple binding and release steps, until protein folding is complete. Hence the affinity for the Hsp70 proteins must be low enough (i.e. high dissociation constant K_D where the K_D is the inverse of the equilibrium binding constant as we will see in Chapter 5) for the target protein for many binding and release steps. The affinity of the SBD for the target protein is regulated by the nucleotide binding domain (NBD). When ADP is bound, the Hsp70 is in a **closed state** and has a high affinity for substrate, so the release rate for bound target protein is slow. When ATP is bound, Hsp 70 is in the **open state** in which the rate of association increases 100 fold and the rate of dissociation increases 1000 fold, so the affinity is lowered significantly. This enables multiple cycles of release, folding, and rebinding to occur.

As mentioned above, Hsp70s have intrinsic but low ATPase activity. When ATP is cleaved to ADP, Hsp70 returns to the closed, high-affinity state. Hence ATP hydrolysis affects both the binding of protein substrate) and kinetics of folding. This activity increases with the binding of protein substrates as well as with the binding of co-chaperones containing what is called a J domain (as is found in the bacterial co-chaperone DnaJ). Hence considerable conformational changing and signaling occur between the NBD and SBDs in a process we will call allosterism in Chapter xx.

The allosteric Hsp70 catalytic cycle is shown in Figure 4.8.22

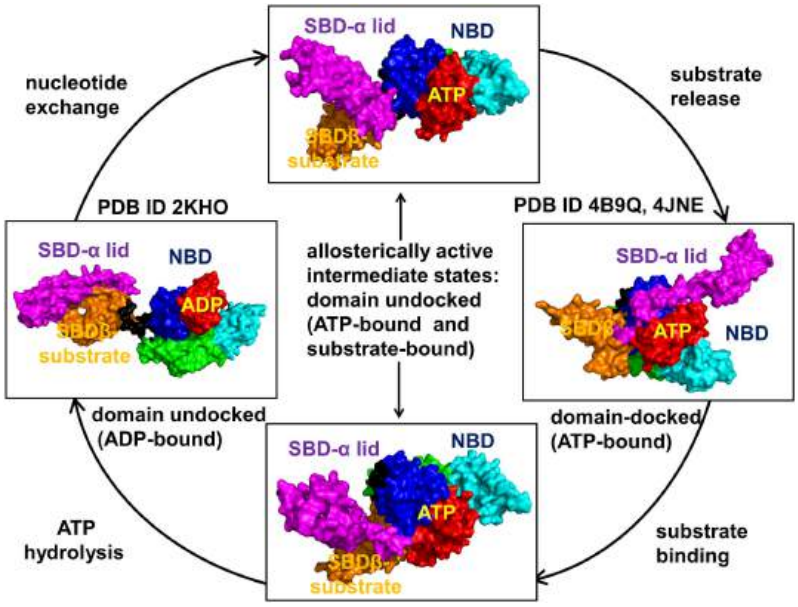


Figure 4.8.22: Hsp70 Catalytic Cycle for Protein Folding. Stetz G, Verkhivker GM (2015) Dancing through Life: Molecular Dynamics Simulations and Network-Centric Modeling of Allosteric Mechanisms in Hsp70 and Hsp110 Chaperone Proteins. PLoS ONE 10(11): e0143752. https://doi.org/10.1371/journal.pone.0143752. Creative Commons Attribution License,

Let's explore the catalytic cycle starting the left structure. (note that NBD subdomains are colored as follows: IA (in blue), IB (in red), IIA (in green), IIB (in cyan))

- Left structure (9 o'clock): ADP is bound so Hsp70 is in the closed (high affinity for protein substrate)state. The SBD consists of two parts or subdomains, the α-lid (magenta) section and the subdomain comprised of β-sheets (orange), where the actual protein substrate binds (not shown here). The α-helical lid of the SBD helps to keep the protein substrate from dissociating. The PDB ID for this structure is 2KHO, but if you explore this structure, no protein substrate is evident. The two domains of Hsp70, the SBD and the NBD, are connected by the flexible interdomain linker (black), are not interacting so they are shown in the **domain undocked** state.
- Top (12 o'clock): This represents an "intermediate" in which ATP is exchanging for bound ADP and the complex is changing to the open form, but without the release of protein substrate. Note that the α-lid is no longer interacting as tightly with the β-sheet subdomain of the SBD.
- Right (3 o'clock): The protein substrate (again not shown) is fully released, the α-lid is fully disengaged from the β-sheet subdomain of the SBD, and Hsp70 is in the open (to substrate binding) conformation. The SBD β-sheet subdomain and the NBD domains are still engaged and are shown in the domain-docked conformation.
- Bottom (6 o'clock): Weak protein substrate binding leads to an "intermediate" in which the α-lid is starting to engage with the β-sheet subdomain of the SBD. Substate binding promotes ATP hydrolysis, which occurs as Hsp70 returns to the full closed conformation (left structure, 9 o'clock) with bound ADP

Figure 4.8.23 shows the incredible conformational change that occurs on movement from the ADP-bound higher affinity closed form of E. Coli Hsp70 (DnaK) (2kho) to the ATP-bound lower affinity, domain-docked open(to substrate binding) form (4jne). Note how the alpha-helical lid in the left-hand side RBD moves to engage with the NBD on the right-hand side.

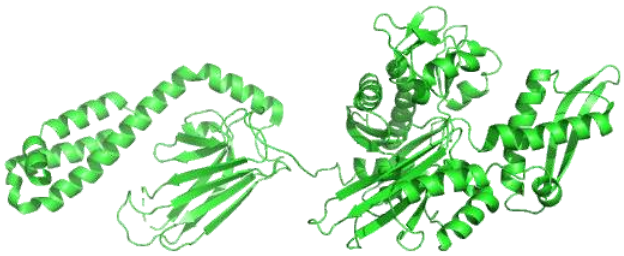


Figure 4.8.23: Animation of conformational changes in the receptor binding domain of Hsp70

Figure 4.8.24how the chaperonins and Hsp70/40 (DnaK/DnaJ) work in concert.

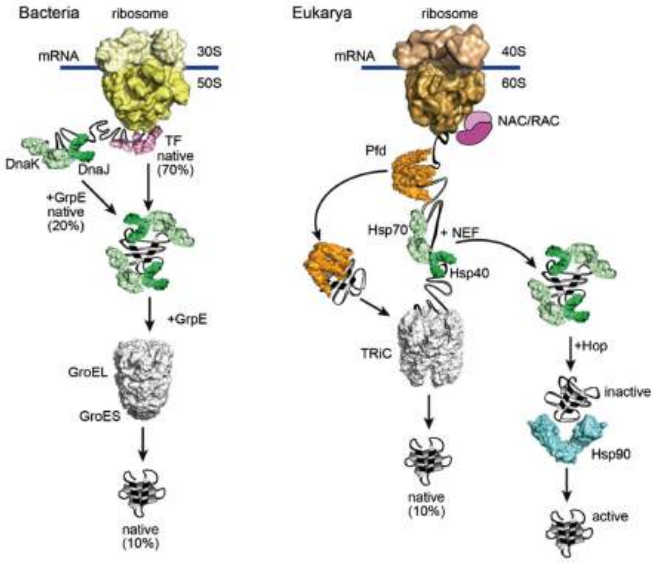


Figure 4.8.24: Comparison of folding pathways in bacteria and euakaryotes F. Ulrich Hartl Mol Biol Cell. 2017 Nov 1; 28(22): 2919–2923. doi: 10.1091/mbc.E17-07-0480. Attribution–Noncommercial–Share Alike 3.0 Unported Creative Commons License (http://creativecommons.org/licenses/by-nc-sa/3.0).

A few other chaperones acting downstream from Hsp70/Hsp40 and DnaK/DnaJ are shown in the image:

- TF (Trigger factor) and NAC/RAC: these are ribosome-binding chaperones
- Prefoldin (Pfd): bind unfolding proteins and shuffles them off to TRic.
- Hop: mediates interactions between Hsp70/Hsp40 and Hsp90
- Hsp90: This chaperone targets protein receptors (including steroid hormone receptors, protein kinases, and proteins with **n**ucleotide-binding site and **l**eucine-rich **r**epeat (NLR) domains as are found in inflammasomes) and has a cycle similar to Hsp70 with some differences. Hsp90 has an N-terminal nucleotide binding domains, middle domains that interact with protein substrates, and C-terminal dimerization domains so the protein exists as a homodimer, When bound to ADP or in the absence of nucleotides, the homodimer forms an **open** V-shape conformation (as shown in the figure above). When ATP binds a huge conformation change occurs that closes the V shaped clamp. The protein has intrinsic ATPase activity, which reverses the conformational change. Many different co-chaperones bond to the HSP90 dimer and modulate the cycle at different points.

Additional Proteins also catalyze protein folding at key steps in the process. Here are two examples.

- Protein Disulfide Isomerase (PDI) - catalyzes the conversion of incorrect to correct disulfides. The active site consists of 2 sets of the the following sequence - Cys-Gly-His-Cys, in which the pKa of the cysteines are much lower (7.3) than normal (8.5). How would this facilitate disulfide isomerization?
- Peptidyl Prolyl-Isomerase - catalyzes X-Pro isomerization, by a mechanism, which probably involves bending the X-Pro peptide bond. How would this facilitate the process?

Many proteins have been found to possess PPI activity. One class is the immunophilins. These are small proteins found in the cytoplasm that bind anti-rejections drugs used to prevent tissue rejection after transplantation. The immunophilin FK506 binding protein (FKBP) binds FK506 while the protein cyclophilin binds that anti-rejection drug cyclosporin. The complex of cyclophilin:cyclosporin or FKBP:FK506 binds to an inhibits calcineurin, an important protein (with phosphatase activity) in immune cells (T cells) required for T cell function. In this case, immunophilin:drug binding to calcineurin inhibits the activity of the T cell, preventing immune attack on the transplanted tissue and rejection. The immunosuppressant drugs (FK506 and cyclosporin) inhibit the PPI activity of their respective immunophilin. The extent to which the PPI activity of cyclophiin is required for its activity is unclear, but it seems to be important for some of its biological effects.

Given the complexity of protein folding and the large number of chaperones involved in in vivo folding, it should not be a surprise that chaperone proteins bind to protein substrates through more than just hydrophobic interactions. Some chaperones (Spy and Trigger Factor) bind to charged regions on the surface of proteins, while the ER proteins calnexin and calreticulin bind to carbohydrates on glycoproteins. GroEL/GroES and TRiC/CCT also interact through electrostatic attractions with protein substrates. Nucleoplasm, a chaperone protein found in the nucleus, binds to histones which are strongly costively charged DNA binding proteins.

4.8.5: Unfolded Protein Response

As the site responsible for the folding of membrane proteins and proteins destined for secretion, as well as the major site for lipid synthesis, the endoplasmic reticulum (ER) must be able to maintain homeostatic conditions to ensure proper protein formation. Plasma cells that synthesize antibodies for secretion as part of the immune activation, show large increases in protein chaperones and ER membrane size

The main pathway controlling ER biology is the **unfolded protein response (UPR)** signaling pathway. If demand for protein synthesis in the ER exceeds capacity, unfolded proteins accumulate. This ER stress condition activates a protein called IRE1, a transmembrane Ser/Thr protein kinase (which phosphorylates proteins). IRE1 activates a transcription factor that controls the transcription of many genes associated with protein folding in the ER. Another protein, ERAD (ER-associated degradation) moves unfolded proteins back into the cytoplasm where they are degraded by the proteasome. Proteins involved in lipid synthesis are also activated as lipids are needed for membranes as the ER increases in size. If the stress can not be mitigated the signaling pathway leads to programmed cell death (apoptosis).

Schuck at al investigated the specific role and importance of UPR in the homeostasis of ER as modeled by the yeast Saccharomyces cervisiae. The UPR signaling pathway was analyzed using light and electron microscopy to visualize and quantify ER growth under various stress conditions. Western blotting procedures were performed to determine chaperone protein concentrations after stress induction and association with ER expansion after the ER was exposed to various treatment conditions. The authors found ER membrane expansion occurred through lipid synthesis since stress induction increased concentrations of proteins responsible for promoting lipid synthesis and expansion failed when the proteins were absent and lipid concentration was low. In addition, these lipid synthesis proteins were activated by the UPR signaling pathway. By separating ER size control and UPR signaling, they found that expansion occurred regardless of chaperone protein concentrations. However, if lipid synthesis genes were not available, raising the ER chaperone level helped alleviate stress levels in ER.

4.8.6: Redox Chemistry and Protein Folding

In general, we envision the interior of a cell to be in a reducing environment. Cells have sufficient concentrations of "b-mercaptoethanol"-like molecules (used to reduce disulfide bonds in proteins in vitro) such as glutathione (g-Glu-Cys-Gly) and reduced thioredoxin (with an active site Cys) to prevent disulfide bond formation in cytoplasmic proteins. For disulfide bonds to occur in a protein, a free sulfhydryl reacts with another one on a protein to form the more oxidized disulfide bond. This reaction occurs more readily if one of the Cys side chains had a lowered pKa (due to its immediate environment) making it a better nucleophile in the reaction. Most cytoplasmic proteins contain Cys with side chain pKa > 8, which would minimize disulfide bond formation as the Cys are predominantly protonated at that pH.

Disulfide bonds in proteins are typically found in extracellular proteins, where they serve to keep multisubunit proteins together as they become diluted in the extracellular milieu. These proteins destined for secretion are cotranslationally inserted into the endoplasmic reticulum (see below) which presents an oxidizing environment to the folding protein and where sugars are covalently attached to the folding protein and disulfide bonds are formed (see Chapter 3D: Glycoproteins - Biosynthesis and Function). Protein enzymes involved in disulfide bond formation contain free Cys which form mixed disulfides with their target substrate proteins. The enzymes (thiol-disulfide oxidoreductases, protein disulfide isomerases) have a Cys-XY-Cys motif and can promote

disulfide bond formation or their reduction to free sulfhydryls. They are especially redox-sensitive since their Cys side chains must cycle between and free disulfide forms.

Intracellular disulfide bonds are found in proteins in the periplasm of prokaryotes and in the endoplasmic reticulum (ER) and mitochondrial intermembrane space (IMS) of eukaryotes. For these proteins, the beginning stage of protein synthesis (in the cytoplasm) is separated temporally and spatially from the site of disulfide bond formation and final folding. Disulfide bonds can be generated in a target protein by concomitant reduction of a disulfide in a protein catalyst, leaving the net number of disulfides constant (unless the enzyme is reoxidized by an independent process). Alternatively, a disulfide can be formed by the transfer of electrons to oxidizing agents such as dioxygen.

In the ER, disulfide bond formation is catalyzed by proteins in the disulfide isomerase family (PDI). To function as catalysts in this process, the PDIs must be in an oxidized state capable of accepting electrons from the protein target for disulfide bond formation. A flavoprotein, Ero1, recycles PDI back to an oxidized state, and the reduced Ero1 is regenerated by passing electrons to dioxygen to form hydrogen peroxide. In summary, on formation of disulfides in the ER, electrons flow from the nascent protein to PDIs to the flavin protein Ero1 to dioxgen (i.e. to better and better electron acceptors). The first step is really a disulfide shuffle, which, when coupled to subsequent steps, leads to de novo disulfide bond formation.

In the mitochondria, disulfide bond formation occurs in the intermembrane space (IMS) and is guided by the mitochondria disulfide relay system. This system requires two important proteins: Mia40 and Erv1. Mia40 contains a redox-active disulfide bond cys-pro-cys and oxidizes cys residues in polypeptide chains. Erv1 can then reoxidize Mia40 which can in turn get reoxized by the heme in cytochrome c. Reduced cytochrome C is oxidized by cytochrome C oxidase of electron transport through the transfer of electrons to dioxygen to form water. The importance of IMS protein oxidation is less understood, but it is believed that the oxidative stress caused by a dysfunction could lead to neurodegenerative diseases.

A recent review by Riemer et al compares the ER and mitochondrial processes for disulfide bond formation:

- Many more and diverse proteins form disulfides in the ER compared to the IMS. Most in the IMS have low molecular mass and have two disulfide bonds between helix-turn-helix motifs. These protein substrates include chaperones that facilitate the localization of proteins in the inner membrane, and in proteins involved in electron transport in the inner membrane.
- There are many PDIs in the ER, probably reflecting the structural diversity of protein substrates in the ER. However, Mia40 appears to be the only PDI in the IMS.
- "De novo" disulfide bond formation is initiated by Ero1 in the ER and Erv1 in the IMS. Convergent evolution led to similar structures for both - a 4-helix bundle that binds FAD with two proximal Cys.
- The mitochondria pathway leads to water formation on reduction of dioxygen, not hydrogen peroxide, minimizing the formation of reactive oxygen species in the mitochondria. The peroxide formed in the ER is presumably converted to an inert form.
- The IMS is in more intimate contact with the cytoplasm through outer membrane proteins called porins which would allow some glutathione access. The IMS presents a more oxidizing environment than the cytoplasm (with more glutathione). The ER, without a porin analog, would be more oxidizing.
- The reversible formation of disulfides in the ER regulates protein activity.

Disulfide bond regulation in the Periplasmic Space of Bacteria

The redox sensitivity of the Cys side chain found in disulfide bonds is important in regulating protein activity. In particular, the thiol group of the amino acid Cys, an important nucleophile often found in the active site, can be modified to control protein activity. The formation of a disulfide bond or the oxidation of free thiols to sulfenic acid or further to sulfinic or sulfonic acid can block protein activity. The E. Coli periplasmic protein DsbA (disulfide bond A) converts adjacent free thiols into disulfide-linked Cystine, in the process becoming reduced. DsbB is reoxidized by DsbA back to its catalytically active form. What about periplasmic protein like YbiS with an active site Cys? Since the environment of the periplasm is oxidizing, YbiS is protected from oxidative conversion of the free Cys to either sulfinic or sulfonic acids causing the protein to become inactive. The mechanism involves two periplasmic proteins known as DsbG and DsbC which are similar to thioredoxin. These two proteins are able to donate electrons to the unprotected thiol preventing it from becoming oxidized, which allows YbiS to remain active in the periplasm. To maintain activity, DsbG and DsbC are reduced by another periplasmic protein, DsbD.

4.8.7: Protein Transport Across Membranes

How does a protein "decide" its final location after synthesis? Protein synthesis occurs in the cytoplasm, but proteins may end up outside of the cell, in cell membranes, internalized into various organelles, or remain in the cytoplasm. How is the decision made?

There must be signals in the protein which target proteins to various sites in a cell, where processing can occur. Proteins that are destined for secretion or plasma membrane insertion typically have a signal peptide at the N-terminus which binds to a signal recognition particle in a cotranslational process, which temporarily arrests translation and nascent folding. This complex docks to signal recognition complex docking sites in the endoplasmic reticulum membrane, where translation continues as the nascent polypeptide extends through a protein pore in the ER membrane. Transport across the ER membrane can also occur partially or fully in a post-translational process if nascent proteins are partially or fully folded through interactions with cytoplasmic chaperones such as Hsp70/40. Figure 4.8.25 shows the cotranslational (a) and post-translation (b) pathways for uptake into the ER lumen.

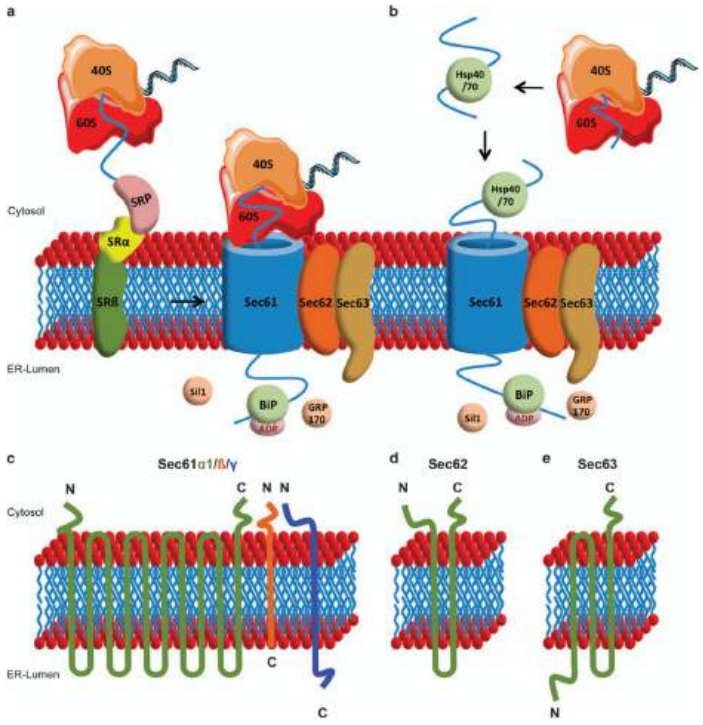


Figure 4.8.25: Co- and Post-translational transport of proteins across the ER membrane. Linxweiler, M. et al. *Sig Transduct Target Ther* **2,** 17002 (2017). https://doi.org/10.1038/sigtrans.2017.2 Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/

In both (a) and (b), the protein is shown during synthesis as it is bound to the ribosome (40S/60S) nanoparticles. The cytoplasmic signal recognition particle (SRP) binds to the hydrophobic signal sequence of the nascent protein. The signal is typically at or near the N-terminus of the growing protein. Either co- or post-translationally, the nascent protein is delivered to Sec61, which is both a ribosome receptor and a gated pore for passages of the target protein. Sec61 is part of a larger ER translocon complex which also includes Sec62 and Sec63. The membrane topology and subunit structure of the Sec proteins are shown in (c). Addition proteins including the chaperones BiP Hsp70/40 are also shown.

Protein transport across the endoplasmic reticulum membrane. Mechanism of (**a**) co-translational and (**b**) posttranslational transport of precursor proteins through the Sec61 channel. (**c**) Topological domains of Sec61α1/ß/γ, (**d**) Sec62 and (**e**) Sec63. We note that (i) Sec63 interacts with Sec62 involving a cluster of negatively charged amino-acid residues near the C terminus of Sec63 and positively charged cluster in the N-terminal domain of Sec62, (ii) Sec62 interacts with the N-terminal domain of Sec61α via its C-terminal domain, (iii) BiP can bind to ER luminal loop 7 of Sec61 α via its substrate-binding domain and mediated by the ATPase domain of BiP and the J-domain in the ER luminal loop of Sec63, (iv) Ca^{2+}-CaM can bind to an IQ motif in the N-terminal domain of Sec61α and (v) LC3 can bind to a LIR motif in the C-terminal domain of Sec62. 40S, 40S ribosome subunit; 60S, 60S ribosome subunit; SR, heterodimeric SRP receptor; SRP, signal recognition particle.

Figure 4.8.26 shows an interactive iCn3D model of the Sec Complex from yeast (6ND1).

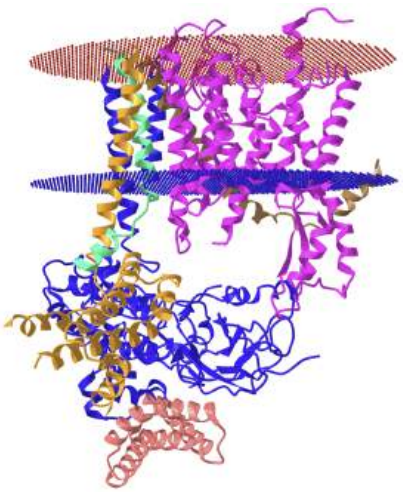


Figure 4.8.26: Sec Complex from yeast (6ND1). (Copyright; author via source).

Click the image for a popup (which loads slowly) or use this external link: https://structure.ncbi.nlm.nih.gov/i...6VoV8gTiUCQko8

The model shows Sec 61 and SEc 63 from Figure 25 above. Sec 61 is the ER protein-conducting channel (the analog in prokaryotes is Sec Y).

The model is colored as follows:

- Protein transport protein SEC61 (B)- magenta
- Protein translocation protein SEC63 (A) - blue
- Protein transport protein SSS1 - darker brown
- Protein transport protein SBH1 - green
- Translocation protein SEC66 - gold
- Translocation protein SEC72 - salmon

If the signal sequence of the protein to be imported not very hydrophobic, it doesn't bind the signal recognition particle, and hence is imported post-translationally. For these proteins, the Sec61 channel requires additional proteins, Sec62 and Sec63. These required also the BiP Hsp70/40 ATPase for import. Sec63 opens a gate on Sec61 leading to a wide opening, allowing proteins into the lipid bilayer.

If destined for secretion, a protein enters the lumen of the ER. Proteins destined for insertion into the cell surface membrane gets "stuck" in the ER membrane, and through a process of vesiculation merges with the Golgi and eventually with the cell surface membrane. Proteins that are taken into organelles like mitochondria are done so in a post-translational process that requires facilitation by protein chaperones. Final protein folding occurs inside the organelle. In both cases, nonnative proteins pass through the membrane after which final folding occurs.

An intriguing question is how the decision is made to keep a protein either in the membrane or allow it to pass through completely (in the case of proteins destined for secretion). Hessa et al investigated this "decision-making" process by studying the eukaryotic membrane pore protein complex, Sec 61 translocon (show in the above figures), whose activity must be closely regulated with the folding of the growing protein. In studying this process, they considered three local regions in a membrane: the hydrophobic region comprised of the nonpolar acyl tails of membrane lipids, the interfacial region in the vicinity of the polar head groups, and the aqueous regions (bulk water) on each side of the head groups. A 19 amino acid peptide was used as the experimental model protein which was added to the translocon. This size was chosen since it is just long enough to span the hydrophobic part of the membrane if the peptide were in an alpha-helical conformation (which is common in membrane-spanning proteins). They varied the proportion of amino acids that tend to partition into each of three regions and studied the disposition of the peptide after interaction with membrane and translocon. To test if the results were consistent with the thermodynamics of amino acid partitioning into nonpolar environments (and not kinetic considerations), they used the Wimley and White hydrophobicity scale, based on the free energy of transfer of amino acid side chains into nonpolar environments, to predict target peptide disposition with the membrane. The table below shows the propensity of amino acids to be in each region at equilibrium, based on this hydrophobicity scale.

Table: Amino Acid Partitioning Into Membrane Regions

Region	Amino Acids
Bulk water	Arg, Asn, Asp, Gln, Glu, His, Lys, Pro
Bulk water + interfacial	Ala, Cys, Gly, Ser, Thr
Interfacial	Tyr
Hydrophobic	Ile, Leu, Met, Phe, Trp, Val

Their experimental results were in concordance with those predicted using the above scale. If a polyalanine 19 mer was used, no insertion was observed. With five leucines in the peptide, almost 90% was inserted into the membrane. The results would be modeled using a two-state equilibrium: Peptide inserted ↔ Peptide translocated.
They then substituted each of the twenty amino acids into a given position into a target peptide and used the results to develop an empirical scale for membrane transfer, not one based on the simple transfer to nonpolar medium. This new scale matched the hydophobicity scale, suggesting insertion and transfer decisions were based on the thermodynamics of side chain partitioning. They also varied the position of the varied amino acid in the test peptide. If the amino acid favored the bulk and/or interfacial region, the peptide would be inserted if that amino acid were at the end of the peptide, not the middle. For translocation, the peptide had to be amphiphilic with one face polar and the other nonpolar.

They developed a simple equilibrium model to show the processes involved, as shown below in a top-down view of the membrane in Figure 4.8.27.

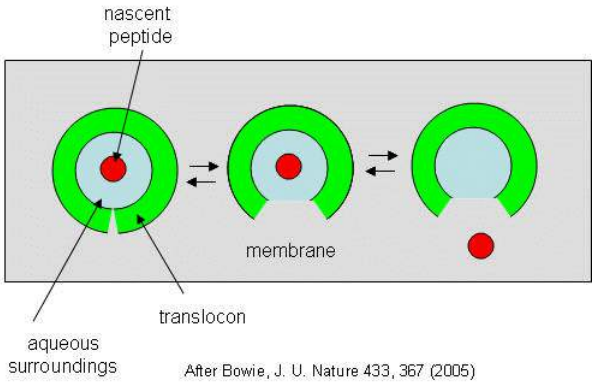


Figure 4.8.27: Translocon Equilibrium Model

The translocon, shown in green, has a water-filled pore but also a sidewise opening toward the membrane interior. The target peptide enters the pore. Transient conformational changes in the pore expose the peptide to the nonpolar membrane core. The target peptide samples both the aqueous and nonpolar environments and partitions into them based on the considerations mentioned above. If it partitions more favorably into the hydrophobic core, it will do so and cause the peptide to become membrane-bound. Otherwise, it will pass through to the other side. This can be modeled as an equilibrium process if the rate of translocation is slow compared to the rates of translocon conformational change and environmental sampling by the peptide. Obviously, the process becomes more complicated if the target is a large protein.

Bacterial toxin proteins also have evolved ways to pass through a cell membrane, again in a nonnative state, through a protein channel in the membrane. Krantz et al have recently worked out details of how the anthrax toxin protein moves through eukaryotic cell membranes. Three anthrax proteins are involved. One is a "prepore" protein that binds to specific proteins on the cell membrane, where it is activated by limited proteolysis to form a pore protein, which assembles into the homoheptamer prepore in the membrane. Two other proteins secreted by the bacteria, lethal factor and edema factor, bind to the heptamer complex and the whole assembly is then taken up into the cell by invagination to form a vesicle with the pore complex in the membrane. This vesicle fuses with a lysosome in the cell, and upon acidification, a conformational change occurs in the prepore complex to activate it. The lethal and edema factors unfold partially, possibly to a molten globule state, and are then passed through the pore into the cell where they exert their toxic influences. An electrochemical potential gradient (which we will discuss later in the semester) is required for the passage of the factors through the membrane. The active pore further unravels the factor protein, facilitating transport.

Krantz et al. studied the pore protein by mutating two amino acids, Phe427 and Ser 429, on each monomer of the pore to Cys. They then modified the Cys with [2-(trimethylammonium) ethylmethanethiosulfonate and observed effects on ion conductance of the

pore and pore conformations. They noted that when both residues were mutated and chemically modified, that ion conductance was blocked, suggesting that these side chains were localized in the narrowest part of the channel. When Phe 427 alone was mutated to smaller side chains (Ala), ion conductance increased but the transfer of peptides from the factor proteins was inhibited. This suggested that an aromatic ring in the narrow part of the channel opening participates in the translocation of bacterial proteins through the membrane. They then analyzed the transport of a variety of small molecules with varying hydrophobicity through the wild-type pore. Their results were consistent with the binding of the molecules through hydrophobic and aromatic electron interactions. They suggest a mechanism of transport consistent with their data in which the unfolded protein "ratchets" through the pore, which promotes factor protein unfolding to expose more hydrophobic groups to the nonpolar aromatic ring in the pore. This mechanism is similar to how the chaperone complex GroEL/GroES unfolds protein in its large central cavity in a process which requires hydrolysis of ATP, not a transmembrane potential. In addition, the Sec61 translocon in the inner membrane of bacteria and in eukaryotic ER membranes also has a pore containing a ring of hydrophobic groups (Ile).

4.9: Protein Stability - Thermodynamics

4.9.1: Introduction to Protein Stability

This material is not easy, and is perhaps the most intellectually challenging of the entire book. Much of the organizational framework for this section comes from an article by Ken Dill, Biochemistry, 29, 7133-7155 (1990) as that article so clearly defined the factors that contribute to protein stability.

Extrapolating from the results of studies of the transfer of small molecule H bond donors/acceptors and hydrophobic molecules from water to nonpolar solvents, it would appear that H bond interactions (as well as ion ..ion interactions) do not drive protein folding per se. Rather, the biggest contributors to the stabilization of the native state are the hydrophobic effect and the van der Waals interactions among the tightly packed buried atoms of the protein. However, using actual data from wild-type and mutant proteins of known structure, it appears that H bonds contribute significantly to protein folding and stability, and may make a greater contribution to the stability of the native state than the hydrophobic effect. The main factor which opposes folding is chain conformational entropy as folding proceeds from very populated denatured states to a "single" folded state. These positive and negative factors sum up to a small negative ΔG favoring protein folding, implying marginal stability of the native protein at normal temperatures.

What types of noncovalent forces might act within a protein and between proteins and solvent molecules that would cause a protein to fold spontaneously to a unique 3D structure? These forces can be long range (ion-ion, ion-dipole, or dipole-dipole) or short range (van der Waals repulsive and attractive forces). The interactions can be local (between adjacent amino acids in the linear sequence) or nonlocal (between sequences separated in the linear sequence but brought close together in 3D space). Clues as to what stabilizes the tertiary structure of a native protein can be gained by subjecting proteins to agents that unfold or denature proteins. Such agents include extremes of pH, high concentrations of some salt solutions or organic solvents, and temperature extremes. Such experiments show that native proteins are only marginally stable (about 0.4 kJ/mol amino acid - or around - 10 kcal/mol (-42 kJ/mol) for a protein of molecular weight of 10,000 - about 100 amino acids). This is equivalent to the stability provided by 2 H bonds. We will consider the different types of noncovalent attractions (ion-ion, H bonds, van der Waals), and stabilizing influences arising from the hydrophobic effect. and ask if each is a significant driving force for protein folding. Figure 4.9.1 shows the relative contributions to the ΔG for protein folding.

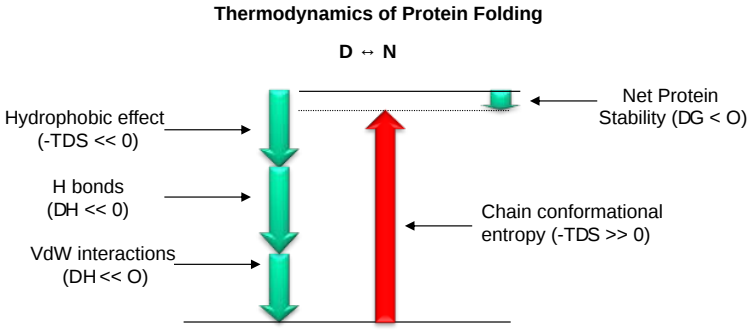


Figure 4.9.1: Relative contributions to ΔG for protein folding

Most of this chapter will deal with H bonding and the hydrophobic effect. A theme of any biochemistry course is that if you can understand the interactions among small molecules, you can apply that knowledge to the understanding of larger molecules like proteins. To understand if H bonds within proteins, often buried in the more hydrophobic interior of the protein, drive protein folding, we will first examine the thermodynamics of H bond formation of a small molecule, N-methylacetamide, in water and in a nonpolar solvent. To understand if the hydrophobic effect, mediated by the burying of nonpolar side chains within the more nonpolar center of the protein, drives protein folding, we will examine the thermodynamics of benzene solubility in water. Most recent studies involve the creation of specific mutants at amino acid positions that might reveal the contributions of H bonding and the hydrophobic effect to folding and protein stability.

4.9.2: Ion - Ion Interactions

These could be investigated by altering pH or ionic strength. Why is that?

a. General Charge Interactions - Proteins denature at either low or high pHs where they have a maximal positive charge or negative charge, respectively. Small proteins that seem to fit a simple two-state folding model (F ↔ U) have a characteristic melting temperature (T_M) at which the 50% of the proteins in the population are unfolded. The higher the T_M, the more stable the protein. The graphs of the T_M for different proteins at different pH values are shown in Figure 4.9.2.

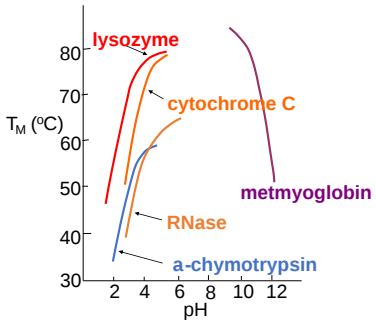


Figure 4.9.2: T_M for small proteins at different pHs. After Dill, Biochemistry, 29, 7133-7155 (1990)

Under extremes of pH (but not so great as to catalyze peptide bond cleavage, electrostatic repulsions would cause the protein to denature. The folded, compact state has an increasing charge density (q/Vol) at pH extremes, which could be alleviated by unfolding to a less dense state. But what about specific charge pair interactions? In contrast to the general charge interactions, these might actually stabilize a protein. Are they the predominant factor that determines stability?

b. Specific Charge Interactions (charge pairs) - If ion pairs are the source of protein stability, you would expect that high salt could disrupt them, and lead to denaturation. Although some salts do denature proteins, others stabilize them. Other evidence argues against this idea. Ion pairs are not conserved in evolution. In addition, the number of ion pairs in proteins is small (approx. 5/150 residues, with one of those on average buried). Also, the stability of a protein shows little dependence on pH or salt concentration (at low concentrations) near the isoelectric point, the pH at which proteins have a net zero charge.

The overall charge state affects not only the stability of a protein but also its solubility. Proteins are most insoluble at their isoelectric point pI since at that pH value (where they have a net 0 charge) the proteins experience the least electrostatic repulsion and are most likely to aggregate and precipitate. Low salt concentration also promotes insolubility. Mutagenesis studies show that solubility can be increased by replacing nonpolar groups on the surface with polar ones. Pace (2009) cites studies on RNase Sa which has a maximally exposed Thr 76. If it is replaced with aspartic acid, the solubility increased to 43 mg/ml but if it is replaced with tryptophan, it decreased to 3.6 mg/ml. His, Asn, Thr and Gln have a negative effect on solubility near the pI compared to Ala, a surprising result. Similar results were obtained compared to Ala when Arg and Lys were used. Smaller side chains, Asp and Ser, at position 76 increased the solubility over Ala.

4.9.3: Hydrogen Bonding

Linus Pauling first suggested that H bonds (between water and the protein and within the protein itself) would play a dominant role in protein folding and stability. It would seem to make sense since amino acids are dipolar and secondary structure is common. Remember, however, the H bonds would be found not only in the native state but also in the denatured state. Likewise, there are H bonds between water and proteins in both the native and denatured states. Do H bonds in proteins contribute differently to the stability of the D vs N states? Many experimental and theoretical studies have been performed investigating helix ↔ (random) coil transitions in small peptides. Remember all the intrachain H bonds in the helix? Are they collectively more stable than H bonds between water and the peptide in a (random) coil?

In this section, we will explore the effects of H bonds on small molecules and infer from them the likely contributions of H bonds to protein stability. After all, if we can't understand small molecules and their interaction, how can we understand the same interactions in big ones like proteins? We will see however that when we use modern tools of site-directed mutagenesis to explore H bonds in proteins, we come to different conclusions as to their relative roles. So consider this immediate section an exercise in model building and critical thinking.

In thinking about conformational studies involving small peptides, it is useful to apply Le Chatelier's Principle to the equilibrium below:

random coil ↔ helix

Anything perturbant (small molecules, solvent, etc) that would preferentially interact with the helical form would push the equilibrium to the helical form.

Early models assumed that intrachain H bonds were energetically (enthalpically) more favorable than H bonds between peptide and water. But to form a hydrogen bond requires an entropy payback since a helix is much more ordered (lower entropy) than a random coil (higher entropy). At low temperatures, enthalpy predominates and helix formation in solution is favored. At high temperatures, the helix is disfavored entropically. Imagine the increased vibrational and rotational states permitted to the atoms at higher temperatures. Theoretical studies on helix-coil transitions predicted the following:

- as the chain length increases, the helix gets more stable;
- increasing the charge on the molecule destabilizes the helix, since the coil, compared to the more compact helix, has a lower charge density;
- solvents that protonate the carbonyl oxygen (like formic acid) destabilize the helix; and
- solvents that form strong H bonds compete with the peptide and destabilize the helix. In contrast, solvents such as $CHCl_3$ and dimethylformamide (a nonprotic solvent), stabilize the helix. Likewise, 2-chloroethanol and trifluoroethanol, which form none or weaker H bonds to the peptide than water, stabilize the helix. (In the case of trifluoroethanol, molecular dynamics simulations have shown that TFE preferentially interacts with (solvates) the peptide, which inhibits H bonds from the peptide backbone to water, stabilizing the intrahelical H bonds.

These helix-coil studies suggest that H bonds are important in stabilizing a protein.

But do they really? Why should these H bonds differ from those in water? It's difficult to figure out whether they are since there are so many possible H bonds (between protein and water, water and water, and protein and protein), and their strength depends on their orientation and the dielectric constant of the medium in which they are located.

If intrachain H bonds in a protein are not that much different in energy than intermolecular H bonds between the protein and water, and given that proteins are marginally stable at physiological temperatures, then it follows that the folded state must contain about as many intramolecular hydrogen bonds within the protein as possible intermolecular H bonds between the protein and water, otherwise the protein would unfold.

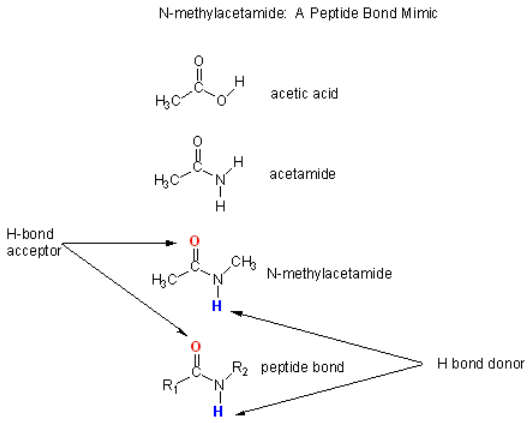
To resolve this issue, and determine the relative strength of H bonds between the varying possible donors and acceptors, many studies have been conducted to compare the energy of H bonds between small molecules in water with the energy of H bonds between the same small molecules but in a nonpolar solvent. The rationale goes like this. If the interior of a protein is more nonpolar than water (lower dielectric constant than water), then intrastrand H bonds in a protein might be modeled by looking at the H bonds between small molecules in nonpolar solvents and asking the question, is the free energy change for the following process less than zero:

$$Dw + Aw \leftrightarrow (DA)n,\ \Delta G^o,\ K$$

where D is a hydrogen bond donor (like the H in N-H) and A is a hydrogen bond acceptor, (like the O in C=O), w is water (i.e. donor and acceptor are in water), and n is a nonpolar solvent, and ΔG^o and K are the standard free energy change and the equilibrium constant, respectively, for the formation of a H-bond in a nonpolar solvent from a donor and acceptor in water. This reaction simulates H-bond contributions to protein folding, where a buried H-bond is mimicked by a H-bond in a nonpolar solvent. The reaction written above is really based on a thought experiment since it would be hard to set up the necessary conditions to make the measurement. However, we can calculate the ΔG^o for this reaction since it is a state function and it really doesn't matter how one accomplishes this process.

You learned about state functions in introductory chemistry courses so this should not be a new concept. Remember a state function is one in which the variable describing two connected states does not depend on the path between the states. A simple example involves potential energy and height. Imagine you climb a mountain to the peak. It doesn't matter which path you take up the mountain, since the difference in height (Δ height) is the same if you start at the same point and end up at the peak. Your change in potential energy is also the same. ΔG^o for protein folding is also a state function for a given set of conditions (temperature, pressure, solvent)

Let's consider a specific example, the formation of H bonds between two molecules of N-methylacetamide (NMA) in water and in a nonpolar solvent. N-methylacetamide is a good mimic for the H bond donors and acceptors of the peptide bond of a polypeptide chain as shown in Figure \(\PageIndex{3}\) below.



(\PageIndex{3}\): N-methylacetamide as a mimic for the peptide backbone

The reaction scheme shown in Figure 4.9.4 describes a set of reactions (a thermodynamic cycle) involving the formation of H-bonded dimers of NMA . A and B are both molecules of NMA, in either water (w) or a nonpolar solvent (n). N-methylacetamide is a good mimic for the H bond donors and acceptors of the peptide bond of a polypeptide chain as .

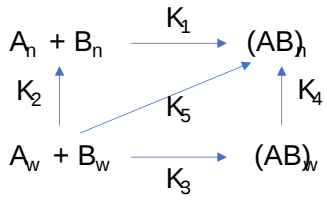


Figure 4.9.4: Thermodynamic cycle for the formation of H bond in nonpolar solvent from donor and solvent in water.

K_1 is the equilibrium constant for the dimerization of NMA in a nonpolar medium. This can be readily determined, and is >1, implying that $\Delta G^o < 0$. (Remember, ΔG^o = -RTlnKeq) For the dimerization of NMA in CCl_4, ΔG^o_1 = -2.4 kcal/mol (-10 kJ/mol).

K_2 is the equilibrium constant (think of it as a partition coefficient) for the transfer of two NMA molecules from water to a nonpolar solvent (again easily measurable). For NMA transferring from water to CCl_4, ΔG^o_2 = + 6.12 kcal/mol (+25.6 kJ/mol).

K_3 is the equilibrium constant for the dimerization of NMA in water. This can be readily be determined and is <1, implying that $\Delta G^o > 0$. For the dimerization of NMA in water, ΔG^o_3 = +3.1 kcal/mol (+13 kJ/mol).

K_4 is the equilibrium constant (think of it as a partition coefficient) for the transfer of a hydrogen-bonded dimer of NMA from water to a nonpolar solvent. You try to think of a way to measure that! I can't. This is where thermodynamic cycles come in so nicely. You don't have to measure it. You can calculate it from $K_{1\text{-}3}$ since ΔG^o is a state function!

$$\begin{aligned} &\Delta G_2^0 + \Delta G^0{}_1 = \Delta G_3^0 + \Delta G_4^0 \text{ or} \\ &-\mathrm{RTInK_2} + -\mathrm{RT}\ln \mathrm{K_1} = -\mathrm{RT}\ln \mathrm{K_3} + -\mathrm{RT}\ln \mathrm{K_4} \\ &\mathrm{InK_2} + \ln \mathrm{K_1} = \ln \mathrm{K_3} + \ln \mathrm{K_4} = \ln(\mathrm{K_2\ K_1}) = \ln(\mathrm{K_3\ K_4}) \text{ or} \\ &\mathrm{K_2\ K_1} = \mathrm{K_3\ K_4} \end{aligned} \tag{4.9.1}$$

For NMA transferring from water to CCl_4, ΔG^o_4 = + 0.62 kcal/mol (+2.6 kJ/mol).

(Note: Biochemists like to talk about thermodynamic cycles which may seem new to you. However, believe it or not, you have seen them before in introductory chemistry in the form of Hess's Law!)

From $K_{1\text{-}4}$ and the corresponding ΔG^o values, we can now calculate ΔG^o_5 for the formation of H-bonded NMA dimers in a nonpolar solvent from two molecules of NMA(aq). This reaction, which we hope simulates the formation of buried intrachain H bonds in proteins on protein folding, is:

Dw + Aw ↔ (DA)n, for which ΔG^o_5= +3.72 (i.e. disfavored).

If this model is a good mimic for studying H bond formation on protein folding, it suggests that the formation of buried H bonds during protein folding does **not** drive protein folding.

However, if the transfer of D and A (from a large protein) from water to the nonpolar medium (modeled by K_2) is driven by other effects (such as the hydrophobic effect), the negative value for ΔG^o_1 will significantly favor buried H bond formation. So, if this happens in proteins, it is clear why so many intrachain H bonds form, since K_1 is favored. H bonds may not assist the collapse of a protein, but would favor internal organization within a compact protein. That is, H bonds don't drive protein folding per se, but form so that the folded protein would not be destabilized by too many unsatisfied H bonds.

There are potential problems with this simple model. The interior of a protein is not homogeneous (i.e. the effective dielectric within the protein will vary). H bond strength is also very sensitive to geometry. Also, there are many H bonds within a protein, so slight errors in the estimation of H bond strength would lead to large errors in the determination of protein stability.

Another argument against H bonds being the determining factor in protein folding and stability comes from solvent denaturation studies. If intrachain H bonds are so important, then should not solvents that can H bond to the backbone denature the protein? Shouldn't water (55 M) act as a denaturant? It doesn't, however. Dioxane (5 membered heterocyclic ring with O) which has only a H bond acceptor wouldn't be expected to denature proteins, but it does. H bonds also increase in nonpolar solvents. Peptides, which have random structures in water, can be induced to form helices when placed in alcohol solutions (trifluorethanol, for example), which are more nonpolar than water, as explained above in the helix-coil studies. If H bonds are the dominant factor in protein stability, the alcohols would stabilize proteins. At low concentrations of alcohol, proteins are destabilized.

4.9.4: Hydrophobic Interactions: Introduction

We have studied the role of the hydrophobic effect (involving the favorable entropic release of caged water molecules about solvent-exposed hydrophobic groups) in driving nonpolar molecule solubility in Chapter 2.5. Does this also drive protein folding? To explore this question, we will study the thermodynamics of small nonpolar molecules, especially benzene, with water and ask whether the thermodynamic parameter associated with benzene solubility are similar to those associated with protein stability. If this analogy holds, anything that will promote benzene solubility will lead to increased hydrophobic amino acid side chain exposure to water and hence protein denaturation. What is the evidence to support this?

a. Crystal structures: PDB crystal structures show that most nonpolar side chains are buried inside a protein, which is tightly packed and which excludes water. Studies show that as the surface area of amino acid side chains increases, the free energy of transfer of amino acids from water to ethanol becomes more negative as shown in Figure 4.9.5.

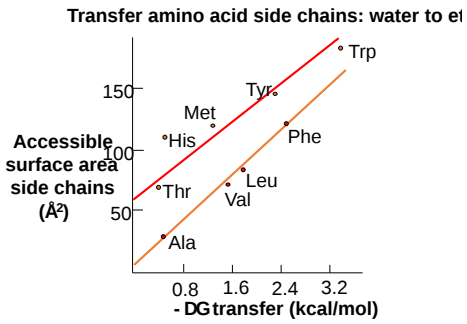


Figure 4.9.5: Transfer free energy of amino acids from water to ethanol

b. Low-temperature denaturation of proteins - It has been observed that proteins can denature at low temperatures (less than 0^o C), suggesting that nonpolar residues become more "soluble" in water at low temperatures (i.e. they move from the more hydrophobic interior of a protein to the more polar outside). Compare the solubility of nonpolar gases like CO_2 or N_2, which are more soluble at low temperatures. As you heat solutions of nonpolar gases in water, the gases become less soluble as evidenced by bubble formation (i.e. phase separation of dissolved gases as they become more insoluble, as you have observed with a bottle of soda). If protein behavior is governed by this same behavior (greater solubility of nonpolar groups at low temperatures), it would suggest that proteins might denature at low temperatures (leading to increased exposure to water of the nonpolar side chains). This phenomenum has been observed.

c. Protein stability affected by different salt species - Over 100 years ago, Hofmeister determined the effectiveness of different cations and anions of salts to precipitate blood serum proteins in the 0.01 - 1 M concentration ranges. The series is shown below:

Cations: $NH_4^+ > K^+ > Na^+ > Li^+ > Mg^{2+} > Ca^{2+} >$ guanidinium

Anions: $SO_4^{2-} > HPO_4^{2-} >$ acetate > citrate > $Cl^- > NO_3^- > ClO_3^- > I^- > ClO_4^- > SCN^-$

- A salt of pairs of the first ions in these series, for example, $(NH_4)_2SO_4$, when added to aqueous solutions of proteins, precipitate the native form of the protein. We must account for the fact that it precipitates the protein, and that the protein is precipitated in the native, not denatured state. More on why it precipitates proteins in a moment. The first ion in each series increases the surface tension of water (making it harder to make a cavity in the water to fit the nonpolar molecule). This decreases the solubility of nonpolar molecules. These "salt-out" nonpolar molecules, promoting not dissolution in water but aggregation followed by a phase separation. By analogy, they will stabilize the native state since the buried hydrophobic side chains would have a decreased propensity to move out into the aqueous environment.
- The last ions of the series have less effect on surface tension, and hence increase the **solubility** of nonpolar molecules ("salt-in"). By analogy, they will destabilize the native state since the buried hydrophobic side chains would have an increased propensity to move out into the aqueous environment.

The Hofmeister Series and its effects on the chemical properties of water and solutes are shown in Figure 4.9.6.

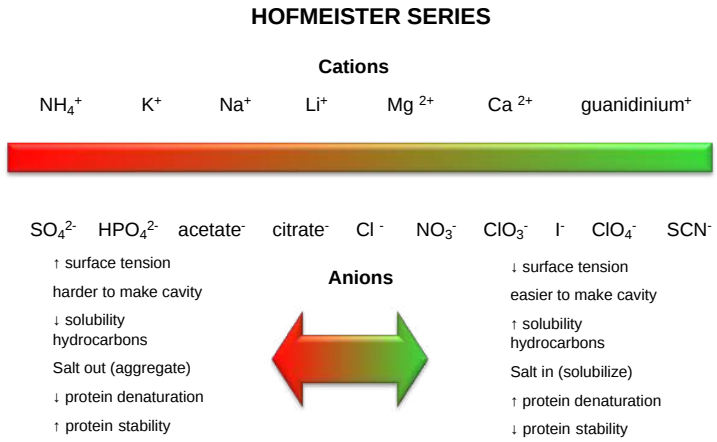


Figure 4.9.6: Hofmeister Series

The solubility of benzene in aqueous salt solutions of this series increases from left to right, just as native protein stability decreases from left to right (i.e. the protein's nonpolar core residues become more "soluble" in water, leading to its denaturation).

d. Conservation of hydrophobic core residues - These residues are highly conserved and correlated with structure.

e. Urea denatures proteins - Another additive, urea ($H_2N(C{=}O{-})NH_2$), at high concentrations is often used to denature proteins. People used to think that urea competed with the intrachain H bonds and hence unraveled the protein. The arguments above with H bonds dispute this contention since water should then denature protein. How does urea denature proteins? It has been shown that the free energy of transfer of the nonpolar amino acids into 8M urea is increasingly negative as the side chains become bigger and more nonpolar. This is also true for denaturation by guanidine hydrochloride. Both also increase the solubility of nonpolar molecules in a manner proportional to their surface area. The structure of urea and guanidinium, along with the side chains of arginine and the Type II diabetes drug metformin, are shown in Figure 4.9.7.

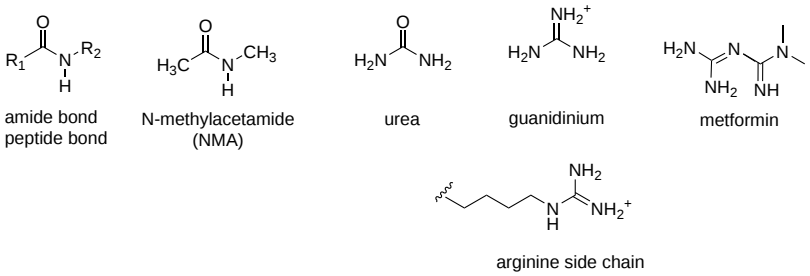


Figure 4.9.7 : Structures of urea, guanidine HCl, arginine and metformin

Figure 4.9.8 shows the free energy of transfer of the nonpolar amino acids into 8 M urea and 6 M guanidine HCl as a function of the surface area of amino acid side chains.

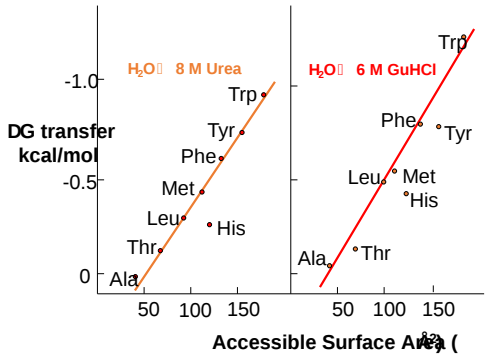


Figure 4.9.8: Free energy of transfer of the nonpolar amino acids into 8 M urea and 6 M guanidine HCl

4.9.5: Additives and Their Interactions with Protein Surfaces

Additives to proteins that increase the stability of the folded state of the protein also tend to decrease their solubilities. These additives are excluded from the preferential water hydration sphere around the protein (negative binding of these agents). Denaturants in contrast tend to increase protein solubility and interact preferentially with the protein surface. In their presence, proteins respond by increasing their surface area by denaturation. For stabilizers, proteins try to minimize their surface area by staying "native" and aggregating to form a precipitate, both of which minimize the surface area from which the stabilizing agent is excluded.

The main effect of dissolved ions on water structure has been thought to involve changes in H bonds (either enhancers/structure makers or inhibitors/structure breakers) which correlate with the salting-in or salting-out effects of various ions. Many techniques have been used to study these interactions:

- viscosity: inferential information on the structure
- diffraction (x-rays/neutrons): gives information on the coordination number of solvation shell (static information)
- NMR: information on average relaxation of bulk and hydration sphere water around ions (dynamic information)
- molecular dynamics simulations: which gives insight into short but not long-range interactions between ions and water.

Recent studies have provided conflicting support for the notion of structure makers/breakers. Omta et al. 2003 used femtosecond mid-infrared pump-probe spectroscopy to study actual H-bonds between water molecules in salt solutions ($Mg(ClO_4)_2$, $NaClO_4$, and Na_2SO_4). In pump-probe spectroscopy, a sample is excited with a short pulse (pump) and after a short time lag, with another pulse (probe), which interacts with the excited state. The linear-polarized infrared pulses (pump) were used to excite OH groups in solution, followed by a probe pulse which was polarized 45 degrees compared to the pump pulse. Only those excited OH groups that had rotated in the time interval between the pump and probe would be excited by the probe. Using this technique, the time frame for reorientation of the OH groups, which is related to the "stiffness" of the H bonds, can be determined. The salts had no effect on the rotational motion of bulk water outside of the first hydration shell, which suggests that salts have no effects on the H bond networks in bulk water. Mg^{2+} ions are considered structure makers, as the ions greatly increase the viscosity of water, brought about supposedly by increased H bonds among water molecules. This study does not support this model. Increased viscosity of Mg solutions must be attributed to those ions directly interacting with water molecules. The solution can be modeled as bulk water with small rigid spheres of ion + first hydration sphere. Clearly, much more experimental and theoretical work must be performed to gain structural insight into the role of salts on water structure. Until then, we will continue to try to understand the effects of different salt on water structure in descriptive terms and with the use of thermodynamic quantities.

Studies have shown that urea binds preferentially to the protein surface, and hence tends to increase the protein's surface area and hydrophobic exposure, and denature proteins. However, note in the figure below, that glycerol, a bigger polar but uncharged molecule, stabilizes the native state. This pair of uncharged additives have correspondingly similar effects on protein stability as do the charged guanidine HCl/ammonium sulfate pair.

Figure 4.9.9 shows how solution additives might interact with the surface of the protein.

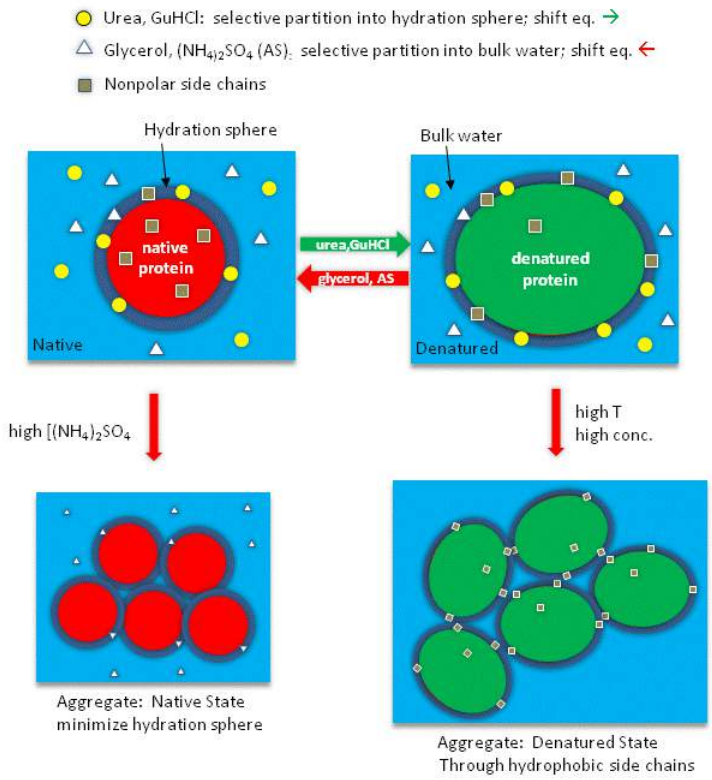


Figure 4.9.9: Interact of solution additives with the protein surface.

Figure 4.9.10shows a thermodynamic cycle for urea denaturation of proteins

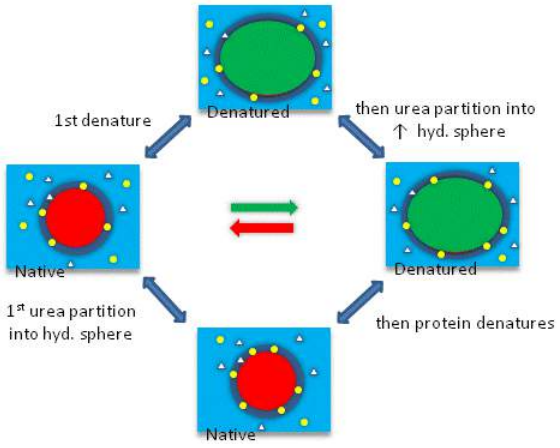


Figure 4.9.10: Thermodynamic cycle for urea denaturation of proteins

A graphical summary of the types of surface changes that result from protein denaturation is given Figure 4.9.11.

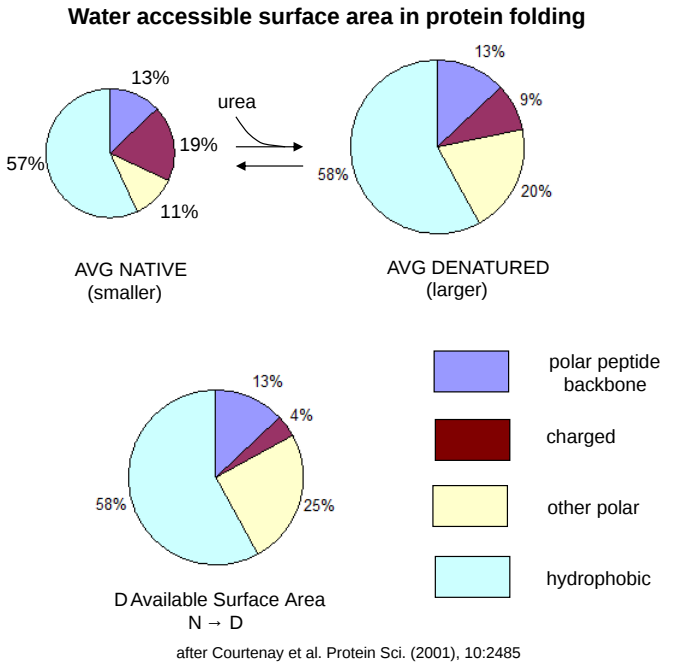


Figure 4.9.11: Water Accessible Surface Area in Protein Unfolding

4.9.6: The Hydrophobic Effect and Change in Heat Capacity

Our understanding of hydrophobic interactions has changed dramatically in the last several years. This is not reflected in most textbooks. The hydrophobic effect means different things to different people. Some refer to the transfer of nonpolar solutes/solvents to aqueous solution. Some refer to the same phenomena only if the effects have a unique temperature dependency. Other refer to the ordering of water around nonpolar residues. The most recent explanation centers around the unique temperature dependencies of the transfer reactions. Before we can understand it, here is an interesting bit of data. If you dissolve one mole of methane in hexane, the volume of 1 L of hexane changes 60 ml, but if done in water, the water volume changes 37 ml, indicating that water molecules pack more efficiently around nonpolar molecules than in its absence.

Let's now consider the thermodynamic aspects of the hydrophobic effect, as we did for micelle and bilayer formation. In a brief summary, we found that the free energy of transfer of a hydrocarbon or alcohol from aqueous solution to the pure hydrocarbon or alcohol, for example, was disfavored enthalpically (unexpectedly) but favored entropically (also unexpectedly until we included solvent in our model). These experiments were done at one temperature and gave us our first initial understanding of the hydrophobic effect. We will expand on this view by looking at the enthalpic and entropic contribution to the transfer of benzene into water as a function of temperature. This will lead us to a more modern view of the hydrophobic effect. To do this we have to think about the thermodynamics of mixing two substances with different properties, such as polarity which affects solubility.

If you mix two substances A and B that aren't very soluble in each other, two opposing thermodynamic factors are relevant.

- The tendency to mix is driven by an increase in entropy.
- The mixing is usually opposed by enthalpy.

The later makes "intuitive" sense since you might expect that van der Waal interactions between A-B might be less than those of A-A and B-B (i.e. the old adage "like dissolves like"). If AA and BB self-interactions are stronger, A would not dissolve in B and vice/versa. You would also expect no significant changes in entropy and enthalpy as a function of temperature in this ideal mixing.

The most modern understanding of the hydrophobic interactions involves the mixing of A and B which is characterized by a unique temperature dependency for the value of the change in entropy and enthalpies. At room temperature, if one corrects the entropy changes for effects due just to mixing, the "excess" entropy is what principally opposes taking a nonpolar molecule into water. Enthalpy changes are small. We have modeled this effect using structured water around the nonpolar residues. We will now further our understanding of the hydrophobic effect by using benzene transfer from water as a model system (much as we did with N-methylacetamide for H bonding).

Before we discuss entropy and enthalpy changes accompanying protein folding/unfolding, let's try to learn about the thermodynamic aspects governing benzene solubility in water. What happens to benzene solubility in water and the corresponding thermodynamic parameters as you raise the temperature? The graph in Figure 4.9.12 shows the change in G, H and -TS for taking benzene from pure benzene to water, $B_B \rightarrow B_W$. This is real data. Note that the graph shows -TΔS not just ΔS. When the values of each of these terms, ΔG, ΔH and -TΔS are negative, the transfer of benzene into water is favored. The blue highlighted region shows more values in a more physiologically relevant temperature range (0-100^0 C).

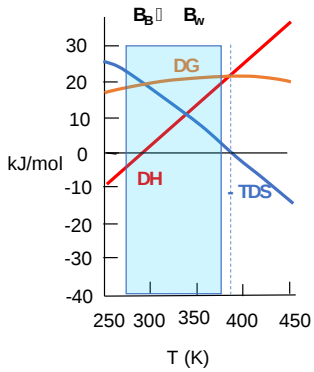
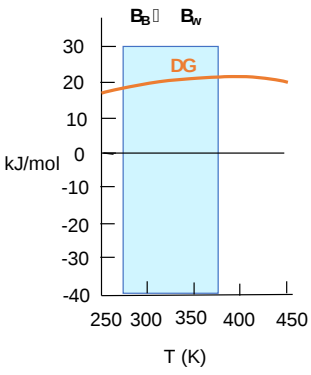


Figure 4.9.12: ΔG, ΔH and -TΔS for transfer of benzene from pure benzene to water.

Two things to note for right now. ΔG for $B_B \rightarrow B_W$ is always positive and hence always disfavored. Another key feature is shown in the vertical dotted line at the temperature at which benzene has its greatest "aversion" to being in water, where the slope of the ΔG vs T curve (dΔG/dT) = 0. At this point, entropy plays no role in the solubility of benzene in water. This goes again the notion of the hydrophobic effect we discussed in Chapter 2.5. This maximum aversion to water is completely driven by enthalpy.

If these axes showed three lines describing the number of ducks, geese and lunes in a lake as a function of lake temperature, no further explanations might be needed. But many years of teaching experience and research shows that thermodynamics parameters are examples of threshold concepts that students struggle with for many years. A graph showing three simultaneously might be uninterpretable to many. So let's deconstruct this graph into a series of stepwise graphs with targeted explanations to facilitate your understanding. These graphs with explanations are shown below.

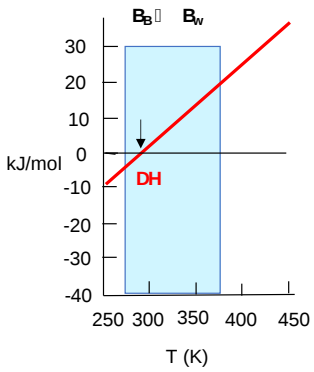
A graph of ΔG vs T for $B_B \rightarrow B_W$



ΔG for $B_B \rightarrow B_W$ is positive and hence **disfavored** over the entire temperature range shown. However note several important features:

- There is a temperature (around 375^{0}K) when ΔG is at a maximum (i.e. in the language of math, the slope of the curve ΔG/ΔT or dΔG/dT is a maximum.
- As the temperature decreases from T_{max}, the transfer of benzene becomes less disfavored. If you were to extrapolate the curve to really low but unrealistic temperatures, it looks like it would become favorable.

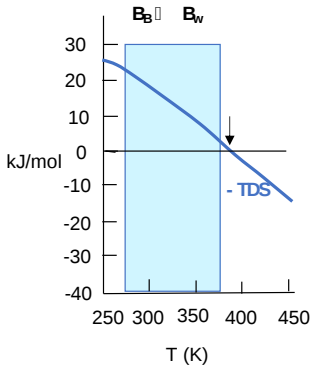
A graph of ΔH vs T for $B_B \rightarrow B_W$



The graph of ΔH vs T for or $B_B \rightarrow B_W$ is linear with a positive slope and crosses the T axis at the point indicated by the arrow.

- At temperatures below the point indicated by the arrow, the transfer of $B_B \rightarrow B_W$ is **enthalpically** favored. We saw this unexpected finding in Chapter 2.5 in considering the transfer of butanol, for example, from the pure liquid alcohol to water at one temperature, $25^0C = 298K$. That data is consistent with this curve.
- At temperatures above the point indicated by the arrow, the transfer of $B_B \rightarrow B_W$ is increasingly disfavored.
- The slope of this line, $\Delta H/\Delta T$, has the units kJ/mol/K, which is the heat capacity of the system.

A graph of $-T\Delta S$ vs T for $B_B \rightarrow B_W$



The graph of $-T\Delta S$ vs T for for $B_B \rightarrow B_W$ is almost linear with a negative slope which shows that it becomes less entropically disfavored. The curve crosses the T- axis at the point indicated by the arrow.

- At temperatures below the point indicated by the arrow, the transfer of $B_B \rightarrow B_W$ is entropically **disfavored.** We saw this in Chapter 2.5 for the transfer of alcohols from the pure liquid into water. We called this the "**hydrophobic effect**" which is the main thermodynamic drive to move organic molecules **out of water.** We attribute this to the formation of a cavity in water for the nonpolar parts of the alcohol in a process that forms a more ordered set of H bonds compared to bulk water and which reduced the number of microstates available to the associated water molecules.
- However, at the temperature indicated by the arrow, there is **no entropic barrie**r to move benzene into water. Above those temperatures, it is favored. This seems to destroy our simple definition of the hydrophobic effect.

Let's add a bit more to our molecular interpretation of this seemingly anomalous favorable entropy at higher temperatures. As the temperature is raised, the available positional and thermal entropy of water increases significantly. It would seem logical that to then put a nonpolar residue into this system of water would become easier than putting it into more structured water (characterized by fewer microstates and lower positional and thermal entropy) at a lower temperature! (Remember from our review of thermodynamics that If the T_{surr} is high, a given heat transfer to or from the surroundings will have a smaller effect on the ΔS_{surr}. Conversely, if the T_{surr} is low, the effect on ΔS_{surr} will be greater. Atkins, in a recent General Chemistry, uses the analogy of the effect of a sneeze in library compared to one in a crowded street. An American Chemistry General Chemistry text uses the analogy of giving $5 to a friend with $1000 compared to one who has just $10.

But look at the other temperature anomaly. It becomes increasingly difficult from an enthalpic point of view to put benzene in water. At a high temperature, $-T\Delta S$ becomes zero, and there is no entropic barrier to putting benzene into water. The barrier is completely enthalpic. This is why a more nuanced definition of the hydrophobic effect has emerged.

If you sum ΔH and $-T\Delta S$ at each temperature, you get the curve shown for the total ΔG to take benzene from pure benzene to water. Notice that it is always positive, so it is always disfavored. The ΔG function is curved. It increases at low temperatures, and decreases at very high temperatures, implying that there will be a temperature at which there is a minimum solubility of benzene in water (a maximum in the positive ΔG). The minimum solubility of benzene (the max. positive ΔG) occurs when $d\Delta G/dT = 0$. This occurs when $\Delta S = 0$, and the maximal aversion of benzene to water is completely enthalpic, a statement not in accord with our initial understanding of the hydrophobic effect.

For those who have taken physical chemistry ...

For just PV work, dG = VdP - SdT, so that dG/dT = V dP/dT - S. At the maximum of the curve of G vs T, dG/dT = 0 so

V dP/dT = S. For this system dP/dT = 0 so S = 0. This is observed in the top graph above.

Here is another equation for heat capacity that you derived in physical chemistry:

Cp = dH/dT= TdS/dT.

The last equality stems from Maxwell's relationships, which Physical Chemistry students should remember.

Let's review some more thermodynamics that the physical chemists in the crowd should remember. Even introductory chemistry students should to a degree (a pun). (Molar) heat capacity, C_p, is defined as the heat required to raise the temperature of a mole of a substance 1^oC. It has units of kJ (or kcal)/mol. The slope of the enthalpy curve ΔH vs T for $B_B \rightarrow B_W$ has units of kJ/mol/K and is the **heat capacity,** Cp, where Cp = dH/dT.

To refresh your mind and demystify the term heat capacity, let's look at the heat absorbed as ice is converted to gas as a function of temperature (i.e. the phase transition for water). Figure 4.9.13 a graph of the heat absorbed, Q_{abs}/gram vs T (top), and a derivative plot of the top graph (slope of the top graph = heat capacity at each point on the T axis).

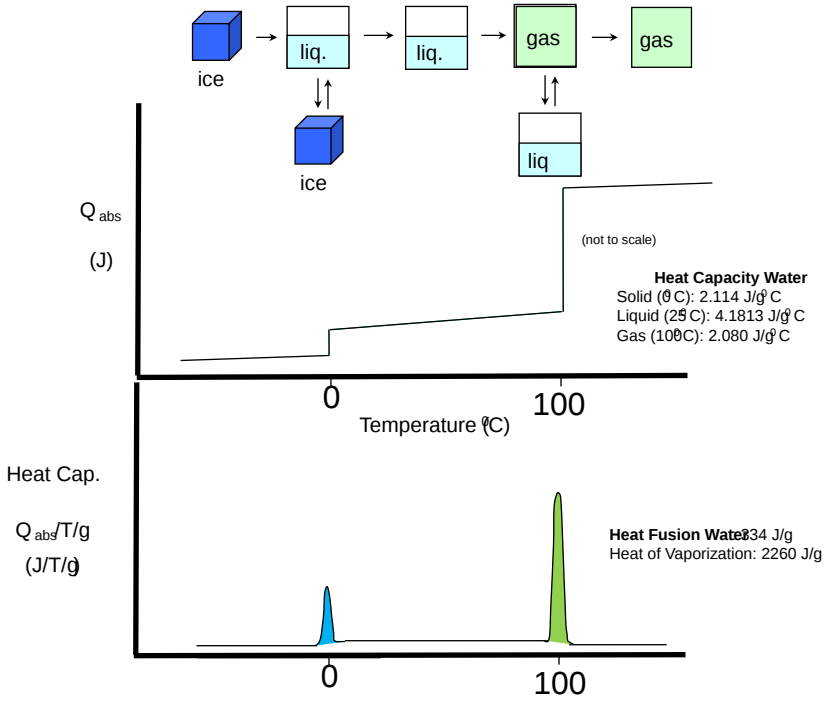


Figure 4.9.13: Q vs T and Cp vs T plots for phase transitions in water.

We can now learn one final concept from the graphs for benzene transfer into water. The graph of ΔH vs T is linear for the transfer of benzene from the pure liquid into water. We now know that the slope of that graph, ΔC_p is a constant positive number for the process when a hydrophobe is exposed to water.

This positive C_p observed when a hydrophobic group is transferred to water, is the signature of our new understanding of the hydrophobic effect.

A positive C_p occurs when both H and S are dependent on temperature, which is observed when a hydrophobe is transferred from a more nonpolar environment to water. Likewise, a negative C_p is observed when hydrophobes in water are transferred to a more nonpolar environment.

4.9.7: Hydrophobic Effect Applied to Proteins

Now let's return to the world of proteins and see how we might apply what we learned from the transfer of benzene from the pure liquid to water. An analogous transfer experiment would be the denaturation of proteins in which buried hydrophobic amino acids

are "transferred" to water. Figure 4.9.14 shows the derivative plot of heat absorbed vs T (i.e a plot of heat capacity Cp vs T) for the thermal denaturation of a protein, obtained using differential scanning calorimetry.

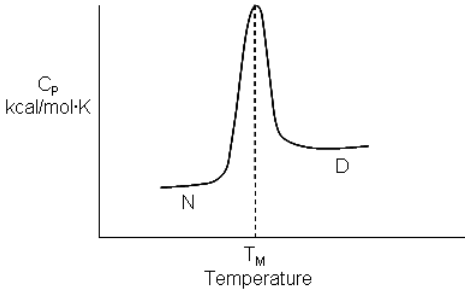


Figure 4.9.14: Cp vs T for the thermal denaturation of a protein

Look at the graph in Figure 4.9.15 which shows the heat capacity of the protein lysozyme at various pHs vs temperature obtained using differential scanning calorimetry.

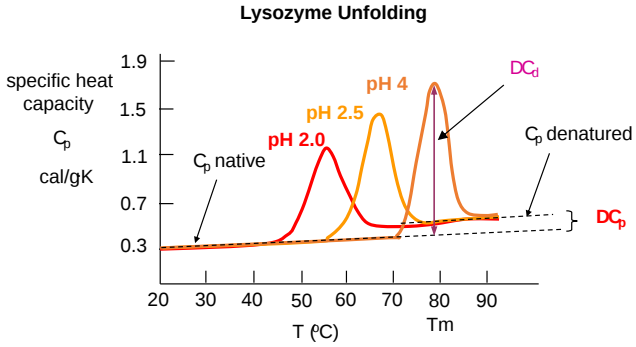


Figure 4.9.15: Heat capacity vs T for lysozyme at various pHs

As the protein is heated, it reaches a temperature at which a large amount of heat is suddenly absorbed, as the protein unfolds. The area under the curve represents the heat absorbed on denaturation (units of cal/g·K x K = cal/g). The temperature at the midpoint is the Tm of the protein. (Why would the Tm be dependent on the pH of the solution?) Notice that there are two ΔCps shown in the graph. One, $\Delta \mathbf{C}_d$, is associated with the actual denaturation process and is analogous to the change in heat capacity observed in phase changes, such as solid to liquid water. The other is ΔC_p which is the difference in heat capacity between the denatured and native state. As was the case for the transfer of benzene to water, the ΔC_p for protein denaturation is also positive, suggesting that in protein denaturation, hydrophobes are transferred from the interior of the protein to water.

Figure 4.9.16 attempts to give a molecular description that applies to different regions of the denaturation curve and helps explain the positive ΔC_p observed in thermal denaturation curves of proteins.

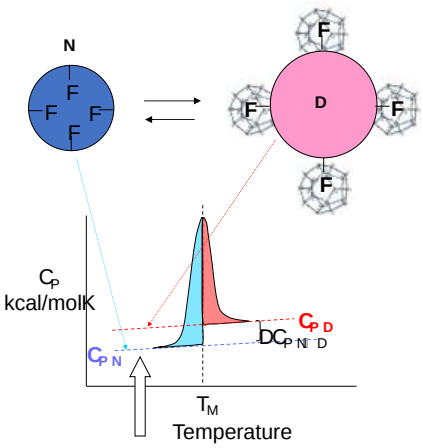


Figure 4.9.16: Molecular Descriptors associated with temperature denaturation of proteins

What is the molecular basis for this large heat capacity change of transfer for benzene. One can show that the Cp is also proportional to the surface area of the nonpolar solute. Figure 4.9.16 shows the smaller, more compact native state, with buried Phe (F) side chains denaturing to the more open D state with exposed F side chains. Since these are nonpolar, we can envision a "clathrate" or cage of ordered water around them. The heat capacity curves for both the native and denatured states are extrapolated into the region where T < Tm (even though there is very little denatured state in that region).

The "caged" water around the exposed F in the D state is low energy due to the "ice-like" H bond network. More heat would be absorbed (as the temperature is increased) to break up that cage compared to the same amount of heat applied to the N state. Hence, C_p D > C_p N. At room temperature, water molecules surrounding the nonpolar residue are low in energy (lots of H bonds) and low in entropy (thermal and positional, fewer available microstates). As the temperature is raised, water populates higher energy states (fewer H bonds) and higher entropy (thermal and positional, more available microstates. The increase in temperature causes "melting" of surrounding water structure in so far as energy and entropy are concerned. The two different energetic states of water provide an energy storage mechanism.

Here is another explanation that might help. Water molecules form "iceberg"-like cage of water around nonpolar molecules, which is often called a clathrate. The water is fully H-bonded (to itself, not to the nonpolar molecule) in a fashion analogous to ice but the geometry of the H bonds is nonideal. This structuring of water decreases its entropy. With increasing temperature, the structured water "melts" which produces the large heat capacity of a solution of a nonpolar molecule in water, just as the actual melting of ice showed a large heat capacity. This large heat capacity is the signature thermodynamic feature of the solution of a nonpolar molecule in water.

For the actual part of the graph in which denaturation occurs (N ↔ D), the following equations can be derived.

$$\begin{aligned} \mathrm{C_p} = \mathrm{C_d} &= \frac{d\mathrm{H}}{dt} \\ d\mathrm{H_d} &= \mathrm{C_d}\,\mathrm{dt} \\ \int_{\mathrm{T_1}}^{\mathrm{T_S}} \mathrm{dH_d} =< \Delta \mathrm{H_d} >= &\int_{\mathrm{T_1}}^{\mathrm{T_2}} \mathrm{C_d}\,\mathrm{dT} \end{aligned} \tag{4.9.2}$$

to help calculate the calorimetric enthalpy changes ($\Delta H_{cal} = \Delta H_d$) of denaturation from differential scanning calorimetry.

where $<\Delta H_d>$ is the average enthalpy change that occurs on denaturation which is represented by the blue and red shaded areas in the curve above (between T1 and T2). Compare this to the van 't Hoff enthalpy discussed in Laboratory Determination of ΔG^o for Protein Folding/Unfolding.

For the calorimetric determination of ΔC_d and ΔH_d, the values are temperature dependent, and no assumption about the temperature independence need be made, as in the case of the ΔH_{vhoff} plot.

H_2O Clathrate Cages - An Extension

Is there a more quantitative description of the ordered water than a cage? Sharp et al investigated the "structure" of the caged water around nonpolar and polar molecules in a theoretical analysis supported by molecular dynamics (Monte Carlo) simulations. The average bond angles and lengths of water-water H bonds in the first hydration sphere around a nonpolar molecule like benzene decreased, but increased for polar ones such as potassium ions. The average changes noted arose from two types of H bonds compared to bulk water, those that were shorter and more linear, and those that were longer and more bent. They stated that "nonpolar groups do not so much increase the ordering of water as decrease the disordering".

A recent review by Silverstein suggests that an immobile clathrate cage is **not** a good representation for water surrounding a hydrophobe. Although we like to envision molecular models that allow us to "explain" experimental thermodynamic findings, such models themselves should be subjected to rigorous experimentation. An alternative explanation hinges on water's small size (compared to other solvents), its tight packing and high density. Consider the density of water compared to more nonpolar liquid solvents as seen in the table below.

Table: Density of common solvents

Solvent	Volume (Emin) (A3) (Spartan)	Density (g/ml)
H_2O	19.17	1.00
methanol CH_3OH	40.66	0.791
ethanol CH_3CH_2OH	59.08	0.789
n-propanol $CH_3CH_2CH_2OH$	75.35	0.804
n-butanol $CH_3CH_2CH_2CH_2OH$	95.68	0.810
hexane	124.8	0.654

Let's consider the density of water surrounding an exposed nonpolar. If we envision the surrounding water as a clathrate, we might assume it is "ice" like. So what are the physical properties of ice and water that might give us insight into the water surrounding a nonpolar molecule?

Ice, of course, has a lower density than liquid water. This can't be simply explained by the number of H bonds since experimental evidence shows that ice has an average of 4 H bonds per water molecule compared with liquid water, with an average of 2.4. Experimental data also shows that to accommodate water molecules into a rigid network of interacting waters with tetrahedral symmetry, the H-O-H bond angle increases to 106 from 104.5. Liquid water molecules, with fewer packing constraints, can self-organize to maximize packing and hence macroscopic density. Studies suggest that ten water molecules solvate a buried methyl group and infrared studies show that four of these have significant barriers to rotational diffusion, suggesting they are effectively immobilized and hence "ice-like". Silverstein suggests then that the water surrounding a nonpolar group on the solution should be considered in a dynamic sense with some immobilized (as in ice) and the remaining more fluid-like.

Now let's review the benzene graph and apply it to protein unfolding one more time. The graph for $B_B \rightarrow B_W$ (Figure 4.9.12) shows a maximum in benzene insolubility. As the temperature is decreased from that maximum, benzene becomes more soluble in water. Alternatively, as temperatures rises to that temperature of maximal insolubility, the solubility of benzene decreases (just like nonpolar gases become increasingly insoluble with increasing temperature). If you extrapolate the ΔG curve in this range of decreasing temperature past the range shown on the graph, it would cross the X-axis and become <0, implying benzene would be favored to dissolve in water. Does the low temperature behavior of benzene/water interactions (becoming more soluble as the temperature is decreased from the maximum temperature for its insolubility) extend to and predict protein behavior at low temperature? Figure 4.9.17 shows the analogy between benzene solubility in water at cold temperatures and cold temperature denaturation of proteins.

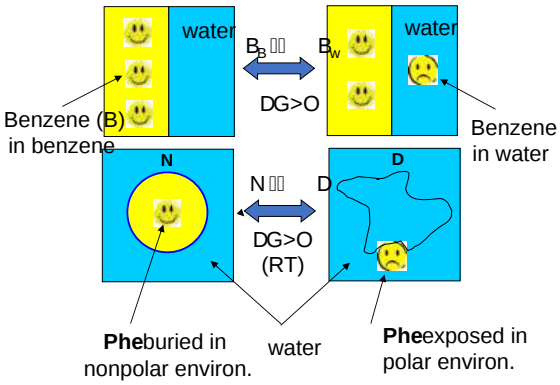


Figure 4.9.17:An analogy between benzene solubility in water at cold temperatures and cold temperature denaturation of proteins

In the figure, F stands for a Phe side chain, which can be buried, sequestered from water as it would be in the native state of the protein, and exposed to water, as it might be in the denatured state.) The answer is our question is yes, at low temperatures. The analogy to benzene being more soluble at low temperatures is the hydrophobic side chains in a protein becoming more likely to flip into water, denaturing the protein. The low temperature behavior would predict low temperature protein denaturation. This phenomena has been observed. Note that it doesn't require a change to a state when the nonpolar side chains prefer to be in water, just a change in that direction might be enough to tip the balance and lead to denaturation of the marginally stable protein.

Please note that we are attempting to extrapolate the thermodynamic parameters associated with benzene solubility in water to the denaturation of a protein, **NOT TO THE SOLUBILITY OF A PROTEIN IN WATER**!

What about high temperature? Proteins denature as the temperature increases in the range that the ΔG curve for benzene reaches a peak. If the buried hydrophobic residues behave like benzene, they would increasingly "prefer" to **not** flip out into water as the temperature rises to the maximum in the ΔG vs T curve. Hence the benzene predicts that the protein should become **more** stable. What then explains the observed denaturation of proteins at high temperatures? Another factor must account for it. What is it?

As the temperature is increased, more protein conformational states will become available and occupied. At low temperatures, let's say that there is only one native state available and occupied and (to pick a number), maybe 10 non-native states that are energetically accessible. At high temperatures, there is still only one native state, but perhaps 1000s of accessible nonnative states. More accurately, think of the protein existing in an ensemble of conformations. As the temperature increases, more non-native states can be populated, compared to at lower temperatures, leading to an entropic driving force favoring unfolding. Which way would this **chain conformational entropy** drive the protein at high temperature? Clearly, it would be driven to the most number of states - to the denatured state. Hence a modern definition of the hydrophobic effect can explain low-temperature denaturation, but not high-temperature denaturation.

4.9.8: Summary of studies from small molecules (N-methyacetamide and benzene)

It is clear that proteins are not all that stable, and many contributions of varying magnitudes must sum to give the proteins marginal stability under physiological conditions. The hydrophobic effect clearly plays a major role in protein stability. Also, since proteins are so highly packed compared to a loser-packed denatured state, collective Van der Waals interactions must also play a significant part. Remember these interactions, especially induced dipole-induced dipole interactions, are short range and become most significant under conditions of closest packing. Opposing folding is the chain conformational entropy just described. Since proteins are so marginally stable, even one unpaired buried ionic side chain, or 1-2 unpaired buried H bond donors and acceptors in the protein may be enough to "unravel" the native structure, leading to the denatured state.

4.9.9: Mutagenesis and Protein Stability

In the last decade, the contributions to the overall stability of a protein from the hydrophobic effect and H bonds have been studied using site-directed mutagenesis. In this technique, the DNA coding sequence for a given amino acid in a gene can be altered so that the new mutant protein differs from the normal (wild-type) protein by one amino acid. To probe the hydrophobic effect, for example, a buried hydrophobic amino acid like Ile could be changed to Gly which is much smaller, and offers a lower hydrophobic contribution to the stability of the native state. The result of this mutation might leave a "hole" in the protein (not unlike the vacant holes in crystal structures of salts). This "hole" might be diminished in size by subtle rearrangement of the protein structure in the vicinity. Certain amino acids would not be used as replacements in such studies. For instance, an Ile would not be replaced with a positively-charge Arg which would clearly destabilize the protein. The extent of destabilization in mutant proteins can be determined by calculating the ΔG^o for the native to denatured transition using urea as the denaturing agent as discussed in another section.

Previously, the following statistics were presented concerning the distribution of amino acids in the tertiary structure of a protein. New values are shown below in red, based on much more crystallographic data, as summarized in Pace's article.

- The side chain location varies with polarity. Nonpolar side chains, such as Val, Leu, Ile, Met, and Phe, are nearly always (83%) in the interior of the protein.
- Charged polar side chains are almost invariably on the surface of the protein. (54% - Asp, Glu, His, Arg, Lys are buried away from water, a bit startling!)
- Uncharged polar groups such as Ser, Thr, Asn, Gln, Tyr, and Trp are usually on the surface, but frequently in the interior. If they are inside, they are almost always H bonded (63% buried - Asn, Gln, Ser, Thr, Tyr, again startling) .

- Globular proteins are quite compact, with water excluded. The packing density (Vvdw/Vtot) is about 0.75, which is like the NaCl crystal and equals the closest packing density of 0.74. This compares to organic liquids, whose density is about 0.6-0.7.

Two articles by Pace suggeststhat Dills "influential review (from which much of the above derives) that concluded that hydrophobicity is the dominant force in protein folding" should be rethought. Using site-directed mutagenesis to change Asn (which can H bond through its side chain) to Ala (which can't) in a variety of proteins, he has shown that approximately 80 cal/mol/A^3 of stability is gained if a side chain (in this case Asn) can form buried H bonds to buried amide links of the protein backbone. Similar studies of mutants in which Leu is replaced with Ala, and Ile with Val, suggest that only 50 cal/mol/A^3 is gained from burying a hydrophobic -CH2- methylene group. Extending these results to protein folding suggests that protein stability is determined more by the formation of buried H bonds than by the hydrophobic effect!

The investigators measured ΔG^o for the N ↔ D transition (presumably by varying the urea concentration and extrapolating the ΔG^o for unfolding to 0 M urea (see: 4.12: Appendix - Laboratory Determination of the Thermodynamic Parameters for Protein Denaturation). For the reaction as written, $\Delta G^o_{unfolding} > 0$ at room temperature and 0 M urea. The mutant protein, since they are destabilized, would have a less positive value for $\Delta G^o_{unfolding}$ (They would also have a less negative value for folding since they are less stable). The difference in $\Delta G^o_{unfolding}$ between the wild type and mutant ($\Delta\Delta G$) is expressed as:

$\Delta\Delta G = \Delta G^o_{unfolding}$ wild-type - $\Delta G^o_{unfolding}$ mutant > 0

$\Delta\Delta G > 0$ since $\Delta G^o_{unfolding}$ wild-type > $\Delta G^o_{unfolding}$ mutant. The more positive the $\Delta\Delta G$, the more the mutant is destabilized in comparison to the wild type. The data for a series of mutants is shown below.

Analysis of Mutants: H Bonds in Protein Folding

mutation	ΔVol side chain ($Å^3$)	% buried	ΔΔG kcal/mol (kJ/mol) (destabilized)	ΔΔG cal/mol/$Å^3$ (J/mol/$Å^3$) (destabilized)
Asn to Ala	37.4	95	2.9 (12)	78 (326)
Leu to Ala	74.5	99	3.6 (15)	48 (200)
Ile to Val	25.8	100	1.3 (5.4)	50 (209)

What leads to protein stabilization/destabilization when Asn is changed to Ala?

One possible contributor to stability is the side chain conformational entropy. Since in the mutant, the Ala would find itself in a larger "hole" and have greater freedom for motion, it would have more conformational entropy that would stabilize the mutant over the wild type. Hence this effect can NOT explain the observed destabilization of the Asn to Ala mutant.

In the proteins he studied, only one of eight Asn to Ala mutations involved an Asn in a helix, so the average change could not be attributable to differences in helix propensities for the two amino acids.

In the mutants, assuming no rearrangement of the remaining side chains, there is an "unnecessary" and unoccupied 37 $Å^3$ cavity. The creation of this cavity is thermodynamically unfavorable (about 22 cal/mol/$Å^3$ obtained from values for hydrophobic mutations). If the same penalty were applied here, the Asn to Ala mutant would be destabilized by 0.8 kcal/mol (22 x 37.4), This is significantly less than the observed destabilization (2.9 kcal/mol, 12.1 kJ/mol), so this effect also could not account for the destabilization of the Asn to Ala mutants.

If there were compensatory changes to minimize the cavity size, this would only help to stabilize the protein and hence can not account for the observed destabilization.

Possible Explanation of Destabilization of Asn to Ala Mutants

possible reasons	explanation	effect on mutant	support observed destab. of mutant?
residue conformational entropy	Ala in a bigger hole: more freedom of motion; favored entropically	stabilize mutant	NO
free energy change excess cavity formation	energy penalty to make an unoccupied cavity approx. 0. 8 kcal/mol (3.3 kJ/mol)	destabilize	yes but of insufficient size compared to the observed effect (2.9 kcal/mol, 12 kJ/mol)

free energy change protein conformational changes	rearrange protein to fill the cavity	stabilize mutant	NO

Hence these alternative sources to explain the destabilization of the mutant can't account for the data and we're left with the explanation that the stability of the native protein over the mutant is accounted for by burying the H bond donor and acceptors of the amide group and associated changes in van der Waals interactions.

Pace argues that burying the amide group of Asn is similar to burying the peptide bond of the main chain. Their sizes are very comparable. Free amide groups can form four H bonds, but peptide (amide) groups can only form three. Even if the value of 78 for the $\Delta\Delta G$ (cal/mol/$Å^3$) is adjusted for this, the new value of 62 is still larger than that for burying a methylene group. Analysis of 108 folded proteins has shown that hydrophobic groups contribute 118,200 $Å^3$ of buried volume, compared to 92,000 $Å^3$ for peptide groups. Multiplying these figures by 78 and 49 (from the above table) suggests that overall, burying peptide groups contributes more to protein stability than burying hydrophobic groups.

Would electrostatic interactions of the buried peptide group with the surrounding environment destabilize a protein? Pace argues that this would be more than compensated for by favorable van der Waal's interactions (short range) at the buried site. This can be illustrated by comparing the ΔG transfer of an amide from water to the vapor (11.2 kcal/mol, 47 kJ/mol) compared from water to cyclohexane (7.6 kcal/mol, 32 kJ/mol). Transfer to the vapor is more unfavored (due to the desolvation required when it moves to the gas phase) than to cyclohexane, even though a cavity must be created in the cyclohexane (a process which would be unfavored entropically). Transfer to octanol is even more favored (1.4 kcal/mol, 5.9 kJ/mol) but all these values are still positive (disfavored). Similar experiments with the transfer of a methylene group (-CH2-) are negative, given the hydrophobic effect and the collective close packing van der Waal's interactions possible. These suggest that van der Waals interactions formed on burying an amide in any solvent are stabilizing. Now consider the packing density of atoms for various substances:

Packing Densities

substance	packing density
water	0.36
cyclohexane	0.44
closest packed spheres	0.71
protein interiors	0.75

From this table, it should be apparent that collective van der Waals interactions (short range) will be more stabilizing in the interior of the protein compared to the same groups in bulk water (or in the denatured state). Carbonyl groups are more polarizable than methylene groups, which should contribute to van der Waals interactions.

One other addition. It has been noted that Gly peptides are not very soluble in water. The backbone, even with the polar peptide bonds appears to be solvophobic. If the backbone of any polymer can't interact well with the solvent - i.e. the solvent is "poor" - then the backbone interacts with itself, which drives collapse. If the backbone interacts well with a "good" solvent, it won't collapse as readily.

4.9.10: Protein Stability in Thermophilic Organisms

What kinds of modifications are made to the sequence of a protein as the temperatures at which the organism thrives increase? A recent study by Szilagyi and Zavodszky (Structure, 8, pg 493, 2000) studied 93 structures of 25 proteins, 29 from organisms that live at elevated temperatures (thermophiles, >50^oC for optimal growth) and 64 at nominal temperatures (mesophiles). Here are their results:

- the number of H-bonds and secondary structure elements do not correlate with temperature, but the number of salt bridges do.
- in hyperthermophiiles (>80^oC for optimal growth organisms) that thrive at very high temperatures (100^oC), few internal cavities were found.
- in those that thrive at intermediate high temperatures (45-80^oC), the surface had more polar residues.
- in general, there was an increase in weaker ion pairs (increased distance between the charged side chains) in the hot group, but increases in strong and weak ion-ion bonds in the very hot group.

Kashefi and Lovley recently reported the identification of a bacteria obtained from a hydrothermal vent in the northeastern Pacific ocean. In a laboratory setting, the strain grew in water at a temperature of 121^{o}C under high pressure. These are the same conditions used in autoclaves to produce sterile samples. Cell doubling took place under these conditions in 24 hours. The authors suggest that this strain would be useful to determine molecules and their properties necessary for such high temperature growth.

Using a computational program called Rossetta Design (PNAS, 97, 10383 (2003)), Korkegian et al determined mutations in buried side chains of the homodimer cytosine deaminase. Buried residues are presumably are important in the stability of a protein and are targets for mutagenesis experiments that would increase the melting temperature (Tm) of the protein. In the program, a target sequence was "threaded" onto the sequence of the template protein (the wild-type protein) and changes were made to side chains in the random sequence. Energies were calculated and those changes resulting in lower energies were saved. Target residues (88) within 4 angstroms of the active site and the dimer interface were fixed to those in the wild-type template in order to minimize alterations in the catalytic activity of the enzyme, cytosine deaminase, that they chose to study. Remember, the goal of the study was not to increase the catalytic activity of the enzyme, but rather to increase its themostability. The rest (65) were changed and energies calculated. 49% of the amino acids subjected to random change produced no change in amino acid compared to the template (wild-type) side chain. 16 changes on the surface were ignored. Two sets of changes were observed, one involving amino acids packed between an alpha helix and beta strands, and the other set between two alpha-helices. These later mutants, when prepared in the lab using recombinant DNA technology, were soluble at high protein concentrations, and could be studied. Three different mutants (A23L, I140L, V108I) were made which increased the T_M by about 2 degrees. However, a triple mutation had T_M values 10 degrees higher than the wild-type protein and a 30-fold longer half life ($t_{1/2}$) at50 degrees C. When the triple mutant was introduced into bacteria, the bacteria grew better at higher temperatures. Crystal structures of both the wild-type and triple mutants show essentially an identical fold, with about 70 Å^2 of additional surface area buried in the mutant protein.

Beeby et al. analyzed sequence and structural data from P. aerophilum (archea) and Thermus thermophilus (thermophilic bacteria) and found that disulfide bonds stabilized proteins from these species. Cytoplasmic protein from eukaryotes don't have disulfides due to the presence of reducing agents (such as glutathione) in the cell. In those thermophiles with disulfides in proteins, a novel protein, protein disulfide oxidoreductase, was found, which catalyzes the formation of sulfide bonds. Finally, Berezovsky and Shakhnovich have also analyzed proteins from hyperthermophilic archaea and bacteria and compared them to analogous proteins from mesophilic bacteria. They found two types of stabilizations of hyperthermophilic proteins, depending on the evolutionary history of the organism. Proteins from cells that originally evolved in high temperature conditions (Archaea) were very compact (maximizing van der Waals interactions, had a high number of contacts per residue, and a high percentage of hydrophobic residues), but did not use specific structural stabilizing interactions (like electrostatic in salt bridges). In contrast, proteins from cells the originally evolved under mesophilic conditions, but later adapted to hyperthermophilic conditions had proteins that evolved specific sequences features that stabilized electrostatic interactions (more charged residues, salt bridges. .

4.10: Protein Aggregates - Amyloids, Prions and Intracellular Granules

4.10.1: Introduction to Protein Aggregates

We have studied different types of protein aggregation, including aggregation of the native state (to form dimers, trimers, multimers, and filaments). We've also studied how misfolded proteins can aggregate and how a whole family of molecular chaperones help newly synthesized and misfolded proteins fold correctly. But what happens if a protein can fold two reasonably stable or metastable structures of starkly different conformation? A class of proteins called **prions** or **amyloid proteins** has this characteristic. their "alternative" conformation structures can bind to the "normal" structure and cause it to flip to the alternative conformation. This then can seed a continuation of the process which ends in the formation of large aggregates that are fibrillar in structure. This process leads to aberrant cell function and when it occurs in neurons can lead to a variety of brain diseases such as Alzheimer's Disease.

Protein aggregates complicate the lives of people who study protein folding in vitro and who try to express human proteins in prokaryotes like E. Coli in vivo, which often end up in large protein aggregates called inclusion bodies. Instead of viewing these aggregates as unwanted "junk", some study them avidly. It turns out that these aggregates are not as non-specific as earlier believed. In addition, an understanding of how and when they form will give us clues into the etiology and treatment of some of the most debilitating and feared diseases.

4.10.2: Specificity of Aggregate Formation

In the early 1970s it was shown that chymotrypsinogen could not be folded in vitro without aggregates forming. An intermediate was presumed to have formed that if present in high concentration would aggregate irreversibly instead of folding to the native state. Refolding of tryptophanase showed that it aggregated only with itself, suggesting specificity. In the 1980s, a single amino acid folding mutant was found in a viral protein. Both the normal and mutant viral proteins unfolded at high temperature, but only the mutant would aggregate at high temperatures, suggesting that aggregation could be programmed into or out of a gene. Also, a single amino acid change in bovine growth hormone completely prevented aggregation without affecting correct folding.

This knowledge of protein folding and aggregation was soon turned toward understanding several diseases in which protein aggregates were observed which either initiated or were associated with diseases. These protein aggregates were termed "**amyloid** deposits" and seemed to be associated and, perhaps causative of several neurodegenerative diseases. The name **amyloid** was first used by a German pathologist, Rudolf Virchow, who in 1853 described waxy tissue deposits associated with eosinophils (a type of immune cell). These deposits seemed to resemble starch (made of amylose and amylopectin) so he termed them amyloid. All known amyloid deposits are, however, composed of protein, not starch.

It now appears that these diseases are likely caused by improper protein folding and subsequent aggregation. Except in certain rare inherited diseases, the amyloid deposits are composed of normal wild-type proteins (not mutants), which seem to undergo conformational changes to form monomers, which catalyze the formation of more altered normal monomers into the altered form, which polymerize into fibrils. Sometimes, in inherited conditions, or when mutations appear in a specific protein, the amyloid protein deposits consist of the mutant protein. The proteins in these deposited fibers are composed predominantly of β sheets which are perpendicular to the fiber axis. In some cases, the monomeric "normal conformation" of the protein has little beta sheet structure.

Figure 4.10.1 shows a simplified model of how a normal protein with a "normal" conformation enriched in this case for illustration in alpha-helices can form fibrils of abnormal monomers which are highly enriched in beta sheets. These can self-associate to form large insoluble protein "amyloid" fibers. Note the green arrows representing beta-strands.

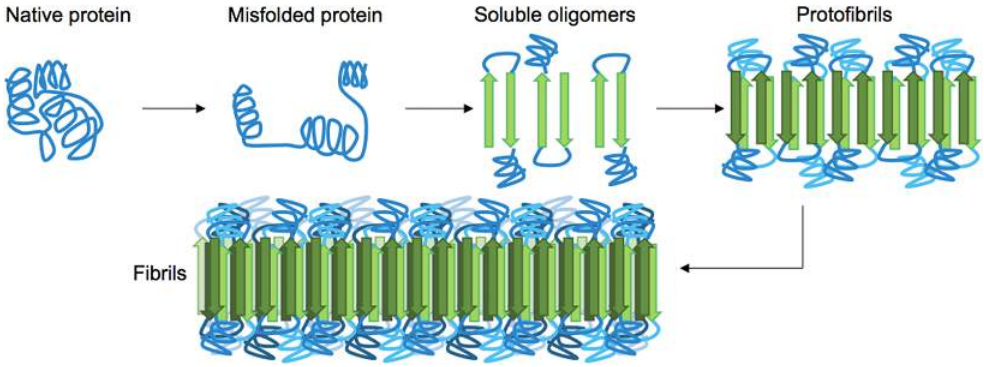


Figure 4.10.1: Simplified pathway toward amyloid fibril formation.Kaminski et al. *Neurophotonics* 3(4), 041807 (29 June 2016). https://doi.org/10.1117/1.NPh.3.4.041807 . Creative Commons Attribution 3.0 Unported License.

In the misfolded state, proteins have an increased propensity to oligomerize through the association of their metastable beta-sheet domains. These can convert into more stable beta-sheet states and the ensuing oligomers act as the nuclei for the subsequent elongation reaction, which leads to the formation of so-called protofibrils. The final amyloid fibril usually consists of a number of intertwined protofibrils.

4.10.3: Diseases of Protein Aggregates

We'll now describe a series of diseases that are caused by or highly associated with fibril formation from normal soluble proteins. For each one, we will present the best available structure of the amyloid fiber obtained mostly through cryo-EM. What's amazing is that at first glance all of the amyloid fiber structures have an astonishingly similar structural appearance. We present them not to be redundant but to illustrate how natural processes can render from a great diversity of protein structures a common structural and often lethal outcome. That the amyloid fibers are so structurally similar suggests that a common therapy to prevent their formation may be developed.

4.10.3.1: Familial amyloidotic polyneuropathy (FAP)

This affects 1/10,00 to 1/100,000 people. The monomer protein involved is called **transthyretin** (147 amino acids, MW 15,887), which normally exists in the blood as a homotetramer (a dimer of dimers). Figure 4.10.2 shows the structure of the dimer (6fxu). Note each monomer contains mostly beta structure. The protein binds L-thyroxine and around 40% of blood plasma transthyretin is bound to retinol-binding protein.

Figure 4.10.2: Structure of the transthyretin dimer (amino acids 21-147) (6fxu)

In mildly acid conditions in vitro, the equilibrium between tetramer and monomer is shifted to monomer, which can aggregate into fibrils. Monomer aggregation could be promoted by a possible transition to a molten globule (discussed previously with lactalbumin) like state. This has secondary structure but loosely-packed tertiary structure with more exposed hydrophobic groups. If the concentration is high enough the molten globules aggregate. In people with the disease, mutations in the protein destabilize the tetramer, pushing the equilibrium to the monomer, which presumably increases molten globule formation and aggregation. Specifically, Val30Met and Leu55Pro mutations promote the dissociation of the tetramer and the formation of aggregates. Conversely, Thr119Met inhibits tetramer dissociation. The aggregates deposit in the heart, lungs, kidney, etc, leading to death. Figure 4.10.3 shows a pictorial representation of how the transthyretin dimer could be stabilized to form monomers or dimers leading to fibril formation.

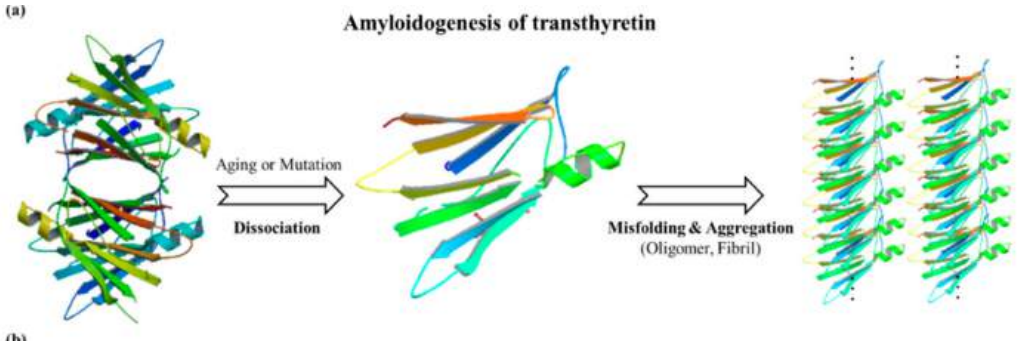


Figure 4.10.3: Conversion of the transthyretin dimer to fibril formation. Park, G.Y.; Jamerlan, A.; Shim, K.H.; An, S.S.A. Diagnostic and Treatment Approaches Involving Transthyretin in Amyloidogenic Diseases. *Int. J. Mol. Sci.* **2019**, *20*, 2982. https://doi.org/10.3390/ijms20122982. Creative Commons Attribution (CC BY) license (**http://creativecommons.org/licenses/by/4.0/**)

Figure 4.10.4 shows the intramolecular hydrogen bonds bonds within monomers (-----) and intermolecular hydrogen bonds between monomers (-----) in the transthyretin dimer. We present this figure to remind readers that in all of the complex amyloid fibrillar structures presented below, the bulk of hydrogen bonds are "intermolecular" between adjacent monomers in their extended states (phi/psi angles consistent with beta-structure) within the fibrillar structure.

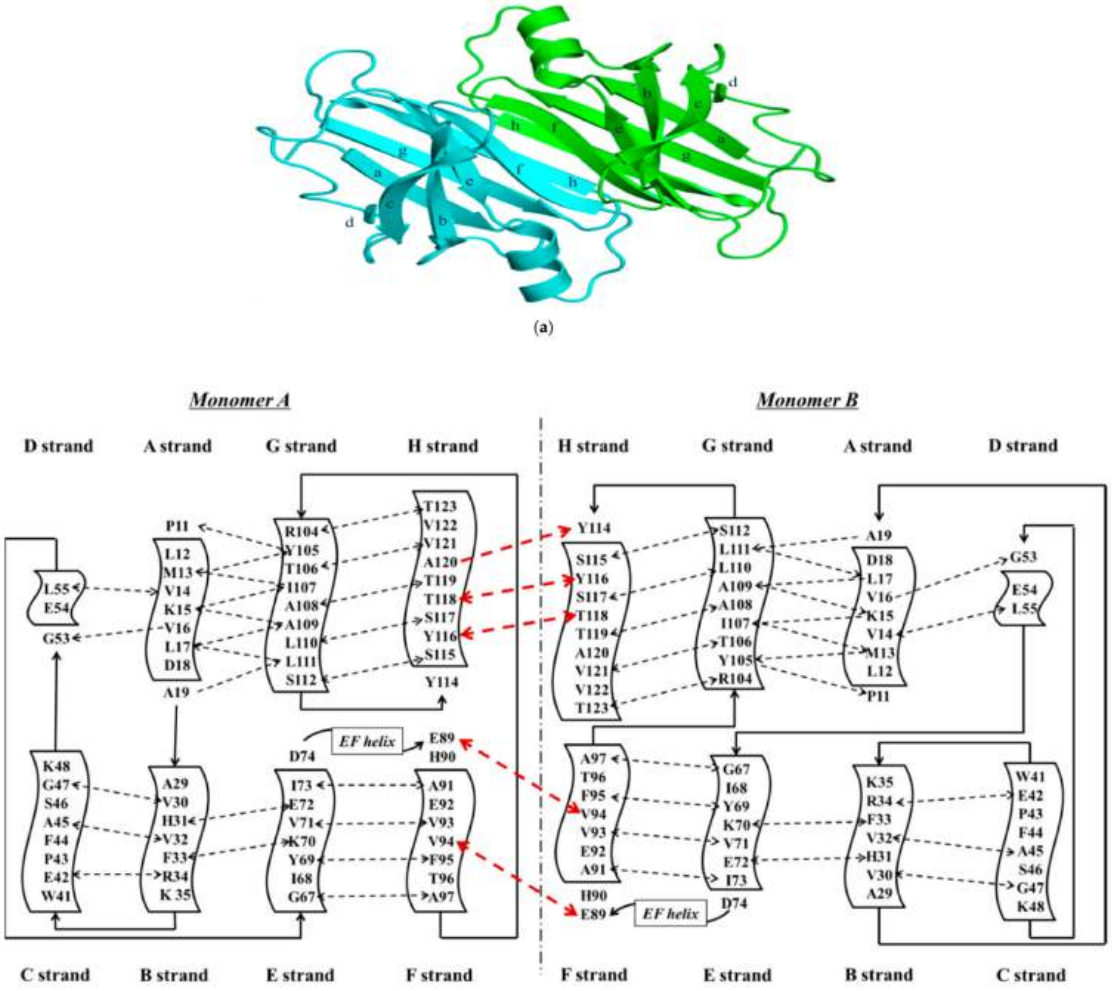


Figure 4.10.4: (**a**) 3-Dimensional structure of ATTRwt (PDBe ID code; 1BMZ [18]) in dimeric form; Hydrogen bonds within monomers (-----) and **between** monomers in the dimer (-----) in the the **transthyretin dimer.** Park et al., ibid

Figure 4.10.5 shows an interactive iCn3D model of Cryo-EM structure of a transthyretin-derived amyloid fibril from a patient with hereditary ATTR amyloidosis (6SDZ). Each separate monomer is shown in a different color. The static image below shows arrows indicating beta strands, which form hydrogen bonds from the main chain amide Hs and carbonyls Os to another main chain atoms on an adjacent monomer.

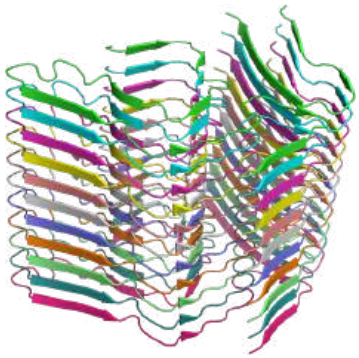


Figure 4.10.5: Click the image for a popup or use this external link: Cryo-EM structure of a transthyretin-derived amyloid fibril from a patient with hereditary ATTR amyloidosis (6SDZ). (Copyright; author via source) Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...wMtaWWJ5KJkQ99

Note the beautiful but unfortunately deadly array of adjacent extended monomeric chains (each colored differently) that form hydrogen bonds with adjacent extend monomers to stabilize the fibrillar structure.

4.10.3.2: Light Chain Amyloidosis - AL amyloidosis (amyloidosis from the light chain)

The light chain (MW approx 25,000) is a normal component of circulating immunoglobulin antibody (protein) molecules. Each contains two light and two heavy chains. We will discuss the structure of antibodies in detail in Chapter 5.5. Needless to say, antibodies are very diverse molecules. Antibodies can be generated by the immune system to recognize almost any foreign molecules. The light chains of antibodies hence are incredibly diverse and variable, although they all have two **immunoglobulin domains**. Each domain has about 110 amino acids containing two layers of β-sheets each with 3-5 antiparallel β-strands with a disulfide bond connecting the two layers so they start with significant beta structure in the monomeric form. The large diversity in light chains is generated in part by the recombination of gene fragments to produce individual light chains. Some variants of the light chains (λ1, λ2, λ3, λ6, and κ1) are associated with AL amyloidosis, a potentially fatal disease. Mutants in the light chain can cause a destabilization of the native state to a state similar to a molten globule, which then conformationally converts to a structure that aggregates into amyloid fibers. These can deposit in various tissues.

The pathway that a simple 25K monomer takes to produce such a complex but regular structure as shown above is must start with simple dimer formation between two monomers. Figure 4.10.6 shows an interactive iCn3D model of the normal conformation of a λ6a light chain **dimer** (6mg4). The λ6a light chain variant is more prone to aggregate to form fibrils.

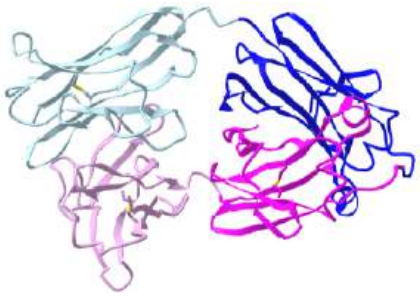
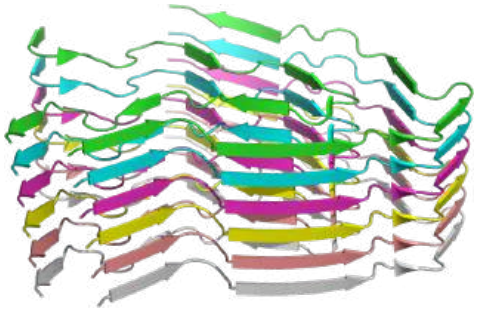


Figure 4.10.6 Full-length human lambda-6A light chain dimer showing IgG fold domains (6mg4). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...MNsiddm4RAdKn8

One light chain in the dimer is shown in blue, with one IgG fold domain in light blue and the other in dark blue. The other light in the dimer is shown in shades of magenta to show the two domains. Normally the light chains don't form dimers but rather associate with heavy chains to form full IgG antibodies. If not associated with a heavy chain, free light chains will form dimers, which can alter conformation and form amyloid fibrils.

Figure 4.10.7 shows an interactive iCn3D model of the AL amyloid fibril from a lambda 3 light chain in conformation A (6Z1O). Each separate monomer is again shown in a different color.



NCBI iCn3D Figure 4.10.7: AL amyloid fibril from a lambda 3 light chain in conformation A (6Z1O). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...pSu2Nm714gvjh7

The Figure 4.10.8 shows a comparison of the secondary structures of the native conformation of the light chain and those found in the two different fibrillar forms (A and B) of the amyloid fibrils (panel a). Remember that each light chain has two IgG domains, each having two sets of β-sheets containing 3-5 antiparallel β-strands).

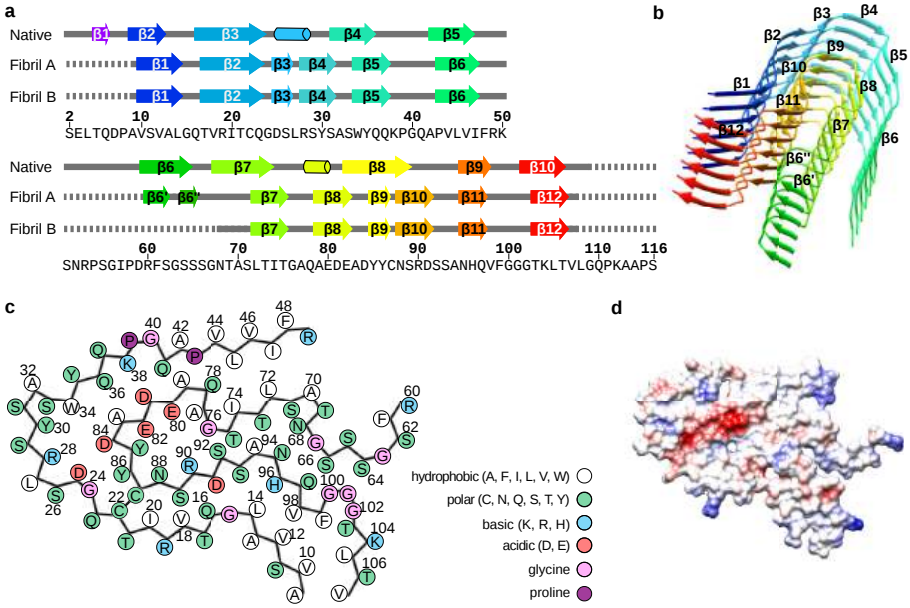


Figure 4.10.8: Secondary structure and folding of the fibril proteins. Lynn Radamaker et al. *Nat Commun* **12,** 875 (2021). https://doi.org/10.1038/s41467-021-21126-2 . Creative Commons Attribution 4.0 International License

Panel b shows the alignment of six fibrillar light chain proteins in conformation A. Panel c shows the trace of the backbones and the local environment of the side chains in a **single** light chain from the fibrillar A conformations. It shows cross-β-sheets interactions with parallel, hydrogen bonds across strands. Note the disulfide at connecting Cys 22 and Cys 87 on adjacent stands. The circles show surface side chains. Note that they are enriched in green (polar) and blue/red (basic/acidic) with some hydrophobic patches (example V10-L14). Some are pairs such as D24 and R28. Panel D shows the electrostatic surface of a single light chain from the fibrillar A conformation with red indicating negative and blue positive.

4.10.3.3: Alzheimers' Disease (AD)

This disease accounts for 60-80% of dementia cases. Brains in Alzheimer's patients have amyloid deposits of the **amyloid beta (Aβ) protein** as well as aggregates of a protein called **tau.** The origin of this disease is not fully resolved. Some believe aggregates of the amyloid beta protein cause the disease. Others suspect the role of tau in forming tau bundles. Still others suggest an infectious disease agent (more on that later). Irrespective of the fundamental cause, aggregates of the amyloid beta protein are neurotoxic and at minimum correlative if not causative of the disease.

The amyloid aggregates in Alzheimer's start with a change in a monomeric protein normally found in the membrane of neurons. The protein, called **β-amyloid precursor protein** (**BAPP** or simply **APP**), is a transmembrane protein. A slightly truncated, soluble form is also found secreted from cells and is found in the extracellular fluid (such as cerebrospinal fluid and blood). The normal function of these APP proteins is not yet clear. An endoprotease cleaves a small 40-42 amino acid fragment from this protein named the **amyloid beta (Aβ) protein**. Figure 4.10.9 shows the normal processing of the amyloid precursor protein APP (left) and the abnormal, amyloidogenic form (right).

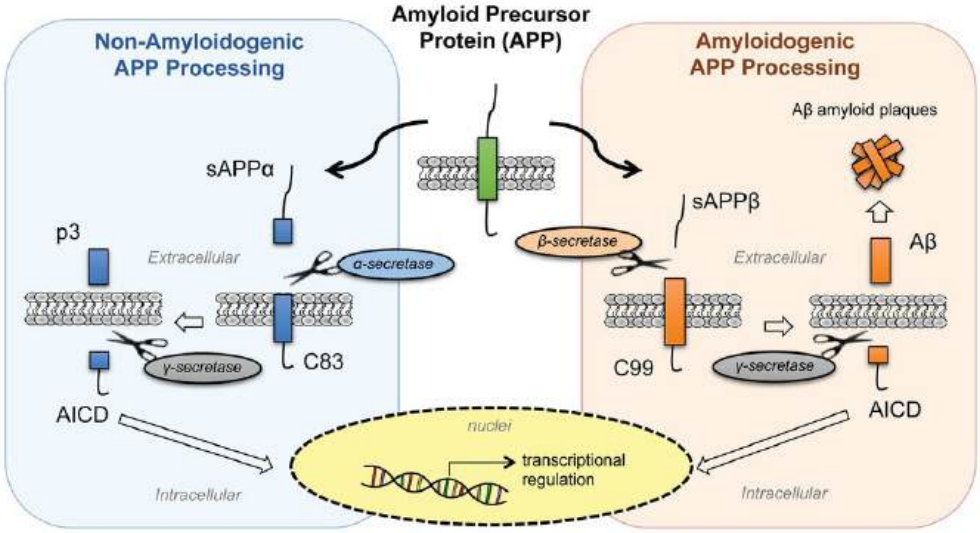


Figure 4.10.9: Processing of amyloid precursor protein (APP) in normal and abnormal conditions. Ristori et al. Front. Physiol., 27 August 2020 | https://doi.org/10.3389/fphys.2020.01056. Creative Commons Attribution License (CC BY).

In a "normal" processing pathway, the proteases α- and γ-secretase release two variant peptides, soluble amyloid precursor protein cleaved by α-secretase (sAPPα) and p3 fragments, into the extracellular environment. In the amyloidogenic processing pathway, β- and γ-secretases release soluble amyloid precursor protein cleaved by β-secretase (sAPPβ) and β-amyloid (Aβ) peptides. Both pathways release the same intracellular domain, AICD, which moves to the nucleus and acts as a transcription factor to regulate gene expression. The β-amyloid (Aβ) peptides aggregate to form the fibrillar aggregate plaque.

It is the amyloid beta (Aβ) protein or a mutant form of it that aggregates to form beta-sheet containing fibrils in Alzheimer's disease. The NMR solution structure of the monomer amyloid beta-peptide (1-42) is shown in Figure 4.10.10 Note the absence of any beta structure.

Figure 4.10.10: Solution structure of the amyloid beta-peptide (1-42) 6SZF.

Several mutations in different proteins have been linked to Alzheimer's, but they all seem to increase the production or deposition or both of the amyloid beta protein. These deposited plaques are extracellular, and have been shown to cause neuronal damage. They are found in areas of the brain required for memory and cognition. The APP gene is found on chromosome 21, the same chromosome which is present in an extra copy (trisomy 21) in Downs Syndrome, whose symptoms include presenile dementia and amyloid plaques. Aggregate formation appears to be driven by increased expression of APP and hence amyloid beta protein. In addition, some mutants may serve to destabilize the amyloid beta protein, increasing its aggregation.

Figure 4.10.11 shows an interactive iCn3D model of the prevalent amyloid-beta fibril structure from Alzheimer's disease brain tissue (6W0O).

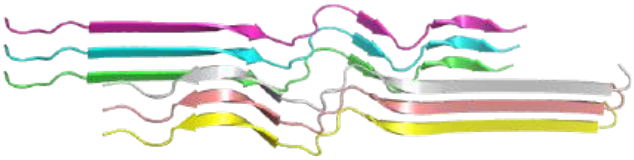


Figure 4.10.11: Amyloid-beta(1-40) fibril derived from Alzheimer's disease cortical tissue (6W0O). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...NZCX9sAhnFPiB6

The **tau** protein (758 amino acids, MW 78,928), which is much larger than the other proteins we discuss in this chapter, has also been implicated as a cause or factor in Alzheimer's Disease. It facilitates microtubule assembly and stability (see Chapter 5) along with other functions in neurons. Since its C-terminus binds microtubules in axons and the N-terminus binds the plasma membrane, it might link both. The cytoskeleton of neurons is disrupted in the neurons in Alzheimers' and in other neurodegenerative diseases patients. When tau tangles are involved, these diseases can also be termed **tauopathies**. One tauopathy is **chronic traumatic encephalopathy (CTE)** caused by repetitive head impacts (from contact sports or physical abuse) or concussions arising from explosions in combat. No full-length structure of tau has been determined. The structures of a predicted model of solution phase monomeric tau and tau fibers from the brain of patients with neurodegenerative diseases (Alzheimer's, CTE and Corticobasal Degeneration - CBD) are shown in Figure 4.10.12

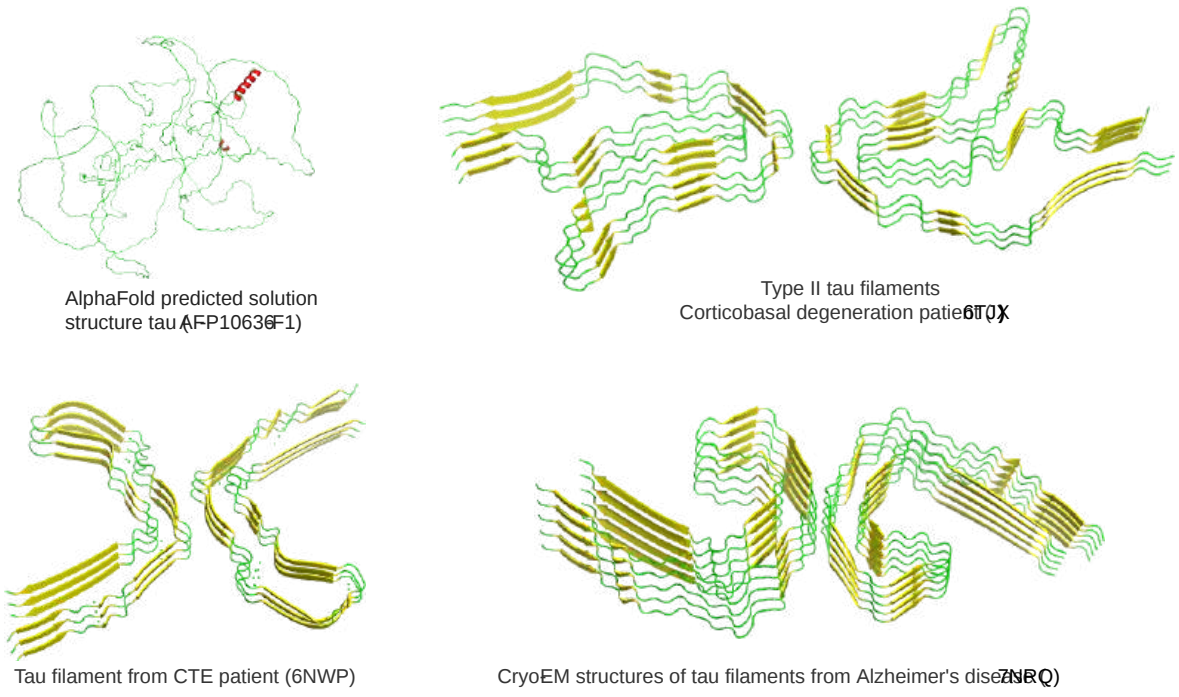


Figure 4.10.12: Structure of solution phase tau and tau filaments from various tauopathies.

The predicted structure of the monmeric protein (using AlphaFold) is almost completely devoid of secondary structure. The structures of tau from the other tauopathies are similar but clearly distinct. In CBD tau fibers there are 4 microtubule-binding repeats (4R). Picks Disease, another tauopathy, has three repeats (3R) while taus in AD and CTE are 3R or 4R.

The fibril cores of tau in both CBD (Lys 274-Glu380, which contains the end of R1 and R2-R4) and Alzheimer's Disease (CD) contain around 13% glycine residue. These allow the main chain flexibility and intersheet packing to allow the large conformation changes necessary to adopt beta structures and fibril formation. Repeats of PGGG motif allow sharp turns or extended chains. Valine (around 10%) and isoleucine, leucine, and phenylalanine facilitate inter-sheet packing through induced dipole-induced dipole interactions as well as through the hydrophobic effect. Certain tau fibrils in CBD contain hydrophilic pockets that bind molecules yet to be elucidated that might seed the nucleation of fibrils. The cavity has three lysines and a histidine so the molecule with the pocket is probably linked to histidine might be a glycan or an ADP-ribose. In addition, recent evidence shows that

different combinations of post-translational modification by ubiquitinylation and acetylation of lysines 311, 317, 321, 343, and 353 might lead to different tau fibril structures. CTE tau protein in tangles appears to be hyperphosphorylated.

Figure 4.10.13 shows an interactive iCn3D model of a paired helical **tau** filament from Alzheimer's Disease human brain tissue (6VHL).

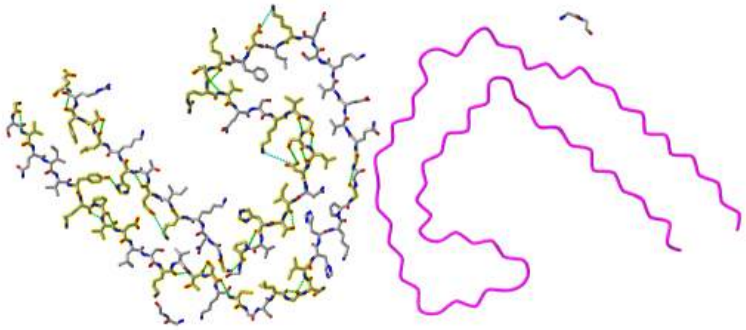


Figure 4.10.13 Paired helical filament of Tau (6VHL) (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...NHsZdhDJNentg9

Zoom into the interactive images to see the H bonds within one of the filaments. Note that the hydrogen bonds (green dotted lines) in the left filament are between side chains and not between backbone amide Hs and carbonyl Os. These are pointing above and below the plane of the backbone where they could interact with other of the chains above and below to create the multi-chain fibers seen in all of the examples above.

Figure 4.10.14 shows an interactive iCn3D model of a singlet Tau fibril obtained from corticobasal degenerated human brain tissue (6VHA).

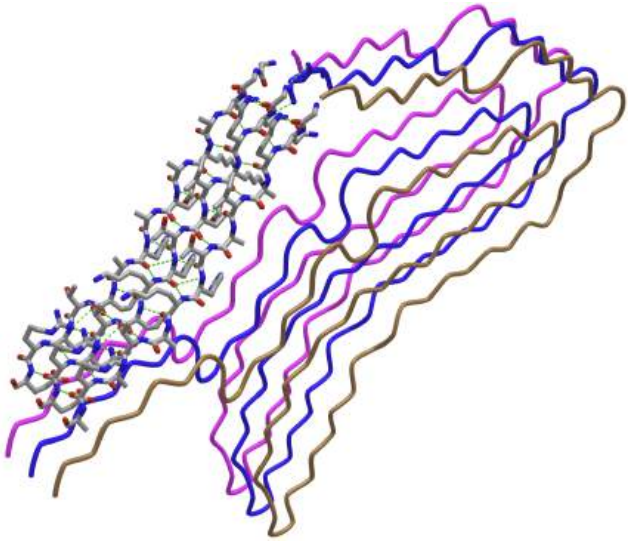


Figure 4.10.14 Singlet Tau fibril from corticobasal degeneration human brain tissue (6VHA). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...Bbxr2Z1caGqnG8

The backbone of the C-terminal amino acid (367-380) of each of the three separate chains of the fibril is shown in CPK colors and hydrogen bonds between the strands that form the parallel beta strands are shown as green dashes.

Progress has been incredibly slow on ways to treat Alzheimer's. Most methods focus on reducing amyloid beta production and aggregation by finding small molecules that inhibit steps in its production, including secretase cleavage, and the resulting conformational steps necessary to produce amyloid fibers. But what if the amyloid aggregates are secondary to the primary cause? What if the overproduction of the amyloid beta protein was the brain's response to defend against the cause?

4.10.3.4: Lewy Bodies and Parkinson Disease

α-Synuclein (140 amino acids, MW 14,460) is expressed in the brain and in presynaptic terminals in the central nervous system and is involved in the regulation of neurotransmitter release and in the synaptic vesicles that hold them. Its aggregation is a cause or consequence of Parkinson's Disease and Lewy Body Dementia. It's found in the cytoplasm and the nucleus and is secreted as well. Figure 4.10.15 shows an NMR solution structure (left) and AlphaFold-predicted structure (right) for this protein which in solutions is so disordered that no full crystal structure has been determined.

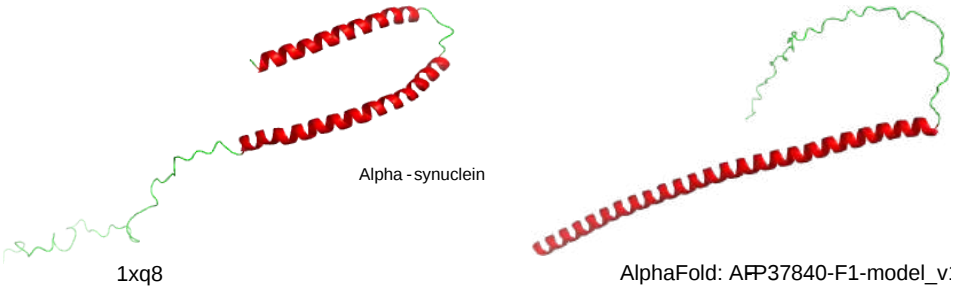


Figure 4.10.15: Alpha-synuclein structure (left, NMR solution structure, right AlphaFold predicted structure)

Figure 4.10.16 shows an interactive iCn3D model of an amyloid fibril structure of alpha-synuclein determined by cryo-electron microscopy (6A6B). Two protofilaments with clearly Greek key topologies are shown.

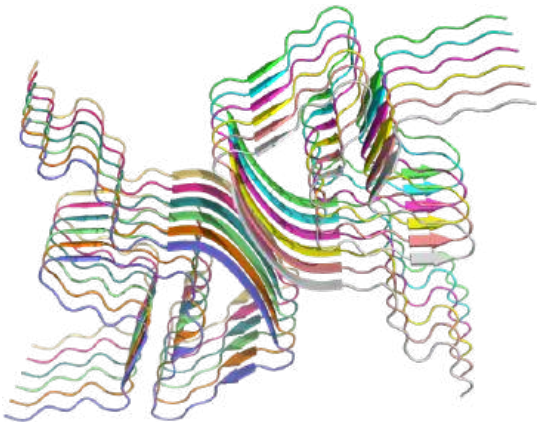


Figure 4.10.16: Amyloid fibril structure of alpha-synuclein determined by cryo-electron microscopy (6A6B). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...SbQq7fhFpQ9K4A

4.10.3.5: Transmissible spongiform encephalopathies (TSEs) - Prion Diseases

Prion diseases are another set of brain diseases resulting from aggregates of monomers of the prion protein (PrP^c). As with the examples above the aggregates form amyloid beta fibrils. The prion diseases include scrapie in sheep, bovine spongiform encephalopathy (mad cow disease), and in humans Creutzfeld-Jacob Disease (CJD), Fatal Familial Insomnia (FFI), Gerstman-Straussler-Scheinker Syndrome, and Kuru (associated with cannibalism). In these fatal diseases, the brain, on autopsy, resembles a sponge with holes (hence the name spongiform). In contrast to the diseases described above, these diseases can be transmitted from one animal to another, but typically not between species. (However, consider the controversy with mad cow disease.) Also, the infectious agent can self-replicate in vivo. The logical conclusion is that a virus (slow-acting) is the causative agent. However, the infectious agent survives radiation, heat, chemical agents, and enzymes designed to kill viruses and their associated nucleic acids. Mathematical analyses suggested that the infectious agent in such diseases could be nothing more than a protein. Stanley B. Prusiner in the 80's isolated just such a protein which he named a **prion**, for **proteinaceous infectious agent**. In October 1997 he was awarded the Nobel Prize in Medicine.

The normal monomeric prion protein, PrP^c (253 amino acids, MW 27,661), is highly conserved in mammals, and is widely expressed in embryogenesis. Expression is highest in the central nervous system. The normal function of the protein is still unclear. It is a physiological substrate to a particular membrane receptor (the Gpr 126 G protein-coupled receptor). Knocking out the gene

shows that the normal protein is involved in synapse structure/function, myelination of neurons and circadian rhythms probably by acting as a transcription factor. It also helps regulate Cu^{2+} and Zn^{2+} levels in the central nervous system. The protein is cleaved and a 209 amino acid fragment is bound to the extracellular side of the neuron membrane-anchored by attachment of a lipid (GPI) anchor.

The protein N-terminal residues (23-124) are flexibleand are followed by residues 125-231 which are mostly alpha helical. There is a disulfide between Cys179 and Cys214. The PrP^c (without the PI link) is water soluble, a monomer, protease-sensitive, and consists of around 45% alpha helix and 3% beta sheet. No full-length crystal structure of the protein has been determined given its highly disordered structure. The solution structure of residues 125-231 has been determined by NMR and the structure of the full protein has been modeled with Alpha Fold. These structures are shown in Figure 4.10.17.

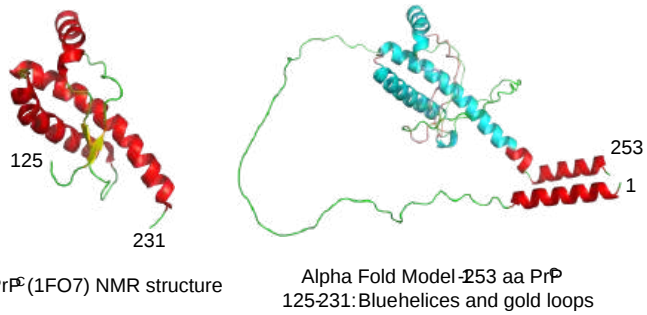
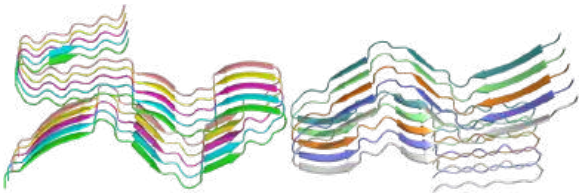


Figure 4.10.17: Structure of the soluble form of PrP^C (left NMR, 1FO7, NMR structure, right Alpha Fold predicted model)

The blue helices and gold loops in the computer model consists of the same amino acids (125-231) as the NMR solution model in the left hand side of Figure 4.10.2

The problem in transmissible spongiform encephalopathies (TSE's) is that amyloid-like protein aggregates form, which are neurotoxic. The protein found in the plaques (in cases other than those that are inherited) has the same primary sequence as the PrPc but a different secondary and presumably tertiary structure. The protein found in the plaques, called the PrP^{sc} (the scrapie form of the normal protein) is insoluble in aqueous solution, protease-resistant, and has a high beta sheet content (43%) and lower alpha helix content (30%) than the normal version of the protein PrPc.

Figure 4.10.18shows an interactive iCn3D model of the cryo-EM structure of an amyloid fibril formed by full-length human prion protein (6LNI). Each line represents a PrP^{sc} chain from amino acids 170-229,which is the core of the fibril.



NCBI iCn3D Figure 4.10.18 Cryo-EM structure of an amyloid fibril formed by full-length human prion protein. (6LNI) (Copyright; author via source).

Click the image for a popup or use this external link https://structure.ncbi.nlm.nih.gov/i...ByT4rL33brHRY7

The aligned zig-zag lines clearly indicated beta strands aligned through hydrogen bonding the adjacent strands. The two alpha helices in the C-terminal domain become beta strands .

A genetic, inheritable form of the disease also exists, in which a mutant form of the PrP^c occurs, whose normal structure is destabilized by the mutation. The aggregates caused by the mutant form of the disease are understandable in light of the other diseases which we discussed above. The question is how does the normal PrP^c form PrP^{sc} . Evidence shows that if radiolabeled PrP^{*c} from scrapie-free cells is added to unlabeled PrP^{sc} from scrapie-infected cells, the PrP^{*c} is converted to PrP^{*sc}! It appears that the PrP^c protein has two forms not that much different in energy, one composed of mostly alpha helix and the other of beta sheet. A dimer of PrP^c·PrP^{sc} might form, which destabilizes the PrP^{*c} causing a conformational shift to the PrP^{sc} form, which would then aggregate. Exposure to the PrP^{sc} form would then catalyze the conversion of normal PrP^c to PrP^{sc} . Hence, it would be transmissible by contact with just the PrP^{sc} form of the protein. Likewise, species specificity could be explained if only dimers of

PrP^c.PrPsc formed from proteins of the same species could occur. The inherited form of the disease would be explained since the mutant form of the normal protein would more easily form the beta structure found in the aggregate.

It has recently been found that the very same mutation in PrP^c, Asp178Asn can cause two different diseases - CJD and FFI. Which disease you get depends on if you have 1 of two naturally occurring, nonharmful variants at amino acid 129 of the normal PrPc gene. If you have a Met at that position and acquire the Asp178Asn mutation, you get CJD. If, on the other hand, you have a Val at amino acid 129 and acquire the Asp178Asn mutation, you get FFI. This disease was first observed in 1986 and has been reported in five families in the world. It occurs in the late 50's, equally in men and women. It is characterized by a progressive loss of the ability to sleep and disrupted circadian rhythms. The brain shows neuronal losses. It is known that amino acids 129 and 178 occur at the start of alpha helices, as predicted from propensity calculations. Chronic exposure to micromolar levels of synthetic fragment 106-126 of PrP^c kills hippocampal neurons. This peptide also has the greatest tendency to aggregate synthetic PrP^c peptides.

Kuru killed many members of the Fore tribe in New Guinea until the cannibalistic practice of eating dead relatives was stopped. Analysis of the genes for the prion protein in the Fore tribe and other ethnic groups in the world show two versions differing by just one amino acid in all people (remember that a single gene is represented in both maternal and paternal chromosomes. That these two forms exist throughout the world suggests that they have been selected for by evolution and confer some biological advantage. People who have just one form of the protein are more susceptible to the development of prion diseases. Mead and Collinge have shown that about 75% of older Fore women (who had lived through cannibalistic practices) had two different prion genes, compared to about 15% of women from other ethnic groups. This high percentage suggests that these women were protected from the disease, leading through natural selection to a high percentage of heterozygotes in this defined population. The general presence of two forms of the prion gene (which probably offers protection from prion disease) suggests that cannibalism might have been widespread in our early ancestors.

There appears to be one main difference between the formation of amyloid fibers from prion proteins and others such as mutant lysozymes. If you add mutant lysozyme to normal lysozyme, the amyloid fibers contain only the mutant protein. However, if you incubate mutant prion proteins with normal prions, the normal proteins become pathological.

4.10.4: Misfolding and Aggregation Summary

Recent work has shown the proteins considered to be completely harmless can generate misfolded intermediates that aggregate to produce pre-fibril structures that are toxic to cells. This process is usually prevented in the cell by the interaction of nascent forms of the proteins with chaperones, which sequester exposed hydrophobic patches and prevent aggregation. (Obviously, prion proteins and the others mentioned above are exceptions). Amyloid fibers (characterized by subunits with an abnormal amount of beta-structure) can be made from many different types of proteins as noted above. Is this property specific to just a handful of proteins, or is it more common than expected from the limited examples noted so far? The new studies show that when a bacterial protein HypF is incubated at pH 5.5 in the presence of trifluoroethanol, aggregates (but not fibrils) form with enhanced beta structure. These aggregates slowly form into fibrils characteristic of amyloid protein fibers. The early aggregates (before fibril formation) proved cytotoxic. Similar results were seen with dimers and trimers (prefibril states) of the amyloid-b peptide released from cultured neurons.

A diverse group of proteins that do not share significant secondary or tertiary structures can form amyloid-like protein aggregates. Even though their monomer forms share little in common, the insoluble amyloid aggregates have a common structure in which the monomer in the aggregates has significant beta structure with the strands running perpendicular to the aggregate axis. Since it has recently been shown that almost any protein, under the "right" set of conditions, can form such aggregates, the stabilizing feature of protein aggregates must be potentially found in any protein. Evidence suggests that it is the polypeptide backbone, and not the side chains, that are key in the formation of stable interstrand H-bonds in beta secondary structures in amyloid aggregates. In contrast, native, nonamyloid forms of normal proteins arise through specific interactions of unique side chain sequence and structure, which out-competes nonspecific interactions among backbone atoms found in amyloid structures. Nonspecific aggregation becomes more prevalent when buried hydrophobic side chains and buried main chain atoms become more solvent exposed. Such exposure occurs when native proteins form intermediate molten globule states when subjected to altered solvent conditions or when destabilizing mutants of the wild-type protein arise. Some mutations may alter the cooperativity of folding which would increase the fraction of nonnative protein states. Other mutations that decrease the charge on the protein or increase their hydrophobicity might enhance aggregation. In addition, chemical modifications to proteins (such as oxidation or deamination) might destabilize the native state, leading to the formation of the molten globule state. Once formed, this state may aggregate through sequestering exposed side chain hydrophobes or through inter-main chain H bond formation. Aggregate formation appears to proceed through the initial formation of soluble units (which may or not be more toxic to cells than the final aggregate). Aggregates are kinetically stable

species. Since amyloid aggregates are cytotoxic and almost any protein can form them, albeit with different propensities, nature, through evolutionary selection, has presumably disfavored proteins with high tendencies to form such aggregates.

Clearly, accurate protein folding is required for cell viability. Aberrant protein folding clearly can be the cause of serious illness. Given the extraordinary nature of the task and its failure, the process governing protein folding must be highly regulated. Figure 4.10.19shows the steps that determine intracellular concentrations and locations of normal and aberrant protein structures.

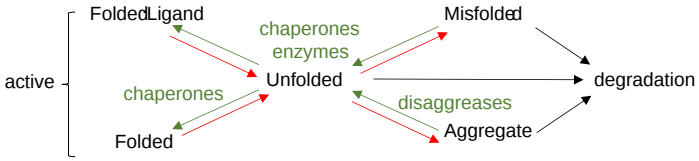


Figure 4.10.19: Conversion of active folded proeins to misfolded and aggregated proteins (after Balch et al, Science, 319, 916, 2008)

Potential therapies for diseases of proteostasis include replacing aberrant proteins, shifting the equilibria toward active forms with small ligands, or modulating the pathways with agents that influence pathways such as signal transduction, transcription, translation, degradation, and translocation using molecules like siRNAs to modulate concentrations of chaperons, disaggregases, and signal pathways.

4.10.5: Binding, Intracellular Granules and Droplets

The above structures are fascinating aggregates of specific proteins. The aggregates are quite large. In Alzheimer's Disease, they vary from around 150-500 μm^2, which would give a length of 12-22 μm if they were squares. In comparison, intracellular "**granules**" are much smaller with diameters from 200-500 nm or 0.2-0.5 μm. The term granule describes particles in cells that are just barely visible by light microscopy. Granules are found in many cells and mostly contain protein. Platelet granules contain many proteins involved in clotting. Pancreatic beta cell granules contain insulin for secretion. Other types of granules in germ-line cells are called various names such as dense bodies, perinuclear P granules in Caenorhabditis elegans, germinal granules in Xenopus laevis, chromatoid bodies in mice, and polar granules in Drosophila. The contains RNA as well as proteins. Those are often called ribonucleoprotein (RNP) granules. Plants and livers also contain starch (a carbohydrate) granules. The granules don't appear to be surrounded by a membrane. Rather they are just aggregates of proteins, or RNA and proteins. In Chapters 10 and 11, we will see analogous particles for lipids, nonpolar "insoluble" molecules that self-aggregate into micelles and membrane **bi**layers. Lipid droplets, which contain TAGs and cholesterol esters, in contrast to the granules mentioned above, are surrounded by a phospholipid **mon**olayer with adsorbed protein. Maybe an understanding of the structure and properties of phase-separated granules can shed light on the aggregates formed in neurodegenerative diseases.

How do these granules form? What principle underlies the specificity of protein and RNA found in them? The aggregates are not toxic compared to the beta-amyloid aggregates discussed above. A quick review of the Cell Tutorial (scroll to bottom) shows granule formation can be caused by a classic "phase transition", not unlike gaseous water can self-associate through hydrogen bonds to form liquid drops, which can freeze with the formation of more hydrogen bonds to form solids. Soluble biomolecules in cells can **reversibly** aggregate through the summation of multiple, weak noncovalent interactions to form storage granules. This balance might be perturbed if storage granules aggregate further in a potentially irreversible process with health consequences as we saw in neurodegenerative diseases. Let's delve into new insights into the processes involved in droplet formation.

Imagine small amounts of a sparing soluble oil added to an aqueous solution. Initially, it is in solution, but at a higher concentration, induced dipole-induced dipole interactions and the “hydrophobic effect” would drive the oil out of the solution into liquid drops. This phase separation could also be called **liquid-liquid demixing** as two liquids (solubilized oil in water and separated oil drops) separate. from each other. This process has been shown to produce many types of non-membrane bound droplets (not to be confused with membrane-bound vesicles) in the cell.

This phenomenon has also been seen with intrinsically disordered proteins and proteins with such domains. These are characterized by amorphous structures with repeated, often positively charged amino acids and/or contain a limited number of different types of amino acids. An example of a protein with a domain that has low sequence complexity is the SP1 transcription factor, a DNA binding protein. One of its tranactivation domains is comprised of almost 20% glutamines with regions within it having even higher percent abundances. It has been estimated that up to 20% of eukaryotic proteins don't have a stable shape as they are in part intrinsically disordered and contain **low complexity domains (LCDs)**. They are found in the N- and C-terminal ends of all mammalian intermediate filament proteins, almost all RNA binding proteins, lining the nuclear pore and in the cytoplasmic faces of

mitochondrial, lysosomal, peroxisomal and Golgi integral membrane proteins. "They decorate both ends of all 75 intermediate filament proteins found in mammals, fill the central channel of nuclear pores, adorn almost all RNA-binding proteins, and occur on the cytoplasmic faces of integral membrane proteins associated with mitochondria, neuronal vesicles, peroxisomes, lysosomes, and the Golgi apparatus. They are the target of up to 3/4s of posttranslational modifications. LCDs hence appear to facilitate the promiscuous binding of a variety of proteins, especially those that lead to or remove covalent tags.

Under the right condition, these can aggregate and “precipitate” from the solution. What is the nature of the precipitate? It might have properties more like distinct liquid droplets so this process could be called **liquid-liquid demixing**.

Properties of demixed drops would include reduced rates of diffusion of material into an out of the drop, coupled movements of materials in the drop, and probable weak hydrophobic-dependent aggregation making drops sensitive to agents like detergents. Liquid-like diffusion inside the drop is observed as evident by the rapid recovery of fluorescence from partially photobleached internal components of the drop.

As with the formation of a crystalline solid from a liquid solution, the process must be seeded. For intrinsically disordered proteins, this process can be “catalyzed” by poly-(ADP-ribose), a nucleic acid-like polyanion. The negative charges would counter the positive charges in the disordered protein domain, which without neutralization, would interfere with protein/protein contacts necessary for aggregation/droplet formation and demixing. Aggregation in these cases may arise from hydrophobic interactions (even though hydrophobic side chains are underrepresented in the disordered domains).

The solubility of proteins in cells is a fascinating topic in itself. High concentration of ATP (5 mM) in the cell actually helps to solubilize proteins. ATP is considered a **hydrotrope**. It’s a small molecule with a very distinct polar part (polyphosphate and ribose) and a more nonpolar part (the adenosine ring). Hence it acts sort of like a mini-detergent (an amphiphile) but it doesn’t form micelles. It does help stabilize more nonpolar parts of proteins in solution and has been shown to inhibit aggregate formation and also disaggregate some aggregates. Figure 4.10.20 shows a nonprotonated form of energy-minimized ATP with its dipole moment shown as an arrow from + to - end. The dipole moment would only be larger if the ATP was deprotonated and had negative charges.

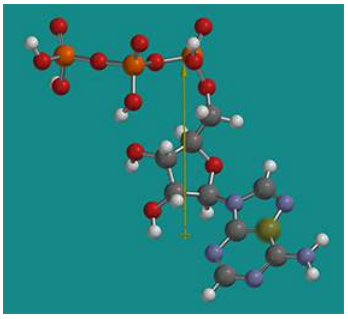


Figure 4.10.20: Dipole moment of ATP

Biochemists also use the term gel (examples include polyacrylamide gel or fibrin blood clots which are chemically cross-linked) and a "gel" form of a bilayer (Chapter 10), when they wish to describe a structure that is neither clearly solid nor liquid. Structures like the cytoskeleton or the actin-myosin network would be examples of the latter.

Noncovalent gels would be characterized by the regulatable dissociation of subunits and hence short half-lives. A gel (either covalent or noncovalent) with a high-water content could be called a hydrogel which would contain hydrophilic components. An example would be RNA and protein-containing particles

4.10.5.1: RNA granules

Granules that contain RNA and proteins are called ribonucleoprotein bodies (RNPs) or RNA granules. Specific examples of these include cytoplasmic processing bodies, neuronal and germ granules, as well as nuclear Cajal bodies, nucleoli and nuclear dots/bodies). Some granules just contain proteins, including inclusion bodies with misfolded and aggregated proteins and those with active proteins involved in biosynthesis, including purinosome (for purine biosynthesis) and cellusomes (for cellulose degradation).

Another feature found in some neurodegenerative diseases is a trinucleotide repeat. In Fragile X syndrome, there 230-4000 repeats of the CGG codon in the noncoding parts of the genome, compared to less than 50 in the normal gene. In Huntington’s disease, the

repeat CAG is found in the protein-coding part of the affected gene. The translated protein has a string of glutamines which probably causes protein aggregation. Specific proteins may also bind to the string of CAGs.

If the trinucleotide expansion is in the nonprotein-coding intronic DNA, deleterious effects are not associated with translated proteins but with the transcribed RNA in the nucleus. The intronic repeats would be spliced out of the primary RNA transcript. A CTG DNA repeat would produce a poly CUG containing RNAs (found in myotonic dystrophy), which could aggregate through non-perfect base pairing.

In vitro experiments show that small complexes are soluble, but as the size increases, a liquid-liquid demixing phase separation (or alternatively a liquid-gel transition) can occur, forming spherical droplets of RNA particles. This would explain the observation that pathologies occur above a certain repeat length. If misfolded proteins are also present, these particles might combine to form larger gels.

In the control experiment, when the repeats were scrambled, demixing and spherical particle formation were not observed. In an experiment similar to the addition of 1,6-hexanediol to intrinsically disordered proteins, if small antisense trinucleotide repeats, such as $(CTG)_8$, which could interfere with the weak H bonds between G and C in the aggregates, were added, the size of RNA drops (foci) were reduced. In vivo experiments showed characteristic drop-like structures but only if the repeats were of sufficient size.

Researchers found that in vitro, RNA drop formation was inhibited by monovalent cations. In the presence of 0.1 M ammonium acetate, which permeates cells without affecting pH, CAG RNA droplets in vitro disappeared.

Aggregation of mRNA might be one way to regulate its translation and hence indirectly regulate gene activity. There are advantages to regulating the translation of a protein from mRNA, especially if the "activity" of the mRNA could be dynamically regulated. This would be useful if new protein synthesis was immediately required. Hence one way to regulate mRNA activity (other than degradation) is through reversible aggregation.

4.10.5.2: Protein drops and granules

The cytoskeletal proteins actin and tubulin (heterodimer of alpha and beta chains) can exist in soluble (by analogy to water gaseous) states or in condensed, filamentous states (actin filaments and microtubules respectively). GTP hydrolysis is required for tubulin formation. Actin binds ATP which is necessary for filament formation but ATP cleavage is required for depolymerization. Hence nucleotide binding/hydrolysis regulates the filament equilibrium which differentiates from simple phase changes such as in water.

Since only certain proteins form granules, they must have similar structural features that facilitate reversible binding interactions. These proteins have multiple, weak-binding sites, but if they act collectively provide multivalent (multiple) binding interactions that allow robust but not irreversible granule formation. Here are some characteristics of proteins found in granules:

- the protein NCK has 3 repeated domains (SH3) that bind to proline-rich motifs (PRMs) in the protein NWASP. These proteins are involved in actin polymerization. In high concentrations they precipitate from the solution and coalesce to form larger droplets;
- repeating interaction domains are widely found especially among RNA-binding proteins;
- some proteins contain Phe-Gly (FG) repeats separated by hydrophilic amino acids in portions of the protein that are intrinsically disordered.
- a biotinylated derivative of 5-aryl-isoxazole-3-carboxyamide (Figure 4.10.21) precipitates proteins, which are enriched in those that bind RNA (RBPs). In general, the precipitate proteins were intrinsically disordered and characterized by low complexity sequences (LCS). One such example contained 27 repeats of the tripeptide sequence (G/S)Y(G/S). The proteins could also form hydrogels (made of hydrophilic polymers and crosslinks) and transition between soluble and gel phases with extensive hydrogen bond networks. The hydrogel gel phase gave x-ray diffraction patterns similar to beta structure-enriched amyloid proteins. Short-range, weak interactions between LCS might then drive reversible condensation to gel-like granule states characterized by extensive hydrogen bonding (again similar to hydrogen bonding on ice formation). If this process goes awry, more continued and irreversible formation of a solid fibril (as seen in neurodegenerative diseases) might occur from the hydrogel state;

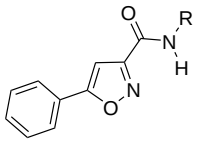


Figure 4.10.21: Structure of 5-aryl-isoxazole-3-carboxyamide

- RNAs appear in granules when proteins bind them through their RNA binding domains, which interact through low complexity sequences leading to phase separation and hydrogel-like formation of granules. Around 500 RNA binding proteins have been found in the human RNA interactome. They are enriched in LCSs and have more tyrosines than average proteins in the whole proteome in which the Tyr are often found in the (G/S)Y(G/S) motif. Phosphorylation of tyrosines (Y) in LCS may decrease association and hydrogel stability.

Given that so many neurodegenerative diseases are associated with unfolded/misfolded protein aggregates, the high protein concentrations in protein-containing liquid drops might pose problems to cells. If high enough, the equilibrium might progress from the liquid drop to a solid precipitate, which would have severe cellular consequences. The progression to the solid state may irreversibly affect the cell.

4.10.6: Low complexity domains (LCD) and neurodegenerative disease

The aggregation of alternatively-folded proteins is clearly associated with neurodegenerative disease. Mutations that lead to diseases lead to the association of low-complexity domains and aggregate formation, which is increasingly being described as phase separation. The demixed phases are stabilized by interchain backbone hydrogen bond as shown in the many beta-sheet aggregates described above. Evidence suggests that labile structures with potential for interchain H bonds and beta strand formation lead to fibril formation. The nascent interactions would involve short stretches of interchain H bonds. If so, mutations that enrich such nascent structural interaction would promote fibril formation while those that inhibit the nascent interactions would inhibit fibril formation. A study (Zhou et al, Science, 377, 2022. DOI: 10.1126/science.abn5582) verifies this.

The investigators made single amino acid variants of the low complexity domains of an RNA binding protein TDP-43 RNA that prevented that single amino acid within a region involved in interchain beta strand formation from forming a hydrogen bond through its amide hydrogen. They did this by methylating single main chain amide nitrogen, which prevents its participation in a hydrogen bond. The modification is shown in Figure 4.10.22

Figure 4.10.22 Methylation of a single backbone nitrogen in a region involved in interchain hydrogen bond and beta sheet formation in low complexity domain of proteins

Of the 23 variants they made, 9 within a continuous stretch inhibited phase separation. These 9 were at the same sites as hydrogen bonds between adjacent chains of the TDP-43 as determined by cryo-EM.

Next, they looked at other proteins with low-complexity domains that form aggregates/polymers. The proteins were the neurofilament light (NFL) chain protein, the microtubule-associated tau protein, and the heterogeneous nuclear RNPA2 (hnRNPA2) RNA-binding protein. They found 10 mutations in LCDs that were known to be associated with neurological disease. Indeed these mutations allow one extra single hydrogen bond in the low complexity domain sequences, and display enhanced aggregate formation mediated presumably through the extra interchain H bond. Specifically, the known mutations replace individual prolines, a cyclic amino acid that lacks an amide H and hence cannot donate a hydrogen bond, with another amino acid, which allows one additional hydrogen bond. Each of the known mutations was associated with neurological disease and increased stable aggregate/polymer formation. This increased aggregation/polymer (phase separation) was reversed in vitro by chemical methylation of the single amino acid change in the mutant which prevented it from forming hydrogen bonds. The site-specific methylation was performed by linking synthetic peptides containing the single, N^{α}-methyl amino acid to the other synthesized peptides that comprise the protein. The semisynthetic NFL protein, for example, was incubated under conditions conducive to the assembly of mature intermediate filaments.

In vitro experiments were conducted using different synthetic head domains of the neurofilament light (NFL) chain protein in which the P8 residue contained a different amino acid at those positions. The variant amino acids are shown in Figure 4.10.23

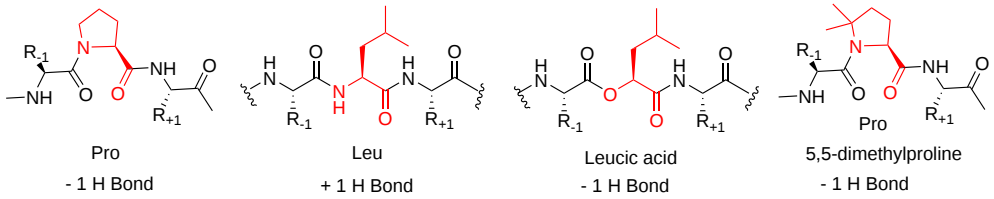


Figure 4.10.23 Variant amino acids used to at position P8 in the low complexity head domain of the neurofilament light (NFL) chain protein (after Zhou et al, ibid)

Only variants containing Leu at position P8 were able to form filaments as measured by in vitro fluorescent studies. Experiments like this are key in ascertaining where phase separation/aggregation causes and are not merely correlated with the development of complex neurodegenerative diseases.

4.11: Biomolecular Visualization - Conceptions and Misconceptions

Structure determines everything in biology and chemistry. Since you learned to represent molecules with Lewis structures, it's been drilled into you that the structure of a molecule determines its physical and chemical properties. Physical properties would include melting points, boiling points, and solubility. In contrast, chemical properties include acid/base, redox, precipitation, and general chemical reactivity determined by the presence of Lewis/Brønsted acids/bases and nucleophiles/electrophiles. More modern analyzes of reactivity would include molecular orbital theory descriptions of bonding.

What makes chemistry and its fundamentally interconnected fields of biology and biochemistry so difficult to many is that we can't see molecules but make inferences from data (x-ray crystallography, NMR spectroscopy, and cryo-EM) about the structure of a molecule (atom type, atom/bond connectivity, and geometry). As the molecules get bigger (consider the muscle protein titin, also called connectin, with a molecular weight of around 3.8 million), we must use computer visualization to understand the structure and infer from it the resulting function and activity of the protein. As with small molecules, we can render the molecule in different ways to better understand various attributes of the molecule that confer function and activity. We ask students to view a biomolecule and infer its properties from the rendering without giving devoted attention and instruction as to how to do that.

4.11.1: Small molecules

Let's start with a small molecule like oleic acid, a long chain carboxylic acid with 18 carbon atoms and one cis (Z) double bond between carbons 9 and 10 as shown in Figure 4.11.1.

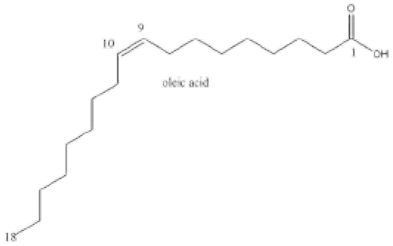


Figure 4.11.1: Line drawing of oleic acid

Table 4.11.1 below shows multiple ways to render the molecule. Each rendering offers insight in the function/activity of the molecule but might at the same time leave students with difficulties in interpreting them and also reinforce or install misconceptions. Each representation below shows the very same molecule. The top row shows representations without H atoms, which the bottom row shows them.

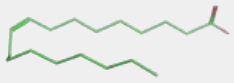
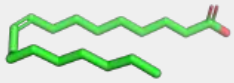
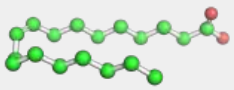
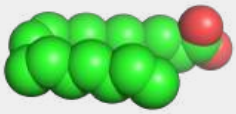
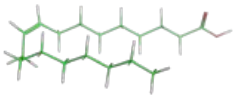
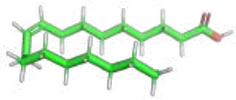
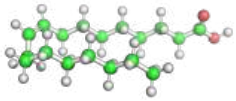
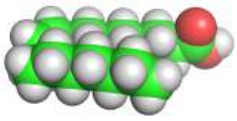
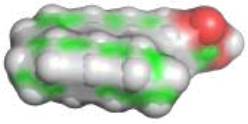
line	stick	ball sphere	sphere	surface
				
				

Table 4.11.1: Different renderings of oleic acid

Here are some important things to remember about biomolecular structures including both small and large molecules:

- Structures obtained using x-ray crystallography are constructed from relative electron densities calculated from diffraction patterns. Computer programs calculate the structure based on these electron density maps and known bond lengths, bond angles, atom types. Most structures used in this book are found in the Protein Data Bank. The structures seen in computer models are visualized data and can contain mistakes (missing atoms, steric conflicts, wrong atoms), although structural refinement techniques minimize such problems.
- Structures derived from X-ray and cryomicroscopy analyzes are static structures and represent only one of a large ensemble of possible conformational structures. As you learned from the study of simple molecules in organic chemistry, bond lengths and

angles changes can change within molecules. Bonds connecting two atoms can stretch, angles connecting three atoms can bend, and the torsional angle around the center bond in a four-atom, three-bond system can rotate to form eclipsed and staggered (gauche and anti) conformers.

- PDB structures obtained by x-ray crystallography contain no H atoms as they are too small and contain too few electrons to diffract/scatter x-rays. So get used to adding them in your mind when you see a structure. Programs are used to calculate and show H bonds between a slightly positive H atom on an O or N atom in a protein, and another slightly negative Os or Ns on the same or different molecule. The H bonds are often shown between N and O atoms. You should look at the atoms involved and the distance between them and visualize a hydrogen atom connected to one of them. Figure 4.11.2 shows an example of an H bond between two base pairs in a DNA molecule.

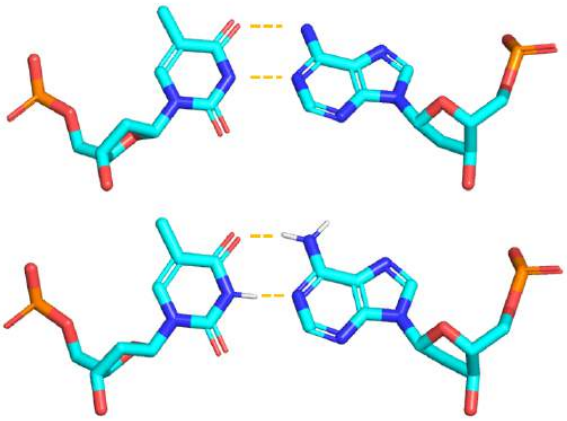


Figure 4.11.2: Hydrogen bonds between two base pairs of DNA, with one showing actual hydrogen atoms

- Double bonds or likely resonance structures are not typically shown in PDB structures.
- Line, ball and stick, and stick renderings are useful for showing connectivity between atoms and bond angles. However, they are not particularly useful in showing how atom size might affect the molecules' structure and properties. This type of information is better shown when spacefill renderings that show the sizes of the atoms (based on their Van der Waals radii) is used, or when the surface of the molecule, calculated from contact surface created between the van der Waals surface of the atoms and a rolling probe (often an O atom mimicking water) is displayed.

4.11.2: Large molecules

As molecules get bigger, line, stick, and ball and stick renderings are increasingly useless. New ways of visualizing the structural features of the molecules become needed. The importance of multiple renderings to clarify structure/function relationships becomes apparent when you wish to understand protein structure. Various renderings of the protein superoxide dismutase (2sod) are shown in Table 4.11.2 below.

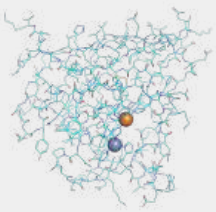
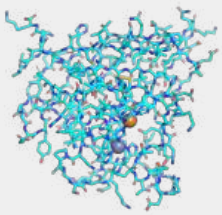
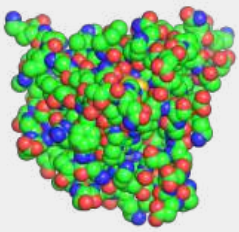
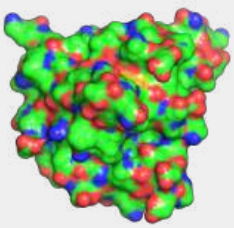
line	stick	cartoon	sphere	surface
				

Table 4.11.2: Multiple renderings of the protein superoxide dismutase

- The same features and limitations described above for small molecules apply to large ones. It is best to leave out most of the atoms and use mixed renderings within a single display to reveal important structural feature of the biomolecule. The cartoon rendering of superoxide dismutase shows one tiny alpha helices (red) and many beta strands (yellow). All side chain and

backbone atoms have been removed. The green line is the trace through the backbone of those amino acids not involved in secondary structure. The Cu and Zn ions are shown as spheres.

Another type of surface rending, the electrostatic potential surface, is beneficial. Figure 4.11.3 shows the electrostatic potential surface of superoxide dismutase taken from two different angles after simply rotating the protein.

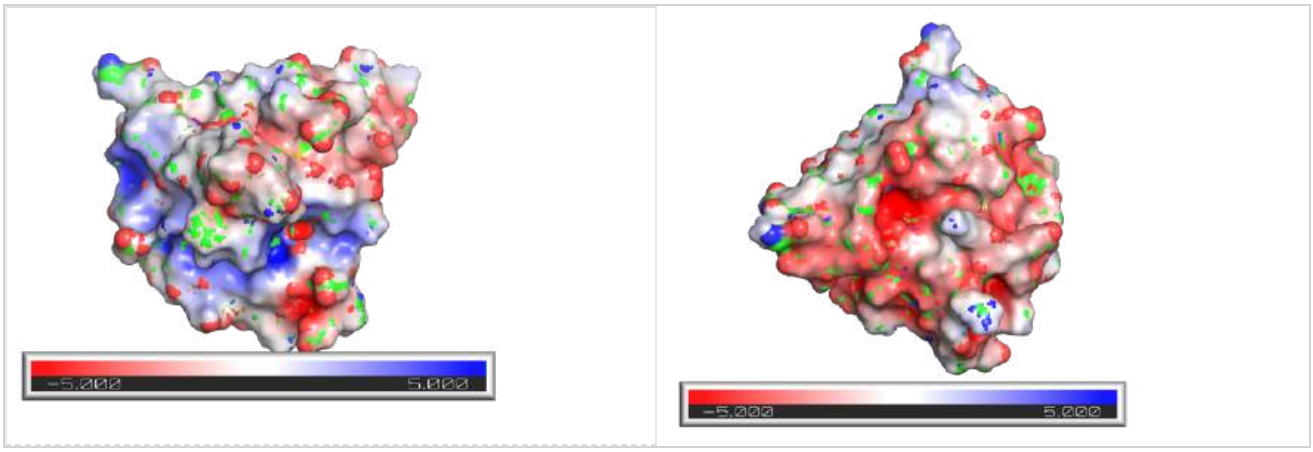


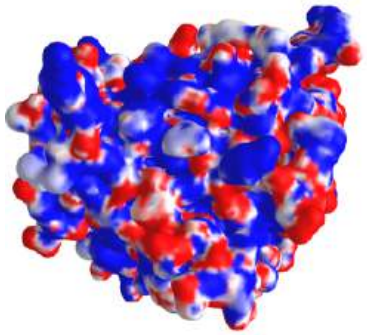
Figure 4.11.3: Electrostatic potential surface of superoxidase dismutase from two perspectives

The red represents minimal (most negative) potential. This part of the structure would be enriched in slightly negative Os and Ns or fully negative Os (i.e. have the highest electron density). The representesents the positive potential, centered in areas of the surface containing slight or full positive charge (I,.e.the lowest electron density). This enzyme binds superoxide, O_2^-, a toxic free radical reduction product of dioxygen, O_2. It catalyzes this reaction:

$$2\,O_2^- + 2\,H^+ \rightarrow O_2 + H_2O.$$

The enzyme can effectively scavenge superoxide in its vicinity as the negative superoxide is drawn into the active site with the Cu and Zn atoms by the positive potential surrounding the active site, enhancing the normal diffusion encounter rate of the reactant with the enzyme pocket.

Figure 4.11.4shows an interactive iCn3D model of the electrostatic potential of superoxidase dismutase (2sod).



NCBI iCn3D Figure 4.11.4: Electrostatic potential of superoxidase dismutase (2sod) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...QQy2hAXehWLBW8

4.12: Laboratory Determination of the Thermodynamic Parameters for Protein Denaturation

4.12.1: Introduction

Multiple methods can be used to investigate the denaturation of a protein. These include UV, fluorescence, CD, and viscosity measurement. In all these methods the dependent variable (y) is measured as a function of the independent variable, which is often temperature (for thermal denaturation curves) or denaturant (such as urea, guanidine hydrochloride) concentration. From these curves, we would like to calculate the standard free energy of unfolding (ΔG^{O}) for the protein (for the reaction N ↔ D). It is relatively to calculate if the denaturation curves show a sigmoidal, cooperative transition from the native to the denatured state, indicating a two-state transition. The dependent variable can also be normalized to show fractional denaturation (fD). An idealized denaturation curve is shown below in Figure 4.12.1.

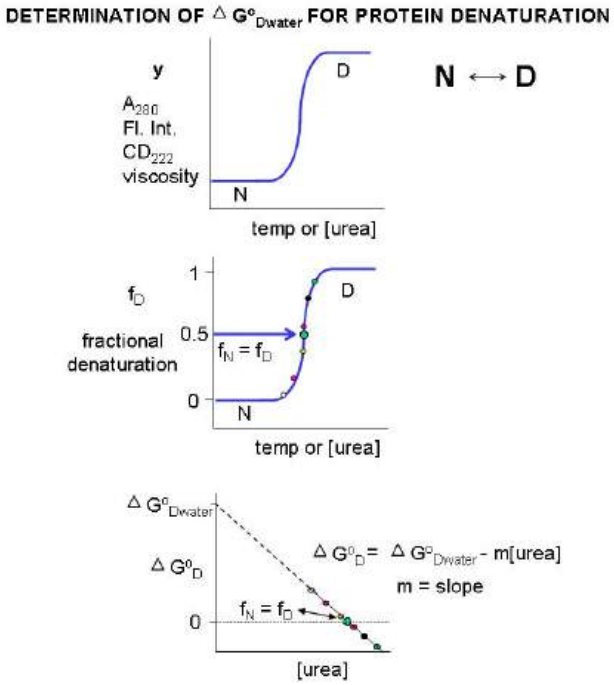


Figure 4.12.1: Grapcial analysis of protein denaturation

A more realistic denaturation curve might show a small linear change in the dependent variable (fluorescence intensity for example) for temperature or denaturant concentrations well before the major unfolding transition, as well as above those at which it is unfolded. In these cases, the mathematical analyses presented at the end is required.

4.12.2: Denaturation with urea or guanidine hydrochloride

For each curve, the value of y (either A_{280}, fluorescence intensity, viscosity, etc.) can be thought of as the sum contributed by the native state and from the denatured states, which are both present in different fractional concentrations from 0 - 1. Hence the following equation should be reasonably intuitive.

$$y = (f_N y_N) + (f_D y_D) \tag{4.12.1}$$

where f_N is the fraction native and y_N is the contribution to the dependent variable y from the native state, and f_D is the fraction denatured and y_D is the contribution to the dependent variable y from the denatured state. Conservation gives the following equation.

$$1 = f_N + f_D \text{ or } f_N = 1 - f_D \tag{4.12.2}$$

Substituting 4.12.2 into 4.12.1 gives

$$y = (1 - f_D)\, y_N + (f_D y_D) = y_N - f_D y_N + (f_D y_D)$$

Rearranging this equation gives

$$f_D = \frac{y - y_N}{y_D - y_N} \tag{4.12.3}$$

Notice the right-hand side of the equations contains variables that are easily measured.

By substituting 4.12.2 and 4.12.3 into the expression for the equilibrium constant for the reaction $N \rightleftharpoons D$ we get:

$$K_{eq} = \frac{[D]_{eq}}{[N]_{eq}} = \frac{f_D}{f_N} = \frac{f_D}{1 - f_D} \tag{4.12.4}$$

From this, we can calculate ΔG^0.

$$\Delta G^0 = -R\,T \ln K_{eq} = -R\,T \ln \left[\frac{f_D}{1 - f_D} \right] \tag{4.12.5}$$

Remember that ΔG^O (and hence K_{eq}) depends only on the intrinsic stability of the native vs denatured state for a given set of conditions. They vary as a function of temperature and solvent conditions. At low temperatures and low urea/guanidine HCl concentration, the native state is favored, and for the N ↔ D transition, $\Delta G^O > 0$ (i.e., denaturation is NOT favored). The denatured state is favored at high temperatures and urea/guanidine-HCl concentrations and $\Delta G^O < 0$.

At some temperature or urea/guanidine concentration value, both the native and denatured states would be equally favored. At this value, $K_{eq} = 1$ and $\Delta G^O = 0$.

The temperature at this point is called the melting temperature (T_m) of the protein. It is analogous to the T_m in the heat capacity vs temperature graphs for protein denaturation, as we saw in Chapter 4.9. This is Illustrated in Figure 4.12.2

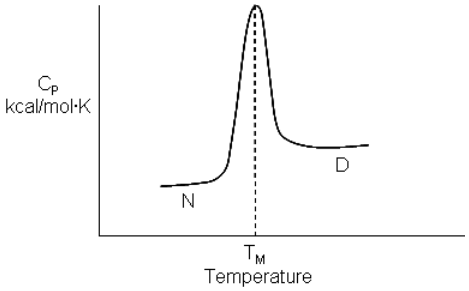


Figure 4.12.2: T_m for thermal denaturation of a protein

Ordinarily, at a temperature much below the T_m for the protein or at a low urea concentration, so little of the protein would be in the D state that it would be extremely difficult to determine the protein concentration in the D state. Hence it would be difficult to determine the K_{eq} or ΔG^o for the reaction N ↔ D. However, in the range where the protein denatures (either with urea or increasing temperature), it is possible to measure f_D/f_N.and hence ΔG^o at each urea or temperature.

Now we can calculate the $\Delta \mathbf{G}^o_w$ for N ↔ D in water without urea. For a simple two state N ↔ D, a plot of ΔG^o vs [urea] is linear and given by the following equation, which should be evident from (\PageIndex{1}\).

$$\Delta G^0 = \Delta G^0_w - m[\text{ urea }] \tag{4.12.6}$$

It is important to know the K_{eq} and ΔG^0 for the N ↔ D transition in the absence of urea and under "physiological conditions". A comparison of the calculated values of ΔG^0_w in the absence of urea for a series of similar proteins (such as those varying by a single amino acid prepared by site-specific mutagenesis of the normal or wild-type gene, would indicate to what extent the mutants were stabilized or destabilized compared to the wild-type protein. ΔG^0_w for the N ↔ D transition of the protein in the **absence** of denaturant (i.e in water) can be determined by extrapolating the straight line to [urea] = 0. Admittedly, this is a long extrapolation, but with high-quality data and a high correlation coefficient for the linear regression analysis of the best-fit line, reasonable values can be obtained.

4.12.3: Denaturation with heat

Calculation of $\Delta \mathbf{H}^o$ and $\Delta \mathbf{S}^o$ for N <=>D at room temperature

K_{eq} values can be calculated from thermal denaturation curves in the same way as described above using urea as a denaturant by monitoring change in an observable (spectra signal for example) vs temperature. Knowing K_{eq}, ΔH^0, DS^0 can be calculated from equation x below since a semi-log plot of $\ln K_{eq}$ vs 1/T is a straight line with a slope of - $\Delta H^0 R$ and a y-intercept of + $\Delta S^0/R$, where R is the ideal gas constant.

$$\begin{aligned} \Delta G^0 = \Delta H^0 - T\Delta S^0 = -RT\ln K_{eq} \\ \ln K_{eq} = -\frac{\Delta H^0 - T\Delta S^0}{RT} \\ \ln K_{eq} = -\frac{\Delta H^0}{RT} + \frac{\Delta S^0}{R} \end{aligned} \tag{4.12.7}$$

From these equations, it should be evident that all the major thermodynamics constants (ΔG^0, ΔH^0 and ΔS^0) for the N ↔ D transition can be calculated from thermal denaturation curves.

Equation (9) below shows that the derivative of equation (8) with respect to 1/T (i.e. the slope of equation 8 plotted as $\ln K_{eq}$ vs 1/T) is indeed $-\Delta H^0/R$. Equation (9) is the van 't Hoff equation, and the calculated value of the enthalpy change is termed the van 't Hoff enthalpy, $\Delta H^0 v_{Hoff}$.

$$\frac{d\ln K_{eq}}{d(1/T)} = -\frac{\Delta H^0}{R} = -\frac{\Delta H^0_{vHoff}}{R} \tag{4.12.8}$$

It is useful to compare the van 't Hoff enthalpy, $\Delta H^0 v_{Hoff}$, with the enthalpy change determined directly using differential scanning calorimetry by analyzing a plot of Cp vs T. (Note that the area under the Cp vs T curve as the protein transitions to the unfolded state has units of kcal or kJ. The ΔH^0 for the unfolding is inversely proportional to the width of the curve.)

In contrast to the long extrapolation of the ΔG^0 vs [urea] to [urea] = 0 to get ΔG^0 (the y-intercept) in the absence of urea, which has some physical meaning, extrapolation of the straight line from the van 't Hoff plot from equation 8 to get $\Delta S^0/R$, the y-intercept, has little meaning since the 1/T value at the y-intercept is 0, which occurs when T approaches infinity. ΔS^0 can be calculated at any reasonable temperature from the calculated value of ΔG^0 at that temperature and the calculated ΔH^0_{vHoff}.

4.13: Predicting Structure and Function of Biomolecules Through Natural Language Processing Tools

Recent Updates: New Chapter Section 11/1/23

Written by Logan Hallee and Henry Jakubowski

4.13.1: Introduction

So far in this chapter, you have learned about protein structure and its determination in the laboratory. After decades of work in modeling protein structure and properties, the life science community has built massive databases organizing this information. While sequencing DNA and discovering protein sequences has become relatively cheap, the actual characterization of protein structure and function is still time and cost-intensive. Instead, researchers look to model and predict protein properties from their amino acid sequence alone to speed up the work necessary in the lab. The most recent and effective tool in this quest is the **protein language model (pLM)**, which models proteins as a biological language of amino acids. At their core, pLMs are **transformers** with a protein vocabulary.

The transformer, an attention-based neural network, emerged as a game-changer for the scientific community with the iconic 2017 paper “Attention Is All You Need.” The crux of this work is the revolutionary idea that by strategically organizing simple neural networks, performance can be enhanced beyond merely scaling up a singular neural network. A neural network is a type of AI/machine learning process, often described as deep learning, which is patterned after the brain with nodes (neurons) that are interconnected by lines (axon/dendritic connections) as in the brain.

Especially adept at processing sequential data like time series or sentences, transformers have become the bedrock of modern **natural language processing (NLP)**. This technology is grounded in understanding the semantic and contextual intricacies of their trained vocabulary.

4.13.2: The Essence of Protein Language Models

Tokenization & Token Embedding

The fundamental problem in NLP is encoding text, a string datatype, into a meaningful numerical representation; it is awfully hard to do math on words. One approach to the problem is to give every sub-part of the vocabulary, say a word, a unique integer. That way, any sequence of words can be turned into a vector of integers and we can easily do math on vectors; think back to physics or math doing the dot product on collections of numbers that have *direction*. This process creates a look-up table called a **tokenizer** because it translates tokens (strings of words, letters, or characters) to integers.

? Vectors - A Simplified Review

To understand how AI/Machine Learning can be used to predict structure and function, we need to know about vectors and their use in physics and mathematics, particularly in matrices. Most students likely need a refresher. Click the link below for a guided view that will allow you to get a better understanding of the rest of the material in this section.

The Review!

Vectors – A Simplified Review

Most biochemistry students have taken physics in high school and college. In those courses, you were introduced to scalar and vector quantities. Scalar quantities like distance and work have no direction. Vector quantities like displacement and force have both a magnitude and a direction. Vectors are shown as arrows with the length representing the magnitude and the direction by an arrow at the end of the vector.

Let’s review a simple concept from elementary physics, work. Work is a scalar quantity and you probably remember that mechanical work is done on an object when an external force moves an object a given distance. Consider a force **F** (bold represents a vector) applied to a block which causes it to move a distance along the surface. The distance and direction together are described as **displacement d**, as shown in Figure 4.13.1below.

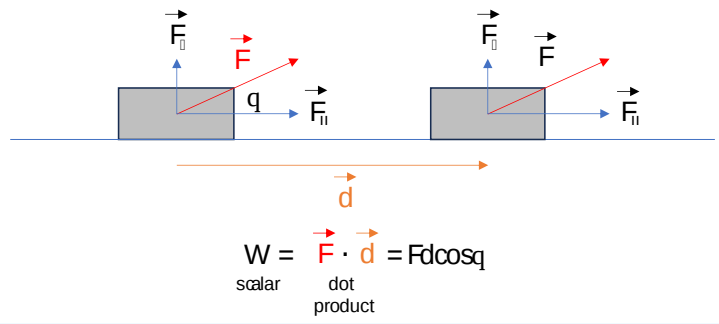


Figure 4.13.1: Forces on a block moving along a surface

If the force was applied vertically, the block would not move along the surface, so no work is done on the box. If the force is applied at some angle, only the horizontal component of the force would cause the block to slide.

The horizontal component of the force is described as **F**cosθ. (When $\theta = 90^0$, $\cos\theta = 0$ so no work is done.) Hence the work W = Fcosqd, which is the "dot product" of the vectors **F** and **d**.

Vectors are also used in math and can be considered as directed line segments. Let's consider the equations for circles and spheres. These equations are based on the Pythagorean Theorem. As you learned in high school geometry, the Cartesian equation of a circle is:

$x^2 + y^2 = r^2$

To generate a circle, set the r-value to a fixed number (such as 1 for a "unit" circle"), and for a multiple number of x values, calculate y values (where $-1 \leq x,y \leq +1$) from the equation. Then plot the x and y coordinate pairs and presto you have a circle, as shown in Figure 4.13.2below.

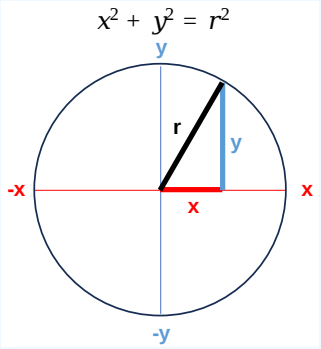


Figure 4.13.2 Cartesian graph of a circle

Each of the (x,y) pairs can be considered a 2D vector with the origin at 0,0, a magnitude of 1 (the fixed radius in this example), and a direction, described by the specific x,y points that fall on the circle. Two simple (x,y) pairs are (1,0) and (0,1) for the unit circle. Another x,y pair that satisfies the Pythagorean theorem is (0.5, 0,866)

The Pythagorean Theorem can be extended to three dimensions to give the Cartesian equation for a sphere:

$x^2 + y^2 + z^2 = r^2$

To generate a sphere simply solve for z for a multitude of x and y values and a fixed r value. Then plot the x, y, and z values and presto you get a sphere, as shown in Figure 4.13.3below for a "unit" sphere of radius 1.

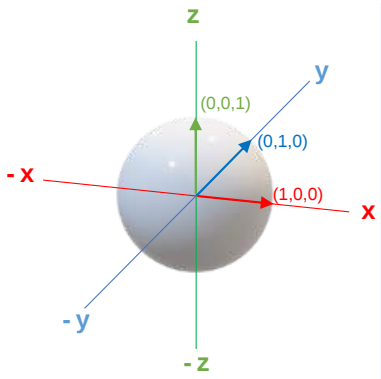


Figure 4.13.3 Cartesian graph of a sphere

The sets of x, y, and z points that land on the surface of the surface are vectors (directed line segments).

Now vectors are also used to describe matrices. Matrices are two-dimensional arrays of numbers. A matrix with just one row or one column is called a row vector or column vector, respectively, as shown in Figure 4.13.4below.

$\begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$ Row vector

$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ Column vector

$\begin{bmatrix} 1 \end{bmatrix}$ 1 row, 1 column, scalar

$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ Square matrix = # columns/rows

Figure 4.13.4

You should now see that all of the vectors that define a sphere can be written as a large matrix. Figure 4.13.5 below shows a x3 square matrix that represents the unit vectors that lie along the x, y, and z axes.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Figure 4.13.5

In this example, any vector will lie on the surface of the unit sphere if the three components have the relationship of the Cartesian equation above.

These examples showcase how values in vectors can represent a position in space, but they can also be measurements of an object. For example, a car with 4 tires, 4 cylinders, and 180 horsepower could be represented as (4, 4, 180), where a more detailed description will lead to a longer vector with more components.

Another use of vectors and matrices comes from more complicated mathematical graphs. Graphs are arbitrary objects with nodes and edges, where edges connect nodes. A popular example of a mathematical graph is the social network Facebook. You can represent the entirety of the Facebook network with each member as a node and an edge between each node when the nodes are friends on the site. An example of a social networking graph is shown in Figure 4.13.6below.

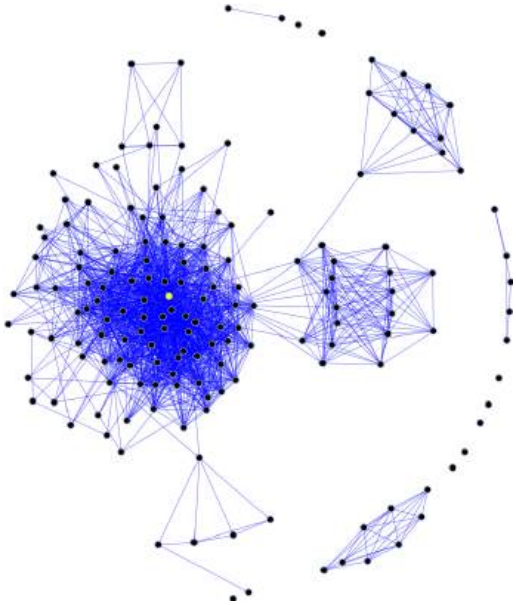


Figure 4.13.6 Graph showing social relationships using graph theory. Darwin Peacock. CC BY 3.0, https://commons.wikimedia.org/w/inde...?curid=6057981

This example is an unweighted undirected graph because the edges do not have a specific value or direction. We could make a weighted graph that represents Facebook by utilizing weighted edges, perhaps the number of mutual friends between different members. This is still an undirected example.

Graphs are an important concept in computational biochemistry because molecules can be represented as graphs, with atoms as nodes and bonds as edges. This is illustrated in Figure 4.13.7 for a multidentate adsorbate complex.

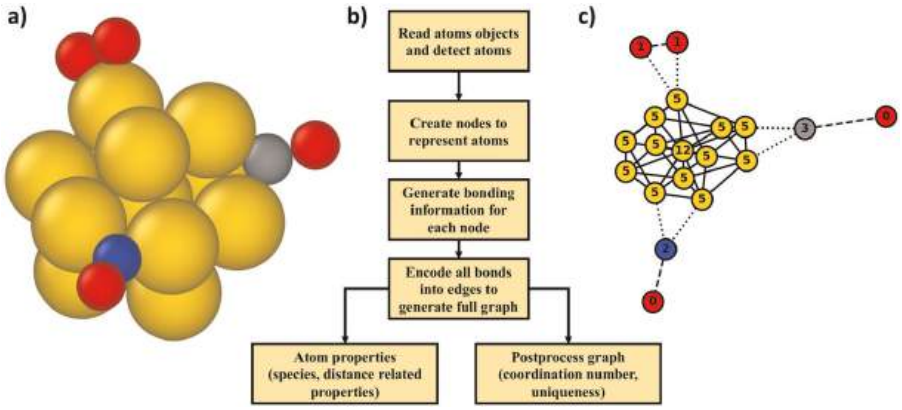


Figure 4.13.7: Graph theory-based algorithm to generate graphs for a given atomic model. Deshpande, S., Maxson, T. & Greeley, J. Graph theory approach to determine configurations of multidentate and high coverage adsorbates for heterogeneous catalysis. npj Comput Mater 6, 79 (2020). https://doi.org/10.1038/s41524-020-0345-2. http://creativecommons.org/licenses/by/4.0/. Creative Commons Attribution 4.0 International License.

Panel a shows an atomic model for a simple nanoparticle with adsorbates. Panel b is an algorithm to generate graph-based representations. Panel c shows the generated graph mode

Figure 4.13.8 below shows a protein structure graph (right) for a short stretch of an alpha helix (left).

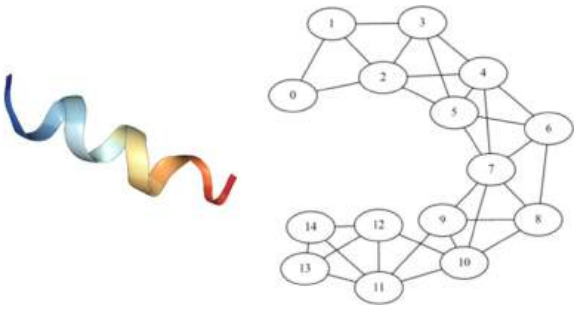


Figure 4.13.8: Protein and protein graph. Baranwal, M., Magner, A., Saldinger, J. et al. Struct2Graph: a graph attention network for structure-based predictions of protein–protein interactions. BMC Bioinformatics 23, 370 (2022). https://doi.org/10.1186/s12859-022-04910-9. http://creativecommons.org/licenses/by/4.0/

Figure 4.13.9 shows a graph not of a single protein but of a small protein:protein interaction network.

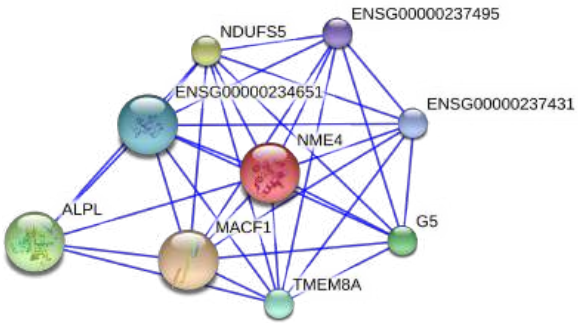


Figure 4.13.9: protein interactions of TMEM8A in humans. https://commons.wikimedia.org/wiki/F...for_TMEM8A.png

You can even produce graphs that represent entire networks of molecules and their relationships. A directed molecular graph might showcase proteins and their substrate. Having a direction in an edge is important in this distinction because a protein may use a substrate for a chemical reaction but a substrate might not act on a protein on its own. Many molecular relationships are weighted and directed. A weight in the protein substrate case might be the relative affinity of binding between the molecules.

These types of graphs contain three types of edges, undirected, directed, and weighted as illustrated in Figure 4.13.10 below.

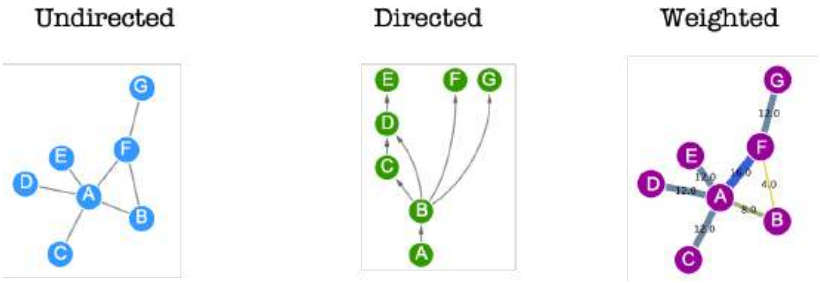


Figure 4.13.10: The main types of edges found in a network. https://www.ebi.ac.uk/training/online/courses/network-analysis-of-protein-interaction-data-an-introduction/introduction-to-graph-theory/graph-theory-graph-types-and-edge-properties/ . Attribution 4.0 International (CC BY 4.0) license

- Undirected edges: Connections in protein-protein interactions, as shown in Figure 9 above, are examples. The proteins are connected through binding but without implied flow between them;
- Directed edges: These are found in metabolic and signaling pathways when arrows indicate the flow of reactants/products in a pathway. These can be arranged in complex hierarchies as those familiar with metabolic and signaling pathways know.
- Weighted edges: Either undirected or directed edges can have a quantitative weight associated with them. These many reflect affinities, similarities between genes, fold effects, etc.

These examples are great for showcasing the versatility of mathematical graphs, but how can we use matrices to represent them? Enter adjacency matrices.

Adjacency matrices state which notes are connected. Each node can have multiple features. Figure 4.13.11 below shows a network of 5 nodes, the adjacency matrix, and a features matrix with each node having features. For example, if an atom is a node, the features could be electronegativity, partial charge, size, etc.

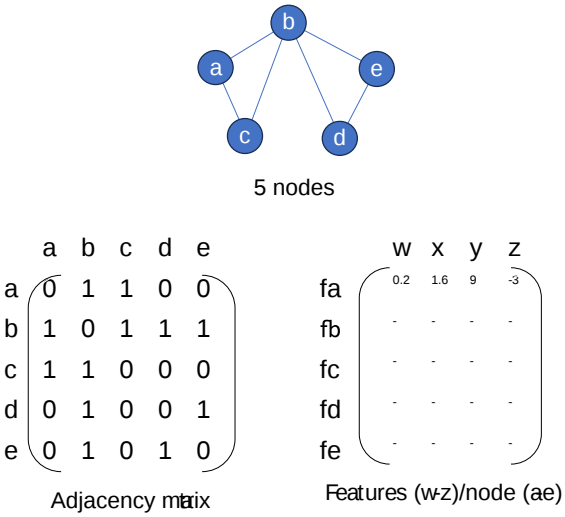


Figure 4.13.11: Properties of a 5 node network

An adjacency matrix stacks n vectors together for a graph that has n nodes. The vectors are also n long, so the resultant matrix is n by n. At the ith jth index of the matrix is a number dictating how many edges the node shares. So if the 1st node has an edge to the 2nd node the 1st row and 2nd column of the adjacency matrix will have a 1. These tend to be symmetric, and in this example, there would also be a 1 at the 2nd row and 1st column. Figure 4.13.12below does a great job at explaining:

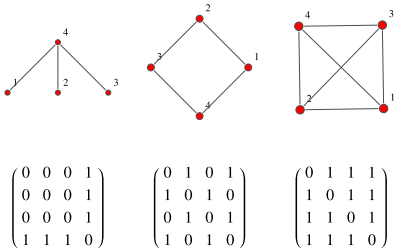


Figure 4.13.12 https://mathworld.wolfram.com/AdjacencyMatrix.html

If there is a weight on an edge the number in the adjacency matrix can be used to store the weight instead of the count of edges between nodes. This can also hold a direction by allowing for positive and negative entries.

Any complex network can be described mathematically as an adjacency matrix with rows and columns indicating nodes and an edge as a number. Unweighted and undirected edges lead to symmetric matrices with just 0 and 1. Directed and weighed can be more complicated with different numbers used to show relationships like affinity. +/- values can be used where + is an activation and a – is an inhibition. These matrices can be manipulated using linear algebra. Examples of adjacency matrices for Undirectrf, Directrf, and Weighted networks are shown in Figure 4.13.13below.

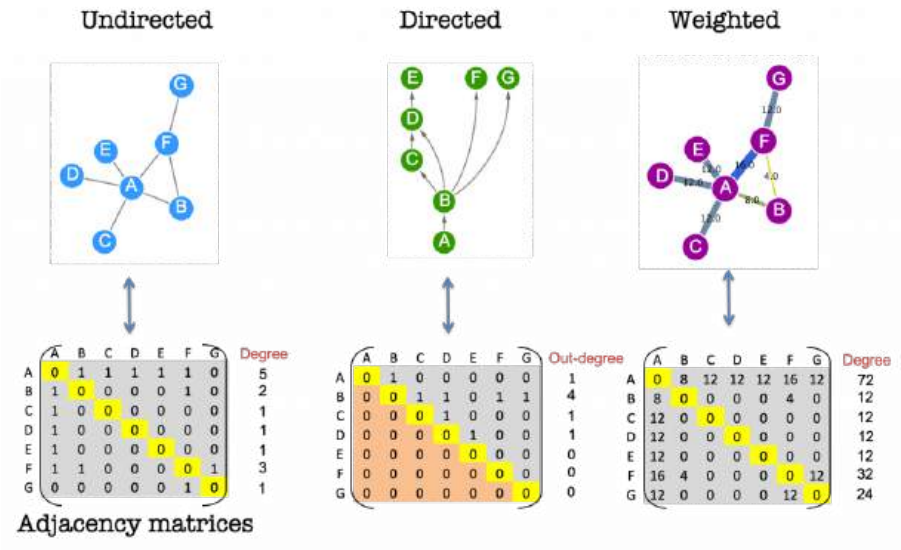


Figure 4.13.13 Adjacency matrices from undirected, directed and weighted networks. https://www.ebi.ac.uk/training/online/courses/network-analysis-of-protein-interaction-data-an-introduction/introduction-to-graph-theory/graph-theory-adjacency-matrices/ . Attribution 4.0 International (CC BY 4.0) license

Hopefully, after this brief intro to vectors and matrices, you can now understand how amino acid position and their semantic properties (polarity, charge, size, etc) could be described as large matrices and their associated vectors.

4.13.3: Back to Tokens

Here are a few types of tokenizers commonly used in NLP:

1. Word Tokenizers: Word tokenizers split text into individual words based on spaces or punctuation marks. This approach assumes that words are the primary units of meaning in a language. For example, given the sentence "The cat is sleeping," a word tokenizer would split it into tokens: ["The", "cat", "is", "sleeping"].
2. Subword Tokenizers: Subword tokenizers split text into subword units that capture partial linguistic information. This approach is useful for handling out-of-vocabulary words like abbreviations or reducing vocabulary size which saves computational resources. Popular subword tokenization algorithms include Byte Pair Encoding (BPE), Unigram Language Model, and SentencePiece. Subword tokenizers can be complicated, but one possible subword tokenization or our example above would be [“The#”, “cat#”, “is#”, “sleep”, “ing#”] where # has been added to showcase the ending of a word.
3. Character Tokenizers: Character tokenizers treat each character as a separate token. This approach is beneficial when dealing with languages without explicit word boundaries or for character-level modeling tasks. Our sleepy cat is now [“T”, “h”, “e”, “\s”, “c”, “a”, “t”, “\s”] and so on. Here, we need to add a space token so the model can tell where the words start and finish.

For pLMs, researchers typically treat amino acids as tokens or “words” and protein sequences as “sentences.” Ex. a protein sequence like "MVKLTA" would be tokenized into individual amino acids: 'M', 'V', 'K', 'L', 'T', and 'A'.

The main problem with tokenizing sentences or sequences is that this numerical space has no *semantic* meaning. The grammar and word meanings are not encoded here. To work on this semantic information storage we will define another look-up table. However, this time, we will make it a large matrix with some arbitrary dimension, say, 768. Let us assume there are 50,000 or so English words we want to model, so our matrix is 50,000 x 768. Now, we will simply connect our tokenizer and our matrix. If the word “dog” corresponds to the token 5, the *vector* that represents the dog will be the 5th row of our matrix. Maybe, “protein” is token 800. The *vector* that represents protein is the 800th row in our matrix.

These vectors that represent each word are where we will store the semantic information behind each word, the meaning, and related grammar within the vocabulary. Unlike an object like a car, with easy numerical features to pick out (4 door, 4 wheel, 6 cylinder, etc.) we need to learn the features of a word. And so, within the transformer neural network, we have a vector that represents each word that contains learnable weights. Through the process of gradient descent, these weights will be adjusted from a random starting point to contain the necessary information for language modeling. One neat fact is that within a well-trained

English language model, the vectors for words King - Man + Woman roughly equals Queen. How amazing, that concepts behind gender and royalty can be encoded in a meaningful numerical space!

Of course, many words have different meanings in different contexts. This is still the case in proteins, where specific residues may be important because of their charge, or maybe just because of the space they occupy. That is why we need a portion of our transformer that handles the contextual understanding. This is where **attention** comes in.

4.13.4: Attention

Attention is used to signify the importance of a word or its part. If high attention is given to a word, for example, less overall information is needed to predict the word. Attention can be conferred by weighting the importance of the word/token. In real language, we attend to certain words in sentences and weigh them more to provide context and make predictions. Consider this sentence referring to a TV show or movie. Predict which show/movie is it from.

..and Spock said that the needs of the many outweigh the needs of the few.

The answer comes quickly (if you are a Star Trek fan) by attending to the word Spock. You could maybe predict the entire sentence using databases by perhaps attending to another word such as outweigh.

Using attention enables the transformers we discussed above to long-term memory and concentrate (attend) to previously generated tokens.

The attention mechanism identifies dependencies and relationships between tokens within an input sequence. But what makes attention especially compelling is its ability to dive deep into a relational space and then seamlessly bridge back into the token embedding space we previously discussed.

Another feature is self-attention which focuses on the relationships with a given sentence.

Self-attention, where every token in a sequence attends to every other token, allows for a dynamic weighting of significance.

Three key inputs are required. The Query is what you are asking (such as the input in a web search box), the Key is the search results, and the Values are the returned content in each search result. In a web search, the program has to find the best matches between the query and the keys.

Now consider both the query and key to be vectors. The similarity between them can be determined using a cosine similarity function which is a bit similar to the dot product of the vectors = A·K/|A||K|= Akcosθ/|A||K|. The denominator is the product of the vector lengths. The cosine function makes sense since it varies between +1 (when the vectors are in the same direction, q = 1) and -1 (when they point in opposite directions, q = -1). This gives the degree of similarity.

Mathematically the relationships among query, key, and values matrices are formulated as

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{4.13.1}$$

Here, the Q, K, and V stand for query, key, and value matrices respectively, and d_k is the dimension of the queries and keys. These come directly by multiplying learned weights with our token embeddings.

$$Q = W_e \times W_Q, \quad K = W_e \times W_K, \quad V = W_e \times W_V, \tag{4.13.2}$$

The softmax function turns a vector of K real values into a vector of K real values that sum to 1, in effect converting them into probabilities (that sum to 1).

W_e symbolizes the token embedding matrix, or, at least the portion extracted from our tokenizer and token embedding process. d_K denotes the dimension of the key matrix, which is used to normalize the values.

During the attention operation, there is a length-wise representation and an embedding-wise representation. The length-wise representation is typically referred to as an “attention matrix” and is often analyzed for researchers to look at how a model makes decisions. Coming from the Softmax(QK^T/d_K) during the attention operation is this attention matrix: A_{LxL} where L symbolizes the sequence length and values are scaled from 0 to 1. When $A_{\{i,j\}}$ is close to 1 signifies that the i^{th} amino acid bears contextual relevance to the j^{th} one. In pLMs, such associations often resonate with spatial or chemical bonds—perhaps the i^{th} amino acid is nestled close to the j^{th} one in the protein's 3D structure. Notably, $A_{\{i,j\}}$ scores are perpetually high, emphasizing an amino acid's

inherent significance to itself. When A is multiplied further by V, we go from this relational space back to the embedding space, where the matrix is L x embedding dimension. In our previous example, this was 768 but this number is completely arbitrary. A smaller embedding dimension essentially leads to a smaller model with a worse theoretical top-end performance, with the advantage that it is much cheaper to use and train. A larger embedding dimension and model will have a theoretically high top-end performance, require more data and computational resources to train, and be more expensive to use.

Diving deeper, multi-head attention further refines this process. It transforms the input sequence into multiple smaller queries, keys, and values—each an independent attention head brandishing unique weights. Each head, in isolation, computes its own version of the attention matrix on a section of the input. These matrices are then woven together, under an additional learned linear transformation, into a final cohesive representation. This multi-headed approach has proven important for pLMs, where an increased headcount is correlated with better performance.

Bringing it all together ...

Lastly, bridging back from the realm of relational depths to the embeddings, transformers don't solely rely on attention. They also encompass feed-forward layers which further improve the generalized modeling capabilities. Together, these components are the essence of the transformer; a multi-headed attention layer and a linear layer. These transformer layers are simply stacked on top of each other, so the input of one is the output of another. These transformer stacks are referred to as large language models (LLMs). LLMs have a couple of main forms and some other important parts that we did not discuss. For example, there are transformer encoders (BERT) and decoders (GPT) that have different organizations of layers. There are also some additional embeddings called positional embeddings, which help transformers learn the intended order of input sequences because the native mathematical operations are position invariant. Also, there are different types of normalizations and skip connections that are also extremely important for protein modeling. So while this introduction to transformers is enough to get your feet wet, we have included several other resources to further refine your knowledge if desired.

4.13.5: Training transformers

At a high level, transformers are trained just like any other neural network. For a large corpus of training data, there is an input and a known desired output called the ground truth. The input is fed to the model and the output is compared against the ground truth with a loss function. This is all set up in a way so that the loss function measures some type of error between the output and the ground truth so that the loss function, and corresponding error, are minimized with optimization techniques. Regardless of the technical details, it is essentially strategic trial and error, where the model is rewarded for producing outputs close to or equal to the ground truth. This way, through optimization, the error is reduced over time and the model learns how to do the specified task effectively.

However, because transformers have such a large parameter count they take a lot of data to successfully adjust for a specific task. Thus, we need a way to produce a ton of labeled data without wasting too much human time annotating data. Luckily, sequences of strings are perfect for this. If we feed a transformer a sequence we can simply hide a token and ask the model to predict what word went there.

Now, in the input, let's replace some words (or tokens) with **Mask** tokens randomly. Mask tokens are used to hide tokens from the model so that we can train the model to recover the missing words, where the desired output is the original sequence. Such a task forces the model to learn what words are around it; hence building the model's semantic and contextual understanding. This is called denoising because we artificially injected noise into our input and our model used the surrounding context to get rid of the noise. Another popular training objective is next token prediction, where part of a sequence is input and the model predicts what goes next. Different organizations of transformer layers are better or worse at these various tasks.

The subject of a MASK token brings up the broader topic of special tokens. Special tokens are tokens that are added to a specific vocabulary that serve a specific purpose. As we discussed, a model can learn to replace a MASK token with a correct token that belongs in the sequence.

Some other popular tokens are **CLS** and **SEP**, which stand for **classification** and **separator** respectively.

CLS tokens are typically prepended to the beginning of sequences so that models can learn to summarize the entire input in a single vector, which is useful for classification tasks. Separator tokens are typically placed at the end of sequences or between sequences if more than one was fed to the model at a time. This way, the model can treat separate sequences as individual entities even if they are input at the same time. Researchers will often create new special tokens to use alongside model training for specific tasks. For example, a pLM called ProstT5 has a special token that indicates a translation from structure to amino acid sequence and an

additional token that does the reverse. The respective token was prepended to the necessary inputs during training so that the model could intuit which task it was supposed to be performing. This is particularly useful if you are designing a transformer model that needs to do multiple distinct tasks.

4.13.6: Applications of protein language models

Now that we have learned substantially about how modern language modeling works we can explore how researchers utilize these techniques to further advance our understanding of biochemistry.

Prediction 3D Structure

By in large the most famous pLM is AlphaFold: The deep learning model that first successfully mapped amino acid sequences to protein structure at a large scale. Since then, other pLMs have also been able to learn from a large corpus of sequences and structure data to perform well on unseen sequences. The fundamental organization of sequence-to-structure models is made up of

- a transformer (pLM) that builds a semantic and contextual understanding of sequences;
- a structure module that maps the latent sequence representation to 3D coordinates.

AlphaFold is special in a few ways. Firstly, the transformer used for AlphaFold is actually two transformers in one! Let us call the first transformer T1. T1 is an MSA transformer, which works on several sequences at once instead of a single sequence. This comes from the concept of multiple sequence alignment (MSA), which is a common search method in bioinformatics that compares amino acid strings based on an evolutionary-informed algorithm. The output of an MSA is a list of strings aligned based on similarity and substitution probability. Figure 4.13.13below …

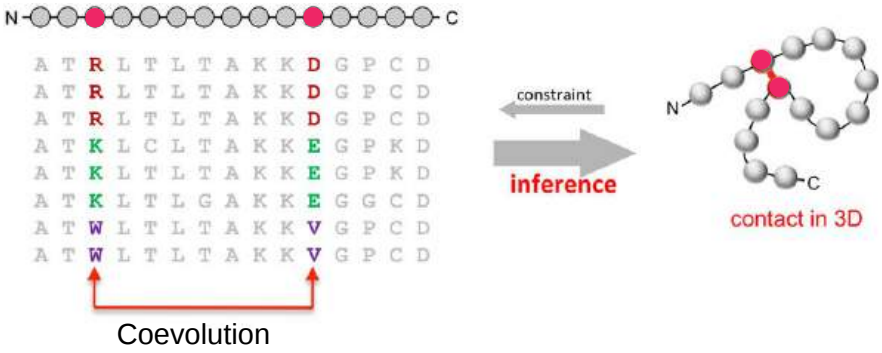


Figure 4.13.13 Correlated mutations carry information about distance relationships in protein structure.https://www.blopig.com/blog/2021/07/alphafold-2-is-here-whats-behind-the-structure-prediction-miracle/

The sequence of the protein for which the 3D structure is to be predicted (each circle is an amino acid residue, typical sequence length is 50–250 residues) is part of an evolutionarily related family of sequences (amino acid residue types in standard one-letter code) that are presumed to have essentially the same fold (iso-structural family). Evolutionary variation in the sequences is constrained by many requirements, including the maintenance of favorable interactions and indirect residue-residue contacts (red line, right). The inverse problem of protein folding prediction from sequence addressed here exploits pair correlations in the multiple sequence alignment (left) to deduce which residue pairs are likely to be close to each other in the three-dimensional structure (right). A subset of the predicted residue contact pairs is subsequently used to fold up any protein in the family into an approximate predicted 3D shape ('fold') which is then refined using standard molecular physics techniques, yielding a predicted all-atom 3D structure of the protein of interest

Including this information in the model allows the model to learn more about the protein than it could from a single sequence because of the concept of coevolution. Coevolution is a simple process that necessitates meaningful substitutions in amino acid sequences through various mutations if an organism is going to stay fit.

However, as we have discussed, transformers are incredibly adept at processing and understanding single sequences; how would one process many sequences simultaneously? The MSA input to the pLM is many sequences stacked on top of each other, so the attention mechanism needs to be modified. The MSA goes through row-wise attention, which picks out the important residues, and column-wise attention which picks out the important sequences. This creates a protein latent space built from an MSA instead of a single sequence.

The other transformer, T2, is also a transformer with modified attention (triangular self-attention). T2 is for computing over a pair-wise representation of the single input sequence. It builds a representation similar to a distogram, which is a matrix representation of the distance between every residue and every other residue. T1 and T2 make up the section of AlphaFold known as the EvoFormer, and there are 48 of these EvoFormer layers in total. The latent representation of the MSA input from T1 and the pair-wise representation from T2 are both inputs structure module.

The structure module also has some fancy attention (invariant point attention) that merges the information from T1 and T2. A simple computer-vision-inspired architecture called a ResNet used the attention output to predict side chain and backbone torsion angles at the atomic level. From these outputs, the atom coordinates for the entire protein are calculated and the structure is relaxed with Amber, which removes any structural violations based on the charges and locations of the atoms.

This entire process is repeated three times, where the structure information and MSA information can inform each other through skip connections and linear transformations. This recycling greatly improves the final 3D structure. Figure 4.13.14below.

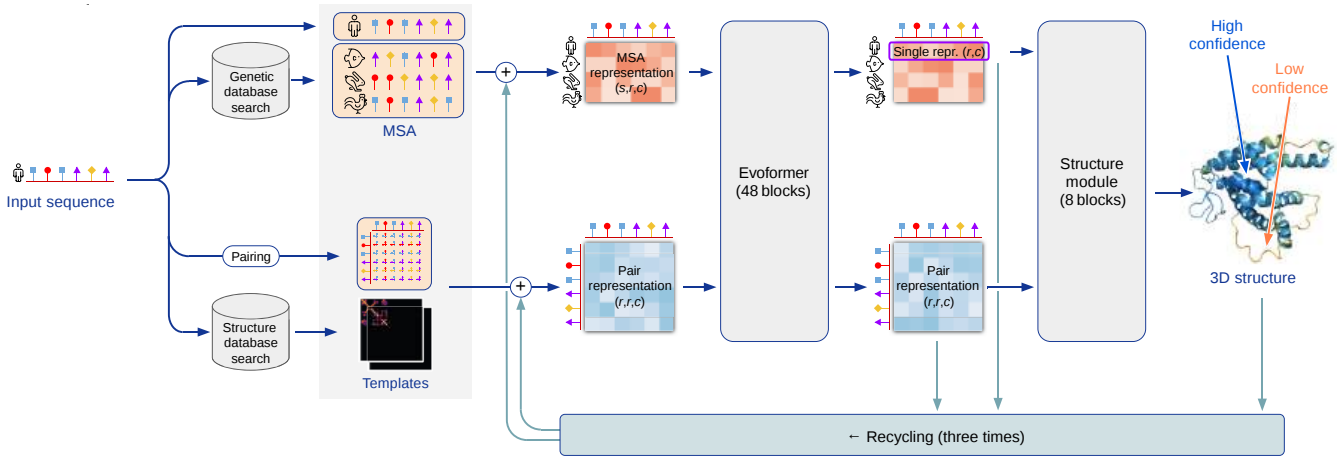


Figure 4.13.14 Model architecture. Jumper, J., Evans, R., Pritzel, A. et al. Highly accurate protein structure prediction with AlphaFold. Nature 596, 583–589 (2021). https://doi.org/10.1038/s41586-021-03819-2. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Arrows show the information flow among the various components. Array shapes are shown in parentheses with s, the number of sequences; r, the number of residues; c, the number of channels.

In summary, AlphaFold is the combination of three main networks that talk to each other and work together. The input sequence is searched with MSA through a large database that pulls out similar sequences. These similar sequences are input to T1 which builds a semantic and contextual understanding of the amino acids called the latent space. This latent space informs a pair-wise distogram that tracks the distances between amino acid residues in the original input sequence. All of this information is utilized by the structure module which does some fancy math to calculate the 3D coordinates of each atom. The entire process is repeated so the structure information can inform the MSA and the MSA can further inform the structure. Throughout the process, a ResNet and Amber are used to prevent any weird side chain or backbone angles that cannot exist in nature.

Other projects have also taken the MSA approach to structure prediction. RoseTTAFold and xtrimoGPLM utilize an MSA with different networks and attention variants to perform with similar accuracy to AlphaFold. However, MSA can be quite computationally expensive and requires a database where similar sequences exist. For so-called protein orphans, who have little to no sequence homology in recorded repositories, these methodologies fall short.

To address the computational and protein orphan concern various projects like ESM, OmegaFold, and Ember have utilized a standard pLM trained through mask denoising and a structure module to also obtain high accuracy on structure prediction, outperforming MSA-based method on protein orphans. Breakthroughs from AlphaFold like recycling are standard practice throughout these different approaches.

Some researchers utilize these various projects in parallel to get multiple predictions of structure for the same sequence. This ensemble approach allows accelerated progress by leveraging the advantages of each model and averaging out the disadvantages. Low confidence regions or disagreements in models may also be an indicator of intrinsically disordered regions of protein structure, which is incredibly important in biological function.

All in all, sequence-to-structure mapping has been effectively correlated with modern computational methods. The backbone of pLMs and structure modules has enabled the large-scale annotation of protein sequences with high-quality structure, something the scientific community can utilize for accelerated breakthroughs. However, structure prediction is not the only thing you can do with pLMs.

Protein function prediction

The latent space learned from vast mask denoising on corpora of amino acid sequences correlates extremely highly with protein structure, which is vital for the study and annotation of proteins (link to other parts of the textbook). Interestingly, this latent space also correlates highly with other useful types of annotation like function. By averaging across the length of the last hidden state output of pLMs one can build an effective vector representation of a protein, called a vector embedding. This way, every protein has the same size numerical representation and can easily be fed to machine learning classifiers like support vector machines, k-nearest neighbors, random forests, and more. The pLM can also be fine-tuned as a classifier given enough annotated data. Precomputed protein embeddings from popular models like ProtT5 can be downloaded from UniProt, as well as large amounts of annotated data if you would like to try this for yourself.

The types of annotations researchers are typically interested in come down to details like EC and GO classes. EC stands for Enzyme Commission, and EC numbers break down protein functionalities into a hierarchical organization scheme delimited by what type of reaction the proteins catalyze.

For example, the hierarchy of an EC number can be illustrated as:

- 1st digit: Represents one of the six primary classes of enzymes, e.g., '1' stands for oxidoreductases.
- 2nd digit: Describes a subclass within the primary class. If an enzyme is a '1.1', it specifically deals with acting on the CH-OH group of donors.
- 3rd digit: Categorizes the enzyme even further by specifying the acceptor. For instance, '1.1.1' would mean that the enzyme acts on the CH-OH group with NAD+ or NADP+ as the acceptor.
- 4th digit: Provides a unique identifier for each enzyme within its specific class, subclass, and sub-subclass. So, '1.1.1.1' is the EC number for alcohol dehydrogenase.

In total, there are currently over 8000 unique EC numbers! pLMs have shown remarkable competency in predicting them from sequence alone, often achieving between 80-90+% accuracy on unseen data.

Another type of popular annotation is Gene Ontology (GO) which labels genes based on what their protein products do in a biological context. GO is split into three main categories

1. Biological Process (BP): Describes a series of events accomplished by one or more ordered assemblies of molecular functions. For instance:

   GO:0006955 - Immune response

   GO:0006958 - Complement activation, classical pathway

   GO:0045087 - Innate immune response

   … and so on

2. Cellular Component (CC): Describes parts of a cell or environment a protein product likely localizes to. For example:

   GO:0005634 - Nucleus

   GO:0005654 - Nucleoplasm

   GO:0005694 - Chromosome

   … and others

3. Molecular Function (MF): Describes catalytic activities, such as binding or catalysis, that occur at the molecular level. This subcategory is very similar to EC numbers. For instance:

   GO:0003824 - Catalytic activity

   GO:0016491 - Oxidoreductase activity

   GO:0016614 - Oxidoreductase activity, acting on CH-OH group of donors

… and further subcategories

The main difference in organization is that GO terms have parent-child relationships that are not unique. In simpler terms, a GO term can have multiple parent terms while an EC number may only have one. Regardless, pLMs also show a wide breadth of impressive performances in predicting GO terms from sequence alone.

Additionally, researchers are interested in the complex interplay of proteins in the cell. How do proteins modify each other and their surrounding cellular components? Why does a specific gene expression cause a disease state? Which chaperones or post-translational modifications can contribute to homeostasis and which ones are detrimental? All of such questions can be picked away by building an understanding of protein-protein interactions and networks.

Protein-protein interactions (PPIs) can be defined in a variety of ways, but the literature typically focuses on some mediated chemical or conformational changes when one protein comes in contact with another. Some other terms are often added to the definition requiring an interaction to have a nonredundant function in some sense. Regardless, researchers typically drastically simplify the problem by treating it as a binary classification: Proteins either interact or not. In biological context, it is much more complicated than this but good PPI classifiers are still informative towards the questions we mentioned above.

Recently, pLMs have received increasing attention for their ability to compare protein sequences and guess about interaction in biological contexts. Vast databases of positive interactors enable this type of analysis. However, confirming that two proteins for certain never interact is a much harder problem. Clever data science towards training PPI classifiers with massive inherent class imbalance is challenging, but there are many promising modern approaches. Hence, PPI classification is another way where protein function can be partially uncovered computationally via pLMs.

Protein sequence generation

Lastly, we discuss one more general application of pLMs in protein sequence design and generation. As we discussed above, transformer decoders (or GPT models) are often trained to predict the next token given some other tokens for context. The popularized ChatGPT does this incredibly well for the English language. Generative pLMs perform the identical task on amino acid tokens, generating sequences from scratch or completing sequence prompts. Many pLM projects are notable in this space.

- ProtGPT2: Stacked transformer decoders that generate viable nature-like sequences from scratch.
- ProGen: Stacked transformers decoders that generate plausible sequences given control tags for context, thus being able to generate sequences of a particular family or ontology.
- ANKH: A general-purpose encoder-decoder pLM that can generate proteins of a specific superfamily or plausible variants with possible increased functionality.
- ProtDT: A pLM fine-tuned by contrasting vector embeddings with an English language model to enable protein generation based on English natural language input.
- xtrimoGPLM: A massive general-purpose pLM that has an extremely capable sequence design. It can even generate sequences with nearly identical structures that have almost no sequence similarity.
- PostT5: A fine-tuned version of ProtT5, which is also capable of sequence generation. This encoder-decoder architecture can generate a sequence given a structure input that approximates said structure. This is a bilingual model with amino acid and structure-based tokens.
- SAProt: An encoder-only system with a similar bilingual vocabulary as ProstT5 that can also translate between sequence and structure, enabling sequence generation based on a structure.

4.13.7: Concluding remarks

Protein language modeling is an interdisciplinary science at the intersection of bioinformatics, biochemistry, and computational sciences. Such modeling techniques are becoming an integral part of biochemical research through the fast-paced progress of NLP and computational hardware. It is easy to recognize the potential of protein language modeling in general life sciences: The generation of novel sequences for therapeutics, industrial catalysts, and synthetic biology, all the while annotating newly sequenced and generated proteins alike.

Importantly, amino acid-based vocabularies are not the only biochemically relevant use of NLP models. DNA, codon, and even atom-wise vocabularies are being explored in many applications: For example, genomics, phylogenetics, and small-molecule-to-protein interactions. There are many avenues ready to explore in the biological NLP field.

As someone learning competency in biochemistry, it is important to keep in mind the capabilities of protein language modeling while recognizing it is a somewhat new and accelerating field. Computational protein modeling may look vastly different in two

years and will likely be completely different a decade from now. Regardless, computational tools serve as a helper to biochemists, not replacements. Standard biochemical assays to determine the structure and function of proteins will always be necessary to confirm and further inform the findings of a computational domain. This computational domain can simply weed out the plausible options for a given problem away from the implausible ones, enabling the impossible-seeming search of biochemical space.

4.13.8: References

Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, L.; Polosukhin, I. Attention Is All You Need. arXiv December 5, 2017. https://doi.org/10.48550/arXiv.1706.03762.

EMBL-EBI. Introduction to graph theory | Network analysis of protein interaction data. https://www.ebi.ac.uk/training/onlin...-graph-theory/ (accessed 2023-10-28).

Geetansh Kalra. Attention Networks: A simple way to understand Self Attention. https://medium.com/@geetkal67/attent...n-f5fb363c736d

DeepFindr. Understanding Graph Attention Networks. https://www.youtube.com/watch?v=A-yKQamf2Fc

Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. Nature 2021, 596 (7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2.

Lin, Z.; Akin, H.; Rao, R.; Hie, B.; Zhu, Z.; Lu, W.; Smetanin, N.; Verkuil, R.; Kabeli, O.; Shmueli, Y.; dos Santos Costa, A.; Fazel-Zarandi, M.; Sercu, T.; Candido, S.; Rives, A. Evolutionary-Scale Prediction of Atomic-Level Protein Structure with a Language Model. Science 2023, 379 (6637), 1123–1130. https://doi.org/10.1126/science.ade2574.

Chen, B.; Cheng, X.; Geng, Y.; Li, S.; Zeng, X.; Wang, B.; Gong, J.; Liu, C.; Zeng, A.; Dong, Y.; Tang, J.; Song, L. xTrimoPGLM: Unified 100B-Scale Pre-Trained Transformer for Deciphering the Language of Protein. bioRxiv July 6, 2023, p 2023.07.05.547496. https://doi.org/10.1101/2023.07.05.547496.

Elnaggar, A.; Heinzinger, M.; Dallago, C.; Rehawi, G.; Wang, Y.; Jones, L.; Gibbs, T.; Feher, T.; Angerer, C.; Steinegger, M.; Bhowmik, D.; Rost, B. ProtTrans: Toward Understanding the Language of Life Through Self-Supervised Learning. IEEE Trans Pattern Anal Mach Intell 2022, 44 (10), 7112–7127. https://doi.org/10.1109/TPAMI.2021.3095381.

Elnaggar, A.; Essam, H.; Salah-Eldin, W.; Moustafa, W.; Elkerdawy, M.; Rochereau, C.; Rost, B. Ankh: Optimized Protein Language Model Unlocks General-Purpose Modelling. arXiv January 16, 2023. https://doi.org/10.48550/arXiv.2301.06568.

Su, J.; Han, C.; Zhou, Y.; Shan, J.; Zhou, X.; Yuan, F. SaProt: Protein Language Modeling with Structure-Aware Vocabulary. bioRxiv October 2, 2023, p 2023.10.01.560349. https://doi.org/10.1101/2023.10.01.560349.

Heinzinger, M.; Weissenow, K.; Sanchez, J. G.; Henkel, A.; Steinegger, M.; Rost, B. ProstT5: Bilingual Language Model for Protein Sequence and Structure. bioRxiv July 25, 2023, p 2023.07.23.550085. https://doi.org/10.1101/2023.07.23.550085.

Hallee, L.; Rafailidis, N.; Gleghorn, J. P. cdsBERT - Extending Protein Language Models with Codon Awareness. bioRxiv September 17, 2023, p 2023.09.15.558027. https://doi.org/10.1101/2023.09.15.558027.

Hallee, L.; Gleghorn, J. P. Protein-Protein Interaction Prediction Is Achievable with Large Language Models. bioRxiv June 9, 2023, p 2023.06.07.544109. https://doi.org/10.1101/2023.06.07.544109.

Ferruz, N.; Schmidt, S.; Höcker, B. ProtGPT2 Is a Deep Unsupervised Language Model for Protein Design. Nat Commun 2022, 13 (1), 4348. https://doi.org/10.1038/s41467-022-32007-7.

Liu, S.; Zhu, Y.; Lu, J.; Xu, Z.; Nie, W.; Gitter, A.; Xiao, C.; Tang, J.; Guo, H.; Anandkumar, A. A Text-Guided Protein Design Framework. arXiv February 9, 2023. http://arxiv.org/abs/2302.04611 (accessed 2023-02-14).

4.13: Chapter 4 Questions

Section 1 Questions:

Q1) Newman projection

Q2) Explain what conformation peptide bonds predominately fall into, cis or trans. Can you think of a situation in which the nonpredominate form could be found in a biological system?
A2) Most peptide bonds are in the trans conformation which helps to reduce steric hindrance R-groups of the amino acids but also decreases the strain on the peptide linkage. The cis form can be found in peptides, but at a low percentage. As the cis form causes strain on the peptide backbone. Enzymes can utilize this bond strain to overcome the activation energy for catalysis.

Q3) Ramachandran Plot example (define the regions)

Section 2 Questions:

Q1) Which amino acid(s) is(are) considered to be alpha helix breaker(s)? Explain why this occurs.
A1) Proline is consider to be the amino acids that "breaks" the secondary structure of an alpha helix. This occurs due to the amine ring in the proline R-group. This R-group is unable to participate in H-bonding needed for helix stabilization, and also is sterically locked in one conformation. Glycine is also an alpha helix breaker. Glycine has only a hydrogen as its R-group and therefore has no constrains for its Φ and Ψ angles. This makes glycine too flexible to hold the structure needed for am alpha helix.

Q2) Helical wheel question

Q3) Beta sheet/alpha helix counting from PDB file

Q4)

CHAPTER OVERVIEW

5: Protein Function

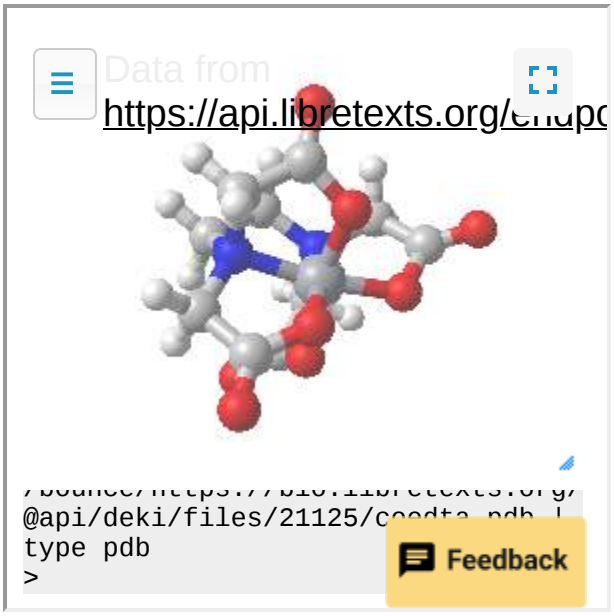
5.1: Binding - The First Step Towards Protein Function
5.2: Techniques to Measure Binding
5.3: Oxygen-Binding Proteins and Allosterism
5.4: Complementary Interactions between Proteins and Ligands - The Immune System and Immunoglobulins
5.5: Protein Interactions Modulated by Chemical Energy- Actin, Myosin, and Molecular Motors
5.6: Binding - Conformational Selections and Intrinsically Disordered Proteins
5.7: Binding - Enzyme Linked Immunosorbant Assays (ELISAs)

5.1: Binding - The First Step Towards Protein Function

5.1.1: Reversible Binding of a Ligand to a Macromolecule

Reversible, noncovalent binding of two or molecules is the first step in the expression of the biological properties of almost all biomacromolecules. If one of the molecules is small, it's often called a ligand. Ligands are often referred to by other names. Substrates are the reactants that bind to the active sites of enzymes. Hormones and neurotransmitters bind to solution phase or membrane-bound receptor proteins. Metal ions (simple like Ca^{2+} or molecular like $CH_3CO_2^-$) are also considered ligands when bound to proteins or nucleic acids.

You might be more familiar with the term ligand when it's applied to the coordination of a transition metal complex by electron pair donors (Lewis acids) on single or multidentate molecules, which for transition metal complexes are called ligands. Here is an interactive molecule model of a cobalt ion binding to EDTA, a multidentate ligand.



The cobalt ion (dark grey ball) is octahedrally-coordinated to the multidentate ligand EDTA.

Whether a macromolecule M and a ligand L bind to each other depends on their relative concentrations and how tightly they bind. Compare this to an acid. Its pKa and the pH of the medium determine if it deprotonates.

Biochemists rarely talk about equilibrium constants to describe the strength of a binding interaction, but rather their reciprocals - the dissociation constants, K_D. For the reactions $M + L \leftrightarrow ML$, where M is free macromolecule, L is free ligand, and ML is macromolecule-ligand complex (which is held together by intermolecular forces, not covalent forces), the K_D is given by

$$K_D = [M]_{eq}\,[L]_{eq} / [ML]_{eq} \tag{5.1.1}$$

Figure 5.1.1 shows free and bound M and L.

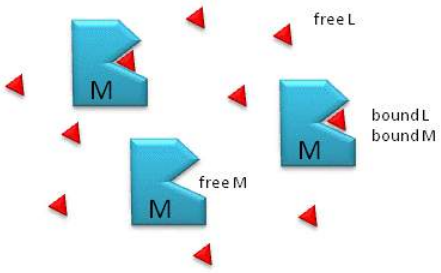


Figure 5.1.1: Free macromolecule (M), free ligand (red triangle) and bound complex ML

Notice the unit of K_D is molarity, M.

- The lower the K_D (i.e. the higher the [ML] at any given M and L), the tighter the binding.
- The higher the K_D, the looser the binding. K_Ds for biological molecules are finely tuned to their environments.

K_D values vary from about 1 mM (weak interactions) for some enzyme-substrate complex, to pM - fM levels. Examples of very tight, non-covalent interactions include the avidin (an egg protein)-biotin (a vitamin) and thrombin (enzyme initiating clotting)-hirudin (a leech salivary protein) complexes. The values are "tuned" so that the relative concentration of free and bound M and L are appropriate for a biological setting.

To understand binding, it is important not only to know the noncovalent, intermolecular forces (IMFs) that lead to binding, but also to ask the simple question: are the macromolecule and ligand bound? If so, to what extent? To know if M or L is bound, we must use simple mathematics that you would have learned in Introductory or Analytical Chemistry courses. We'll start with the mathematical description which is harder for students to understand than the IMFs.

We will start with three basic equations:

For the Dissociation constant:

$$K_D = ([M]eq[L]eq)/[ML]eq = ([M][L])/[ML] \tag{5.1.2}$$

(note that K_D has units of molarity);

For Mass Balance of M:

$$M_0 = M + ML \tag{5.1.3}$$

where M_0 is the total amount of macromolecule. (note: brackets and the eq subscript will be left off if the resulting equation is nonambiguous)

For Mass Balance of L:

$$L_0 = L + ML \tag{5.1.4}$$

where L_0 is the total amount of ligand

We would like to derive equations which give ML as a function of known or measurable values. The K_D equation (5.1) shows

that ML depends on free M and free L. From the equations above we can two derive two fundamental and equally valid equations which are useful under different experimental conditions.

5.1.1.1: Case 1:

This applies when you can readily measure **free** L OR when experimental conditions are such the Lo >> Mo (so L= Lo), which is often encountered in a lab setting. You don't have to measure free L since, for this case, it is approximately the total ligand that was added to the system.

Substitute 5.1.3 into 5.1.1 gives

$$\begin{aligned} K_D = ([M][L])/[ML] = [Mo - ML][L])\,/[ML \\ (ML)K_D = (M_o)\,L - (ML)L \\ (ML)K_D + (ML)L = (M_o)\,L \\ (ML)\,(K_D + L) = (M_0)\,L \end{aligned} \tag{5.1.5}$$

or

$$(ML) = \frac{(M_0)\,L}{K_D + L} \tag{5.1.6}$$

This equation is ALWAYS TRUE for the chemical equation written above. L is the free ligand concentration at equilibrium.

An interactive plot of the concentration of the ML complex (ML) vs free L (L) is shown below. Vary the sliders and note the changes in the graph.

If L_0 >> M_0, then the equations simplify to:

$$ML = \frac{(M_0)(L_0)}{K_D + L} \tag{5.1.7}$$

Dividing this equation by Mo gives the fractional saturation Y of the macromolecule M.

$$Y = [ML]/M_0 = \frac{L}{K_D + L} \tag{5.1.8}$$

where Y can vary from 0 (when L = 0) to 1 (when L >> K_D)

Note that the interactive graph above and graphs of ML vs L (equation 5.1.10) and Y vs L (equation 5.1.11) are all HYPERBOLAs

To get a "gut" level understanding of the graphs of $(ML) = (M_0)(L)/(K_D + L)$ and $Y = L/(K_D + L)$, let's consider 3 different values or sets of values of free ligand:

1. L = 0: This obviously gives ML = 0
2. L = K_D: $(ML) = (M_0)(L)/(L + L) = (M_0)(L)/(2L) = Mo/2$ which indicates that M is half saturated. In fact the operational definition of K_D is the ligand concentration at which the M is half saturated.
3. L >> K_D: ML = M_0 and the macromolecule is saturated with ligand.

5.1.1.2: Case 2 (more general):

This applies when you know K_D, but don't know free L or haven't measured it, and you just wish to calculate how much ML is present at equilibrium, given a K_D value. In this case, L_0 does not have to be much greater than M_0. If where, like it is often in an experimental system, you would know that free L = L_0 and you could use Case 1.

In this case, we will substitute mass balance equations for both M_0 (Eq 5.1.2) and L_0 (Eq 5.1.3)and into the equation for K_D (Eq. 5.1.1). This gives:

$$\begin{gathered} K_D = ([M][L])/[ML] = [M_0 - ML]\,[L_0 - ML]\,/[ML] \\ (ML)K_D = (M_0 - ML)\,(L_0 - ML) \\ (ML)K_D = (M_0)\,(L_0) - (ML)\,(L_0) - (ML)\,(M_0) + (ML)^2 \end{gathered} \tag{5.1.9}$$

or

$$(ML)^2 - (L_0 + M_0 + K_D)\,(ML) + (M_0)\,(L_0) = 0 \tag{5.1.10}$$

This can be rearranged into the form $ax^2 + bx + c = 0$ where

- a = 1
- b = - (L_0 + M_0 +K_D)
- c = (M_0)(L_0)

with the well known solution $x = [(-b) - (b^2 - 4(a)(c))^{1/2}]/2a$. Therefore,

$$(ML) = \left[(L_0 + M_0 + K_D) - \left((L_0 + M_0 + K_D)^{\,2} - 4\,(M_0)\,(L_0)\right)^{\,1/2}\right]/2 \tag{5.1.11}$$

An interactive plot of the Y, fractional saturation, vs total L (L_0) is shown below. Vary the sliders and note the changes in the graph.

In the derivations, we came up with two equations for ML, Eq 5.1.10 which gives ML vs L and Eq 5.1.16 which gives ML vs L_0.

Both equations are valid. In the first, you must know free L which is often L_0 if M_0 << L_0. In the second, you don't need to know free M or L at all. At a given Lo, Mo, and K_D, you can calculate ML, which should be the same ML you get from the first equation if you know free L.

Equations 5.1.10 and 5.1.16 are useful in several circumstances. They can be used to

- calculate the concentration of ML if K_D, M_0, and L (for Eq. 5.1.10) or if K_D, M_0, and L_0 (for Eq. 5.1.16) are known. This is analogous to the use of the Henderson-Hasselbalch equation to calculate the protonation state (HA) and hence the charge state of an acid at various pH values. In the former bind case, we are measuring the concentration of a reversibly bound ligand (ML) and in the latter case, the concentration of covalently bound protons (HA).
- calculate K_D if ML, M_0, and L (for Eq. 5.1.10) or if ML, M_0, and L_0 (for Eq. 5.1.16) are known. Techniques to extract the K_D from binding data will be discussed in A separate chapter section.

5.1.2: Interpretation of Binding Analyzes

It is important to get a mathematical understanding of the binding equations and graphs. It is equally important to get an intuitive understanding of their properties. Just as we used the +/- 2 pH rule in determining at a glance the charge state of an acid, you need to be able to determine the extent of binding (how much of M is bound with L) given their relative concentrations and the K_D. The usual situation is that $[M_0]$ is << $[L_0]$. What happens to the binding curves for M + L ↔ ML if the K_D gets progressively lower? Intuitively, you should expect that binding will increase, especially as L gets greater. The curves below should help you develop the intuition you need with respect to binding equilibria. Figure 5.1.2 show Y vs L_0 at Varying K_Ds

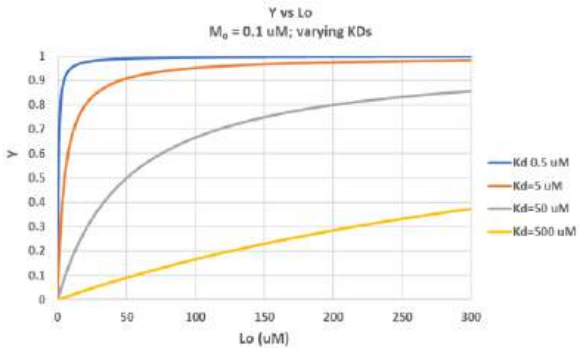


Figure 5.1.2: Fractional saturation Y vs L_0 at Varying K_Ds

Figure 5.1.3 shows Y vs L_0 at a very low K_D (0.001 uM = 1 nM, resulting in a sharp "titration" curve. Any increment of L added is bound so effectively none is present in the free form. The graph abruptly changes to a horizontal line when all the macromolecule is bound. This curve could be used to determine $[M_0]$!

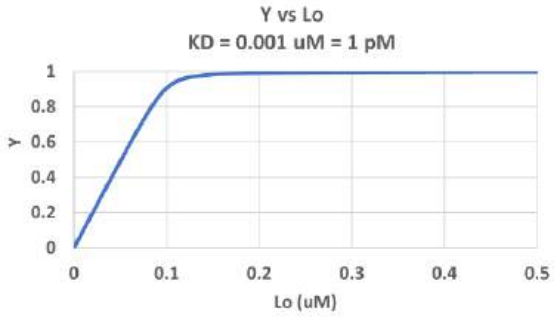


Figure 5.1.3: Fractional saturation Y vs L_0 for very low K_D binding interactions

Note that in the last graph, given the same M_0 and L_0 concentrations, the "titration curves" for a binding equilibrium characterized by even tighter binding (for example, a K_D = 0.1 pM or 0.01 pM) would be indistinguishable from the graph when K_D = 1 pM. It should be apparent that for all of these K_D values, all of the added ligand is bound until $[L_0] > [M_0]$. To differentiate these cases, much lower ligand concentrations would be required such that on the addition of ligand, all is not bound. Also note that this curve is NOT hyperbolic, which makes sense since the graph is of Y vs L_0, not Y vs L, and since L_0 is not >> M_0.

The interactive graph below shows fractional saturation Y vs L at two different K_D values

It is quite interesting to compare graphs of Y (fractional saturation) vs L (free) and Y vs Lo (total L) in the special case when L_0 is not >> M_0. Figure 5.1.4 when M_0 = 4 μM, Kd = 0.19 μM . Under the ligand concentration used, it should be apparent the L can't be approximated by L_0

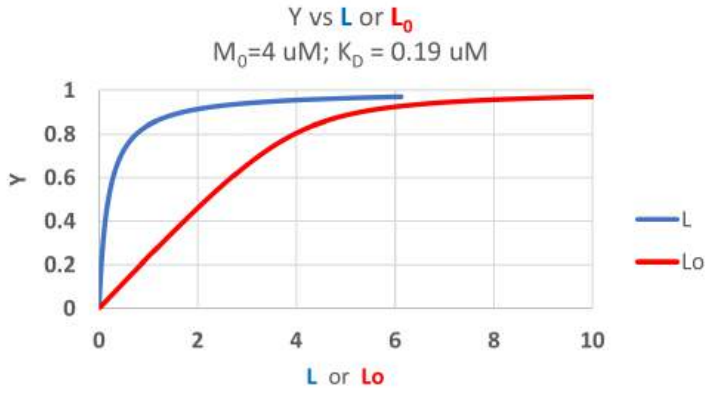


Figure 5.1.4: Fractional saturation Y vs L or L_0 when L_0 is not >> M_0

Two points should be evident from these graphs when L is not approximated by Lo:

- a graph of Y vs L_0 is not truly hyperbolic, but it does saturate
- a K_D value (ligand concentration at half-maximal binding) can not be estimated by inspection from the Y vs L_0, but it can be from the Y vs L graph.

Figure 5.1.5 shows a comparison of the extent of covalent binding of a proton to an acid at pH values around the pKa and by analogy the extent of noncovalent binding of a ligand at log[L] values around the log K_D.

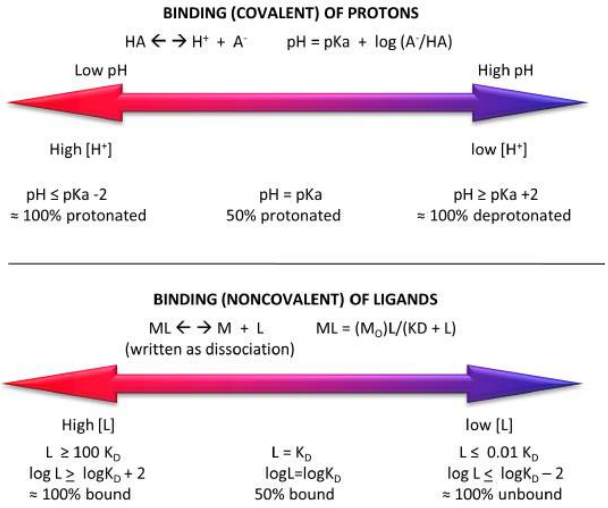


Figure 5.1.5: Analogy between covalent binding of protons and noncovalent binding of ligands to a molecule

5.1.3: Different Graphical Analyzes of Binding

In addition to the hyperbolic plots of [ML] vs [L] and fractional saturation Y vs [L], a variety of derivative plots are often encountered. The equations and their graphs (for two different K_D values, are shown below. The graphs are in the form of Y vs L_0, when L_0 is approximately equal to free L.

Hyperbolic saturation plot:

$$\mathrm{Y} = \frac{\mathrm{L}}{\mathrm{K_D} + \mathrm{L}} \tag{5.1.12}$$

Double reciprocal plot:

$$\frac{1}{\mathrm{Y}} = \frac{\mathrm{K_D} + \mathrm{L}}{\mathrm{L}} = \frac{\mathrm{K_D}}{\mathrm{L}} + 1 = \mathrm{K_D}\left(\frac{1}{\mathrm{L}}\right) + 1 \tag{5.1.13}$$

A plot of 1/Y vs 1/L has a slope of K_D and a y intercept of 1 (which is the number of binding sites for this simple mechanism)

The Scatchard plot:

$$\begin{aligned} \mathrm{Y}\,(\mathrm{K_D} + \mathrm{L}) &= \mathrm{L} \\ Y\,(\,\mathrm{K_D}) + YL &= L \\ Y\,(\,\mathrm{K_D}) = L - \mathrm{YL} &= \mathrm{L}(1 - \mathrm{Y}) \end{aligned} \tag{5.1.14}$$

which gives the final Scatchard plot equation:

$$\frac{Y}{\mathrm{L}} = \frac{1 - \mathrm{Y}}{\mathrm{K_D}} = -\frac{\mathrm{Y}}{\mathrm{K_D}} + \frac{1}{\mathrm{K_D}} \tag{5.1.15}$$

A plot of Y/L vs Y has a slope of -1/K_D and a y intercept of 1/K_D.

Y vs logL

Plotting Y vs **L** give a hyperbola, but a plot of Y vs **log L** give a **sigmoidal** plot. Plots of Y vs log L are often used in the research literature instead of traditional hyperbolic plots of Y vs L. There are several reasons for this:

- the log [L] is more fundamentally related to the thermodynamic expression that relates ΔG^0 and K_{eq} or K_D, namely

$$\Delta G^0 = -RT \ln K_{eq} = R\,\mathrm{T} \ln K_D \tag{5.1.16}$$

- plots of Y vs L plateau over a very large range of [L], but given the compression of the X axis values in a semilog plot, the plots reach a saturation plateau over a much narrower range of log [L]. A range from 1-100 on the [L] scale becomes 0-2 on the log

[L] scale. Since it takes a very high concentration of ligand to truly reach saturation (100xK_D), it's much easier to see if saturation is achieved in semilog plots.

- multiple plots of binding data for different K_D values have exactly the same shape on a semilog plot. Binding data for a ligand to a wild type and mutant proteins, all with different K_Ds, will give identical plots with curves for higher K_D values shifted to the right. Semilog plots are also used routinely to display multiple plots of binding in the absence and presence of a binding inhibitor.

Figure 5.1.6 show different graphs for ligand binding to a macromolecule

5.1.4:

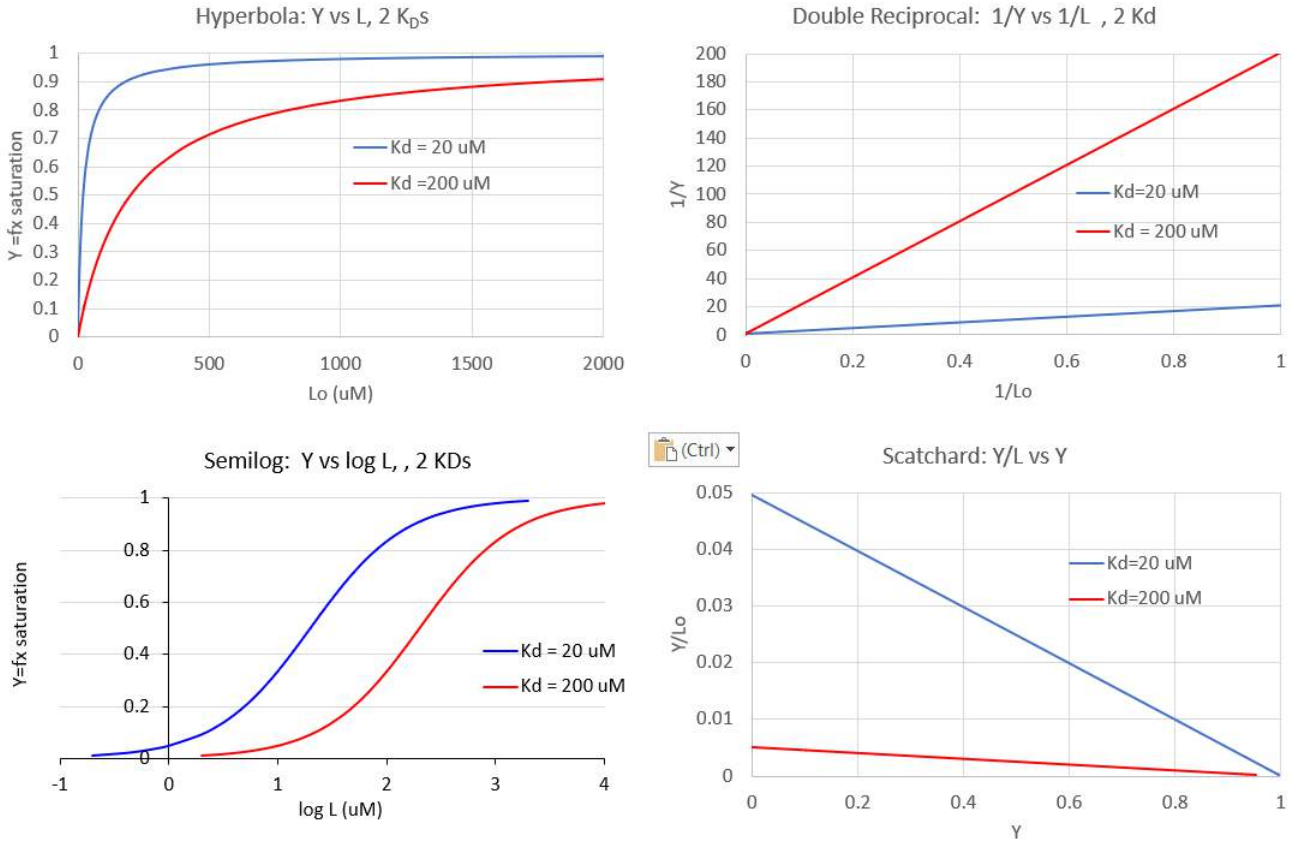


Figure 5.1.6: Different graphs for ligand binding to a macromolecule

Sigmoidal binding curves: A note of caution

The graph of Y vs log [L] is sigmoidal but the same data would give a hyperbola if plotted as Y vs [L]. However, as we will see in section 5.3, there are some occasions when the graph of Y vs **[L]** is sigmoidal. For example, this can occur when the binding of a ligand to a multimeric binding protein affects the binding of the ligand to additional sites on the protein. This is an example of allosteric binding, which we will explore in great detail in section 5.3. So if you see a sigmoidal plot, be careful to examine the graph to see if it is a regular or semilog plot.

Straight line transformations of the hyperbolic binding equations are useful to get approximate values of K_D, but linear regression analysis to get slopes and intercepts is not statistically optimal as the errors in the y variable (Y) and in the y and x variables in the Scatchard plot are not identical across values. To determine K_D, it is best to fit experimental data to the nonlinear function for the hyperbola.

5.1.5: Dimerization and Multiple Binding Sites

In the previous examples, we considered the case of a macromolecule M binding a ligand L at a single site, as described in the equation below:

M + L ↔ ML

where K_D = [M][L]/[ML]

We saw that the binding curves (ML vs L or Y vs L are hyperbolic, with a K_D = L at half-maximal binding. But there are many other chemical equilibria than can mechanistically explain binding data. We'll consider just two cases here.

Dimerization

A special, yet common example of this equilibrium occurs when a macromolecule binds itself to form a dimer (D), as shown below:

M + M ↔ M_2 or D

where D is the dimer, and where

$$K_D = [M][M]/[D] = [M]^2/[D] \tag{5.1.17}$$

At first glance, you would expect a graph of [D] vs [M] to be hyperbolic, with the K_D again equaling the [M] at half-maximal dimer concentration. This turns out to be true, but a simple derivation is in order. In the case of dimer formation, Mo, which superficially represents both M and L in the earlier derived expression, are both changing. So we have to invoke mass balance of M again: $[Mo] = [M] + 2[D]$, where the coefficient 2 is necessary since there are 2 M in each dimer.

More generally, for the case of the formation of trimers (Tri), tetramers (Tetra), and other oligomers, $[Mo] = [M] + 2[D] + 3[Tri] + 4[Tetra] + \ldots$.

Rearranging (12) and solving for D gives $D = ([M_0] - [M])/2$. Substituting this into the K_D expression (1) gives

$$K_D = \frac{M^2}{\frac{[M_0]-[M]}{2}} = \frac{2M^2}{[M_0]-[M]} = \tag{5.1.18}$$

This can be rearranged into quadratic form for M (not D):

$$2M^2 + K_D(M) - K_D\,(M_0) = 0 \tag{5.1.19}$$

which is of the form $y = ax^2+bx+c$.

Solving the quadratic equation gives [M] and with M_0, D can be calculated from $D = ([M_0] - [M])/2$.

A value Y, similar to fractional saturation, can be calculated, where Y is the fraction of total possible D, which can vary from 0-1: $Y = 2D/M_0$

A graph of Y vs Mo with a dimerization dissociation constant K_D = 25 uM, is shown in Figure 5.1.7.

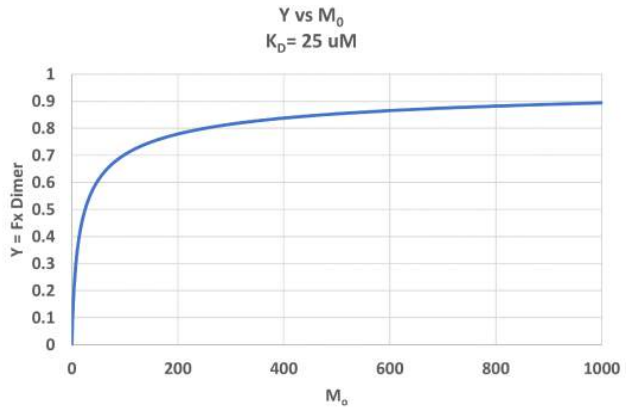


Figure 5.1.7: Fractional saturation Y vs M_0 for protein dimerization

Note that the curve appears somewhat hyperbolic. Half-maximal dimer formation does occur at a total M concentration $M_0 = K_D$. Also note, however, that even at M_0 = 1000 uM, which is 40x K_D, only 90% of the total possible D is formed (Y = 0.90). For the simple M + L ↔ ML equilibrium, if L_0 = 40x the K_D and $M_0 << L_0$, $Y = L/(K_D + L) = L/[(L/40) + L] = 0.976$

An interactive graph showing Y (the fraction of dimers) vs M_0 is show. Move the sliders to show how changes in "K_D" affect the dimerization.

The aggregation state of a protein monomer is closely linked with its biological activity. For proteins that can form dimers, some are active in the monomeric state, while others are active as a dimer. High concentrations, found under conditions when proteins are crystallized for x-ray structure analysis, can drive proteins into the dimeric state, which may lead to the false conclusion that the active protein is a dimer. Determination of the actual physiological concentration of [Mo] and K_D gives investigators knowledge of the Y value which can be correlated with biological activity. For example, interleukin 8, a chemokine, which binds certain immune cells, exists as a dimer in x-ray and NMR structural determinations, but as a monomer at physiological concentrations. Hence the monomer, not the dimer, binds its receptors on immune cells. Viral proteases (herpes viral protease, HIV protease) are active in dimeric form, in which the active site is formed at the dimer interface.

Binding of a ligand to two independent sites

What if a ligand L binds to two different sites on the same biomacromolecule? Assuming that the binding of ligand L to one site does not affect the binding of the ligand to the other site (and vice versa), the following equation can be simply derived:

$$Y = \left[\frac{L}{(K_{D1} + L)} + \frac{L}{(K_{D2} + L)} \right] /2 \tag{5.1.20}$$

The numerator of the equation has a term for the fractional saturation of site 1 characterized by K_{D1} and a term for the fractional saturation of site 2 characterized by K_{D2}. These two terms are divided by 2 so that the fractional saturation of all sites is 1 at

saturating values of free ligand. Note that there can only be one free ligand concentration in solution so the only thing that differs in the two terms in the numerator is the K_D value.

The interactive graph below shows such binding to two independent sites with different K_D values. Again, we'll assume the binding of one ligand does NOT influence the binding of the other.

5.1.6: The Binding Continuum

Binding affinities give us a way to measure the relative strength of binding between two substances. But how "tight" is tight binding? Weak binding? Let us exam that issue by considering a binding continuum. Consider two substances, A and B that might interact. Over what range of strengths can they actually bind to each other? It would helpful to set up the extremes of the binding continuum. At one end is no binding at all. At the other end, consider two things that bind covalently. We have discussed how Kd reflects binding strength. Remember, $K_D = 1/K_{eq}$. Also, we know that K_{eq} is related to ΔG^o by the equations:

$$\Delta G^0 = -R\,\mathrm{T}\ln K_{eq} = R\,\mathrm{T}\ln K_D \tag{5.1.21}$$

Given these simple equations, you should be able to interconvert between K_{eq}, K_D, and ΔG^0. (Keep your units straight.).

No interaction: One end of the binding continuum represents no interaction. Let's assume that K_{eq} is tiny (K_D large), for example $K_{eq} \sim 2.4 \times 10^{-72}$. Plugging this into the equation \(ΔG^0 = - RTlnK_{eq}\), where R = 2.00 cal/mol·K, and T is about 300K, the $\Delta G^0 \sim +100$ kcal/mol (418 kJ/mo). That is, if we add A + B, there is no drive to form AB. If AB did form, then it would immediately fall apart.

Covalent interaction: At the other end of the continuum consider the interaction of 1H atom with another to form H_2. From a general chemistry book we can get ΔG^0_{form}. Using simple thermodynamics, we can calculate ΔG^o for H-H formation. ($\Delta G^o = \Sigma \Delta G^0_{form}$ prod. - $\Sigma \Delta G^0_{form}$ react.) Doing this gives a value of -97 kcal/mol (-406 kJ/mol).

Specific and Nonspecific Binding: Consider the interaction of a protein, the lambda repressor (R), with a small oligonucleotide to which it binds tightly (called the operator DNA, O). This is an example of a biologically tight, but reversible interaction. R can bind to many short oligonucleotides due to electrostatic interactions and H bonds between the positively charged protein and the negatively charged nucleic acid backbone. The tight binding interaction, however, involves oligonucleotides of a specific base sequence. Hence we can distinguish between tight binding, which usually involves specific DNA sequences and weak binding which involves nonspecific sequences. Likewise, we will speak of specific and nonspecific binding. R and O, which bind with a K_D of 1 pM, represent an example of specific binding, while R and nonspecific DNA (D), which bind mostly through electrostatic interactions with a K_D of 1 mM, are an example of nonspecific binding. You might expect any positively charged protein, like mitochondrial cytochrome C, would bind negatively charged DNA. This nonspecific interaction would presumably have no biological significance since the two are localized in different compartments of the cell. In contrast, the interaction between positively charged histone proteins, bound to DNA in the nucleus, would be specific.

Rate constants for association and dissociation: When the reaction
M + L ↔ ML is at equilibrium, the rate of the forward reaction is equal to the rate of the reverse reaction. From General Chemistry, the forward reaction is biomolecular and second order. Hence the v_f, the rate in the forward direction is proportional to [M][L], or $v_f = k_f[M][L]$, where k_f is the rate constant in the forward direction. The rate of the reverse reaction, v_r is first order, proportional to [ML], and is given by $v_r = k_r[ML]$, where k_r is the rate constant for the reverse reaction. Notice that the units of k_f are $M^{-1}s^{-1}$, while units of k_r are s^{-1}. At equilibrium, $v_f = v_r$, or

$$k_f[M][L] = k_r[ML] \tag{5.1.22}$$

Rearranging the equation gives

$$[ML]/[M][L] = k_f/k_r = K_{eq} \tag{5.1.23}$$

Hence K_{eq} is given by the ratio of rate constants. For tight binding interactions, $K_{eq} >> 1$, $K_D << 1$, and k_f is very large (in the order of 10^{8-9}) and k_r must be very small (10^{-2} - 10^{-4} s^{-1}).

To get a more intuitive understanding of K_Ds, it is often easier to think about the rate constants which contribute to binding and dissociation. Let us assume that k_r is the rate constant that describes the dissociation reaction. It is often times called k_{off}. Likewise, k_f is often called the on rate (k_{on}). It can be shown mathematically that the rate at which two simple molecules associate depends on their radius and effective molecular weight. The maximal rate at which they will associate is the maximal rate at which diffusion will lead them together. Let us assume that the rate at which M and L associate is diffusion limited. The theoretical k_{on} is about 10^8 $M^{-1}s^{-1}$. Knowing this, the K_D and the fact that $k_{on}/ k_{off} = K_{eq} = 1/K_D$, we can calculate k_{off}, which is a first-order rate constant.

We can also determine k_{off} experimentally. Imagine the following example. Adjust the concentrations of M and L such that Mo << Lo and Lo>> Kd. Under these conditions of ligand excess, M is entirely in the bound from, ML. Now at t = 0, dilute the solution so that Lo << Kd. The only process that will occur here is dissociation, since negligible association can occur given the new condition. If you can measure the biological activity of ML, then you could measure the rate of disappearance of ML with time, and get k_{off}. Alternatively, if you could measure the biological activity of M, the rate at which activity returns will give you k_{off}.

Now you will remember from Introductory Chemistry that for a first-order rate constant, the half-life ($t_{1/2}$) of the reaction can be calculated by the expression: $k = 0.693/t_{1/2}$. Hence given k_{off}, you can determine the $t_{1/2}$ for the associated specie's existence. That is, how long will a complex of ML last before it dissociates? Given ΔG^o or K_D, and assuming a k_{on} (10^8 $M^{-1}s^{-1}$), you should be able to calculate k_{off} and $t_{1/2}$. Or, you could be able to determine k_{off} experimentally, and then calculate $t_{1/2}$. Applying these principles, you can calculate the binding parameters. Table 5.1.1 below shows calculated k_{off} and $t_{1/2}$ for binary complexes assuming diffusion-controlled k_{on}.

Complex	K_D (M)	k_{off} (s^{-1})	t½
H_2	1×10^{-71}	1×10^{-63}	2×10^{55} yr
RtV3 : Rt'L3(a)	10^{-17}	1×10^{-9}	2 yr
Avidin:biotin	10^{-15}	1×10^{-7}	80 days
thrombin:hirudin(b)	5×10^{-14}	5 x 10-6	2 days

lacrep:DNAoper(c)	1 x 10^{-13}	1 x 10^{-5}	0.8 days
Zif268:DNA(d)	10^{-11}	1 x 10^{-3}	700 s
GroEL:r-lactalbumin(e)	10^{-9}	0.1	7 s
TBP:TATA(f)	2 x 10^{-9}	2 x 10-1	3 s
TBP:TBP	4 x 10^{-9}	4 x 10^{-1}	2 s
LDH (pig): NADH(g)	7.1x10^{-7}(j)	7.1 x 10^{1}	10 ms
profilin: CaATP-G-actin	1.2 x 10-6	1.2 x 10^{2}	6 ms
TBP: DNAnonspec(h)	5 x 10^{-6}	5 x 10^{2}	1 ms
TCR(i): cyto C peptide	7X10^{-5}	7X10^{3}	100 us
lacrep:DNAnonspec(h)	1 x 10^{-4}	1 X10^{4}	70 us
uridine-3P: RNase	1.4x10^{-4} (j)	1.4X104	50 us
Creatine Kinase: ADP	8.2x10-4 (j)	8.2X10^{4}	10 us
Acetylcholine:Esterase	1.2 x 10^{-3}	1.2 x 10^{5}	6 us
no interaction	4 x 10^{73}	4 x 10^{81}	-

Table 5.1.1: Calculated k_{off} and $t_{1/2}$ for binary complexes assuming diffusion-controlled k_{on}

a. Trivalent Vancomycin derivative RtV3 + Trivalent D-Ala-D-Ala deriv, Rt'L3'
b. Hirudin is a potent thrombin inhibitor from leach saliva
c. lac rep is the E. Coli lac operon repressor protein, and DNAoper is the specific DNA binding region in the E. Coli genome that binds to the repressor
d. Zif268 is a mouse zinc-finger binding protein
e. GroEL is a chaperone protein; r-lactalbumin is the reduced form of lactalbumin
f. TBP is the TATA Binding Protein that binds to the TATA box consensus sequence
g. LDH is lactate dehydrogenase
h. DNAnonspec is DNA which does not contain the specific DNA sequence region involved in specific binding to a DNA binding protein
i. TCR is the T-cell receptor
j. calculated from equation: KD = k_{off}/k_{on}.

What is usually measured is K_D and/or k_{off} (if the k_{off} is reasonable). This analysis is very simplified. Electrostatic forces and other orientation factors may significantly change k_{on}, while conformational changes in the complex may prevent ready unbinding of the bound ligand, dramatically altering k_{off}.

Figure 5.1.8 shows an interactive iCn3D model of one of the tightest binding complexes, avidin and biotin (2avi).

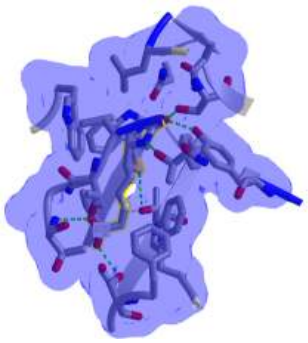


Figure 5.1.8: Avidin-Biotin Complex (2avi) (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...ZssSe5GoUfr9Q6

The blue shows the surface around the complex, which shows that the avidin in completely buried. Hydrogen bonds are shown to biotin (labeled as BTN) in green dashes.

It is important to note that even reactions characterized by high K_D (weak binding) can be specific. Specificity is ultimately defined as a binding interaction between a macromolecule and ligand that can be co-localized in the same environment and for which a biological function is elaborated upon binding.

5.2: Techniques to Measure Binding

It is often essential to determine the K_D for a ML complex since given that number and the concentrations of M and L in the system, we can then predict if M is bound under physiological conditions. Again, this is important since whether M is bound or free will govern its activity. To determine K_D, you need to determine ML and L at equilibrium. How can we differentiate free from bound ligand? The following techniques allow such a differentiation.

5.2.1: Techniques that require the separation of bound from the free ligand.

Care must be given to ensure that the equilibrium of M + L ↔ ML is not shifted during the separation technique.

Gel filtration chromatography

Add M to a given concentration of L. Then elute the mixture on a gel filtration column, eluting with the free ligand at the same concentration. The ML complex will elute first and can be quantitated. If you measure the free ligand coming off the column, it will be constant after the ML elutes, with the exception of a single dip near where the free L would elute if the column were eluted without free L in the buffer solution. This dip represents the amount of ligand bound by M.

Membrane filtration

Add M to radiolabelled L, equilibrate, and then filter through a filter that binds M and ML. For instance, a nitrocellulose membrane binds proteins irreversibly. Determine the amount of radiolabeled L on the membrane, which equals [ML].

Precipitation

Add a precipitating agent like ammonium sulfate, which precipitates proteins (both M and ML). Determine the amount of ML.

5.2.2: Techniques that do not require the separation of bound from free ligand.

Equilibrium dialysis

Place M in a dialysis bag and dialyze against a solution containing a ligand whose concentration can be determined using radioisotopic or spectroscopic techniques. At equilibrium, determine free L by sampling the solution surrounding the bag. By mass balance, determine the amount of bound ligand, which for a 1:1 stoichiometry gives ML. Repeat at many different ligand concentrations

Spectroscopy

Find a ligand whose absorbance or fluorescence spectra changes when bound to M. Alternatively, monitor a group on M whose absorbance or fluorescence spectra changes when bound to L.

Isothermal titration calorimetry (ITC)

In ITC, a high-concentration solution of an analyte (ligand) is injected into a cell containing a solution of a binding partner (typically a macromolecule like a protein, nucleic acid, or vesicle). Figure 5.2.1 shows an isothermal titration calorimeter cell,

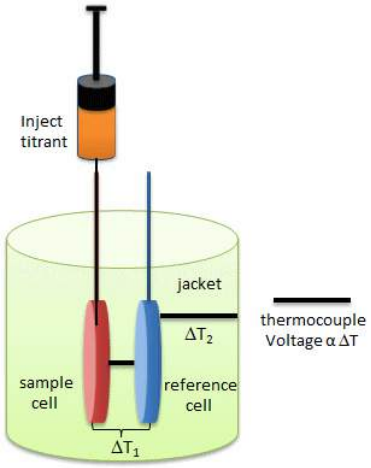


Figure 5.2.1: Isothermal Titration Calorimeter Cell

On binding, heat is either released (exothermic reaction) or adsorbed, causing a small temperature change in the sample cell compared to the reference cells containing just a buffer solution. Sensitive thermocouples measure the temperature difference (ΔT_1) between the sample and reference cells and apply a current to maintain the difference at a constant value. Multiple injections are made until the macromolecules are saturated with the ligand. The enthalpy change is directly proportional to the amount of ligand bound at each injection, so the observed signal attenuates with time. The actual enthalpy change observed must be corrected for the change in enthalpy on simple dilution of the ligand into buffer solution alone, determined in a separate experiment. The enthalpy changes observed after the macromolecule is saturated with the ligand should be the same as the enthalpy of dilution of the ligand. A binding curve showing enthalpy change as a function of the molar ratio of ligand to binding partner (L_0/M_0 if $L_0 >> M_0$) is then made and mathematically analyzed to determine K_D and the stoichiometry of binding. Figure 5.2.2 shows a typical isothermal titration calorimetry data and analysis

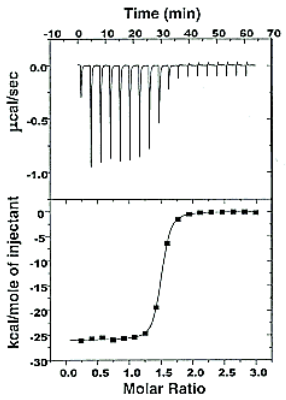


Figure 5.2.2: Typical isothermal titration calorimetry data and analysis. http://www.microcalorimetry.com/index.php?id=312

It should be evident in the example above that the binding reaction is exothermic. But why is the graph of ΔH vs the molar ratio of L_0/M_0 sigmoidal (s-shaped) and not hyperbolic? One clue is that the molar ratio of ligand (titrant) to macromolecule centers around one, so, as explained above, when L_0 is not >> M_0, the graph might not be hyperbolic. The graphs below show a specific example of a K_D and ΔH_0 calculated from the titration calorimetry data. We will use them to show why the graph of ΔH vs molar ratio of L_0/M_0 is sigmoidal.

A specific example illustrates these ideas. A soluble version of the HIV viral membrane protein, gp 120 (4 µM), was placed in the calorimetry cell. A form of its natural ligand, CD4, a membrane receptor protein from T helper cells, was placed in the syringe and titrated into the cell (Myszka et al. 2000). Enthalpy changes/injection were determined. The data was transformed and fit to an equation that shows the ΔH "normalized to the number of moles of ligand injected at each step". Fig."e 5.2.3 shows the raw data (top) and the best-fit model (bottom), assuming a 1:1 stoichiometry of CD4 (the "ligand") to gp 120 (the "macromolecule") and a K_D = 190 nM.

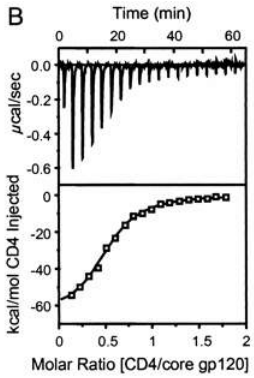


Figure 5.2.3: Titration Calorimetry determination of K_D and ΔH for the interaction of gp120 and CD4

Note that the bottom curve is sigmoidal, not hyperbolic. A single experiment can determine the stoichiometry of binding (n), the K_D, and the ΔH_0. From the value of ΔH^o and K_D and the relationship $\Delta G^o = -RTlnK_{eq} = RTlnK_D = \Delta H_0 - T\Delta S^0$, the ΔG^0 and ΔS^0 values can be calculated. No separation of bound from free is required. Enthalpy changes on binding were calculated to be -62 kcal/mol (260 kJ/mol).

A series of plots can be derived using the standard binding equations (5, 7, and 10 above) to calculate free L and ML at various Lo concentrations and R = L_0/M_0 ratios. Two were shown earlier to illustrate differences in Y vs L and Y vs L_0 when L_0 is not >> M_0. They are shown again below in Figure 5.2.4.

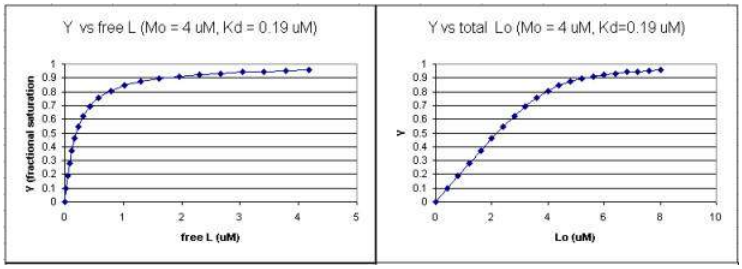


Figure 5.2.4: Y vs L and Y vs L_0 when L_0 is not >> M_0

A more precise understanding of the data analyses is shown in Figure 5.2.5, which shows a plot of ML vs R (= $[L_0]/[M_0]$ (panel A1 right) was made. This curve appears hyperbolic, but it has the same shape as the Y vs L_0 graph shown in Figure 5.2.4 (right). However, if the amount of ligand bound at each injection (calculated by subtracting [ML] for injection i+1 from [ML] for injection i) is plotted vs R (= $[L_0]/[M_0]$), a sigmoidal curve shown in Figure 5.2.5 (panel A2, left) is seen, which resembles the best-fit graph for the experimentally determine enthalpies in Figure 5.2.4. The relative enthalpy change for each injection is shown in red. Note the graph in Figure 5.2.5 (A2) shows the negative of the amount of ligand bound per injection to make the graph look the that in the graph showing the actual titration calorimetry trace and fit above.

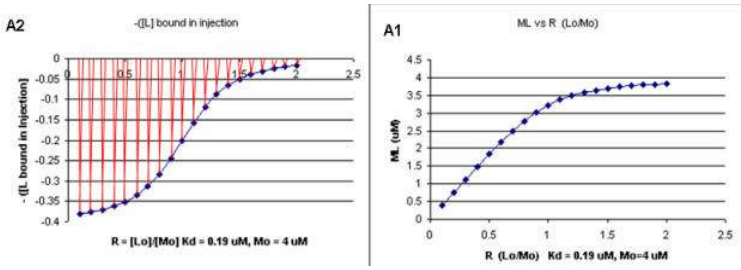


Figure 5.2.5: Binding Curves that Explain Sigmoidal Titration Calorimetry Data for gp120 and CD4

Surface Plasmon Resonance

A newer technique to measure binding is called surface plasmon resonance (SPR), using a sensor chip consisting of a 50 nm layer of gold on a glass surface. A carbohydrate matrix is then added to the gold surface. A macromolecule that contains a binding site

for the ligand is covalently attached to the matrix. The binding site on the macromolecule must not be perturbed significantly. A liquid containing the ligand is passed over the binding surface.

The detection system consists of a light beam that passes through a prism on top of the glass layer. The light is reflected, but another component of the wave, called an evanescent wave, passes into the gold layer, where it can excite the Au electrons. If the correct wavelength and angle are chosen, a resonant wave of excited electrons (plasmon resonance) is produced at the gold surface, decreasing the total intensity of the reflected wave. The angle of the SPR is sensitive to the layers attached to the gold, and binding and dissociation of the ligand are sufficient to change the SPR angle, as seen in Figure 5.2.6.

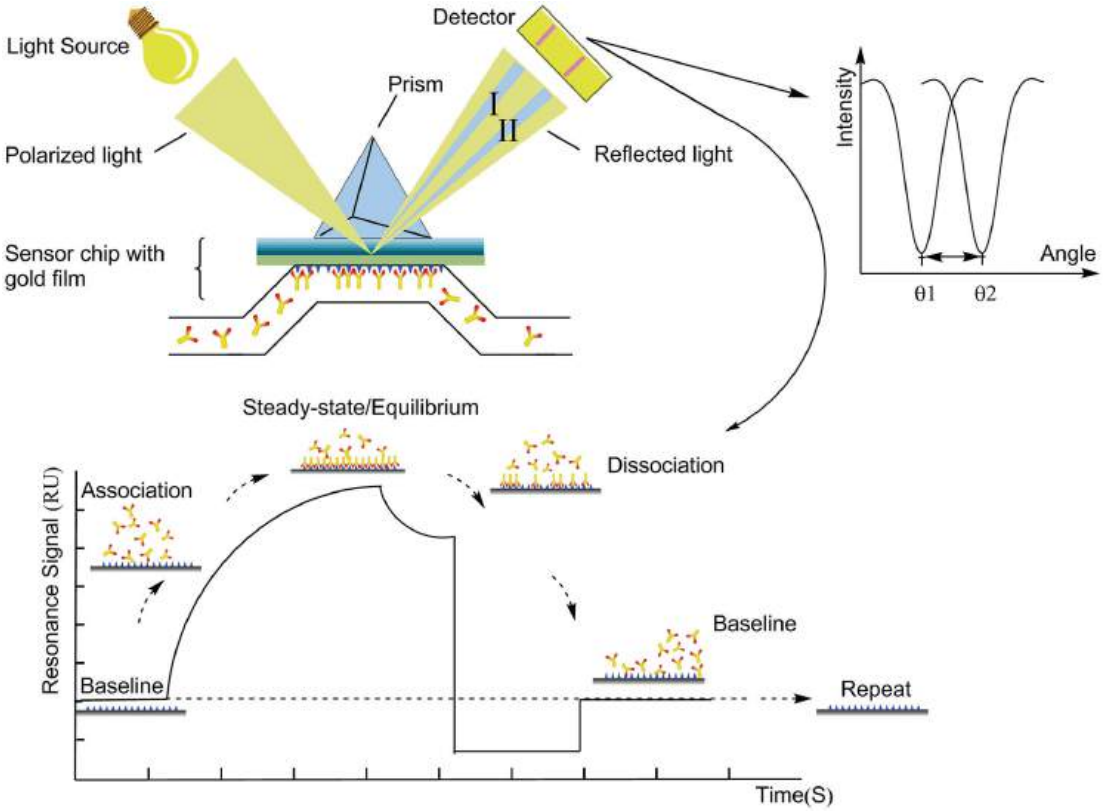


Figure 5.2.6: Surface plasmon resonance (SPR) system. SPR detects changes in the refractive index near the surface layer of a sensor chip. The sensor surface is gold with antibodies attached to it. During the measurement, the chip is irradiated from the bottom with a beam of a wide angle range within that of total internal reflection. The SPR angle shifts (from I to II in the diagram) when biomolecule binding events cause changes in the refractive index at the surface layer. The detector will determine the angle of the intensity decrease. This change in resonant angle can be monitored non-invasively in real-time as a plot of resonance signal (proportional to mass change) versus time. Song, Chengcheng & Zhang, Shaocun & Huang, He. (2015). Choosing a suitable method for the identification of replication origins in microbial genomes. Frontiers in MICROBIOLOGY. 6. 10.3389/fmicb.2015.01049. DOI: 10.3389/fmicb.2015.01049. License CC BY 4.0

This technique can distinguish fast and slow binding/dissociation of ligands (as reflected in on and off rates) and be used to determine K_D values (through measurement of the amount of ligand bond at a given total concentration of ligand or more indirectly through the determination of both k_{on} and k_{off}.

Binding DB: a database of measured binding affinities, focusing chiefly on the interactions of proteiproteinsdered to be drug targets with small, drug-like molecules

PDBBind-CN: a comprehensive collection of the experimentally measured binding affinity data for all biomolecular complexes deposited in the Protein Data Bank (PDB).

5.2.3: Extreme Binding Affinities

An incredibly tight binding interaction has recently been reported for the binding of Cu1+ to the CueR protein from E. Coli. Cu^{1+} ions are usually kept at very low concentrations in cells to prevent toxicity. Yet some enzymes require Cu. Free copper ions must be present in the cell to allow binding to appropriate sites in proteins. How are these competing concerns regulated in the cell? The

total Cu concentration in E. Coli is about 10 μM (10,000 nM), which, given the small size of the bacterium, represents about 10,000 copper ions per cell.

Cells have evolved many mechanisms to control and deliver Cu ions. Copper ions can be delivered to target proteins by copper chaperones (analogs of the chaperone proteins which guide protein folding). CueR in E. Coli appears to regulate the copper-induced expression of genes involved in copper biochemistry (including an enzyme that oxidizes Cu^{1+} to $Cu^{2+,}$ which is less toxic). One particular gene that is up-regulated is copA. CueR increases the transcription of copA in the presence of Cu, Ag, and Au (coinage metal) ions. Changela et al. developed an in vitro assay to determine the extent of expression of CueR-regulated genes under various ion types and concentrations. In the assay, purified CueR was added to a gene construct containing the promoter (a section of DNA immediately upstream of a gene start site where RNA polymerase binds) for copA. Initially, they found that transcription was always on, even in the presence of a ligand, glutathione, which binds Cu^{1+} avidly and should keep free Cu^{1+} levels very low. They switched to an even tighter binding Cu1+ coordinator, cyanide (CN^-), to reduce the free Cu^{1+} levels to even lower levels. Extremely high levels of CN^- (millimolar) stopped transcriptional activation, but if additional Cu1+ was added, activation ensued, suggesting that copper binding to the protein was reversible. At 1 mM CN^-, transcription increased with the addition of copper ions up to a TOTAL Cu^{1+} concentration of 60 μm. Under these conditions, the free Cu^{1+} concentrations were much lower. Given the presence of CN^- concentrations used, half-maximal activation occurred at a TOTAL Cu^{1+} concentration of 0.7 μM. Similar activation was observed by Ag^{1+} and Au^{1+}, but not by Zn and Hg ions, showing the specificity for monovalent cations over divalent cations.

Knowing the pKa of HCN, stability constants for Cu^{1+}:CN^- complexes, and CN- concentrations, Changela et al et al.ced a series of solutions buffered in FREE Cu^{1+} that extended from 10^{-18} to 10^{-23} M (pH 8.0). (For example, the log of the binding constant β, logβ, for the $Cu^{1+} + 2CN^- \leftrightarrow [Cu(CN)_2]^-$ is 21.7. You solved problems involving linked equilibrium if you have taken analytical chemistry.) The free Cu^{1+} concentration at half-maximal activation of gene reporter transcription, a measure of the dissociation constant, K_D, was approximately 1×10^{-21} M (zeptomolar)! Now assume that the volume of the contents of an E. Coli cell is 1.5×10^{-15} L. If there were only one ion of Cu^{1+} in the cell, it would have a concentration of 10^{-9} M. The values suggest that there are no free Cu^{1+} ions in the cell and that only 1 Cu^{+1} ion in the cell is enough to ensure its binding to CueR and subsequent transcriptional activation of copA.

It is essential for survival that bacterial cells get the correct metal to metalloproteins. A recent review by Waldron and Robinson illustrates how. The cell has many mechanisms for restricting specific binding sites, so metals can get to the right proteins. In addition, the natural order of stability for transition metal complexes must be considered in understanding metal affinities. That stability is given by the Irving-William series shown below (along with Group 2A metal ions). The trend parallels the size of the cation (going from largest to smallest):

$$Mn^{2+} < Fe^{2+} < Co^{2+} < Ni^{2+} < Cu^{2+} > Zn^{2+} \text{ (tightest binding)}$$

- The ability of a protein to change shape on ligand binding allows different metals to bind. For example, cyanobacterium has a high demand for copper and manganese. Manganese might bind to a protein, followed by folding, which traps it in the protein. This unstable metal cannot be replaced by copper, which would ordinarily out-compete Mn^{2+} for the site.
- Metal transporters help regulate how many ions of each metal are in the cell. Metal sensors are under the control of these metal transporters, regulating gene expression. Once a specific metal has a sufficient concentration for binding, the metal sensors target mRNA to repress specific genes and halt transcription.
- Another enzyme can also be activated for the metal's export. By restricting the concentrations of the competing metals, weaker metal-binding sites remain available.
- Metal sensors can also help to regulate what protein some metals use based on what is available. For example, E. coli switches metabolism to minimize the number of iron-requiring proteins expressed when iron is less abundant.
- Metals are supplied by multiple pathways (in case a specific enzyme is not present) and are trafficked to the correct protein through many ligand-exchange reactions.
- Certain enzymes bind specific metals that cause preferential conformational changes. Hence, if a metal comes along that binds more tightly but is not preferred by the enzyme, it will not trigger the enzyme because it binds in a different manner.

5.2.4: Molecular Basis of High Affinity Interactions

What differentiates high and low affinity binding at the molecular level? Do high affinity interactions have lots of intramolecular H-bonds, salt bridges, van der Waals interactions, or are hydrophobic interactions most important? Recently, the crystal structures of various antibody-protein complexes were determined to study the basis of affinity maturation of antibody molecules. It is well

known that antibodies elicited on exposure to a foreign molecule (antigen) are initially of lower affinity than antibodies released later in the immune response. An incredible number of different antibodies can be made by antibody-producing B cells due to genetic mechanisms (combining different variable regions of antibody genes through splicing, imprecise splicing, and hypermutation of critical nucleotides in the genes of antigen binding regions of antibodies). Clones of antibody-producing cells with higher affinity are selected through binding and clonal expansion of these cells. Investigators studied the crystal structure of 4 different antibodies which bind to the same site (epitope) on the protein antigen lysozyme. Increased affinity was correlated with increased buried apolar surface area and not with increased numbers of H bonds or salt bridges. The data for these antibodies are shown below in Table 5.2.1.

Table 5.2.1: Characteristics of Antibody:Hen Egg Lysozyme Complexes (HEL) from Li,Y. et al. Nature: Structural Biology. 6, pg 484 (2003)

Antibody	H26-HEL	H63-HEL	H10-HEL	H8-HEL
K_D (nM)	7.14	3.60	0.313	0.200
Intermolecular Interactions				
H bonds	24	25	20	23
VDW contacts	159	144	134	153
salt bridges	1	1	1	1
Buried Surface Area				
ΔASURF (A2)	1,812	1,825	1,824	1,872
ΔASURF-polar (A2)	1,149	1,101	1,075	1,052
ΔASURF-apolar (A2)	663	724	749	820

Electrostatic interactions between biological molecules are still very important, even though we may consider them often nonspecific. Consider the binding of DNA binding proteins with positive domains to the negative polyanion, DNA. The initial encounter will be electrostatic in origin and important in targeting the proteins to DNA where other specific interactions may take place.

In a similar example (Yeung, T et al.), it was recently reported that moderately positively charged proteins are directed to endosomes and lysosomes through interactions with negatively charged membrane phosphatidylserine (PS), whereas more positively charged proteins are targeted to the inner surface of the plasma membrane, which is enriched in PS and phosphorylated phosphatidyl inositol derivatives (PIP_2, PIP_3), as shown in Figure 5.2.7.

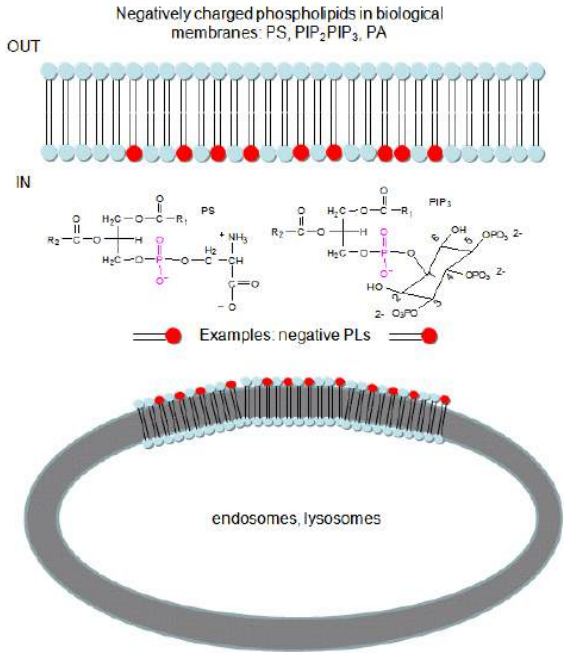


Figure 5.2.7: Negatively charged phospholipids in biological membranes

To study this they used the C2 domain of lactadherin (Lact-C2) from milk that binds PS in the presence of calcium. The C2 domain was covalently linked to the green fluorescent protein. This protein contains an internal fluorophore comprised of three amino acids (Ser65-Tyr66-Gly67) that cyclize spontaneously on folding to produce a fluorophore that emits green light. A fusion gene of Lact-C2 and GFP was introduced in wild-type (WT) and mutant yeast lacking PS. It bound to the inner leaflet in WT cells and to endosome and lysosome vesicles but found diffused through the cytoplasm in mutant cells. They also made cationic probes with farnesyl tails that could anchor the soluble probes to membranes. The most positively charged probes were recruited to the plasma membrane inner leaflet, while less charged ones were recruited to internal vesicles. The authors speculate that PS on cytoplasmic membrane layers can target signal transduction proteins to these regions.

Antibodies with Infinite Affinity. Chmura et al. PNAS. 98, pg 8480 (1998)

5.2.5: Docking

The quantitative methods described above do not elucidate the mechanism of binding. Computer programs have been developed that allow the docking of a ligand (small molecule or even another protein) to another protein. The automatic docking of flexible ligands to proteins can be modeled using free programs such as Autodock. Molecular dynamics simulations can also be used to study the actual binding and unbinding processes.

5.2.6: The Crowded Cell

Most binding studies are performed in vitro with dilution concentrations of both macromolecule and ligand. Are these conditions illustrious of conditions inside a cell? The answer is no! Cells are crowded with organelles, macromolecular complexes, and cytoskeletal components that provide an internal architecture to the cell. Total macromolecule concentration in the cell has been estimated to be as high as 400 g/1L = 400 g/1000 mL = 0.4 g/mL = 400 mg/mL. Try to dissolve a water-soluble protein like albumin at those concentrations! From 5 to 40% of the entire cellular volume is occupied with large molecules, and at the upper range, very little space exists for other large macromolecules. A representation showing the crowdedness of a bacterial cell at the atom level is shown in Figure 5.2.8.

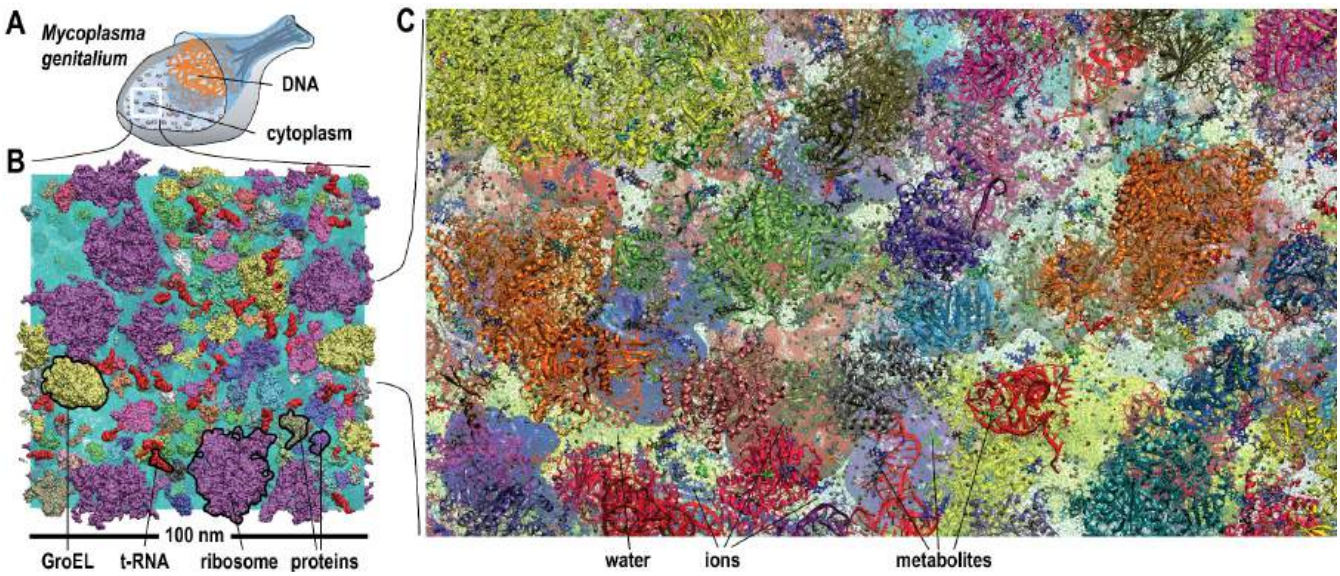


Figure 5.2.7: (**A**) Schematic illustration of *Mycoplasma genitalium* (*MG*). (**B**) Equilibrated MG_h system highlighted with proteins, tRNA, GroEL, and ribosomes. (**C**) MG_h cl ose-up showing atomistic level of detail. Yu et al. (2016) *eLife* **5**:e19274. https://doi.org/10.7554/eLife.19274. Creative Commons Attribution License.

Imagine trying to diffuse through that!

5.3: Oxygen-Binding Proteins and Allosterism

5.3.1: Myoglobin, Hemoglobin, and their Ligands

Almost all biochemistry textbooks start their description of the biological functions of proteins using myoglobin and hemoglobin as exemplars. On the surface this is a rational approach since they have become model systems to describe the binding of simple ligands, like dioxygen (O_2), CO_2, and H^+, and how the structure of a protein determines and is influenced by binding of ligands.

Yet in most ways, these globin-binding "ligands" are dissimilar to the majority of both small ligands, such as substrates (for enzymes), inhibitors and activators as well as large "ligands", such as other proteins, nucleic acids, carbohydrates and lipids that bind to proteins through noncovalent interactions (described in detail in Chapter 2.4). In contrast, dioxygen (O_2), CO_2, and H^+ bind reversibly, but through covalent interactions. Dioxygen binds to a heme Fe^{2+} transition metal through a coordinate covalent or dative bond, protons obviously bind covalently to proton acceptors (Lewis bases like histidine), while CO_2 binds covalently as it forms a carbamate with the N terminus of a hemoglobin chain. In typical covalent bonds, each bonded atom contributes to and shares the two electrons in the bond. In coordinate or dative covalent bonds, the ligand, a Lewis base, contributes both electrons in the bond. For simple analyses and four counting electrons, both electrons can be considered to be "owned" by the ligand and not by the transition metal ion, a Lewis acid, unless you analyze the interactions using molecular orbital (ligand field) theory. Hence the ligand can readily dissociate from the metal ion, much as a ligand bound through classical noncovalent interactions does. This analogy can be extended to protons which are also Lewis acids (with no contributing electrons) as they react with Lewis bases (lone pair donors) on atoms such as nitrogen on a histidine side chain. H^+ readily leave (ionize) from a Lewis acid if the pH of the microenvironment is conducive to ionization.

Even though we disagree with starting the discussion of protein structure and function with the covalent binding of small gaseous and marginally soluble ligands to myoglobin and hemoglobin, we will anyway to make the book consistent with most other texts and allow easier use without shuffling the order of chapters.

Let's start with myoglobin (Mb), a monomeric protein containing 8 α−helices (A-H) and with hemoglobin, a heterotetramer with two α -and two β−subunits, each which also contains 8 α−helices. Both are oxygen binding proteins. Both contain heme (one in myoglobin, and 4 for the four subunits of hemoglobin). Each heme has a central Fe^{2+} ion, which forms a coordinate covalent bond with dioxygen. Dioxygen is transported from lungs, gills, or skin of animals to the capillaries, where it can be delivered to respiring tissue. O_2 has a low solubility in blood (0.1 mM). Whole blood contains 150 g Hb/L, and can achieve a dissolved oxygen concentration of 10 mM. Invertebrates can have alternative proteins for oxygen binding, including hemocyanin, which contains Cu and hemerythrin, a non-heme protein. On binding dioxygen, solutions of Hb change color to bright red. Solutions of hemocyanin and hemerythrin change to blue and burgundy colored, respectively, on binding dioxygen. Some Antarctic fish don't require Hb since dioxygen is more soluble at low temperatures. Myoglobin is found in the muscles, and serves as a storage protein for oxygen transported by hemoglobin.

The structure of heme in myoglobin and hemoglobin, is shown in Figure 5.3.1.

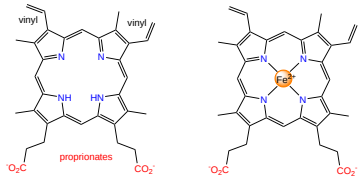


Figure 5.3.1: Heme

The heme group contains protoporphyrin IX, with four tetrapyrrole rings linked by methene bridges. Attached to the tetrapyrrole structure are four methyl, two vinyl, and two propionate groups. These can be arranged in 15 ways, of which only one (IX) occurs in biological systems. Protoporphyrin IX with bonded ferrous (Fe^{2+}) iron is called heme and has a nitrogen atom on each of the four pyrrole rings that form a coordinate covalent bond to Fe^{2+}. The heme fits into a hydrophobic crevice in heme-bindin proteins with the propionate groups exposed to solvent.

5.3.2: Myoglobin (Mb)

Myoglobin is an extremely compact protein with 75% alpha helical structure. It has 8 α−helices labeled A-H. Four are terminated by a proline, a helix breaker. The interior amino acids are almost entirely nonpolar. The only polar amino acids found completely buried are two histidines. One is called the proximal His as it is nearer the heme and serves the 5th ligand to the heme Fe^{2+}. The other is called the distal His, which is too far to coordinate the heme Fe^{2+}. This last potential 6th ligand-binding site forms a coordinate covalent bond with O_2 in oxy-myoglobin.

Figure 5.3.2 shows an interactive iCn3D model of deoxymyoglobin from wild boar. The heme is shown in sticks along with the proximal and distal histidines.

Figure 5.3.2: Deoxymyoglobin (1MWD) (Copyright; author via source). Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...PTroKxtYLedjJ6

5.3.3: Hemoglobin

Hemoglobin has an illustrious history. It is the first protein whose molecular weight was determined and the first assigned a specific function (dioxygen transport). It was the first protein in which a mutation in a single amino acid caused by a single base pair change in the DNA coding sequence was shown to cause a disease (sickle cell trait and disease). The mathematical theories developed to model dioxygen binding are used to explain enzyme activity. It also binds H^+, CO_2, and bisphosphoglcyerate which bind to sites (allosteric) distant from the oxygen binding site which regulates its dioxygen binding affinity.

As with myoglobin, the Fe^{2+} ion is coordinated to 4 Ns on the 4 pyrrole rings, The 5th ligand is supplied by proximal His (the 8th amino acid on helix F) of the protein. In the absence of dioxygen, the 6th ligand is missing. and the geometry of the complex is somewhat square pyramidal, with the Fe slightly above (0.2 Å) the plane of the heme ring. A distal His (E7) is on the opposite side of the heme ring, but too far to coordinate with the Fe^{2+}. When dioxygen binds, it occupies the 6th coordination site and pulls the Fe into the plane of the ring, leading to octahedral geometry. These changes that occur on oxygenation are shown in Figure 5.3.3.

Figure 5.3.3: Changes in heme structure on binding of dioxygen.

The proximal histidine that provides the imidazole nitrogen ligand is shown. Dioxygen is shown as red spheres. Fe^{2+} ion is shown as a small orange sphere. Its size has been dramatically reduced in this image so its movement can be more readily observed.

Carbon monoxide (CO), nitric oxide (NO), and hydrogen sulfide (H_2S) also bind to the sixth coordination site, but with higher affinity than dioxygen, which can lead to CO poisoning for example. The distal histidine keeps these ligands (including dioxygen) bound in a bent, non-optimal geometry. This minimizes the chances of CO poisoning.

Figure 5.3.4 shows an interactive iCn3D model of human oxy-hemoglobin (2dn1)

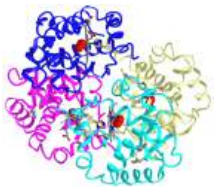


Figure 5.3.4: Human oxy-hemoglobin (2DN1) (Copyright; author via source). Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...yBJD4p8EazRpU7

5.3.3.1: Fe^{2+} ion ligand interactions

When the 6th ligand, dioxygen, binds to heme Fe^{2+}, the geometry of the complex becomes octahedral. The Fe^{2+} ion has 6 electrons in d orbitals. The electronic configuration of atomic Fe is $3d^64s^2$ while the Fe^{2+} ion has a $3d^6$ configuration, as shown in Figure 5.3.5. Each of the orbitals would have the same energy except for the doubly occupied one which would have slightly higher energy due to the extra repulsion of the two electrons in the orbital. This effect is minimal so to a first approximation, the orbitals are considered to have the same energies (they are degenerate). The figure below shows them having the same energy

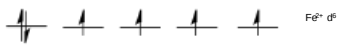


Figure 5.3.5: Electron atomic orbital diagram of Fe^{2+}, a d^6 transition metal ion.

You will remember from introductory chemistry classes that transition metal complexes and their solutions are highly colored. Since oxygenated hemoglobin (found in arteries) appears bright red/orange, it must absorb blue/green light more than deoxyhemoglobin, which is darker red (but still reddish). These absorbed wavelengths are removed from the spectrum, making hemoglobins shades of red. Veins contained more deoxygenated blood returning to the heart to be reoxygenated. The visible veins in your arms and legs appear blue **not** because of the spectral properties of deoxyhemoglobin. Rather, blue light doesn't penetrate into the tissue as far as red, so red is preferentially removed from the remaining light which is reflected, making veins appear blue. Figure 5.3.6 shows a partial absorbance spectrum of deoxy- and oxy-hemoglobin from 280-1000 nm.

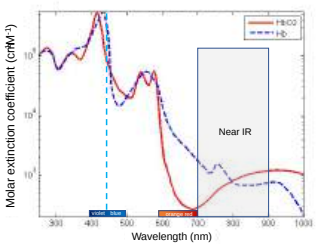


Figure 5.3.6: Spectrum of deoxy- and oxyhemoglobin - log molar extinction coefficient vs wavelength. The red line is the oxyHb spectra, and the dotted blue line shows the deoxyHb spectrum. Modified from en:Bme591wikiproject, Creative Commons Attribution-Share Alike 3.0 Unported .

The dashed light blue vertical line shows the approximate wavelength (around 450 nm) of the largest molar extinction coefficient difference between the two forms. Note that the y axis is a log scale. At 450 nm, oxyhemoglobin has an extinction coefficient of about 600,000 while deoxyhemoglobin is about 60,000, so much more blue light is removed from a solution of oxyHb, making it appear bright red. Note also that the same differences appear in the red region of the spectra, where the extinction coefficients only vary from 3000 for oxyhemoglobin to 200 for deoxyHb. Hence absorption in this region has little effect on the visible color of blood.

The spectrum shown in Figure 5.3.6 also shows the near-infrared region (denoted by a rectangle) of the spectra. Inexpensive pulse oximeters (some built into watches) have been increasingly used by people at home to their measure their oxygenation status during the COVID-19 pandemic. These use two pulsed LEDs, one at 660 nm, where oxyhemoglobin has a higher extinction coefficient, and one at 940 (infrared region), where deoxyhemoglobin has a higher extinction coefficient.

5.3.3.2: Binding of ligand to the heme Fe^{2+}: crystal and ligand fields

Most biochemistry books offer minimal coverage of bioinorganic chemistry, even though a large percentage of proteins bind metal ions. Since biochemistry is an interdisciplinary field, it is important to use past learning in other biology and chemistry courses and apply it to biochemistry. Most students study transition metal chemistry in introductory chemistry courses. You should familiar with transition metal ions, their electronic configuration, crystal field theory, high and low spin states, paramagnetism and diamagnetism. Hence it is appropriate and important to bring these ideas into biochemistry and extend them when necessary. The basis of the material below is modified from Structure & Reactivity in Organic, Biological and Inorganic Chemistry by Chris Schaller (Creative Commons Attribution-NonCommercial 3.0 Unported License).

Let's look at the electronic structure of Fe^{2+} in oxyhemoglobin. The Fe^{2+} is coordinated to six ligands (4 pyrrole rings of the heme in a plane, one axial imidazole ring each from the proximal histidine and bonded dioxygen). The geometry of the electron clouds around the Fe^{2+} is octahedral. The six d orbitals are oriented in the same x, y, and z axes direction as the heme ligands. This geometry and the shapes and orientations of the Fe^{2+} d **atomic** orbitals are illustrated in Figure 5.3.7.

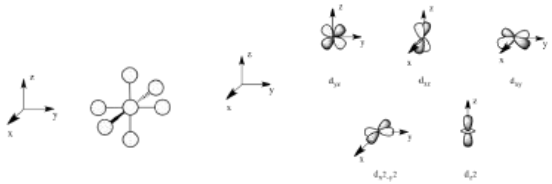


Figure 5.3.7: Fe^{2+} d orbitals and their orientations

Two of the orbitals, d_{z2} and d_{x2-y2}, appear different in that they are oriented directly along the x, y and z axes while the other three are in-between the axes. Now imagine six **anionic** ligands with lone pairs approaching along the axes of the Fe^{2+} atomic orbitals, a postulate of **crystal field theory** used to explain bonding in transition metal complexes. The energy of the Fe electrons in the d_{z2} and d_{x2-y2} atomic orbitals would be raised higher than the others due to electron-electron repulsion. This is illustrated in Figure 5.3.8. Two different outcomes can arise

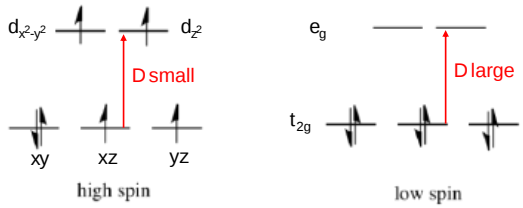


Figure 5.3.8: High spin and low spin electron energy level diagrams for Fe^{2+} ion in an octahedral complex

1 (left panel): If the electrostatic interaction between the atomic orbitals of the Fe ion and the incoming ligands is low, the energy of the d_{z2} and d_{x2-y2} orbitals would be a bit higher (by an amount Δ) due to the great repulsion along those axes from anions oriented along them. This is illustrated in the left part of Figure 5.3.8. When filling the orbitals with the six Fe^{2+} $3d^6$ electrons, you would add one electron to each of the 5 orbitals and then pair one for the sixth orbital. In that case there would be 4 unpaired electrons and the complex would be paramagnetic. We call this a **high spin** state.

2 (right panel): If however, the electrostatic interactions of the incoming ligands with the d atomic orbital electrons is high, the Δ would be large. When filling the orbitals in this case with the six Fe^{2+} $3d^6$ electrons, the electron would be paired in the three lower energy orbitals and there would be no electrons unpaired, so the ion is diamagnetic. This is the **low spin** case.

In either case, if light of energy equal to the Δ interacts with the metal ion, electrons would be promoted to the higher energy level. For the low spin case (larger axial electrostatic interactions, Δ is large, the energy of the required photon is larger (more blue shifted) than in the high spin case. This could make a solution of the molecule in the low spin case appear redder than for the high spin case (as blue light is removed on absorption). This is the case for dioxygen, which interacts strongly with the Fe^{2+} axially-oriented orbitals. Hence oxy-Fe^{2+} heme complex is low spin and diamagnetic.

Ligand Field Theory

You studied basic crystal field theory in introductory chemistry courses. The theory is simplistic in several ways. The ligands might not be anions. More importantly, bonding is best described by molecular orbitals. A more comprehensive **ligand field theory** takes into account the effect of the donor electrons and the d orbitals of the transition metal ion. These atomic orbitals combine to produce **molecular orbitals (MOs)**, which better describes bonding.

Take the simple case of a covalent bond between two singly occupied adjacent p atomic orbitals the two carbons of ethylene (C_2H_4). Simplistically, you could imagine two carbon atoms with their two p orbitals approaching each other. The two p orbitals (or more accurately their wave functions) could combine constructively or destructively to form two new molecular orbitals (MO). One is a pi bonding MO, π, which is lower in energy (promoting bond formation) than the atomic p orbitals. The other is a pi antibonding MO, π^*, which is higher in energy (antagonizing bond formation), as shown in Figure 5.3.9. Two atomic orbitals form from two MOs!

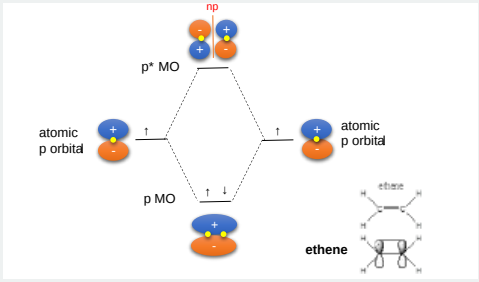


Figure 5.3.9: Formation of pi molecular orbtials for ethene

Let's use ligand field theory with its MOs to describe ligands binding to the heme Fe^{2+} d orbitals. This is illustrated in the MO diagram in Figure 5.3.10.

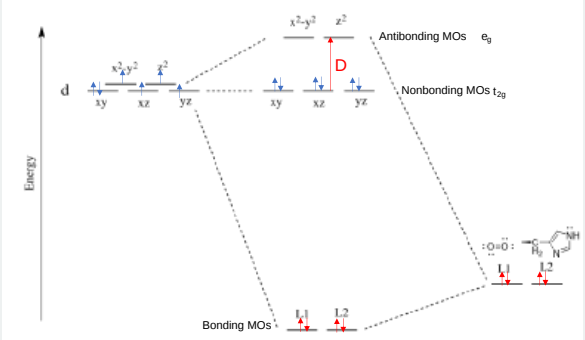


Figure 5.3.10: Ligand field molecular orbtials for the d^6 Fe^{2+} ion for two axial ligands

For simplicity, for Fe^{2+} we will consider only d orbitals and focus on the $d_{z}2$ and d_{x2-y2} (also called the e_g) orbitals, since these are most affected by the ligands as described in crystal field theory. Let's assume these interact with lone pairs (a simplistic assumption as well) in two ligand orbitals, one on the dioxygen and one on the proximal histidine imidazole N. Seven atomic orbitals (5 Fe d^{2+} orbitals and 2 ligand orbitals) combine to produce 7 MOs. Since the orbitals (in sp^2- or sp^3-like orbitals) on the ligands are closer in energy to the lower energy bonding MOs, their electrons would go there. Since O_2 interacts strongly with the Fe^{2+} d orbitals, the system is in a low spin state, with the unoccupied $d_{z}2$ and d_{x2-y2} (e_g orbitals) now considered antibonding orbitals. The fully occupied d_{xy}, d_{xz} and d_{yz} are considered nonbonding orbitals since they have the same relative energy atomic orbitals.

We will return to molecular orbitals occasionally throughout this book when they offer the best explanation for biological events.

Here are some important things to note. When dioxygen binds to the heme iron, the oxidation state of the Fe^{2+} ion does **not** change, even though dioxygen is a great oxidizing agent. Hence the Fe^{2+} ion is a reversible carrier of dioxygen not of electrons. Free heme in solution is oxidized by dioxygen, forming a complex with water which occupies the 6^{th} position, with the iron in the Fe^{3+} state. An intermediate in this process is the formation of a dimer of 2 hemes linked by 1 dioxygen. This can't occur readily when the heme is in Hb or Mb. Other heme proteins (like Cytochrome C), which we will explore in future chapters, are designed to be carriers of electrons. A small amount of the Fe^{2+} ion can get oxidized to Fe^{3+} ion myoglobin and hemoglobin, resulting in met-Hb and met-Mb. The brown color of old meat results in large part from Met-Mb. A enzyme is required to reduced the iron back to the Fe^{2+} state.

The differences between hemoglobin and myoglobin are equally important. Hemoglobin is a heterotetramer of two α and two β subunits held together by noncovalent interactions (an example of quaternary protein structure), with 4 bound hemes, each of which can bind a dioxygen. In a fetus, two other subunits make up hemoglobin (two zeta - ζ and two epsilon - ε subunits -analogous to the two α and two β subunits, respectively). This changes in development to two α and two γ subunits. Fetal hemoglobin has a higher affinity for dioxygen than does adult hemoglobin. Myoglobin is a single polypeptide chain and has a higher affinity for dioxygen than hemoglobin.

The α and β chains of hemoglobin are similar to that of myoglobin, which is unexpected since only 24 of 141 residues in the α and β chains of Hb are identical to amino acids in myoglobin. This suggests that different sequences can fold to similar structures. The globin fold of myoglobin and each chain of hemoglobin is common to vertebrates and must be nature's design for dioxygen carriers. A comparison of the sequence of hemoglobin from 60 species shows much variability of amino acids, with only 9 identical amino founds. These must be important for structure/function. All internal changes are conservative (e.g. changing a nonpolar for a nonpolar amino acid). Not even prolines are conserved, suggesting there are different ways to break helices. The two active site histidines are conserved, as is glycine B6 (required for a reverse turn). http://www.umass.edu/molvis/tutorials/hemoglobin/

5.3.4: Normal and Cooperative Binding of Dioxygen - Structural Analyses

Plots of Y (fractional saturation) vs L (pO_2) are hyperbolic for Mb, but sigmoidal for Hb, suggesting cooperative binding of oxygen to Hb (binding of the first oxygen facilitates binding of second, etc). Figure 5.3.11 shows fractional saturation (Y) binding curves vs dioxygen concentration (P_{O2}) for both myoglobin and hemoglobin.

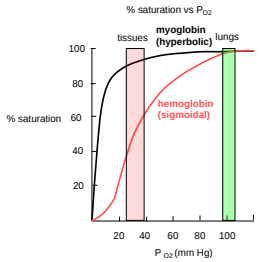


Figure 5.3.11: Plots of Y (fractional saturation) vs L (pO2) are hyperbolic for Mb, but sigmoidal for Hb

Note that hemoglobin is saturated with O_2 at the high concentration found in the lung, but it releases much of its bound O_2 in respiring tissues in which O_2 is much lower. In contrast, myoglobin only releases significant bound oxygen at much lower O_2 concentrations. Hence myoglobin is designed for dioxygen storage.

In another difference, the affinity of Hb for dioxygen, but not Mb, depends on pH. This is called the Bohr effect, after the father of Niels Bohr, who discovered it. Figure 5.3.12 shows binding curves for hemoglobin in the presence of increasing and decreasing concentrations of H^+ (pH) as well as for CO_2 and another ligand, 2,3-disphosphoglycerate (2,3-DPG).

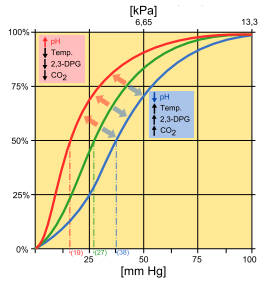


Figure 5.3.12:

Michał Komorniczak (Poland). https://commons.wikimedia.org/wiki/F...ohr_Effect.svg***Creative Commons 3.0. Attribution-ShareAlike (CC BY-SA 3.0)***.

Protons (decreasing pH), carbon dioxide, and bisphosphoglycerate, all allosteric ligands which bind distal to the oxygen binding sites on the heme, shift the binding curves of Hb for oxygen to the right, lowering the apparent affinity of Hb for oxygen. The same effects do not occur for Mb. These ligands regulate the binding of dioxygen to Hb.

From these clues, we wish to discern the

- molecular and mathematical bases for the sigmoidal binding curves
- mechanism for the exquisite regulation of O_2 binding by allosteric ligands.

The two obvious features that differ between Mb and Hb are the tetrameric nature of Hb and its multiple (4) binding sites for oxygen. Regulation of dioxygen binding is associated with conformational changes in hemoglobin.

Based on crystallographic structures, two main conformational states appear to exist for Hb, the deoxy (or T - taut) state, and the oxy (or R -relaxed) state. The major shift in conformation occurs at the alpha-beta interface, where contacts with helices C and G and the FG corner are shifted on oxygenation. Figure 5.3.13 shows conformation changes on O_2 binding to deoxy-hemoglobin (files aligned with DeepView, displayed with Pymol). Dioxygen is shown as red spheres.

Figure 5.3.13: Conformational changes in deoxyhemoglobin on binding dioxygen (red spheres)

The deoxy or T form is stabilized by 8 salt bridges which are broken in the transition to the oxy or R state. This is Illustrated in Figure 5.3.14.

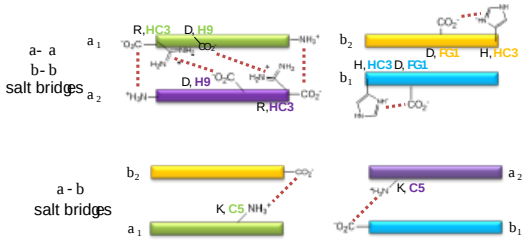


Figure 5.3.14:: Salt Bridges in Deoxy Hb

6 of the salt bridges are between different subunits (as expected from the above analysis), with 4 of those involving the C- or N- terminus.

In addition, crucial H-bonds between Tyr 140 (alpha chain) or 145 (on the beta chain) and the carbonyl O of Val 93 (alpha chain) or 98 (beta chain) are broken. Crystal structures of oxy and deoxy Hb show that the major conformational shift occurs at the interface between the α and subunits. When the heme Fe binds oxygen it is pulled into the plan of the heme ring, a shift of about 0.2 nm. This small shift leads to larger conformational changes since the subunits are packed so tightly that compensatory changes in their arrangement must occur. The proximal His (coordinated to the Fe^{2+}) is pulled toward the heme, which causes the F helix to shift, causing a change in the FG corner (the sequence separating the F and G helices) at the alpha-beta interface as well as the C and G helices at the interface, which all slide past each other to the oxy-or R conformation.

Decreasing pH shifts the oxygen binding curves to the right (to decreased oxygen affinity). Increased [proton] will cause protonation of basic side chains. In the pH range for the Bohr effect, the most likely side chain to get protonated is His (pKa around 6), which then becomes charged. The most likely candidate for protonation is His 146 (on the β chain - HC3), which can form a salt bridge with Asp 94 of the β(FG1) chain. This salt bridge stabilizes the positive charge on the His and raises its pKa compared to the oxyHb state. Carbon dioxide binds covalently to the N-terminus to form a negatively-charge carbamate, which forms a salt bridge with Arg 141 on the alpha chain. Bisphosphoglycerate (BPG), a strongly negatively charged ligand, binds in a pocket lined with Lys 82, His 2, and His 143 (all on the beta chain). It fits into a cavity between the β subunits of the Hb tetramer in the T state. Notice all these allosteric effectors lead to the formation of more salt bridges which stabilize the T or deoxy state. The central cavity where BPG binds between the β subunits becomes much smaller on oxygen binding and the shift to the oxy or R state. Hence BPG is extruded from the cavity.

The binding of H^+ and CO_2 help shift the equilibrium to the deoxyHb form, which facilitates the release of oxygen to the tissues. It is in respiring tissues that CO_2 and H^+ levels are high. CO_2 is produced from oxidation of glucose through glycolysis and the Krebs cycle. In addition, high levels of CO_2 increase H^+ levels through the following equilibrium:

$$H_2O + CO_2 \leftrightarrow H_2CO_3 \leftrightarrow H^+ + HCO_3{}^- \tag{5.3.1}$$

In addition, H^+ increases from weak acids such as pyruvic acid produced in the central metabolic pathway (glycolysis) to produce energy from glucose oxidation.

The binding of CO_2 and H^+ to hemogl serves an additional function: it removes excess CO_2 and H^+ from the tissues where they build up. When deoxyHb with bound H^+ and CO_2 reaches the lungs, they leave as O_2 builds and deoxyHb is converted to oxyHb.

Hemoglobin exhibits **allosterism**. Allosterism occurs when a regulatory ligand (like CO_2 and H^+) binds to a site distal to the binding site of a main ligand (like O_2) and changes the affinity for the main ligand. We will define in our own convention two kinds of allosterism:

Type I occurs when a ligand such as dioxygen binds to multiple ligand binding sites on the same protein and gives sigmoidal binding plot as a function of ligand concentration. Multiple binding sites for a main ligand can be found on a multimeric protein with identical or similar subunits (as in the case of hemoglobin but not myoglobin). In the case of hemoglobin, the main ligand O_2 binds to the same active site in each monomer. This site is called the **orthosteric** binding site. For the hemoglobin tetramer, the orthosteric site of course is the heme Fe^{2+} ion. **Type I allosterism** occurs when an allosteric ligand binds at the active sites of the monomeric subunits of the protein. Most texts call this **homotropic allosterism** and the ligands **homotropic** ligands. Hence O_2 and CO are homotropic ligands for hemoglobin.

Type II occurs when a chemical modulator binds to a site different from the binding site for the main ligand. In doing so, it modulates (activates or inhibits) the binding affinity of the main ligand for the active or orthosteric site and shifts the binding curve for ligand binding. In addition, binding plots at a fixed main ligand concentration with varying modulator concentrations are also sigmoidal. The **modulator** binds at an **allosteric** (other) binding site. In the case of hemoglobin, the allosteric modulators are H^+, CO_2 and BPG. H^+ and CO_2 binding shift the O_2 bind curve in ways that lower the affinity for O_2, leading to its release. Most texts call this type of allosterism **heterotropic**. Protons (H^+) and CO_2 hence are **heterotropic** ligands. We prefer Type I and Type II over the more jargonistic less intuitive terms homo- and heterotropic.

5.3.5: Mathematical Analysis of Cooperative Binding

How do the sigmoidal dioxygen binding curves for Hb arise? Mathematics can offer clues that complement and extend structural information. At least three models (Hill, MWC, and KNF) have been developed that give rise to sigmoidal binding curves. Remember, sigmoidal curves imply cooperative binding of oxygen to hemoglobin. As oxygen binds, the next oxygen seems to bind with higher affinity (lower K_D). We will discuss the mathematics behind two of the models. Both models are routinely applied to binding phenomena that give sigmoidal curves.

Previously we have shown that the binding of oxygen to myoglobin can be described by chemical and mathematical equations.

$$M + L \leftrightarrow ML \tag{5.3.2}$$

$$Y = \frac{L}{K_D + L} \tag{5.3.3}$$

These mathematical equations are that of a hyperbola where Y is fractional saturation. Let's now explore two models that give sigmoidal curves.

5.3.5.1: Hill Model

In this model, we base our mathematical analysis on the fact that the stoichiometry of binding is not 1:1, but rather 4 to 1: Perhaps a more useful equation to express the equilibrium would be M + 4L ↔ ML_4. We can derive an equation analogous to the one above:

$$Y = \frac{L^4}{K_D + L^4} \tag{5.3.4}$$

For any given L and K_D, a corresponding Y can be calculated. A plot of Y vs L is not hyperbolic but sigmoidal (see the next link below). Hence we're getting closer to modeling that actual data. However, there is one problem. This sigmoidal curve does not give a great fit to the actual oxygen binding curve for Hb. Maybe a better fit can be achieved by altering the exponents in the equation. A more general equation for binding might be M + nL ↔ ML_n, which gives the Hill equation:

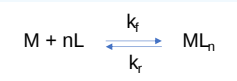
$$Y = \frac{L^n}{K_D + L^n} \tag{5.3.5}$$

? A derivation of the Hill Equation

Some books/sources offer a different Hill equation. The one above is correct. Click the derivation below, which requires only a background from the preceding chapter sections and maybe a bit from Chapter 6.1.

Derivation

Let's consider this reaction:



For a minute, assume n=2, and the reaction was written as M + 2L ↔ ML_2. Also assume that the reaction occurs when M, L, and a second L all collide simultaneously (unlikely) to form the product, ML_2. From introductory chemistry, you would write the rate equation for just the forward "ternary" reaction as:

$$\text{rate}_f = d[ML_2]/dt = k_f[M][L][L] = k_f[M][L]^2$$

The rate of just the reverse reaction would be:

$$\text{rate}_r = d[ML_2]/dt = k_r[ML_2]$$

Subtract the two rates to get the net rate. Now let's switch from n= 2 back to just n and proceed without further explanation.

At equilibrium, the forward and reverse rates are equal, so

$$\frac{d\,[ML_n]}{dt} = k_f[M][L]^n = k_f[M][L]^n - k_r\,[ML_n] = 0 \text{ at equilibrium} \tag{5.3.6}$$

K_D is equal to the ratio of k_r/k_f, so solving for K_D gives

$$K_D = \frac{k_r}{k_f} = \frac{[M][L]^n}{[ML_n]} \tag{5.3.7}$$

Solve for $[ML_n]$ gives

$$[ML_n] = \frac{[M][L]^n}{K_D} \tag{5.3.8}$$

The fractional saturation is given by

$$Y = \frac{[ML_n]}{[M] + [ML_n]} \tag{5.3.9}$$

Plug in the previous equation for $[ML_n]$ to obtain the correct form of the Hill Equation.

$$Y = \frac{\frac{[M][L]^n}{K_D}}{[M] + \frac{[M][L]^n}{K_D}}\left(\frac{K_D}{K_D}\right) = \frac{[M][L]^n}{K_D[M] + [M][L]^n} = \frac{[L]^n}{K_D + [L]^n} \tag{5.3.10}$$

QED!

If n is set to 2.8, the theoretical curve of Y vs L gives the best, but still not perfect, fit to the experimental data. It must seem arbitrary to change the exponent which seems to reflect the stoichiometry of binding. What molecular interpretation could you give to 2.8?

Consider another meaning of the equilibrium described above: M + 4L ↔ ML_4.

One interpretation is that all 4 oxygens bind at once to Hb. Or, alternatively, the first one binds with some low affinity, which through associated conformational changes alters the remaining 3 sites to very high affinity sites. This model implies what is described as **infinitely cooperative binding** of oxygen.

(Notice that the Hill equation becomes: Y = L/[K_D + L], when n =1 (as in the case with myoglobin, and in any equilibrium expression of the form M+L↔ML. Remember plots of ML vs L or Y vs L gives hyperbolas, with K_D = L at Y = 0.5.)

Does K_D = L at Y = 0.5? The oxygen concentration at which Y = 0.5 is defined as P_{50}. We can substitute this value into equation 3 which gives an operational definition of K_D in terms of P_{50}.

$$\mathrm{Y} = 0.5 = \frac{\mathrm{P}_{50}^{\mathrm{n}}}{\mathrm{K_D} + \mathrm{P}_{50}^{\mathrm{n}}} \tag{5.3.11}$$

multiply both sides by 2 give

$$\begin{gathered} 1 = \frac{2\mathrm{P}_{50}^{\mathrm{n}}}{\mathrm{K_D} + \mathrm{P}_{50}^{\mathrm{n}}} \\ \mathrm{K_D} + \mathrm{P}_{50}^{\mathrm{n}} = 2\mathrm{P}_{50}^{\mathrm{n}} \\ \mathrm{K_D} = \mathrm{P}_{50}^{\mathrm{n}} \end{gathered} \tag{5.3.12}$$

Note that for this equation, K_D is **not** the ligand concentration at half-saturation as we saw in the case with hyperbolic binding curves.

This gives a modified version of the Hill equation for hemoglobin binding of dioxygen:

$$Y = \frac{L^n}{P_{50}^n + L^n} \tag{5.3.13}$$

Is the Hill equation still useful?

This Hill equation with the Hill coefficient n that is empirically determined to obtain the best fit to the binding data might seems a bit contrived (especially after seeing the more mechanistically and chemically intuitive MWC equation described below). Is it useful in any other circumstance? Indeed it is and it is used often in modeling more complex interconnected binding and kinetic pathways that show similar "exquisite" sensitivity to concentrations and resulting sigmoidal binding and kinetic plots. We will see its use in Chapter 30.13 when we model protein kinases that regulate the cell cycle!

Use the sliders in the interactive graph below to explore the effect of changes in K_D and n on fractional saturation.

The Hill equation for hemoglobin gives sigmoidal dioxygen binding curve that fit the actual binding data.

5.3.5.2: MWC Symmetry Model

In the MWC (Monod, Wyman, and Changeux) model, in the absence of ligand (oxygen), hemoglobin is assumed to exist in two distinct conformations, the T state (equivalent to the crystal structure of deoxyHb) and the R state (equivalent to the crystal structure of oxyHb without the oxygen). In the absence of dioxygen, the T state (T_0) is greatly favored over the unliganded R state (R_0) at equilibrium. In the presence of increasing oxygen, the R state is favored.

A constant (somewhat equivalent to a dissociation constant) can be defined.

$$\mathrm{L} = \mathrm{T_0}/\mathrm{R_0} \tag{5.3.14}$$

(Note: L is **not** the ligand concentration so don't get confused.) In addition, let us assume that hemoglobin can **not** exist with some of the monomers in the tetramer in the T state while others in the same tetramer are in the R state. Hence this model is often called the **symmetry model**. Finally, let's assume that each oxygen can bind to either the T or R state with the dissociation constants K_T and K_R respectively. These constants do not depend on the number of dioxygens already bound to the tetramer. Hence

$$\mathrm{K_R} = \frac{[\mathrm{R_0}]\,[\mathrm{S}]}{[\mathrm{R_1}]} = \frac{[\mathrm{R_1}]\,[\mathrm{S}]}{[\mathrm{R_2}]} = \ldots \frac{[\mathrm{R_n}]\,[\mathrm{S}]}{[\mathrm{R_{n+1}}]} \tag{5.3.15}$$

and

$$\mathrm{K_T} = \frac{[\mathrm{T_0}]\,[\mathrm{S}]}{[\mathrm{T_1}]} = \frac{[\mathrm{T_1}]\,[\mathrm{S}]}{[\mathrm{T_2}]} = \ldots \frac{[\mathrm{T_n}]\,[\mathrm{S}]}{[\mathrm{T_{n+1}}]} \tag{5.3.16}$$

where the subscript on R and T refers to the number of dioxygens bound to that form of R or T. A cartoon representation of the T and R forms and accompanying dioxygen binding is shown in Figure 5.3.15.

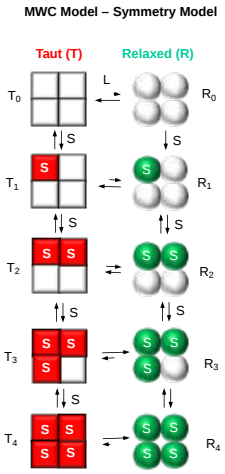


Figure 5.3.15: MWC model for cooperative binding of dioxygen to hemoglobin.

Now define two new parameters:

$$\alpha = \frac{pO_2}{K_R} = \frac{[S]}{K_R} \tag{5.3.17}$$

where α is really a normalized ligand concentration describing how many times the K_R the ligand concentration is, and

$$c = \frac{K_R}{K_T} \tag{5.3.18}$$

the ratio of the dissociation constants for the R and T forms.

If oxygen binds preferentially to the R form of hemoglobin, c would be a small fractional number. In the limiting case, when oxygen didn't bind to the T form, K_T would be infinite, and c = 0.

Using these definitions and equations, the following equation for Y, fractional saturation vs α can be derived, with n, the number of binding sites per molecule, = 4 for Hb.

$$Y = \frac{\alpha(1-\alpha)^{n-1} + Lc\,\alpha(1+c\alpha)^{n-1}}{(1+\alpha)^n + L(1+c\alpha)^n} \tag{5.3.19}$$

Figure 5.3.16 shows how fractional saturation (Y) vs alpha varies with L and c for the MWC model.

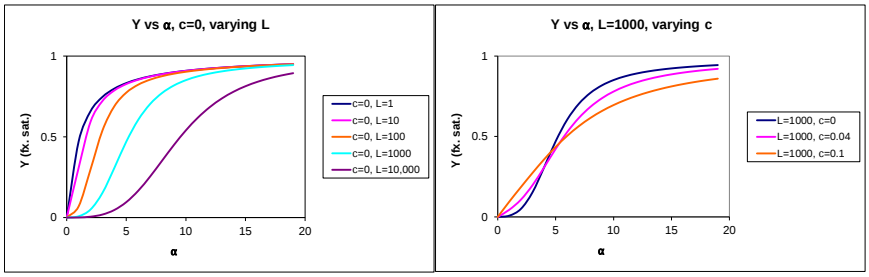


Figure 5.3.16: fractional saturation (Y) vs alpha varies with L and c for the MWC model

When L is set at 9000 and c = 0.014, the Y vs α curve fits the experimental oxygen binding data well. Figure 5.3.17 shows the best experimental dioxygen binding data that we could find (obtained from a graph, not from a table), the best fit of the Y vs L data using a Hill coefficient of n=2.8 (fitting equation 3 above), and the best fit of Y vs L using the MWC model, with L=9000, c=0.014, and Kr = 2.8 torr.

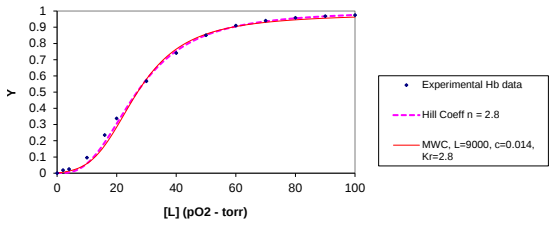


Figure 5.3.17: Hb binding curves: Experimental vs Theoretical Hill and MWC Equations

Use the sliders in the interactive graph below to explore the effect of changes in L and c on fractional saturation.

Hence, like the Hill equation, the MWC equation gives sigmoidal dioxygen binding curves. It does **not** require an empirical Hill-like coefficient, which has no clear physical meaning!

Another way to think about the MWC Model

The MWC model assumes that oxygen binds to either the T or R form of Hb in a noncooperative fashion. Hence K_T and K_R are constant, independent of the number of dioxygens bound to that form. If so, what is the basis of the cooperative oxygen binding curves? The answer can be seen below. The cyan curve might reflect the binding of a ligand to the T form of a macromolecule, with K_D = 100 uM (low affinity), for example. The binding curve looks linear, but it really is just the initial part of a hyperbolic binding cure. Likewise, the magenta curve reflects the binding of a ligand to the R form of the macromolecule with K_D = 10 uM. The T and R forms are linked through the T ↔ R equilibrium. That equilibrium will be shifted to the tighter binding (lower K_D) R form with increasing ligand concentration, assuming the ligand binds preferentially to the R form. This shifts the actual binding curve from that resembling the T form at low ligand (cyan) to one resembling the R form (magenta) as the ligand increases, imparting sigmoidal characteristics to the "observed" binding curve (gray). Figure 5.3.18 shows how sigmoidal binding curve could arise from a switch from a low affinity to high affinity form.

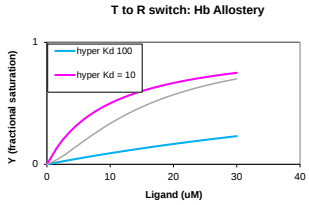


Figure 5.3.18: Cooperativity as transition between two hyperbolic binding curves with different Kd's

5.3.5.3: KNF Sequential Model

The KNF (Koshland, Nemethy, and Filmer) Sequential model was developed to address concerns with the concerted model. One of the major problems with the concerted model is that it seemed unrealistic to expect all the subunits to change conformation together. Why shouldn't there be some differences in subunit conformation? The KNF model also fits the experimental data well. Figure 5.3.19 shows the linked equilibria in the KNF model. Data suggests that the MWC model better explains the transition in proteins on ligand binding and that there is an all-or-none interconversion between the two states.

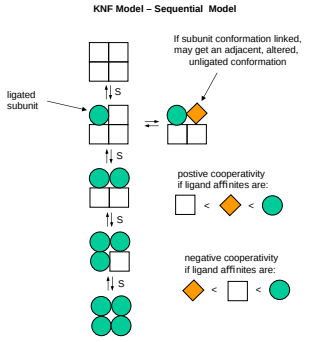


Figure 5.3.19: KNF model for cooperative binding

5.3.6: Allosterism in other multisubunit protein complexes

Changeux (of the MWC model) has written eloquently about the occurrence and effects of allostery in other proteins. We will encounter these proteins in other chapters, but present some here, in advance of the chapter in which they are usually discussed. We do this to show that other proteins display allostery and that the MWC can often be used in describing their behaviors. This offers a rationale to discuss allosterism using hemoglobin with its nonstandard covalent ligands as a model for allosteric binding proteins and enzymes.

Environmental factors such as ligands and allosteric modulators can shift the degree of cooperativity for ligand binding, promote allosteric rearrangements and T ↔ R transitions of proteins other than hemoglobin. We offer several examples of multimeric proteins (complexes) that display allosterism. Many of these allosteric proteins not only bind ligands, but acts as catalysts. One protein, a **ligand-gated ion channel**, moves ions across a membrane. Others catalyze the chemical transformation of a substrate to a product. Another is a structural viral protein. The examples involving catalysis are more complex, since an additional step (transport of ions or alteration in covalent bonds) after binding is added to effect protein function. This extra step can be described as a rate, so we explore rate vs ligand concentration, not just fractional saturation vs ligand concentration curves.

Lactate dehydrogenase (LDH)

LDH is an enzyme that catalyzes the reversible reduction of the 3-carbon carboxylic acid pyruvate to lactate by the oxidizing agent NAD^+, as shown in the reaction below (which is written in reverse as the reduction of pyruvate, the normal function of the enzyme).

$$\text{pyruvate} + \text{NADH} + H^+ \leftrightarrow \text{lactate} + NAD^+$$

Its activity is modulated by the allosteric activator fructose 1,6-bisphosphate (FBP). The kinetics can be modeled using the MWC model, in which the enzyme exists in T (tense/taut) and R (relaxed) allosteric states. FBP binds preferentially to the R state.

Figure 5.3.20 shows an interactive iCn3D model comparing the T state of **bacterial L-lactate dehydrogenase with bound NAD^+ from** Bifidobacterium longum (1LLD), and the R state of the enzyme from Geobacillus stearothermophilus (2LDB) with bound NAD^+ and the allosteric activator fructose 1,6-bisphosphate (F6P). Toggle between the two states using the "a" key.

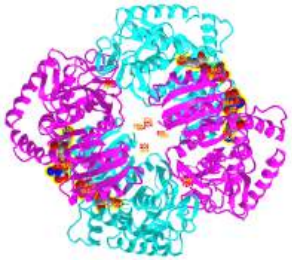


Figure 5.3.20: Comparison of the T (1LLD) and R (2LDB) states of bacterial L-lactate dehydrogenase with bound NAD^+ and allosteric activator F6P (in R state) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...YEXKSp5J1ErFT6

The enzyme substrate NAD^+, the allosteric activator F6P for the R state, and SO_4^{2-} (from ammonium sulfate used to crystallize the protein) are shown in spacefill and labeled.

Aspartate transcarbamylase (ATCase)

This enzyme catalyzes the addition of aspartate and carbamoyl phosphate to form carbamoyl aspartate, the first step in the pathway for the synthesis of the pyrimidine nucleotides cytidine triphosphate (CTP) and uridine triphosphate (UTP).

The end products of the pathway, CTP and UTP, feedback and allosterically inhibit the enzyme. In contrast, ATP is an allosteric activator. This prevents a buildup of pyrimidine nucleotides over purine nucleotides since equal amounts are needed for nucleic acid synthesis.

Figure 5.3.21 shows an interactive iCn3D model comparing the T (**4FYW)** and R (**1D09**) states of asparatate transcarbamylase (ATCase). Toggle between the two states using the "a" key.

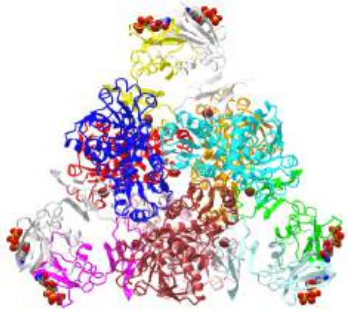


Figure 5.3.21: Comparison of the T tense (**4FYW)** and R (**1D09**) relaxed state of asparatate transcarbamylase (ATCase). Toggle between the two states using the "a" key. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...MHLtC9ALwWVia6

Each of the subunits is shown in a different color. The T state (4FYW) has bound CTPs (at the periphery, shown in spacefill) while the R state has a bound substrate analog, N-(phosphonacetyl)-L-aspartic acid. High levels of the substrate (or substrate analog) shifts the equilibrium to the active R state.

Pentameric ligand-gated ion channels (bacterial)

Protein channels in membrane bilayers are needed to "catalyze" and regulate the flow of ions across the hydrophobic membrane. Hence it makes sense that channels exist in closed and open states. One example is the bacterial GLIC pentameric ligand-gated ion channel, which is opened by ligand binding, often called ligand-gating..

Figure 5.3.22 shows an interactive iCn3D model comparing the GLIC pentameric Ligand-Gated Ion Channel Loop2-22' oxidized mutant in a locally-closed conformation (LC3 subtype) (3TLV) and the A237F mutant channel in the open conformation (3LSV). Toggle between the two states using the "a" key.

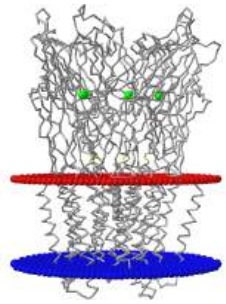


Figure 5.3.22 Comparison of the GLIC pentameric Ligand-Gated Ion Channel Loop2-22' oxidized mutant in a locally-closed conformation (LC3 subtype) (3TLV) and the A237F mutant of the pentameric ligand gated ion channel from Gloeobacter Violaceus in the open conformation (3LSV). Toggle between the two states using the "a" key. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...GKiPY9Jh7tQXTA

Note that the outer (red) and inner (blue) membrane leaflets are shown only in the closed channel (3TLV). The green spheres represent chloride ions.

Nudaurelia capensis ω virus capsid

This hallow viral protein structure surrounds the internal viral genome, so it is an example of allostery in a protein complex that is neither a transporter nor an enzyme. This hetero 480-mer with icosahedral symmetry undergoes a global shape change when the immature capsid undergoes selective and limited proteolysis to form the mature capsid, as illustrated in Figure 5.3.22

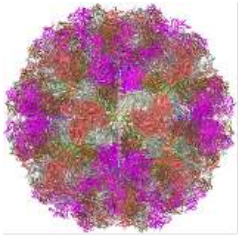


Figure 5.3.22: Comparison of the immature Nudaurelia capensis omega virus capsid: virus-like particles (7ANM) and the mature version (7ATA) (with a hole in the center)

The R form is more open. This is a wonderful example to envision the global conversion of all subunits from a "T" to an "R" state, clearly necessary in this case to preserve the exquisite symmetry!

Now, let's look at binding and rate curves for some multimeric allosteric enzymes. Since this is all a bit complicated, let's review again the difference between what we call Type 1 and Type II allosterism:

Type I: Increasing the amount of a substrate can induce conformational changes in a multisubunit protein to a form that has apparently higher (or potentially lower as well) affinity for the substrate in the remaining unoccupied substrate binding sites. In this case, the substrate is binding to the **orthosteric** site. These sites are where substrates bind but also competitive inhibitors (if the protein is an enzyme), and agonists or competitive antagonists of receptors. We will explore enzymes and receptors later in this book. In Type I allosterism, binding or kinetic curves would show sigmoidal fractional saturation (or kinetic curves) with increasing concentrations of substrate.

Type II: Increasing amounts of a chemical species (an inhibitor or activator) can bind to an **allosteric** site, which affects the binding of the substrate to the **orthosteric** site. The regulators shift and change the shape of the Y or rate curves vs substrate. In experiments to show this kind of allosterism, you wouldn't change the substrate and allosteric effector concentrations simultaneously since the resulting data and graphs would be hard to interpret. You could change the ligand or substrate that binds to the orthosteric site over a large range of concentrations (hopefully over a 1000 - 10,000 fold change, or 4 log units) in several different experiments, with each experiment having a different fixed concentration of the allosteric effector. Alternatively, you could conduct the experiment over a large concentration range of a given allosteric effector (again a 1000-10,000 fold change if possible) in several different fixed concentrations of ligand or substrate in a series of experiments.

5.3.7: Rate vs ligand curves for allosteric proteins that catalyze chemical reactions

Since we have already seen an example of Type I allosteric binding curves (hemoglobin binding dixoygen), let's look at a few examples of Type II allosteric in multisubunit proteins since their graphs are a bit more complicated. We realize the curves below show relative rates of enzymes and not relative fractional saturation of enzymes, but the same principles are present.

Phosphofructokinase

Figure 5.3.23 shows an example of allosteric kinetic curves for Phosphofructokinases A (Pfk A) and B (Pfk B) from *Mycobacterium tuberculosis*. The enzyme catalyzes the phosphorylation of fructose-6-phophate (F6P) by ATP to produce fructose-1,6-bisphosphate (F1,6-BP) and ADP.

$$F6P + ATP \rightarrow F1,6\text{-}BP + ADP$$

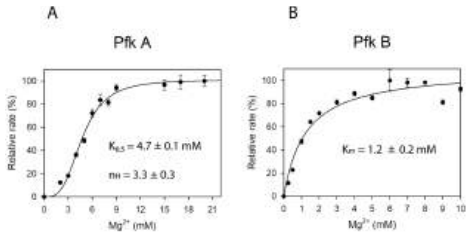


Figure 5.3.23: The dependence of Pfk A and Pfk B activities on concentration of Mg^{2+}. Individual reactions were performed in buffers containing fixed initial concentration for both substrates (1 mM F6P and ATP) with the concentration of Mg^{2+} varied. Snášel, J. et al. *Int. J. Mol. Sci.* **2021**, *22*, 1483. https://doi.org/10.3390/ijms22031483. Creative Commons Attribution (CC BY) license (**http://creativecommons.org/licenses/by/4.0/**).

Only Pfk A shows allosteric activation of the enzyme by Mg^{2+}, run under fixed initial (and probably saturating) concentrations of the substrates F6P and ATP. The Hill coefficient is 3.3 for Pfk A, which suggests that Mg^{2+} is important in maintaining/promoting the active site and formation of the enzyme tetramer. Pfk B shows hyperbolic kinetics and no allosterism, with a Hill coefficient close to 1. These curves are modeled with the Hill equation and not the MWC equation.

Lactate Dehydrogenase

Again this enzyme catalyzes the following reaction:

$$pyruvate + NADH + H^+ \leftrightarrow lactate + NAD^+$$

The graphs in Figure 5.3.24 show relative inhibition (graph A, top) and double-reciprocal plots (C, bottom) for the enzyme lactate dehydrogenase B (LDHB) in the presence of an allosteric inhibitor, AXKO0046. This enzyme catalyzes the reduction of pyruvate by NADH (the substrate) to form lactic acid and NAD^+ (the products). We'll discuss the graphs below.

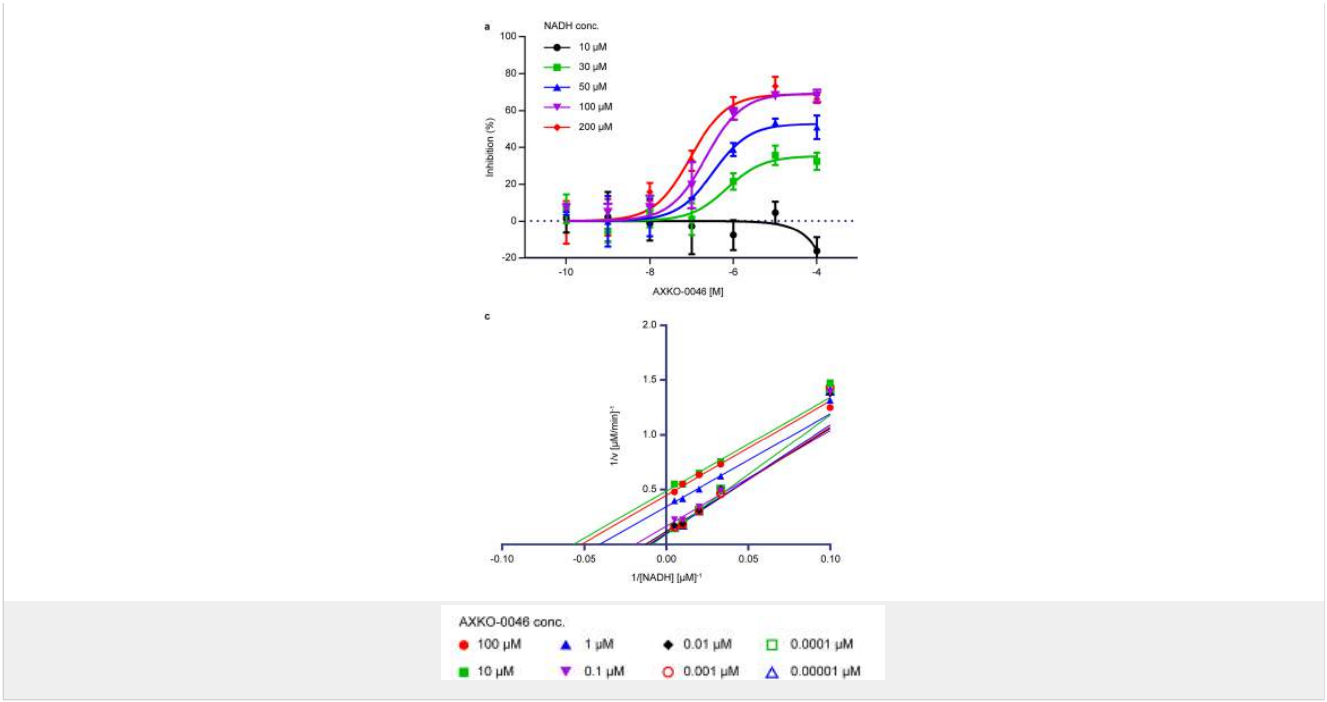


Figure 5.3.24: Biochemical characterization of AXKO-0046. LDHB inhibition by AXKO0046 was studied using varying concentrations of (a) NADH (c) Double reciprocal (Lineweaver-Burk) plots of the kinetic data. Shibata, S. et al. *Sci Rep* **11**, 21353 (2021). https://doi.org/10.1038/s41598-021-00820-7. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Now, these graphs are a bit more complicated since in this case, the initial concentration of one substrate is varied, while the other concentration is fixed (in contrast to the PfkA experiments in which the initial concentrations of both substrates were held constant).

Let's look at graph C (bottom) first. Instead of showing graphs of rate vs [NADH], the authors showed double-reciprocal kinetic plots, run with varying [NADH] and at one fixed concentration of pyruvate (not given) and several fixed concentrations of inhibitor. The graphs look like they are generally straight lines except at the end, which occurs at a low concentration of NADH (which gives the highest value of 1/[NADH] = 0.10). At this point, 1/rate (given as 1/v) data points are higher than the best-fit line would suggest, implying that the rate v is "abnormally" low. That rate accelerates as [NADH] increases (as 1/[NADH] decreases in a manner consistent with allosterism. This suggests that at any given **fixed** concentration of inhibitor and fixed pyruvate concentration, the graphs of rate vs substrate (NADH) (i.e not the double-reciprocal plot) would be sigmoidal. It's different from the graph for PfkA (Fig 23), in which the x-axis variable is the allosteric activator Mg^{2+}.

Look at Graph A (top) which uses varying inhibitor concentrations and several different fixed concentrations of reactant NADH. This is analogous to the graph for PfkA but notice that on the x-axis the log [inhibitor] is plotted instead of the [inhibitor]. These are NOT plots of v vs [substrate], expected to be hyperbolic, or v vs log[substrate] which are expected to be sigmoidal. (A lesson here is to look carefully at the axes). But note something unusual about the curves. The plateau for inhibition is not the same at each NADH concentration. The highest % inhibition (red and purple curves) occurs when the substrate [NADH] is highest. (We will see in the next chapter that this is a sign of what is called uncompetitive inhibition).

X-ray crystal structures show that the inhibitor (AXKO-0046) does bind to an allosteric site, not the active orthosteric site. It appears to bind in the interface of the LDHB tetramer. The graphs show that over four orders of magnitude of inhibitor (4 log units), the inhibition goes from 0 to about 100%, which is expected if the sigmoidal semi-log curves gave hyperbolic curves with [inhibitor] plotted on the x-axis.

This may seem confusing, but such sigmoidal curves are found in plots of rate vs log concentration of allosteric activators and inhibitors, as discussed in Chapter 5.1. So don't immediately jump to the conclusion that a sigmoidal curve implies allosterism. Look at the reactions and relative concentrations carefully.

Consider this example. What if a protein binds a ligand L and an inhibitor I at the same orthosteric site? If one bound, the other couldn't. This is an example of a classical competitive, non-allosteric, inhibition. Now, what if an inhibitor, I, binds to an allosteric site and when bound, it altered the conformation of the orthosteric site such that the ligand could not bind. Binding of L and I would be mutually exclusive. This would produce the same binding curves as the classical competitive inhibition. In either case, at very high ligand concentrations, the effect of the inhibitor would be lost, and full maximal binding would be observed. It would just take higher concentrations of ligand to get the same fractional saturation of the protein in the presence of the inhibitor than in its absence. In the presence of a fixed concentration of these competitive inhibitors, the effective K_D would be higher. Y vs L curves for both would be hyperbolic, and double-reciprocal plots would be linear.

5.3.8: Allostery within monomeric protein

Allosterism can also occur in monomeric proteins:

Type I (again our nomenclature) allosterism can occur in monomers that have more 2 or more binding sites for a ligand/substrate and if the binding of ligand/substrate to one site significantly alters the affinity of the other site for substrate enough to produce a nonhyperbolic, sigmoidal binding/kinetic curve for substrates. This case is different than the binding of substrate to two different preformed substrate binding sites, each with a different fixed affinity, which we discussed in Chapter 5.1 (scroll down to binding of a ligand to two independent sites). We show again the graph of fractional saturation Y vs L for the binding of a ligand to two different preformed sites of different affinities below.

Note that the above graph doesn't look sigmoidal. It is essentially hyperbolic except in the extreme case when one of the K_D is much much less than the other AND at low ligand concentration such that the higher affinity binding leads to an abrupt, titration curve-like saturation of the low K_D site before the second site has much occupancy.

Rec A is an OK example of a "possible" Type I allosteric monomer binding protein (if you have a better example, let us know!). This protein is required for homologous recombination in bacteria. It has ATPase activity and catalyzes ATP-driven homologous pairing and strand exchange of DNA required for DNA repair. The structure is known for the Mycolicibacterium smegmatis apo form of the enzyme, the enzyme:substrate (dATP, a substrate analog) complex, and the enzyme:substrate:allosteric effector (a second dATP and possibly citrate) complex.

The enzyme has three domains (N-terminal 1-30, the major M domain (31-269) and the C-terminal (270-349). The M domain is the catalytic domain which has nucleotide triphosphate hydrolase activity. It binds nucleotides, DNA and also interacts with the N domain of another RecA to promote the polymerization of RecA into a filament. The C-terminal domain is disordered but becomes ordered when bound to a second dATP in the crystal structure.

Figure 5.3.25 shows conformational changes in RecA:dATP (the ES complex) on binding a second dATP (the ESA complex), where A is the likely allosteric activator (the second bound dATP). A result of this ordering on binding is likely the polymerization of the RecA into filaments.

Figure 5.3.25: Conformational changes in RecA:dATP (the ES complex) on binding a second dATP (the ESA complex).

The dATP in the catalytic site is shown in spacefill with CPK colors. The ES complex is shown as a darker gray protein with one bound dATP (spacefill, CPK colors). The ESA complex is shown in lighter gray with dATP bound in the catalytic (orthosteric) site in CPK colors and a second dATP (spacefill, cyan) bound in the putative allosteric site in the C domain.

Type II: increasing amounts of a chemical species (an inhibitor or activator) that binds to an **allosteric** site in a monomeric protein could affect the binding of the substrate to an orthosteric site in the monomer. In this case, as in Type II for multimeric proteins, you could again run two different types of experiments (one with varying substrate at 3-4 different fixed allosteric effector concentrations or vice/versa.

One example is thrombin, the last protease in a cascade of clotting proteins. The proteins are synthesized as inactive precursors (zymogens) that become activated on limited proteolysis. Active thrombin is a procoagulant enzyme in that it can cleave circulating fibrinogen (and other procoagulant molecules) into fibrin. This then self-associates to form a fibrin clot.

Paradoxically, thrombin also has anticoagulant properties in that it can cleave another circulating protein, Protein C, which e inhibits further clotting. Thrombin does so when it binds a transmembrane protein, thrombomodulin, present on endothelial cells that line blood vessels.

These contrasting activities support the notion that thrombin has two interconverting conformations, each stabilized by different ligands or proteins. One such ligand is the simple monatomic ion Na^+. Indeed, thrombin appears to have two main catalytic conformations, a high-activity “fast” form (with bound Na^+) and a low-activity “slow” form (without bound Na^+). The fast form with bound Na^+ (15 Å from the active site) appears to be the procoagulant form while the slow form is the anticoagulant form.

Figure 5.3.26 shows an interactive iCn3D model comparing the anticoagulant slow form of thrombin (1SGI) and the procoagulant sodium-bound fast form of thrombin (1SG8). Toggle between the two states using the "a" key.

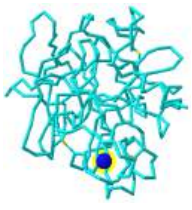


Figure 5.3.26: Anticoagulant slow form (1SGI) and the procoagulant sodium-bound fast form of thrombin (1SG8). Toggle between the two states using the "a" key. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...kxnATvMdVD1566

The magenta represents the slow form and the cyan with bound Na^+ the fast form which has enhanced coagulant activity

Table 5.3.1 below shows a list of monomeric allosteric proteins and their PDB file codes. The proteins (P) are enzymes that bind a substrate (S) and an allosteric effector (A) to form PS, PA, PAS complexes (adapted from Wang et al. J. Phys. Chem. Lett. 2021, 12, 5404−5412)

protein	P	PA	PS	PAS	Effect

protein RecA (RecA)	2OES		2ODN	2G88	activation
mitogen-activated protein kinase 8 (MAP8)	1UKH	3O2M	2XRW		inhibition
cAMP-dependent protein kinase catalytic. subunit alpha (Prkaca)	4NTS	4NTT	4IAF		inhibition
cAMP-dependent protein kinase catalytic. subunit alpha (Prkaca)	4NTS	1BKX	4DG0		activation
cyclin-dependent kinase 2 (CDK2)	3PXR	3PXF	1HCK		inhibition
casein kinase II subunit alpha (CK2α)	5ZN5	3H30	2PVR		inhibition
myosin-2 heavy chain (mhcA)	1FMV		2JJ9	2JHR	inhibition
tyrosine-protein phosphatase. nonreceptor type 1 (PTP1B)	4QBW	1T49	1PTV		inhibition
		1T48			
		1T4J			

5.3.9: Hemoglobin variants and disease

Many diseases have been associated with alternations in the amino acid sequence of hemoglobin. Around 300,000 babies are born each year with a genetic disorder causing an aberrant hemoglobin structure. Over 80% of these are in low/middle-income countries. The worst is sickle cell anemia followed by β-thalassaemias and the less serious α-thalassaemias. We will only focus on sickle cell disease. In this disease, red blood cells become distorted in share from a normal smooth discoid shape to a crescent-like shape under low oxygen concentrations found in capillaries. These impede blood flow as shown in Figure 5.3.25 and lead to symptoms ranging from anemia, episodic pain, swelling of hands and feet, and vision problem. Complications often lead to premature death.

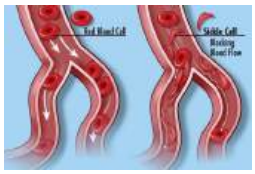


Figure 5.3.25: Sickling of red blood cells in sickle cell disease. https://commons.wikimedia.org/wiki/File:Sickle_Cell_Disease_(27249799083).jpg. National Human Genome Research Institute (NHGRI) from Bethesda, MD, USA. Creative Commons Attribution 2.0 Generic license

Linus Paul and colleagues showed that the isoelectric point (pI) of oxy-Hb and deoxy-Hb in normal blood was 6.87 and 6.68, respectively, but were higher (7.09 and 6.91, respectively) in sickle cell disease. This was the first evidence that a disease was caused by a molecular alteration of a protein. Eventually, we learned that a single negatively charged amino acid, Glu 6, on the β-chain of hemoglobin was mutated to a nonpolar one, valine. This causes hemoglobin, which is present at a concentration of 150 g/L in blood to self-aggregate into a long polymer. This distorts the cell in a sickle shape. Humans have two genes for the beta chain of hemoglobin, one from the egg donor and the ohter from the sperm donor. If only one mutated, the disease is called sickle cell **trait**. If both are mutated, sickle cell **disease** is observed.

The hydrophobic Val 6 in on the surface of both beta chains in sickle cell disease. It can bind to a hydrophobic patch comprised of Ala 70, Phe 85 and Leu 88 on another β-chain on another hemoglobin tetramer. Hence hemoglobin has two Val 6s and two hydrophobic patches, allowing first the formation of a sickle cell hemoglobin "dimer" of tetramers, followed by elongation of the growing fibril. This is a disease of aberrant induced dipole-induced dipole interactions and the "hydrophobic effect".

Figure 5.3.27 shows a "dimer" aggregate of two hemoglobin S tetramers, in which each β-chain has the D6V mutation. α chains are shown in grey, β chains in cyan, Val 6 in red spacefill and the surface hydrophobic patch of A70, F85 and L88 in orange spacefill (1hbs). Note the binding of the two tetramers is mediated by the interaction of the red Val 6 on the right tetramer with the orange surface hydrophobic patch on the left tetramer. Note also that there are the "dimer" aggregate has three more exposed Val 6 (red sphere) and three more hydrophobic patches. This would allow the extension of the dimer and the formation of long fibril-like polymers, with binding mediated by noncovalent interactions.

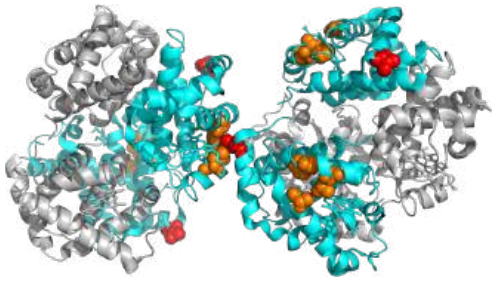


Figure 5.3.26: Dimer of hemoglobin S tetramers showing the interaction of Val 6 (red sphere) with a surface hydrophobic path (orange) on a second tetramer.

Figure 5.3.28 shows an interactive iCn3D model of one tetramer of hemoglobin S. As in Figure 5.3.21, the α chains are shown in grey, the β chains are shown in cyan, and Val 6s are shown in red spacefill. The **surfaces** of the hydrophobic pockets where the Val 6 another HbS tetramer binds, comprised of amino acids A70, F85 and L88, are shown in orange.

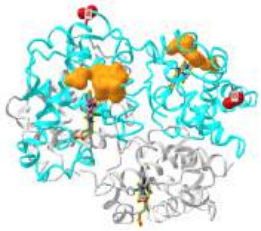


Figure 5.3.28: Hemoglobin S tetramer (1HBS) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...APyWAsDLnbAeR7

Sickle cell disease and trait are endemic in Sub-Saharan Africa, where over 4 million have the disease and over 40 million have the trait. Its geographic distribution, along with that of malaria, is shown in Figure 5.3.29

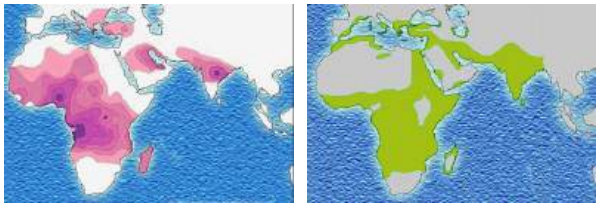


Figure 5.3.29: Maps showing distribution of sickle cell trat and malaria. https://commons.wikimedia.org/wiki/F...stribution.jpg. CC-LAYOUT; CC-BY-SA-2.5,2.0,1.

The Plasmodium parasite reproduces in red bloods cells. Their ability to reproduced is compromised as red blood cells with sickle cell hemoglobin rupture more frequently. Also the parasite uses hemoglobin as a source of amino acids. The endocytose it and hydrolyze it to amino acids in digestive organelles. Sickle cell hemoglobin is more resistant to this process. Hence evolution appears to have maintained the sickle cell gene in these areas as protection against malaria.

Sickle cell disease is a systemic problem, so treatment of the multitude of symptoms is important. Some of these treatments are described in Figure 5.3.30

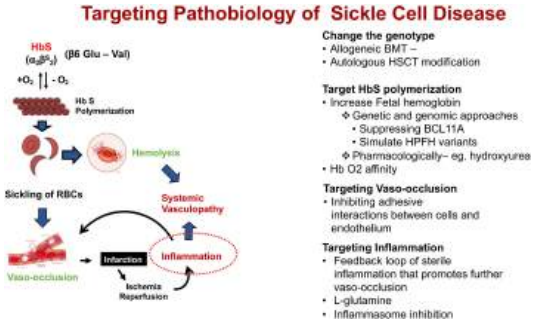


Figure 5.3.30: Treatment of sickle cell disease. Cisternos et al. Front. Physiol., 20 May 2020 | https://doi.org/10.3389/fphys.2020.00435. **Creative Commons Attribution License (CC BY)**.

Of course the cure would be to use DNA editing to change the base pair for the mutated Val 6 gene back to the wild-type Glu 6 as shown in Figure 5.3.31 The amino acid sequence encoded by the DNA and RNA shown below is Pro-Glu-Glu for the normal and Pro-Val-Glu for the sickle cell chain.

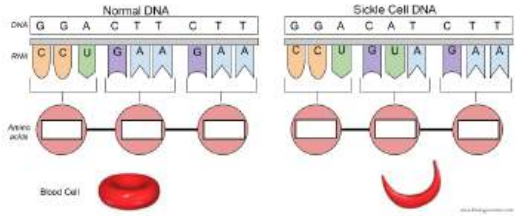


Figure 5.3.31: Mutation in sickle cell disease. https://www.biologycorner.com/worksh...icklecell.html. Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Crisper gene editing trials are now underway to attempt a cure of this dreadful disease.

5.4: Complementary Interactions between Proteins and Ligands - The Immune System and Immunoglobulins

5.4.1: Introduction to the Immune System

Now let's consider the daunting task faced by the immune system - to recognize all possible "foreign" molecules and react to them, either by targeting them for elimination or, paradoxically, to recognize them but not react to them (a process called tolerance). The same can be said of "self-molecules., and the immune system must recognize them but not respond to them, otherwise autoimmune disease might arise in which the body's powerful immune system targets self.

It is virtually impossible to give an in-depth description of the immune system in a short section. Our goal is to illustrate how the immune system recognizes such a vast number of molecules. We will briefly cover the innate and adaptive immune system and their differences, and how some cells (macrophages in particular) in the innate immune system and cells (B and T cells) in the adaptive immune response recognize and respond to target molecules and cells. Finally, we'll discuss how the immune system can respond to similar molecules through the recognition of common molecular patterns. Emphasis will be given to recognition. Ways to simplify the complexities of the immune system have been presented in a fantastic book written by Lauren Sompayrac, How the Immune System Works. (2003, Blackwell Publishing. ISBN: 0-632-04702-X) and adopted here.

We realize that we have not yet reached the chapters on carbohydrates, membrane proteins, and nucleic acids. Nevertheless, we present the material in this section to organize it in one specific location. Users can revisit this page after they studied subsequent chapters.

Before we start, think of the variety of chemical species that the immune system should recognize as foreign:

- a bacterial glycan or glycolipid on the outside of the cell
- a viral surface protein, such as the spike protein of the SARS Coronavirus 2
- bacterial dsDNA (and not host dsDNA)
- viral dsRNA (which is not common in host systems
- a self protein that has been modified in a tumor cell
- a crystal of urea
- extracellular ATP (a place where it is not usually found)
- a silica particle found in particles like asbestos.

How would you design an immune system to bind each of the "enemy" targets above? That is what we will explore in this section - the binding interactions. What happens after the binding is beyond the scope of this section and falls generally in the field of signal transduction - how binding events at the self surface are transferred into intracellular responses.

Three lines of defense protect us from the "enemies", foreign substances (bacteria, viruses and their associated proteins, carbohydrates, and lipids) collectively called antigens.

- physical barriers of cells that line our outside surface and our respiratory, GI tract, and reproductive systems
- , the innate immune system (IS) that all animals have. Composed of scavenger cells like macrophages (MΦ), neutrophils, dendritic cells, and natural killer cells (NK) that can move around the body through the blood and lymph systems and burrow into tissues to meet the enemy where they can engulf and destroy bacteria and "cellular debris". Macrophages start off as immature circulating monocytes, which enter tissues by slipping through blood vessel walls. They differentiate into macrophages. There they lie in wait ready for the enemy.
- the adaptive immune system, which, as its name implies, can change and adapt to new molecular threats. This branch is better at dealing with viruses, which do their damage inside host cells. The adaptive IS is comprised of **B cells** that make and secret protein antibodies that recognize specific foreign molecules, and **T cells**.

In a world experiencing the most deadly pandemic (Covid) of the last 100 years, and with more to come, immune recognition must be an important part of any biochemistry text. This chapter section could be a whole chapter, but we'll leave it as a very long section. Let's start with the adaptive immune system, which we can coopt to make vaccines to the major threats we face.

5.4.2: B Cells and Antibodies

B cells and their differentiated forms (B memory and plasma cells) make antibodies. Antibodies bind to foreign molecules (proteins, glycans, lipids, etc), which might neutralize their effects. For example, an antibody can bind to the hemagglutinin

molecule of the influenza virus and prevent its entry into cells. We all are now familiar with the utility of vaccines that create antibodies to recognize the spike protein of the acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Antibodies also bind to foreign cells like bacteria, which signals other host immune proteins and cells to come in for the kill. Antibodies are secreted by B cells, which also have a membrane-bound form of the antibody on their surface. This antibody acts like a receptor which binds antigen and, through a signal transduction process, helps to activate the B cell. Mature B cells (those that have previously seen antigen) can secrete lots of antibodies quickly. Surface and secreted antibodies can recognize and bind to almost any molecule.

There are many forms of antibody, also called immunoglobins (Ig). These include IgA, IgG, IgM and IgD. We will concentrate on the structure of IgG. It consists of 4 chains (a tetramer) of two light chains and two heavy chains. The light chains form disulfide links with the heavy chains and disulfides also link the heavy chains. Effectively, it's one big protein molecule (about 160 K). Figure 5.4.1 shows a spacefill, secondary structure, and geometric cartoon rendering of a mouse IgG protein (pdb ID 1IGT).

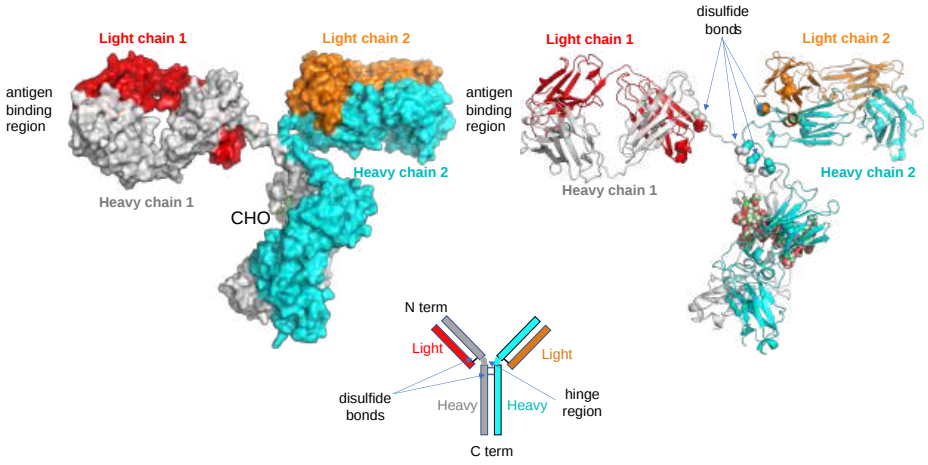


Figure 5.4.1: Renderings of an IgG antibody

The antibody is shaped like a **Y**. Foreign molecules (antigens) bind at the end of the top tips of the Y with both chains contributing to antigen binding. The structures of both are dominated by antiparallel beta sheets.

Each chain consists of a single N-terminal variable domain (V_L or V_H) which participate in antigen recognition. The light chains have an additional constant domain (C_L) while the heavy chains have 3 constant domains (C_{H1}-C_{H3}). The constant domains are not involved in antigen recognition, but they are involved in effector functions (such as the binding of other immune molecules like complement proteins) to the antigen-bound antibody heavy chain constant regions. Each of the domains is about 100 amino acids. A cartoon structure with the domain structures is shown in Figure 5.4.2.

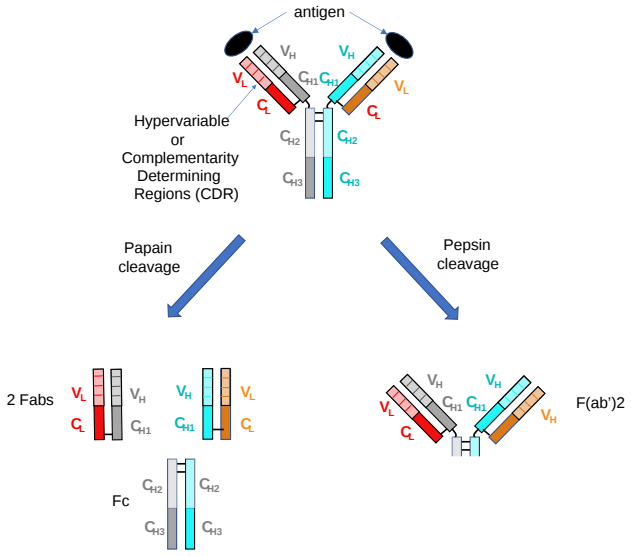


Figure 5.4.2: IgG antibody and fragments

Two other features are depicted in the above figure. In each variable region of both light and heavy changes, there are **hypervariable** regions, which contribute to the unique binding features of a given antibody. The regions are also **called complementarity-determining regions (CDRs)**. Membrane-bound forms of antibodies that serve as "receptor" proteins have additional domains now shown in the figure above. The binding site on the antigen recognized by the antibody is called the **epitope**. The corresponding binding site on the Y-shaped ends of the antibody that recognized the antigen is called the **paratope**.

In Figure 5.4.2, you can see that the intact full IgG molecule has in 12 variable V and constant C domains called the **immunoglobulin domain**. Each has about **110 amino acid** in length, two layers of β-sheets each with 3-5 antiparallel β-strands with a disulfide bond connecting the two layers.

When we discussed domain structure, we indicated that proteins with multiple binding domains can often be selectively cleaved with protease, with cleaved fragments often retaining binding and other functions properties. The same is true with antibodies. Cleavage with either the proteases pepsin or papain forms fragments with binding activities as illustrated in the figure above. Selected protease digestion was used to clarify structure/function relationships in antibody recognition.

When antibodies targeting different antigens were sequenced, it was clear that much variability was found among antibodies in the variable domains of both the light and heavy chains. In those domains there were also hypervariable regions. The origin of the variability and hypervariability arise mostly from an extremely large number of gene segments (also exons) in the gene encoding the variable domains. The exons can be spliced together at both the DNA and RNA levels to produce many different DNA/RNA sequences. These are decoded into the variable and hypervariable regions of the light and heavy chain proteins. Somatic mutations are also enhanced in this region.

In-depth: Generation of Antibody Diversity

We mentioned above that both DNA and RNA splicing occurs as B cells mature to become antibody-secreting cells (plasma cells). For those who have studied the Central Dogma of Biology, splicing for primary RNA transcripts should come as no surprise. What's is surprising is that the DNA genome of B cells changes on their maturation due to splicing of a multiple of exons within the variable chain genes to produce unique coding sequences for each clone of a given B cell. There are sets of exons (V, D and J) or segments within the genes for the variable chain. As the immune cells terminally differentiate, a unique combination of a VDJ segment forms in the DNA genome, so each terminally differentiated B cell is different. When needed (i.e. when their unique antigen binds to membrane forms of the antibody), the cell secretes a monoclonal antibody.

Figure 5.4.3 shows how the different segments become linked in the DNA and how they can be uniquely spiced in the RNA to produce a unique, monoclonal antibody.

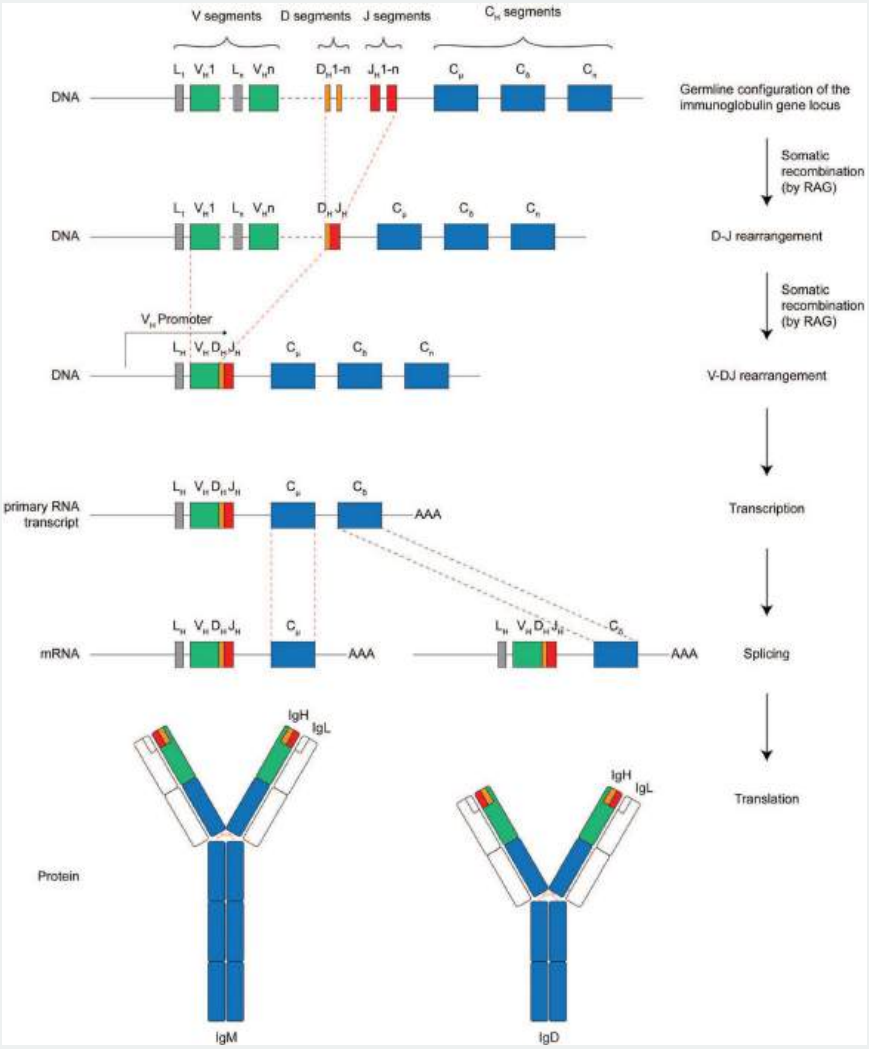


Figure 5.4.3 Generation of Antibody Diversity. Oliver Backhaus DOI: 10.5772/intechopen.72818. Creative Commons Attribution 3.0 License,

The first antibodies produced by the immune system are often of low affinity. Over time, high affinity (low K_D) antibodies are produced. What differentiates high and low affinity binding at the molecular level? Do high affinity interactions have lots of intramolecular H-bonds, salt bridges (ion-ion interactions), or are hydrophobic interactions most important? Crystal structures of many antibody-protein complexes were determined to study the basis of affinity maturation of antibody molecules. Clones of antibody-producing cells with higher affinity are selected through binding and clonal expansion of these cells. Investigators studied the crystal structure of four different antibodies which bound to the same site or **epitope** on the protein antigen lysozyme. Increased affinity was correlated with increased buried apolar surface area and not with increased numbers of H bonds or salt bridges as described in Table 5.4.1 below.

Antibody	H26-HEL	H63-HEL	H10-HEL	H8-HEL
Kd (nM)	7.14	3.60	0.313	0.200
Intermolecular Interactions				
H bonds	24	25	20	23
VDW contacts	159	144	134	153
salt bridges	1	1	1	1
Buried Surface Area				

Antibody	H26-HEL	H63-HEL	H10-HEL	H8-HEL
ΔASURF (A2)	1,812	1,825	1,824	1,872
ΔASURF-polar (A2)	1,149	1,101	1,075	1,052
ΔASURF-apolar (A2)	663	724	749	820

Table 5.4.1: Characteristics of Antibody:Hen Egg Lysozyme Complexes (HEL). Data from Y. et al. Nature: Structural Biology. 6, pg 484 (2003)

Many crystal structures of antibody:antigen complexes have been determined. Especially interesting are those in which the antigen is a protein. It is important to understand antibody:protein antigen interactions to develop vaccines against key epitopes in proteins such as the spike protein of the SARS-CoV-2. Let's look in more detail at the antibody that binds to hen egg white lysozyme (HEWL). The crystal structures of many different IgG antibodies that bind HEWL are known. One recognizes a discontinuous epitope on lysosome consisting of the following amino acids: H15, G16, Y20, R21, T89, N93, K96, K97, I98, S100, D101, G102, W63, R73 AND L75. Most of these amino acids are polar, and five are charged.

Figure 5.4.4 shows the interaction of part of the Fab fragment of an antibody that binds to the HEWL epitope just mentioned (3hfm). The light chain is shown in magenta, the heavy chain in dark blue, and the antigen lysozyme in gray. The side chains of the amino acids in the epitope of HEWL are shown in sticks. Note the complete complementary of HEWL and Fab surfaces. Water is excluded from the interface.

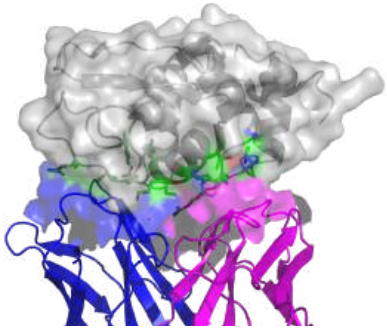


Figure 5.4.4: Surface interactions between a hen egg while lysozyme epitope and an IgG Fab antibody fragment

Figure 5.4.5 shows an interactive iCn3D model of the same HEWL:Fab complex (3hfm). Lysozyme is shown in black.

Figure 5.4.5: Hen egg white lysozyme:Fab complex. (3hfm) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...agQkuAKhgfs177

Here is an external link to an interactive iCn3D model showing a detailed view of the multiple interactions (salt bridges, hydrogen bonds, pi-cation)

5.4.3: T Cells

What happens if a virus makes it into a cell? Antibodies can not bind to them anymore to prevent their entry. Something must be able to recognize a virally-infected cell and eliminate it. What about a cancer cell? Wouldn't it be nice if something could recognize a tumor cell as foreign and eliminate it before it divides too much and metastasizes? Those "something" are T cells. There are many

T cells in a person, and many different kinds, including T helper cells (Th), cytotoxic lymphocytes (CTL), and even suppressor T cell. The express different subsets of proteins that differentiate them and their functions.

T cells also recognize antigens but unlike B cells, these antigens can only be proteins fragments. The membrane proteins that recognize protein fragments are called **T cell receptors**. In addition, they don't recognize protein antigens in isolation. They must be bound to a protein on the surface of an "antigen" presenting cell (such as a macrophage or dendritic cell). The T cell receptor recognizes and binds simultaneously to the foreign protein fragment and to the self "antigen-presenting" protein on the surface of the antigen-presenting cells. The self protein which binds and presents the foreign protein fragments (peptides) is called a **Major Histocompatability Complex (MHC) protein**.

Antigen presenting cells like macrophages and dendritic cells have **MHC Class II** molecules on their surface. These bind protein fragments from engulfed bacteria, for example, and present them on the surface. T cell receptors bind to the peptide:MHC II complex. All cells in the body have **MHC Class I** proteins on their surface. If a cell is infected with a virus, protein fragments from the virus end up bound to the MHC Class I protein on the surface. Now a T cell can bind through its T cell receptor to the peptide:MHC Class I complex. By displaying a viral protein fragment on the surface, the immune cell can recognize a virally-infected cell without getting inside of the cell where the virus is. Sompayrac describes MHC molecules as looking like a hot dog bun. In the grove of the bun lies the peptide fragment - like the hot dog. The T cell receptor recognizes both the bun and the hot dog!

Figure 5.4.6 shows an interactive iCn3D model of a MHC Class Class I heavy chain complexed with beta-2-microglobulin with a peptide fragment of the vesicular stomatitis virus nucleoprotein (2VAA).

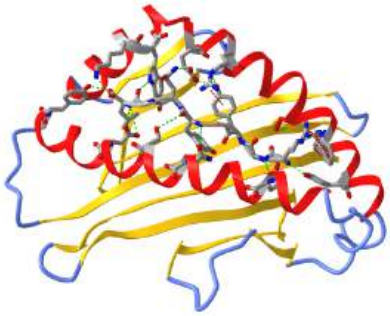


Figure 5.4.6: MHC Class I heavy chain - beta-2-microglobulin - vesicular stomatitis virus nucleoprotein peptide complex (2VAA). (Copyright; author via source). Click the image for a popup or use this external link https://structure.ncbi.nlm.nih.gov/i...NPS9QfLeghCmo8

The **T-cell receptor** consists of two transmembrane protein chains, **alpha** and **beta**, each containing a single variable and constant domain, followed by a transmembrane domain. Hence they are less complicated than an antibody chain. They bind through their extracellular variable domains a peptide fragment bound to a MHC Class I or Class II membrane protein in the target cell.

In-depth: Generation of T-Cell Receptor Diversity

We described above how an undifferentiated B cell has the potential to produce an incredible diversity of different antibodies from a starting genetic sequence. This occurs through both DNA and RNA splicing. The same processes occur with the alpha and beta chains of T-cell receptors. This is illustrated in Figure 5.4.7. Note that the alpha chains have no D (diversity) coding sequences.

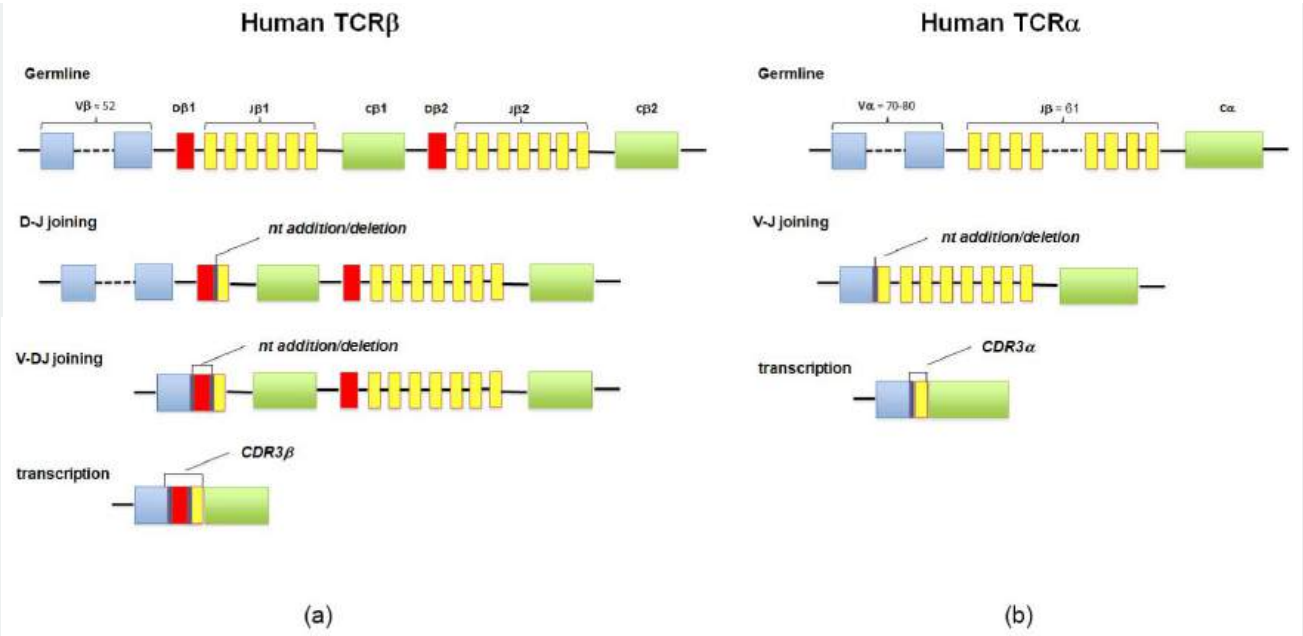


Figure 5.4.7: The diversity of T-cell receptor (TCR)αβ is a result of genetic recombination and diversification mechanisms occurring at the α and β TCR chain loci. Diversity is first created in the germline via recombination of variable V, diversity D (for β chain), and joining J segments. Further diversification occurs through imprecise junctions of these gene segments (addition of P- and N-nucleotides adjacent to the D segment), and the combination of α and β chains

Figure 5.4.8 shows an interactive iCn3D model of the T-cell receptor alpha and beta chains binding to MHC Class 1 protein with a bound peptide (6rp9). The MHC protein complex consists of the histocompatibility antigen, A-2 alpha chain and β-2-microglobulin, an 11K subunit of MHC Class I proteins but not Class II MHC proteins. Bound to it is the 9 amino acid cancer/testis antigen 1 (shown in spacefill). The peptide is sandwiched between the MHC protein complex and the T-cell receptor α and β chains.

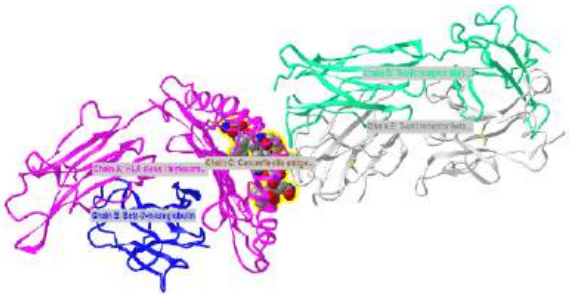


Figure 5.4.8: T-cell receptor alpha and beta chains binding to MHC Class 1 protein with a bound peptide (6rp9) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...txUWWfbNhNXwQ6

The actual functional structure *in vivo* is actually more complicated. The T cell receptor is found within the much larger **T Cell receptor complex (TRC),** which contains two copies of the **CD3** complex, which itself consists of γ, δ, ε and ζ chains, as shown in Figure 5.4.9 (A). Part A shows the variable and C domains of the α and β chains of the T-cell receptor in green and dark. The rest of the T-cell receptor complex includes two copies each of the CD3 complex, which consist of one copy of εδζ chains and one copy of εγζ chains.

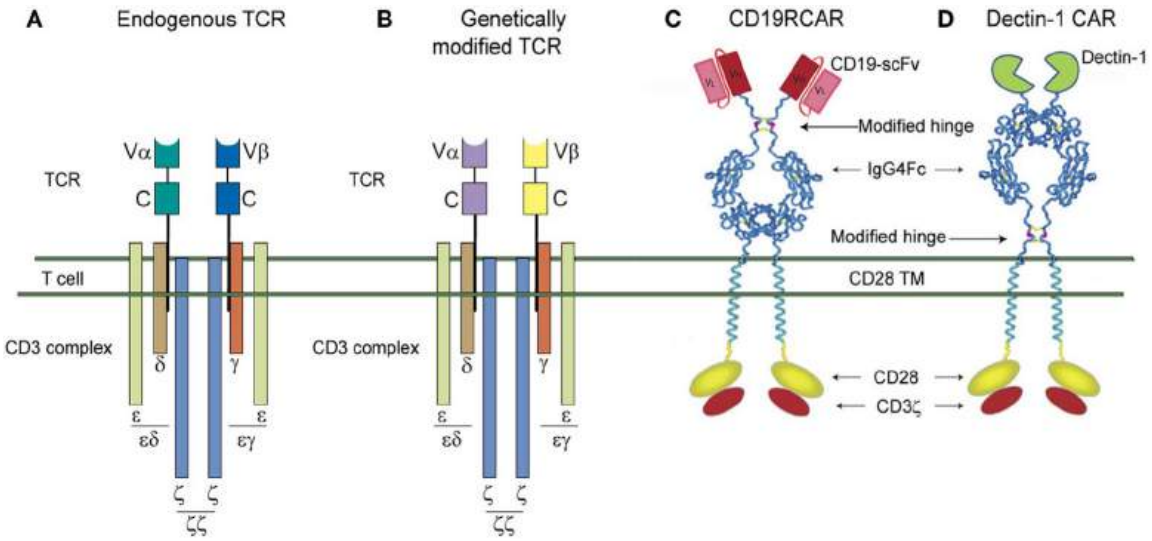


Figure 5.4.9: T cell receptor structor. Kumaresan Pappanaicken R., da Silva Thiago Aparecido, Kontoyiannis Dimitrios P. Methods of Controlling Invasive Fungal Infections Using CD8+ T Cells.
Frontiers in Immunology, 8, 1939 (2018). https://www.frontiersin.org/article/...mmu.2017.01939.
DOI=10.3389/fimmu.2017.01939. **Creative Commons Attribution License (CC BY)**.

As mentioned earlier, one of the functions of the MHC Class I molecules is to present peptides derived from tumor antigens to T-cells, which leads to the activation of other immune cells and hopefully destruction of the tumor cells displaying the tumor antigen. Much work has gone into the study of immune surveillance and ultimate destruction of tumor cells with the hopes of improving on our own immune response to cancer cells. In early work, T-cells that had infiltrated tumors were isolated from a patient, amplified in the lab by adding a **cytokine** (ex interleukin 2 - IL2), a protein growth factor released by activated immune cells, and then re-infusing the tumor-specific T cells along with IL2 back into the patients. This adoptive cell transfer (ACT) therapy led to remissions in some patients but the therapy also could be lethal.

One promising type of immune therapy is **chimeric antigen receptor (CAR) T cell therapy** or (***CAR T***) in which patients are treated with modified versions of their own T-cells. T cells are removed from a cancer patient's own blood. A gene is constructed to mimic the V and C domains of the alpha and beta chain of the T-cell receptor and inserted into the patient's T-cell using a viral vector. The gene construct contains as their tumor antigen binding motif the V and C domain of an **antibody** gene made to recognize the tumor antigen. The receptor is hence a **c**himeric (formed from parts of different proteins) **a**ntigen **r**eceptor (CAR) with antibody and T cell parts. The genetically modified cells are amplified and re-infused back into the patient. Once the collected T cells have been engineered to express the antigen-specific CAR, they are "expanded" in the laboratory into the hundreds of millions.

Compare the structure of the chimeric antigen receptor (CAR) in Figure XX-C with normal T-cell receptors shown in A. The CAR contains two, single-chain variable fragment (scFv) derived from combining the variable domains of the light (V_L) and heavy (V_H) chains of an antibody recognizing the tumor cell. These domains are connected by a linker peptide (10-25 amino acids) enriched in glycines, which confers flexibility and serines/threonines for hydrogen bonding interactions. This is attached to an F_C fragment and other intracellular effector domains to create the receptor. Figure 5.4.10 shows the scFv structure. We'll discuss the addition cytoplasmic CD28 domain in a bit.

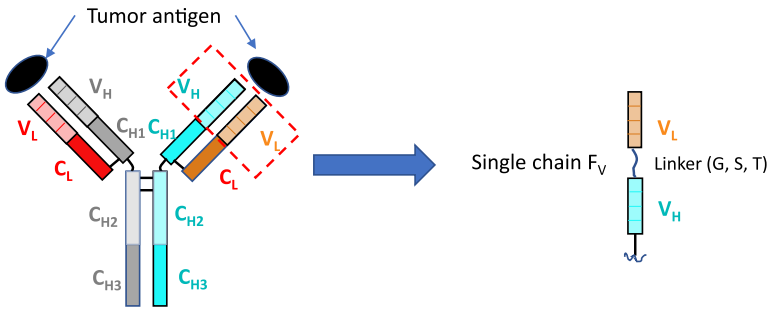


Figure 5.4.10: Single-chain variable fragment structure used in CART

You can imagine this whole T-cell receptor complex involved in the binding of a tumor peptide antigen presented by a MHC I transmembrane protein on a tumor cell, as illustrated in Figure 5.4.11 (in different colors). The entire interacting structure is called the **T cell immunological synapse**.

Figure 5.4.11: T cell immunological synapse of T cell with a cancer cell. Zhao Lijun, Cao Yu J. Engineered T Cell Therapy for Cancer in the Clinic. Frontiers in Immunology, 10, 2250 (2019) https://www.frontiersin.org/article/...mmu.2019.02250. **Creative Commons Attribution License (CC BY)**.

Protection Against Autoimmune Recognition - Coreceptors

How can the immune system recognize and bind to any foreign molecule but not self? The subject of immune tolerance is too specialized to include here, but there are a few features we will discuss.

The MHC Class I proteins do present "self" peptides in their binding pockets. Self-proteins are also cleaved into peptides in the cell by proteasomes. However, the T cell receptor does not recognize and bind to the self-peptide fragment bound to the MHC Class 1 protein. Hence T cells do not recognize self and turn against their own cells. Once and a while they do, however, and autoimmune diseases like MS, rheumatoid arthritis, and lupus result.

B cells and T cells must be activated before they can carry out their function. It is important to regulate the "on" switch. If the cells were activated without need, they might turn against self. In addition to T cell receptor complex binding to foreign peptide:MHC complexes for immune cell activation, they must also bind yet another protein on the antigen-presenting cell.

In the case of T helper cells. the T cell protein **CD28** must also bind the **B7** protein on an antigen-presenting cell like a macrophage expressing an MHC II protein:foreign peptide complex. Hence there is one specific signal (the peptide:MHC complex binding to the T cell receptor complex) and a nonspecific signal (B7 binding CD28). Why are two signals needed for activation? Again, Sompayrac has a great analogy. A safety deposit box at a bank takes two keys, a specific key (which you have) and a "nonspecific" key (which the bank uses for all boxes) to open the box. Think of it as double security. You don't want to activate immune cells for killing unless you really need to do so.

Yet other proteins are involved to ensure correct T-cell activations. We'll consider T-cells expressing either the proteins **CD4 or CD8**. T-cells expressing these expand after antigen stimulation (infection or immunization). It depends on the subtype of T-cell. Let's consider two here:

T cells expressing the protein CD4: After initial simulation, the differentiate and proliferate in T helper cells (T_H1 if they produce the cytokine interferon (IFN)-γ) and T helper type 2 (T_H2) cells (if they produce the cytokine IL4). CD4 is an integral membrane protein and acts as a co-receptor for MHC Class II:peptide complex found on cells like macrophages.

T-cells expressing CD8: These cells produce cytokines (IFN-γ and tumor necrosis factor (TNF)-α) or secrete protein which form pore-forming complex on foreign cells, leading to their lysis of cells such as pathogens or tumor cells. The CD8 protein has an alpha and beta subunit. They serve as co-receptors for MHC Class I:peptide complex found on tumor cells for example. **Cytotoxic T-cells** (a type of T-cell) express CD8.

Figure 5.4.12 shows the multiple co-signals that are required to activate the CD4 T-cell (blue sphere), which has the T-cell receptor complex, the co-receptor CD4 and the CD28 protein. It also displays a cytokine receptor, which binds cytokines released by the antigen-presenting cells (macrophage shown in pink). This leads to the proliferation and differentiation of the activated T-cell.

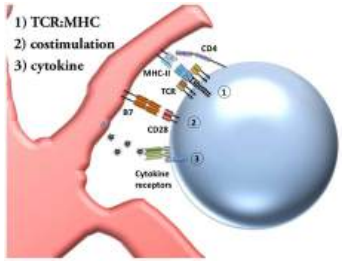


Figure 5.4.12: Co-signals that are required to activate the CD4 T-cell (blue sphere) Salmonella as a Model for Non-Cognate Th1 Cell Stimulation. Frontiers in Immunology 5(621):621 (2014) DOI: 10.3389/fimmu.2014.00621. CC BY 4.0

Sompayrac asks another interesting question. Why is antigen presentation by MHC proteins necessary at all? B cells don't really need presentation since they can bind antigen with membrane antibody molecules. Why do T cells need it? He gives different reasons for Class I and Class II presentations:

Class I MHC (found on most body cells): T cells need to be able to "see" what is going on inside the cell. When virally-infected cells bind foreign peptide fragments and present them on the surface, they can be "seen" by the appropriate T cell. It's a way to get a part of the virus, for example, to the surface. They can't hide out in the cell. T cells don't need to recognize extracellular threats since antibodies from B cells can do that. Presentation is also important since viral protein fragments found outside of the cell might bind to the outer surface of a noninfected cell, targeting them for killing by the immune system. That wouldn't be good. It also helps that peptide fragments are presented on the surface. This allows parts of the protein that are buried and not exposed on the surface, which would be hidden from interaction with outside antibodies, to be used in signaling infection of the cell by a virus.

MHC Class II (found on antigen-presenting cells like macrophages): Two different cells (the presenting cell and the T helper cell) must interact for a signal for immune system activation to be delivered to the body. Again it is a safety mechanism to prevent the nonspecific activation of immune cells. Also, as in the case above, since fragments are presented, more of the foreign "protein" can contribute to the signal to activate the immune system.

5.4.4: Recognition and Response in the Innate Immune System

The B and T cell part of the immune system represents the more sophisticated branch of the immune system called the **adaptive immune system**. It can be trained to recognize any foreign chemical/cellular species. The other branch of the immune system is the **innate immune system**. The system recognizes common molecular structures found all many different organisms, so in this branch, there is no need to adapt to each foreign species individually. The adaptive immune response also must be activated by cells of the innate immune system.

The innate immune system recognizes common structural features in viruses and living cells like bacteria, fungi, and protozoans like amoebas. The cells of the innate system (dendritic cells, macrophages, eosinophils, etc, which we talked about as antigen-presenting calls above) have receptors called **Toll-like Receptors 1-10 (TLRs)** that recognize the common **pathoge- associated molecular patterns (PAMPs)** , which leads to binding, engulfment, signal transduction, maturation (differentiation), antigen presentation, and cytokine/chemokine release from these cells. Dendritic cells, which reside in the peripheral tissues and act as sentinels, are an example. They can bind PAMPs which include:

- CHO/Lipids on bacteria surface (LPS)
- mannose (CHO found in abundance on bacteria,
- yeast dsRNA (from viruses)
- nonmethylated CpG motifs in bacterial DNA

After entering an immune cell, bacterial and viral nucleic acids are recognized by **intracellular TLRs.** Dendritic cells phagocytize microbial and host cells killed through programmed cell death (apoptosis). During maturation, surface protein expression is altered, allowing the cells to leave the peripheral tissue and migrate to the lymph nodes where they activate T cells through the antigen presentation methods described above. They also control lymphocyte movement through the release of chemokines. Figure 5.4.13 shows the TLR family, their binding signals, and intracellular adapter proteins used to transmit signals into the cell.

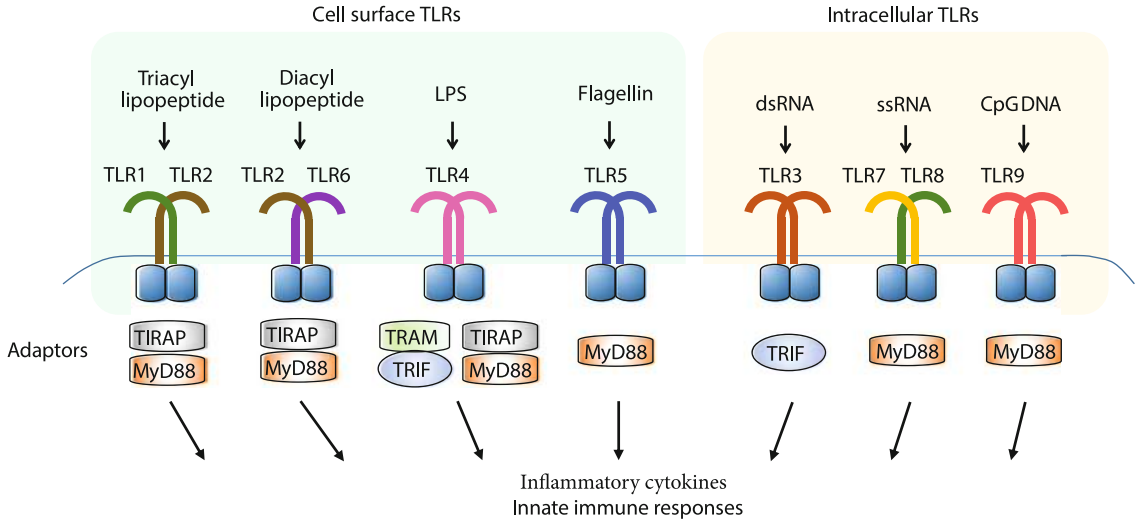


Figure 5.4.13: TLR family, their binding signals, and intracellular adapter proteins Ji-Yoon Noh, Suk Ran Yoon, Tae-Don Kim, Inpyo Choi, Haiyoung Jung, "Toll-Like Receptors in Natural Killer Cells and Their Application for Immunotherapy", *Journal of Immunology Research*, vol. 2020, Article ID 2045860, 9 pages, 2020. https://doi.org/10.1155/2020/2045860. This is an open access article distributed under the Creative Commons Attribution License

Figure 5.4.14 shows an interactive iCn3D model of the mouse Toll-like receptor 3 ectodomain (that sticks out into the cytoplasmic space from an internal organelle) complexed with double-stranded RNA (3CIY).

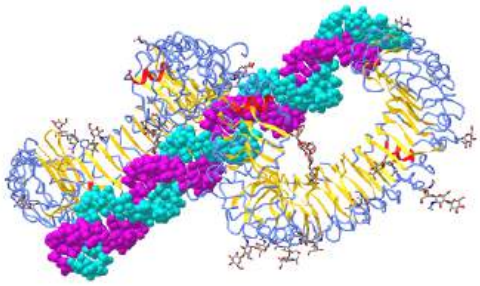


Figure 5.4.14: Mouse Toll-like receptor 3 ectodomain complexed with double-stranded RNA (3CIY) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?otC8JXNQe45xBFx47

Double-stranded RNA is found in the life cycle of many viruses so it makes great sense for evolution to create a binding protein to recognize this common structure (PAMP). The TLR3 ectodomains (ECDs) form dimers when the dsRNA is at least 40-50 nucleotides long. The dsRNA is shown in spacefill (cyan and magenta). Note the extensive glycosylation (colored cubes) in the structure. The protein looks like a "horseshoe-shaped solenoid " with lots of beta structure. It has 23 leucine-rich repeats (LRRs) with some conserved asparagines allowing for extensive hydrogen-bonding. One face appears to be free of carbohydrate residues and may be important in dimerization and function.

TLRS and mRNA vaccines

Messenger RNA vaccines against the SARS-Cov2 spike protein have probably saved up to 20 million lives in the first year of the COVID-19 pandemic. The development of mRNA vaccines will go down as a great scientific achievement that required decades of fundamental and applied research by many scientists.

Vaccines usually are composed of target proteins from a virus, for example. Instead of delivering an actual protein, whose actual development and mass production takes years, why not use mRNA that encodes a viral protein or fragments of it? The idea has been around for a long time. The problem is that RNAs, with their 2'-OH on the ribose ring, are very labile and degrade easily. In addition, injecting RNA into a patient causes a significant immune response to the RNA. We mount immune responses to foreign viral RNA through our TLR receptors (TL3, 7 and 8) but what is needed is an immune response to the protein made from the injected mRNA, not to the RNA. Yet we don't make an immune response against our own RNA. Why?

There are two major problems that had to be solved (and a host of others as well) to make mRNA vaccines, the problem of stability and our immune response against them. A hint comes from the observation that TLRs recognize non- or undermethylated DNA found in bacteria. Methylated CpG motifs in DNA to do not stimulate an immune response. Katalin Karikó,Michael Buckstein, Houping Ni and Drew Weissman reported in 2005 that the incorporation of methylated (m) and modified nucleosides m5C, m6A, m5U, s2U, and pseudouridine ablated the immune response to the RNA. This opened the door to mRNA vaccines. The paper was rejected by Nature and Science but published in Immunity (Vol. 23, 165–175, August, 2005. DOI 10.1016/j.immuni.2005.06.008). The last line from the paper was truly prophetic: "Insights gained from this study could advance our understanding of autoimmune diseases where nucleic acids play a prominent role in the pathogenesis, determine a role for nucleoside modifications in viral RNA, and give future directions into the design of therapeutic RNAs".

Katalin Karikó and Drew Weissman were awarded the 2021 Lasker–DeBakey Clinical Medical Research Award (often a prelude to the Nobel Prize) for their fundamental research that has saved so many of us. They were awarded the Nobel Prize in Medicine in 2023. See Chapter 9.1 for more details about the role of pseudouridine in mRNA vaccines.

5.4.5: Inflammasome

Think of the things you would want your immune system to protect you from. Of course, there are the pathogens like viruses, bacteria, and fungi. And of course, you want to be protected from yourself in that you don't want to activate your immune system with self-antigens. But what about "non-biological" molecules like silica or asbestos whose presence might be deleterious? What about normal biomolecules (proteins, nucleic acids) that suddenly find themselves in the wrong cellular location due to cell death by necrosis or physical injury?

In the previous section, we discussed how innate system immune cells (dendritic cells, macrophages, eosinophils, etc) have receptors that recognize common **pathogen-associated molecular patterns (PAMPs)** such as lipopolysaccharides (LPS) on the surface of bacteria, mannose on bacteria and yeast, flagellin from bacterial flagella, dsRNA (from viruses) and nonmethylated CpG motifs in bacterial DNA. These antigens are recognized by pattern recognition receptors (PRRs) - specifically the **Toll-like Receptors (TLRs)** 1-10. These include plasma membrane TLRs (TL4 for LPS, TL5 for flagellin, TLR 1, 2 and 6 for membrane and wall components of fungi and bacteria) and intracellular endosomal TLRs (TLR3 for dsRNA, TLR 7 and 8 for ssRNA and TLR9 for dsDNA)

Damage-associated molecular patterns (DAMPs) are typically found on molecules released from the cell or intracellular compartments on cellular damage (hence the name DAMP). Many are nuclear or cytoplasmic proteins released from the cells. These would now find themselves in a more oxidizing environment which would further change their properties. Common DAMP proteins include heat shock proteins, histones and high mobility group proteins (both nuclear), and cytoskeletal proteins. Think what non-protein molecules might be released from damaged cells that might pose problems. Here are some other common non-protein DAMPS: ATP, uric acid, heparin sulfate, DNA and cholesterol crystals. In the wrong location, these can be considered danger signals.

If TLRs recognize PAMPs, what recognizes DAMPs? They are recognized by another type of intracellular pattern recognition receptor (PRR) called **NOD** (Nucleotide binding **O**ligomerization **D**omain (**NOD**)- **L**ike **R**eceptors or **NLRs**. NLRs **also** recognize PAMPs. The proteins also are named as the **N**ucleotide-binding domain (NBD) and **L**eucine-**R**ich repeat (LRR)–containing proteins (**NLR**)s. This family of proteins participates in the formation of a large protein structure called the **inflammasome**. (Sorry about the multiple abbreviations and naming systems!)

As both PAMPs and DAMPs pose dangers, it would make sense that once they recognize their cognate PRRs (TLRs and NLRs, respectively), pathways leading from the occupied receptors might converge in a common effector system for the release of inflammatory cytokines from immune cells. Given that uncontrolled immune effector release from cells in an inflammatory response might be dangerous, it would be sometimes helpful to require two signals to trigger cytokine release from the cell. We've seen this two-signal requirement for the activation of T cells.

Two such inflammatory cytokines are Interleukin 1-beta (IL 1-b) and IL-18. Activation of TLRs by a PAMP leads to activation of a potent immune cell transcription factor, NF-kbeta, which leads to transcription of the gene for the precursor of the cytokine, pro-interleukin 1-beta. Without a specific proteolytic cleavage, the active cytokine will not be released from the cell.

The protease required for this cleavage is activated by a signal arising when a DAMP activates a NLR, which then through a sequence of interactions leads to the proteolytic activation of another inactive protease, **procaspase 1**, on a large a multi-protein complex called the **inflammasome**. (In later chapters we will see other such protein complexes with targeted activities - including the spliceosome, which splices RNA to produce mRNA and the proteasome which conducts controlled intracellular proteolysis). The activated inflammasome activates procaspase to produce the active protein **caspase** (a cysteine-aspartic protease).

The convergence of the signals from the PAMP activation of a TLR and DAMP activation of a NLD at the inflammasome is shown in Figure 5.4.15

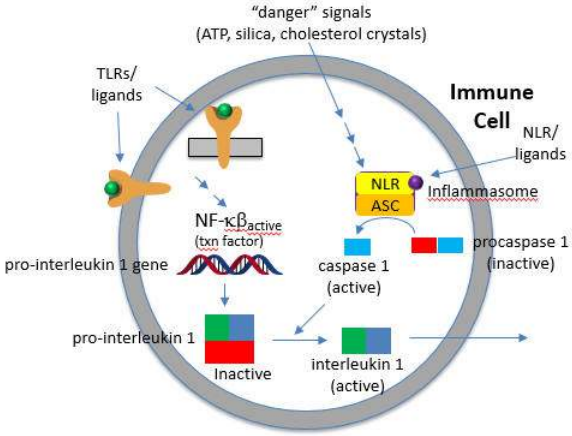


Figure 5.4.15: Signals from PAMP activation of a TLR and DAMP activation of a NLD at the inflammasome

The active cytokine interleukin 1-beta helps recruit innate immune cells to the site of infection. It also affects the activity of immune cells in the adaptive immune response (T and B cels). Active IL-18 leads to the increase of another cytokine, interferon-gamma and it also increases the activity of T cells that kill other cells.

The focus of this chapter is on binding interaction and their biological consequences. From that perspective, this section will address

- the structure and activity of caspases, which activate the pro-cytokine prointerleukin 1 beta,
- the structure and ligands for the NLRs, the structure and properties of the inflammasome, and finally
- how "danger" molecules such as ATP and crystals (cholesterol, silica) activate the inflammasome.

Unfortunately, there are many proteins involved with crazy acronyms for names. These proteins have multiple domains and many of the proteins often have multiple names. Sorry in advance!

5.4.5.1: Caspases

Caspases (Cys-asp-proteases), not to be confused with Cas9 (CRISPR associated protein 9, an RNA-guided DNA endonuclease) is a protease which when active can lead to cell death, or in a less austere fashion initiate the inflammatory response (sometimes good, often bad or even fatal). They have an active site nucleophilic Cys and cleave peptide bonds after an Asp in target proteins. All caspases (13 in humans) have an N-terminal pro-domain followed by large and small protease catalytic domain subunits. As with other proteases, it is found as an inactive zymogen. Why is this important?

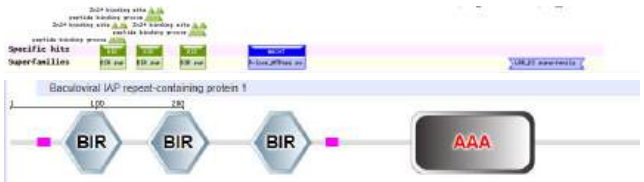
To become activated they are recruited to a scaffolding protein where they are activated by removal of the N-terminal domain of the zymogen and then a second cut between the large and small catalytic subunits. The enzyme that does this is caspase itself in an autocatalytic step. . There are 3 kinds of caspases, two of which are involved in programmed cell death. We'll discuss the inflammatory cytokine processing Caspase-1. Once activated, the initiators activate other effector (executioner) caspases in the cell) . Caspase 1 is activated by the inflammasome.

Two major domains are found in Caspase 1, the **caspase recruitment domain (CARD)** which mediate self-interaction with scaffold and adaptor proteins in the inflammasome for activation, and a proteolytic catalytic domain, as shown in Figure 5.4.16. All domain structures in the section were obtained using Conserved Domains from the NCBI (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) or the Simple Modular Architecture Research Tool (SMART) at the EMBL (http://smart.embl-heidelberg.de/smart/set_mode.cgi?NORMAL=1). Uniprot was used for protein (FASTA) sequences (http://www.uniprot.org/uniprot/). We will see the CARD domain often.

Figure 5.4.16: Domains structure of Caspase 1

5.4.5.2: NOD-like receptor proteins (NLRPs)

The **NOD-like receptor proteins** (NLRPs) are a family of proteins with similar domain structures. The structures and abbreviations used for the molecular players in inflammasome activation are very complicated and confusing. Different programs show different domains, which adds to the complexity. We will attempt to reduce the confusion by just showing domain structure diagrams, even if some show different domains for the same protein. Remember domains are calculated from structure so different algorithms using different databases might return different domain structures. Table 5.4.2 belows shows the domain structure for NLRPs.

NLRP1,2,3	
NLRP1 (NALP1)	
NLRP3 (NALP3)	
NLRP4	
NAIP1	

NAIP2	

Table 5.4.2: Domain structures of the NLRPs

Another protein in the NLR family is NAIP (neuronal apoptosis inhibitor protein). The domain structure for NAIP1 is shown in Table 5.4.2 above. In contrast to the other NLR for which specific ligands have not yet been found, several NAIPs have been shown to bind specific PAMPs. NAIP1 binds the needle protein CprI from *C.violaceum* which starts to drive the assembly of the NLRC4 inflammasome. NAIP2 binds the inner rod protein of the bacterial type III secretion system (which for *Salmonella typhimurium* is the protein PrgJ). NAIP5 and NAIP6 bind bacterial flagellin (which for *Salmonella typhimurium* is the protein FliC). AAA in the second domain representation is **A**TP-**a**ssociated **a**ctivities in the cell (otherwise denoted as the NACHT domain in the top representation).

NAIP2 interacts with another adapter NLR family protein, NLRC4 (NLR family CARD domain-containing protein), to form the inflammasome. The domain structure of NAIP2 is shown in Figure 5.4.x below:

Note that many of these proteins share common domains:

- Pyrin-NALP - Pyrin domains on different proteins self-associate through inter-protein Pyrin:Pryin interactions
- NACHT - This domain contains about 300-400 amino acids and can bind ATP and may cleave it (i.e. act as an ATPase)
- LRR - for Leucine Rich Repeat. These 20-30 amino acid repeats may occur up to 45 times in a given protein. They fold into an arc shape and seem to facilitate protein:protein interactions. On the concave side of the arc they have a parallel beta sheet while on the convex side they have an alpha helix. They also appear to be involved in the binding of PAMPS and DAMPs;
- CARD - for caspase activation and recruitment. CARD domains on different proteins self-associate through inter-protein CARD:CARD interactions;
- BIR - Baculoviral inhibition of apoptosis protein repeat;
- ASC - Apoptosis-associated speck-like protein containing a CARD Adapter domain, allowing it to interact with other proteins with a CARD domain.

5.4.5.3: **ASC Adaptor Protein**

Small adapter proteins like ASC with a CARD domain mediate binding of caspases in the apoptosome (involved in apoptosis or programmed cell death) and in the inflammasome. This smaller protein has two domains, a pyrin domain and a CARD domain as shown in Figure 5.4.17. It is required for the recruitment of caspase-1 to some inflammasomes (for example, ones that contain NLRP2 and NLRP3

Figure 5.4.17: Domain structure of the small adapter proteins ASC

5.4.6: **The Active Inflammasome**

The active inflammasome, in general, consists of three different kinds of proteins, some present in multiple copies: NLRPs, adapter proteins like ASC, and procaspases. They may also contain additional recruitment and ligand sensor proteins. We'll discuss two types using different NLRPs, the NLRP4 and NLRP3 inflammasome.

5.4.6.1: NLRP4 Inflammasome

Some of the best structures (obtained by cryomicroscopy) are for the NAIP2:NLRP4 inflammasome. Figure 5.4.18 shows part of the complex consisting of 11 NLRP4 subunits arranged in a large ring. The actual biological complex has 1 NAIP2 subunit and 10 NLRP4s.

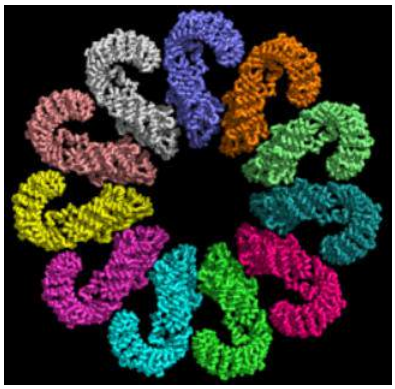


Figure 5.4.18: Part of the complex of the NAIP2:NLRP4 inflammasome showing 11 NLRP4 subunits arranged in a large ring

Figure 5.4.19 shows an interactive iCn3D model of the activated NAIP2-NLRC4 inflammasome (3JBL))

Figure 5.4.19: Activated NAIP2-NLRC4 inflammasome (3JBL). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...8Ae7vY4aX1SLX8

How does this structure arise? Presumably, it would not exist in the absence of a PAMP or DAMP in order to minimize immune-mediated inflammatory damage. Data suggests that the bacterial protein PrgJ (denoted PrgX in the figure below) binds to its receptor, NAIP2, altering its conformation as shown in Figure 5.4.20. This binary complex presents an asymmetric electrostatic surface which allows a loose association with NLRP4, which leads, after a conformational change, to a tighter binding interaction. Nine more NLRP4s bind in a similar fashion to form the 11-subunit ring structure.

Figure 5.4.20: Formation of the NAIP2:NLRP4 ring

The complementary electrostatic interactions between two of the many NLRC4 subunit monomers in the NLRP4 inflammasome are depicted in Figure 5.4.21.

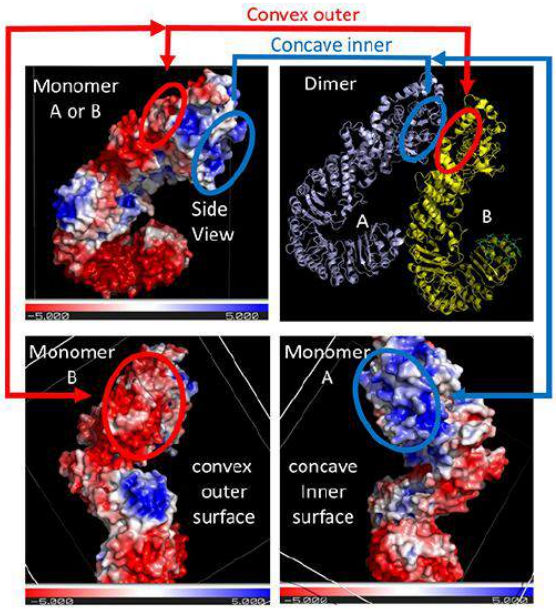


Figure 5.4.21: Complementary electrostatic interactions between two of the many NLRC4 subunit monomers in the NLRP4 inflammasome

The top right panel shows two of the NLRC 4monomers (of the 12 in the Jsmol model) bound to each other, with the concave inner face of the A subunit interacting with the convex outer face of the B subunit. For simplicity, only the interactions at the top of the dimers are highlighted. The other panels show the electrostatic potential surface (red indicating negative and blue positive) for each monomer on a sliding scale of -5 to +5 (images created using the PDB2PQR Server and Pymol).

The depicted negative (red) electrostatic potential outer surface on one NLRC4 monomer that is complementary to the positive (inner) surface on the other NLRC4 subunit are outlined in each panel in red or blue ellipses, respectively. The curved tertiary structure of the proteins and the opposing electrostatic surface potentials of opposite faces commit the subunits to form a large ringed 12-mer core of the nucleosome.

Note that this assembled ring brings together the CARD (caspase recruitment domain, yellow circles/spheres) which can interact with the CARD domain of the procaspase protein through CARD:CARD inter-protein interactions (think of a stack of playing cards all stuck together in a deck of cards) as shown in Figure 5.4.22

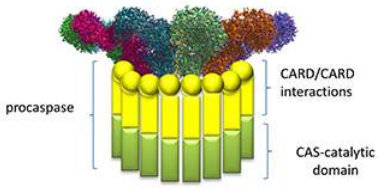


Figure 5.4.22: Interactions of procaspase with the NAIP2:NLRP4 ring

Once assembled, proximal procaspases autocatalytically convert procaspase 1 into active caspase 1, which can activate, by proteolysis, the cytokines interleukin 1 beta and interleukin 18 to form active cytokines which are released from the cell. Remember that the procytokines are present only if their genes have been transcribed following activation of the transcription factor NF-kappa beta through PAMP binding to a TLR.

5.4.6.2: NLRP3 Inflammasomes

Figure 5.4.23shows an interactive iCn3D model of the NLRP3 double-ring cage, 6-fold (12-mer) (7LFH)

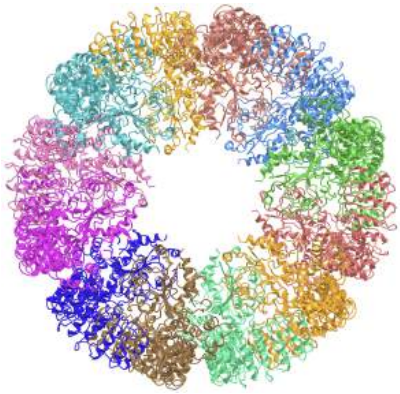


Figure 5.4.23: NLRP3 double-ring cage, 6-fold (12-mer) (7LFH). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?DEbdkUoBtqRQ9bu59

The full-length mouse NLRP3 consists of 12- to 16-mer organized in a double-ring cage. It is held together by interactions between the leucine-rich repeats (LRR) domains. The pyrin domains are shielded by the structure, so they will not be activated without appropriate signals. The complex is also localized to the membrane.

In contrast to NLRP4 inflammasomes, which require specific PAMPs/DAMPs for activation, NLRP3 inflammasomes seem to be activated by cellular distress as well as cell exposure to pathogens. It is one of the main responders to a variety of microbial infections. Given the large number of microbes that lead to NLRP3 inflammasome activation, it has been suggested that the actual signal that triggers NLRP3 is indirect. One such indirect signal is K^+ ion levels in cells.

In normal cells, K^+ ions are higher in the cytoplasm than in the outside of the cell. Potassium ion decreases in cells caused by efflux can activate NLRP3 inflammasomes. Other conditions include the rupture of lysosomes (perhaps associated with the cellular uptake of particles like silica, uric acid, cholesterol crystals, and other "nanoparticles"), altered mitochondrial metabolism (which can lead to reactive oxygen species within the cell), etc. Obviously, all of these danger triggers don't bind to NLPR3 but somehow lead to downstream activation of it. NLRP3 hence probably works by being a general sensor for cell stress.

Inappropriate and chronic activation of inflammation has been associated with many disease such as cancer, cardiovascular disease, diabetes and autoimmune diseases. Given the multiple types of signals that can activate the NLRP3 inflammasome, this complex is the focus for active drug development to find inhibitors that would stop undesired inflammation. These inflammasomes are found in granulocytes, monocytes (macrophages), megakaryocytes, and dendritic cells.

Activated NLRP3 recruits the ASC Adaptor Protein, which leads to recruitment and activation of procaspase 1. NLRP3 has a pyrin, NACHT, and LRR domain. ASC has a pyrin and CARD domain. Active LRP3 can then recruit ASC through pyrin:pyrin inter-protein domain interactions. This then allows the CARD domain of bound ASC to recruit procaspase through CARD:CARD interactions (remember that procaspase has a CARD domain as well), forming the active NLRP3 inflammasome. An added feature of NLRP3 inflammasome activation occurs when the transcription factor NFkb, which is activated by PAMPs (signal 1), leads to the transcription of both the procytokines (IL-1 beta and IL 18) and of NLRP3 itself.

Hence two signals are again needed:

Signal 1

The first signals are the bacterial and viral (influenza virus, poliovirus, enterovirus, rhinovirus, human respiratory syncytial virus, etc) PAMPs, which bind to TLRs and lead to the activation of the NFkb transcription factor. This activates not only the transcription of pro-interleukin 1-beat and interleukin 18, but also to the transcription of the NLRP3 sensor itself.

Signal 2

Signal 2 is delivered by PAMPs and DAMPs indirectly to the sensor NLRP3. This leads to the assembly of the inflammasome. These DAMPs appear to prime the activation of NLRP3 protein and subsequent formation of the active NLRP3 inflammasome. But what activates NLR3P3? After many studies, it became clear that the typical bacterial ligands that would activate TLRs and perhaps NLRs only prime NLRP3 for activation. They don't bind to it directly.

Extracellular ATP is a major activator of NLRP3. Nanoparticles are known to release ATP as well. Most studies show that K^+ efflux from the cell is an early signal and that the NEK7, a protein that phosphorylates other proteins, binds to NLRP3 after potassium ion efflux and activates it. Removing NEK7 stopped NLRP3 but not NLRP4 inflammasome activation. Although NLRP3 bound to NEK7 through the NEK7 catalytic domain, the activity of the catalytic domain of NEK7 was not needed.

What leads to K^+ efflux? Let's back up to possiblde upstream events that could lead to efflux and try to find a link to ATP. The background for some of this material will be explored in future chapters. The following steps occur as shown in the figure and information below:

- solids such as silica, cholesterol crystals, uric acid crystals, and even aggregated proteins such as prions can be engulfed by monocytes/macrophages (much as they engulf bacteria as part of their immune function) in a process called phagocytosis. The particles are enveloped in plasma bilayer-derived membrane which buds off into the cell. This vesicle merges with a lysosome which gets damaged in the process. They then release ATP into the cytoplasm;

- cytoplasmic ATP can then move outside of the cell through the glycoprotein membrane channel called pannexin 1;

- extracellular ATP can bind to another membrane protein called the P2X7 purinoceptor. This protein now becomes a cation channel which allows K^+ efflux since the ion has a higher concentration inside the cell than outside, as illustrated in Figure 5.4.24. The extracellular ATP "gates" open the P2X7 cation channel. The pore-forming toxin nigericin from *Streptomyces hygroscopicus* also leads to potassium ion efflux. Likewise, pore-forming proteins from *S. aureus* (hemolysins) lead to potassium ion efflux and activation of the NLRP3 inflammasome. We will discuss membrane protein in great detail in a latter chapter.

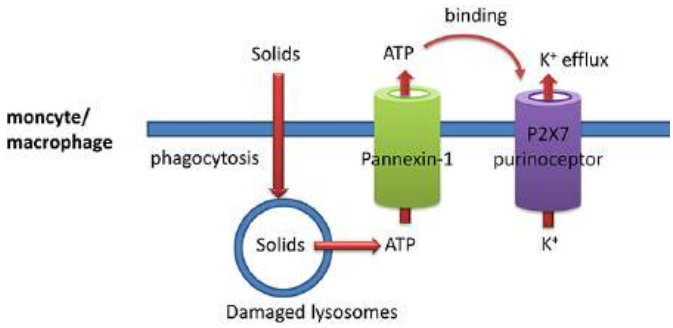


Figure 5.4.24:Extracellular ATP triggers K^+ efflux from cell

Other signals also activate the NLRP3 inflammasome. These include mitochondrial damage and the release of reactive oxygen species.

5.5: Protein Interactions Modulated by Chemical Energy- Actin, Myosin, and Molecular Motors

9/21/21: The material below is derived from (shortened and summarized): Overview of the mechanism of cytoskeletal motors based on structure. Kato et al. Biophys Rev. 2018 Apr; 10(2): 571–581. Published online 2017 Dec 12. doi: 10.1007/s12551-017-0368-1. Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0).

5.5.1: Muscle Contraction

The biochemistry of muscle structure and contraction involves many proteins, including myosin heavy and light chains, actin, three types of troponins, tropomyosin, and other regulatory proteins. These form complexes which engage in cyclic conformation changes leading to muscle contraction in a process that requires chemical energy in the form of ATP hydrolysis. In the process, chemical energy is converted to mechanical energy. Those who view biochemistry through a more chemical lens might find the biology somewhat confusing. We'll provide a short introduction before looking at the proteins in more detail.

First let's start with a simplified version of the structure of myosin and actin to the basic repeating structural unit of the muscle, the **sarcomere**, which is shown in panel A of Figure 5.5.1. Myosin forms a function dimer of two intertwined heavy chains (**blue**), each associated with two different light chains. These are named the essential light chain (**red**) and the regulatory light chain (**yellow**). Overall the functional dimer of myosin is a hetero 6-mer.

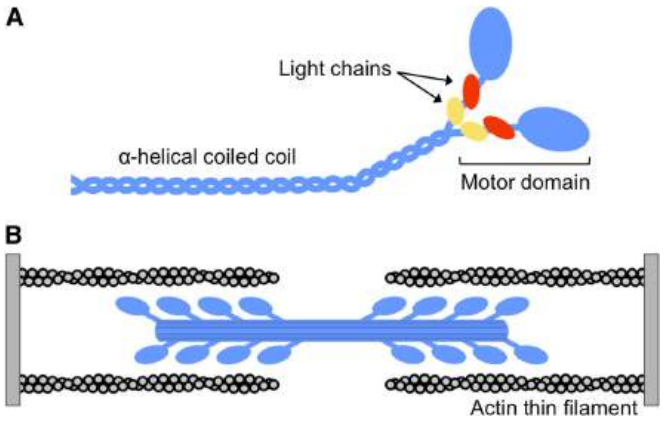


Figure 5.5.1: Myosin and Actin form the basic components of the sarcomere. Lee et al. Skeletal Muscle (2019) 9:7 https://doi.org/10.1186/s13395-019-0192-3. Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/).

The heavy chain of myosin consists of two major domains and other regions as shown in panel A of Figure 5.5.1. It contains a

- motor (head) domain (blue ellipse), which binds and cleaves ATP (i.e. it is an ATPase) and which also binds actin;
- an alpha-helical tail domain that binds the tail domain of another myosin heavy chain to form the functional a coiled-coil structure in the hetero 6-mer.
- an α-helical lever arm, which binds the two light chains.

In the muscle sarcomere, many hetero 6-mers aggregate through their coiled-coil tails to form the large blue structure shown in panel B of Figure 5.5.1. The protruding heads that radiate away from the filament axis interact with actin filaments. This is shown in panel B of Figure 5.5.1. The actin filaments are shown as two intertwined actin strands (small gray circles).

Note also that in panel B, the myosin thick filaments appear to be bipolar as their motor head domain are pointed in opposite direction (antiparallel) along the center of the thick filament that has no heads in the center. The central region without motor domains has aggregated hetero 6-mers pointed in opposite directions. However, the rest are packed in a parallel fashion and extend the rest of the way.

The head domain of the myosin heavy chain can be removed, along with the associated light chains, from the rest of the using selective proteolysis to form the S1 (head), or S1 and part of the myosin light chain (called heavy meromyosin -HMM). This has enabled studies of these molecules under simpler conditions. The myosin heavy chain fragments are shown in Figure 5.5.2.

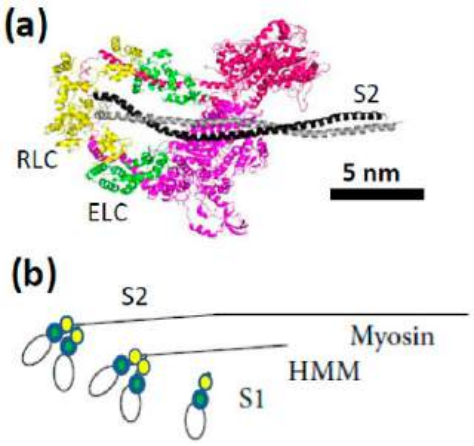


Figure 5.5.2: Proteolytic fragment of the myosin hetero 6-meer. Månsson et al. Int. J. Mol. Sci. 2018, 19, 1863; doi:10.3390/ijms19071863. Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

Going in the direction from the rod domain to the motor domain, the myosin heavy chain has a long alpha-helical lever arm which binds the two myosin light chains (the regulatory light chain (yellow above) and the essential light chain (green above). This is followed by the converter domain and finally the motor domain, which binds ATP.

The actin thin filament consists of two intertwined antiparallel strands of many G-actin (globular) monomers (1atn, molecular weight of 42K), which self-assemble in a process requiring ATP hydrolysis, to form the F (filamentous) actin structure. The structure of G-actin and the overall arrangement of the asymmetric monomers in the F filament are shown in Figure 5.5.3.

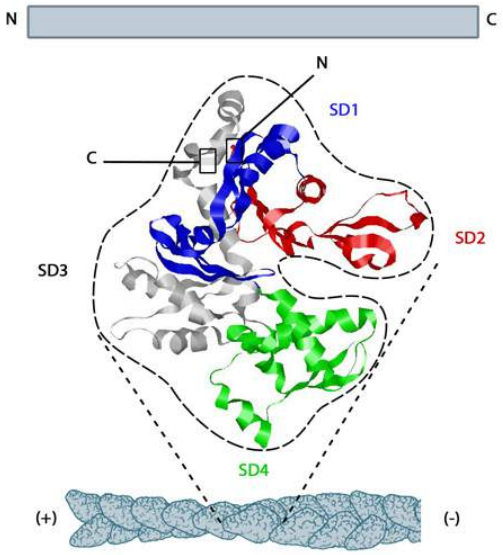


Figure 5.5.3: The structure of monomeric actin in the F-actin thin filament. https://www.mechanobio.info/cytoskel...ilaments-grow/. Creative Commons Attribution-NonCommercial 4.0 International License.

It would be nice if the structure of the sarcomere were as simple as illustrated in Figure 5.5.1. Many more proteins are involved which regulate its contraction. Let's add some more proteins (troponin subunits T, C and I as well as tropomyosin) to get a deeper understanding of the structure as shown in the cartoon representation shown in Figure 5.5.4.

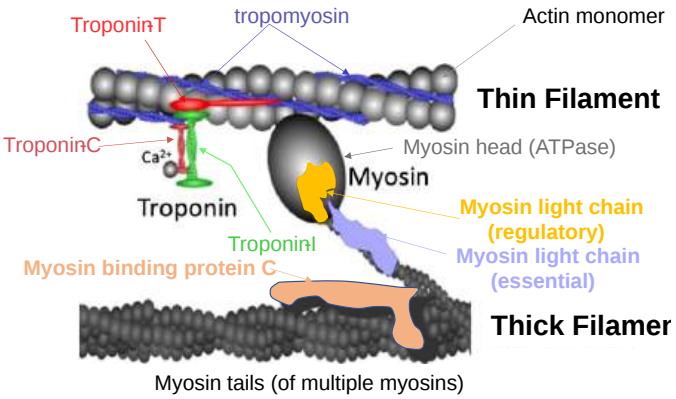


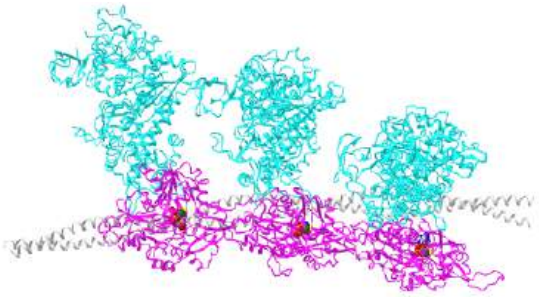
Figure 5.5.4: Modified from de Tombe PP et al. Global Cardiology Science and Practice 2016:21 http://dx.doi.org/10.21542/gcsp.2016.21. Creative Commons Attribution 4.0 International License.

The trimeric troponin complex consists of

- troponin-C which binds Ca^{2+} ions
- troponin-I which inhibits contraction
- troponsin-T which attaches the trimeric troponin to the thin actin filament.

Ca^{2+} ions released from stores in the sarcoplasmic reticulum (the muscle equivalent of the endoplasmic reticulum) on neural stimulation of muscle cells by the neurotransmitter acetylcholine initiates sarcomere contraction. The Ca^{2+} ions binds to troponin-C (dark red) which leads to conformational changes that rotates tropomyosin, which occupies the groove between the two actin chains in the thin filament. This exposes a binding site on actin for the motor domain heads on the myosin thick filament.

An interactive iCn3D model of the Actin-Myosin-Tropomyosin ADP complex (6X5Z) is shown in Figure 5.5.5. - actin magenta, myosin cyan, tropomyosin gray, ADP and Mg spacefill



NCBI iCn3D Figure 5.5.5: Actin-Myosin-Tropomyosin ADP complex (6X5Z). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...Q8KSdXLhjtNzt6

Now let's show an even larger cartoon view, which shows what is observed under a microscope. Figure 5.5.6 shows a cartoon representation of a sarcomere.

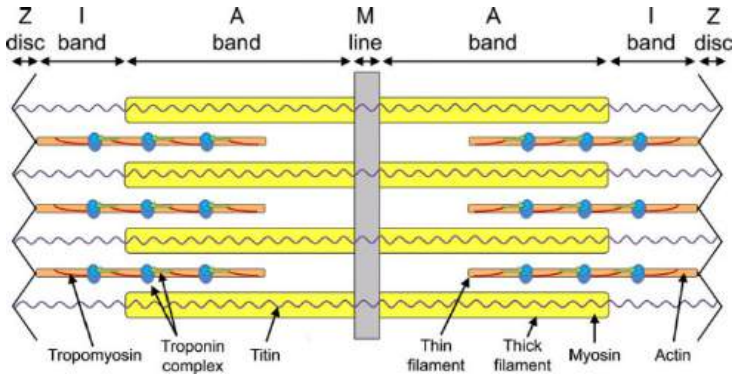


Figure 5.5.6: England and Loughna. Cell. Mol. Life Sci. (2013) 70:1221–1239. DOI 10.1007/s00018-012-1131-1Creative Commons Attribution License

- the myosin thick filament is shown as a yellow rectangle;
- the actin thin filament is shown as a orange rectangle with tropomysin shown as a **red** line and the trimeric troponin complex shown in blue

The M line provides anchors for the thick filament while the Z disc attaches to and anchors the thin filaments. The myosin filaments are crosslinked and anchored at the end M band, which contains a series of proteins including myomesin, creatine kinase and M-protein. The Z disc/line contains many proteins including alpha-actinin (predominant), actin, many other proteins, and the end of titin. This cartoon shows the features shown in an electron micrograph of the vertebrate striated muscle sarcomere, which extends between the two Z-lines (or disks), as shown in Figure 5.5.7. A single bidirectional myosin thick filament is shown as an orangish rectangle, while single thin actin filaments are shown in blue (attached to the Z disc).

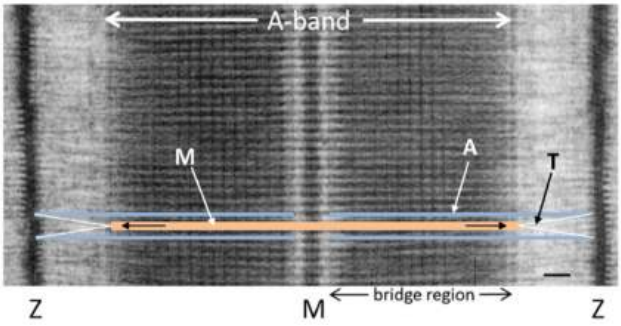


Figure 5.5.7: Electron micrograph of the sarcomere. John Squire. Int. J. Mol. Sci. 2019, 20, 5715; doi:10.3390/ijms20225715. Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

Again note that the rod domains of the thick filaments are packed in an antiparallel fashion (ie. pointing in opposite directions) in the middle of the sarcomere near the end M line band, so in this region that are no myosin heads. However, the rest are packed in a parallel fashion and extend through the rest of the A-band. These myosin heavy chains are packed in a somewhat (quasi) symmetrical fashion, which repeats. Looking down the myosin fiber, the motor domains radiate outward with a 3-fold rotational symmetry and are staggered at 120^0.

Now let's visualize the contraction of the sarcomere since it's hard to display in a single figure. Imagine it contracting somewhat like an accordion. The M line, devoid of myosin head groups and where the myosin heavy chains extend in different directions toward the left and right Z-lines, remains stationary. Since the myosin chains in the thick filament are attached to it, they remain stationary as well. Imagine the left and right Z line (disk), which anchor the actin thin filaments, moving towards the center line as the actin thin filaments slide along the stationary myosin thick filaments. This contracts the sarcomere. The sliding is powered, of course, by ATP hydrolysis in the motor heads of myosin in a series of steps involving attachment, detachment, and reattachment of the flexible myosin head domain further along with the actin thin filament.

Especially note the elongated protein called titin shown in cartoon form in Figure 5.5.6. This links the ends of the myosin filament and the Z-line. Titin is the largest single polypeptide known. Its molecular weight is close to 4 million with over 34,000 amino acids. It is highly elastic. It helps balance the forces on each side of the sarcomere. Its repetitive structure is shown in Figure 5.5.8.

A series of structured Ig domains, comprised mostly of beta-strands are separated by flexible hinge structures which allow rearrangements and stretching.

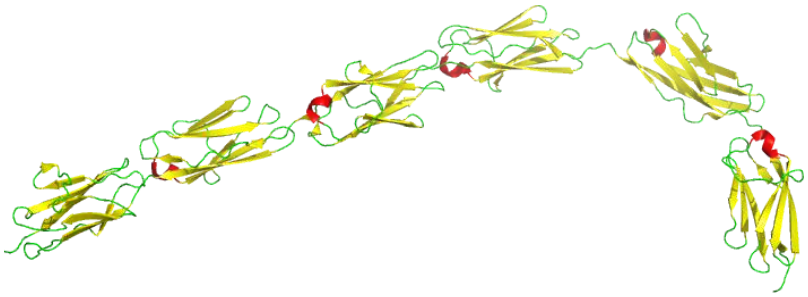


Figure 5.5.8: 6-Ig fragment I65-I70 from the elastic I-band fraction of titin(3b43)

The chemical cycle of adenine nucleotides consists of discrete steps which repeat: binding of ATP, hydrolysis, and dissociation of ADP and P_i. These are linked to a mechanical cycle involving attachment and detachment of myosin heads with the actin thin filaments and sliding of the thin filaments with respect to the thick myosin filaments.

Figure 5.5.9 shows the linked chemical and mechanical states in the cycle detailing how force is generated.

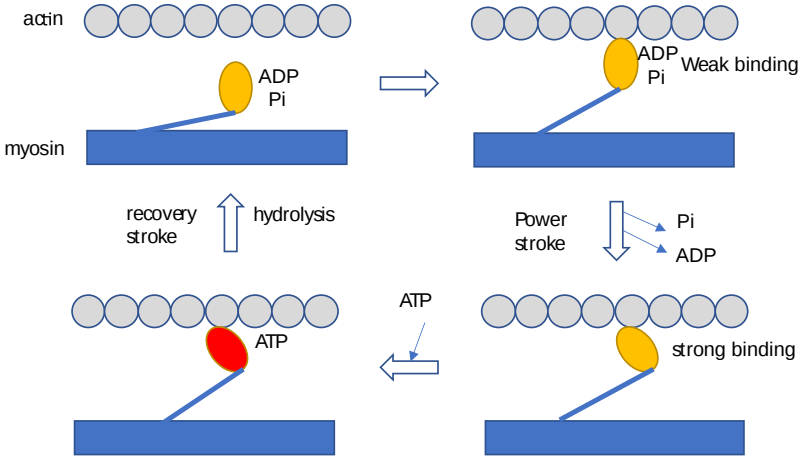


Figure 5.5.9: Chemical and Mechanical Cross-Bridge Cycle for muscle contraction. Adapted from www.pnas.org/cgi/doi/10.1073/pnas.0811630106 and https://www.nature.com/scitable/topi...tion-14567666/

The myosin S1 head is shown at different angles. Starting with the bottom left in the ATP bound form, the head is bent backward to the left toward the rod domain. On hydrolysis of bound ATP, the head rotates into a vertical "cocked" position (top left). In this orientation, the S1 head binds weakly to actin (top right). A power stroke ensues as ADP and Pi dissociate from the myosin S1 head, reorienting the head to an orientation similar to the ATP-bound form.

As it is a cycle, you can start any place along it to follow it. Figure 5.5.10 shows a "linear" presentation of the cycle which starts with Ca^{2+} binding to troponin C. This exposes the myosin binding site on actin and allows the actin binding to the myosin S1 head to form the weakly bound M·ADP·Pi state. which is shown in the upper left of Figure 5.5.9.

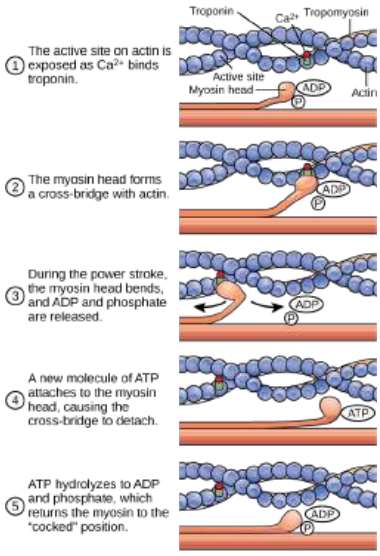


Figure 5.5.10: **Cross-bridge muscle contraction cycle**: The cross-bridge muscle contraction cycle, which is triggered by Ca^{2+} binding to the actin active site, is shown. With each contraction cycle, actin moves relative to myosin. https://bio.libretexts.org/Bookshelv...le_Contraction

In this diagram:

1. In the resting (extended) sarcomere state, ADP and P_i are bound to the motor head of myosin. On neural stimulation of the muscle, Ca^{2+} ions are released from the sarcoplasmic reticulum and bind troponin-C. This leads to the movement of tropomyosin, which exposes the myosin motor head binding site on actin.
2. On binding of actin, the power stroke results, as the actin filaments slide in contraction, leading to the dissociation of ADP and P_i.
3. ATP then binds to the motor head, leading to the detachment of the myosin heads from the actin filament and ATP hydrolysis
4. The motor head cocks to a position where it can interact again with actin filaments and start the process again.

More states also exist including an ATP-bound state that is detached from actin (post rigor), an ADP-bound state (M·ADP) without P_i, and a nucleotide-free myosin state that is bound to actin (rigor). P_i dissociation is the key step that leads to the movement of the lever arm when bound to actin, resulting in the power stroke. ADP dissociates slowly after this so the M·ADP lasts longer.

Figure 5.5.11 shows one last alternative figure with some of the states missing from the diagrams above and with corresponding structural features.

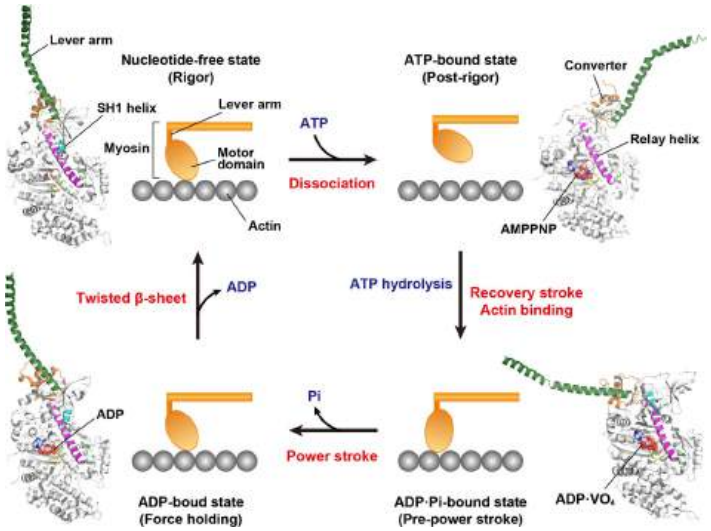
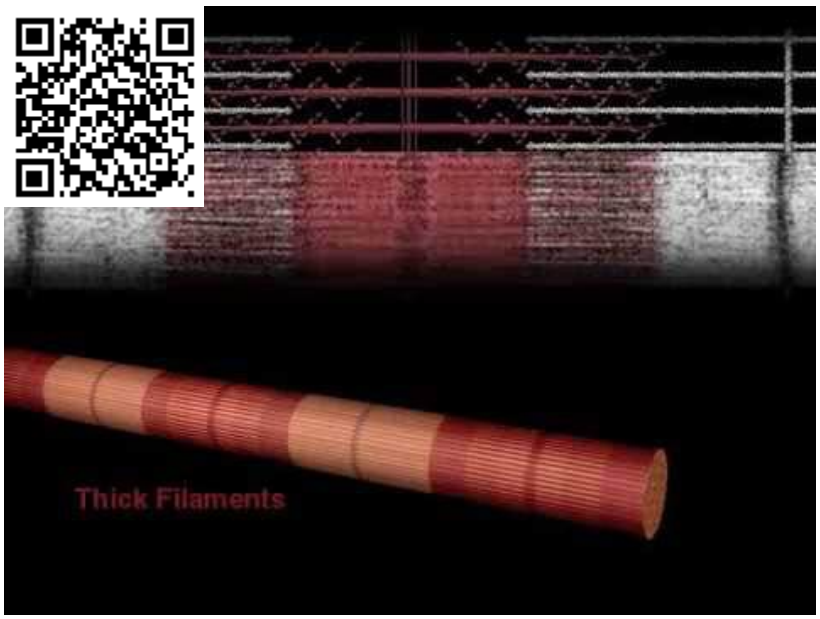


Figure 5.5.11: The structural models by the crystal structures of myosin II (scallop myosin S1; upper left, PDB 1SR6; upper right, PDB 1KQM; lower right, 1DFL; lower left, 3I5F). The lever arm (dark green), converter subdomain (orange) and SH1 helix (cyan) adopt different orientations and conformations in each state. The nucleotide-binding regions, including the P-loop, switch I and switch II, are colored in red, yellow and green, respectively. Each nucleotide in the motor domain is represented by a spacefill model. A power stroke corresponds to a step of force generation". Kato et al, Biophys Rev. 2018 Apr; 10(2): 571–581. Published online 2017 Dec 12. doi: 10.1007/s12551-017-0368-1. Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0)

To make matters more complicated, there are 35 different classes of myosin (MW 227K) and it is involved in functions (vesicle transport for example) other than muscle contraction. We'll explore the structure of the S1 head more closely when we discuss another motor protein, kinesin.

No two-dimensional illustration of the process of sarcomere/muscle contraction can rival the insight gained by viewing a quality simulation of the molecular process. The video found below offers such a simulation. We don't routinely incorporate public videos into this book, in part since the links might not last and they often come with advertisements. This video, which is not narrated but richly annotated so clearly illustrates the ideas presented above, that we choose to incorporate it.



Here is a link to another video: Sliding movement of actin filaments

5.5.2: The cytoskeleton: An Overview

Before we explore two other motor proteins, kinesin and dynein, whose structures are very similar to myosin, we will offer some background on the proteins that make up the interior skeleton of the cell. Once again, the structure of the cytoskeleton is a bit amorphous and constantly changing. It structurally supports the inner organelles and other components of the cell. It facilitates the movement of these structures within the cell. It must dramatically rearrange when cells divide or when they move. More importantly for us, they interact with the motor proteins kinesin and dynein. If we don't know a bit about them, our understanding of motor proteins would lack context. Much of the following overview comes, unless otherwise noted, from a BioLibre introductory biology text by Easlon.

The cytoskeleton is a network of different protein fibers that provides many functions: it maintains or changes the shape of the cell; it secures some organelles in specific positions; it enables movement of cytoplasm and vesicles within the cell; and it enables the cell to move in response to stimuli. There are three types of fibers within the cytoskeleton: microfilaments, intermediate filaments, and microtubules, as shown in Figure 5.5.12 Some of the cytoskeletal fibers work in conjunction with molecular motors, which move along the fibers within the cell to carry out a diverse set of functions. There are two main families of cytoskeletally-associated *molecular motors: dyneins* and *kinesins*.

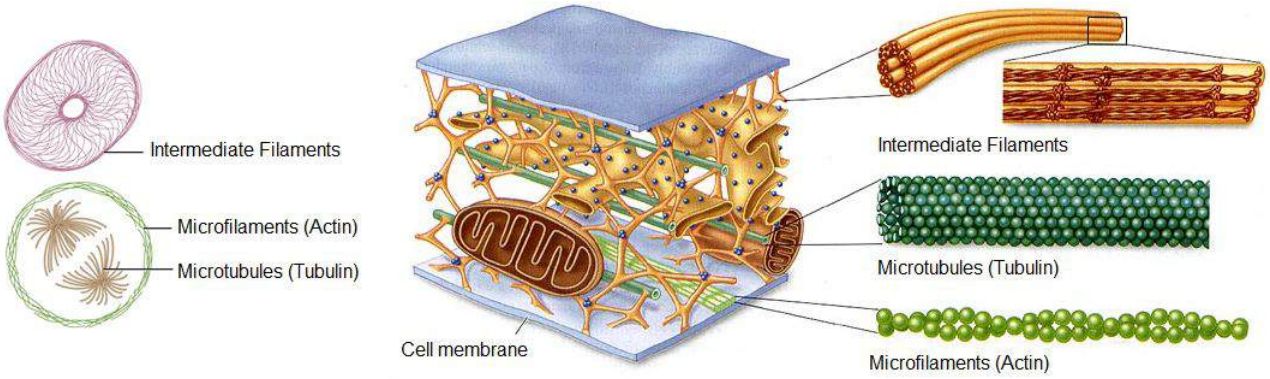


Figure 5.5.12: *Microfilaments thicken the cortex around the inner edge of a cell; like rubber bands, they resist tension. Microtubules are found in the interior of the cell where they maintain cell shape by resisting compressive forces. Intermediate filaments are found throughout the cell and hold organelles in place.*

How can the cell purposely move and control the location of materials between these organelles? More specifically, how can a eukaryotic cell transport compounds from their place of origin (in most cases the cytoplasm) to where they are needed (perhaps the nucleus, the mitochondria, or the cell surface)?

One possible solution is for the cell to create a network that can connect all the different parts of the cell together. This network could be used not only as a scaffold to hold components in place but also as a reference for direction. For example, we can use a map to determine the direction we need to travel and the roads to connect and travel from home to campus. Likewise, an interconnecting network inside the cell can be used to direct and move compounds from one location to a final destination. Some of the required characteristics of this network are listed below. Can you add to this list?

Here are some characteristics of the cytoskeleton:

- The network needs to be extensive, and connect every area of the cell.
- The network needs to be flexible, able to change and adapt as the cell grows larger, divides into two cells, or physically moves from one environment to another.
- The network needs to be strong, and able to hold up to mechanical pressure from inside or outside of the cell.
- The network needs to be composed of different fibers and each of these fibers needs to be for a specific connection in the cell. For example, certain fibers might be involved in holding organelles in place, and other fibers would be involved in connecting two different organelles.
- The fibers need to have directionality (or polarity), meaning they need to have a defined starting point and a defined end to help direct movement from one location to another.
- The fibers need to work with proteins that can convert chemical energy into kinetic energy, to actively transport compounds along the fibers.

Figure 5.5.13 is a bovine pulmonary artery endothelial cells that as been stained to show thin filaments composed of actin (red), and microtubules (green), composed of tubulin.

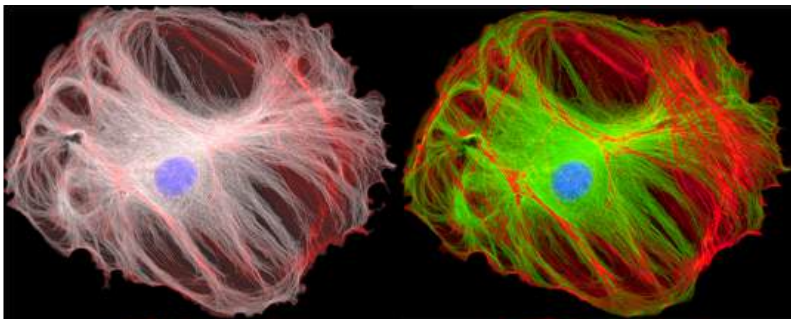


Figure 5.5.13: Actin and microtubules in endothelial cells. https://sitn.hms.harvard.edu/art/201...tin-four-ways/. Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

5.5.3: Microfilaments

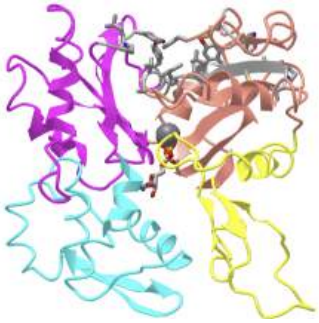
Microfilaments are cytoskeleton fibers composed of *actin* subunits. Actin is the most abundant protein in the cytosol of eukaryotic cells and comprises 20% of total cellular protein by weight in muscle cells. The actin amino acid sequence is highly conserved in eukaryotic cells, meaning that the protein amino acid sequence, and therefore its final 3-D shape, has changed little over the course of evolution, maintaining more than 80% similarity between algae and humans. We have already seen actin filaments as part of the sarcomere.

Actin can be present as either a free monomer called G-actin (globular) or as part of a polymer microfilament called F-actin ("F" for filamentous). Actin must be bound to ATP to assemble into its filamentous form and maintain the structural integrity of the filament. (Note that we discussed above ATP binding to the myosin head and its subsequent hydrolysis during the chemomechanical sarcomere contraction, but we didn't discuss the role of ATP in F-actin formation.)

Recent Updates: Actin - 7/16/23

G-actin has 2 outer subdomains (1 and 2) and two inner ones (3 and 4). Between the outer and inner subdomains are a nucleotide-binding cleft and an opposing hydrophobic cleft involved in actin/actin and actin/actin-binding protein interactions necessary for filament formation.

An interactive iCn3D model of the rabbit F-Actin-ADP complex (2ZWG) is shown in Figure 5.5.14 below.



NCBI iCn3D Figure 5.5.14: Rabbit F-Actin-ADP complex (2ZWG). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...n4poXDwJuqmASA

The large cleft between subdomains 2 (yellow) and 4 (cyan) is the nucleotide-binding cleft. The hydrophobic cleft is between subdomain 1 (salmon) and 3 (magenta). Many of the hydrophobic amino acid side chains in hydrophobic cleft are shown as gray sticks. Note that this form of Actin-ADP is flat on rotation.,

Figure 5.5.15 below shows an animation of the transition from G-Actin-ATP/Ca^{2+} complex (magenta) to F-Actin-ADP/Ca^{2+} complex (cyan)

Figure 5.5.15 below shows an animation of the transition from G-Actin-ATP/Ca^{2+} complex (magenta) to F-Actin-ADP/Ca^{2+} complex (cyan)

Note the conformational changes and the changes in the cleft on the conversion of the G-ATP complex to the F-ADP one. Although the G in G-actin is for "globular", it is not really spherically. Both G- and F-actin are flat. Imagine the proteins as resembling the letter "H" but with the top "cleft" of the H (representing the nucleotide-binding domain) bigger than the bottom "cleft" (representing the hydrophobic cleft). In G actin, one of the long sides of the letter "H" is tilted compared to the other, while in F-actin, both sides are parallel, as shown in Figure 5.5.16 below. Hence F-actin is flatter than G-actin.

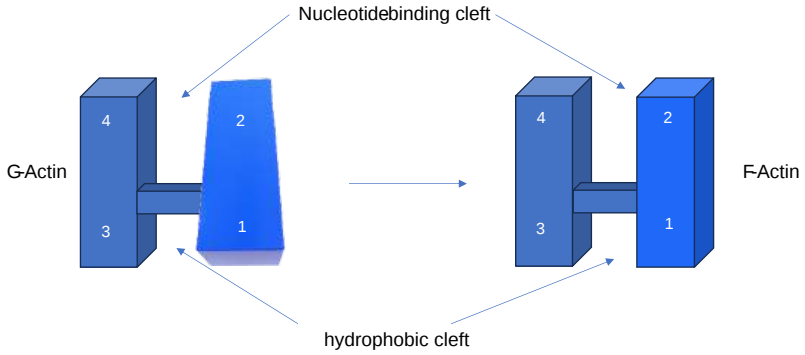
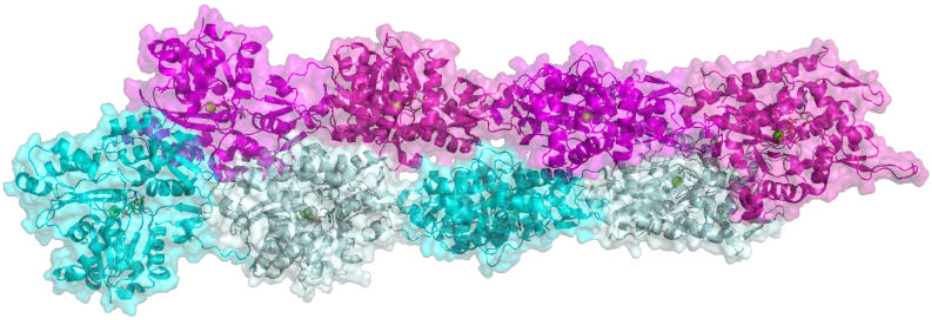


Figure 5.5.16: cartoon showing structural changes in G- to F-Actin transition
(after https://www.science.org/doi/10.1126/science.adg6812)

The actin filament itself has structural polarity implying that it has two distinct ends to the filament. These ends are called the "(-)" end and the "(+)" end. At the "(+)" or "**barbed**" end, **actin·ATP** subunits are added to the elongating filament. At the "(-)" or "**pointed**" end, **actin·ADP** subunits dissociate from the filament. Actin hence is an ATPase. G-actin is a slow ATPase while F-actin is a fast ATPase due to a conformational change in the protein forming a flatter subunit conformation with a more catalytically productive active site.

This process of assembly and disassembly is controlled by the ATP to ADP ratio in the cytoplasm. The general structure of bare F-actin from Oryctolagus cuniculus is shown in an interactive iCn3D model Figure 5.5.17. The bottom strand (repeating subunits shown in cyan and light cyan) winds around the top strand (subunits shown in magenta and light magenta). They twist around each other in a right-handed sense. A complete turn is formed after in a 13-mer. You can see the twist in the figure below which shows only 4 monomers in each strand.



NCBI iCn3D Figure 5.5.17: The structure of bare F Actin (6bno). (Copyright; author via source).
Click the image for a popup (slow load) or use this external link: https://structure.ncbi.nlm.nih.gov/i...QcXoreeyVaHhB7

Figure 5.5.18 below shows the free barbed end (magneta, 8F8R) and the free pointed end (cyan, 8F8S) of actin.

Figure 5.5.18: Free barbed end (magneta, 8F8R) and the free pointed end (cyan, 8F8S) of actin

Proteins interacting at the barbed and pointed ends include CapZ and tropomodulin, respectively. The CapZ binds to the F-actin flat "H" form with the CapZ undergoing significant conformational changes and minor one for the actin end. At the free pointed end, the monomers are more like G-actin in the twisted "H" form. On binding of tropomodulin, the second to the end subunit forms a more F actin conformation.

In the process of filament "treadmilling" (polymerization), actin-ATP binds to the + (barbed) end of F-actin. The bound ATP is quickly hydrolyzed with the Pi slowly released. Actin-ADP dissociates from the - (pointed_ end. Free actin can bind ATP and be ready to add to the + (barbed) end. Proteins like forming can catalyze the addition of actin monomers to the + (barbed) end while proteins like cofillin can promote dissociation from the - (pointed) end. Also both ends can be capped to stop actin addition at the + (barbed) end (by CapZ) or removal from the - (pointed) end (by tropomodulin).

As with myosin-actin interactions, hydrolysis of ATP and release of P_i are key steps. Multiple nucleotide states of actin include A·ATP, A·ADP·P_i, and A·ADP. As the polymer form, a gradient of these nucleotide states exist. On fiber formation, it appears that the key general base, His161, is positioned closer to the labile terminal phosphate on ATP, enabling hydrolysis.

Actin participates in many cellular processes, including muscle contraction, cell motility, cytokinesis during cell division, vesicle, and organelle movement, and the maintenance of cell shape. Actin filaments serve as a track for the movement of a family of motor proteins called *myosins* discussed above

5.5.4: Intermediate filaments

Intermediate filaments are made of several strands of fibrous proteins that are wound together as shown in Figure 5.5.19. These elements of the cytoskeleton get their name from the fact that their diameter, eight to ten nm, is between those of the smaller microfilaments and the larger microtubules. The intermediate filaments are the most diverse group of cytoskeletal elements. Several types of fibrous proteins are found in the intermediate filaments. You are probably most familiar with keratin, the fibrous protein that strengthens your hair, nails, and the epidermis of the skin.

Figure 5.5.19: *Intermediate filaments consist of several intertwined strands of fibrous proteins.*

Intermediate filaments have no role in cell movement. Their function is purely structural. They bear tension, thus maintaining the shape of the cell, and anchor the nucleus and other organelles in place. The figure above shows how intermediate filaments create a cable-like supportive scaffolding inside the cell.

5.5.5: Microtubules

Microtubules are the largest component of the cytoskeleton and are found throughout the cytoplasm. These polymers are made up of globular protein subunits called *α-tubulin* and *β-tubulin*. Microtubules are found not only in eukaryotic cells but in some bacteria as well.

Both the α-tubulin and β-tubulin subunits bind to GTP instead of GTP. When bound to GTP, the microtubule starts to form with the first tubulins "nucleating" the growth of the growing tubule. As more GTP tubulin dimers assemble onto the filament, GTP is

slowly hydrolyzed by β-tubulin to form GDP. Tubulin bound to GDP is less structurally robust and can lead to disassembly of the microtubule.

Much like the actin filaments discussed above, microtubules also have a distinct polarity that is critical for their biological function. Tubulin polymerizes end to end, with the β-subunits of one tubulin dimer contacting the α-subunits of the next dimer. These differences lead to different subunits being exposed on the two ends of the filament. The ends are designated the "(−)" and "(+)" ends. Unlike actin filaments, microtubules can elongate at both the "(+)" and "(-)" ends, but elongation is significantly more rapid at the "(+)" end. The structure of tubulin is shown in Figure 5.5.20. One turn of the tubulin (7diz) in cortical microtubules from the human parasite Toxoplasma gondii with proteins inside the tubulin polymers is shown. The alpha subunits are shown in cyan, beta in magenta and the inner proteins specific to this tubulin polymer in gray.

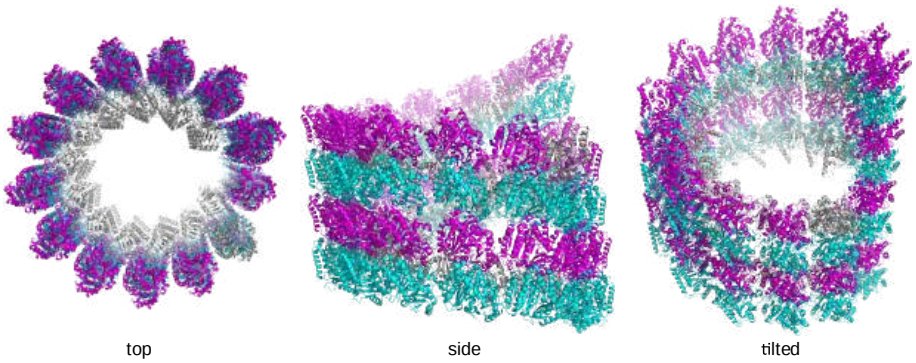


Figure 5.5.20: The structure of microtubules. Microtubules are hollow. Their walls consist of 13 polymerized dimers of α-tubulin and β-tubulin (right image). The left image shows the molecular structure of the tube.

Microtubules help the cell resist compression, provide a track along which vesicles move through the cell, pull replicated chromosomes to opposite ends of a dividing cell, and are the structural elements of flagella, cilia, and centrioles (the latter are the two perpendicular bodies of the centrosome). In fact, in animal cells, the centrosome is the microtubule organizing center. In eukaryotic cells, flagella and cilia are quite different structurally from their counterparts in bacteria, as discussed below.

A key function of microtubules is to move molecular "cargo" along the microtubule which acts like a "railroad" track. Two key motor proteins are involved in binding to target cargos, and hauling them along the polymer. These motor proteins are kinesin, which moves cargo towards the (+) end, and dynein, which moves it to the (-) end. These interactions are illustrated in Figure 5.5.21.

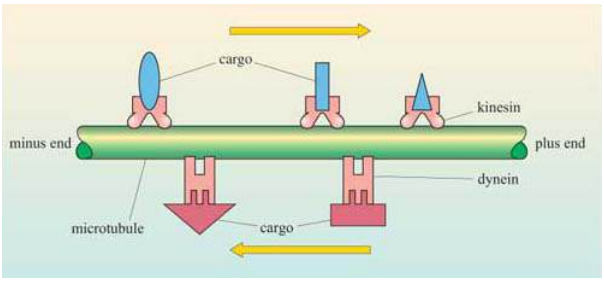


Figure 5.5.21: Kinesins and dyneins carry cargo along microtubules. Kinesin moves towards the plus end of the microtubule, whereas the dyneins move towards the minus end.

Here is an animation of the dynamics of micotubule and cargo movement.

5.5.6: Motor Proteins - Cargo movent along microtubule tracks

We will explore the motors protein kinesin and dynein in more detail. As they serve similar functions, you might expect them to have similar structures. Let's see!

Some of the material is from https://www.open.edu/openlearn/scien...nt-section-5.2

5.5.6.1: Kinesins

Motor proteins bind to vesicles and organelles and use energy from ATP to move them along the microtubule or microfilament network. Two families of motor proteins, the **kinesins** and **dyneins,** move vesicles along microtubules, and members of the **myosin** family move them along microfilaments. The myosin family is also important in cell movement.

The kinesin superfamily (KIF) consists of over 40 different proteins, divided into about 16 families. They bind both microtubules and ATP and are molecular motors. The first ones identified were involved in cargo (organelles, mRNAs, proteins) transport along the long axons of neuronal cells. This method of transport would be vastly superior than simple diffusion down the axons. The kinesin-1 family (members include KIF5A, KIF5B and KIF5C) are heterotetramers with two heavy and two light chains, somewhat similar to myosin.

As with myosin, kinesins have a motor head domain, which binds and hydrolyses ATP. They have a microtubule-binding domain instead of an actin-binding region. The N-kinesins have their motor domain at the end terminal, while the domain for M- and C-kinesins are in the middle or carboxy end, respectively. Both N-kinesins and C-kinesins are responsible for plus end and minus end-directed motility, and M-kinesins are used for depolymerization of microtubules in tubulin molecules. As with myosin, the neck stalk and tail are coiled coils of the heavy chains. Cargos bind through intermediary adapter proteins although some bind directly to kinesin tails.

The direction of movement of bound vesicles along the cytoskeleton is absolutely dependent on the polarity of the microfilaments and microtubules. Some motor proteins move from the minus end to the plus end and others in the opposite direction. For example, of the various myosins that have been discovered throughout the animal and plant kingdoms, all but one (myosin VI) move towards the plus end of the filament.

Kinesins have a tertiary structure that is similar to myosin II, even though there is no significant similarity in the primary structure. Both molecules have two heads with motor domains formed around an ATP-binding core, and a coiled tail that binds to the cargo, as shown in Figure 5.5.30. A number of other molecules are related to kinesin, and all of them share the kinesin motor domain, but very little else. These are the **kinesin-related proteins**. Kinesin itself moves towards the plus end of microtubules, but other members of the kinesin family move to the plus or minus end depending on the protein. Some of the kinesin-related proteins are involved in moving microtubules during mitosis – in this way the motor protein and the microtubule act in an analogous way to myosin and microfilaments in cell movement.

Figure 5.5.22 below shows an interactive iCn3D model of the neck and motor domains of dimeric Kinesin-3 KIF13B from rat (6A1Z)

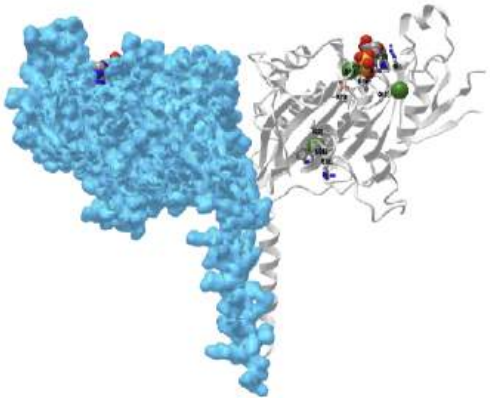


Figure 5.5.22: Neck and motor domains of dimeric Kinesin-3 KIF13B from rat (6A1Z). (Copyright; author via source). Click the image for a popup (slow load) or use this external link: https://structure.ncbi.nlm.nih.gov/i...LTKsQN4xcN9B99

One monomer is shown as a blue surface and the other as a gray cartoon. ADP is shown in colored spacefill. Amino acid side chains involved in nucleotide and microtubule binding are shown as coloreds sticks and labeled.

(Dyneins, which we will explore below, are unrelated to either kinesins or myosins, and they move towards the minus end of microtubules. Each is composed of two or three heavy chains, with the cytoplasmic dyneins having two chains, each of which forms a large motor domain. In nerve cells, the axonemal dyneins, which have two or three motor domains, transport vesicles along microtubule bundles in the axons.)

The speed of the movement mediated by dyneins and kinesins is signficant. In vitro, kinesins can move along microtubules at 2 μm s^{-1} and dyneins at up to 14 μm s^{-1}. Although these high rates of movement would not be achieved in the complex environment of a cell, they can explain, for example, how caveolar transcytosis of molecules (from the luminal side of the blood vessel to the subendothelial space) across an endothelial cell can occur in 1–2 minutes. Movement and force generation by both classes of proteins involves ATP hydrolysis and allosteric shifts in the orientation of the motor domains, so that the proteins are thought to 'step' progressively down the microtubule.

Figure 5.5.23 shows a cartoon of a kinesin dimer attached to a microtubule.

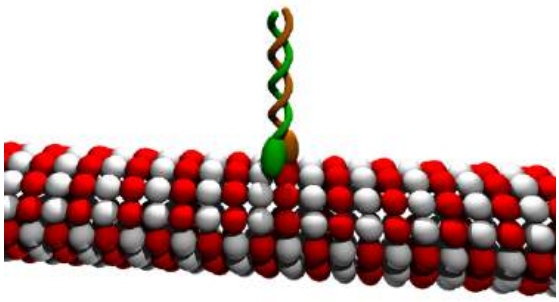


Figure 5.5.23: Cartoon of kinesin dimer attached to a microtubule. Moez. https://commons.wikimedia.org/wiki/F...in_cartoon.png. Creative Commons Attribution-Share Alike 3.0 Unported

Figure 5.5.24 shows an animated cartoon of Kinesin walking on microtubule.

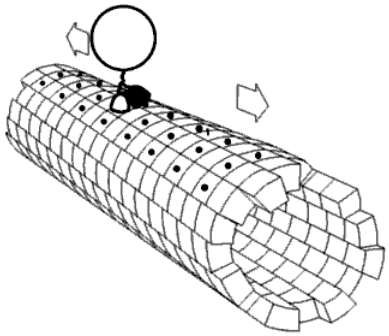


Figure 5.5.24: Cartoon of Kinesin walking on microtubule. By Jzp706 - Own work, CC0, https://commons.wikimedia.org/w/inde...=13188271jdkfj

Figure 5.5.25 shows an animation showing a kinesin dimer bound to a tubulin microtubule .

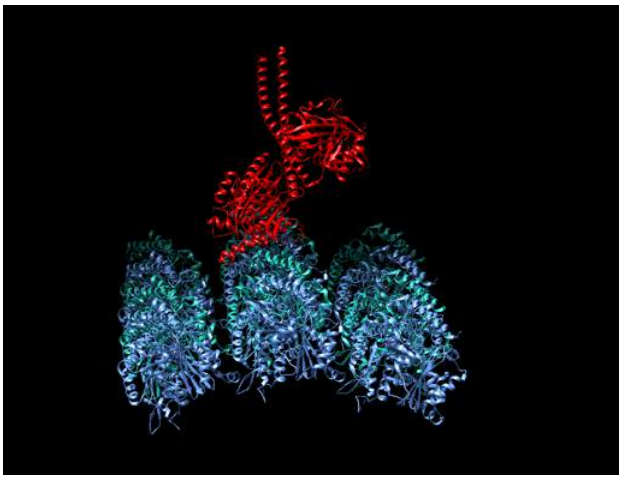


Figure 5.5.25: Kinesin dimer bound to a tubulin microtubules. Kaden.Rawson, MarisaAsadian, Stevengparkin, CC BY-SA 4.0. https://creativecommons.org/licenses/by-sa/4.0, via Wikimedia Commons

Figure 5.5.26 shows two motor domains of kinesin-1 bind to the microtubule lattice simultaneously.

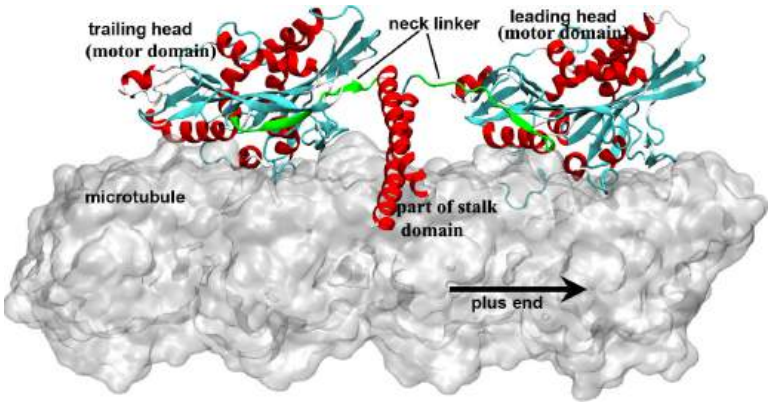


Figure 5.5.26: Two motor domains of kinesin-1 bind to the microtubule lattice simultaneously. Qin, J.; Zhang, H.; Geng, Y.; Ji, Q. How Kinesin-1 Utilize the Energy of Nucleotide: The Conformational Changes and Mechanochemical Coupling in the Unidirectional Motion of Kinesin-1. *Int. J. Mol. Sci.* **2020**, *21*, 6977. https://doi.org/10.3390/ijms21186977. Creative Commons Attribution (CC BY) license (**http://creativecommons.org/licenses/by/4.0/**)

The neck linkers of the kinesin-1 dimer are colored green. The leading head is in the nucleotide-free state and has an undocked neck linker, which points to the minus end of the microtubule. The trailing head is in the ADP·Pi/ADP-bound state and has a plus-

end pointed neck linker.

Notice however that the ATP-binding site of kinesin is located at the distal tip of the motor domain, whereas in myosin the equivalent site is deep within the motor domain and covered by the myosin's actin-binding site. Therefore the mechanism of stepping is different in the two molecules. In particular, the α-helical linking region connecting the two motor domains of kinesin appears to transfer allosteric changes between them to coordinate ATPase activity and hence the stepping motion of the protein. It is interesting that the motor domains of kinesin-related proteins that move to the plus end and those that move to the minus end of the microtubule are similar but the linkage between them is quite different. Kinesin has its motor domain near the N-terminus, whereas the molecule Ncd, which moves to the minus end of the microtubule, has its motor domain located near the C-terminus. It seems that whether the protein is directed to the plus or the minus end is dependent on the configuration of domains and the coordination.

Figure 5.5.27: shows a comparison of structures of the motor head domain of myosin (panel A) and kinesin·ADP (panel B) and kinesin·ATP (panel C). T the structure of these proteins varies with the form of the bound nucleotide.

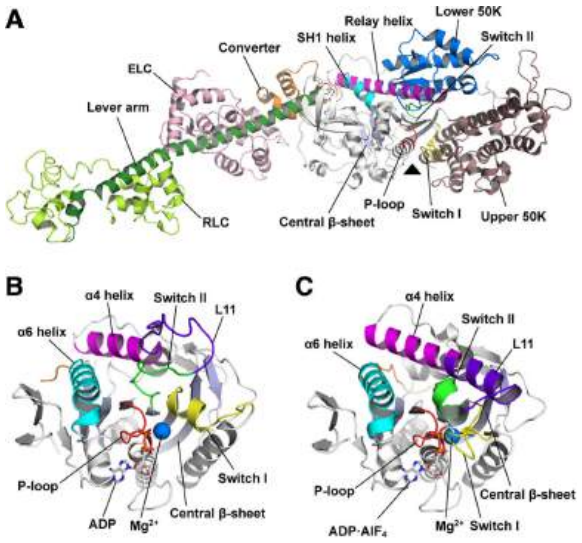


Figure 5.5.27: Structures of myosin and kinesin. **a** Ribbon diagram of the motor domain, lever arm, essential light chain (ELC) and the regulatory light chain (RLC) of myosin II (scallop myosin S1; PDB 1SR6). The triangle indicates the nucleotide-binding site. **b** The motor domain of ADP-bound kinesin (KIF5B; PDB 1BG2). ADP and Mg^{2+} are represented by stick and sphere models, respectively. **c** The motor domain of ATP-bound KIF5B (PDB 4HNA). ADP·AlF_4 Kato et al., ibid

The myosin head has a 7-stranded central β-sheet and 3 α-helices on each side of the sheet. Nucleotides bind in the β-sheet regions which as a P (phosphate binding)-loop and two "switches". On the back side of the β-sheet are two 50K domains, in between which actin binds when exposed. A converter and lever arm follow leading to the coiled rod domain. The motor domain of kinesin is smaller but homologous with the tubulin binding site also on the back side of the β-sheet. These structures suggest similar mechanisms of action for myosin and kinesins. However, in contrast to the importance of P_i release in the power stroke for the sarcomere, it appears that ADP/ATP are the key players. When ADP is released, the central β-sheet twists, leading to binding to microtubules, implying that microtubules act to exchange ADP/ATP to generate the force. The linkage of conformational change and binding in microtubule:kinesin chemomechanical coupling is shown in Figure 5.5.28

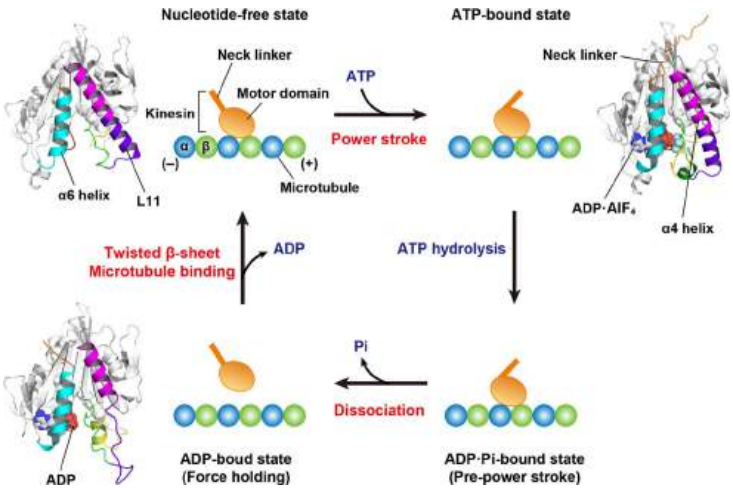


Figure 5.5.28: The mechanochemical cycle of N-type kinesin. The structural models with ribbon diagrams are created by the crystal structures of kinesin-1 (KIF5B; upper left, PDB 4LNU; upper right, PDB 4HNA; lower left, 1BG2). The structures in individual steps adopt the different conformations of the neck linker (orange), α6 helix (cyan), α4 helix (magenta), L11 loop (violet), switch I (yellow), switch II (green) and P-loop (red). Each nucleotide in the motor domain is represented by a spacefill model. Microtubules are composed of hetero-oligomers composed of α-tubulin (α) and β-tubulin (β). The neck linker moves from the minus-end direction (left) along a microtubule to the plus-end direction (right) via the power stroke, which corresponds to the step of force generation. Kato et al., ibid

Kinesin is permanently attached to the microtubule, by either one head or the other. By comparison, the myosin heads (which are arranged in bundles in myofibrils) are only in contact with actin filaments for about 5% of each movement cycle.

5.5.6.2: Dynein

Like myosin and kinesin, dynein is a motor protein that undergoes chemomechanical cycles to move cargo on microtubules in the direction of the (-) end. Less is known about dynein given its much larger size (MW 1.4M) than myosin and kinesin, and the difficulty in studying it. There are cytoplasmic and flagellar versions of the protein. The latter generates a force that leads to flagellar motion and propulsion of the cell. They are part of the **A**TPases **A**ssociated with diverse cellular **A**ctivities (**AAA+** proteins) superfamily. These often form hexameric ring complexes that are involved in the nucleotide-dependent changes in protein structure. Some acts as protease, chaperone, transcription factors, etc. Dynein is different from other AAA+ proteins as it is a single protein that includes six domains that arrange to form a hexameric ring structure. The domain structure and model of dynein bound to a microtubule are shown in Figure 5.5.29

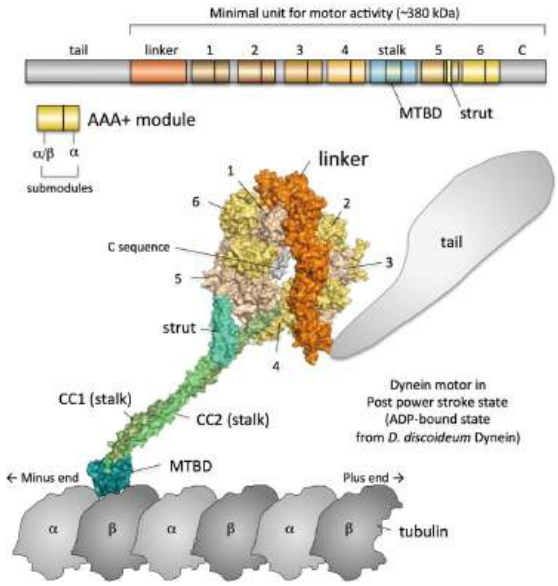


Figure 5.5.29: Overall structure of dynein. *Upper panel* Domain composition of dynein. The numbers of AAA+ modules (AAA1–AAA6) are indicated. 'C' indicates the C sequence. *Lower panel* The overall structure of dynein in the post-power stroke state. CC1 and CC2 indicate coiled-coil helices 1 and 2 of the stalk, respectively. Kato et al., ibid.

The structure contains:

- six AAA+ module domains with different sequences that nevertheless form a hexameric ring (labeled 1-6 in Figure 5.5.x above;
- a lever-like linker domain that stabilizes the ring domains, connects to the tail domain and produces the power stroke;
- a coiled-coil antiparallel stalk domain contains two intertwined alpha-helices;
- a globular microtubule-binding domain (MTBD) that in the primary sequences of the protein separates the two alpha-helices in the stalk and which binds the "track" (microtubule).
- a tail-domain that binds to cargos, or if multiple dyneins interact, other dyneins.

It is quite amazing that dynein has a 3D structure (globular cytoskeleton binding domain - stalk - tail) that mimics that of myosin and kinesin, but all in a single protein. The power stroke is linked to lever motion and ATP hydrolysis. We must use this structural information to understand ATP binding and hydrolysis, conformational changes in the linker during the power stroke, transmission of conformational changes through the stalk and ensuing changes in the interaction of the microtubule-binding domain with the microtubule.

Nucleotide binding and hydrolysis

- the lever-like linker swings between AAA+2 and the stalk base(Roberts et al. 2009).
- AAA+ modules 1-4 can bind an cleave ATP so they acts as ATPases. This contrasts to both myosin and kinesis which have only 1 ATP binding site per heavy chain. AAA+ 5 and 6 have mutations, which prevent ATP hydrolysis.
- The ATPase site of AAA1 is between AAA1 and AAA2 binds ADP with higher affinity than ATP since an arginine side chain on AAA2 is not close enough to interact with the gamma-phosphate of ATP. The ADP-bound form is called the open form.
- Crystals structure of ADP·vanadate (an equivalent ATP analog) shows a closed active site in which the arginine side chains interact with the vanadate, a proxy for the gamma-phosphate of ATP. Hence ATP binding closes this active site. A

Figure 5.5.30shows the open (ADP bound) and closed (ATP equivalent) bound state of dynein.

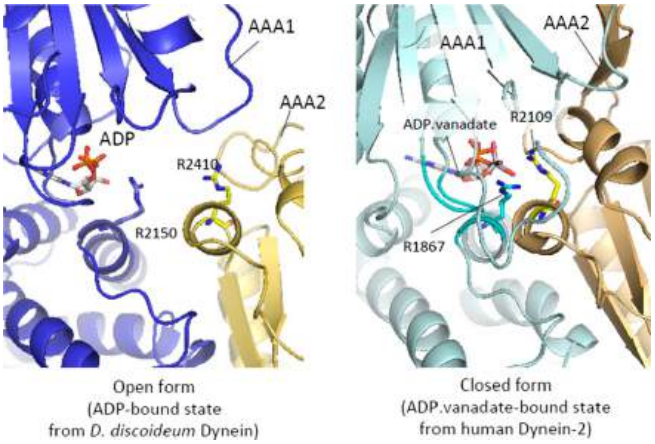


Figure 5.5.30: Cleft closure with an ADP.Pi analog. *Left panel* The distance between ADP and R2410 is long, and AAA1 (blue) and AAA2 (gold) are detached from each other. *Right panel* The distance between ADP and R2109 corresponding to R2410 in the left panel is short, and the cleft between AAA1 (pale cyan) and AAA2 (light orange) is closed. Kato et al., ibid.

Linker conformational changes in power stroke

Linker motion produces the power stroke and the ADP bound form represents a post-stroke state in which the linker has minimal contact with the ring domains and is mostly detached. A long α-helix, H10, within the linker appears to be involved in the power stroke as it extended in th ADP bound state, but bent in the ATP proxy state, where it is posed to initiate the power stroke. These conformations are illustrated in Figure 5.5.31.

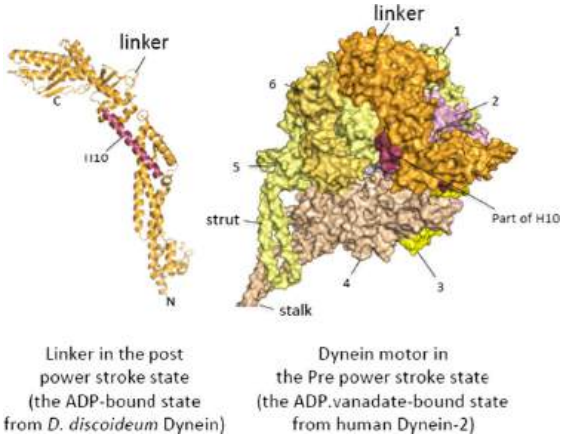


Figure 5.5.31: Structural change of the linker. *Left panel* Ribbon model of the linker. The hinge helix H10 is highlighted. *Right panel* The AAA+ ring in the pre-power stroke state has the linker bent at H10. Kato et al, ibid.

Communication between the domains for ATP hydrolysis and track binding

Obviously, there needs to be coordination and correcting timing between changes in the ATPase hexameric ring and its communication through the linker to the stalk and to the microtubule-binding domain (MTBD), which is 140 Å distant. This communication is easier for kinesins and myosins since the motor domain, where ATP cleavage occurs, is also the domain that binds actin and tubulin, respectively. The communication between the distal domains occurs through the coiled-coiled stalk. This makes sense as the MTBD is in between the two helices of the stalk.

The two antiparallel helices presumably slide over each other to form different alignments or registries between the helices. Indeed different alignments appear to affect the affinity of the MTBD for tubulin. This was experimentally verified by creating a new protein with the coiled-coil stalk of dynein with another protein, seryl-tRNA synthetase, as illustrated in Figure 5.5.32

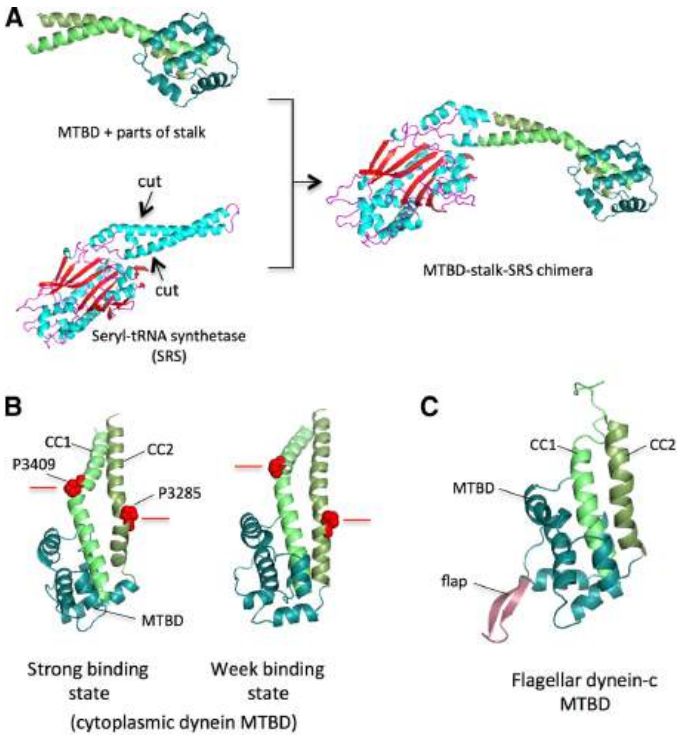


Figure 5.5.32: Registry change and structure of the microtubule-binding domain (MTBD). **a** The coiled-coil of seryl-tRNA synthetase (SRS) was ligated with the stalk of the MTBD to produce constructs with various coiled-coil registries. **b** Coiled-coil registries were different between the strong and weak binding states of the MTBD. In the figure, the height of P3409 relative to that of P3285 changes in two different states. **c** The MTBD of flagellar dynein-c contains the characteristic flap. Kato et al., ibid.

Interactions with tubulin

There also appear to be two main conformations for the MTBD for its interaction with tubulin. Figure 5.5.x below shows the domain structure of dynein and various synthetic constructs used in experiments (panel A), a model structure for the interaction of dynein with the α/β-tubulin dimer (panel B) and a detailed rendering of the strong (alpha-helix registry) and weak (beta structure registry) of the MTBD and the α/β-tubulin dimer. During each full cycle, the MTBD must dissociate from and reassociate with the microtubule as it marches along the polymer (hence the need for weak and strong binding interactions). Figure 5.5.33 below shows the domain organization of dynein and low- and high-affinity bound states of the MTBD.

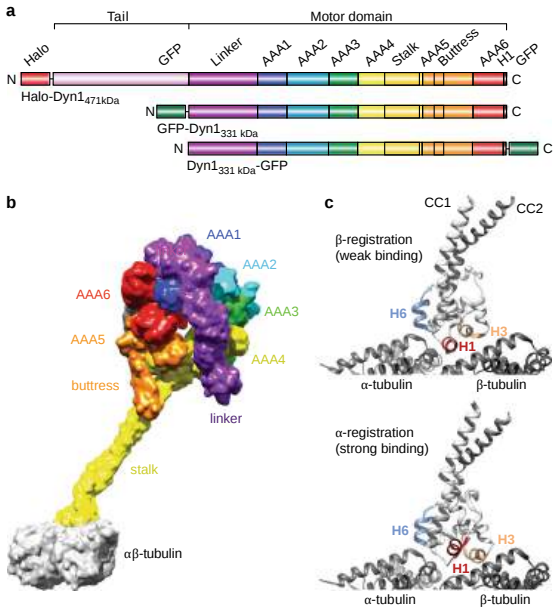


Figure 5.5.33: Domain organization of cytoplasmic dynein and low- and high-affinity MT-bound states of the dynein MTBD. **a** Organization of the full-length cytoplasmic dynein heavy chain (HC) with an N-terminal Halo-Tag, Halo-Dyn1471kDa (a.a. 1–4092) and the tail-truncated monomeric constructs, GFP-Dyn1331kDa and Dyn1331kDa-GFP (a.a. 1219–4092). **b** Dynein MD bound to α/β-tubulin in the strong binding state (merged from PDB entries 3VKG and 3J1T; see Supplementary Note 1). **c** MT-bound MTBD in the weak MT-binding β-registration of the stalk helices (top, PDB entry 3J1U) and in the strong MT-binding α-registration (bottom, PDB entry 3J1T). Lu Rao et al. Nature Communications (2019) 10:3332 | https://doi.org/10.1038/s41467-019-11231-8. Creative Commons Attribution 4.0 International License.

5.6: Binding - Conformational Selections and Intrinsically Disordered Proteins

5.6.1: Conformational Selection

In our study of hemoglobin structure in the MWC model, we developed the idea that there were two forms of hemoglobin in solution, the taut and relaxed form, which are pre-existing and interconvertible even in the absence of dioxygen. Oxygen was presumed to bind preferentially to the relaxed form. In the KNF model we saw that ligand binding can induce conformational changes in adjacent subunits, promoting cooperative binding of ligand. In general, these two models distill down to combinations of two simpler models. The first might be called the conformational selection in which ligand binds tightly to a preexisting conformations in a "lock and key manner" without inducing subsequent macromolecular conformational change. Alternatively, the ligand might bind loosely and alter the macromolecular conformation to produce tighter binding, an example of an induced fit model. For the binding of dioxygen to hemoglobin, thermodynamic cycles could be drawn showing either the binding of ligand and subsequent conformational changes in protein structure or conformational changes in protein structure proceeding binding. Is there additional evidence to support the conformational selection model of binding of ligand to a protein that can exist in two conformations in the absence of ligand? The answer is yes.

Antibodies are immune system protein molecules than can bind "foreign" molecules and target them for biological neutralization. Many crystal structures have been determined of antibodies in the presence or absence of a "foreign" ligand molecule. In these cases, the conformation of the bound antibody is different from that of the free. Either an induced fit model for ligand binding or a lock and key model of binding of ligand to one of two different pre-existing antibody conformations could account for this observation. These different mechanisms could be differentiated experimentally by stop-flow kinetic technique since both display slow and fast phases that are affected differently by ligand concentration. Theoretically, in the induced fit model, only one ligand type could bind to the antibody which would undergo a conformational rearrangement to produce tighter binding. However, a different structural ligand might bind to the two main antibody conformations in the two preexisting conformational models. James et al. have recently shown through stop flow kinetics techniques (to investigate binding) and x-ray crystallography (to investigate final structures) that one antibody molecule can, through existing in two different preexisting conformations, bind two different ligands (antigens). One antibody conformation binds small aromatic molecules with low affinity (including the small molecule 2,4-dinitrophenol, the immunizing molecule or hapten) and then rearranges to produce a high affinity binding complex in which the DNP is bound in a narrow cavity (reducing the effective off rate (k_{off}) of the bound ligand. A second antibody conformation binds a protein ligand over a broad, flat binding site of the antibody molecule.

Lange (2008) et al, using a NMR technique, residual dipolar coupling, that allows sampling of structures in the microsecond time scale, have shown that the solution structure of ubiquitin (which we modeled in our first lab), in the absence of ligand, exists in an ensemble of conformational states. More importantly, these different conformational states are identical to those found in the 46 crystal structure of ligand complexed to various protein ligands, strongly supporting the concept of conformational selection. In all likelihood, a combination of both induced fit and conformational selection probably occurs within a 3D energy landscape in which an initial binding encounter by either a lock and key fit to the "optimal fit" conformer or to a higher energy conformer in which the bound state relaxes to a lower energy through the induction of shape changes in the binding protein.

Figure 5.6.1 shows a cartoon illustrating the differences between conformational selection and induced fit binding (after Boehr and Wright, Science 320, 1429 (2008)).

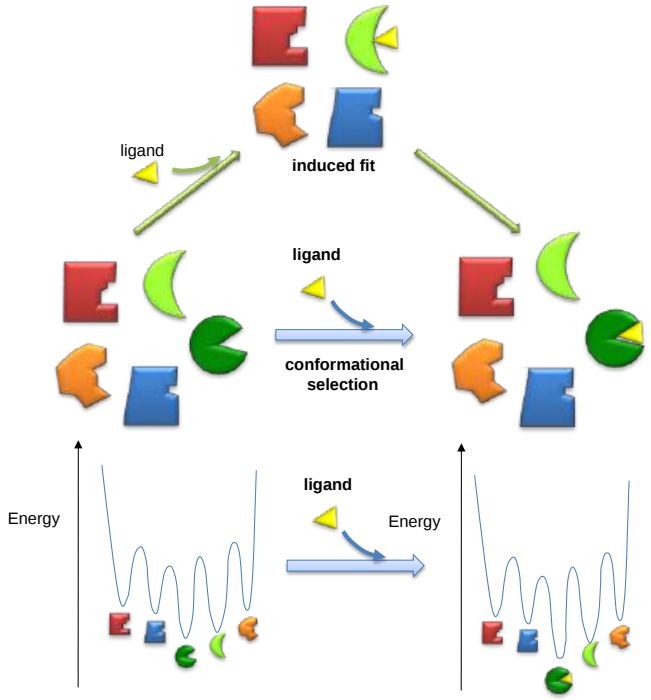


Figure 5.6.1: Conformational Selection vs Induced Fit Binding (after Boehr and Wright, Science 320, 1429 (2008))

Rea et al. offered an interesting experimental model to distinguish conformational selection versus induced ligand binding. They studied rabbit ileal bile acid binding protein (I-BABP). The wild-type protein has a helix-turn-helix motif at its N terminus. They produced a mutant (Δa-I-BABP) that replaced this motif with a Gly-Gly-Ser-Gly linker, causing the protein to unfold. Next, they conducted binding and folding studies on addition of taurochenodeoxycholate (TCDC) using stopped-flow fluorescence to measure the binding behavior. They wished to distinguish between two distinct mechanisms – folding before binding (or conformational selection) and binding before folding (or induced-fit model). The data support a two-phase model. One phase did not depend on ligand and one did, suggesting binding followed by a conformational change).

Conformational Selection

Equation5.6.1 below describes the equilibria involved in the conformation selection model. The forward rate constants are shown as k_n while the reverse ones are shown as k_{-n}.

$$P \underset{k-1}{\overset{k_1}{\leftrightarrow}} P^* + L \underset{k_{-2}}{\overset{k_2}{\leftrightarrow}} P^*L \tag{5.6.1}$$

P* in the conformational selection model represents a high affinity, pre-existing conformation of the protein.

Induced Fit

Equation5.6.2 below describes the equilibria involved in the induced fit model.

$$P + L \underset{k-1}{\overset{k_1}{\leftrightarrow}} PL \underset{k_{-2}}{\overset{k_2}{\leftrightarrow}} P^*L \tag{5.6.2}$$

P* in the induced fit models results when high ligand shifts the equilibrium to the right.

One way to differentiate these models is to look at the dependency of the different kinetic phases on ligand. In the conformation selection model, the slow step is the formation of the high affinity form of the protein, P*. The first slow step has a nonlinear dependence in L while the fast second step has a linear dependence. The data did not fit this model well.

$$\begin{aligned} k_{\text{slow}} &= k_{-2} + \frac{k_2}{1 + \frac{L}{\left(\frac{k_{-1}}{k_1}\right)}} \\ k_{\text{fast}} &= k_{-1} + k_1 L \end{aligned} \tag{5.6.3}$$

In the induced fit model, the ligand binds to a low affinity and perhaps unfolded form of the protein, which subsequently collapses to the bound form in a slow step.

$$k_{\text{slow}} = k_{-2} + \frac{k_2 L}{\left(L + \frac{k_{-1}}{k_1}\right)} \qquad (5.6.4)$$
$$k_{\text{fast}} = k_{-1} + k_1 [L]$$

Both ligand-dependent and independent phases are evident in the equation for the slow step for the induced fit mechanism. At high ligand concentration (when L >> k_{-1}/k1) , the slow step in the induced fit would be independent of ligand (k_{slow} = k_{-2} + k_2). The authors state the data is consistent with a variant of induced fit called the "fly casting model". In this model, the protein first encounters ligand and forms a hydrophobic collapse intermediate (PL) in a fast step characterized by a linear dependence on ligand concentration. Then the intermediate slowly interconverts into a wild type like complex through conformational re-arrangement. Wild-type protein binds the ligand 1000x as quickly, suggesting entropic barriers to binding of the ligand to the unfolded state and rearrangement of the protein thereafter.

Junker et al used atomic force microscopy (AFM) to observe the effects of ligand binding on the folding/unfolding fluctuations of a single molecule of calmodulin (CaM), a calcium-binding protein that binds amphiphilic helicals peptides, leading to a large conformation change in the protein. To do this, they sandwiched a single CaM molecule between filamins that serve as attachment points for the AFM tip and a surface. A slow pulling force was then applied to the molecule, and the length gain was measured as the protein unfolded. The rapid fluctuations between folded and unfolded states were quantified and used to derive a complete energy landscape for the folding of CaM. They conducted these experiments in the presence of two ligands, Ca^{2+} and mastoparan (Mas), a wasp venom peptide. They found that Mas does not affect the folding rate of CaM, although it does stabilize the already folded form. This suggests that Mas does not bind to the transition state or the unfolded protein, but rather selects a particular conformation from an ensemble of possible choices. Ca^{2+} however, increases the folding rate, which suggests that it stabilizes both the transition state and the folded state. AFM offers a considerable degree of precision in drawing energy landscapes of protein folding and unfolding, and it has several applications that are yet to be explored.

5.6.2: Binding to Intrinsically Disorder Protein and MORFs

As described above, the binding of a protein to a ligand (including another protein) could occur by a lock and key mechanism, possibly through a conformational selection process, or through an induced fit when an initial binding event is followed by a conformation rearrangement to form a more tightly bound complex. But how does binding to completely intrinsically disordered protein (which has been documented) occur? These cases are quite removed from those envisioned in simple induced fit mechanisms. Binding to IDPs might occur through specific **Molecular Recognition Features (MoRFs)**.

MoRFs are typically contiguous but disordered sections of a protein that first encounter a binding partner (a protein for example). Mohan et al conducted a structural study of MoRFs in the Protein Data Bank by selecting short regions (less than 70 amino acids) from mostly disordered proteins that were bound to proteins of greater than 100 amino acids. They chose a sequence size of 70 amino acids and smaller since they would most likely display conformational flexibility before binding to a target. 2512 proteins fit their criteria. For comparison, they created a similar database of ordered monomeric proteins. The analysis showed that after they encounter a binding surface on another protein, the MoRF would adopt or "morp" into several types of new conformations, including alpha-helices (a-MoRFs), beta-strands (b-MoRFs), irregular strands (i-MoRFs) and combined secondary structure (complex-MoRFs), as shown in the figure below.

Figure: Types of Molecular Recognition Features in Intrinsically Disordered Proteins

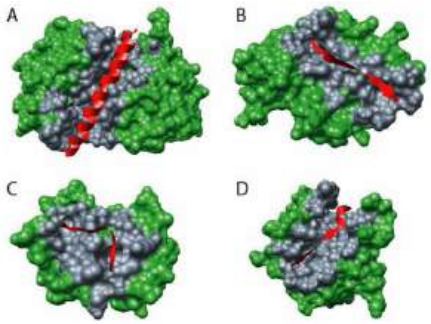
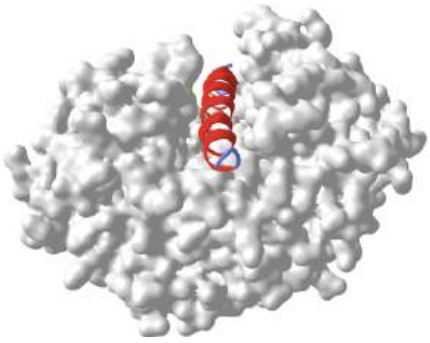


Figure 5.6.2: Types of Molecular Recognition Features in Intrinsically Disordered Proteins. (A) α-MoRF, Proteinase Inhibitor IA3, bound to Proteinase A (PDB entry 1DP5). (B) A β-MoRF, viral protein pVIc, bound to Human Adenovirus 2 Proteinase (PDB entry 1AVP). (C) An ι-MoRF, Amphiphysin, bound to α-adaptin C (PDB entry 1KY7). (D) A complex-MoRF, β-amyloid precursor protein (βAPP), bound to the PTB domain of the neuron specific protein X11 (PDB entry 1X11). Partner interfaces (gray surface) are also indicated. Vacic, V. et al. Journal of Proteome Research 6, 2351 (2007). Permission from Copyright Clearance Center's Rightslink /American Chemical Society

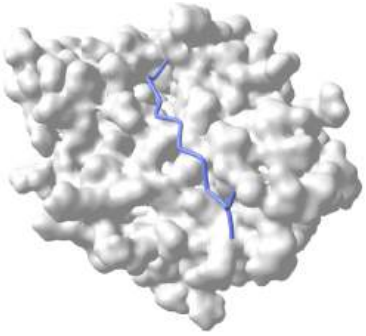
Figure 5.6.8 shows interactive iCn3D models of the types of molecular recognition features in intrinsically disordered pProteins

(A) α-MoRF, Proteinase Inhibitor IA3, bound to Proteinase A (1DP5)



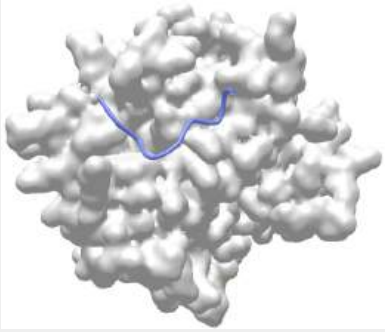
(Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...tk45aWAeMrTct7

(B) A β-MoRF, viral protein pVIc, bound to Human Adenovirus 2 Proteinase (1AVP)



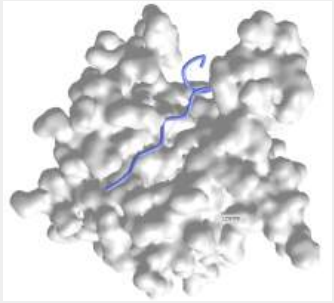
(Copyright; author via source). Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...EFmanRy2T7zjX6

(C) An ι-MoRF, Amphiphysin, bound to α-adaptin C (1KY7)



(Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...JEBC4VHUA9C718

(D) A complex-MoRF, β-amyloid precursor protein (βAPP), bound to the PTB domain of the neuron specific protein X11 (1X11)



(Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...XtQ7retidnksT9

Vacic et al have further characterized the binding surfaces between MoRFs and their binding partners using structural data from PDB files. Interfaces were studied by determining the differences in accessible surface area between MoRFs and their binding partners, and the protein in unbound states. These were compared to ordered protein complexes, including homodimers and antibody-protein antigen interactions that were not characterized by disordered interactions. Their findings are summarized below.

- MoRF interfaces have more hydrophobic groups and fewer polar groups compared to the surface of monomers. This is true even as the overall amino acid composition of intrinsically disordered proteins are enriched in polar amino acids, which leads them to adopt a variety of unfixed solution conformations.
- a-MoRFs have few prolines, which is expected as prolines are helix breakers.

- Methionine is enriched in both MoRFs and in their binding partner interface. Methionine is unbranched, flexibile, and contains sulfur, which is large and polarizable, making it an ideal side chain to be involved in London forces in a hydrophobic environment.
- Even though MoRFs have few residues, their binding interfaces were of similar or larger size than other protein binding interfaces, a result which also applies to IDPs as a whole. MoRFs interfaces also have a larger solvent-exposed surface area, similar to IDPs. This is consistent with the notion that MoRFs are disordered before binding and that a defined structure is not possible with little buried surface area.
- As MoRFs have significant nonpolar character within a IDP that is highly enriched in polar amino acids, MoRFs should be highly predictable by search algorithms.

5.7: Binding - Enzyme Linked Immunosorbant Assays (ELISAs)

5.7.1: Introduction

Enzyme-linked immunosorbent assays (ELISAs) are used widely in biotechnology, pharmaceutical and clinical medicine labs. At the same time, they appear to be underrepresented in chemistry and biochemistry curricula, even though their sensitivity, selectivity, and ease of use would argue for their widespread use

ELISAs use primary antibodies specific to a target analyte (or antigen) as a central part of the assay. One of them is Immobilized in wells of a multiwell plate or on a strip. The species not immobilized is then added, and an immobilized analyte (antigen)-antibody complex forms. After washing the bound complex on the solid phase support to remove unbound species, a second labeled antibody is added for detection. This secondary antibody binds to the distal end (Fc domain) of the bound antibody and not the analyte. The label on the second antibody can either be a fluorophore or an enzyme that can interact with an added substrate that will generate a colored solution. The color development is then measured with a fluorometer, spectrophotometer, or if just visually if quantitation is not needed.

There are several variants of ELISAs, including the traditional ELISA, in which the antigen is bound or fixed to the surface of the solid support, or a sandwich ELISA, in which the antibody is bound to the surface. In the latter case, a second labeled antibody that binds to the antigen must bind at a different site (or epitope) on the antigen. For sandwich ELISAs, the target analyte must be large enough to accommodate two antibodies binding to different sites on the same molecule.

Cartoon diagrams showing the binding interactions in traditional and sandwich ELISAs are shown in Figure 5.7.1.

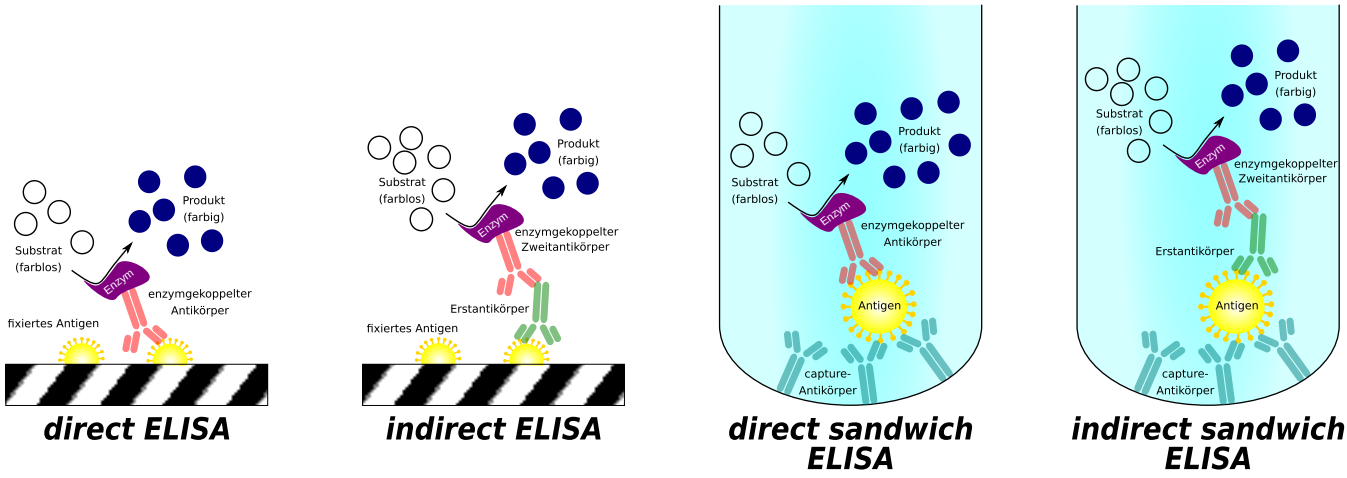


Figure 5.7.1: Binding interactions in traditional and sandwich ELISAs. Abbreviations are: fixiertes (fixed or immbolized), substrat (substrate), farblos (colorless), farbig (colored), enzymegekoppelter (enyzme coupled), antikörper (antibody), zweitantikörper (second antibody), erstantikörper (primary antibody). https://commons.wikimedia.org/wiki/File:ELISA.svg. Creative Commons Attribution 4.0 International license.

Note that both types have direct and indirect versions.

Some of the steps in a traditional ELISA are shown in Figure 5.7.2.

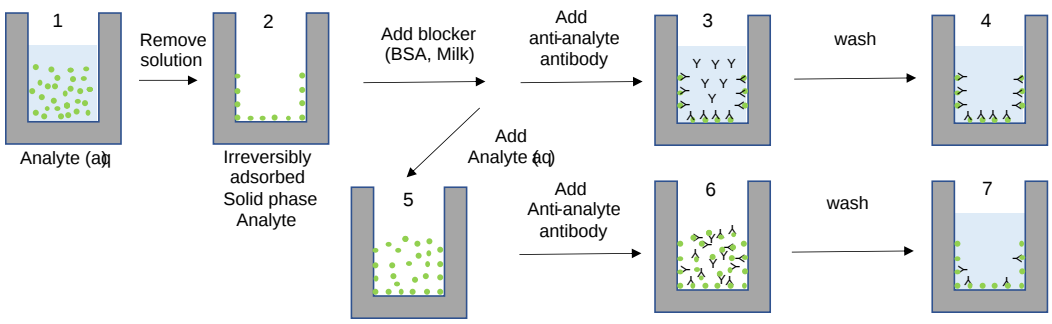


Figure 5.7.2: Steps in a traditional ELISA

In (1), an analyte is added to a well (outlined in gray). The solution is removed after a predefined time, and then a specific amount of analyte (such as protein) is irreversibly adsorbed (2). A blocker such as bovine serum albumin (BSA) or milk is added to bind any sites on the plate that could still bind protein. A primary anti-analyte protein is added. Some bind to the immobilized protein, and the rest is free in solution. After washing, only the primary antibody-immobilized analyte remains.

Now if a test solution (for example from a patient's blood) is added after blocking but before the addition of the primary antibody, it remains in the well unbound (5). If a primary antibody is added to this well, the immobilized antigen competes with the free antigen for binding, so less secondary antibody would be bound to the well (6). After washing, wells 4 and 7 remain. After additional rounds of washing and blocking, the secondary labeled antibody is added. Wells (such as 7) that contain patient analyte will bind less secondary antibody to the immobilized protein. When a substrate for the conjugated enzyme is added, the solution will be less colored after a defined incubation time.

In assays, a standard curve is made using a range of known concentrations of the analyte. The more analyte in the standard, the greater the competition with the immobilized analyte for the solution phase primary antibody, which would lead, after washing away the solution phase antigen:primary antibody complex, to a **lower** absorbance in the well with the higher solution phase antigen in the sample.

Now let's consider a sandwich assay. Instead of immobilizing a protein antigen, an antibody that binds the target antigen is immobilized. For example, an antibody against a surface SARS-Cov2 protein can be immobilized. Next, a specimen (saliva, nasal swab) containing the target surface protein recognized by the immobilized antibody is added. The great the viral load, the more SARS-Cov2 binds to the antibody. Then a second labeled antibody can be added that recognizes a different protein (the spike protein from SARS-Cov2) is added. After washing, the substrate is added, and the color development from the action of the enzyme on the substrate is measured. In this case, the more SARS-Cov2 present in the sample, the **higher** the signal (absorbance or fluorescence).

ELISAs have detection limits varying between 0.01 pg/mL to 100 ng/mL [1]. Although they are extensively used in health fields, they are not widely used in undergraduate biochemistry or chemistry courses, nor are they mentioned in the ACS's Guidelines and Supplements for either Analytical Chemistry or Biochemistry. Given their importance, we choose to discuss them in this ext.

5.7.2: ELISA Data Analysis

The most difficult parts about ELISAs are developing an understanding of the chemical and mathematical equations, choosing/using modeling and analysis software, and evaluating validity/reliability. The typical data analysis is based on the generic Hill equation instead of the classical hyperbolic binding curve analysis we derived for a single ligand to a single site on a macromolecule. The Hill equation has more empirical parameters to use when fitting bind curves.

Here is the Hill Equation equation that we studied before.

$$Y = \frac{L^n}{K_D + L^n} \tag{5.7.1}$$

For more complicated ELISA data, when a standard curve of known concentrations is used, and output signals (fluorescence, absorbance) vary from some minimum to a maximum value, the similar but more empirically useful **Logistic Equation** is used:

$$Y' = M\left(\frac{x^n}{k + x^n}\right) \tag{5.7.2}$$

where Y is the observable signal and M is the maximal observable signal. The maximal signal might not be actually observed in a ELISA assay as was the case in the binding of a ligand to a macromolecule when ligand concentrations >>K_D could not be reached.

Let's use a variant of the Hill equation using the L_{50} value, the ligand concentration at half-maximum binding.

$$\begin{aligned} 0.5 &= \frac{L_{50}^n}{K_D + L_{50}^n} \\ 1 &= \frac{2L_{50}^n}{K_D + {L_{50}}^n} \end{aligned} \tag{5.7.3}$$

hence

$$Y = \frac{L^n}{K_D + L^n} = \frac{L^n}{L_{50}^n + L^n}\left(\frac{\frac{1}{L^n}}{\frac{1}{L^n}}\right) = \frac{1}{\frac{L_{50}^n}{L^n} + 1} = \frac{1}{\left(\frac{L_{50}}{L}\right)^n + 1} \tag{5.7.4}$$

An analogous somewhat similar equation can be derived from the logistic equation:

$$Y = d + \frac{a - d}{1 + \left(\frac{L}{c}\right)^b} \tag{5.7.5}$$

where four empirical parameters define the curve:

- a is the smallest measured absorbance value (blank);
- d is the largest absorbance value when Y = 1;
- c is the inflection point in the curve which can easily be seen to occur when [L]= L50= the ligand concentration at half maximal saturation; and
- b is the slope of the curve at L50 which is the Hill coefficient n. (for many ELISA curves ≈ 1).

An idealized graph of ELISA data is shown in Figure 5.7.3.

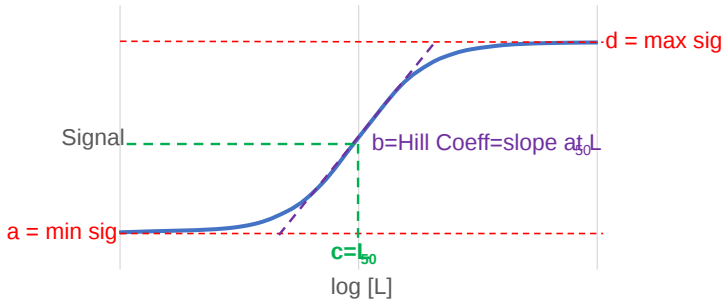


Figure 5.7.3: Idealized ELISA signal (fluorescence, absorbance) vs log [L] curve

The 4-parameter modified Logistic equation is ideal for fitting ELISA data.

An interactive active graph for the 4-parameter Logistic curve is shown in Figure 5.7.4. T

Figure 5.7.4: Interactive active graph for the 4 parameter Logistic curve.

Two of the parameters, the minimum signal **a,** and the maximal signal **d** have been set to 0.01 and 2.0 respectively, and are not changeable in the figure. Note that the greater the value of **b** (slope of the curve at the inflection point), the more "sigmoidal" the semilog curve appears (similar to the Hill binding curves for hemoglobin).

5.7.3: Lateral flow immunoassays

During the COVID pandemic, home test kits (when they were available) were used. These tests are a modified version of the sandwich ELISA described above. They differ in two main ways. The assays were not done in wells but a planar sheet in which the samples flow laterally across the sheet. As it flows across a strip through capillary action, a sample containing the SARS-COVID-2 with its spike protein would first encounter a labeled antibody to the spike protein. It would then flow into a region which contained immobilized test (anti-spike protein) and control antibodies. The bound analyte would stop and build up to sufficient concentration to see an observable colored band on the strip, but only if the sample contained the viral spike protein. These events are illustrated in Figure 5.7.5.

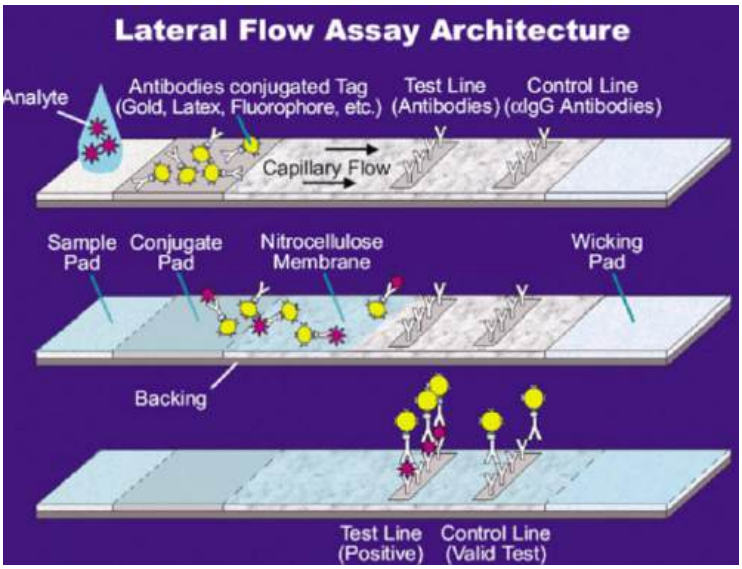


Figure 5.7.5: Lateral flow ELISA assay. https://en.Wikipedia.org/wiki/Latera...Flow_Assay.jpg

Figure 5.7.6 shows a lateral flow assay that detects the presence of serum or potentially salivary **antibodies** (either IgG or IgM) against the SARS-Cov2 proteins.

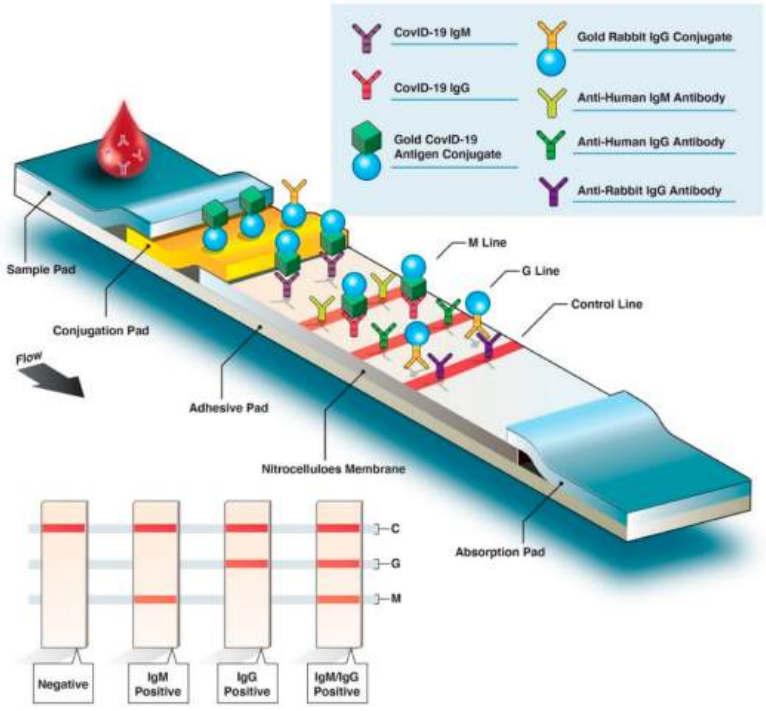


Figure 5.7.6 show a lateral flow assay which detect the presence of serum or potentially salivary **antibodies** (either IgG or IgM) against the SARS-Cov2 proteins. https://commons.wikimedia.org/wiki/F...0453-g002.webp. Creative Commons Attribution 4.0 International license.

The green cube represents the virus with viral proteins. It has been labeled with a gold nanoparticle (blue sphere). Gold is widely used as a labeling reagent in lateral flow immunoassays because it is chemically inert and hence extremely stable. The concentrated gold particles found in positive samples at the end of the strip can be observed visually since the gold nanoparticles absorb light through surface plasmon resonance. In this process, light of certain tunable wavelengths (based on the size of the nanoparticle) is absorbed when it matches the oscillatory frequency of the electron clouds of the metal nanoparticle. Plasmons are the oscillations in

the electrons (hence electromagnetic oscillation) that occur when the nanoparticle's surface interacts with photons, causing oscillations in the electrons at the same frequency as the light (resonance).

CHAPTER OVERVIEW

6: Enzyme Activity

6.1: How Enzymes Work
6.2: Kinetics without Enzymes
6.3: Kinetics with Enzymes
6.4: Enzyme Inhibition
6.5: Enzymatic Reaction Mechanisms
6.6: Enzymes and Protein Regulation
6.7: Ribozymes - RNA Enzymes
6.8: Cofactors and Catalysis - A Little Help From My Friends

6.1: How Enzymes Work

In this section, we will explore chemical and physical factors that speed up reactions, and begin to relate these effects to reactions catalyzed by enzymes. We will see that enzymes employ various chemical strategies to increase the rates of reactions, in addition to physical ones like reactant proximity and the introduction of strain. This can result in reactions that are 10 million (or more!) times faster than the uncatalyzed reaction. To put this 10 million-fold rate enhancement in perspective, if the catalyzed reaction takes one second, the uncatalyzed one would take nearly four months!

This chapter section has been written by Kristen Procko and Henry Jakubowski.

Reactions in solution that are not catalyzed are slow. Consider the hydrolysis of an ester in water, illustrated in Figure 6.1.1. The ester is stabilized by resonance, and is therefore weakly electrophilic; thus, attack by weakly nucleophilic water is a slow process. Examining the transition state, we can see that charge development and separation occurs in the transition state for the uncatalyzed reaction, resulting in an intermediate (**P**) with both positive and negative charges.

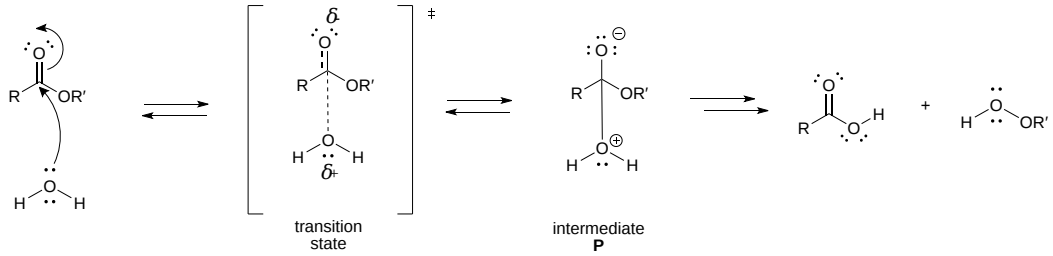


Figure 6.1.1: Charge development in the transition state during ester hydrolysis

When bonds are made or broken, charged intermediates are often formed, which are higher in energy than the reactants. Consider the energy diagram for the first step of a generic endergonic reaction, shown in 6.1.2. The transition state is closer in energy to the intermediae **P,** than it is to the reactant **R.** Therefore, the TS more closely resembles **P** than it does the starting reactants. Applying this analysis to the ester hydrolysis reaction from Figure 6.1.1, the transition state is closer in energy to the charge-separated intermediate **P**, and therefore more closely resembles the charge-separated species. In this example, the intermediate is higher in energy than the reactants, thus the transition state is *even higher* in energy than the intermediate.

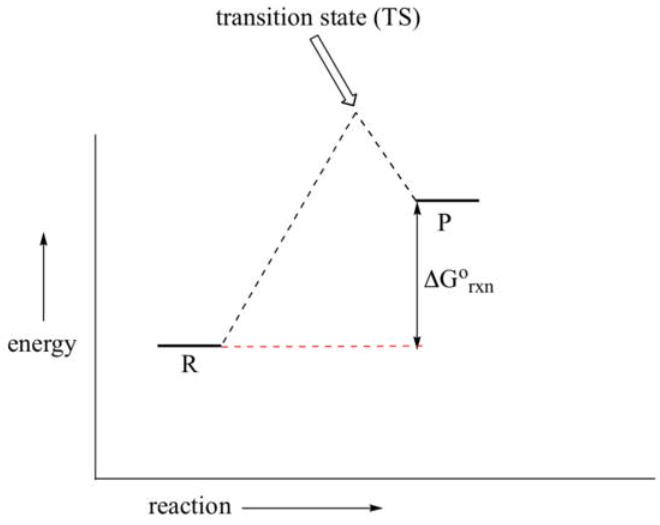


Figure 6.1.2: An energy diagram for an endergonic reaction (Image modified from: Chemistry LibreTexts, 6.9)

Anything that can stabilize the charges on the intermediate will also stabilize the developing charges in the transition state. This lowers the energy of the transition state and catalyzes the reaction. In this section, we will investigate the mechanism underlying the catalysis by small molecules of chemical reactions. Presumably, biological macromolecular catalysts (like protein enzymes) will use similar mechanisms in their catalytic effects (which will be discussed in the next section).

Catalysts, including enzymes, can employ at least 5 different ways to stabilize transition states.

6.1.1: Chemical Strategies for Rate Enhancement

6.1.1.1: General Acid and Base Catalysis

Considering intermediate **P** in Figure 6.1.1, we can envision two strategies to reduce the charge separation: the negative charge on the anionic oxygen could be protonated, or the positive charge on the cationic oxygen could be removed by deprotonation. If the reaction is pH dependent, and the reaction rate solely depends on hydronium ion concentration, [H_3O^+], then **specific acid catalysis** is operative. **Specific acid catalysis** occurs when the hydronium ion concentration is the sole factor determining the reaction rate, and the rate is not influenced by the concentration of any buffer components present in the solution. In other words, the rate of the reaction depends *specifically* on the concentration of hydronium ion. **Specific base catalysis** occurs when the reaction rate is dependent solely on the concentration of hydroxide ion, and is again independent of any buffer components in solution.

By contrast, **general acid catalysis** occurs when the reaction is not solely dependent on the [H_3O^+] concentration; that is, the reaction rate is influenced by the concentration of a buffer component. With general acid catalysis, the charge separation in the transition state is decreased by donation of a proton to the carbonyl from general acids (e.g., acetic acid or a protonated imidazole ring). Proton donation decreases the developing negative in the transition state. In Figure 6.1.3, the first step of the ester hydrolysis mechanism is shown via specific acid catalysis, alongside the general acid catalysis mechanism using the weak acid, acetic acid.

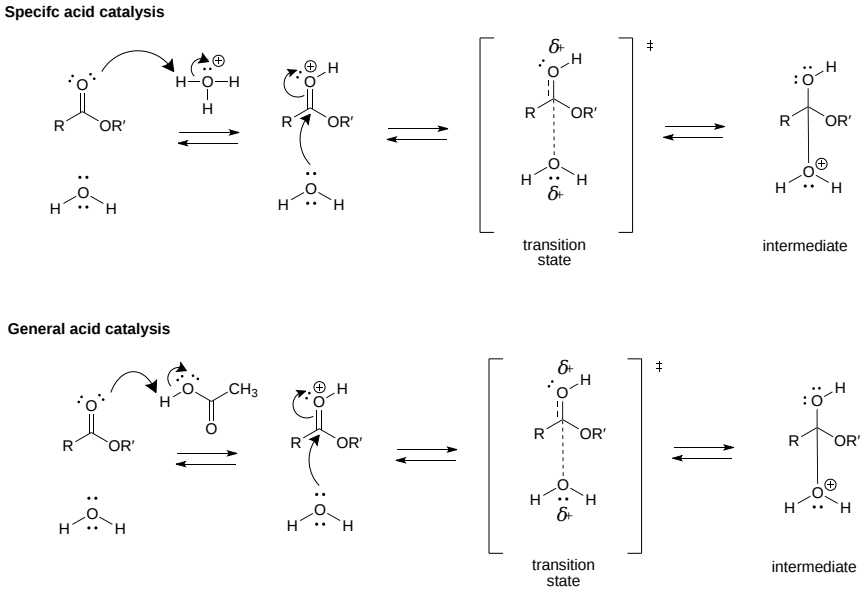


Figure 6.1.3: Specific vs. general acid catalysis

Alternatively, the first step of the ester hydrolysis mechanism can be base-catalyzed, which increases the strength of the nucleophile. In Figure 6.1.1, the attacking water molecule develops a partial positive charge in the transition state as it begins to form a bond to the electrophilic carbon of the carbonyl. In the base-catalyzed mechanism shown in Figure 6.1.4, hydroxide becomes the nucleophile in the specific base-catalyzed mechanism. The energy of the transition state can also be lowered by the presence of a general base (e.g., acetate, a deprotonated imidazole ring). Proton abstraction decreases the developing positive charge.

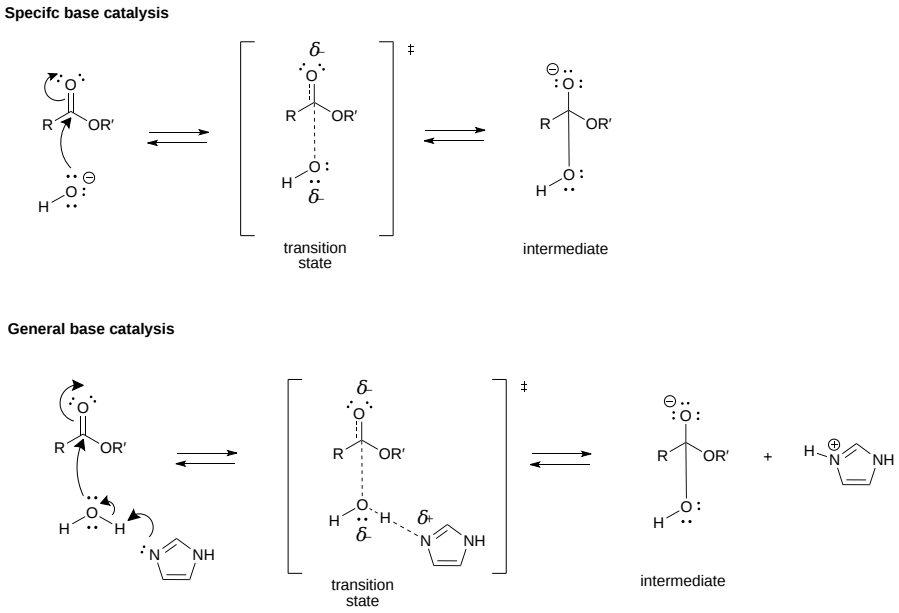


Figure 6.1.4: Specific vs. general base catalysis

General acid/base catalysis is common with enzymes because enzymes often use amino acid side chains to promote acid-base reactions within the **active site**, the region of the enzyme where the chemical reaction takes place. Acetic acid is similar to glutamic and aspartic acid side chains, and the imidazole ring shown in the general base catalysis reaction in Figure 6.1.4 is present in the side chain of the amino acid histidine.

6.1.1.2: Metal Ion or Electrostatic Catalysis

A metal such as Cu^{2+} or Zn^{2+} can also stabilize the transition state. The metal must be able to bind the charged intermediate and hence the transition state. An oxyanion intermediate formed during the reaction of an electrophilic carbonyl C can interact with a metal, especially when there is an O on an adjacent atom which can help coordinate the metal ion. This charge stabilization of the developing negative in the transition state and the full negative in the intermediate is often called **electrostatic catalysis**, and is illustrated in the decarboxylation of a β-keto carboxylic acid Figure 6.1.5. Coordination of Cu^{2+} to the β-keto carboxylic acid increases the electrophilicity of the carbonyl, making it a superior electron acceptor, which facilitates decarboxylation. The intermediate enolate formed in the decarboxylation is protonated, giving the more stable ketone as the product.

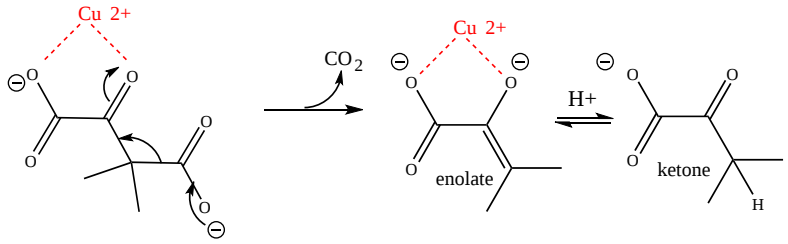


Figure 6.1.5: Metal ion catalysis

Electrostatic catalysis is likely to occur in many enzymes since nearly 1/3 of all enzymes require metal ions. A classic example of an enzyme using metal ion catalysis is carboxypeptidase A. Figure 6.1.6s shows an interactive iCn3D model of Zn and the inhibitor citric acid bound to carboxypeptidase A (3KGQ). Note the histidine and aspartate amino acid side chains of the active site coordinating to the Zn^{2+} ion, along with the carboxylate group of citrate.

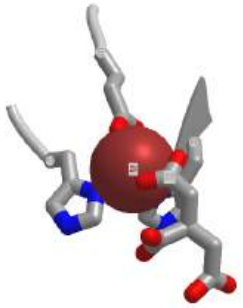


Figure 6.1.6: Zn binding to carboxypeptidase A (3KGQ) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?pH7tydTANZw5ayh9A

Metals can also act in a different way. They may coordinate a water and by further polarizing the H-O bond increase the acidity of the bound water. For instance, a water molecule in the hexaaquairon(III) ion has a pKa of 9.4, compared to pure water, with a pK_a of 14 (6.1.7). The complexed hydroxide is a better nucleophile than bulk water.

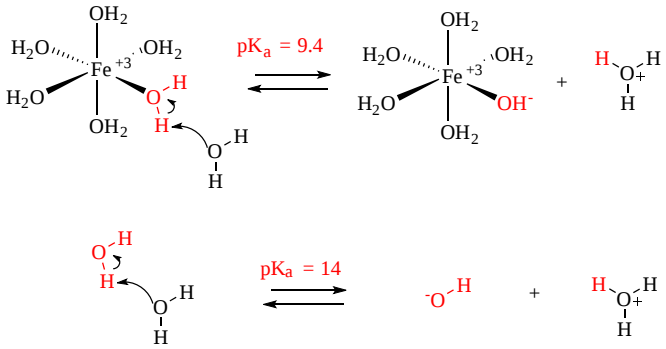


Figure 6.1.7: Metal ion decrease of pK_a of coordinated water

Another enzyme that utilizes Zn^{2+} is carbonic anhydrase. It is among the fastest of all enzymes, with a k_{cat} of 10^6 s^{-1} and a k_{cat}/K_m of 8.3 x 10^7 $M^{-1}s^{-1}$ (reference). It is diffusion controlled at low substrate (CO_2) concentration and converts one million bound CO_2 per second to HCO_3^-! The Zn^{2+} appears to bind a water molecule and reduce its pK_a such that the bound form is OH^-. This is illustrated in Figure 6.1.8, which depicts the local environment of the bound Zn^{2+} (coordinated by 3 histidine side chains and an OH^-) in the absence (left) and presence (right) of CO_2.

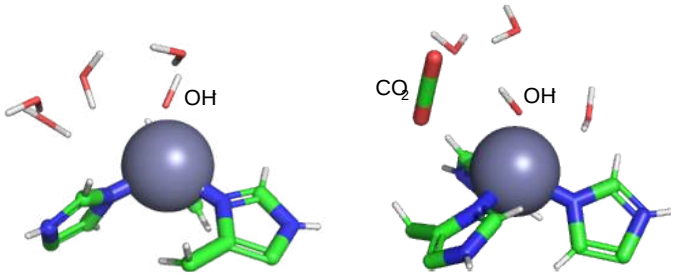


Figure 6.1.8: Coordination of OH^- to Zn^{2+} in carbonic anhydrase in the absence (left) and presence (right) of substrate CO_2.

6.1.1.3: Covalent or Nucleophilic Catalysis

One way to change the activation energy of the reaction is to change the reaction mechanism in ways which introduces new steps with lower activation energy. As shown in Figure 6.1.9, the catalyzed reaction has a new lower energy well, representing formation of the covalent intermediate, and the activation energy is lowered overall. Formation of the new intermediate results in two transition states, represented by the two high energy points of the blue line in the plot.

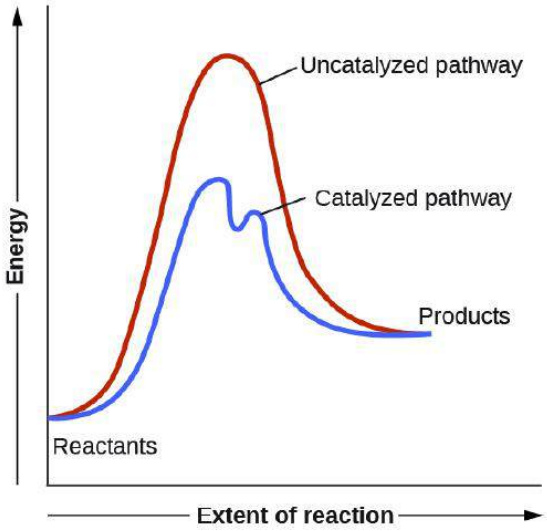


Figure 6.1.9: Energy diagram for an uncatalyzed reaction compared to a catalyzed reaction that utilizes covalent catalysis.

A typical way to achieve covalent catalysis is to add a nucleophilic catalyst, which forms a covalent intermediate with the reactant. Figure 6.1.8 shows how pyridine (red) acts as a nucleophilic or covalent catalyst in the hydrolysis of an anhydride. The anhydride is very reactive to start with, and the charged pyridinium ion intermediate contains a very good leaving group. The desired nucleophile, water, can then interact with the intermediate in a nucleophilic substitution reaction. In these types of reactions in general, as long as the nucleophilic catalyst is a better nucleophile than the ultimate nucleophile (usually water) then the activation energy is lowered, and the reaction is catalyzed. The nucleophilic catalyst and the original nucleophile usually interact with a carbonyl C in a substitution reaction.

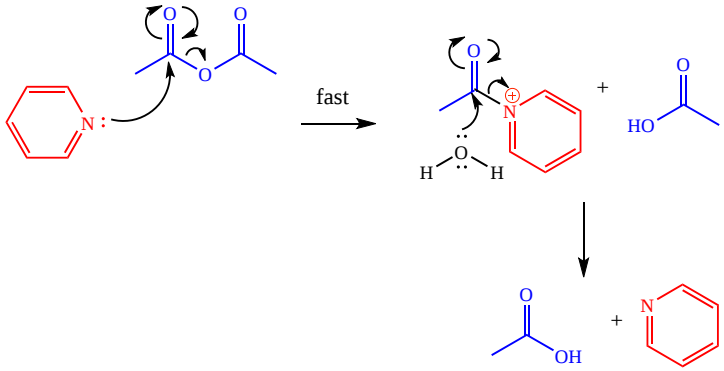


Figure 6.1.10: Nucleophilic covalent catalysis by pyridine

6.1.1.3.1: Reactions involving iminium ions are a recurring theme in biochemistry

Positively-charged nitrogen cations (iminium ions) form as intermediates in many biochemical mechanisms. The iminium ion is a powerful electron acceptor, and can promote the cleavage of bonds that would otherwise be difficult to break, such as C–C and C–H bonds. To begin our analysis of how iminium ions promote such cleavage reactions, let's revisit a common carbon-carbon bond cleaving reaction, the decarboxylation of a β-keto acid, which we examined briefly above in Figure 6.1.5 with metal ion catalysis.

Under acidic conditions, β-keto acids usually decarboxylate with gentle warming. A cyclic transition state is often invoked, and the presence of the carbonyl of the ketone, adjacent to the breaking bond, gives the electrons somewhere to go (Figure 6.1.11). The product of the decarboxylation is an enol, which tautomerizes to the more stable ketone. Under the slightly basic conditions that

characterize the medium for most biochemical reactions, the equilibrium favors the deprotonated carboxylate form. The adjacent carbonyl again gives the electrons somewhere to go in the decarboxylation reaction, and under these basic conditions, an enolate is formed. Protonation gives the final product, a ketone.

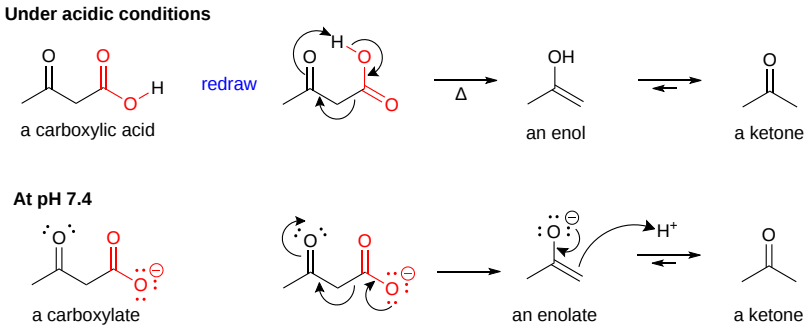


Figure 6.1.11: Decarboxylation of a β-keto acid under acidic conditions, and at pH 7.4

In Figure 6.1.5, we saw that a metal ion can promote the decarboxylation reaction by interacting with the electron-accepting ketone carbonyl, which makes it even more electrophilic. Another strategy to create a better electron acceptor involves forming a full positive charge on the electron accepting atom, which can be done by converting the ketone to an iminium ion. Amines react with aldehydes or ketones to form iminium ions. Figure 6.1.12 illustrates this strategy, which involves covalent catalysis. The amine, RNH_2, reacts to form a new intermediate, the iminium ion, with a full positive charge. The protonated nitrogen serves as an excellent electron "sink" for decarboxylation reactions of beta-keto acids.

This **iminium ion** or protonated **Schiff Base** has a pK_a of about 7, so the protonated iminium and deprotonated imine are in equilibrium near pH 7. Figure 6.1.12 also illustrates a simple way to view reaction mechanisms. Electrons in chemical reactions can be viewed as flowing from a source (such as a carboxyl group) to a sink (such as an electrophilic carbonyl O or a positively charged N in a Schiff base).

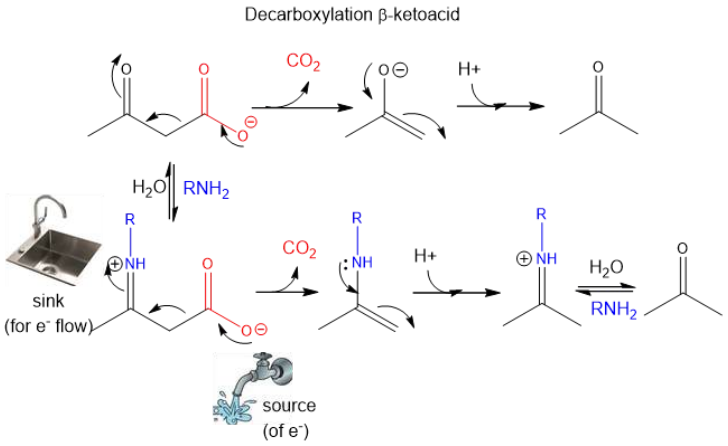


Figure 6.1.12: Schiff-base catalyzed decarboxylation of a β-ketoacid - Source/Sink analogy

Acid- and base-catalyzed reaction mechanisms for Schiff base formation are shown in Figure 6.1.13. An amine is used as the nucleophilic catalyst, forming the initial addition product, a carbinolamine. The carbinolamine dehydrates, since the free pair of electrons on the N are more likely to be shared with the carbon to form a double bond than electrons from the original carbonyl O, which is more electronegative than the N. If catalyzed by a general acid, an iminium ion, and the base-catalyzed reaction forms an imine. Near pH 7.4, the imine easily protonated to form a positively charged N at the former carbonyl O center.

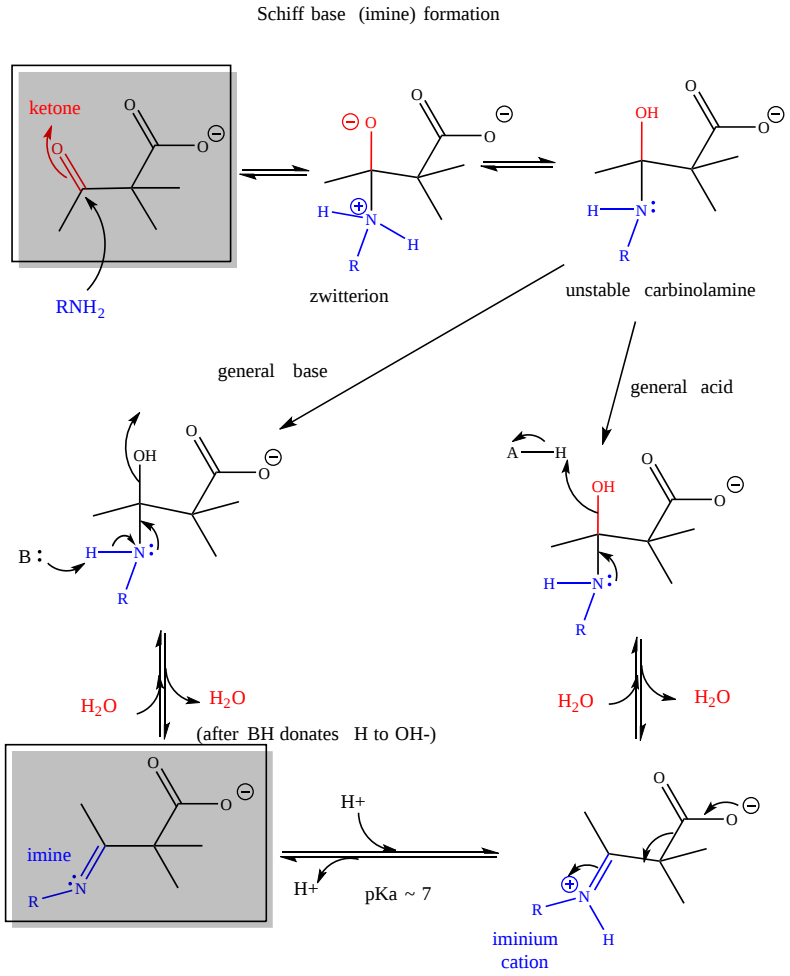


Figure 6.1.13: Schiff Base Formation - reaction mechanism

An actual Schiff base intermediate between fructose-1,6-bisphosphate (2FP-400) and Lys 239 from the enzyme fructose bisphosphate aldolase from Leishmania mexicana is shown in Figure 6.1.14 Only a single bond between the carbon and nitrogen in the Schiff base is shown.

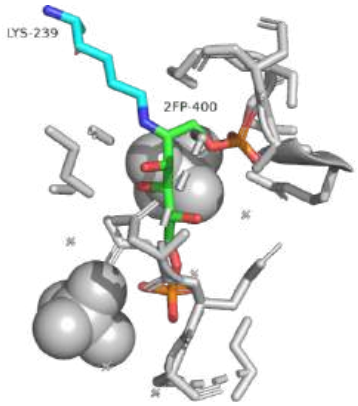


Figure 6.1.14: Schiff base intermediate between fructose-1,6-bisphosphate (2FP-400) and Lys 239 from the enzyme fructose bisphosphate aldolase from Leishmania mexicana(2QDG)

6.1.1.4: Transition State Stabilization

In the middle of the 20th century, Linus Pauling postulated that the only thing that a catalyst must do is bind the transition state more tightly than the substrate. This can be discerned in Figure 6.1.15 and a little math. The diagram shows how the substrate (S) and the transition state (S*) can react with an enzyme (E) to form a complex which then proceeds to product (following the diagram from "start here" in the right-hand direction), or can go to product in the absence of enzyme (E) (following the diagram from "start here" in the left-hand direction). Note that the diagram is arbitrarily drawn such that the standard Gibbs free energy (G°) of the free product P is higher than that of the free substrate, S.

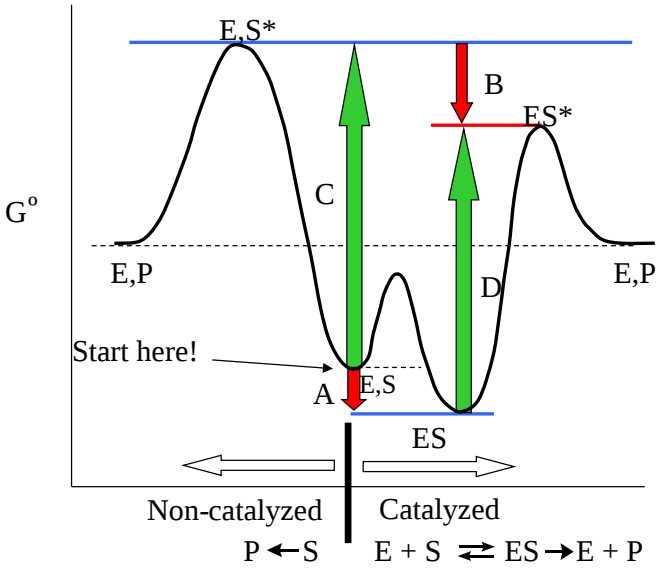


Figure 6.1.15: Enzymes bind the transition state more tightly than the substrate

The colored large vertical arrows represent the ΔG° for the transition shown:

- The red arrows A and B represent the ΔG°s for the binding of E + S (arrow A) and E + S* (arrow B), respectively.
- The green arrows C and D represent the ΔG°s for the activation energy of the free substrate (arrow C) and the enzyme-bound substrate (arrow D).

Now consider the two pairs of arrows, A,C and D,B and add each set like a vector in elementary physics. Since the distance between the two horizontal blue lines is the same for the left hand process (uncatalyzed) and the right hand one (catalyzed), it follows that

$$C - A = D - B \tag{6.1.1}$$

The negative signs for A and B are used since in the diagram both A and B have negative ΔG° values.

Now for an enzyme to be a catalyst, the activation energy D for the reaction in the presence of the enzyme E must be less positive (i.e., smaller) than the activation energy C in the absence of enzyme. Therefore, after rearranging the equation:

$$C - D = A - B > 0 \tag{6.1.2}$$

and substituting in the ΔG° values for A and B, we can directly compare the free energy of binding the substrate (S) vs. binding the transition state (S*):

$$-RT \ln \mathrm{K_{eqS}} - \left(-\mathrm{RT} \ln \mathrm{K_{eqS^*}}\right) > 0 \tag{6.1.3}$$

Hence, the equilibrium constant for binding the transition state is larger than that for binding free substrate:

$$\mathrm{K_{eq}}\, S^* > \mathrm{K_{eq\,S}} \tag{6.1.4}$$

Pauling was right. The enzyme just needs to bind the transition state more tightly than the substrate to catalyze the reaction. This is why chemists synthesize stable transition state analogs as potential tight binding inhibitors of target proteins.

The stability of the transition state also affects the reaction kinetics (which makes sense given that the activation energy clearly affects the speed of a reaction). As you probably remember from organic chemistry, biomolecular nucleophilic substitution (S_N2) reactions are slow when the central atom where the substitution will occur is surrounded by bulky substituents (sterics once again). We discussed this in context to nucleophilic substitution on a sp^2 hybridized carbonyl carbon in carboxylic acid derivatives versus on a sp^3 hybridized phosphorous in phosphoesters and diesters. The explanation for this phenomena has usually been attributed to hindered access of the central atom caused by bulky substituents (intrinsic effects). Is this true? Studies on S_N2 reactions of methylchloroacetonitrile and t-butylchloroacetonitrile (with the reagent labeled with ^{35}Cl) using $^{37}Cl^-$ as the incoming nucleophile in the gas phase shown that the more hindered t-butyl derivative's activation energy was only 1.6 kcal/mol (6.7 kJ/mol) higher than the methyl derivative, but in aqueous solution, the difference is much greater for comparable reactions (Figure 6.1.16).

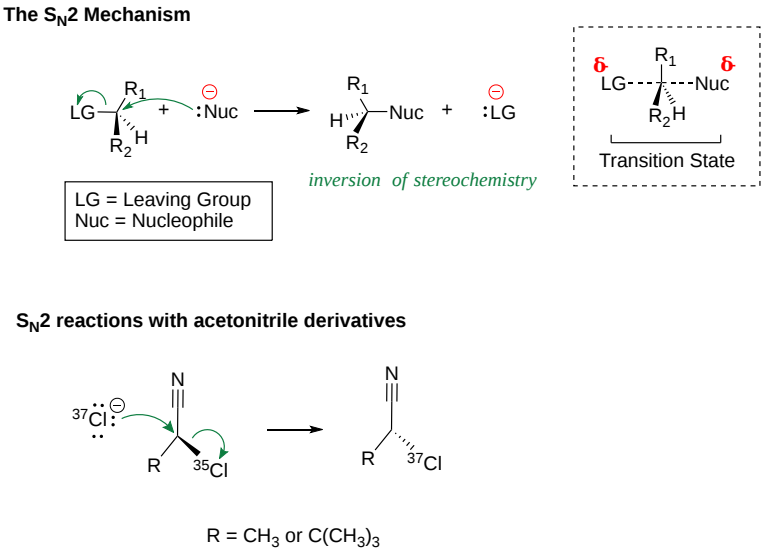


Figure 6.1.16: S_N2 reactions are characterized by a pentavalent transition state

They attributed the differences to solvation effects of the transition state. The bulkier the substituents on the central atom, the more difficult it is to solvate the transition state since water can't reorient around it as well. In effect there is steric hindrance for both reactant and solvent.

What does it take for a macromolecule (M) to be a catalyst - an enzyme? It seems the minimum criterion are:

- M binds a reactant
- M binds the transition state more tightly than the substrate

Anything above these is just "icing on the cake". If different functional group are present in the "active" site of the enzyme that would allow electrostatic, intramolecular, covalent, general acid and/or base catalysis, the better the catalyst.

6.1.1.4.1: A transition state analog case study: Abyzmes (Antibody Catalysis)

Recall that antibodies are immune system proteins that bind foreign molecules (see Chapter 5.4). The usual role of an antibody is to initiate an immune response. When the antigen-binding site, located in the variable region of an antibody, binds to an antigen, it triggers the formation of new antibodies (within B-cells) in an effort to optimize the immune response to that antigen. These new antibodies are made with mutations in the antigen binding region, and those that bind better than the original antibody will be selected to form longer lived memory B-cells, ready for the next time the body encounters the antigen.

Decades ago, Linus Pauling made a hypothesis that antibodies could be produced with an atypical role—to act as catalysts! If antibody molecules could be made to bind to a compound resembling the transition state of a chemical reaction, they should also presumably catalyze the chemical reaction. In 1987, his prediction was verified. Lerner et al. made a transition state analog of an ester. When an ester is hydrolyzed, as shown in Figure 6.1.17, the sp^2 hybridized carbonyl carbon is converted to an sp^3 hybridized center in the intermediate, with the carbonyl oxygen becoming an oxyanion.

The transition state presumably looks more like this unstable intermediate (sp^3, oxyanion). Thus, Lerner synthesized a phosphonate, an ester mimic with a sp^3 hybridized phosphorous replacing the carbonyl C. It also has a negatively charged oxygen as does the intermediate for the ester. This phosphonate ester is very resistant to hydrolysis. When injected into a mouse (after first being covalently attached to a carrier protein so the small molecule becomes "immunogenic"), the mouse makes a protein antibody which binds to the phosphonate. When the corresponding carboxylic acid ester is added to the antibody, it is cleaved with nominal k_{cat} and K_M values. Site specific mutagenesis can then be done to make it an even better catalyst! The antibody enzymes have been called **abzymes**. The structure shown in Figure 6.1.17 shows how phosphonamides act as transition state analogs as well.

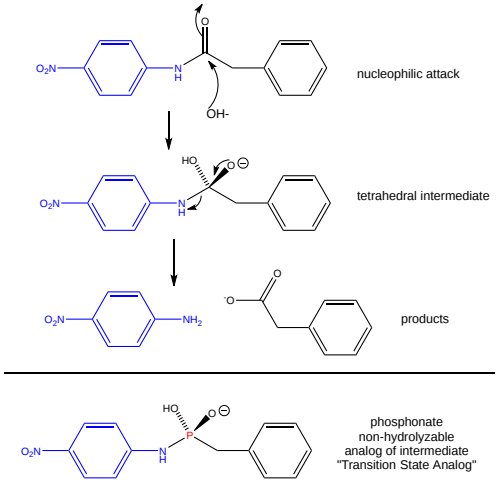


Figure 6.1.17: PHOSPHONAMIDES: TRANSITION STATE ANALOGS

Figure 6.1.18 shows an interactive iCn3D model of transition state analog 5-(para-nitrophenyl phosphonate)-pentanoic acid bound to a mouse Fab antibody fragment with esterase activity (1aj7)

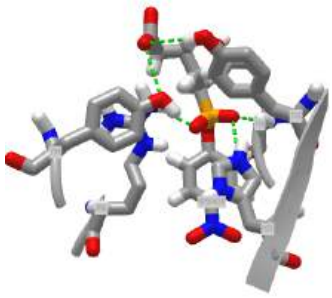


Figure 6.1.18: Transition state analog 5-(Para-Nitrophenyl Phosphonate)-Pentanoic Acid (NPE406) bound to Fab antibody fragment with esterase activity (1aj7) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?zSWAhDag8t4LpJaN7

Transition state theory can be used to more clearly quantify the relationships described in the graphical analysis above. This analysis will use the equilibrium constant (in contrast to the last two chapters which used dissociation constants to characterize macromolecule, receptor, and enzyme binding to ligand). Let assume that a substrate S is in equilibrium with its transition state $S^{\ddagger}$. Hence Keq = $[S^{\ddagger}]/[S]$. The following reaction can be written: S → $S^{\ddagger}$ → P. Based on our previous kinetic analysis and experience in writing differential equations, dP/dt = k1$[S^{\ddagger}]$. By analogy, enzyme bound S (ES) can be converted to ($ES^{\ddagger}$) and then on to product as shown in the following chemical equation:

$$\mathrm{E + S \rightleftharpoons ES \longrightarrow ES^{\dagger} \longrightarrow E + P.}$$

For the non-enzyme catalyzed reaction, transition state theory can be used to show that the first order rate constant $k_1 = kT/h$ where k is the Boltzman's constant, T is the Kelvin temperature, and h is Planck's constant. Hence, using $K_{eq} = [S^{\dagger}]/[S]$, equation 1 can be derived

$$\frac{dP}{dt} = \frac{kT}{h}\left[S^{\dagger}\right] = \frac{kT}{h}K^{\dagger}[S] = k_n[S] \tag{6.1.5}$$

where k_n (hereafter written as k_N) $=(kT/h)K^{\dagger}$ is the effective first order rate for the non-catalyzed rate. Now let's create a more complicated linked equilibrium showing the same reaction in the presence of an enzyme. Figure 6.1.19

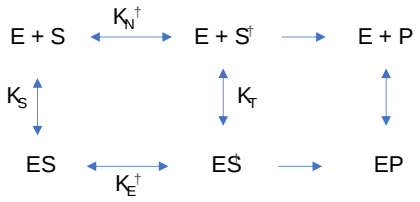


Figure 6.1.19: Linked equilibrium for conversion of free $S^{\dagger}$ and $ES^{\dagger}$ to product

Remember that the K values for this analysis are equilibrium constants not dissociation constants. Note two important equilibrium constants, K_S, the equilibrium constant for the binding of free S to E, and K_T, the equilibrium constant for the binding of free $S^{\dagger}$ to E (assuming that free $S^{\dagger}$ could bind to E before it converted to product). As we have seen for linked equilibrium before, since the K_{eq} values are related to the standard free energy changes which are state functions, the sum of the standard free energies going from E + S to $ES^{\dagger}$ (by either the top or bottom paths) are path independent so the products of the K_{eq} for the top path are equal to those for the bottom paths. This gives the following equation:

$$\frac{K_T}{K_S} = \frac{K_{E^{\dagger}}}{K_{N^{\dagger}}} = \frac{k_E}{k_N} \tag{6.1.6}$$

The right hand side is the ratio of the effective first order rate constant for conversion or $ES^{\dagger} \rightarrow E + P$, k_E divided by the rate constant for the conversion of $S^{\dagger} \rightarrow P$ for the noncatalyzed rate, k_N. The final ratio of rate constants can be derived from the simple relationship that $k_x=(kT/h)K^{\dagger}_x$ where x is either N (non catalyzed) or E (enzyme catalyzed). Equation 2 states that the equilibrium constant for the binding of $S^{\dagger}$ to E, K_T, is greater than the equilibrium constant for the binding of S to E, K_S (as $k_E > k_N$). K_T/K_R ranges from 10^8 - 10^{14}. Given common values for the equilibrium constant for binding of S to E (10^3 – 10^5 M^{-1}) which is equivalent to dissociation constant values K_d = 10 uM – 1 mM, the calculated value of $K_T = 10^{15}$ M^{-1} which gives a dissociation constant for the enzyme and transition state of $K_d = 10^{-15}$ M (1 femtomolar). This is as tight as one of the highest affinity binding interactions in the biological world, the binding of avidin and biotin. As we noted in Chapter 5.1, assuming that the second order rate constant for avidin/biotin binding and as shown above for $E/S^{\dagger}$ is diffusion controlled (about 10^8 $M^{-1}s^{-1}$), the off rate for the avidin-biotin or $ES^{\dagger}$ complex is 10^{-7} s^{-1}, equivalent to a half life of the complex of 80 days. It doesn't get much tighter than that.

Figure 6.1.20 represent an image of an enzyme and three different molecules, 1–3, that could bind to it. Using the analysis above, which molecule do you think represents substrate? Transition state? Product?

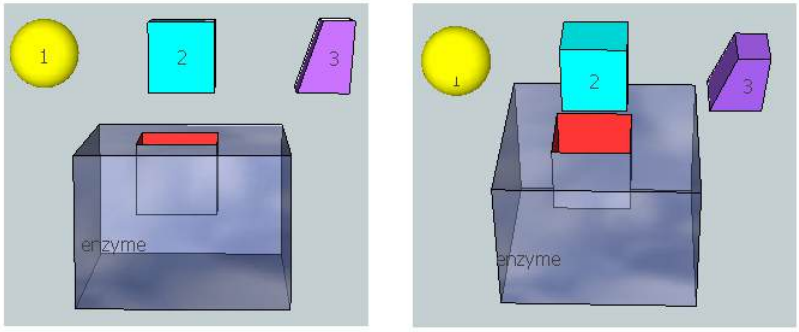


Figure 6.1.20: 3D model for binding substrate, transition state and product to an enzyme

6.1.2: Physical Strategies for Rate Enhancement

6.1.2.1: Intramolecular Catalysis

Consider the hydrolysis of phenylacetate. This reaction, a nucleophilic substitution reaction, could be catalyzed by the addition of the general base acetate to the solution, as described above. Since this reaction rate would double with the doubling of the solution acetate, the reaction is bimolecular (first order in reactant and catalyst). Now consider the same reaction only when the the general base part of the catalyst, the carboxyl group, is part of the reactant phenylacetate. Such a case occurs in the acetylated form of salicylic acid—i.e., aspirin. When the carboxy group is ortho compared to the acetylated phenolic OH, it is in perfect position to

accept a proton from water, decreasing the charge development on the O in the transition state. The general base does not have to diffuse to the appropriate site when it is intramolecular with respect to the carbonyl C of the ester link. The rate of this intramolecular base catalysis is about 100 fold greater than of an intermolecular base catalyst like acetate. It is as if the effective concentration of the intramolecular carboxyl base catalyst is much higher due to its proximity to the reaction site.

Another type of reactions involving a carboxyl group (in addition to simple proton transfer) is when the negatively charged carboxyl O acts as a nucleophile and attacks an electrophilic carbonyl carbon. When the carbonyl is part of an ester, the carboxyl group engages in a nucleophilic substitution reaction, expelling the alcohol part of the ester as a leaving group. The remaining examples below consider the nucleophilic (carboxyl) substitution on phenylesters, with phenolate as the leaving group. The reactions in effect transfer an acyl group to the carboxyl group to create an anhydride.

First consider acyl transfer with aspirin derivatives. Aspirin, as you know, contains a carboxyl group ortho to an ester substituent. Hence the carboxyl group can act as a nucleophile and attack the carbonyl carbon of the ester in a nucleophilic substitution reaction. The net effect is to transfer the acetyl group from the phenolic OH to the carboxyl group converting it to an anhydride. This is an intramolecular reaction. Compare this reaction to a a comparable bimolecular reaction shown in Figure 6.1.21.

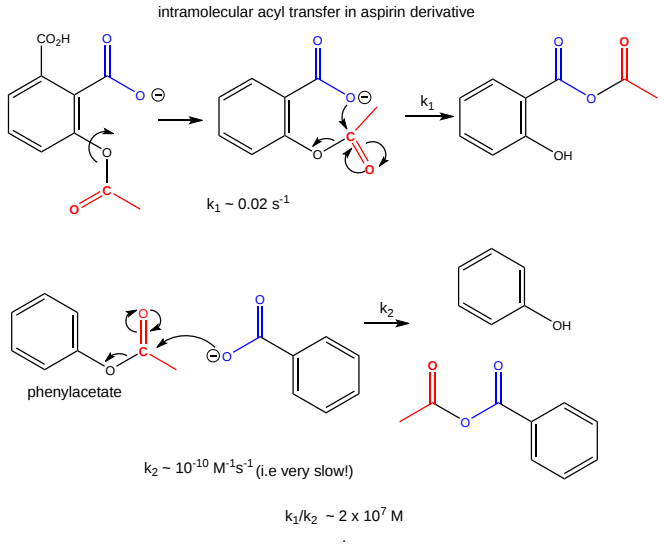


Figure 6.1.21: Acyl transfer in aspirin derivatives.

The first order rate constant of the intramolecular transfer of the acetyl group to the carboxyl group, $k_1 = 0.02\ s^{-1}$. The analogous bimolecular reaction rate constant $k_2 \sim 10^{-10}\ M^{-1}s^{-1}$. Dividing k_1/k_2 gives the relative rate enhancement of the intramolecular over the intermolecular reaction. With units of molarity, this ratio can be interpreted as the relative effective concentration of the intramolecular nucleophile. This makes the effective concentration of the carboxylate in the aspirin derivative 2×10^7 M.

Now consider the cleavage of phenylacetate using acetate as the nucleophile as shown in Figure 6.1.22 The products are acetic anhydride and phenolate. This is a bimolecular reaction (a slow one at that), with a bimolecular rate constant, k_2 which I will arbitrarily set to 1 for comparison to some similar reactions.

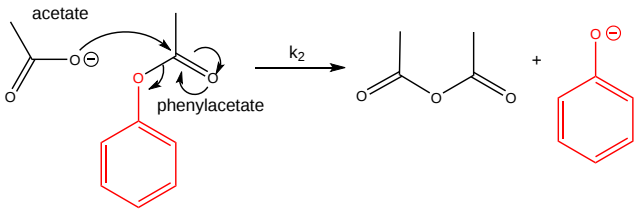


Figure 6.1.22: Reaction of acetate with phenylacetate

Now consider a monoester derivatives of succinic acid - phenyl succinate - in which the free carboxyl group of the ester attacks the carbonyl carbon of the ester derivative, as shown in Figure 6.1.23

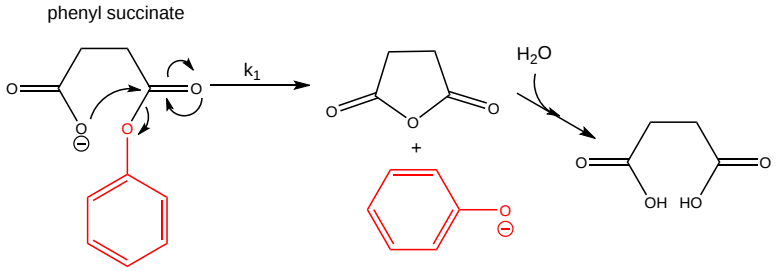


Figure 6.1.23: Intramolecular reaction of phenylsuccinate

If you assign a second order rate constant k_2 = 1 $M^{-1}s^{-1}$ to the analogous intermolecular reaction of acetate with phenylacetate (as described above), the first order rate constant for the intramolecular reaction of phenylsuccinate is 10^5 s^{-1}. The ratio of rate constants, k_1/k_2 = 10^5 M. That is it would take 10^5 M concentration of acetate reacting with 1 M phenylacetate in the first bimolecular reaction to get a reaction as fast as the intramolecular reaction of phenylsuccinate. The intramolecular reaction of an even more sterically restricted bicyclic phenylcarboxylate shown in Figure 6.1.24 has a k_1/k_2 = 10^8 M.

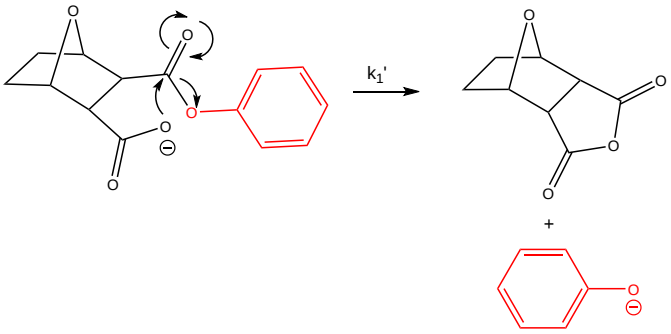


Figure 6.1.24: Intramolecular reaction of bicyclic phenylcarboxylate

Another example is anhydride formation between two carboxyl groups. The ΔG^o for such a reaction is positive, suggesting an unfavorable reaction. Consider two acetic acid molecules condensing to form acetic anhydride. For this intermolecular reaction, K_{eq} = $3x10^{-12}$ M^{-1}. Now consider the analogous intramolecular reaction of the dicarboxylic acid succinic acid. It condenses in an intramolecular reaction to form succinic anhydride with a K_{eq} = $8x10^{-7}$ (no units). The ratio $K_{eq\text{-}intra}/K_{eq\ inter}$ = 3 $x10^5$ M. It is as if the effective concentration of the reacting groups because they do not have to diffuse together to react, is 3 $x10^5$ M.

How does this apply to enzyme catalyzed reaction? Enzymes bind substrates in physical steps which are typically fast. The slow step is often the chemical conversion of the bound substrate, which is effectively intramolecular if the initial binding reaction is fast. These three kinds of reactions, intermolecular, intramolecular, and enzyme-catalyzed can be broken down into two hypothetical steps, a binding followed by catalysis as shown in Figure 6.1.25.

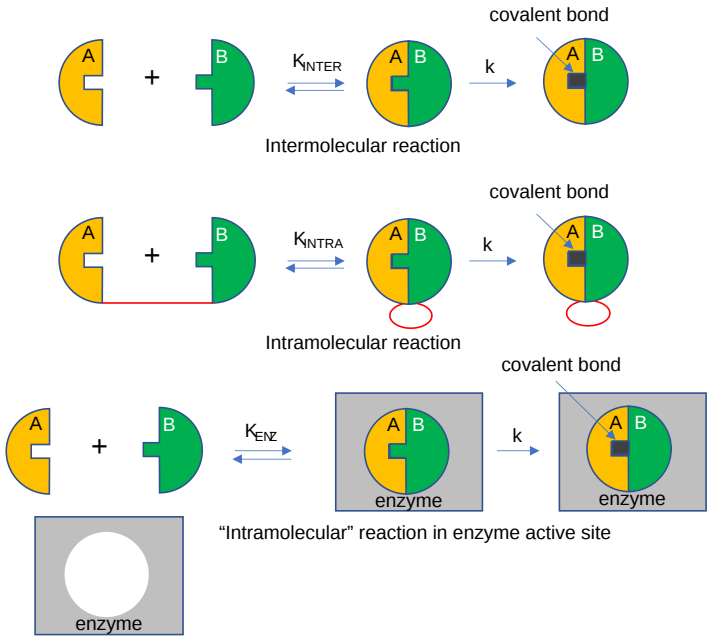


Figure 6.1.25: Intermolecular, intramolecular, and enzyme-bound reactions

If the rate constants for the chemical steps are all identical, the advantage of the intramolecular and enzyme-catalyzed reaction over the intermolecular reaction is K_{INTRA}/K_{INTER} and K_{ENZ}/K_{INTER}, respectively.

The advantage of intramolecular reactions can be seen by studying the Ca-EDTA complex. Calcium in solution exists as a octahedrally coordinated complex with water occupying all the coordination sites. EDTA, a multidentate ligand, first interacts through one of its potential six electron donors to Ca in a reaction which is entropically disfavored from the the Ca-EDTA perspective, although one water is released. Once this first intramolecular complex is formed, the rest of the ligands on the EDTA rapidly coordinate with the Ca and release bound water as illustrate in Figure 6.1.26. The former is no longer entropically disfavored since it is now an intramolecular process while the later is favored through the release of the remaining five water molecules.

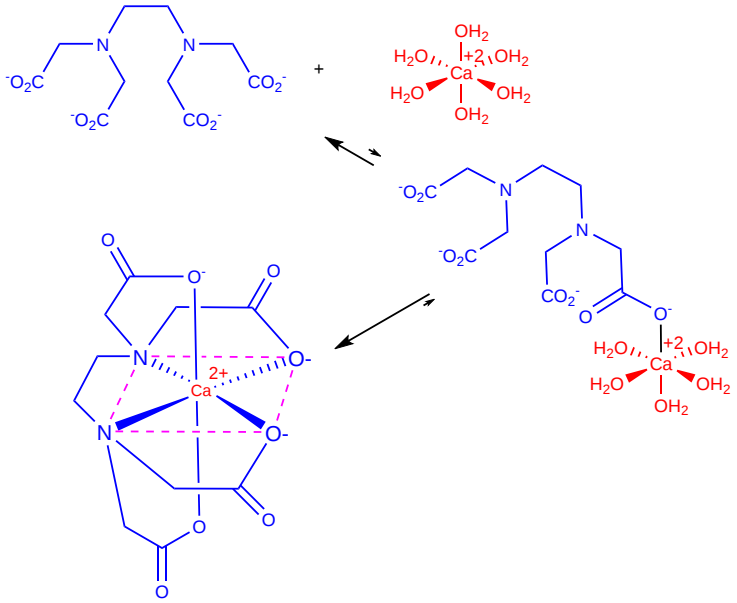


Figure 6.1.26: Binding of Ca^{2+} and EDTA

We've shown above the catalytic advantage offered by intramolecular reaction in terms of a dramatic increase in the effective concentration of reactants, which sometimes reached levels of 10^8 M. Another way is to look at entropy changes associated with dimer formation. The table below shows that an intramolecular reaction is favored over an intermolecular reaction since in the latter, significant decreases in translational and rotation entropy result.

Translational, Rotational, and Internal Entropies for Dimer Formation: A + B ↔ A-B (cal/K·mol)

System	A	B	A-B	ΔS
		Gas		
S trans	30	30	30	-30
S rot	20	20	20	-20
S int	5	5	20	+10
Gas → Solution	-10	-10	-15	
S sol	45	46	55	-35 (Correspond to 10^8-10^9 M)

6.1.2.2: Strain Distortion

In organic chemistry, you learned that certain structures such as three-membered and four-membered ring structures, such as epoxides were highly reactive due to the strain distortion inherent to the unfavored bond angles inherent to the ring. Enzyme active sites can also utilize strain distortion within a bound substrate to increase the reactivity of the molecule and favor the formation of the transition state. Many enzymes that function by the induced fit model also utilize strain distortion within their catalytic mechanism. Within the unbound state they remain in a low catalytic state, however the interaction with the substrate induces the destabilization of the enzyme active site or may induce strain within the substrate causing the initiation of the catalytic activity of the enzyme.

6.1.3: A Note on Asymmetric Catalysis/Organocatalysis

In a subsequent section, we will discuss how protein enzymes use the catalytic strategies described above. An intriguing question arises: how much of the structure of a large protein is really needed for catalysis? Much work has been directed to the development of small molecule catalysis mimetics of large protein enzymes. Just how small can you go in reducing the size of a protein and still get catalysis?

One important feature of enzyme catalysis is that they catalyze reactions in which only one enantiomer is produced. That is, the synthesis is asymmetric. This is typically a consequence of the asymmetric enzyme (itself chiral) binding only one enantiomer as a reactant and/or the imposition of steric restrictions on the possible reactions of the bound substrate. L-Pro alone can act as such an asymmetric catalyst in an aldol condensation reaction. Figure 6.1.27:

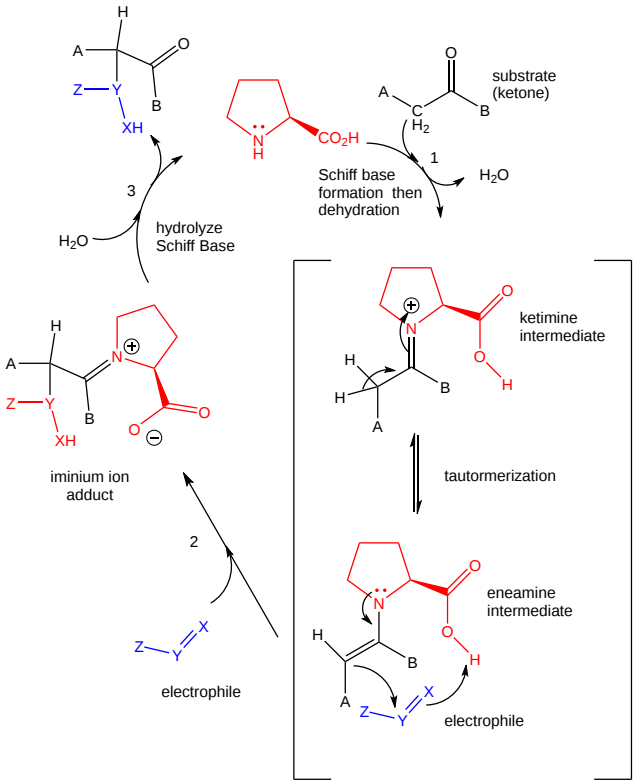


Figure 6.1.27: L-PRO CATALYSIS OF AN ALDOL CONDENSATION: POSSIBLE MECHANISM

Catalysts are vital in biological settings but also in the laboratory synthesis of molecules that sustain our culture and economy. Transition metal and, increasingly, protein enzymes have been used as industrial catalysts. They have now been joined by new **asymmetric catalysts** (a subset of **organocatalysts**). The work of Benjamin List and David MacMillan, who were instrumental in developing the ideas of asymmetric catalyst, has been recognized by the Nobel Commission which awarded the 2021 Nobel Prize in Chemistry to them.

The enzyme triose phosphate isomerase catalyzes an asymmetric reaction, in which only one enantiomer of glyceraldehyde-3-phosphate is produced from the achiral dihydroxyacetone phosphate. Figure 6.1.28 shows this enantiospecific reaction.

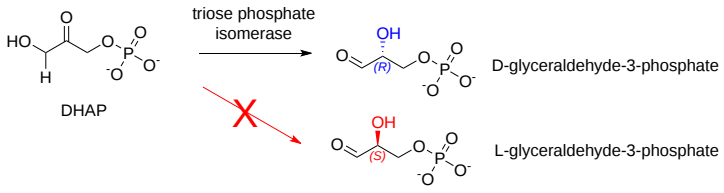


Figure 6.1.28: Reaction catalyzed by triose phosphate isomerase

Figure 6.1.29 contains an interactive iCn3D model of a triose phosphate isomerase from Trypanosoma brucei brucei (1KV5), which shows a conserved active site Pro 168 (spacefill) and amino acid side chains within 4 Å (stick) within the context of one monomer (cartoon) of the dimeric protein.

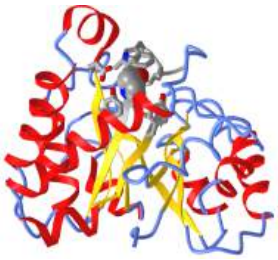


Figure 6.1.29: Active site proline in triose phosphate isomerase from Trypanosoma brucei brucei (1KV5). (Copyright; author via source). Click the image for a popup or use this external link https://structure.ncbi.nlm.nih.gov/icn3d/share.html?1Hnu1xmU5j76Jwrx9

6.1.4: References

Amyes TL, Richard JP. Biochemistry. 2013, 52(12), 2021-35. doi: 10.1021/bi301491r

Ferst and Kirby, J. Am. Chem. Soc. 1967, 89, 19, 4857–4863. https://doi-org.ezproxy.csbsju.edu/10.1021/ja00995a007

Komiyama et al., PNAS, **1977,** 74 (7) 2634-2638. https://doi.org/10.1073/pnas.74.7.263

Lerner, R. A., & Tramontano, A. (1987). Antibodies as enzymes. *Trends in Biochemical Sciences*, *12*, 427-430. https://doi.org/10.1016/0968-0004(87)90208-8

OpenStax College. (2022). *Chemistry*. OpenStax. https://openstax.org/details/books/chemistry-2e

Regan, C. K., Craig, S. L., & Brauman, J. I. *Science.* 2002 *295*(5563), 2245-2247. DOI: 10.1126/science.1068849

6.2: Kinetics without Enzymes

6.2.1: Single Step Reactions

First, we will explore the kinetics of non-catalyzed reactions, which is needed to understand the kinetic of the more complicated enzyme-catalyzed reactions.

Calculus: Derivatives and Integrals

In this book, we will refer to the change of concentration of a chemical species X as a function of time as dX/dt instead of ΔX/Δt, where dX/dt is the derivative of X with respect to time t. This is the language of calculus, which most readers would have studied. The use of calculus will be mainly limited to writing equations of the form dX/dt = f(t), which is a type of differential equation. We will also use a few integrals, but a working knowledge of calculus is not required. For readers who have not studied calculus, replace dX/dt with ΔX/Δt in your mind, and you will derive the same meaning.

You studied two types of kinetic equations in introductory chemistry to analyze kinetic data:

- Initial Rates: In this method, the initial rate, v_0, is measured as a function of the concentration of reactants. The initial velocity, v_0, is the initial slope of a graph of the concentration of reactants or products as a function of time, taken over a range of times such that only a small fraction of A has reacted. Under this condition, [A] over this short time range is approximately constant and equal to Ao. Initial rate graphs are often based on the measurement of product increase with time, ΔP/Δt, so v_0 vs. A plots have positive slopes. The velocity at time t along the A vs. t curve, dA/dt, constantly changes as [A] decreases since the velocity depends on the [A]. To reiterate, the initial velocity of the reaction is the slope of the initial linear part of the decay curve when the rate is essentially linear over a narrow range of [A].
- Integrated Rates and Progress Curves: In this method, a differential equation that gives the change of A or P with time (dA/dt or dP/dt) is integrated to give an equation that shows the concentration of A or P as a function of time. For any given reaction, it is essential to be able to write the integrated rate equation, often called an ordinary differential equation (ODE). In almost every case, we provide the solutions to the ODEs in either a mathematical equation or a fitted graph of A vs. t. As the reactions get more complicated, we let computers solve them numerically and show the output. The ODEs are also called **progress curves** as they show how the concentration of reactants and products change with time.

In either case, the mathematical equations describing the reaction are used to fit experimental data taken in the lab. Statistics are used to see how closely the experimental data fit the rate equations. The best-fit equation give the mostly likely chemical reaction equations for the interconversion of reactant and products.

Most biochemistry textbooks focus on initial velocities when explaining enzyme kinetics. However, in many ways, this method is less intuitive than exploring how the concentration of a molecule changes with time. An analogy is the concept of density (mass/volume), which is more complicated than understanding mass or volume separately. We use both methods to develop both chemical and biological intuitions of kinetic properties.

We will first explore simple irreversible reactions (1st and 2nd order), then make them reversible, and then couple them together to form more complex reaction schemes, much as we are compelled to do for enzyme-catalyzed reactions.

6.2.1.1: First Order Reaction

$$\mathrm{A} \xrightarrow{k_1} \mathrm{P}$$

where k_1 is the first-order rate constant. For these reactions, the velocity of the reaction, v, is directly proportional to [A], or

$$v = -\frac{dA}{dt} = +\frac{dP}{dt} = k_1 A \tag{6.2.1}$$

The negative sign in -d[A]/dt indicates that the concentration of A decreases. The equation could also be written as: $v = \frac{dA}{dt} = -k_1 A$

For the rest of the reactions in this book, we follow the convention of writing all velocities expressed as d[x]/dt as positive numbers. A negative sign for a term on the right-hand side of the differential equation will indicate that the concentration

dependency of that term will lead to a decrease in [x] with time. Likewise, a positive sign for the term on the right-hand side of the equation will indicate that concentration dependency of that term will lead to an increase in [x] with time.

Examples: dA/dt = **-k**A shows that the A will **decrease** with time. dA/dt = **+k**A shows that A will **increase** with time.

Here is the solution to the differential Equation 6.3.A.1 for [A] as a function of t.

$$A = A_0 e^{-k_1 t} \tag{6.2.2}$$

? A derivation of the first order rate equation

Here it is!

Derivation

$$\begin{aligned} \int_{A_O}^{A} \frac{dA}{A} &= -k_1 \int_0^t dt \\ \ln A - \ln A_0 &= -k_1 t \\ \ln A &= \ln A_0 - k_1 t \\ A &= A_0 e^{-k_1 t} \end{aligned} \tag{6.2.3}$$

Equation 6.3.A.2 is an example of an integrated rate equation. The following graphs show plots o.f A vs t and lnA vs. t for a first-order process. Note that the derivative of the graph of A vs. t (dA/dt) is the velocity of the reaction. The graph of ln[A] vs. t is linear with a slope of -k_1. The velocity of the reaction (slope of the A vs t curve) decreases with decreasing A, which is consistent with equation 1. Again, the initial velocity is determined from data taken in the first part of the decay curve when the rate is linear and little A has reacted. That is, [A] is approximately equal to [A_0].

Figure 6.2.1 shows two ways to plot 1st order reaction data. The left graph shows the exponential decay of A with time and the corresponding rise in P when A_0 is 0 and k_1=2. The other shows the linear fall of ln[A] vs time.

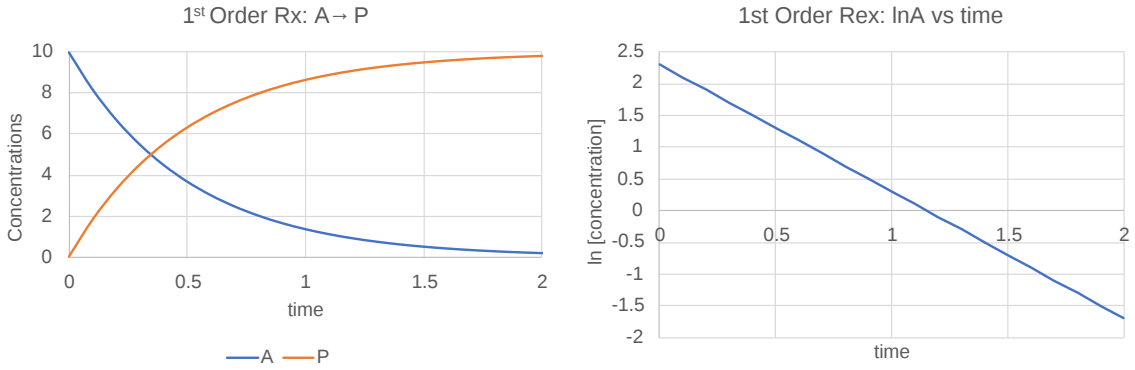


Figure 6.2.1: Two ways to plot 1st order reaction data

Once again, for complete clarification, another way of analyzing the kinetics of a reaction, in addition to following the concentration of a reactant or product as a function of time and fitting the data to an integrated rate equation, is to plot the initial velocity, v_o, of the reaction as a function of the concentration of reactants. The initial velocity is the initial slope of a graph of the concentration of reactants or products as a function of time, taken over a range of times such that only a small fraction of A has reacted, so [A] is approximately constant = A_o. From the first-order graph of A vs. t above, the slope approaches 0 with increasing time as [A] approaches 0, which indicates that the reaction velocity depends on A. For this first-order order process, two equivalent equations can be written showing the

- disappearance of A as v = -d[A]/dt = k_1[A], while
- appearance of A as v = d[A]/dt = -k_1[A],

Both equations show that v is directly proportional to A. As [A] is doubled, the initial velocity is doubled.
Velocity graphs used by biochemists often show the initial velocity of product formation (not reactant decrease) as a function of reactant concentration. Hence, as product concentration increases, the initial velocity are positive. A graph of v (= dP/dt) vs [A] for a first order process would have a positive slope and be interpreted as showing that the rate of appearance of P depends linearly on [A].

6.2.1.2: Second Order/Pseudo First Order Reactions

$$A + B \xrightarrow{k_2} P \text{ or } A + A \xrightarrow{k_2} P \tag{6.2.4}$$

where k_2 is the second-order rate constant. For the first of these irreversible reactions, the velocity of the reaction, v, is directly proportional to [A] and [B], or

$$v = \frac{dA}{dt} = -k_2[A][B] \tag{6.2.5}$$

We will consider two special cases of this reaction type:

1. [B] >> [A]. Under these conditions, [B] never changes, so Equation 5 becomes

$$v = -\left(k_2[B]\right)[A] = -k_1'[A] \tag{6.2.6}$$

where $\mathbf{k_1'}$ is the pseudo first order rate constant (= k2[B]) for the reaction. The reaction appears to be first order, depending only on [A].

2. As illustrated in the second reaction above, the only reactant is A, which must collide with another A to form P.

The following differential equation can be written and solved to find [A] as a function of t.

$$v = \frac{dA}{dt} = 2\frac{dP}{dt} = -k_2A^2 \tag{6.2.7}$$

Solving the differential equation for A gives the following:

$$\frac{1}{A} = \frac{1}{A_0} + k_2t \tag{6.2.8}$$

? A derivation of the second order rate equation

Here it is!

Derivation

$$\begin{gathered}
\frac{dA}{dt} = -k_2A^2 \\
\int_{A_0}^{A}\frac{dA}{A^2} = \int_{A_0}^{A}A^{-2}dA = \int_{A_0}^{A}A^{-2}dA = -k_2\int_0^t dt \\
\left.\frac{A^{n+1}}{n+1}\right]_{A_0}^{A} = \left.\frac{A^{-1}}{-1}\right]_{A_0}^{A} = \left.-k_2t\right]_{A_0}^{A} \\
-\frac{1}{A} - \left(\frac{1}{A_0}\right) = -k_2t
\end{gathered} \tag{6.2.9}$$

Figure 6.2.2 shows plots of A vs. t and 1/A vs. t for a second-order process when A_0 is 0 and k_2=1. The right graph shows the linear rise of 1/[A] with time.

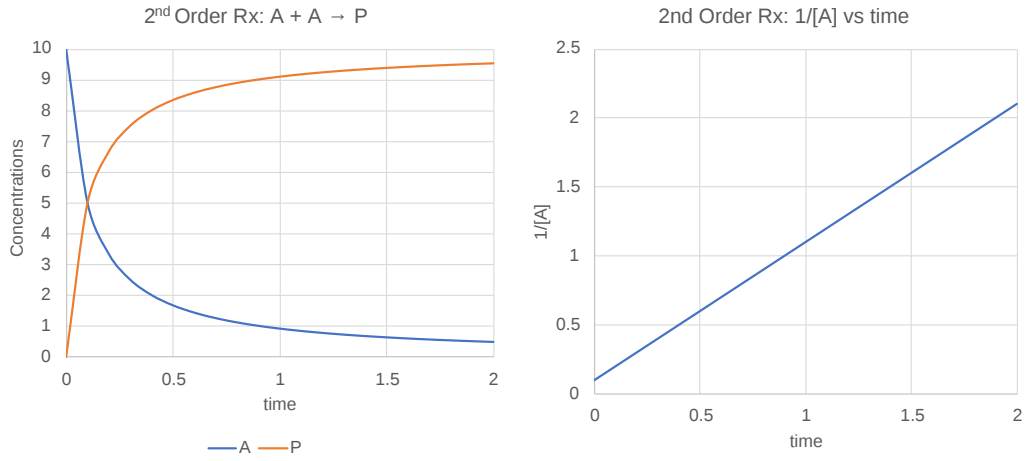


Figure 6.2.2: Plots of A vs t and 1/A vs t for a second order process when A_0 is 0 and k_2=1

Note that just from a plot of A vs. t, it would be difficult to distinguish a first from a second-order reaction. If the plots were superimposed, you would observe that at the same concentration of A (10, for example), the v_o of a first-order reaction would be proportional to 10, but for a second-order reaction, to 10^2 or 100. Therefore, the second-order reaction is faster (assuming similarity in the relative magnitude of the rate constants), as indicated by the steeper negative slope of the curve. However, at low A (0.1 example), the v_o of a first-order reaction would be proportional to 0.1 but second-order order reaction to 0.1^2 or 0.01. Therefore, at low A, the second-order reaction is slower.

The interactive graphs below show the first and second-order conversion of reactant A to product. Change the sliders to see how the curves are different.

By comparing these curves, you should see how difficult it is to differentiate between a 1st and 2nd order process unless the reaction progresses to almost completion.

6.2.2: Multi-Step Reactions

6.2.2.1: Reversible First Order Reactions

$$A \underset{k_2}{\overset{k_1}{\rightleftarrows}} P \tag{6.2.10}$$

Here is the differential equation for dA/dt..

$$v = \frac{dA}{dt} = -k_1\ A + k_2 P \tag{6.2.11}$$

Here are the solution for both A and P as a function of time

$$\begin{aligned} A &= \frac{A_0\left(k_2 + k_1\left[e^{-(k_1+k_2)t}\right]\right.}{k_1 + k_2} \\ P &= A_0 - \left(\frac{A_0\left(k_2 + k_1\left[e^{-(k_1+k_2)t}\right]\right.}{k_1 + k_2}\right) \end{aligned} \tag{6.2.12}$$

Figure 6.2.3 shows graphs of A and P vs t for this reaction at two different sets of values of k_1 and k_2.

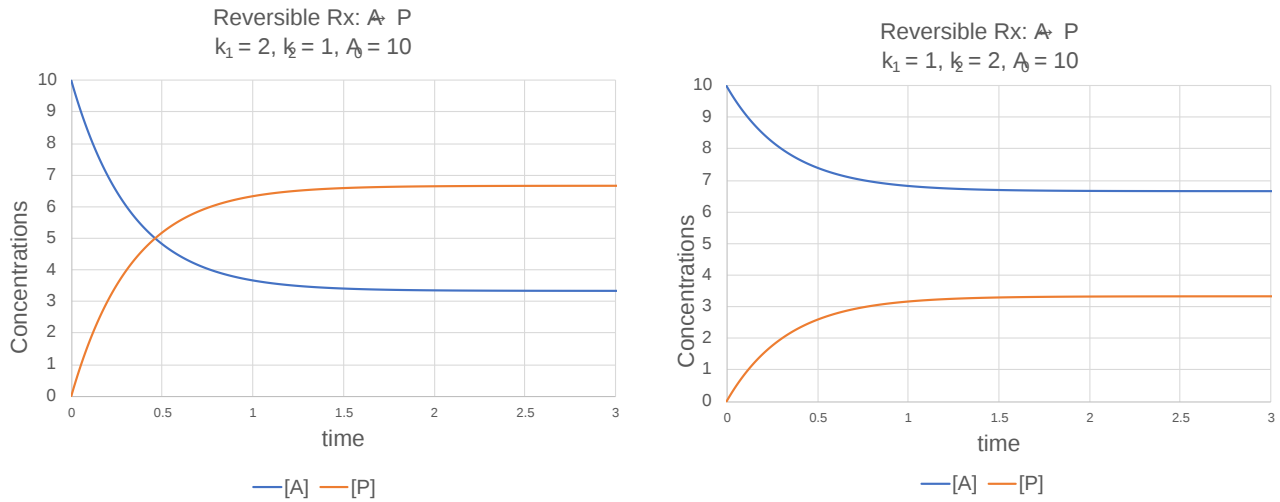


Figure 6.2.3: Graphs of A and P vs t for the reversible reaction A ↔ P at two different sets of values of k_1 and k_2

Change the sliders on the interactive graph below of a reversible reaction to see how changing the relative values of the forward and reverse rate constants affects the concentrations at which the concentration plateaus are reached.

We all grew up on mathematical graphs that give you valuable insight into textual descriptions and data tables from which the graphs were made. These graphs are enhanced when you can use sliders to change constants as for the reversible reaction A ↔ P above. Even then, you might not infer that when the reaction has reached equilibrium, product is still being made from reactant,

and reactant from product, since the equilibrium is dynamic. To add insight into simple and complex reactions, animations showing the continual disappearance of reactants and products are valuable.

This book will incorporate many animations to visually show the changes in the reactant and product concentrations. Hui Liu and Shraddha Nayak (Animation Lab, University of Utah) made all the animations in this book using mathematical solutions to the progress curves for the reactions. Multiple modes of presentation are useful as learners with different backgrounds and preferred ways of learning attempt to understand complex materials.

It is relatively simple to write the differential equations (differential) to show how the rate of disappearance of a reactant A (for example), dA/dt, depends on the concentration of its immediate participants in the reaction. It is not so easy to solve the equations (as we did above) for the progress curve, which shows how [A] changes with time t (i.e. [A] = f(t)). Luckily, many programs have been developed that produce numerical solutions to the differential equations and give progress curve graphs like [A] = f(t). Two interrelated, freely available programs, Copasi and Virtual cell (Vcell), can solve all the equations for hundreds of cellular reactions simultaneously. They use a format called **Systems Biology Markup Language** (**SBML**) for describing and storing computational models. We will use Vcell models in this book as they are straightforward to create. All the coding to describe the reactions is built into Vcell and this book and hidden from you. All you will see are the output results. You can change the progress curves by moving sliders to change constants and see the resulting changes in graphical outputs.

The VCell models use a reaction diagram that shows all of the interconnected species. The first Vcell model we will run is for the reversible conversion of A to P (A ↔P), which we just discussed and displayed in the graph above. Here is the Vcell reaction diagram and a description of its components.

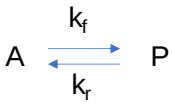
- The reactant A and product P are called species and are shown as green spheres.
- The yellow square indicates a reaction node connecting A to P.
- Lines connect the species that participate in the reaction. The arrows appear unidirectional, **BUT** the equations describing the concentrations of A and P are derived assuming a reversible reaction with rate constants k_f (forward reaction) and k_r (reverse reaction).

The program calculates A and P as a function of time (i.e. it solves the differential equations for both species). The graphs of concentration vs time are called progress curves. It also can calculate **fluxes (J)** (velocities) for each species. The flux at any given time is the slope of the concentration vs time curve at any given time. When we get to metabolism, we will talk about fluxes of metabolites through pathways. Also, fluxes are used to describe the rate of movement of solute through membranes. Here is the result of the simulation run in Vcell, exported as a sbml file, and displayed in the book using a program called MiniSideWinder.

VCell MODEL

Reversible reaction A ↔ P

Vcell reaction diagram (1-way arrows defined as reversible in actual mathematical model) and chemical equation



Initial parameter values: $k_f = 2$, $k_r = 4$

Select Load [model name] below

Load AtoP_Rev_Uncatalyzed

Select **Start** to begin the simulation.

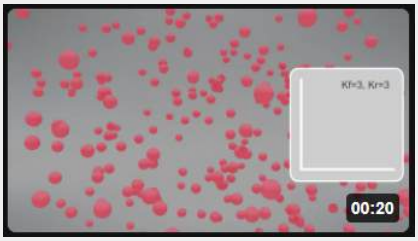
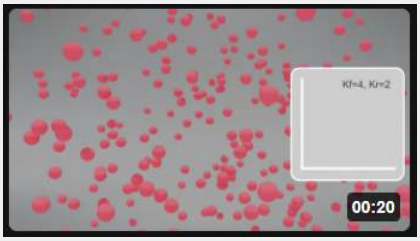
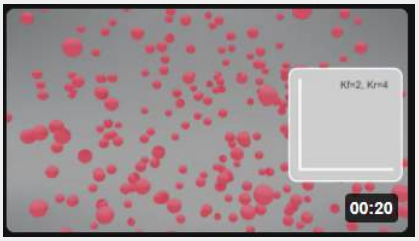
Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

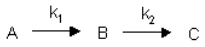
Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

Animations

The video animations show particles representing A (red) and P (cyan) interconverting in a reversible process with embedded progress curves showing A (red) and P (cyan) vs. time.

Reversible Rx A (red) ↔ B (cyan): k_F = k_R = 3	Reversible Rx A (red) ↔ B (cyan): k_F=4, k_R=2	Reversible Rx A (red) ↔ B (cyan): k_F=2, k_R=4
		

6.2.2.2: Consecutive Irreversible First Order Reactions



The following differential equations can be written for these reactions:

$$\begin{aligned} \frac{dA}{dt} &= -k_1\, A \\ \frac{dB}{dt} &= k_1\, A - k_2\, B \\ \frac{dC}{dt} &= k_2\, B \end{aligned} \tag{6.2.13}$$

Here are the solutions to the differential equations:

$$\begin{aligned} A &= A_0 e^{-k_1 t} \\ B &= \frac{k_1 A_0}{k_2 - k_1}\left(e^{-k_1 t} - e^{-k_2 t}\right) \\ C &= A_0 - A - B = A_0\left[1 + \frac{1}{k_1 - k_2}\left(k_2 e^{-k_1 t} - k_1 e^{-k_2 t}\right]\right. \end{aligned} \tag{6.2.14}$$

Figure 6.2.4 shows graphs of A, B, and C vs. t for these reactions for a fixed value of k_1 and k_2.

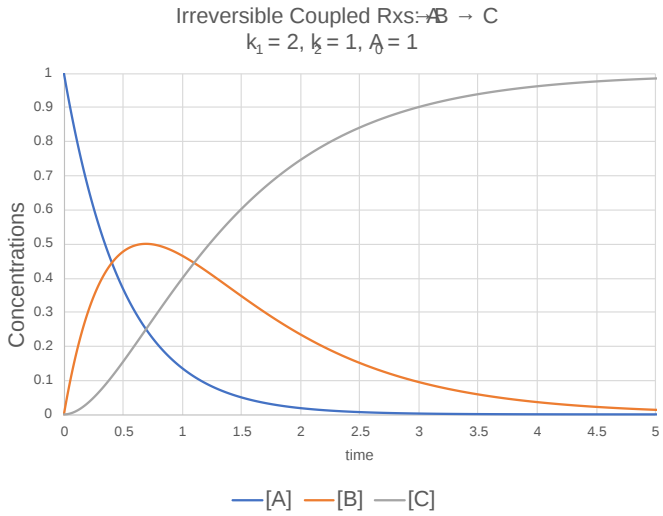
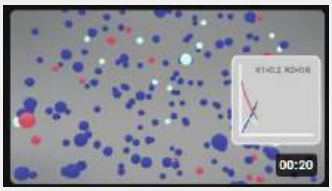


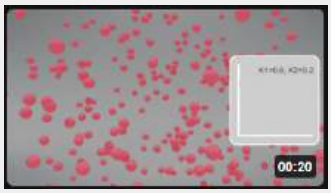
Figure 6.2.4: Graphs of A, B, and C vs. t for the irreversible reactions A → B → C for a fixed value of k_1 and k_2.

Change the sliders on the interactive graph below to see how changing the forward and reverse rate constants affect the curves.

Here are two different animations for the **irreversible** reaction using different rate constants.

Irreversible Rx A (red) → B (cyan) → C (blue) k_1 = 0.2, k_2 = 0.6	Irreversible Rx A (red) → B (cyan)→ C (blue) k_1= 0.6, k_2 = 0.2





6.2.2.3: Consecutive Reversible First Order Reactions

You can imagine that solving the equations for the completely reversible reactions of A ↔ B ↔ C would be very difficult. However, writing the differential equations for each step is straight-forward and can be done easily in Vcell by choosing the built-in equations for each separate reaction based on mass action. The program can then solve the equations numerically to produce progress curve graphs.

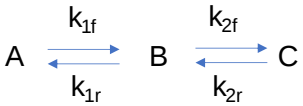
Now let's look at the simulation for the fully **reversible** reactions A ↔ B ↔ C. Again, the model was built and solve in VCell, and then exported in the system's biology markup language (sbml) format. The interactive graphs are made using a program called miniSideWinder.

A note: Arrows in VCell Diagrams - In the reaction diagram for the reversible reaction A ↔ B ↔ C below, the arrows go in only one direction, left to right, and simply show that the species are connected. However, in the Vcell program, the equations for the reversible reaction were used to produce the graphs below. To run the simulation of the irreversible reaction, the rate constants for the reverse reaction would be set to 0.

VCell MODEL

Reversible reaction A ↔ B ↔ C.

Vcell reaction diagram (1-way arrows defined as reversible in actual mathematical model)



Initial parameter values: k_{1f} = 0.2, k_{1r} = 0.1, k_{2f} = 0.6, k2r = 0.3 A_0 = 1

Select Load [model name] below

Load A to B to C_RevNoEnz

Select **Start** to begin the simulation.

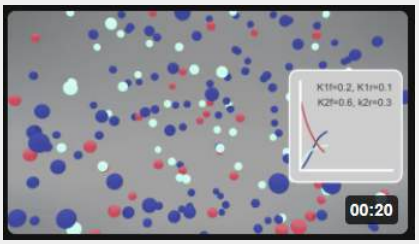
Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

Here is the corresponding animation for the fully reversible reaction A ↔ B ↔ C.

Reversible Rx A (red) ↔ B (cyan) ↔ C Blue)
k_{1f} = 0.2, k_{1r} = 0.1; k_{2f} =0.6, k_{2r} = 0.3



6.3: Kinetics with Enzymes

An enzyme alters the pathways for converting a reactant to a product by binding to the reactant and facilitating the intramolecular conversion of bound substrate to bound product before it releases the product. Enzymes do not affect the thermodynamics of reactions. For reversible reactions (as an example), the equilibrium constant, $K_{eq,}$ is unchanged. What is charged is the rate at which equilibrium is achieved. Enzymes lower the activation energy for bound transition states and change the reaction mechanism.

Figure 6.3.1 shows the simplest chemical reaction that can be written to show how an enzyme catalyzes a reaction.

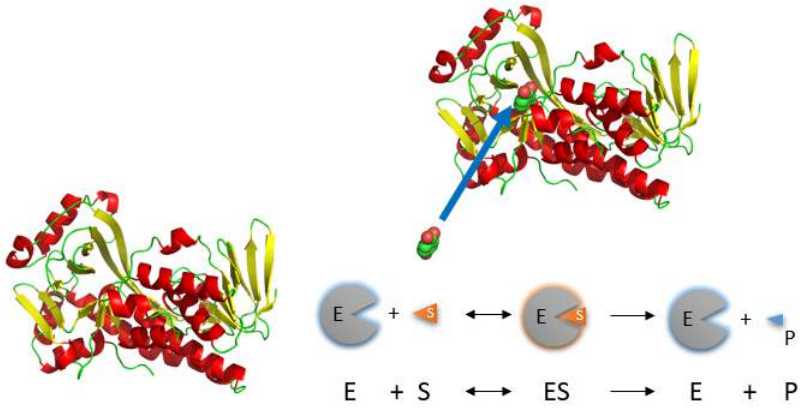


Figure 6.3.1: Chemcial equation for a simple enzyme-catalyzed reaction

6.3.1: Rapid Equilibrium Enzyme-Catalyzed Reactions

We have previously derived equations for the reversible binding of a ligand to a macromolecule. Next, we derived equations for the receptor-mediated facilitated transport of a molecule through a semipermeable membrane. This latter case extended the former case by adding a physical transport step. Now, in what hopefully will seem like deja vu, we will derive almost identical equations for the chemical transformation of a ligand, commonly referred to as a substrate, into a product by an enzyme. We will study two scenarios based on two different assumptions, each enabling a straightforward mathematical derivation of kinetic equations:

1. **Rapid Equilibrium Assumption** - enzyme E (macromolecule) and substrate S (ligand) concentrations can be determined using the dissociation constant since E, S, and ES are in rapid equilibrium, as we previously used in our derivation of the equations for facilitated transport. Sorry about the switch from A to S in the designation of the substrate. Biochemists use S to represent the substrate (ligand) and A, B, P, and Q to represent reactants and products in the case of multi-substrate and multi-product reactions.
2. **Steady State Assumption** (more general) - enzyme and substrate concentrations are not those determined using the dissociation constant.

Enzyme kinetics experiments, as we will see in the following chapters, must be used to determine the detailed mechanism of the catalyzed reaction. Using kinetic analysis, you can determine the order of binding/dissociation of substrates and products, the rate constants for individual steps, and clues to the mechanism used by the enzyme in catalysis.

Consider the following reaction mechanism for the enzyme-catalyzed conversion of substrate S into product P. (We will assume that the catalyzed rate is much greater than the noncatalyzed rate.)

$$\mathrm{E} + \mathrm{S} \overset{K_s}{\rightleftharpoons} \mathrm{ES} \overset{k_3}{\longrightarrow} \mathrm{E} + \mathrm{P}$$

As we did for the derivation of the equations for the facilitated transport reactions under rapid equilibrium conditions, this derivation is based on the assumption that the relative concentrations of S, E, and ES can be determined by the dissociation constant, K_S, for the interactions and the concentrations of each species during the early part of the reaction (i.e. under initial rate conditions). Assume also the S >> E_0. Remember that under these conditions, S does not change much with time. Is this a valid assumption? Examine the mechanism shown above. S binds to E with a second-order rate constant k_1. ES has two fates. It can dissociate with a first-order rate constant k_2 to S + E, or it can be converted to product with a first-order rate constant of k_3 to give P + E. If we assume that k_2 >> k_3 (i.e. that the complex falls apart much more quickly than S is converted to P), then the relative ratios of S, E, and ES can be described by Ks. Alternatively, you can think about it this way. If S binds to E, most of S will

dissociate, and a small amount will be converted to P. If it does, then E is now free, and will quickly bind S and reequilibrate since the most likely fate of bound S is to dissociate, not be converted to P (since $k_3 \ll k_2$). This also makes sense if you consider that the physical step, characterized by k_2, is likely to be quicker than the chemical step, characterized by k_3. Hence the following assumptions have been used:

- $S \gg E_0$
- $P_0 = 0$
- k_3 is rate limiting (i.e. the slow step)

We will derive equations showing the initial velocity v as a function of the initial substrate concentration, S_0, assuming that P is negligible over the time period used to measure the initial velocity. Also assume that $v_{catalyzed} \gg v_{noncatalyzed}$. In contrast to the first-order reaction of S to P in the absence of E, v is not proportional to S_0 but rather to S_{bound}. Therefore, $v \propto [ES]$, or

$$v = const[ES] = k_3[ES] \tag{6.3.1}$$

where v is the velocity (i.e., reaction rate).

Now, let's get ES from the dissociation constant K_S (assuming rapid equilibrium of E, S and ES) and mass balance for E (E_0 = E + ES , so E = E_0 - ES). We will use mass balance for S when we derive the equation of the steady state. In many cases, S is approximately equal to S_0 or the total amount of substrate. This makes sense if you consider that the enzyme is a catalyst acts repeatedly to produce the product, so you don't need significant amounts of the enzyme. The resulting equation of [ES] is shown below.

$$(ES) = \frac{(E_0)\,(S)}{K_S + (S)} \tag{6.3.2}$$

? A derivation of [ES] under rapid equilibrium conditions

Here it is!

Derivation

$$\begin{aligned} K_S = \frac{[E][S]}{[ES]} &= \left(\frac{([E_0] - [ES])\,[S])}{[ES]} \right. \\ (ES)K_S &= (E_0)\,(S) - (ES)(S) \\ (ES)K_S + (ES)(S) &= (E_0)\,(S) \\ (ES)\,(K_S + (S)) &= (E_0)\,(S) \end{aligned} \tag{6.3.3}$$

This derivation assumes that we know S (which is equal to S_0) and E_{tot} (which is E_0).

Let us assume that S is much greater than E, as is the likely biological case. We can calculate ES using the following equations and the same procedure we used for the derivation of the binding equation, which gives the equation below:

$$[ES] = \frac{[E_0]\,[S]}{K_S + S} \tag{6.3.4}$$

which is analogous to

$$[ML] = \frac{[M_0]\,[S]}{K_D + L} \tag{6.3.5}$$

This leads to

$$v_0 = \frac{(k_3)\,(E_0)\,(S)}{K_S + (S)} = \frac{(V_M)\,(S)}{K_S + (S)} \tag{6.3.6}$$

where

$$V_M = k_3 E_0 \tag{6.3.7}$$

This is the world-famous Henri-Michaelis-Menten Equation. It is a hyperbola just like the graph for binding of a ligand to a macromolecule with a given dissociation constant, K_D.

Move the sliders in the interactive graph below to see how changing Km and Vm alters the graph. Note that this graph is identical to the graph for M + L ↔ ML.

× Choose File No file chosen

- Examples
  - Space Curves
  - Implicit Surfaces
  - Parametric Surfaces
  - Vector Fields

Space Curves:

Implicit Surfaces:

Parametric Surfaces:

Vector Fields:

☰

Welcome to CalcPlot3D!

Just as in the case with noncatalyzed first-order decay, it is easiest to measure the initial velocity of the reaction when [S] does not change much with time and the velocity is constant (i.e. the slope of the dP/dt curve is constant). A plot of [P] vs t (called a progress curve) is made for each different substrate concentration studied. From these curves, the initial rates at each [S] is determined.

Alternatively, one reaction time that gives a linear rise in [P] with time is determined for all the different substrate concentrations. At that specified time, the reaction can be stopped (quenched) with a reagent that does not cause any change in S or P. Then initial rates can be easily calculated for each [S] from a single data point.

Under these conditions:

- a plot of v vs S is hyperbolic
- v = 0 when S = 0
- v is a linear function of S when S<<Ks.
- $v = V_{max}$ (or V_M) when S is much greater than Ks
- $S = K_S$ when $v = V_M/2$.

These are the same conditions we detailed for our understanding of the binding equation

$$(ML) = \frac{(M_0)\,L}{K_D + L} \tag{6.3.8}$$

Note that when S is not >> K_S, the graph does not reach saturation and does not look hyperbolic. It should be apparent from the graph that only if S >> KS (or when S is approximately 100 x > K_s) will saturation be achieved.

The K_S constant is usually called the **Michaelis constant, K_M**. We will see in a bit that the K_M for most enzyme-catalyzed reactions is not equal to the dissociation constant for ES, which we called K_S.

Very often, these graphs are transformed into **double reciprocal or Lineweaver-Burk** plots as shown below.

$$\frac{1}{v_0} = \frac{K_M + S}{V_M S} = \left(\frac{K_M}{V_M}\right)\frac{1}{S} + \frac{1}{V_M} \tag{6.3.9}$$

These plots are used to estimate V_M from the 1/v intercept ($1/V_M$) and K_M from the 1/S axis ($-1/K_M$). These values should be used as "seed" values for a nonlinear fit to the hyperbola that models the actual v vs S curve. An interactive graph of $1/v_0$ vs 1/[S] is shown below. Change the sliders for K_M and V_M and note the change in the slope and intercepts of the plot

× Choose File No file chosen

- Examples
  - Space Curves
  - Implicit Surfaces
  - Parametric Surfaces
  - Vector Fields

Space Curves:

Implicit Surfaces:

Parametric Surfaces:

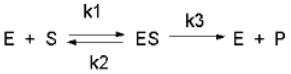
Vector Fields:

Welcome to CalcPlot3D!

Help

As we saw in the graph of A or P vs t for a noncatalyzed, first-order reaction, the velocity of the reaction, given as the slope of those curves, is always changing. Which velocity should we use in Equation 5? The answer invariably is the initial velocity, v_0, measured in the early part of the reaction when little substrate is depleted. Hence v vs S curves for enzyme-catalyzed reactions invariably are really v_0 vs [S] curves.

6.3.2: Steady State Enzyme-Catalyzed Reactions



In this derivation, we will consider the following equations and all the rate constants, and will not arbitrarily assume that $k_2 >> k_3$. We will still assume that S >> E_0 and that P_0 = 0. An added assumption, however, is that d[ES]/dt is approximately 0. Look at this assumption this way. When an excess of S is added to E, ES is formed. In the rapid equilibrium assumption, we assumed that it would fall back to E + S (a physical step) faster than it would go onto product (a chemical step). In the steady state case, we will assume that ES might go on to product either less or more quickly than it will fall back to E + S. In either case, a steady state concentration of ES arises within a few milliseconds, and its concentration does not change significantly during the initial part of the reaction under which the initial rates are measured. Therefore, d[ES]/dt is about 0. For the rapid equilibrium derivation, v = k_3[ES]. We then solved for ES using K_S and mass balance of E. In the steady state assumption, the equation v= k_3[ES] still holds, but now we will solve for [ES] using the steady state assumption that d[ES]/dt =0.

$$\frac{d[ES]}{\mathrm{dt}} = \mathrm{k_1}[E][S] - \mathrm{k_2}[ES] - \mathrm{k_3}[ES] = 0 \tag{6.3.10}$$

We can solve this and obtain the Michaelis-Menten equation for reaction.

$$\mathrm{v} = \mathrm{k_3}[\mathrm{ES}] = \frac{\mathrm{k_3}\,[\mathrm{E_0}]\,[\mathrm{S}]}{\frac{\mathrm{k_2}+\mathrm{k_3}}{\mathrm{k_1}} + \mathrm{S}} = \frac{\mathrm{V_M}[\mathrm{S}]}{\mathrm{K_M} + \mathrm{S}} \tag{6.3.11}$$

? Derivation of the Michaelis-Menten Equation for the steady state

To see how to derive this, click below.

Answer

Applying mass balance for E (i.e E = E_0-ES) in appropriate term below gives

$$\begin{aligned}
&\mathrm{k_1}[E][S] = (\mathrm{k_2} + \mathrm{k_3})\,[ES] \\
&\mathrm{k_1}\,[\mathrm{E_0} - \mathrm{ES}]\,[S] = (\mathrm{k_2} + \mathrm{k_3})\,[ES] \\
&\mathrm{k_1}\,[\mathrm{E_0}]\,[S] - \mathrm{k_1}[\mathrm{ES}][S] = (\mathrm{k_2} + \mathrm{k_3})\,[ES] \\
&\mathrm{k_1}\,[\mathrm{E_0}]\,[S] = (\mathrm{k_2} + \mathrm{k_3})\,[ES] + \mathrm{k_1}[\mathrm{ES}][S] \\
&\mathrm{k_1}\,[\mathrm{E_0}]\,[S] = [ES]\,(\mathrm{k_2} + \mathrm{k_3} + \mathrm{k_1 S})
\end{aligned} \tag{6.3.12}$$

Solving for [ES] gives

$$[\mathrm{ES}] = \left(\frac{[\mathrm{E_0}]\,[\mathrm{S}]}{\frac{\mathrm{k_2}+\mathrm{k_3}}{\mathrm{k_1}} + \mathrm{S}} \right) \tag{6.3.13}$$

Substituting into the Henri Michaelis Menten equation gives

$$\mathrm{v} = \mathrm{k_3}[\mathrm{ES}] = \frac{\mathrm{k_3}\,[\mathrm{E_0}]\,[\mathrm{S}]}{\frac{\mathrm{k_2}+\mathrm{k_3}}{\mathrm{k_1}} + \mathrm{S}} = \frac{\mathrm{V_M}[\mathrm{S}]}{\mathrm{K_M} + \mathrm{S}} \tag{6.3.14}$$

Note that

$$V_M = k_3 E_0 \tag{6.3.15}$$

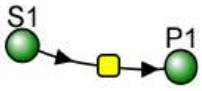
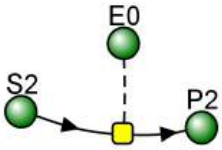
and

$$K_M = \frac{k_2 + k_3}{k_1} \tag{6.3.16}$$

Now let's look at a progress curve simulation that compares the Michaelis-Menten equation derived from rapid equilibrium assumptions (when k_2>>k_3) and in which K_M is the actual dissociation constant to the steady state approximation, when k_2 is not >>k_3.

VCell MODEL

A comparison of the rapid equilibrium (left) and steady state (right)Michaels-Menten reaction

	
$v_0 = \frac{k_3[E_O][S]}{K_M + S} = \frac{V_M[S]}{K_M + S}$ (6.3.17)	$v_0 = \frac{k_3[E_O][S]}{\frac{k_2+k_3}{k_1} + S} = \frac{V_M[S]}{K_M + S}$ (6.3.18)
Initial Condition: $V_M = 10$; $K_M = 10$; $S = 50$	Initial Conditions: $k_1 = 10$; $k_2 = 90$; $k_3 = 10$; $E_0 = 1$ uM; $S = 50$ $V_M = k_3E_0 = 10$; $K_M = (k_2 + k_3)/k_2 = 10$)

Select Load [model name] below

Load MMvsSteadyState

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

The initial conditions in the graph are set so the graphs of the rapid equilibrium and steady state are identical.

Recommendations:

- Scale the y-axis to 50
- change the slider k_{2_r2} for the steady state graph to other values. Watch the curves separate. Although these plots are only for 1 substrate concentration (50 uM), the effects of changing k_2 for the steady state plot are very dramatic. This should convince you that in general, unless $k_3 << k_2$, the calculated value of K_M is **not** equal to the thermodynamic dissociation constant, K_D.

6.3.3: Analysis of the General Michaelis-Menten Equation

This equation can be simplified and studied under different conditions. First, notice that $(k_2 + k_3)/k_1$ is a constant which is a function of relevant rate constants. This term is usually replaced by K_M which is called the **Michaelis constant**. Likewise, when S approaches infinity (i.e. $S >> K_M$, equation 5 becomes $v = k_3(E_0)$ which is also a constant, called V_M for maximal velocity. Substituting V_M and K_M into equation 5 gives the simplified equation:

$$v = \frac{V_M[S]}{K_M + S} \tag{6.3.19}$$

It is extremely important to note that K_M in the general equation does **n**ot equal the K_S, the dissociation constant used in the rapid equilibrium assumption! K_M and K_S have the same units of molarity, however. A closer examination of K_M shows that under the limiting case when $k_2 >> k_3$ (the rapid equilibrium assumption) then,

$$K_M = \frac{k_2 + k_3}{k_1} = \frac{k_2}{k_1} = K_D = K_S \tag{6.3.20}$$

If we examine these equations under several different scenarios, we can better understand the equation and the kinetic parameters:

- when $S = 0$, $v = 0$.
- when $S >> K_M$, $v = V_M = k_3E_0$. (i.e. v is zero order with respect to S and first order in E. Remember, k_3 has units of s^{-1} since it is a first-order rate constant. k_3 is often called the **turnover number,** because it describes how many molecules of S "turn over"

to product per second.

- v = V_M2, when S = K_M.
- when S << K_M, v = V_MS/K_M = k_3E_0S/K_M (i.e. the reaction is bimolecular, dependent on both on S and E. k3/K_M has units of $M^{-1}s^{-1}$, the same as a second order rate constant.

6.3.4: More Complicated Enzyme-catalyzed Reactions

6.3.4.1: A reversibly-catalyzed reaction

You have learned previously that enzymes don't change the equilibrium constant for a reaction, but rather lowers the activation energy barrier to move from reactants to products. This implies that the activation energy to move in the reverse direction, from products to reactants, is also lowered. Hence the enzyme speeds up both the forward and reverse reaction. We haven't accounted for that yet in our kinetic equations. Many reactions in metabolic reactions that do not have large negative values of ΔG (ie. they are not significantly favored) are reversible, allowing the enzyme to be used in the reverse direction. Take for example the pathway to break down glucose to pyruvate (glycolysis). It has 9 different steps, of which 5 are reversible, allowing them to be used in the reverse pathway to take pyruvate to glucose (gluconeogenesis).

Let's set up the equations for the reversible reaction of substrate S to product P catalyzed by enzyme E. Assume that the K_M for the forward reaction is K_{MS} (or K_S) and for the reverse reaction is K_{MP} (or K_P) as shown in the reaction scheme in Figure 6.3.2. The rate constant k_2 is the k_{cat} (forward rate constant) for conversion of ES to EP and k_{-2} is the k_{cat} (reverse rate constant) for conversion of EP to ES

$$\begin{array}{ccccccc} & & K_{MS} & & K_{MP} & & \\ & & K_S & k_2 & K_P & & \\ E + S & \leftrightarrow & ES & \leftrightarrow & EP & \leftrightarrow & E + P \\ & & & k_{-2} & & & \end{array}$$

Figure 6.3.2: Reaction scheme for a simple enzyme-catalyzed reaction

The following simple Michaelis-Menten equations can be written for just the forward reaction and for the reverse reaction:

$$\begin{aligned} v_f &= \frac{V_f S}{K_{MS} + S} = \frac{\frac{V_f S}{K_{MS}}}{1 + \frac{S}{K_{MS}}} \\ v_r &= \frac{V_r P}{K_{MP} + P} = \frac{\frac{V_r P}{K_{MP}}}{1 + \frac{P}{K_{MP}}} \end{aligned} \tag{6.3.21}$$

Now you might think that simply substracting the two would give the net velocity in the forward direction, but that is **NOT** the case.

$$v \neq \left[\frac{\frac{V_f S}{K_{MS}}}{1 + \frac{S}{K_{MS}}} - \frac{\frac{V_r P}{K_{MP}}}{1 + \frac{P}{K_{MP}}} \right] \tag{6.3.22}$$

The reason is that in the derivation, the equations for both the forward and reverse rates must have terms for the reverse and forward reactions, respectively.

A simple derivation shows that this is the equation for the reversible conversion of substrate to product.

$$v = k_2[ES] - k_{-2}[EP] = \frac{V_f \frac{[S]}{K_S}}{\left[1 + \frac{[S]}{K_S} + \frac{[P]}{K_p}\right]} - \frac{V_r \frac{[P]}{K_P}}{\left[1 + \frac{[S]}{K_S} + \frac{[P]}{K_P}\right]} = \frac{V_f \frac{[S]}{K_S} - V_r \frac{[P]}{K_P}}{\left[1 + \frac{[S]}{K_S} + \frac{[P]}{K_P}\right]} \tag{6.3.23}$$

This reversible form of the Michaelis-Menten equation and other equations, written in the format shown in equation 6.24, are commonly used in programs such as VCell and Copasi to model the kinetics of whole pathways of biological interactions and reactions. The figures below show a reaction diagram, graphical results showing S and P vs time for the selected K_M and V_M values shown, and animations for the reaction. The chemical (S and P) are shown as green spheres connected by a line. The red dot again

represents the enzyme (shown as a node of connection between S and P). Equation 6.24 was used to model the reversible reaction (even though the arrows shown between S and P are unidirectional.

VCell MODEL

Reversible Enzyme-Catalyzed Reaction: E + S ↔ ES ↔ EP ↔ E + P .

Vcell reaction diagram (1-way arrows defined as reversible in the actual mathematical model) and chemical equation

J	reaction rate	$\frac{\left(\frac{S \cdot VmaxFwd}{KmFwd} - \frac{P \cdot VmaxRev}{KmRev}\right)}{\left(1.0 + \frac{S}{KmFwd} + \frac{P}{KmRev}\right)}$
KmFwd	Km forward	10.0
VmaxFwd	max forward rate	10.0
KmRev	Km reverese	20.0
VmaxRev	max reverse rate	5.0
Size	size	1.0
initConc	initial concentration for S	5.0
initConc	initial concentration for P	0.0

Initial parameter values: as shown in above

Select Load [Enz Rev] below

Load Enz Rev

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

If you reflect on it, this reaction is very similar to the reversible reaction of A ↔ P in the **absence** of an enzyme, which we explored in Chapter section 6.2. Just for comparison, the graph for that reaction is shown below. Change the sliders to produce a curve similar to the enzyme-catalyzed reaction shown in the figure above.

6.3.4.2: Reaction with intermediates

Not all reactions can be characterized so simply as a simple substrate interacting with an enzyme to form an ES complex, which then turns over to form product. Sometimes, intermediates form. For example, a substrate S might interact with E to form a complex, which then is cleaved to products P and Q. Q is released from the enzyme, but P might stay covalently attached. This happens often in the hydrolytic cleavage of a peptide bond by a protease, when an activated nucleophile like Ser reacts with the sessile peptide bond in a nucleophilic substitution reaction, releasing the amine end of the former peptide bond as the leaving group while the carboxy end of the peptide bond remains bonded to the Ser as an Ser-acyl intermediate. Water then enters and cleaves the acyl intermediate, freeing the carboxyl end of the original peptide bond. This is shown in the written reaction in Figure 6.3.3:

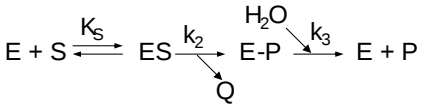


Figure 6.3.3: Reaction scheme for enzyme-catalyzed reaction with an intermediate

Even for this seemingly complicated reaction, you get the standard Michaelis-Menten equation.

To simplify the derivation of the kinetic equation, let's assume that E, S, and ES are in rapid equilibrium defined by the dissociation constant, Ks. Assume Q has a visible absorbance, so it is easy to monitor. Assume from the steady state assumption that:

$$\frac{d[E-P]}{dt} = k_2[ES] - k_3[E-P] = 0 \tag{6.3.24}$$

assuming that k_3 is a pseudo first-order rate constant and that $[H_2O]$ doesn't change.

The velocity depends on which step is rate-limiting. If $k_3 << k_2$, then the k_3 step is rate-limiting. Then

$$v = k_3[E-P] \tag{6.3.25}$$

If $k_2 << k_3$, then the k_2 step is rate-limiting. Then

\

$$v = k_2 [ES] \tag{6.3.26}$$

The following kinetic equation for this reaction can be derived, assuming $v = k_2[ES]$.

$$v = \frac{\frac{k_2 k_3}{k_2 + k_3} [E_0] [S]}{\left(\frac{k_3}{k_2 + k_3}\right) K_S + S} = \frac{k_{cat} [E_0] [S]}{K_M + S} = \frac{V_M S}{K_M + S} \tag{6.3.27}$$

You can verify that you get the same equation if you assume that $v = k_3[E\text{-}P]$.he Derivation:

This equation looks quite complicated, especially if you substitute for Ks, k_{-1}/k_1. All the kinetic constants can be expressed as functions of the individual rate constants. However, this equation can be simplified by realizing the following:

- When $S >> \frac{K_S k_3}{k_2 + k_3}$, $v = \left(\frac{k_2 k_3}{k_2 + k_3}\right) E_0 = V_M$
- $\frac{K_S k_3}{k_2 + k_3} = \text{constant} = K_M$

Substituting these into equation 7 gives:

$$v = \frac{V_M S}{K_M + S}$$

This again is the general form of the Michaelis-Menten equation

The expression for V_M in the first bulleted expression above is more complicated than our earlier definition of $V_M = k_3E_0$. They are similar in that the term E_0 is multiplied by a constant which is itself a function of rate constant(s). The rate constants are generally lumped together into a generic constant called k_{cat}.

- For the simple reaction $\mathbf{k_{cat} = k3}$
- For the more complicated reaction above with a covalent intermediate, $k_{cat} = \frac{k_2 k_3}{k_2 + k_3}$
- For all reactions, $V_M = k_{cat}E_0$.

Figure 6.3.4 compares the Michaelis-Menten kinetic equations for the rapid equilibrium, steady state assumptions, and covalent intermediate cases .

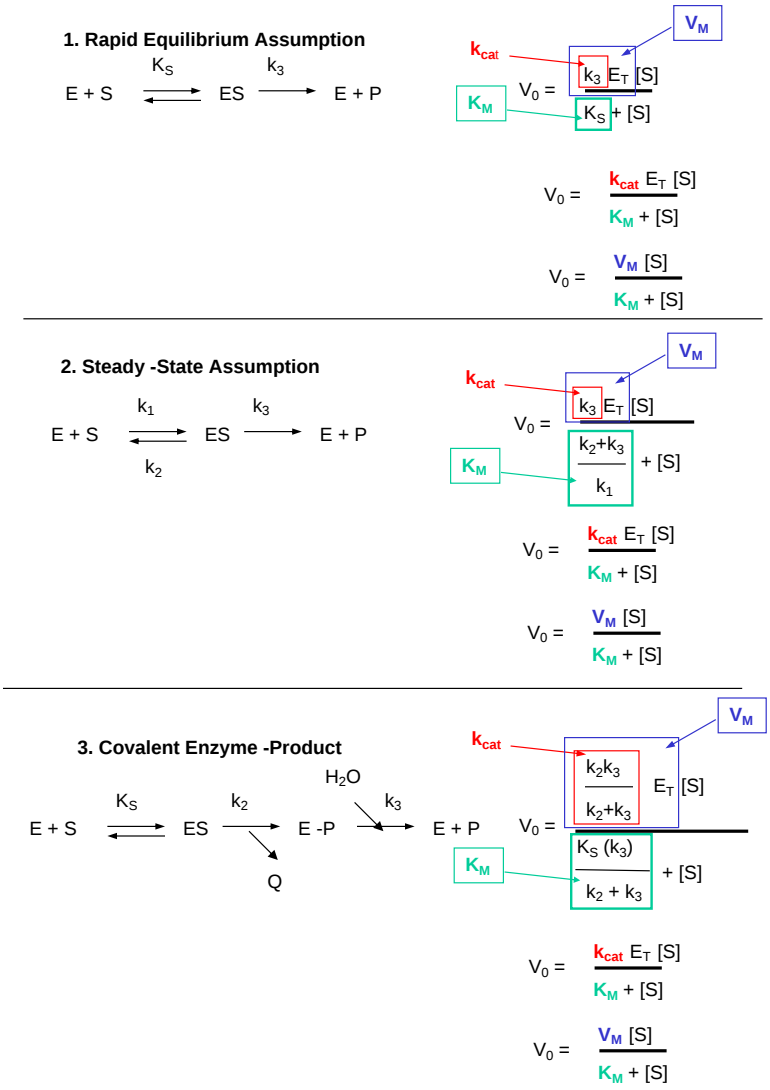


Figure 6.3.4: Michaelis-Menten kinetic equations for the rapid equilibrium, steady state assumptions, and covalent intermediate cases

6.3.5: Meaning of Kinetic Constants

It is important to get a "gut-level" understanding of the significance of the rate constants. Here they are:

- K_M: The Michaelis constant with units of molarity (M), is operationally defined as the substrate concentration at which the initial velocity is half of V_M. It is equal to the dissociation constant of E and S only in if E, S and ES are in rapid equilibrium. It can be thought of as an "**effective**" (but not actual) K_D in other cases.
- k_{cat}: The catalytic rate constants, with units of s^{-1} is often called the **turnover number**. It is a measure of how many bound substrate molecules "turnover" or form product in 1 second. This is evident from equation $v_0 = k_{cat}[ES]$
- k_{cat}/K_M: Under condition when [S] << K_M, the Michaelis-Menten equation becomes $v_0 = (k_{cat}/K_M)[E_0][S]$. This really describes a biomolecular rate constant (k_{cat}/K_M), with units of $M^{-1}s^{-1}$, for conversion of free substrate to product. Some enzymes have k_{cat}/Km values around 10^8, indicating that they are diffusion controlled. That implies that the reaction is essentially done as soon as the enzyme and substrate collide. The constant k_{cat}/Km is also referred to as the **specificity constant** in that it describes how well an enzyme can differentiate between two different competing substrates. (We will show this mathematically in the next chapter.)

Table 6.3.1 below shows K_M and k_{cat} values for various enzymes

K_M values		
enzyme	substrate	Km (mM)
catalase	H_2O_2	25
hexokinase (brain)	ATP	0.4
	D-Glucose	0.05
	D-Fructose	1.5
carbonic anhydrase	HCO_3^-	9
chymotrypsin	glycyltyrosinylglycine	108
	N-benzoyltyrosinamide	2.5
b-galactosidase	D-lactose	4.0
threonine dehydratase	L-Thr	5.0
k_{cat} values		
enzyme	substrate	k_{cat} (s^{-1})
catalase	H_2O_2	40,000,000
carbonic anhydrase	HCO_3^-	400,000
acetylcholinesterase	acetylcholine	140,000
b-lactamase	benzylpenicillin	2,000
fumarase	fumarate	800
RecA protein (ATPase)	ATP	0.4

Table 6.3.1: K_M and k_{cat} values for various enzymes

Table 6.3.2 below show k_{cat}, K_M and k_{cat}/K_M values for diffusion-controlled enzymes

Enzymes with k_{cat}/K_M values close to diffusion controlled (10^8 - 10^9 $M^{-1}s^{-1}$)					
enzyme	substrate		kcat (s^{-1})	Km (M)	kcat/Km ($M^{-1}s^{-1}$)
acetylcholinesterase	acetylcholine		1.4×10^4	9×10^{-5}	1.6×10^8
carbonic anhydrase	CO_2		1×10^6	1.2×10^{-2}	8.3×10^7
	HCO_3^-		4×10^5	2.6 x 10-2	1.5×10^7
catalase	H_2O_2		4×10^7	1.1	4×10^7
crotonase	crotonyl-CoA		5.7×10^3	2×10^{-5}	2.8×10^8
fumarase	fumarate		8×10^2	5×10^{-6}	1.6×10^8
	malate		9×10^2	2.5×10^{-5}	3.6×10^7
triose phosphate isomerase	glyceraldehyde-3-P		4.3×10^3	4.7×10^{-4}	2.4×10^8
b-lactamase	benzylpenicillin		2.0×10^3	2×10^{-4}	1×10^8

Table 6.3.2 below show k_{cat}, K_M and k_{cat}/K_M values for diffusion controlled enzymes

6.3.6: Experimental Determination of V_M and K_M

How can V_M and K_M be determined from experimental data?

6.3.6.1: From initial rate data

The most common way to determine V_M and K_M is through initial rates, v_0, obtained from P or S vs time curves. Hyperbolic graphs of v_0 vs [S] can be fitted or transformed as we explored with the different mathematical transformations of the hyperbolic binding equation to determine K_D. These included:

- Michaelis-Menten plot: nonlinear hyperbolic fit
- Lineweaver-Burk double reciprocal plot
- Scatchard plot
- Eadie-Hofstee plot

We discussed all of these plots except for the **Eadie-Hofstee** plot, in the chapter on binding. The Eadie-Hofstee plot is another linearized version of Michaelis-Menten equation

Here is a derivation of that equation, which starts with each side of the double-reciprocal plot being multiplied by v_0V_M.

$$\begin{aligned} (v_0 V_M)\,\frac{1}{v_0} &= (v_0 V_M)\left(\frac{K_M}{V_M}\right)\frac{1}{S} + (v_0 V_M)\,\frac{1}{V_M} \\ V_M &= (v_0)\,K_M\,\frac{1}{S} + (v_0) \\ v_0 &= -K_M\left(\frac{v_0}{S}\right) + V_M \end{aligned} \tag{6.3.28}$$

Note that a graph of v_0 vs v_0/S is linear so slopes and intercepts can be used to obtain values for V_M and K_M.

The double-reciprocal plot is commonly used to analyze initial velocity vs substrate concentration data. When used for such purposes, the graphs are referred to as **Lineweaver-Burk** plots, where plots of 1/v vs 1/S are straight lines with slope m = K_M/V_M, and y-intercept b = $1/V_M$. Figure 6.3.5 common graphs used to display initial rate enzyme kinetic data.

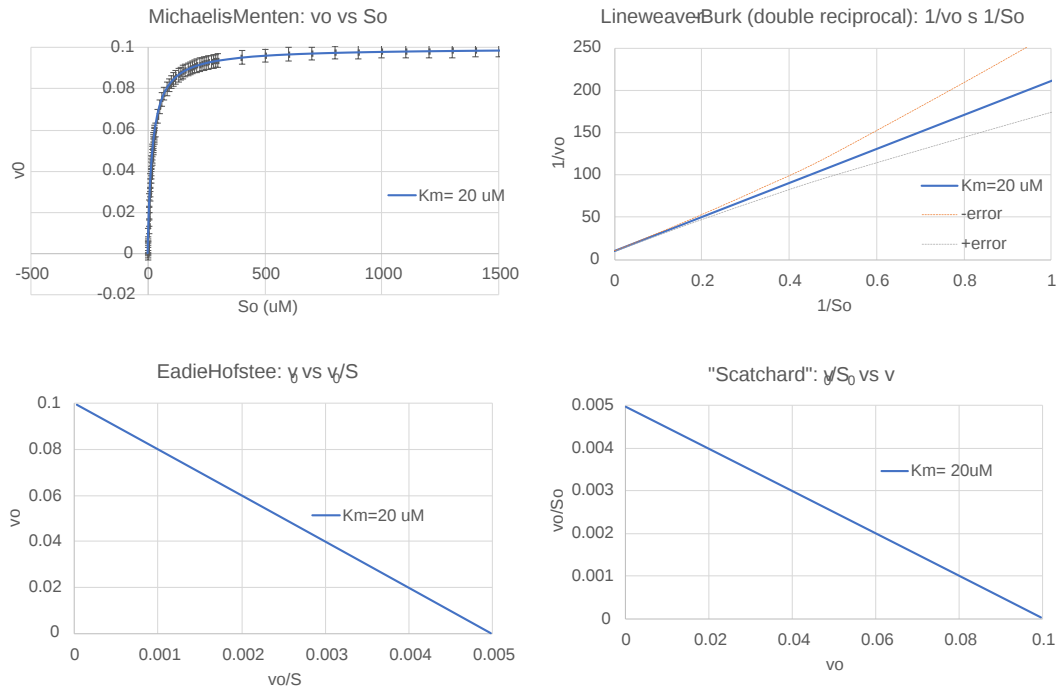


Figure 6.3.5: Common graphs used to display initial rate enzyme kinetic data

The straight-line plots shown above should **not** be analyzed using linear regression, since simple linear regression assumes constant error in v_0 values. A weighted linear regression or even better, a nonlinear fit to a hyperbolic equation should be used. (Common Error in Biochemistry Textbooks: The Shape of the Hyperbola). A rearrangement of the corresponding Scatchard equations in the Eadie-Hofstee plot is also commonly used, especially to visualize enzyme inhibition data as we will see in the next chapter.

6.3.7: An Extension: k_{cat} and V_M from integrated rate equations

K_M and V_M could be theoretically extracted from progress curves of A or P as a function of t at one single A concentration by deriving an integrated rate equation for A or P as a function of t, as we did in equation 2 (the integrated rate equation for the conversion of A → P in the absence of enzyme). In principle, this method would be better than the initial rates methods. Why? It is not easy to be certain you are measuring the initial rate for each and every [S] which should vary over a wide range. It's also time intensive. In addition, think how much data is discarded if you take an entire progress curve at each substrate concentration, especially if you quench the reaction at a given time point, which effectively limits the data to one time point per substrate.

In practice, the mathematics is complicated and it is not possible to get a simple explicit function of [P] or [S] as a function of time. A slight variant of a progress curve can be derived. Let us consider the simple case of a single substrate S (or A) being converted to product P in an enzyme-catalyzed reaction. The analogous equations for first-order, noncatalyzed rates were $A=A_0e^{-k_1t}$ or $P = A_0(1-e^{-k_1t})$.

We can derive the equation for the enzyme-catalyzed reaction shown below.

$$E + S \underset{k_2}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_3} E + P$$

Here it is!

$$\frac{P}{t} = \frac{K_M \ln\left(\frac{S_0 - P}{S_0}\right)}{t} + V_M \tag{6.3.29}$$

? Derivation: k_{cat} and V_M from integrated rate equation

Click below to see the derivation

Derivation

$$v = -\frac{dS}{dt} = +\frac{dP}{dt} = \frac{V_M S}{K_M + S} \tag{6.3.30}$$

$$\int_{S_0}^{S} \frac{K_M + S}{V_M S} dS = -\int_0^t t$$

$$-t = \frac{S + K_M \ln S - S_0 - K_M \ln S_0}{V_M} \tag{6.3.31}$$

On rearrangement, this gives:

$$S_0 - S + K_M \ln \frac{S_0}{S} = V_M t \tag{6.3.32}$$

This equation is an implicit equation, not an explicit one, as it does NOT give S(t) explicitly as a function of t.

Equation yy can be written with respect to product P as follows:

$$P = S_0 - S'' \text{ or } ''S = S_0 - P$$

$$S_0 - (S_0 - P) + K_M \ln \frac{S_0}{S_0 - P} = V_M t \tag{6.3.33}$$

$$P - K_M \ln\left(\frac{S_0 - P}{S_0}\right) = V_M t$$

Rearranging this gives

$$\frac{P}{t} = \frac{K_M \ln\left(\frac{S_0 - P}{S_0}\right)}{t} + V_M \tag{6.3.34}$$

S(t) explicitly as a function of t.

This equation does not give P(t) explicitly as a function of time. Rather one can get a graph of P/t vs [ln (1-P/S_0)]/t (shown below) from the derived equation, which *does* give a straight line with a slope of Km and a y-intercept of Vm. Note that the calculated values of V_M and K_M are derived from only one substrate concentration, and the values may be affected by product inhibition.

Figure 6.3.6 shows the comparison of a first-order noncatalyzed conversion of A → P to the enzyme-catalyzed rate. The V_M for the enzyme-catalyzed reaction was chosen to be small to make the two graph comparable.

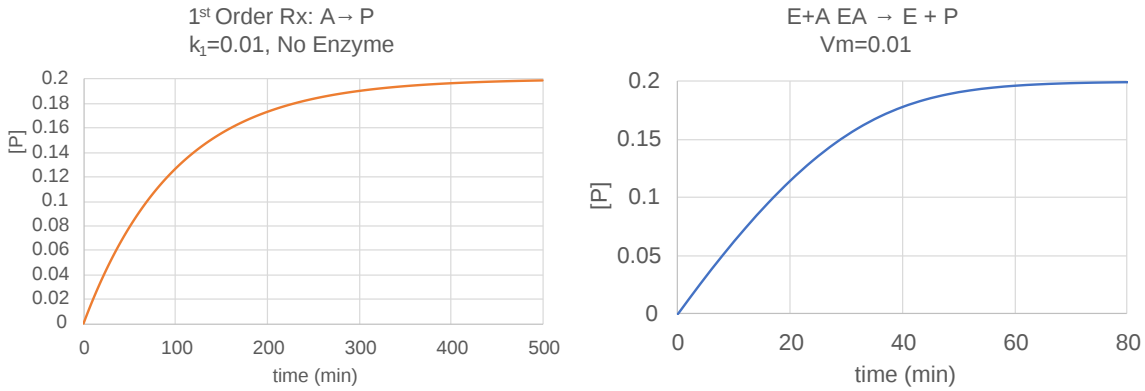


Figure 6.3.6: Comparison of a first order noncatalyzed conversion of A → P to the enzyme catalyzed rate

Note that the curves are similar but not identical. If you didn't know an enzyme was present, you could fit the data to a first-order rise in [P] with time, but it would not be the optimal fit. The progress curves are a lot more complicated to analyze if the product, which shares structural similarities with the substrate, binds the enzyme tightly and inhibits it (called product inhibition).

6.3.7.1: Comparison of progress Curves of uncatalyzed and catalyzed reactions.

Let's explore progress curves for enzyme-catalyzed reactions a bit further. Students usually see v_0 vs [S] Michaelis-Menten plots in textbooks. These plots are in some ways less intuitive than seeing P vs t curves, which are more in line with how we might contemplate how a reaction proceeds. Hence, it would be illuminating to compare progress curve graphs of A → P (irreversible) for the uncatalyzed and S → P for the enzyme-catalyzed reactions. What might you expect? We saw one example in Figure 6.3.6.

In each case, P should increase with time. In the uncatalyzed reaction, S exponentially decreases to 0 and P concomitantly rises to a value of P = S_0. You also get a rise in P vs t for the enzyme-catalyzed rate, but you would think it would be a faster rise with time since the enzyme catalyze the reaction. Figure 6.3.7 shows a comparison of the progress curves for the uncatalyzed first-order reaction of A → P_1 (red) and S → P_2 (blue) for enzym- catalyzed reaction (blue, right) for these conditions: uncatalyzed reaction A → P_1, k_1 = 0.1; Catalyzed reaction: S → P_2, V_M=10, K_M=5. Note that the rate at which bound S (i.e ES) goes to P for the catalyzed rate is 100x faster than the rate constant for the catalyzed rate.

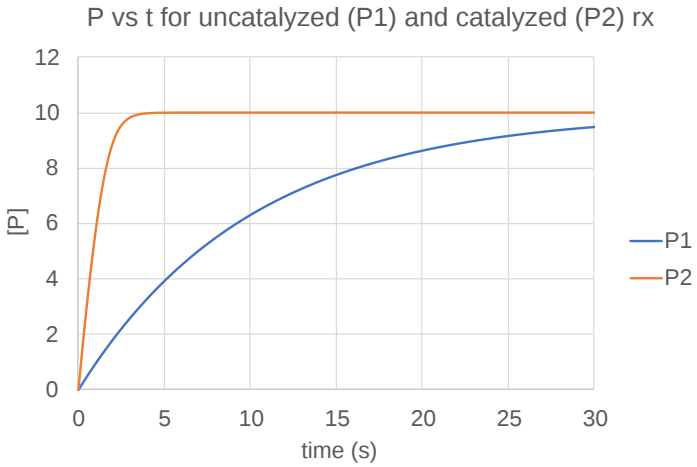


Figure 6.3.7: Another comparison of progress curves (P vs t) for a catalyzed and uncatalyzed conversion of A → P.

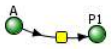
Note that the curves are somewhat similar in shape but also clearly different in comparison to the one shown in Figure 6.3.6.

Now let's use Vcell to compare the reactions for different values for the kinetic constants for the uncatalyzed and the enzyme-catalyzed reaction. Change the constants and try to find a set of conditions so that the catalyzed and catalyzed rates for conversion of reactions to products are superimposable. How can that be?

VCell MODEL

Irreversible reactions: A → P1 and E + S ↔ ES → E + P

Vcell reaction diagram (1-way arrows defined as reversible in actual mathematical model) and chemical equation



Initial conditions

J	reaction rate	(Kf · A - Kr · P1)
Kf	forward rate constant	0.1
Kr	reverse rate constant	0.0
J	reaction rate	$\frac{Vmax \cdot S}{(Km + S)}$
Km	Km (1/2 max)	5.0
Vmax	max reaction rate	10.0
Size	size	1.0
initConc	initial concentration for A	10.0
initConc	initial concentration for P1	0.0
initConc	initial concentration for S	10.0
initConc	initial concentration for P2	0.0

Select Load [model name] below

Load Compare Irrev Rx No EnzW ithEnz

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

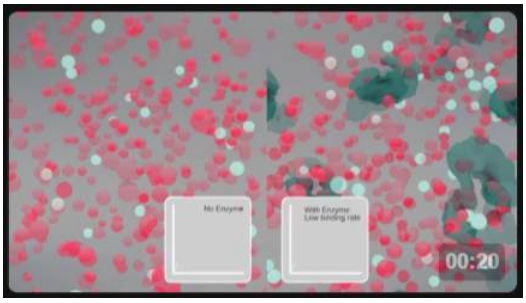
Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

Animations

Now let's look at an animation of the same irreversible reactions in which enzyme-catalyzed reaction is no faster than the noncatalyzed rate (a worthless enzyme!). Here are the reactions:

$A \rightarrow P$, $k_1 = 0.1$; $S \rightarrow P$, $V_M = ?$, $K_M = ?$.

The animations show just the accumulation of product. Animations are by Shraddha Nakak and Hui Liu.



? Exercise 6.3.1

Use the Vcell model above to find a set of values for K_M and V_M that would make the two graphs superimposable - i.e. when the graphs in the absence and presence of E are identical. (Hint: that would be a really bad enzyme if it didn't increase the reaction over the uncatalyzed rate!)

Answer

$K_M = 96$, $V_M = 10$

6.4: Enzyme Inhibition

6.4.1: Irreversible Covalent Inhibition

Given what you already know about protein structure, it should be easy to determine how to inhibit an enzyme. Since structure mediates function, anything that would significantly alter the structure of an enzyme would inhibit the activity of the enzyme. Hence extremes of pH and high temperature, all of which can denature the enzyme, would irreversibly inhibit the enzyme unless it could refold properly. Alternatively, we could add a small molecule, which interacts noncovalently with the enzyme to either change its conformation or directly prevent substrate binding. Finally, we could covalently modify certain side chains, that if they are essential to enzymatic activity, would irreversibly inhibit the enzyme.

We discussed previously the types of reagents that would chemically modify specific side chains that might be critical for enzymatic activity. For example, iodoacetamide might abolish enzyme activity if a cysteine side chain is required for activity. These reagents will usually modify several side chains, however. Determining which is critical for binding or catalytic conversion of the substrate can be difficult. One way would be to protect the active site with saturating concentrations of a ligand that binds reversibly at the active site. Then the chemical modification can be performed at varying reaction times. The critical side chain would be protected from the chemical modification with the extent of protection depending on the K_D, the concentration of the protecting ligand., and the length of the reaction.

The rest of the chapter will deal with reversible, noncovalent inhibition

6.4.2: Competitive Inhibition

Reversible Competitive inhibition occurs when substrate (S) and inhibitor (I) both bind to the same site on the enzyme. In effect, they compete for the active site and bind in a mutually exclusive fashion. This is illustrated in the chemical equations and molecular cartoons shown in Figure 6.4.1.

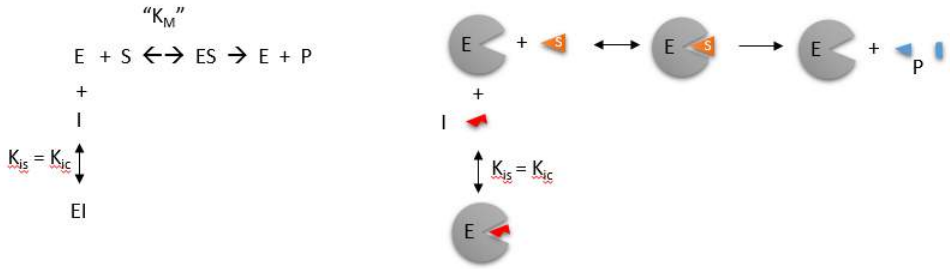


Figure 6.4.1: Competitive Inhibition

$$v_0 = \frac{V_M S}{K_M \left(1 + \frac{I}{Kis}\right) + S} \tag{6.4.1}$$

There is another type of inhibition that would give the same kinetic data. If S and I bound to different sites, and S bound to E and produced a conformational change in E such that I could not bind (and vice versa), then the binding of S and I would be mutually exclusive. This is called allosteric competitive inhibition. Inhibition studies are usually done at several fixed and non-saturating concentrations of I and varying S concentrations.

The key kinetic parameters to understand are V_M and K_M. Let us assume for ease of equation derivation that I binds reversibly, and with rapid equilibrium to E, with a dissociation constant K_{IS}. The "s" in the subscript "is" indicates that the slope of the 1/v vs 1/S Lineweaver-Burk plot changes while the y-intercept stays constant. K_{IS} is also named K_{IC} where the subscript "c" stands for competitive inhibition constant.

A look at the top mechanism shows that even in the presence of I, as S increases to infinity, all E is converted to ES. That is, there is no free E to which I could bind. Now, remember that $V_M = k_{cat}E_0$. Under these conditions, ES = E_0; hence v = V_M. V_M is not changed. However, the apparent K_M, K_{Mapp}, will change. We can use LaChatelier's principle to understand this. If I binds to E alone and not ES, it will shift the equilibrium of E + S → ES to the left. This would increase the K_{Mapp} (i.e. it would appear that the affinity of E and S has decreased.). The double reciprocal plot (Lineweaver-Burk plot) offers a great way to visualize the inhibition as shown in Figure 6.4.2.

$$v_0 = \frac{V_M S}{K_M\left(1 + \frac{I}{Kis}\right) + S}$$

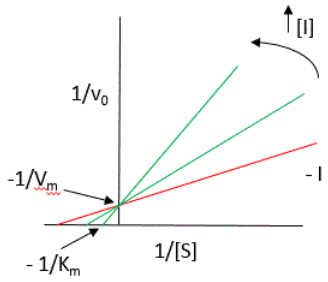
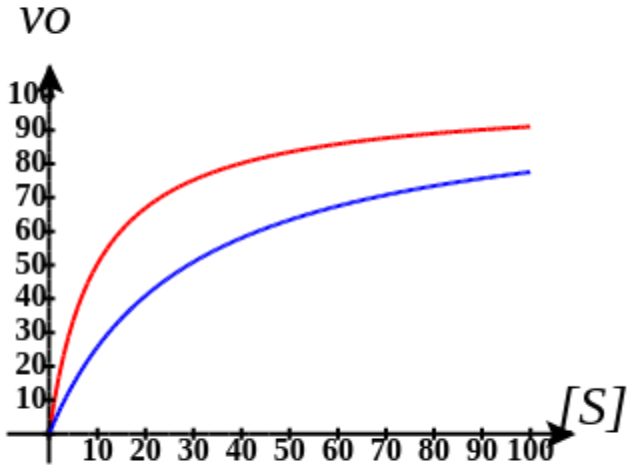


Figure 6.4.2: Competitive Inhibition: Lineweaver-Burk plots

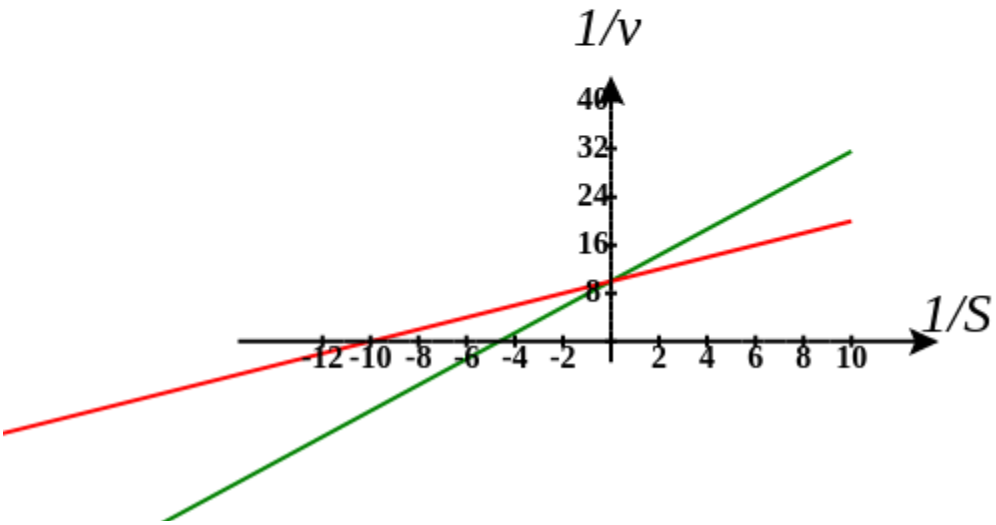
In the presence of I, V_M does not change, but K_M appears to increase. Therefore, $1/K_M$, the x-intercept on the plot will get smaller, and closer to 0. Therefore the plots will consist of a series of lines, with the same y-intercept ($1/V_M$), and the x-intercepts ($-1/K_M$) closer and closer to 0 as I increases. These intersecting plots are the hallmark of competitive inhibition.

Here is an interactive graph showing v_0 vs [S] for competitive inhibition with Vm and Km both set to 100. Change the sliders for [I] and Kis and see the effect on the graph.



Here is the interactive graphs showing $1/v_0$ vs 1/[S] for competitive inhibition, with Vm and Km both set to 10.

Competitive Inhibition



Note that in the first three inhibition models discussed in this section, the Lineweaver-Burk plots are linear in the presence and absence of an inhibitor. This suggests that plots of v vs S in each case would be hyperbolic and conform to the usual form of the Michaelis Menton equation, each with potentially different apparent V_M and K_M values.

An equation for v_0 in the presence of a competitive inhibitor is shown in the above figure. The only change compared to the equation for the initial velocity in the absence of the inhibitor is that the K_M term is multiplied by the factor $1+I/K_{is}$. Hence $K_{Mapp} = K_M(1+I/Kis)$. This shows that the apparent K_M does increase as we predicted. K_{IS} is the inhibitor dissociation constant in which the inhibitor affects the slope of the double reciprocal plot.

If the data were plotted as v_0 vs log S, the plots would be sigmoidal, as we saw for plots of ML vs log L in Chapter 5B. In the case of a competitive inhibitor, the plot of v_0 vs log S in the presence of different fixed concentrations of inhibitor would consist of a series of sigmoidal curves, each with the same V_M, but with different apparent K_M values (where $K_{Mapp} = K_M(1+I/K_{is})$, progressively shifted to the right. Enzyme kinetic data is rarely plotted this way. These plots are mostly used for simple binding data for the M + L ↔ ML equilibrium, in the presence of different inhibitor concentrations.

Reconsider our discussion of the simple binding equilibrium, M + L ↔ ML. For fractional saturation Y vs a log L graphs, we considered three examples:

1. $L = 0.01\ K_D$ (i.e. $L << K_D$), which implies that $K_D = 100L$. Then $Y = L/[K_D+L] = L/[100L + L] \approx 1/100$. This implies that irrespective of the actual [L], if $L = 0.01\ K_D$, then $Y \approx 0.01$.
2. $L = 100\ K_D$ (i.e. $L >> K_D$), which implies that $K_D = L/100$. Then $Y = L/[K_D+L] = L/[(L/100) + L] = 100L/101L \approx 1$. This implies that irrespective of the actual [L], if L = 100 Kd, then Y ≈1.
3. $L = K_D$, then Y = 0.5

These scenarios show that if L varies over 4 orders of magnitude ($0.01K_D < K_D < 100K_D$), or, in log terms, from $-2 + \log K_D < \log K_D < 2 + \log Kd$), irrespective of the magnitude of the K_D, that Y varies from approximately 0 - 1.

In other words, Y varies from 0-1 when L varies from log K_D by +2. Hence, plots of Y vs log L for a series of binding reactions of increasingly higher K_D (lower affinity) would reveal a series of identical sigmoidal curves shifted progressively to the right, as shown below in Figure 6.4.3.

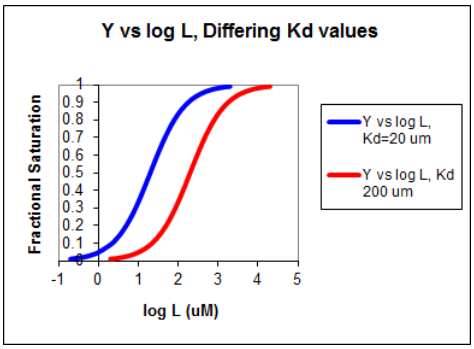


Figure 6.4.3: Plot of fractional saturation Y vs log L for two different K_D values.

The same would be true of v_0 vs S in the presence of different concentrations of a competitive inhibitor, for initial flux, Jo vs ligand outside, in the presence of a competitive inhibitor, or ML vs L (or Y vs L) in the presence of a competitive inhibitor.

In many ways plots of v_0 vs lnS are easier to visually interpret than plots of v_0 vs S . As noted for simple binding plots, textbook illustrations of hyperbolas are often misdrawn, showing curves that level off too quickly as a function of [S] as compared to plots of v_0 vs lnS, in which it is easy to see if saturation has been achieved. In addition, as the curves above show, multiple complete plots of v_0 vs lnS at varying fixed inhibitor concentrations or for variant enzyme forms (different isoforms, site-specific mutants) over a broad range of lnS can be made which facilitates comparisons of the experimental kinetics under these different conditions. This is especially true if Km values differ widely.

Now that you are more familiar with binding and enzyme kinetics curves, in the presence and absence of inhibitors, you should be able to apply the above analysis to inhibition curves where the binding or the initial velocity is plotted at varying competitive inhibitor concentrations at different fixed nonsaturating concentrations of ligand or substrate. Consider the activity of an enzyme. Let's say that at some reasonable concentration of substrate (not infinite), the enzyme is approximately 100% active. If a competitive inhibitor is added, the activity of the enzyme decreases until at saturating (infinite) I, no activity would remain. Graphs showing this are shown below in Figure 6.4.4.

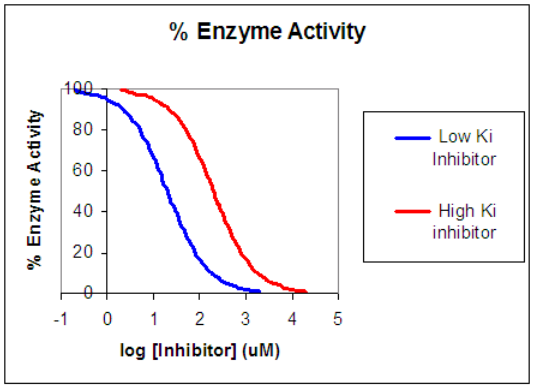


Figure 6.4.4: Inhibition of Enzyme Activity - % Activity vs log [Inhibitor]

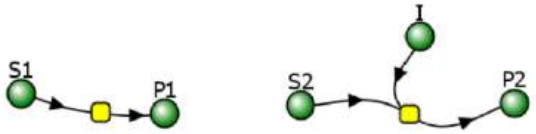
6.4.2.1: Progress Curves for Competitive Inhibition

In the previous section, we explored how important progress curve (Product vs time) analyses are in understanding both uncatalyzed and enzyme-catalyzed reactions. We are aware of no textbooks which cover progress curves for enzyme inhibition. Yet progress curves are what most investigators record and analyze to determine initial rates v_0 and to calculate V_M, K_M and inhibition constants, as described above. We will use Vcell to produce progress curves for reversibly inhibited enzyme-catalyzed reactions.

VCell MODEL

Competitive Inhibition with constant [I]:

No inhibition (left) and competitive inhibition (right)



Initial conditions for no inhibition

Name	Description	Global	
J	reaction rate	☐	$\frac{Vmax \cdot S1}{(Km + S1)}$
Km	Km (1/2 max)	☐	5.0
Vmax	max reaction rate	☐	10.0
S1	Species Concentration	☑	Variable

Initial conditions for competitive inhibition

Name	Description	Global	
J	reaction rate	☐	$\frac{Vmax \cdot S2}{(Km + S2 + \frac{Km \cdot I}{Ki})}$
Km	user defined	☐	5.0
Vmax	user defined	☐	10.0
Ki	user defined	☐	5.0
S2	Species Concentration	☑	Variable
I	Species Concentration	☑	Variable

I is fixed for each simulation (as it is not converted to a product) but can be changed in the simulation below.

Select Load [model name] below

Load NoInhib_CompInhib_MM_I

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed. For this model, select Vm, Km, Ki and I | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

The graphs from your initial run show the concentrations of S, P and I as a function of time for just the initial conditions shown above. In typical initial rate laboratory analyzes, of competitive inhibition, at least three sets of reactions are run with the same varying substrate concentrations and different fixed concentrations of inhibitor. In the analyses above, [I] is fixed at 5 uM.

Conduct a series of run at different values of I. Vary the K_I, the dissociation constant for the EI complex, as follows:

- I << K_I, the dissociation constant for the EI complex
- I >> K_I, the dissociation constant for the EI complex. Then download the data and determine the initial rate for each of the initial conditions.

Figure 6.4.5 shows an interactive iCn3D model of human low molecular weight phosphotyrosyl phosphatase bound to a competitive inhibitor (5PNT)

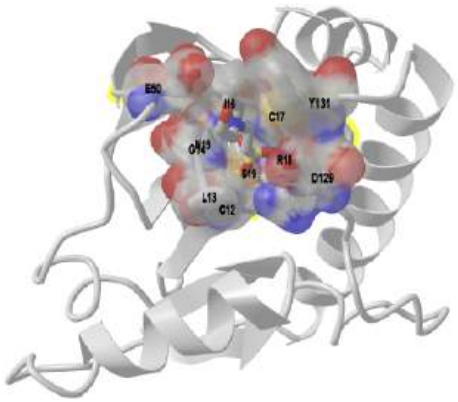


Figure 6.4.5: Human low molecular weight phosphotyrosyl phosphatase bound to a competitive inhibitor (5PNT). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...XsEacG2tixDDi9

The competitive inhibitor, the deprotonated form of 2-(N-morpholino)-ethanesulfonic acid (MES), is actually the conjugate base of the weak acid (pKa = 6.15) of a commonly used component of a buffered solution. It is shown in color sticks with the negatively charged sulfonate sitting at the bottom of the active site pocket. The amino acids comprising the active site binding pocket are shown as color sticks underneath the transparent colored surface of the binding pocket. The normal substrates for the enzyme are proteins phosphorylated on tyrosine side chains so the sulfonate is a mimic of the negatively charged phosphate group of the phosphoprotein target.

6.4.2.2: Two specials cases of competition inhibition

Product Inhibition

Let's look at an enzyme that converts reactant S to product P. Since P arises from S, they may have structural similarities. For example, what if GTP was the reactant and GDP was a product? If so, then P might also bind in the active site and inhibit the conversion of S to P. This is called **product inhibition**. It probably occurs in most enzymes, and when it does occur it will start bending downward the beginning part of the progress curve for P formation. If the product binds very tightly, it might cause a significant underestimation of the initial velocity (v_0) or flux (J_0) of the enzyme. Let's use Vcell to explore product inhibition. The model will explore two reactions:

- E + R ↔ ER → E + Q (no product inhibition)
- E + S ↔ ES → E + P (with product inhibition)

Note that the chemical equation above does not explicitly show the product P binding the enzyme to form an EP complex. An actual reaction diagram showing the inhibition of an enzyme by an inhibitor I and by the product P is shown in Figure 6.4.6 below.

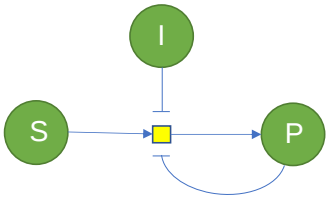


Figure 6.4.6: reaction diagram showing inhibition of an enzyme by an inhibitor I and by the product P

Vcell uses much simpler diagrams since it is most often used for modeling whole pathways or even entire cells. In the simpler Vcell reaction diagrams, the inhibitor is typically not shown since the inhibition is built into the equation for the enzyme, represented by the node or yellow square in the figure above.

Let's now explore product inhibition in Vcell. R and Q are the reactant and product, respectively, in the reaction *without* product inhibition. S and P are used for the reaction *with* product P inhibition.

VCell MODEL

Irreversible MM Kinetics - Without (left rx 1) and With (right, rx 2) Product Inhibition

Initial Conditions: No product inhibition

Name	Description	Global	
J	reaction rate	☐	$\frac{Vmax \cdot S1}{(Km + S1)}$
Km	Km (1/2 max)	☐	5.0
Vmax	max reaction rate	☐	10.0
S1	Species Concentration	☑	Variable

Initial Conditions: With product inhibition

Name	Description	Global	
J	reaction rate	☐	$\frac{Vmax \cdot S2}{(Km + S2 + \frac{Km \cdot P2}{Ki})}$
Km	user defined	☐	5.0
Vmax	user defined	☐	10.0
Ki	user defined	☐	5.0
S2	Species Concentration	☑	Variable
P2	Species Concentration	☑	Variable

Select Load [model name] below

Load ProdInhib_NoProdInhib

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

Inhibition by a competing substrate - the specificity constant

In the previous chapter, the specificity constant was defined as k_{cat}/K_M which we also described as the second-order rate constant associated with the bimolecular reaction of E and S when S << K_M. It also describes how good an enzyme is in differentiating between different substrates. If an enzyme encounters two different substrates, one can be considered to be a competitive inhibitor of the other. The following equation gives the ratio of initial velocities for two competing substrates at the same concentration is equal to the ratio of their k_{cat}/K_M values.

$$\frac{v_A}{v_B} = \frac{\frac{k_{cat\,A}}{K_A}A}{\frac{k_{cat\,B}}{K_B}B} \tag{6.4.2}$$

? A derivation of the specificity constant for an enzyme with competing substrates

Here it is!

Derivation

$$v_A = \frac{V_A A}{K_A\left(1+\frac{B}{K_B}\right)+A} \quad v_B = \frac{V_B B}{K_B\left(1+\frac{A}{K_A}\right)+B} \tag{6.4.3}$$

$$\frac{v_A}{v_B} = \frac{\frac{V_A A}{K_A\left(1+\frac{B}{K_B}\right)+A}}{\frac{V_B B}{K_B\left(1+\frac{A}{K_A}\right)+B}} = \frac{\frac{V_A A}{K_A+\frac{K_A B}{K_B}+A}}{\frac{V_B B}{K_B+\frac{K_B A}{K_A}+B}} \tag{6.4.4}$$

Now in the above equation:

multiple the top half of the right-hand expression by

$$\frac{\frac{1}{K_A}}{\frac{1}{K_A}} \tag{6.4.5}$$

multiple the bottom half of the right-hand expression by

$$\frac{\frac{1}{K_B}}{\frac{1}{K_B}} \tag{6.4.6}$$

replace V_A with $k_{catA}E_0$ and V_B with $k_{catB}E_0$

This gives the following expression for v_A/v_B:

$$\frac{v_A}{v_B} = \frac{\frac{k_{catA}}{K_A}A}{\frac{k_{catB}}{K_B}B} \tag{6.4.7}$$

6.4.3: Uncompetitive Inhibition

Reversible uncompetitive inhibition occurs when I binds only to ES and not free E. One can hypothesize that on binding S, a conformational change in E occurs which presents a binding site for I. Inhibition occurs since ESI can not form the product. It is a dead-end complex that has only one fate, to return to ES. This is illustrated in the chemical equations and molecular cartoon shown in Figure 6.4.7.

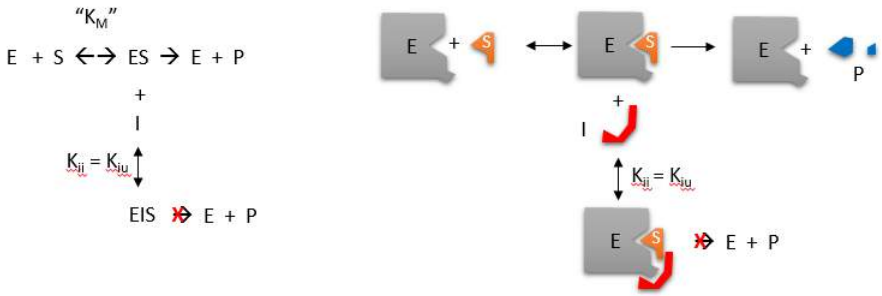


Figure 6.4.7: Uncompetitive Inhibition

$$v_0 = \frac{V_M S}{K_M + S\left(1+\frac{I}{Kii}\right)} = \frac{\left(\frac{V_M}{1+\frac{I}{Kii}}\right)S}{\left(\frac{K_M}{1+\frac{I}{Kii}}\right)+S} \tag{6.4.8}$$

Let us assume that I binds reversibly to ES with a dissociation constant K_{ii}. The second "i" in the subscript "ii" indicates that the intercept of the 1/v vs 1/S Lineweaver-Burk plot changes while the slope stays constant. K_{ii} is also named Kiu, where the subscript "u" stands for the uncompetitive inhibition constant.

A look at the top mechanism shows that in the presence of I, as S increases to infinity, not all of E is converted to ES. That is, there is a finite amount of ESI, even at infinite S. Now remember that $V_m = k_{cat}E_0$ if and only if all E is in the form ES. Under these conditions, the apparent V_m, V_{mapp} is less than the real Vm without the inhibitor. In addition, the apparent K_m, K_{mapp}, will change. We can use LaChatelier's principle to understand this. If I binds to ES alone, and not E, it will shift the equilibrium of E + S ↔ ES to the right, which would have the effect of decreasing the K_{mapp} (i.e. it would appear that the affinity of E and S has increased.). The double reciprocal plot (Lineweaver Burk plot) offers a great way to visualize the inhibition. In the presence of I, both Vm and Km decrease. Therefore, $-1/K_m$, the x-intercept on the plot, will get more negative, and $1/V_m$ will get more positive. It turns out that they change to the same extent. Therefore the plots will consist of a series of parallel lines, which is the hallmark of uncompetitive inhibition, as shown in Figure 6.4.8.

$$v_0 = \frac{V_M S}{K_M + S\left(1 + \frac{I}{Kii}\right)} = \frac{\left(\frac{V_M}{1 + \frac{I}{Kii}}\right) S}{\left(\frac{K_M}{1 + \frac{I}{Kii}}\right) + S}$$

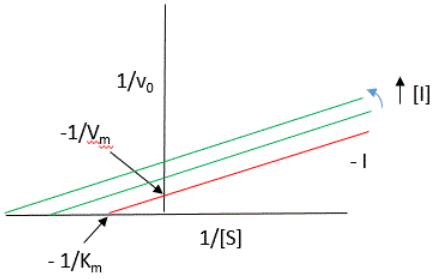
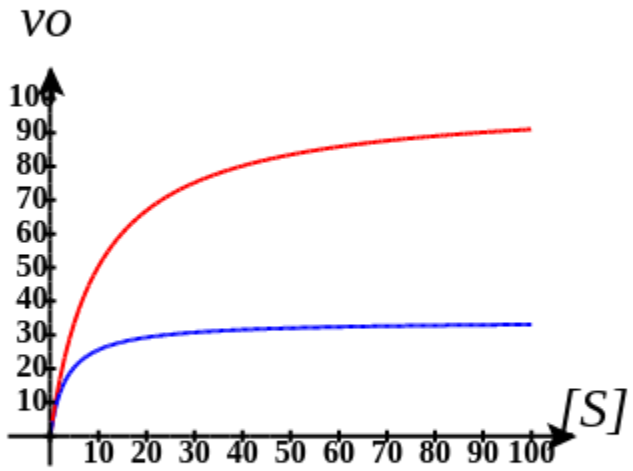


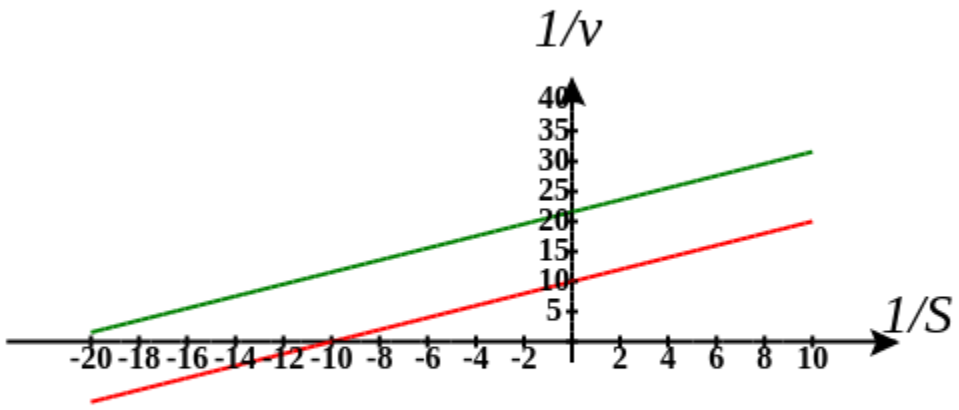
Figure 6.4.8: Uncompetitive Inhibition: Lineweaver-Burk plots

Here is an interactive graph showing v_0 vs [S] for uncompetitive inhibition with Vm and Km both set to 100. Change the sliders for [I] and Kis and see the effect on the graph.



Here is an interactive graph showing uncompetitive inhibition with Vm and Km both set to 10. Change the sliders for [I] and Kii and see the effect on the graph

Uncompetitive Inhibition



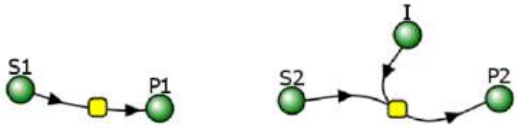
An equation, shown in the diagram above, can be derived which shows the effect of the uncompetitive inhibitor on the velocity of the reaction. The only change is that the S term in the denominator is multiplied by the factor 1+I/Kii. We would like to rearrange this equation to show how Km and Vm are affected by the inhibitor, not S, which obviously isn't. Rearranging the equation above shows that Kmapp = Km/(1+I/Kii) and Vmapp = Vm/(1+I/Kii). This shows that the apparent Km and Vm do decrease as we predicted. Kii is the inhibitor dissociation constant in which the inhibitor affects the intercept of the double reciprocal plot. Note that if I is zero, Km and Vm are unchanged.

6.4.3.1: Progress Curves: Uncompetitive Inhibition

Now let's compare the progress curves for an enzyme-catalyzed reaction in the absence and presence of an uncompetitive inhibitor.

VCell MODEL

No inhibition (left) and Uncompetitive Inhibition (right)



(Note the the Vcell reaction diagram is the same as for competitive inhibition. However, the mathematical equations differ as shown below.

Initial values No Inhibition

Name	Description	Global	
J	reaction rate	☐	$\frac{Vmax \cdot S1}{(Km + S1)}$
Km	Km (1/2 max)	☐	5.0
Vmax	max reaction rate	☐	10.0
S1	Species Concentration	☑	Variable

Initial values With Uncompetitive Inhibitor

Name	Description	Global	
J	reaction rate	☐	$\frac{Vmax \cdot S2}{(Km + S2 \cdot (1.0 + \frac{I}{Ki}))}$
Km	user defined	☐	5.0
Vmax	user defined	☐	10.0
Ki	user defined	☐	5.0
S2	Species Concentration	☑	Variable
I	Species Concentration	☑	Variable

I is fixed for each simulation (as it is not converted to a product) but can be changed in the simulation below.

Select Load [model name] below

Load NoInhib_UncompInhibMM

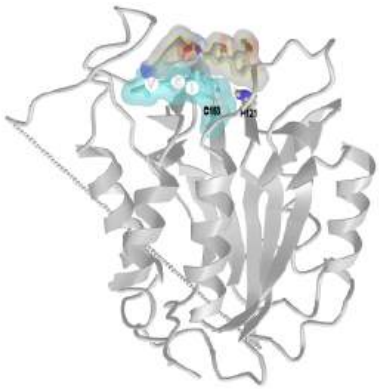
Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

Figure 6.4.9 shows an interactive iCn3D model of an uncompetitive inhibitor of the cysteine protease caspase-6 (4HVA)



NCBI iCn3D Figure 6.4.9: Uncompetitive inhibitor of the cysteine protease caspase-6 (4HVA) . (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...TVfKio7NxcpX28

The "substrate" in this model (cyan, transparent surface, with labels on V, E, and I) is a substrate analog, VEI-CHO in which the tripeptide substrate VEI ends not in a free carboxyl or amide group but an aldehyde, which causes this "substrate" to become covalently attached to the enzyme and act as an inhibitor. The uncompetitive inhibitor (gray transparent surface) binds externally

to the blue surface. Hence it binds to the ES complex. Two active site residues, Cys 163, the active site nucleophile, and His 121, a catalytic acid/base are shown in colored sticks and labeled.

Figure 6.4.10 below shows a Lineweaver-Burk plot showing $1/v_o$ vs 1/[S], where the substrate is the divalent compound $(VEID)_2R110$. Its N-terminus is capped with a benzyloxy (Z) group. R110 is a rhodamine-type fluorophore, which on cleavage gives a strong fluorescent signal that was initially quenched by the benzyloxy group in the uncleaved substrate.

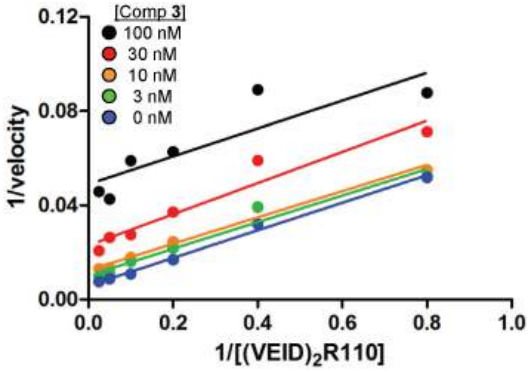


Figure 6.4.10: Double-reciprocal Lineweaver-Burk plot of compound 3 with $(VEID)_2R110$ substrate showing uncompetitive inhibition. Heise CE et al. (2012) Mechanistic and Structural Understanding of Uncompetitive Inhibitors of Caspase-6. PLoS ONE 7(12): e50864. https://doi.org/10.1371/journal.pone.0050864. Creative Commons Attribution License.

6.4.4: Noncompetitive and Mixed Inhibition

Reversible noncompetitive inhibition occurs when I binds to both E and ES. We will look at only the special case in which the dissociation constants of I for E and ES are the same. This is called noncompetitive inhibition. It is quite rare as it would be difficult to imagine a large inhibitor that inhibits the turnover of a bound substrate having no effect on the binding of S to E. However covalent interaction of protons with both E and ES can lead to noncompetitive inhibition. In the more general case, the Kd's are different, and the inhibition is called mixed. Since inhibition occurs, we will hypothesize that ESI can not form the product. It is a dead-end complex that has only one fate, to return to ES or EI. This is illustrated in the chemical equations and in the molecular cartoon in Figure 6.4.11.

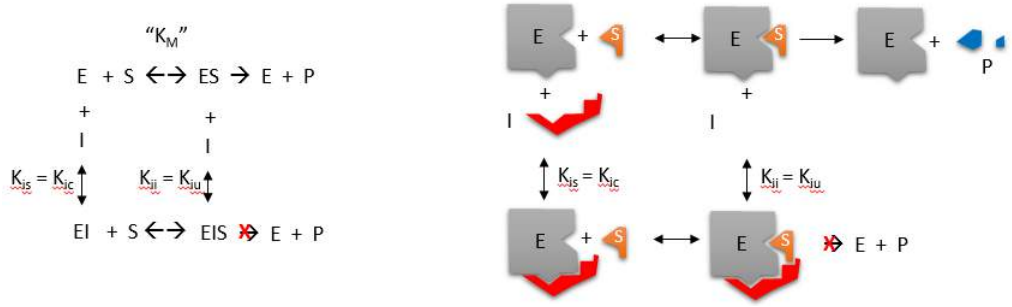


Figure 6.4.11: Noncompetitive and Mixed Inhibition

Let us assume for ease of equation derivation that I binds reversibly to E with a dissociation constant of K_{is} (as we denoted for competitive inhibition) and to ES with a dissociation constant K_{ii} (as we noted for uncompetitive inhibition). Assume for noncompetitive inhibition that $K_{is} = K_{ii}$. A look at the top mechanism shows that in the presence of I, as S increases to infinity, not all of E is converted to ES. That is, there is a finite amount of ESI, even at infinite S. Now remember that $V_m = k_{cat}E_0$ if and only if all E is in the form ES . Under these conditions, the apparent V_m, V_{mapp} is less than the real Vm without an inhibitor. In contrast, the apparent K_m, K_{mapp}, will not change since I binds to both E and ES with the same affinity, and hence will not perturb that equilibrium, as deduced from LaChatelier's principle. The double reciprocal plot (Lineweaver Burk plot) offers a great way to visualize the inhibition. In the presence of I, just V_m will decrease. Therefore, $-1/K_m$, the x-intercept will stay the same, and $1/V_m$ will get more positive. Therefore the plots will consist of a series of lines intersecting on the x-axis, which is the hallmark of noncompetitive inhibition. You should be able to figure out how the plots would appear if K_{is} is different from K_{ii} (mixed inhibition).

$$v_0 = \frac{V_M S}{K_M\left(1+\frac{I}{Kis}\right)+S\left(1+\frac{I}{Kii}\right)} \tag{6.4.9}$$

An equation, shown in the diagram above can be derived which shows the effect of the noncompetitive inhibitor on the velocity of the reaction. In the denominator, Km is multiplied by $1+I/K_{is}$, and S by $1+I/K_{ii}$. We would like to rearrange this equation to show how Km and Vm are affected by the inhibitor, not S, which obviously isn't. Rearranging the equation as shown above shows that $K_{mapp} = K_m(1+I/K_{is})/(1+I/K_{ii}) = K_m$ when $K_{is}=K_{ii}$, and $V_{mapp} = V_m/(1+I/K_{ii})$. This shows that the Km is unchanged and Vm decreases as we predicted. The plot shows a series of lines intersecting on the x-axis as shown in Figure 6.4.12 Both the slope and the y-intercept are changed, which are reflected in the names of the two dissociation constants, K_{is} and K_{ii}. Note that if I is zero, $K_{mapp} = K_m$ and $V_{mapp} = V_m$. Sometimes the K_{is} and K_{ii} inhibition dissociation constants are referred to as K_c and $\mathbf{K_u}$ (competitive and uncompetitive inhibition dissociation constants.

$$v_0 = \frac{V_M S}{K_M\left(1+\frac{I}{KiS}\right)+S\left(1+\frac{I}{Kii}\right)} = \frac{\left(\frac{V_M}{1+\frac{I}{Kii}}\right)S}{K_M+S}$$

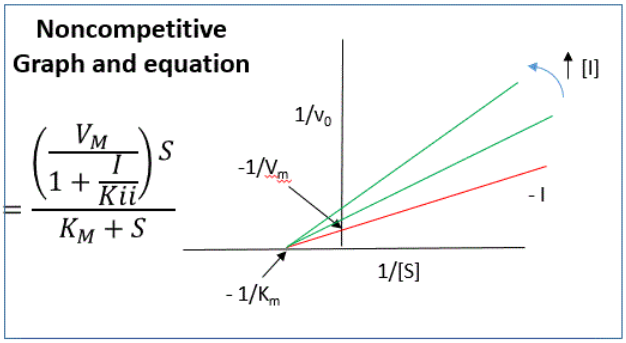
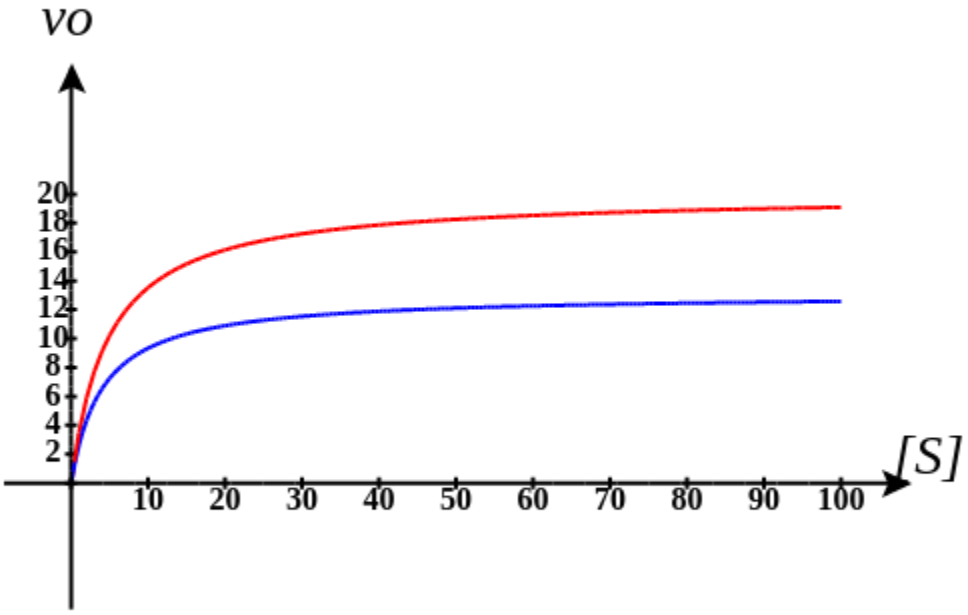


Figure 6.4.12: Noncompetitive Inhibition: Lineweaver-Burk plots

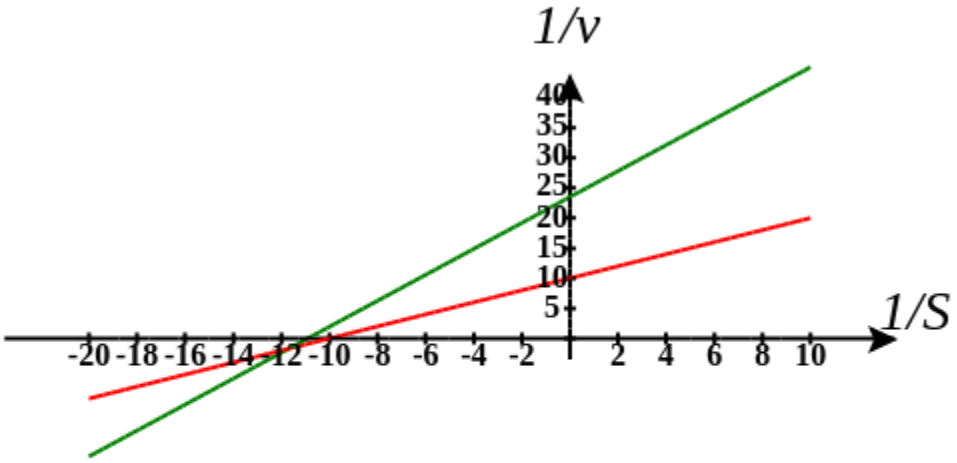
Move the sliders on this interactive graph below to show changes in K_{is} and K_{ii} affect position on the graph where the lines intersect. Try to change their values to move the intersections of the graphs from the left top quadrant to the x-axis to the left bottom quadrant.

Here is an interactive graph showing v_0 vs [S] for mixed inhibition with Vm and Km both set to 100. Change the sliders for [I] and Kis and see the effect on the graph



Here is an interactive graph showing mixed inhibition with Vm and Km both set to 10. Change the sliders for [I] and Kii and see the effect on the graph

Mixed Inhibition

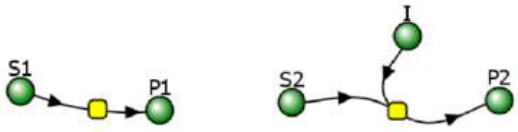


6.4.4.1: Progress Curves: Mixed/Noncompetitive Inhibition

Now let's compare the progress curves for an enzyme-catalyzed reaction in the absence and presence of a mixed inhibitor.

VCell MODEL

No inhibition (left) and Uncompetitive Inhibition (right)



Note that the Vcell reaction diagram is the same as for competitive and uncompetive inhibition. It doesn't explicitly show that the mixed inhibitor binds to both free and substrate-bound enzymes. Those interactions are addressed in the mathematical equations for mixed inhibition.

Initial values No Inhibition

Name	Description	Global	
J	reaction rate	☐	$\frac{Vmax \cdot S1}{(Km + S1)}$
Km	Km (1/2 max)	☐	5.0
Vmax	max reaction rate	☐	10.0
S1	Species Concentration	☑	Variable

Initial values With Uncompetitive Inhibitor

Name	Description	Global	
J	reaction rate		$\frac{Vmax \cdot S2}{(Km \cdot (1.0 + \frac{I}{Kc}) + S2 \cdot (1.0 + \frac{I}{Ki}))}$
Km	user defined	☐	5.0
Vmax	user defined	☐	10.0
Kc	user defined	☐	1.0
Ki	user defined	☐	1.0
S2	Species Concentration	☑	Variable
I	Species Concentration	☑	Variable

I is fixed for each simulation (as it is not converted to a product) but can be changed in the simulation below.

Select Load [model name] below

Load NoInhib_MIXEDInhib_MM

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

Mixed and noncompetitive inhibition (as shown by the mechanism above) differ from competitive and uncompetitive inhibition in that the inhibitor binding is not simply a dead-end reaction in which the inhibitor can only dissociate in a single reverse step. In the above equilibrium, S can dissociate from ESI to form EI so the system may not be at equilibrium. With dead-end steps, no flux of reactants occurs through the dead-end complex so the equilibrium for the dead-end step is not perturbed.

Other mechanisms can commonly give mixed inhibition. For example, the product released in a ping pong mechanism (discussed in the next chapter) can give mixed inhibition. A ping pong reaction mechanism is shown and superficially explained in Figure 6.4.13

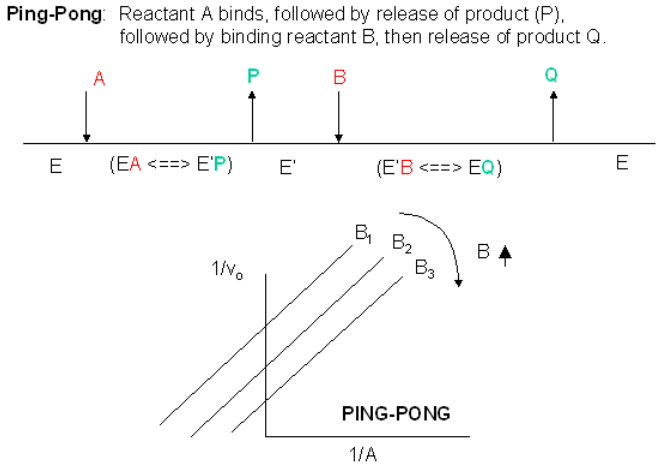


Figure 6.4.13: Ping pong reaction mechanism

If P, acting as a product inhibitor, can bind to two different forms of the enzyme (E' and also E), it will act as a mixed inhibitor.

Figure 6.4.14 shows an interactive iCn3D model of a noncompetitive inhibitor of M. tuberculosis's class IIa fructose 1,6-bisphosphate aldolase (4LV4).

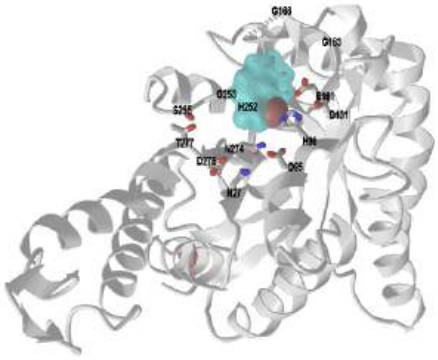


Figure 6.4.14: Noncompetitive inhibitor of M. tuberculosis's class IIa fructose 1,6-bisphosphate aldolase (4LV4). (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...wZDtkvtwqnrVy7

The non-competitive inhibitor, 8-hydroxyquinoline carboxylic acid (HCA), is shown as with a cyan surface. Active site side chains of the Zn^{2+}- containing enzyme are shown as colored sticks and labeled.

The noncompetitive inhibitor does not enter into the pocket where the substrate (not shown in the model above) binds. The binding of the noncompetitive inhibitor causes a conformational change that moves the Zn-binding loop (Z-loop) which contains His 212 which coordinates the Zn^{2+} ion. Hence the Zn-loop is part of the active site and when it moves on binding of the inhibitor, the access to the empty pocket where the substrate binds is hindered. When the Zn^{2+} moves away from the active site, it can no longer engage in catalytic activity.

6.4.5: Enzyme Inhibition *in Vivo*

The pharmaceutical industry is devoted to finding drug molecules that affect biological processes. Typically this means the development of small molecule inhibitors of target proteins. Recent work has expanded to the development of inhibitory RNA molecules that affect DNA transcription and mRNA translation. Using combinatorial synthetic techniques and computational modeling, it has gotten easier to develop small molecule inhibitors (especially competitive ones) that inhibit proteins in vitro using purified enzymes, substrates, and inhibitors in lab testing. Assuming that the inhibitor could pass through the membrane and accumulate to a sufficient enough concentration, would it have the same inhibitory properties in the cell as in the test tube? The answer turns out to be maybe. Remember that a cell is tightly packed with a multitude of other small molecules and macromolecules. In addition, the enzyme targeted for inhibition is most likely part of a pathway of enzymes that feeds reactant into the enzyme and removes the product. Hence, the flux of substrate and product is controlled by the entire pathway and not just the single target enzyme. The concentration of the product of the target enzyme is determined by kinetic parameters for the enzyme and available substrate concentration.

The conditions under which the enzymes are studied (in vitro) and operate (in vivo) are very different.

- In vitro (in the lab), the enzyme is held at a constant concentration while the substrate is varied (i.e the substrate concentration is the independent variable). The velocity is determined by the substrate concentration. When inhibition is studied, the substrate is varied while the inhibitor is held constant at several different fixed concentrations.
- In vivo (in the cell), the velocity might be held at a relatively fixed level with the substrate determined by the velocity. To avoid a bottleneck in flux, the substrate can't build up at the enzyme, so the enzyme processes it in a steady state fashion to produce the product as determined by the Michael-Menten equation.

What happens when an inhibitor is added *in vivo*? Let's assume that the enzyme is running at $v = V_m/2$. How might in vivo inhibition plots look at constant velocity (for example $v=V_m/2$) when both I and S can vary and in which S for an enzyme in the middle of a pathway is determined by v?

The equations and graph below show the ratio of S/K_m vs I/K_{ix} for inhibition at constant v, a condition encountered when an enzyme in a metabolic pathway is subject to flux controls imposed by the entire pathway. The x-axis reflects the ratio of inhibitor concentration to its inhibition constant. Likewise, the y-axis reflects the relative amount of substrate compared to its K_m. The graph for *in vivo* competitive inhibition is linear, but it "blows up" for uncompetitive inhibition as shown in Figure 6.4.15.

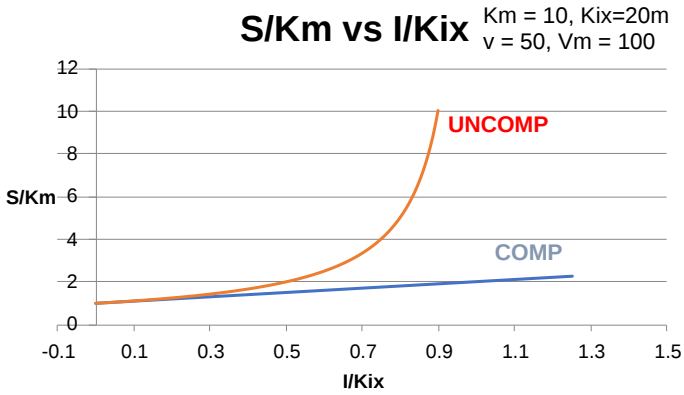


Figure 6.4.15: S/K_m vs $1/K_{ix}$ plots for competitive and uncompetitive inhibition

Here are derivations used to produce the graphs in Figure 6.4.13

? A derivation of the competitive and uncompetitive inhibition in vivo when v = constant

Here it is!

Derivation

Competitive Inhibition at constant velocity v:

Let's start with the equation of competitive inhibition.

$$\mathrm{v} = \frac{\mathrm{V_M S}}{\mathrm{K_M}\left(1 + \frac{\mathrm{I}}{\mathrm{K_{is}}}\right) + \mathrm{S}} \tag{6.4.10}$$

Let

$$\left(1 + \frac{\mathrm{I}}{\mathrm{K_{is}}}\right) = \mathrm{y} \tag{6.4.11}$$

Then,

$$\begin{gathered} v = \frac{V_M S}{K_M y + S} \\ v\left(K_M y + S\right) = V_M S \\ v K_M y + v S = V_M S \\ v K_M y = V_M S - v S = S\left(V_M - v\right) \\ K_M = \frac{S\left(V_M - v\right)}{v y} \end{gathered} \tag{6.4.12}$$

which gives:

$$\begin{gathered} \frac{S}{K_M} = \frac{v y}{\left(V_M - v\right)} = \frac{y}{\frac{V_M}{v} - 1} = \frac{1 + \frac{I}{K_{is}}}{\frac{V_M}{v} - 1} = \frac{1}{\frac{V_M}{v} - 1} + \frac{\frac{I}{K_{is}}}{\frac{V_M}{v} - 1} \\ \frac{S}{K_M} = \frac{1}{\frac{V_M}{v} - 1} + \left(\frac{1}{\frac{V_M}{v} - 1}\right) \frac{I}{K_{is}} \end{gathered} \tag{6.4.13}$$

Note from the last equation that the graph of S/K_M vs I/K_{is} is linear (at a fixed v), as shown in the above figure.

Uncompetitive Inhibition at constant velocity v:

Let's start with the equation of uncompetitive inhibition.

$$\mathrm{v} = \frac{\mathrm{V_M S}}{\mathrm{K_M + S\left(1 + \frac{I}{K_{ii}}\right)}} \tag{6.4.14}$$

Let

$$\left(1 + \frac{1}{\mathrm{K_{ii}}}\right) = \mathrm{y} \tag{6.4.15}$$

then:

$$\begin{aligned} v &= \frac{V_M S}{K_M + Sy} \\ v\left(K_M + Sy\right) &= V_M S \\ vK_M = V_M S - vSy &= S\left(V_M - vy\right) \end{aligned} \tag{6.4.16}$$

which gives:

$$\frac{S}{K_M} = \frac{v}{V_M - vy}\left(\frac{\frac{1}{v}}{\frac{1}{v}}\right) = \frac{1}{\frac{V_M}{v} - y} = \frac{1}{v - 1 - \frac{I}{K_{ii}}} \tag{6.4.17}$$

This graph is not a linear function of I/K_{ii} as it was for in vivo competitive inhibition.

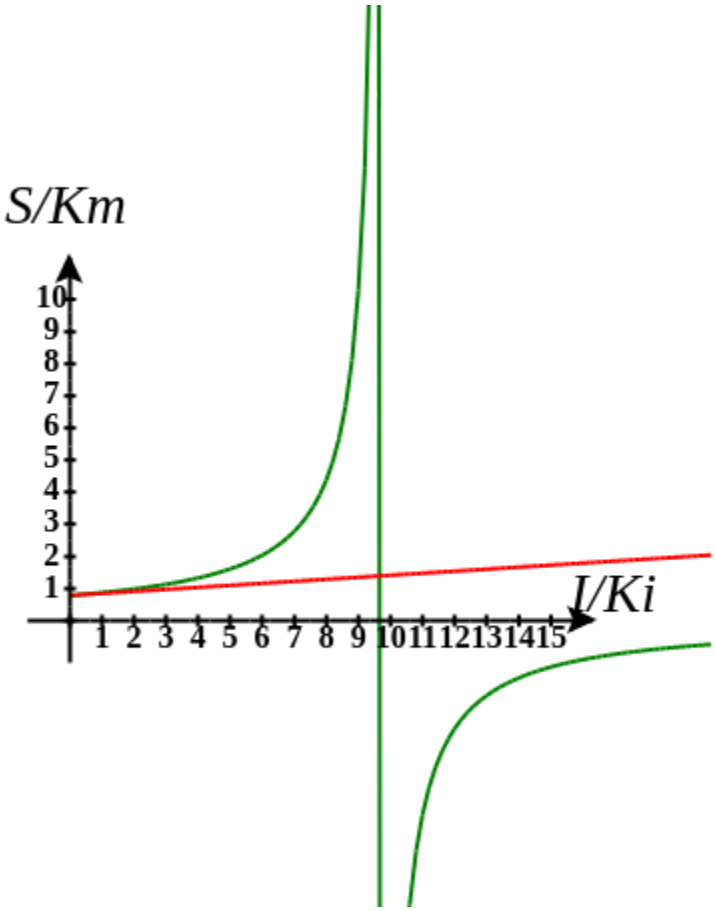
- for competitive inhibition, the graph of S/K_M is a linear function of I/K_{is}
- for uncompetitive inhibition, the graph of S/K_M is **NOT** a linear function of I/K_{ii} but rather "blows up" to infinity.

These graphs and associated equations are **dramatically** different from the very similar forms of inhibition equations and curves for in vitro inhibition at varying S and different fixed values of inhibitor. Consider the uncompetitive graph and equation. In the absence of an inhibitor, if S=Km, then Vm/v = 2 so the calculated value from the equation above of S/Km = 1. The y-intercept of the graph above is 1 for uncompetitive (and competitive) inhibition. If I is allowed to increase to a value of Kii (so I/Kii = 1), again at constant v=Vm/2, then the right-hand side goes to infinity.

In a linked series of reactions, if the middle reaction is inhibited, the substrate for that enzyme builds, whether the inhibition is competitive or uncompetitive. With competitive inhibition, the substrate concentration can be raised to meet the requirements of the enzyme. But as the above figure shows, this can't happen for uncompetitive inhibition since as more substrate accumulates, the reaction reaches a point where the steady state is lost.

Obviously, this limiting case can't be realistically reached but it does suggest that uncompetitive inhibitors would be more effective *in vivo* in controlling a metabolic pathway than competitive inhibitors. Cornish-Bowden argues that purely uncompetitive inhibitors are rare in nature because of the degree of inhibition they can hypothetically produce (1986). Likewise, he suggests that medicinal chemists should synthesize uncompetitive inhibitors if their goal is to maximally inhibit a metabolic pathway under the kind of flux control described above. Although it is more difficult to synthesize a purely uncompetitive inhibitor (as it can't be easily modeled after the structure of a natural ligand that binds to the active site and are competitive inhibitors), he notes that synthesizing mixed (and noncompetitive) inhibitors whose Kii values are of reasonable size compared to their Kis values, would be one approach.

Move the sliders on the interactive graph below to show how the graphs change. Disregard the graph in the right lower quadrant as that location would require negative values of either S or K.



6.4.6: Inhibition by Temperature and pH Changes

From 0 to about 40-50^{o} C, enzyme activity usually increases, as do the rates of most reactions in the absence of catalysts. (Remember the general rule of thumb that reaction velocities double for each increase of 10^{o}C.). At higher temperatures, the activity decreases dramatically as the enzyme denatures. These features are illustrated in Figure 6.4.16.

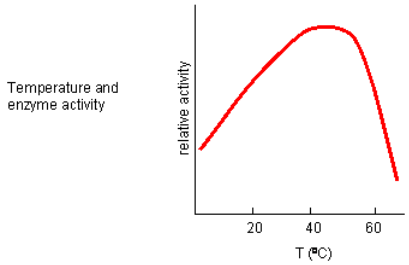


Figure 6.4.16: Temperature effects on enzyme activity

Likewise, pH has a marked effect on the velocity of enzyme-catalyzed reactions as illustrated in Figure 6.4.17.

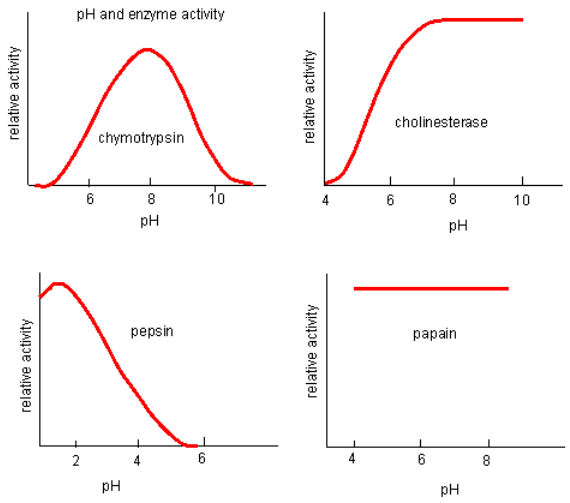


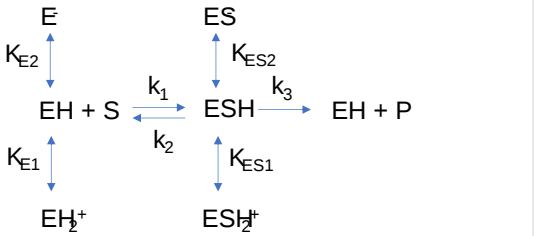
Figure 6.4.17: pH effects on enzyme activity

Think of all the things that pH changes might affect. It might

- affect E in ways to alter the binding of S to E, which would affect Km
- affect E in ways to alter the actual catalysis of bound S, which would affect kcat
- affect E by globally changing the conformation of the protein
- affect S by altering the protonation state of the substrate

The easiest assumption is that key side chains necessary for catalysis must be in the correct protonation state. Thus, a sidechain, with an apparent pKa of around 6, must be deprotonated for optimal activity of trypsin which shows increases in activity with pH centered at pH 6. Which amino acid side chain would be a likely candidate?

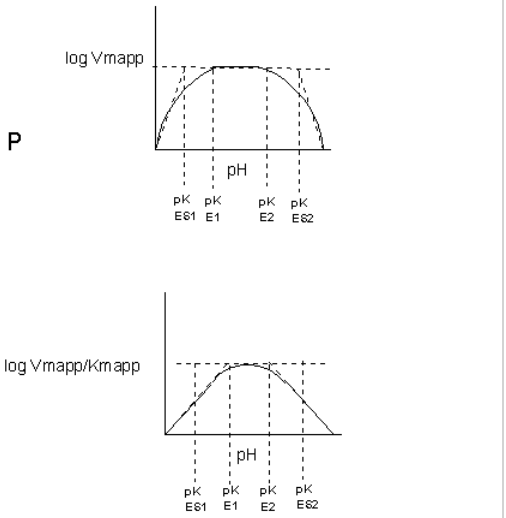
Table 6.4.1 below shows how pH effects on enzyme kinetics can be modeled at the chemical and mathematical level.



$$v_0 = \frac{V_{mapp} S}{K_{m\ app}} + S$$

$$V_{m\ app} = \frac{V_m}{1 + \frac{H^+}{K_{ES1}} + \frac{K_{ES2}}{H^+}} \qquad (6.4.18)$$

$$K_{m\ app} = \frac{K_m \left(1 + \frac{H^+}{K_{ES1}} + \frac{K_{ES2}}{H^+}\right)}{1 + \frac{H^+}{K_{ES1}} + \frac{K_{ES2}}{H^+}}$$



6.5: Enzymatic Reaction Mechanisms

This chapter section has been written by Kristen Procko and Henry Jakubowski.

We can apply what we learned about catalysis by small molecules (e.g., acids and bases) to enzyme-catalyzed reactions. To understand the mechanism of an enzyme-catalyzed reaction, we try to alter as many variables, one at a time, and ascertain the effects of the changes on the activity of the enzyme. Kinetic methods can be used to obtain data, from which inferences about the mechanism can be made. Crystal structures of the enzyme in the presence and absence of a competitive inhibitor give abundant information about possible mechanisms. It is amazing, however, how much information about enzyme mechanism can be gained even if all you have is a blender, a stopwatch, an impure enzyme, and a few substrates and inhibiting reagents.

6.5.1: Introduction to Enzymes Mechanisms

Almost every chemical reaction in the biological world is catalyzed by protein enzymes. The human genome encodes for over 20,000 different proteins, thousands of which are enzymes. The total number of different enzymes in the biosphere must be staggering. Yet at the same time, all of these enzymes catalyze different sets of similar reactions. To bring order to the world of enzyme catalysis, the IUBMB has classified enzymes based on the type of chemical reactions they catalyze. There are 7 main categories as shown in the expandable Table 6.5.1 below. Each reaction type is given a four digit **Enzyme Commission** number. For example, alcohol dehydrogenase, the enzyme that catalyzes the oxidation of ethanol, a primary alcohol, to acetaldehyde using an oxidizing agent called NAD^+, is given the enzyme commission number EC 1.1.1.1. Other enzymes that oxidize primary alcohols to aldehydes or secondary alcohols to ketones are also give the same EC number.

Subclass	Description
Oxidoreductases	Redox reactions
Transferases	Transfer/exchange of group from one molecule to another

Hydrolases	**EC 3**	hydrolysis reactions
Lyases	**EC 4**	elimination forming double bond
Isomerases	**EC 5**	conversions of geometric, stereo- or constitutive isomers
Ligases	**EC 6**	condensation of two molecules into one
Translocases	**EC 7**	movement of species across a semipermeable membrane

Table 6.5.1: ExplorEnz database that for the curation and dissemination of the International Union of Biochemistry and Molecular Biology (IUBMB) Enzyme Nomenclature. (Source: https://academic.oup.com/nar/article...1/D593/1000297)

Enzymes with the same or similar EC numbers probably have similar reaction mechanisms. Throughout this book, we will explore the reaction mechanisms of many enzymes, but we can't and shouldn't explore all of them. You can take your acquired understanding of the reaction mechanism for key representative enzymes and apply them to others. Of course, experimental evidence is needed to validate a given mechanism. In this chapter section, we will focus on the mechanisms of a few transferases (EC2) and hydrolases (EC3) as prototypical examples.

6.5.2: Arrow Pushing Conventions in Biochemical Mechanisms

The rules for electron pushing in biochemical mechanisms mirror those from organic chemistry. However, because of the length of some biochemical mechanisms, abbreviated mechanisms are often accepted and will be presented in the literature and textbooks. The aim of this section is to review arrow pushing by presenting some simple biochemical mechanisms, and to familiarize you with acceptable alternative ways to show arrow pushing.

6.5.2.1: Class I Methyltransferases

Coenzymes are organic molecules that participate in some enzyme-catalyzed reactions (see Section 6.8 for a detailed discussion). Often, these "enzyme helpers" impart reactivity that an enzyme would not have on its own. We'll begin our investigation with a mechanism catalyzed by class I methyltransferases, which bind the coenzyme S-adenosylmethionine (SAM). SAM itself is formed from the reaction of methionine with ATP, resulting in the positively charged sulfur shown in Figure 6.5.1. The blue methyl group attached to the sulfur is very electrophilic due to the sulfur cation, and is transferred to a nucleophilic substrate.

The reaction in Figure 6.5.1 shows the SAM-promoted conversion of norepinephrine to epinephrine. It is catalyzed by the enzyme phenylethanolamine N-methyltransferase (EC: 2.1.1.28). The nucleophilic nitrogen atom in norepinephrine is brought into proximity of the electrophilic methyl group when it binds to the enzyme. In an S_N2 reaction, nitrogen attacks, and we show a second arrow to keep track of the electrons from the carbon-sulfur bond, becoming a lone pair on the sulfur atom.

An amine has a pK_a close to 30, but a protonated amine has a pK_a around 8–10. Specifically, the conjugate acid of epinephrine has a pK_a near 8. Therefore, following the attack step, a protonated amine is a reasonable product to show at physiological pH (near 7.4). However, biochemical products are often shown uncharged when depicting an overall reaction. Therefore, a general base can be shown deprotonating the ammonium ion. The general base could be an amino acid side chain, or a general base within the buffer components contained in a cell; we will represent the general base as B: here.

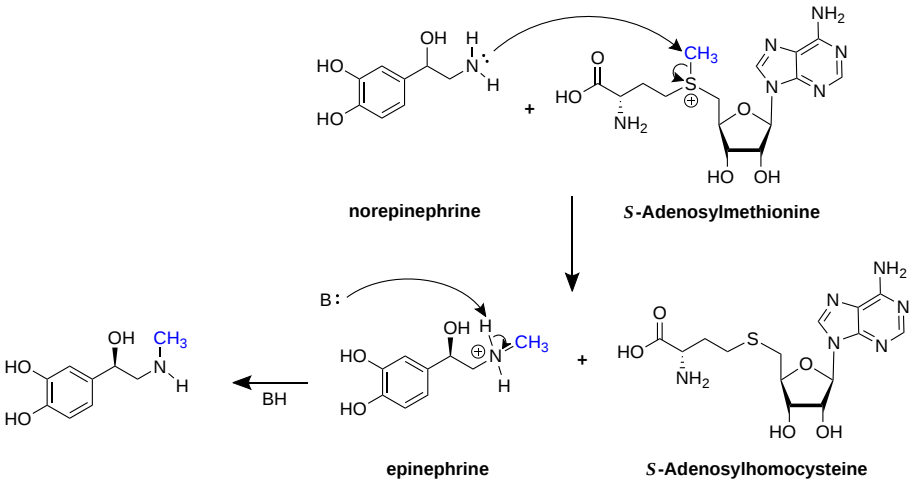


Figure 6.5.1: Class I methyltransferase mechanism with substrate norepinephrine

To simplify the mechanism in Figure 6.5.1, biochemists may abbreviate the arrow pushing steps. As shown by the one-step deprotonation in Figure 6.5.2, the general base deprotates the amine. The electrons from that bond are used to attack, and deprotonated epinephrine is produced in a single step with that arrow pushing. It is important to keep in mind that an anionic nitrogen is not produced in the reaction mechanism. With pK_a values near 30, most amines are only deprotonated in practice using very strong organic bases, such as butyllithiums.

Because of this, the authors prefer the alternative arrow pushing mechanism shown at the bottom of Figure 6.5.2. Partial bond formation between the nitrogen and the methyl group must occur before a biological general base is able to deprotonate norepinephrine, so showing the arrow coming from the lone pair on nitrogen is perhaps a better depiction of the biological reality. The arrow pushing shown in each of the methyltransferase mechanisms is acceptable, and you may see each of these in different contexts.

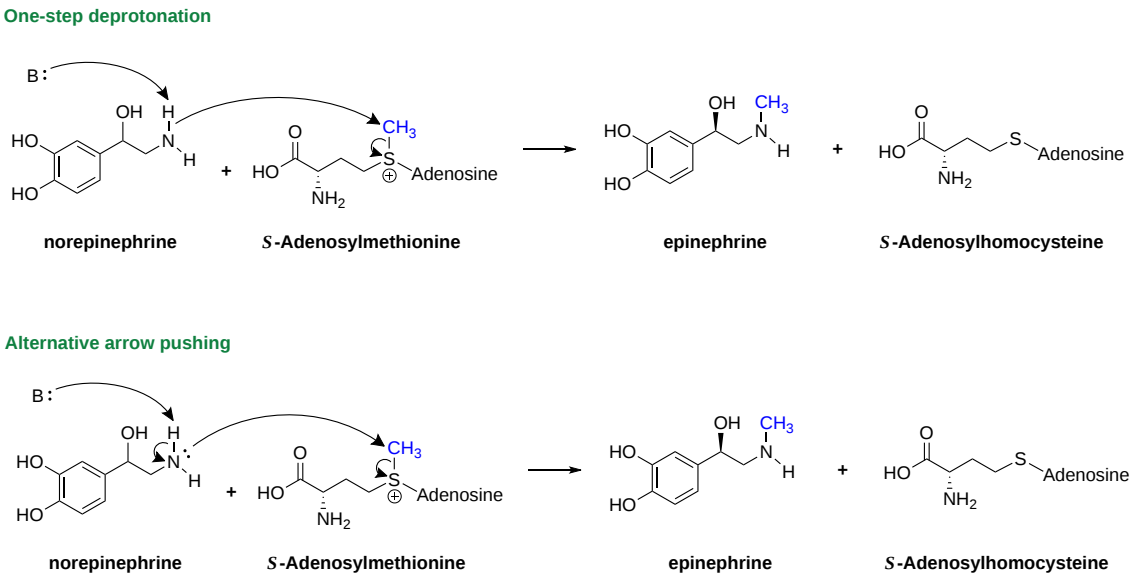


Figure 6.5.2: Class I methyltransferase mechanism alternative arrow pushing

6.5.2.2: Kinases

Transfer of a phosphate group is common in biochemistry. The phosphate group is often transferred from adenosine triphosphate (ATP) via kinase enzymes. The following excerpt from Chemistry LibreTexts describes the phosphate transfer mechanism:

In a phosphate transfer reaction, a phosphate group is transferred from a **phosphate group** donor molecule to a phosphate group acceptor molecule as shown in Figure 6.5.3.

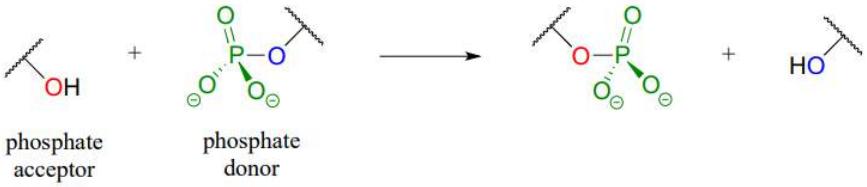


Figure 6.5.3: Phosphate transfer to an acceptor

A very important aspect of biological phosphate transfer reactions is that the electrophilicity of the phosphorus atom is usually enhanced by the Lewis acid (electron-accepting) effect of one or more magnesium ions. Phosphate transfer enzymes generally contain a Mg^{2+} ion bound in the active site in a position where it can interact with non-bridging phosphate oxygens on the substrate (Figure 6.5.4). The magnesium ion pulls electron density away from the phosphorus atom, making it more electrophilic.

Figure 6.5.4: Phosphate interaction with magnesium ion

Without this metal ion interaction, a phosphate is actually a poor electrophile, as the negatively-charged oxygens shield the phosphorus center from attack by a nucleophile.

Note: For the sake of simplicity and clarity, we may omit the magnesium ion or other active site groups interacting with phosphate oxygens in some of the figures that follow—but it is important to keep in mind that these interactions play an integral role in phosphate transfer reactions.

Mechanistically speaking, a phosphate transfer reaction at a phosphorus center can be thought of as much like a S_N2 reaction at a carbon center. Just like in an S_N2 reaction, the nucleophile in a phosphoryl transfer approaches the electrophilic center opposite the leaving group, from the backside as shown in Figure 6.5.5.

Concerted model:

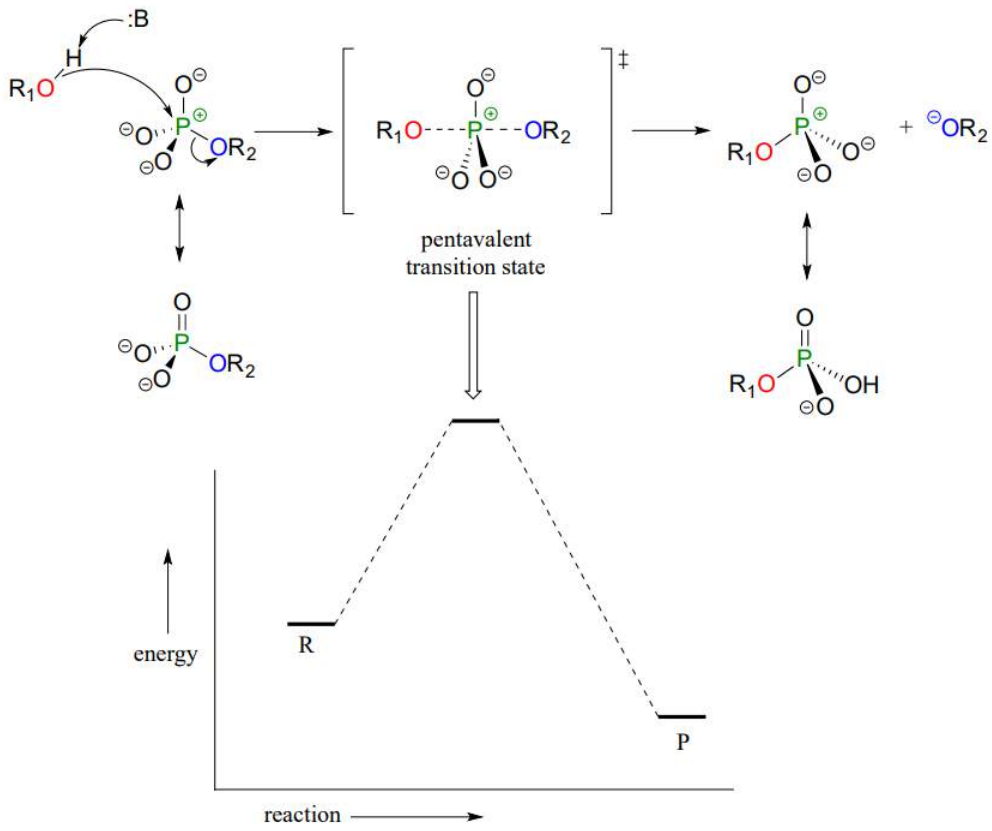


Figure 6.5.5: Phosphate transfer mechanism and energy diagram

As the nucleophile gets closer and the leaving group begins its departure, the bonding geometry at the phosphorus atom changes from tetrahedral to trigonal bipyramidal at the pentavalent (5-bond) transition state (Figure 6.5.6). As the phosphorus-nucleophile bond gets shorter and the phosphorus-leaving group bond grows longer, the bonding picture around the phosphorus atom returns to its original tetrahedral state, but the stereochemical configuration has been 'flipped', or inverted.

In the trigonal bipyramidal transition state, the five substituents are not equivalent: the three non-bridging oxygens are said to be equatorial (forming the base of a trigonal bipyramid), while the nucleophile and the leaving group are said to be apical (occupying the tips of the two pyramids).

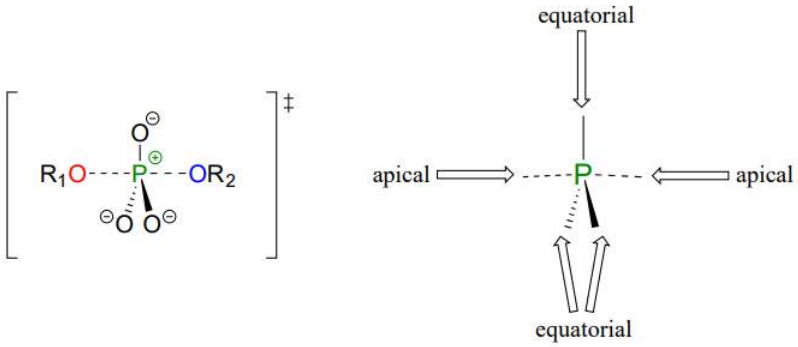


Figure 6.5.6: Transition state of the phosphate transfer reaction

Although stereochemical inversion in phosphoryl transfer reactions is predicted by theory, the fact that phosphoryl groups are achiral made it impossible to observe the phenomenon directly until 1978, when a group of researchers was able to synthesize

organic phosphate esters in which stable oxygen isotopes ^{17}O and ^{18}O were specifically incorporated (Figure 6.5.7). This created a chiral phosphate center.

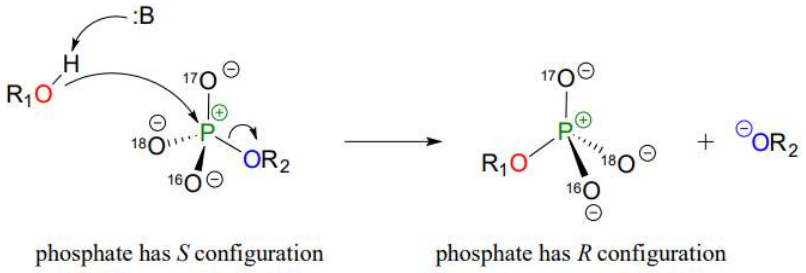


Figure 6.5.7: Inversion of stereochemistry in the phosphate transfer reaction

Subsequent experiments with phosphoryl transfer-catalyzing enzymes confirmed that these reactions proceed with stereochemical inversion. (Nature 1978 275, 564; Ann Rev Biochem 1980 49, 877). (The previous excerpt has been adapted from Chemistry LibreTexts.)

We should note that, although the charge-separated resonance form shown above contributes to the structure and therefore contributes to the resonance hybrid, the phosphate group is almost always shown with a double bonded oxygen atom to the phosphorous. Note that we have also depicted the mechanism below with our preferred biochemical arrow pushing, where the attack of oxygen at phosphorous facilitates the deprotonation of the alcohol (R_1OH) by a general base.

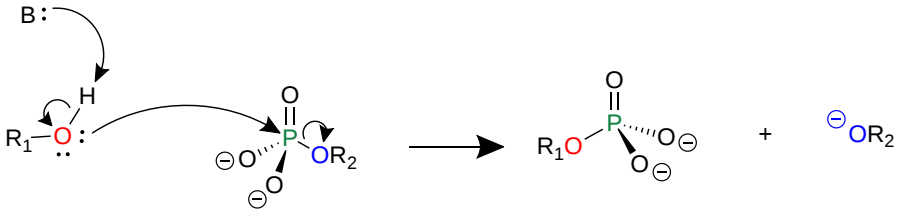


Figure 6.5.8: Common arrow pushing for phosphate transfer

With a formal charge of negative four, ATP would be an extremely poor electrophile. Therefore, the phosphoryl transfer from ATP, shown for hexokinase (EC 2.7.1.1) in Figure 6.5.9, requires binding of a magnesium ion with ATP. The magnesium ion partially neutralizes the negative charge, allowing for nucleophilic attack by the oxygen atom of glucose.

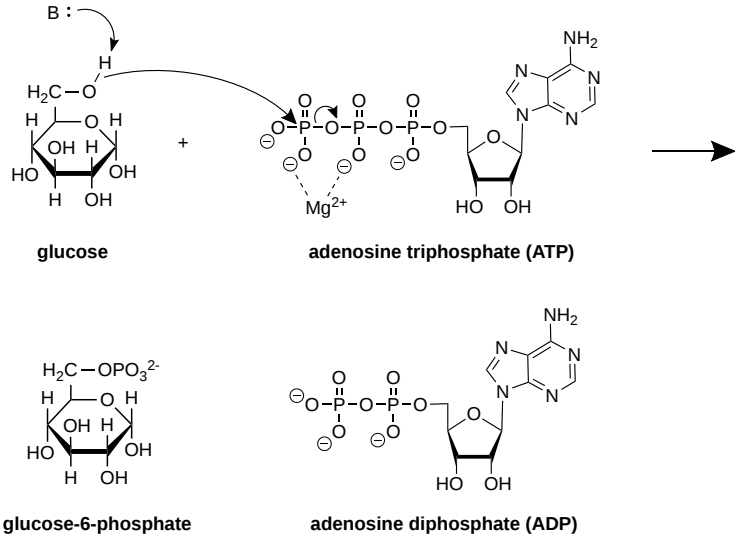


Figure 6.5.9: Hexokinase mechanism with phosphate transfer from ATP

A crystal structure of the enzyme was solved with the substrate, glucose-6-phosphate, bound. The magnesium cofactor also occupies the active site, along with a non-hydrolyzable analog of ATP. In the substrate analog, called ANP, the oxygen of the

terminal phosphoanhydride bond of ATP is replaced with a nitrogen atom, and allows us to view probable interactions that occur between the enzyme of the catalytic complex. In the iCn3D image shown in Figure 6.5.10, ANP and glucose are bound in the active site. The magnesium ion is shown in green.

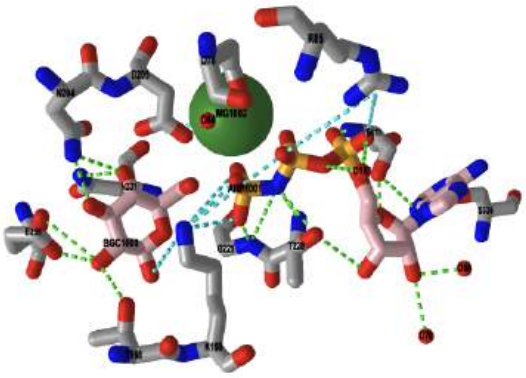


Figure 6.5.10: Interactive iCn3D image of the catalytic complex of human glucokinase (a hexokinase isoform, 3FGU). (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?EajeSCpF8GM16HYq5

6.5.3: Elucidation of Reaction Mechanisms Using Kinetic Data

To to this point, we've presented mechanisms with the support of PDB structures alone. However, much was known about enzyme mechanisms prior to ready access to crystal structures in the Protein Data Bank. Systematically, the kineticists, medicinal chemists and molecular biologists (i.e., a well trained chemist) can change:

A. the substrate - for example, changing the leaving group or substituents of a hydrolyzable substrate;
B. the pH or ionic strength - which can give data about general acids/bases in the active site;
C. the enzyme - by chemical modification of specific amino acids, or through site-specific mutagenesis;
D. the solvent - an odd idea on the surface but it leads to new insights into enzyme catalysis.

For the following enzymes, we will concentrate on reaction mechanisms based on a mix of structural data, alongside kinetic data to hypothesize a reaction mechanism consistent with the findings. Even with lots of data, there are often different proposed mechanisms for a given reaction. Kinetic data is vital as it can help to determine:

- the order of binding/dissociation of substrates and products;
- the rate constants for individual steps;
- and clues to the nature of catalytic groups found in the enzyme.

6.5.3.1: Chymotrypsin and Other Endoproteases

Chymotrypsin (EC 3.4.21.1), an endoprotease, cleaves an internal peptide bond after aromatic side chains by hydrolysis. It also cleaves small ester and amide substrates after aromatic residues. As an example, in Figure 6.5.11, cleavage occurs on the C-terminal side of the tyrosine residue, giving two peptide fragments.

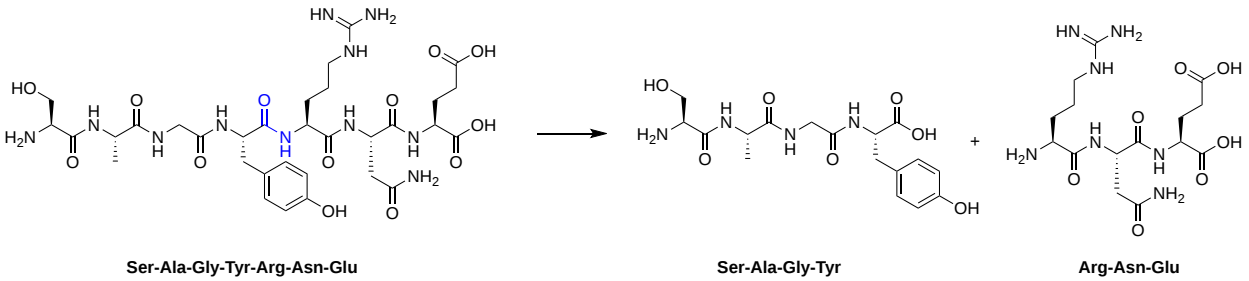


Figure 6.5.11: Chymotrypsin cleavage of an example peptide substrate

Chymotrypsin has a similar mechanism to a multitude of other proteases that used the same catalytic triad, Ser 195, Asp 102 and His 95, so we'll study it significant detail. In determining the mechanism of an enzyme, you have to change an experimental

variable and see how catalytic activity changes. What can be changed? Turns out everything including the solvent! Let's explore these changes and how they affect chymotrypsin activity.

A. **Changing the substrate** (for example changing the leaving group or acyl substituents of a hydrolyzable substrate):

In the lab, it's easier to study the enzyme using small substrate mimics of a protein than to use a full protein substrate. The mimics include both esters and amides. Data from the cleavage of small amide and ester substrates shown in Figure 6.5.12 suggest that a covalent intermediate is formed during chymotrypsin catalyzed cleavage.

Figure 6.5.12: Small amide and ester substrates of chymotrypsin

Table 6.5.2 below shows kinetic data for the cleavage of these substrates.

Chymotrypsin substrate cleavage, 25 °C, pH 7.9			
kinetic constants	Acetyl-Tyr-Gly-amide	Acetyl-Tyr-O Ethylester	Ester/Amide
k_{cat} (s^{-1})	0.50	193	390
K_m (M)	0.023	0.0007	0.03
k_{cat}/K_m ($M^{-1}s^{-1}$)	22	280,000	12,700

Kinetic constants for chymotrypsin cleavage of N-acetyl-L-Trp Derivatives - N-acetyl-L-Trp-X		
X	k_{cat} (s^{-1})	K_m x 103 (M)
$-OCH_2CH_3$	27	0.097
$-OCH_3$	28	0.095
-p-nitrophenol	31	0.002
$-NH_2$	0.026	7.3

Table 6.5.2: Cleavage of peptides and ethylester substrate analogs by chymotrypsin

Here's how these data can be interpreted.

1. The k_{cat} and k_{cat}/K_m are larger and the K_m smaller for ester substrates compared to amide substrates, suggesting that amides are more difficult to hydrolyze (Table 2 above). This is expected given the poorer leaving group of the amide.
2. The k_{cat} for the hydrolysis of ester substrates doesn't depend on the nature of the leaving group (i.e., whether it is a poorer leaving group such as methoxy or a better leaving group such as p-nitrophenolate) suggesting that this step is not the rate limiting step for ester cleavage. Without the enzyme, p-nitrophenyl esters are cleaved much more rapidly than methyl esters. Therefore deacylation must be rate limiting. But deacylation of what? If water was the nucleophile, release of the leaving group would result in both products, the free carboxyl group and the amine being formed simultaneously. Since they are not released simultaneously, this suggests an acyl-enzyme covalent intermediate.

When the acyl end of the ester substrate is changed, without changing the leaving group (a p-nitrophenyl group), a covalent intermediate can be trapped. Specifically, the deacylation of a trimethyacetyl group is much slower than an acetyl group. It is so slow that a ^{14}C-labeled trimethylacetyl-labeled chymotrypsin intermediate can be isolated after incubation of chymotrypsin with ^{14}C-labeled p-nitrophenyltrimethylacetate using gel filtration chromatography.

We have seen a kinetic mechanism previously consistent with these ideas before. The data suggest a mechanism based on the chemical equations shown in Figure 6.5.13:

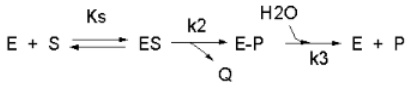


Figure 6.5.13: Chemical equations for chymotrypsin hydrolysis of a substrate involved a covalent intermediate with ping-pong kinetics.

In this reaction, a substrate S might interact with E to form a complex, which then is cleaved to products P and Q. Q is released from the enzyme, but P might stay covalently attached, until it is expelled. This conforms exactly to the mechanism described above. For chymotrypsin-catalyzed cleavage, the step characterized by $\mathbf{k}_2$ is the acylation step. The step characterized by $\mathbf{k}_3$ is the deacylation step in which water attacks the acyl enzyme to release product P (free phosphate in Lab 5). The mathematical equation for this reaction is shown below (without derivation)

$$v_0 = \frac{\left(\frac{k_2 k_3}{k_2+k_3}\right) E_0\, S}{K_S\left(\frac{k_3}{k_2+k_3}\right)+S} \tag{6.5.1}$$

For hydrolysis of ester substrates, which have better leaving groups compared to amides, deacylation is rate limiting, ($k_3 << k_2$). For amide hydrolysis, as mentioned above, acylation can be rate-limiting ($k_2 << k_3$). From this, equation 6.5.1 can be simplified as shown in Table 6.5.3 below for ester and amide hydrolysis.

Ester hydrolysis (deacylation rate limiting, $k_3 << k_2$)		Amide hydrolysis (deacylation rate limiting, $k_2 << k_3$)	
$v_0 = \frac{k_3 E_0\, S}{K_S\left(\frac{k_3}{k_2}\right)+S}$	(6.5.2)	$v_0 = \frac{k_2 E_0\, S}{K_S + S}$	(6.5.3)
$V_M = k_3 E_0$	(6.5.4)	$V_M = k_2 E_0$	(6.5.5)
$K_M = K_S\left(\frac{k_3}{k_2}\right)$	(6.5.6)	$K_M = K_S$	(6.5.7)

Table 6.5.3: Simplification of equation 6.5.1

Just as we saw before for the rapid equilibrium assumption (when ES falls apart to E + S more quickly than it goes to product, Chapter 6.3), K_M = Ks in the case of amide hydrolysis.

B. **Changing the pH or ionic strength** - which can give data about general acids/bases in the active site:

- a graph of k_{cat} as a function of pH indicates that a group of pK_a of approximately 6 must be deprotonated to express activity (i.e., $V_{max}/2$ is at about pH 6). This suggests that an active site histidine is necessary, which, if it must be deprotonated to express activity, must be acting as a general base.
- a graph of k_{cat}/K_m shows a bell-shaped curve indicating the necessity of a deprotonated side chain with a pK_a of about 6 (i.e., the same His above) and a group which must be protonated with a pK_a of about 10. This turns out to be an N terminal Ile in chymotrypsin, which must be protonated to form a stabilizing salt bridge in the protein. *Note*: This N-terminal Ile is actually at the 16 position in the inactive precursor of chymotrypsin (called chymotrypsinogen); upon activation of chymotrypsinogen loses the first 15 amino acids by selective proteolysis.

(*Note*: The PKAD is a database of experimentally measured pK_a values of ionizable groups in proteins. It is searchable by the PDB ID.)

C. **Changing the enzyme** - by chemical modification of specific amino acids, or through site-specific mutagenesis:

Here are some specific examples.

1. Modification of chymotrypsin (and many other proteases) with diisopropylphosphofluoridate (DIPF) modifies only one (Ser 195) of many serines in the protein, suggesting that it is hypernucleophilic and probably the amino acid that attacks the carbonyl C in the substrate, forming the acyl-intermediate. This reaction is illustrated in Figure 6.5.13. The figure also shows analagous molecules used in common insecticides, which act through a similar mechanism.

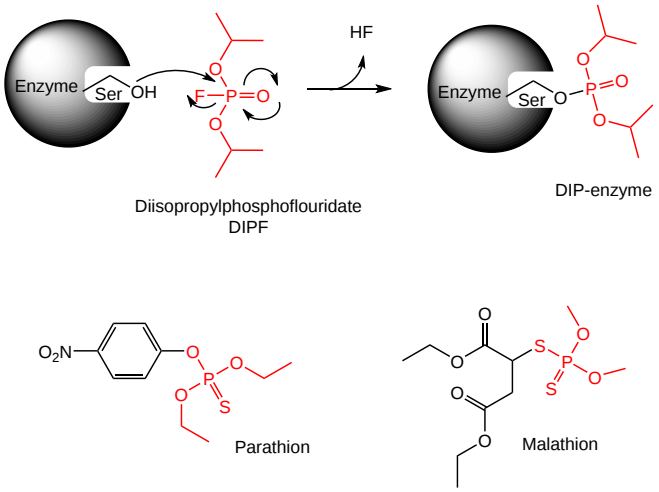


Figure 6.5.13: Mechanism of inhibition of chymotrypin by covalent modification by diisopropylphosphflouridate.

2. Modification of the enzyme with tos-L-Phe-chloromethyl ketone inactives the enzyme with a 1:1 stoichiometry which results in a modified His, as shown in Figure 6.5.14

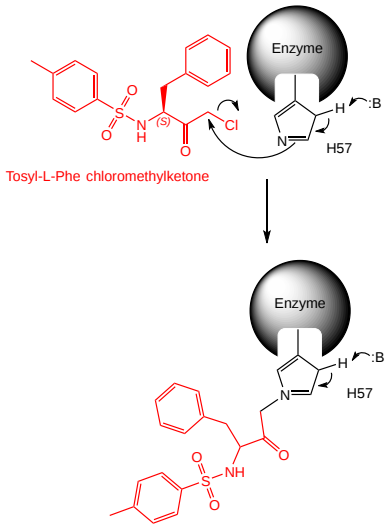


Figure 6.5.14: Reaction of chymotrypsin and other serine protease with chloromethy ketones.

3. Comparison of the primary sequence of many proteases show that three residues are invariant: a Ser, a His, and an Asp residue.
4. Site-specific mutagenesis show that if Ser 195 is changed to Ala 195, the enzymatic activity is almost reduced to background. The strongly suggests that Ser 195 is an active site nucleophile.

D. **Changing the solvent**. Yes indeed you can take chymotrypsin and show that it is active in anhydrous organic solvents. Surely this is impossible you say! It is true and we will explore it at the end of chapter since its challenging enough to understand chymotrypsin activity in aqueous solution. No new chemistry is needed, just a change in what your mind can conceptualize.

6.5.3.1.1: The Chymotrypsin Arrow-Pushing Mechanism

The chymotrypsin mechanism will be presented to explore the different types of acceptable arrow pushing one can show for this nucleophilic acyl substitution mechanism. In the mechanism, only the peptide bond will be shown. It is only necessary to focus on this small portion of the molecule shown in Figure 6.5.15 to show the arrow pushing.

Figure 6.5.15: Abbreviated chymotrypsin peptide cleavage reaction

The active site of chymotrypsin contains a catalytic triad, three amino acids working together to carry out the reaction that cleaves the peptide bond. The amino acids involved are the aspartate, histidine, and serine residues mentioned earlier (Figure 6.5.16).

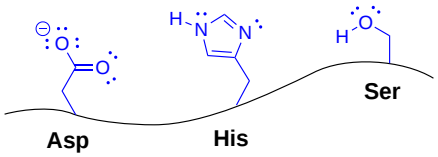


Figure 6.5.16: Serine protease catalytic triad

The deprotonated aspartate side chain acts to increase histidine's basicity, allowing it to accept a proton from serine in the catalytic mechanism. In some mechanisms, Asp is shown accepting a proton from histidine; however, simplified arrow pushing can be shown without Asp acting as a proton acceptor, and that is how the mechanism will be represented here.

In the first stage of the mechanism, histidine deprotonates serine, which acts as a nucleophile and attacks the partially electropositive carbon atom of the carbonyl functional group (Figure 6.5.17). In the simplest form of arrow pushing that can be shown, histidine deprotonates the nucleophilic using the lone pair on nitrogen, and the electrons from the hydrogen-oxygen bond are shown attacking. The carbonyl double bond breaks, shifting the electrons onto oxygen.

This forms a tetrahedral sp^3 hybridized carbon atom from the sp^2 hybridized carbonyl group, and is therefore called a tetrahedral intermediate.

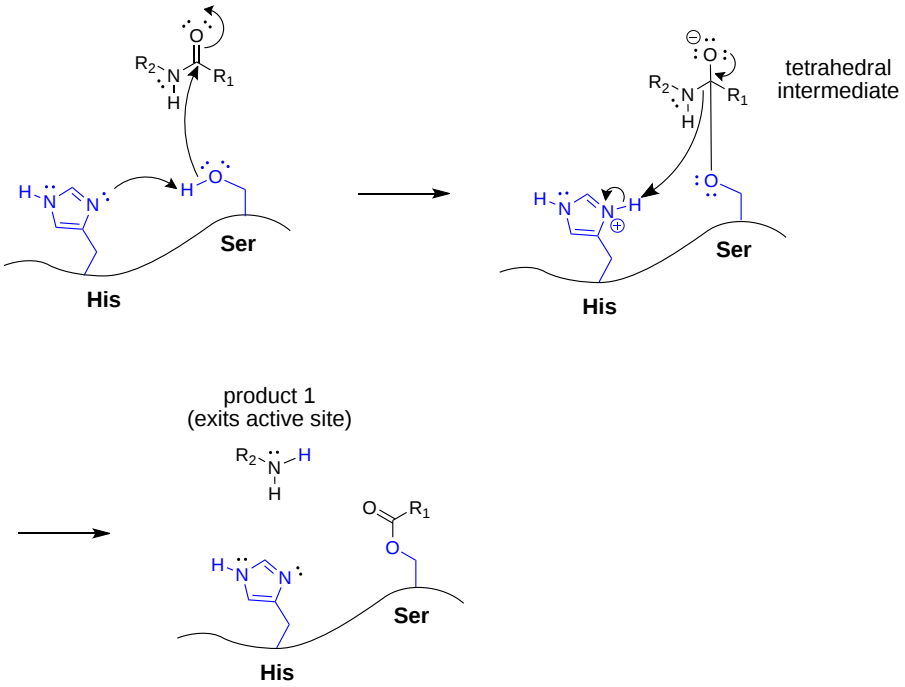


Figure 6.5.17: First stage of the chymotrypsin mechanism

Using the simple arrow pushing model again, the electrons from the negatively-charged tetrahedral intermediate reform a double bond, kicking out the amine leaving group, which accepts an H^+ from protonated histidine, neutralizing the charge on histidine. The arrow from the N-H bond neutralizes the positive charge on histidine; please note that this arrow is important to show. In all deprotonations, it is a convention to show electrons from a breaking bond becoming a lone pair on the atom receiving them.

Another acceptable form of arrow pushing for this stage of the reaction requires more arrows, but better depicts how the enzyme is interacting with the substrate in the active site (Figure 6.5.18. Because an active site often uses entropy reduction, bringing substrates close together in a reactive orientation, the lone pairs on heteroatoms are already interacting favorably to form the new bond. Additionally, in the first mechanism, it almost appears that a serine alkoxide attacks the carbon of the peptide bond, and that R-NH^- (a poor leaving group) departs the molecule, picking up a proton after the bond breaks.

Therefore, in this second mechanism option, these subtleties are considered. In the first step, the proton on serine is deprotonated by histidine, and those electrons are pushed toward the serine oxygen atom. The serine oxygen is positioned close to the carbonyl carbon of the amide bond, and an arrow originating from the serine lone pair depicts the attack. This type of arrow pushing implies the attack and deprotonation steps are happening in concert.

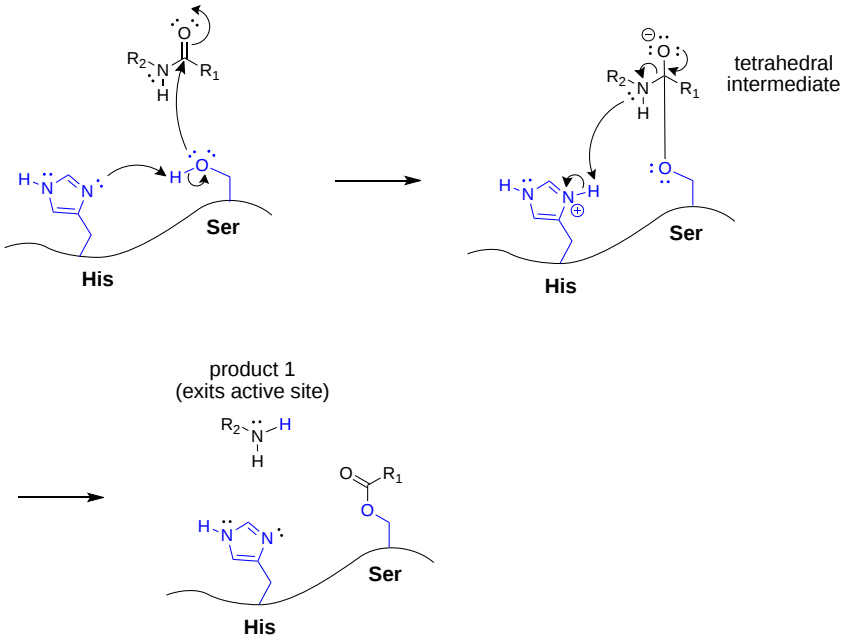


Figure 6.5.18: Alternate arrow pushing for the first stage of the chymotrypsin mechanism

In the second step, the electrons that push down from the negatively charged oxygen atom break the N-H bond, becoming a lone pair on oxygen. Simultaneously, the lone pair already present on nitrogen is shown deprotonating histidine.

It is important to note that, because biochemical mechanisms are often lengthy, they may be shown with the tetrahedral intermediate omitted. This arrow pushing for the first step, which may be shown using either convention described above, shows the serine oxygen attacking and the amine leaving group departure in one step, as shown in Figure 6.5.19 below. Note that this abbreviated style of arrow pushing, which does not show the tetrahedral intermediate, often uses generic acids and bases, so it does not keep track of protonation states, or account for the amino acid residues performing protonation and deprotonation steps (which is quite important in the chymotrypsin mechanism).

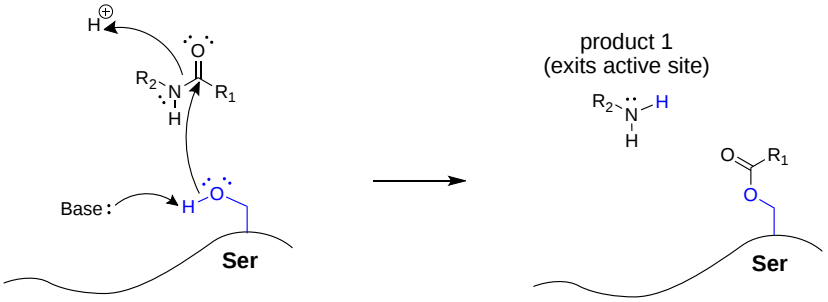


Figure 6.5.19: Abbreviated arrow pushing for the first stage of the chymotrypsin mechanism

The covalent intermediate must be released from the enzyme in order for chymotrypsin to catalyze another reaction. This second nucleophilic acyl substitution also proceeds through a tetrahedral intermediate; this time, water is the nucleophile, as shown in Figure 6.5.20.

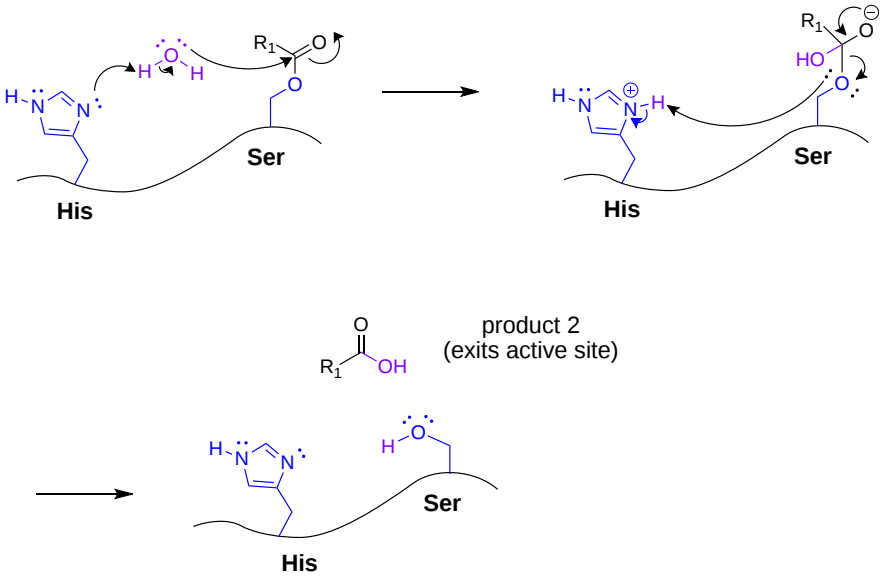


Figure 6.5.20: Second stage of the chymotrypsin mechanism

Now, that we've seen the steps in detail, let's put all this together to show the full mechanism for serine protease cleavage of protein, shown in Figure 6.5.21.

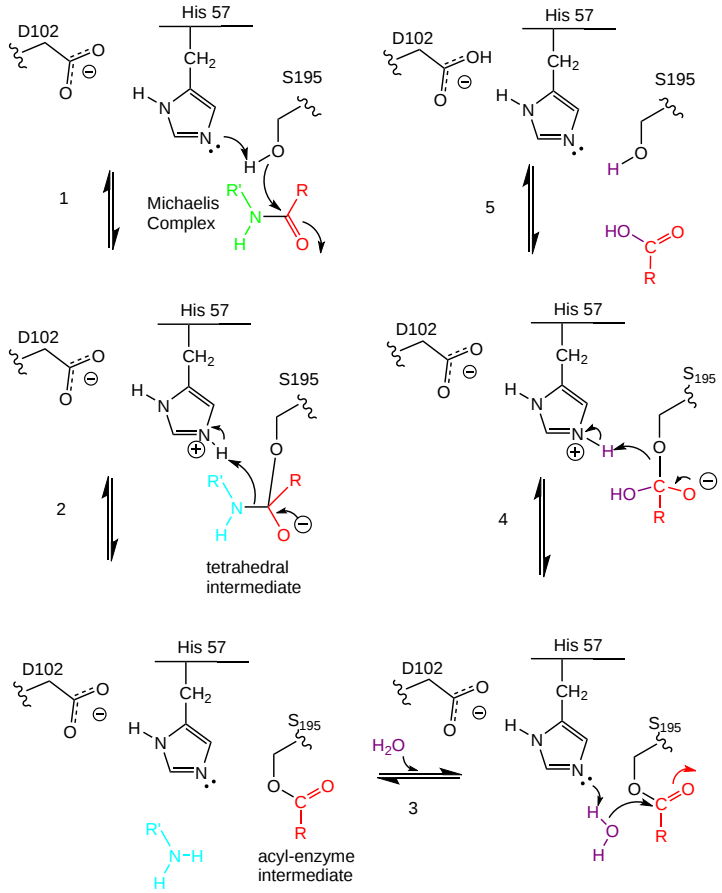


Figure 6.5.21: General mechanism of peptide bond cleavage by chymotrypsin and other serine proteases

Here a summary of what the Figure 6.5.21 mechanism shows:

- The deprotonated His 57 acts as a general base to abstract a proton from Ser 195, enhancing its nucleophilicity as it attacks the electrophilic C of the amide or ester link, creating the oxyanion tetrahedral intermediate. Asp 102 acts electrostatically to

stabilize the positive charge on the His.

- The oxyanion collapses back to form a double bond between the O and the original carbonyl C, with the amine product as the leaving group. The protonated His 57 acts as a general acid donating a proton to the amine leaving group, regenerating the unprotonated His 57.
- The mechanism repeats itself, only now with water as the nucleophile, which attacks the acyl-enzyme intermediate, to form the tetrahedral intermediate.
- The intermediate collapses again, releasing the E-SerO- as the leaving group which gets reprotonated by His 57, regenerating both His 57 and Ser 195 in the normal protonation state. The enzyme is now ready for another catalytic round of activity.
- The mechanism for the first nucleophilic attack (by Ser) is the same as for the second (by water). The reverse mechanism of condensation of two peptide would be the reverse of the above mechanism, and is an example of the **principle of microscopic reversibility.**

In short, many of the catalytic mechanisms we encountered previously are at play in chymotrypsin catalysis. These include nucleophilic catalysis (with the Ser 195 forming a covalent intermediate with the substrates), general acid/base catalysis with His 57, and loosely, electrostatic catalysis with Asp 102 stabilizing not the transition state or intermediate, but the protonated form of His 57. An important point to note is that His, as a general acid and base catalyst, not only stabilizes developing charges in the transition state, but also provides a path for proton transfer, without which, reactions would have difficulty in proceeding.

One final mechanism is at work. The enzyme does indeed bind the transition state more tightly than the substrate. Crystal structures with poor "pseudo"-substrates that get trapped as partial tetrahedrally-distorted substrates of the enzyme and with inhibitors show that the oxyanion intermediate, and hence presumably the TS, can form H-bonds with the amide H (from the main chain) of Gly 193 and Ser 195. These **cannot be made** to the trigonal, sp^2 hybridized substrate. In the enzyme alone, the hole into which the oxyanion intermediate and TS would be placed is not occupied. This oxyanion hole is occupied in the tetrahedral intermediate.

A crystal structure of a relative of chymotrypsin, **trypsin**, which cleaves after positively charged lysine and arginine side chains, has been determined with a bound transition state analog inhibitor. The transition state inhibitor is t-butoxy-Ala-Val-boro-Lys methyl ester shown in Figure 6.5.22

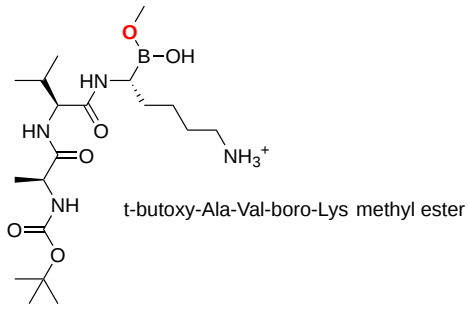


Figure 6.5.22: The transition state inhibitor of trypsin, t-butoxy-Ala-Val-boro-Lys methyl ester

Recall from introductory chemistry that neutral boron compounds like BH_3 and BF_3 are trigonal planar (sp^2) and electron deficient. Although the boron is not charged, it has a significant partial positive charge (δ^+) so it is electrophilic. The nucleophilic oxygen of Ser 195 can then attack the boron to form a tetrahedral intermediate. This intermediate is not an oxyanion, but one of the attached oxygens with a δ^- charge occupies the oxyanion hole.

Figure 6.5.23 show the active site group in trypsin interacting with part of the transition state analog (1BTZ). The serine 195 side chain O is covalently attached to the boron, so the boron is now tetrahedral (sp^3).

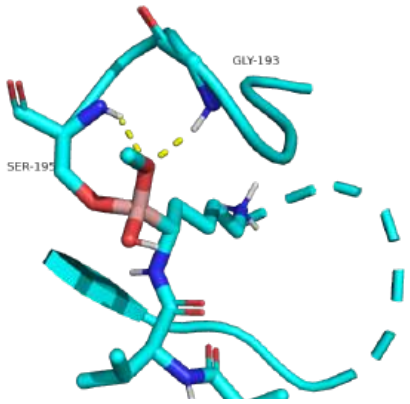
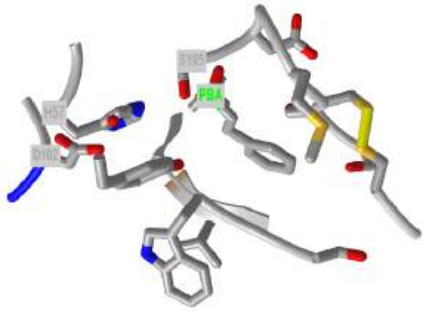


Figure 6.5.23: Stabilization of a transition state analog by the oxyanion hole

The yellow dotted lines show hydrogen bonding between the **backbone** amide hydrogens of Ser 195 and Gly 193 with the methoxy oxygen of the now tetrahedral borate transition state analog inhibitor. The boron atom is the yellow/orange sp^3 atom connected to 3 oxygen (red) atoms and one carbon (cyan) atom. Normally, the oxyanion O^- from the tetrahedral intermediate in amide or ester cleavage would occupy the oxyanion hole.

Figure 6.5.24 shows an interactive iCn3D model of the active site of the phenylethane boronic acid (PBA) complex of alpha-chymotrypsin (6cha).



NCBI iCn3D Figure 6.5.24: Active site of the phenylethane boronic acid (PBA) complex of alpha-chymotrypsin (6cha). (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...qXUxbrakurYmx6

Many enzymes have active site serines which act as nucleophilic catalysts in nucleophilic substitution reactions (usually hydrolysis). One such enzyme is acetylcholinesterase, which cleaves the neurotransmitter acetylcholine in the synapse of the neuromuscular junction (Figure 6.5.25).

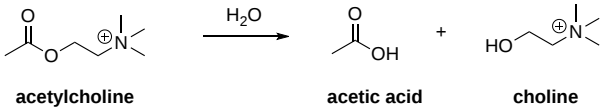


Figure 6.5.25: Reaction catalyzed by acetylcholinesterase

The neurotransmitter leads to muscle contraction when it binds its receptor on the muscle cell surface. The transmitter must not reside too long in the synapse, otherwise muscle contraction will continue in an uncontrolled fashion. To prevent this, a hydrolytic enzyme, acetylcholinesterase, a serine esterase found in the synapse, cleaves the transmitter, at rates close to diffusion controlled. Diisopropylphosphofluoridate (DIPF) also inhibits this enzyme, which effectively makes it a potent chemical warfare agent. Another fluoride-based inhibitor of this enzyme, sarin (Figure 6.5.26), is the most potent lethal chemical agent of this class known. Only 1 mg is necessary to kill a human being.

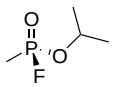


Figure 6.5.26: Sarin

Serine proteases have unique specificities to allow cleavage after a different subset of side chains. They cleave the peptide bond on the carboxylic acid side of specific amino acids and the specificity is determined by the size/shape/charge of amino acid side chain that fits into the enzyme's S1 binding pocket (Figure 6.5.27). Three chymotrypsin-like family members that share high sequence homology are the pancreatic digestive enzymes, trypsin, chymotrypsin and elastase. The protein cleavage sites of these enzymes vary. Trypsin cleaves proteins on the carboxylic side of basic residues, such as lysine and arginine, while chymotrypsin cleaves after aromatic hydrophobic amino acids, such as phenylalanine, tyrosine, and tryptophan. Elastase cleaves after small, hydrophobic residues, such as glycine, alanine, and valine. As shown in Figure 6.5.27, variations in the amino acid residues within the binding pocket of these proteases, enables electrostatic interactions with the substrate and determines sequence specificity.

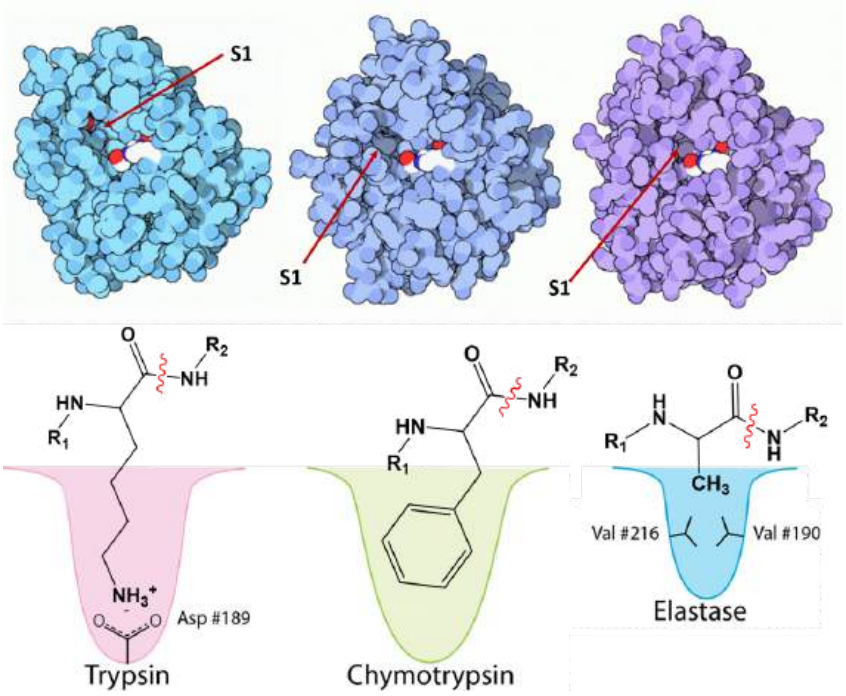


Figure 6.5.27: Substrate Specificity of Trypsin, Chymotrypsin, and Elastase. The *upper panel* shows the space-filling crystal structures of Trypsin, Chymotrypsin, and Elastase, respectively, with the S1 substrate binding pocket indicated. The *lower panel* depicts the S1 binding domains of each protease in more detail with important amino acid R-groups indicated. For Trypsin, an aspartate residue in the lower portion of the S1 pocket aid in electrostatic interactions with basic residues of the substrate. The Chymotrypsin S1 binding pocket is large and hydrophobic in nature accommodating aromatic residues of the substrate, while the Elastase S1 binding pocket is small and hydrophobic, only allowing other small and hydrophobic R-groups to dock in this location. Image modified from: Goodsell, D. (2012) Molecule of the Month, Protein Database and Aleia Kim

A schematic nomenclature developed by Berger and Schechter is often used to show the sites on the substrate (labeled P3, P2, P1, P1', P2' and P3') referring to the products made after cleavage of the peptide/protein that is cleaved between P1 and P' (the scissile bond) and the corresponding sites on the protease (S3, S2, S1, S1', S2' and S3'). This is illustrated in Figure 6.5.28

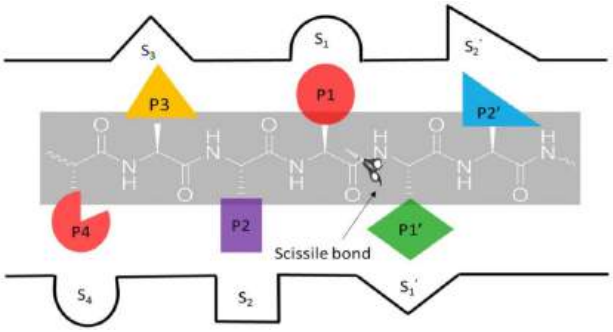


Figure 6.5.28: Nomenclature of denote key interacting groups in the enzyme (S_n) and in the substrate (P_n) around the scissile bond for protein cleavage

Serine proteases are just one type of endoproteases. However, they are extremely abundant in both prokaryotes and eukaryotes. Protease A, a chymotrypin-like protease from Stremptomyces griseus, has a very different primary sequence than chymotrypsin, but its overall tertiary structure is quite similar to chymotrypsin. The positions of the catalytic triad amino acids in the primary

sequences of the protein are very similar, indicating that the genes for the proteins diverged from a common precursor gene. In contrast, subtilisin, a serine protease from B. Subtilis, has both limited sequence and tertiary structure homology to chymotrypsin. However, when folded it also has a catalytic triad (Ser 221 - His 64 - Asp 32) similar to that of chymotrypsin (Ser 195 - His 57 - Asp 102). The alignment of the core structures of chymotrypsin (5cha, magenta) and subtilisin (1sbc, cyan), are shown in Figure 6.5.29

Figure 6.5.29: Convergent Evolution of the Serine Proteases, Chymotrypsin and Subtilisin.

The list of serine proteases is quite long. They are grouped in two broad categories - 1) those that are chymotrypsin-like and 2) those that are subtilisin-like. Though subtilisin-type and chymotrypsin-like enzymes use the same mechanism of action, including the catalytic triad, the enzymes are otherwise not related to each other by sequence and appear to have evolved independently. They are, thus, an example of ***convergent evolution*** - a process where evolution of different forms converge on a structure to provide a common function.

Proteases have multiple functions, other than in digestion, including degrading old or misfolded proteins and activating precursor proteins (such as clotting proteases and proteases involved in programmed cell death). In general, four different classes of proteases have been found, based on residues found in their active sites. Proteases can also be integral membrane proteins, and carry out their activities in the hydrophobic environment of the membrane. For example, aberrant cleavage of the amyloid precursor protein by the membrane protease presenillin can lead to the development of Alzheimer's Disease.

Table 6.5.4 below shows a classification of proteases based on their active site nucleophiles.

Class (active site)	Active Site Nucleophile	Location	Examples
Serine/Threonine Hydrolases	Ser/Thr	soluble	trypsin, chymotrypsin, subtilisin, elastase, clotting enzymes, proteasome
		membrane	Rhomboid family
Aspartic Hydrolases	H_2O activated by 2 Asps	soluble	pepsin, cathepsin, renin, HIV protease
		membrane	β-secretase (BACE), presenilin I, signal peptide peptidase
Cysteinyl Hydrolases	Cys	soluble	bromelain, papain, cathespsins, caspases
		membrane	?
Metallo Hydrolases	H_2O activated by 1 or 2 metal ions	soluble	thermolysin, angiotensin converting enzyme
		membrane	S2P family
Glutamate Hydrolases	Glu	.	eqolysins (fungal)

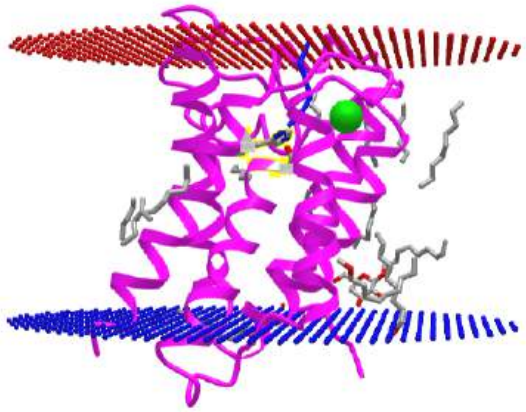
Asparagine Lysases (EC4) (elimination rx which are self-cleavage and hence not catalytic)	Asn	.	Tsh autotransporter E. Coli

Table 6.5.4: Protease classification

How do integral membrane protease catalyze the hydrolysis (using water) of transmembrane domains in proteins, given the hydrophobic environment of the bilayer? The **rhomboid** class of membrane proteases, which are found in prokaryotic and eukaryotic cells, is one of the most conserved membrane proteins in nature. Instead of using a catalytic triad, these serine proteases use a dyad of Ser 201 as a nucleophile and His 254 as a general acid/base.

The chief requirement for protein substrates of rhomboids is the presence of a transmembrane domain in the target protein. No specific amino acid sequence seems to be required for specificity of one particular substrate, the drosophila transmembrane protein spitz found in Golgi membranes. On cleavage of this protein, the remaining part of the protein is released as a water soluble protein to the lumen of the Golgi where it can eventually be released from the cell. The soluble protein fragment that is released from the cell contains an epidermal growth factor domain.

The structure of a rhomboid protease, GlpG (EC:3.4.21.105), from E. Coli, was determined. It is a serine protease with a catalytic dyad (Ser 201 and His 254) instead of a triad as in most serine proteases. This transmembrane protein has 6 transmembrane helices. The enzyme has a polar active site at the bottom of a V-shape opening situated laterally in the membrane. The active site His and Ser residues are deep in this V-shaped cleft, well below the surface of the membrane. Access to the transmembrane strand of the protein substrate is blocked by a loop, which must be gated open to allow substrate access between the V-shaped gap between helices S1 and S3. Ser 201 (nucleophile) and His 254 (general base/acid) are essential for activity. The active site His 254 can be covalently modified with different chloromethylketone peptide derivatives. Figure 6.5.30 shows an interactive iCn3D model of the Rhomboid intramembrane protease GlpG 4QO2.



NCBI iCn3D Figure 6.5.30: Rhomboid intramembrane protease GlpG (4QO2). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...ft9vFn7WqcujM7

Proteolytic enzymes (also termed peptidases, proteases and proteinases) are found in all living organisms, from viruses to animals and humans. Proteolytic enzymes have great medical and pharmaceutical importance due to their key role in biological processes and in the life-cycle of many pathogens. Proteases are extensively applied enzymes in several sectors of industry and biotechnology, furthermore, numerous research applications require their use, including production of Klenow fragments, peptide synthesis, digestion of unwanted proteins during nucleic acid purification, cell culturing and tissue dissociation, preparation of recombinant antibody fragments for research, diagnostics and therapy, and the exploration of the structure-function relationships.

Proteolytic enzymes belong to the hydrolase class of enzymes and are grouped into the subclass of the peptide hydrolases or peptidases. Depending on the site of enzyme action, the proteases can also be subdivided into exopeptidases (like chymotrypsin) or endopeptidases (like carboxypeptidase A) as we will discuss next. Exopeptidases, such as aminopeptidases and carboxypeptidases catalyze the hydrolysis of the peptide bonds near the N- or C-terminal ends of the substrate, respectively. Endopeptidases cleave peptide bonds at internal locations within the peptide sequence. These differences are illustrated in Figure 6.5.31. Proteases may also be nonspecific and cleave all peptide bonds equally or they may be highly sequence specific and only cleave peptides after certain residues or within specific localized sequences.

Figure 6.5.31: Peptide bond cleavage by various proteases (Figure from Mótyán, J.A., et al. (2013) *Biomolecules 3*(4), 923-942).

The action of proteolytic enzymes is essential in many physiological processes. For example, proteases function in the digestion of food proteins, protein turnover, cell division, the blood-clotting cascade, signal transduction, processing of polypeptide hormones, apoptosis and the life-cycle of several disease-causing organisms including the replication of retroviruses such as the human immunodeficiency virus (HIV). Due to their key role in the life-cycle of many hosts and pathogens they have great medical, pharmaceutical, and academic importance.

It was estimated previously that about 2% of the human genes encode proteolytic enzymes and due to their necessity in many biological processes, proteases have become important therapeutic targets. They are intensively studied to explore their structure-function relationships, to investigate their interactions with the substrates and inhibitors, to develop therapeutic agents for antiviral therapies or to improve their thermostability, efficiency and to change their specificity by protein engineering for industrial or therapeutic purposes.

The following section material (between the two horizontal lines) is not found in most biochemistry textbooks, so it could be considered optional. At the same time, it offers another unique way of understanding enzymes that ultimately will lead to a better understanding of how they function.

6.5.3.1.2: Enzyme catalysis in organic solvents

In our earlier lists, we mentioned changing the solvent and exploring its affect on enzyme catalysis. It might seem a bit wild, but as we saw with the rhomboid protease, some enzymes work in hydrophobic environments. Also, lipases work at the boundary between the aqueous and hydrophobic worlds. For those interested, let's see what happens when we change solvents. These including putting the enzyme in various solvents, or mixtures of solvents, as described below:

- Water miscible solvents like ethanol and acetone were added. If the water concentration was high enough, activity remained.
- Biphasic mixtures in which an aqueous solution of an enzyme was emulsified in a water immiscible solvent like chloroform or ethylacetate. The substrate would partition into both phases, while the product hopefully would end up into the organic phase.
- Nearly nonaqueous solvents, with a few % water at less than the solubility limits of water.
- Anhydrous organic solvents (0.01% water). It is this case that is most astonishing since enzymatic activity is often retained.

It is important to realize that in this last case, the enzyme is **not in solution**. It is rather in **suspension** and acts as a **heterogeneous catalyst**, much like palladium acts as a heterogeneous catalyst in the hydrogenation of alkenes. The suspension must be mixed vigorously and then sonicated to produce small suspended particles, so diffusion of reactants into the enzyme and out is not rate limiting. Let's explore the activity of chymotrypsin in a nonpolar solvent.

Why aren't the enzymes inactive? Surely it must seem ridiculous that they aren't, since as we learned earlier, proteins are not that stable. A 100 amino acid protein on average is stabilized only about 10 kcal/mol (41 kJ/mol) over the denatured state, or the equivalent of a few H bonds. Surely the hydrophobic effect, one of the dominant contributors to protein folding and stability, would not stabilize the native structure of enzymes in nonpolar organic solvents, and the protein would denature. It doesn't however! Maybe the real question should be not whether water is necessary, but rather how much water is necessary. The enzyme can't "see" more than a monolayer or so of water around it. The data suggests that the nature of the organic solvent is very important. The most hydrophobic solvents are best in terms of their ability to maintain active enzymes! Chymotrypsin retains 10^4 more activity in octane than pyridine (see $\mathbf{k_{cat}/K_m}$ below), which is more hydrophilic than octane. The more polar the solvent, the more it can strip bound water away from the protein. If you add 1.5% water to acetone, the bound water increases from 1.2 to 2.4%, and the activity of chymotrypsin increases 1000 fold.

Table 6.5.5 below shows chymotrypsin activity in organic solvents.

Solvent	Structure	$\mathbf{k_{cat}}/K_m$ ($M^{-1}min^{-1}$)	relative ratio k_{cat}/K_m	$\mathbf{H_2O}$ bound to enzyme (%, w/w)
Octane		63	15000x	2.5
Toluene		4.4	1000x	2.3
Tetrahydrofuran		0.27	175x	1.6
Acetone		0.022	5.5x	1.2
Pyridine		<0.004	1x (.004)	1.0

Table 6.5.5: Chymotrypsin activity in organic solvents

Consider the following questions:

- *How much water do the enzymes need?* 1 molecule of chymotrypsin in octane has less than 50 molecules of water associated and can demonstrate activity. To form a monolayer requires about 500 water molecules. Water can be added which presumably leads to more bound water and higher activity.
- *How stable are the enzymes?* Denaturation requires conformational flexibility, which apparently requires water. The half-life of chymotrypsin in water at 60 ^oC is minutes, but in octane at 100 ^oC it is hours. At 20 ^oC, the half-life in water is a few days, but in octane it is greater than 6 months. Remember two factors contribute to stability: 1. The protein can denature at high temperatures. 2. Chymotrypsin is a protease, it can cleave itself in an autoproteolytic reaction.

Table 6.5.6 below shows the half-life of chymotrypsin activity in water and octane

Solvent	60^{o}C	100^{o}C	20^{o}C
water	minutes	-	few days

octane	-	hours	> 6 months

Table 6.5.6: Half-life of chymotrypsin activity in water and octane at different temperatures

- *Is the enzyme specificity changed?* The net binding energy is a function of the binding energy of the substrate - the binding energy of the water, since water must be displaced from the active site on binding. In an anhydrous solvent, specificity changes must be expected. For chymotrypsin, the driving force for binding of substrates in water is mostly hydrophobic. In water, the k_{cat}/K_M for the reaction of N-acetyl-L-Ser-esters is reduced 50,000 times compared to the Phe ester. However, in octane, chymotrypsin is three times more active toward Ser esters than Phe esters.

Table 6.5.7 shows specificity changes in chymotrypsin in water and octane

Substrate	k_{cat}/K_m	
	solvent: H2O	solvent: Octane
N-acetyl-L-Ser-ester	1x	3x
N-acetyl-L-Phe-ester	50,000x	1x

Table 6.5.7: Specificity changes in chymotrypsin in water and octane

Now, consider competitive inhibitors. Naphthalene binds 18 times more tightly than 1-naphthoic acid, but in octane, chymotrypsin binds naphthoic acid 310 times as tightly. Likewise the ratio of [k_{cat}/K_m (L isomer)]/[k_{cat}/K_m (D isomer)] of N-acetyl-D- or N-acetyl-L-Ala-chloroethyl esters is 1000-10,000 in water, but less than 10 in octane.

Table 6.5.8 shows chymotrypsin inhibition constants in water and octane.

Inhibitor	Inhibition Constant Ki (nM)	
	In water	In Octane
Benzene	21	1000
Benzoic acid	140	40
Toluene	12	1200
Phenylacetic acid	160	25
Naphthalene	0.4	1100
1-Naphthoic acid	7.2	3

Table 6.5.8: Chymotrypsin inhibition constants in water and octane

Can new reactions be carried out in nonpolar solvents? The quick answers is yes, since reactions in aqueous solutions can be unfavorable due to low K_{eq} values, side reactions, or insolubility of reactants. Consider lipases, which cleave fatty acid esters by hydrolysis in aqueous solutions. In nonaqueous solutions, reactions such as transesterification or ammonolysis can be performed.

Enzymes are clearly active in organic solvents which appears to contradict our central concepts of protein stability. Two reasons could could explain this stability:

1. It is possible that from a thermodynamic view, the enzyme is stable in organic solvents. However, as was discussed above, this is **inconceivable** given the delicate balance of noncovalent and hydrophobic interactions required for protein stability.
2. The second reason must win the day: the protein is unable to unfold from a **kinetic** point of view. Conformational flexibility is required for denaturation. This must require water as the solvent. Denaturation in organic solvents is kinetically, not

thermodynamically controlled.

A specific example helps illustrate the effects of different solvents on chymotrypsin activity. Dry chymotrypsin can be dissolved in DMSO, a water miscible solvent. In this solvent it is completely and irreversibly denatured. If it is now diluted 50X with acetone with 3% water, no activity is observed. (In the final dilution, the concentrations of solvents are 98% acetone, 2.9% water, and 2% DMSO.) However, if dry chymotrypsin was added to a mixture of 98% acetone, 2.9% water, and 2% DMSO, the enzyme is very active. We end up with the same final solvent state, but in the first case the enzyme has no activity while in the second case it retains activity. These ideas are illustrated in Figure 6.5.32

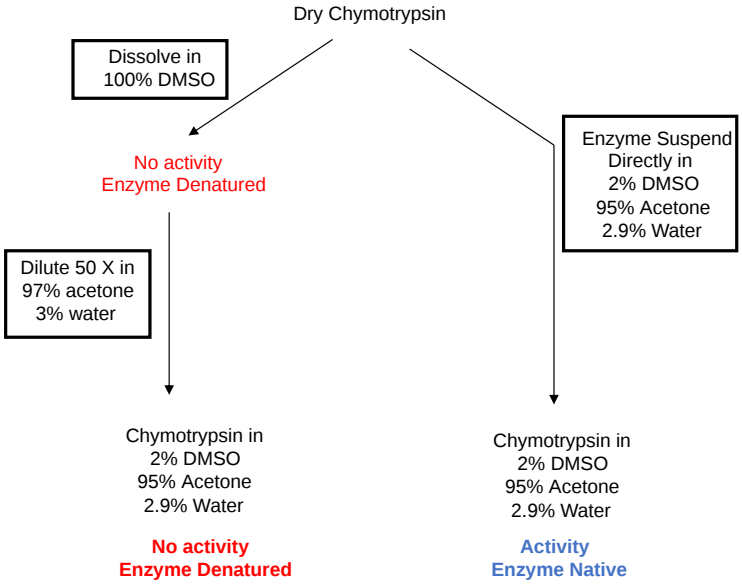


Figure 6.5.32: Chymotrypsin activity in acetone depends on order of solvent addition

Dry enzymes added to a concentrated water-miscible organic solvent (like DMSO) will dissolve and surely denature, but will retain activity when added to a concentrated water-immiscible solvent (like octane), in which the enzyme will not dissolve but stay in suspension.

It appears the enzymes have very restricted conformational mobility in nonpolar solvents. By lyophilizing (freeze-drying) the enzyme against a specific ligand, a given conformation of a protein can be trapped or literally imprinted onto the enzyme. For example, if the enzyme is dialyzed against a competitive inhibitor (which can be extracted by the organic solvent), freeze-dried to remove water, and then added to a nonpolar solvent, the enzyme activity of the "imprinted" enzyme in nonpolar solvents is as much as 100x as great as when no inhibitor was present during the dialysis. If chymotrypsin is lyophilized from solutions of different pHs, the resulting curve of V/K_m for ester hydrolysis in octane is bell-shaped with the initial rise in activity reaching half-maximum activity at a pH of around 6.0 and a fall in activity reaching half-maximum at pH of approximately 9.

Use of enzymes in organic solvent allows new routes to organic synthesis. Enzymes, which are so useful in synthetic reactions, are:

- stereoselective - can differentiate between enantiomers and between prochiral substrates
- regioselective - can differentiate between identical functional groups in a single substrate
- chemoselective - can differentiate between different functional groups in a substrate (such as between a hydroxyl group and an amine for an acylation reaction)

Enzyme in anhydrous organic solvents are useful (from a synthetic point) not only since new types of reactions can be catalyzed (such as transesterification, ammonolysis, thiolysis) but also because the stereoselectivity, regioselectivity, and chemoselectivity of the enzyme often changes from activities of the enzyme in water.

Organic reactions are usually conducted in organic solvents, since many organic molecules react with water, and the reagents and products are usually not soluble in water. In a manner analogous to using an enzyme as a heterogeneous catalyst in nonpolar solvent, Sharpless is pioneering a technique to conduct organic reactions in water. They (Narayan et al.) have shown that many unimolecular and bimolecular reactions occur faster in water than in organic solvents. As in enzyme catalysis in nonpolar solvent, the reactions must be mixed vigorously to disperse reactants in micro-drops (a suspension) in water, greatly increasing the surface

area that might allow water to act on transition states or intermediates to stabilize them through hydrogen bonding. They called these reactions "on water" reactions since reactants usually float on water. They have performed cycloadditions, alkene reactions, Claisen rearrangments, and nucleophilic substitution reactions using this process. One cycloaddition reaction went to completion in ten minutes at room temperature, compared to 18 hours in methanol and 120 in toluene. Adding nonpolar solvent at certain times greatly increased the rate of the reaction.

6.5.3.2: Carboxypeptidase A

This enzyme (EC 3.4.17.1) cleaves the C-terminal amino acid from a protein through a hydrolysis reaction. As such it is an **exoprotease** (not an endoprotease which cleaves proteins internally within the sequence). In terms of selectivity toward C-terminal amino acids, its activity is increased if the C-terminal side chain group is aromatic or branched aliphatic (Phe, Tyr, Trp, Leu or Ile). X-Ray structures of the enzyme with and without a competitive inhibitor show a large conformational change at the active site when inhibitor or substrate is bound. Without inhibitor, several waters occupy the active site. When an inhibitor (and presumably, by extension, a substrate) is bound, the water leaves (which is entropically favored), and Tyr 248 swings around from near the surface of the protein into the active site to interact with the carboxyl group of the bound molecule, a distance of motion equal to about 1/4 the diameter of the protein. This effectively closes off the active site and expels the water.

A Zn^{2+} ion is present at the active site. It is bound by His 69, His 196, Glu 72, and finally a water molecule as the fourth ligand. A hydrophobic pocket that interacts with the phenolic group of the substrate accounts for the specificity of the protein. In the catalytic mechanism, Zn^{2+} might have several roles. In one, it may help a coordinated water to be more nucleophilic by either polarizing the water or converting it to a more potent nucleophile OH^-. It might also stabilize developing negative charges in the transition state and in an intermediate. Two possible mechanisms have been offered.

The Water Pathway. In this proposed mechanism, water acts as a nucleophile, and is deprotonated by Glu 270, acting as a general base. Glu 270, along with Zn^{2+}, helps to promote dissociation of a proton from the bound water, making it to a better nucleophile. Water attacks the electrophilic carbon of the sessile bond, forming a tetrahedral intermediate. The tetrahedral intermediate then collapses, expelling the alkoxy leaving group, which picks up a proton from Glu 270, now acting as a general acid catalyst. People used to believe that Tyr 248 acted as a general acid, but mutagenesis showed that Tyr 248 can be replaced with Phe 248 without significant effect on the rate of the reaction. A simplified reaction reaction is shown in Figure 6.5.33

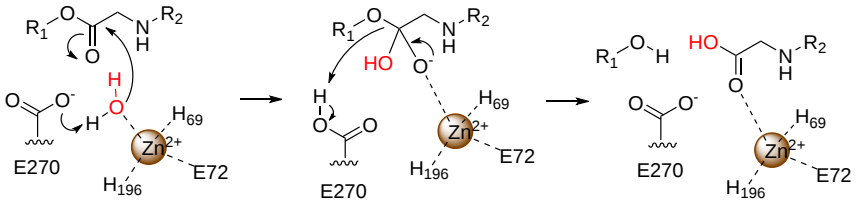


Figure 6.5.33: Water pathway mechanism for carboxypeptide A. After Wu et al. J Phys Chem B. 2010 July 22; 114(28): 9259–9267. doi:10.1021/jp101448j

Nucleophilic Pathway. In this pathway, Glu 270 is the primary initial nucleophile in the formation of the initial tetrahedral intermediate. The role of Zn^{2+} is in charge stabilization. This pathway is illustrated in Figure 6.5.34

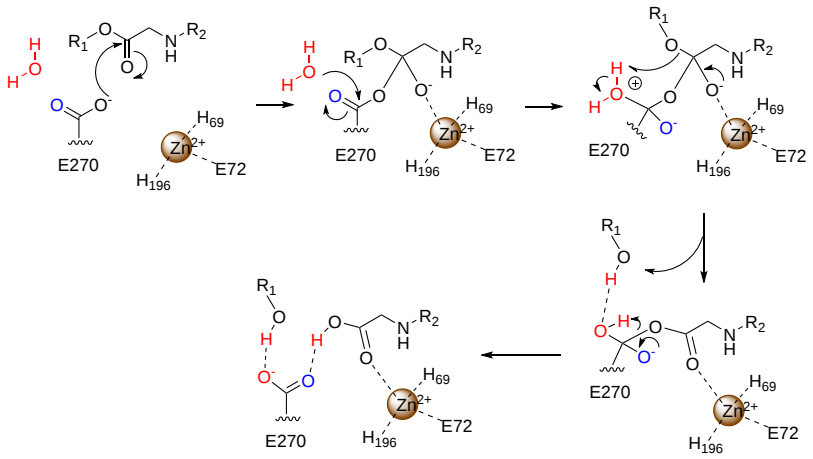
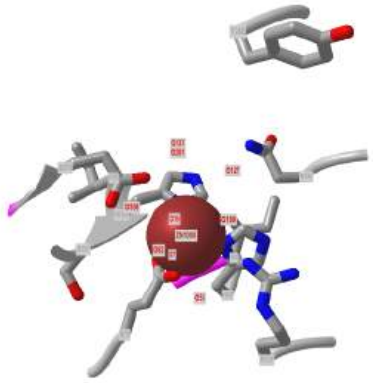


Figure 6.5.34: Nucleophilic Pathway after Wu et al. J Phys Chem B. 2010 July 22; 114(28): 9259–9267. doi:10.1021/jp101448j

Figure 6.5.35 shows an interactive iCn3D model of the active site of bovine carboxypeptidase in the absence of a substrate or inhibitor (1M4L). The Zn^{2+} ion is shown as a red sphere.



NCBI iCn3D Figure 6.5.35: Bovine carboxypeptidase A (1M4L) (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...sX1NYbGav1CFY6

Note how far Tyr248 is away from the active site in the model. Glu72 and Glu270 are negatively charged in the resting state of the enzyme at pH 7.5. The values are much higher (weaker acid) than solution pK_a of the side chain of glutamic acid. Also the water bound to the Zn^{2+} is long enough to suggest that the water is neutral and not in the form of OH^- in this form of the enzyme. If OH^- were present, the distance between it and the Zn^{2+} would be shorter due to the great electrostatic force.

Figure 6.5.36 shows an interactive iCn3D model of the active site of bovine carboxypeptidase bound to the inhibitor aminocarbonylphenylalanine (1HDU). The Zn^{2+} ion is shown as a red sphere.

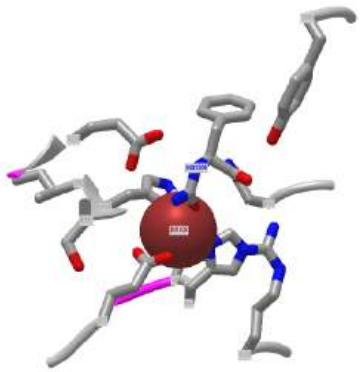



Figure 6.5.36 Bovine carboxypeptidase A bound to bound to the inhibitor aminocarbonylphenylalanine (1HDU). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...vWt2qeZtUrwj78

Note the closer proximity of tyrosine 248 to the active site.

6.5.3.3: Lysozyme

Lysozyme (EC 3.2.1.17), found in cells and secretions of vertebrates but also in viruses which infect bacteria, cleaves peptidoglycan GlcNAc (β-1,4) MurNAc repeat linkages (NAG-NAM) in the cell walls of bacteria and the GlcNAc(β-1,4) GlcNAc (poly-NAG) in chitin, found in the cells walls of certain fungi. Since these polymers are hydrophilic, the active site of the enzyme would be expected to contain a solvent-accessible channel into which the polymer could bind. The crystal structures of lysozyme and complexes of lysozyme and NAG have been solved to high resolution. The inhibitors and substrates form strong H bonds and some hydrophobic interactions with the enzyme cleft. Kinetic studies using $(NAG)_n$ polymers show a sharp increase in k_{cat} as n increases from 4 to 5. The k_{cat} for $(NAG)_6$ and $(NAG\text{-}NAM)_3$ are similar. Models studies have shown that for catalysis to occur, $(NAG\text{-}NAM)_3$ binds to the active site with each sugar in the chair conformation, except the fourth which is distorted to a half chair form. This labilizes the glycosidic link between the 4^{th} and 5^{th} sugars. Additional studies show that if the sugars that fit into the binding site are labeled A-F, then because of the bulky lactyl substituent on the NAM, residues C and E cannot be NAM, which suggests that B, D and F must be NAM residues. Cleavage occurs between residues D and E.

A review of the chemistry of glycosidic bond (an acetal) formation and cleavage shows the acetal cleavage is catalyzed by acids and proceeds by way of an oxonium ion which exists in resonance form as a carbocation. A reaction mechanism of hemiacetal/acetal formation and cleavage is illustrated in Figure 6.5.37.

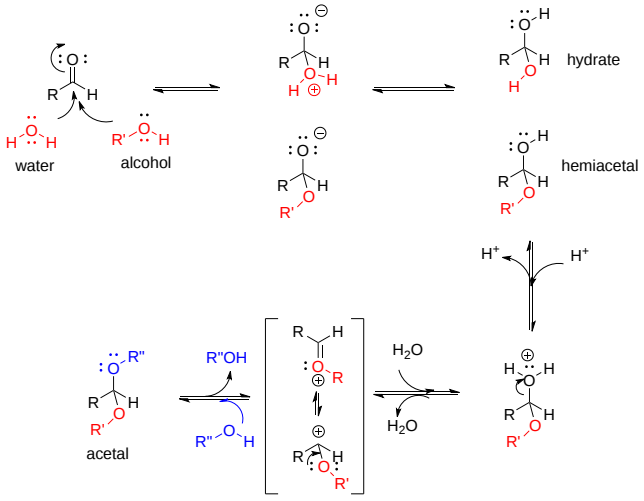


Figure 6.5.37: A reaction mechanism of hemiacetal/acetal formation and cleavage

Catalysis by the enzyme involves Glu 35 and Asp 52 which are in the active site. Asp 52 is surrounded by polar groups but Glu 35 is in a hydrophobic environment. This should increase the apparent pK_a of Glu 35, making it less likely to donate a proton and acquire a negative charge at low pH values, making it a better general acid at higher pH values. Here is a possible general mechanism:

- binding of a hexasaccharide unit of the peptidoglycan with concomitant distortion of the NAM.
- protonation of the sessile acetal O by the general acid Glu 35 (with the elevated pK_a), which facilitates cleavage of the glycosidic link and formation of the resonance stabilized oxonium ion.
- Asp 52 stabilizes the positive oxonium through electrostatic catalysis. The distorted half-chair form of the NAM stabilizes the oxonium which requires co-planarity of the substituents attached to the sp^2 hybridized carbon of the carbocation resonant form (much like we saw with the planar peptide bond).
- water attacks the stabilized carbocation, forming the hemiacetal with release of the extra proton from water to the deprotonated Glu 35 reforming the general acid catalysis.

Part of a mechanism illustrating the roles of Glu 35 and Asp 52 is shown below in Figure 6.5.38

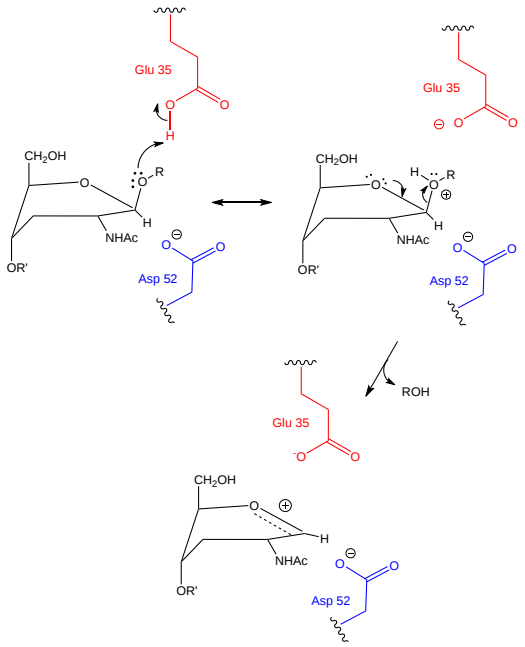


Figure 6.5.38: Mechanism of acetal cleavage by lysozyme

Binding and distortion of the D substituent of the substrate (to the half chair form as shown above) occurs before catalysis. Since this distortion helps stabilize the oxonium ion intermediate, it presumably stabilizes the transition state as well. Hence this enzyme appears to bind the transition state more tightly than the free, undistorted substrate, which is yet another method of catalysis.

pH studies show that side chains with pKa's of 3.5 and 6.3 are required for activity. These presumably correspond to Asp 52 and Glu 35, respectively. If the carboxy groups of lysozyme are chemically modified in the presence of a competitive inhibitor of the enzyme, the only protected carboxy groups are Asp 52 and Glu 35.

In an alternative mechanism, Asp 52 acts as a nucleophilic catalyst and forms a covalent bond with NAM, expelling a NAG leaving group with Glu 35 acting as a general acid as shown in Figure 6.5.39. This alternative mechanism also is consistent with other β-glycosidic bond cleavage enzyme. Substrate distortion is also important in this alternative mechanism.

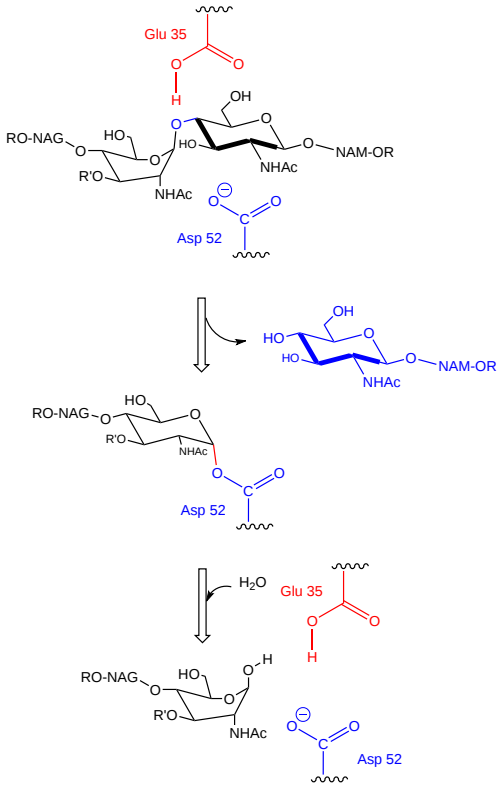
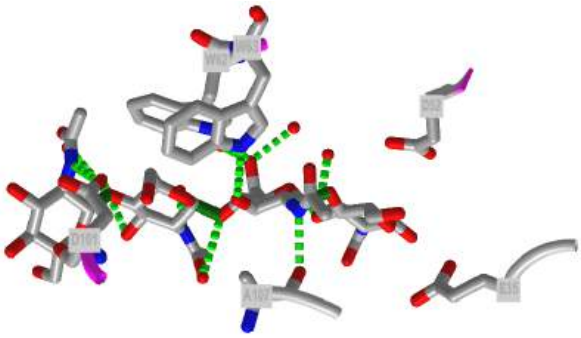


Figure 6.5.39: Alternative mechanism for lysozyme catalysis employing Asp 52 as a nucleophilc catalyst (after Vocadlo et al., Nature, **412** (2001), https://www.nature.com/articles/35090602.

Recent structural work shows that Asp 52 is involved in a strong hydrogen bond network that might preclude its ability to form a covalent bond with the glycan substrate. An earlier structure (1H6M) did show a covalent bond.

Figure 6.5.40 shows an interactive iCn3D model of the active site of hen egg white lysozyme bound to a $(NAG)_4$ glycan (7BR5). Note the positions of E35 and D52.



NCBI iCn3D Figure 6.5.40: Interactions of hen egg white lysozyme with bound $(NAG)_4$ glycan (7BR5) (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?mnZuP4W4pfTxKRmX7

6.5.4: Summary

In this section, we explored some of the biochemical arrow pushing conventions for S_N2 reactions, including methyltransferase enzymes that are dependent on the coenzyme SAM, and kinases, which transfer a phosphate group. We saw that the nucleophilic acyl substitution mechanisms for chymotrypsin and carboxypeptidase can be quite abbreviated such that it takes some biochemical intuition to recognize that a tetrahedral intermediate is operative. There are two major proposed mechanisms for carboxypeptidase, which propose distinct roles for the metal ion in the reaction.

The investigation of chymotrypsin's mechanism by substrate specificity experiments (changing the substrate), altering the pH, mutagenesis and reaction with irreversible inhibitors (changing the protein) gives rich information that can be used to deduce enzyme mechanisms. Coupled with structural data, these investigations reveal key details about enzyme structure-function, and allow biochemists to propose reasonable enzyme mechanisms.

6.6: Enzymes and Protein Regulation

Many different mechanisms control protein activity within a cell. The primary sequence of a protein is a main determinant of protein folding and final conformation as well as biochemical activity, stability, and half-life. However, at any given moment, the proteome, the full complement of proteins within a cell, is up to two or three orders of magnitude more complex than the encoding genomes would predict. This chapter will give an overview of the major mechanisms utilized by biological systems to regulate protein functions after the protein has been synthesized. Note that these mechanisms seldom work in isolation. There are multiple levels of protein control that function at any given time and in response to many different environmental cues and signals. We will discuss many ways to regulate protein activity. At the end, we will discuss regulation through the use of different **isozymes** of an enzyme, which are variants of an enzyme arising from differential splicing of a gene or arising from slightly different genes that arose from a common precursor gene. We will focus on cyclooxygenases, the target of so many medicinal drugs.

6.6.1: Post-Translational Modifications (PTMs).

The human genome contains approximately 20,000 genes. When analyzing the transcriptome, it becomes apparent that the genome becomes amplified by the wide array of splice variants that can occur when primary RNA transcripts are spliced to form mature messenger RNAs. There are about 200,000 expressed transcripts of which about 146,00 encode proteins, with about 7.4 transcripts per gene. A summary of human gene transcripts is shown in Table 6.6.1 below.

Type	GTEx dataset	Expressed all genes	Expressed protein-coding genes
Transcripts per Gene	3.42	3.63	7.43
Gene counts	58,219	53,539	19,591
Transcript counts	199,324	194,146	145,571

Table 6.6.1: **Numbers of transcript and gene in the GTEx dataset.** Tung, KF., Pan, CY., Chen, CH. *et al.* Top-ranked expressed gene transcripts of human protein-coding genes investigated with GTEx dataset. *Sci Rep* **10**, 16245 (2020). https://doi.org/10.1038/s41598-020-73081-5. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

The number of distinct primary structures of proteins (**proteoforms**) from 20,000 human genes is amplified again through posttranslational modifications (PTMs) of proteins, which produce up to a million different variants. PTMs are present in both eukaryotes and prokaryotes, but are more common in eukaryotic cells, in which about 5% of the genome is dedicated to enzymes that carry out posttranslational modifications of proteins. We discussed chemical modification of specific amino acids in Chapter 3.1: Amino Acids and Peptides.

Protein PTM results from the enzymatic or nonenzymatic attachment of specific chemical groups to amino acid side chains. Such modifications occur either following protein translation or concomitant with translation. PTM influences both protein structure and physiological and cellular functions. Examples of enzymatic PTMs include phosphorylation, glycosylation, acetylation, methylation, sumoylation, palmitoylation, biotinylation, ubiquitinylation, nitration, chlorination, and oxidation/reduction. Nonenzymatic PTMs include glycation, nitrosylation, oxidation/reduction, acetylation, and succinylation. Some rare and unconventional PTMs, such as glypiation, neddylation, siderophorylation, AMPylation, and cholesteroylation, are also known to influence protein structure and function. Note that many of these modifications are not made in isolation. It is common for proteins to have several different types of modifications and that these modifications can differ depending on the tissue type and environmental circumstances present. The major PTMs in eukaryotes, their target amino acid residue(s), and the types of enzyme(s) or protein(s) involved are shown in Table 6.6.2 below.

Posttranslational modification	Target amino acid residue(s)	Enzyme(s) or proteins involved
Phosphorylation	Tyrosine, serine, threonine	Kinases, phosphatases
Glycosylation N-linked	Asparagine	Glycosyltransferases, deglycosylases
Glycosylation O-linked	Serine/threonine	Glycosyltransferases, deglycosylases
Acetylation	Lysine	Acetyltransferases, deacetylases
Methylation	Lysine, arginine	Methyltransferases, demethylases
Ubiquitination	Lysine	Ubiquitin-activating enzymes, ubiquitin-conjugating enzymes, ubiquitin ligases, deubiquitinases
Sumoylation	Lysine	Ubiquitin-activating enzymes, ubiquitin-conjugating enzymes, ubiquitin ligases, deubiquitinases
Myristoylation	Glycine	N-Myristoyltransferases
Prenylation	Cysteine	Farnesyltransferases, geranyl geranyltransferases
Palmitoylation	Cysteine	DHHC protein acyltransferases, acyl-protein thioesterases
Sulfation	Tyrosine	Sulfatases, desulfatases
S-Nitrosylation	Cysteine, methionine	
Glycation	Lysine	
Nitration	Tyrosine	Denitrases
Chlorination	Tyrosine	Myeloperoxidases
Oxidation/reduction	Cysteine	Peroxidases, oxidases, glutathione, thioredoxin
Carbonylation	Lysine, proline, arginine, threonine	

Table 6.6.2: **Common Protein Post-Translational Modifications, Their Target Amino Acid Residues, and the Enzyme(s) or Proteins Involved.** *Santos, A.L, and Lindner, A.B. (2017) Oxidatve Medicine and Longevity, Article ID: 5716409*

PTMs can exert their effect through many different structural changes, including opening and closing the active site, changing the conformation and electrostatic properties of binding sites, changing the flexibility of the chain, increasing or decreasing intrinsically disordered regions, altering protein-protein interactions, etc. These effects can also arise from the binding of small molecules.

6.6.1.1: Methods to Detect Protein Posttranslational Modifications

Specific amino acid residues are subjected to PTMs depending on the chemistry of the reaction and the sequence specificity of the enzyme involved. Initially, the detection of PTMs was carried out by various analytical methods, such as radiolabeling of the proteins, thin-layer chromatography, column chromatography, and/or polyacrylamide gel electrophoresis. Other methods, such as protein sequencing by Edman degradation and Western blotting using protein-specific antibodies, have since been developed. Currently, antibody-based detection methods and mass spectrometry-based proteomic analysis are the predominant methods used to detect and analyze PTMs. However, mass spectrometric methods are the only available tool to perform global or large-scale PTM analysis.

Antibody-based methods mostly rely on the availability of antibodies that can specifically recognize a modified amino acid residue within a protein or peptide. Such antibodies can be polyclonal or monoclonal and are developed against either the modified peptide/protein or against the modified amino acid. Moreover, antibody-based detection and quantification of PTMs on protein/peptide samples can be performed by two methods: chemiluminescence-based Western blotting and absorbance/fluorescence-based ELISA. However, the detection of PTMs depends entirely on the recognition site of the antibody used. If the antibody detects only the modified amino acid, additional analysis—for instance, protein/peptide isolation and sequencing—should be performed to detect the sequence context of the modification. However, if the antibody detects the PTM

within a specific sequence context, the presence of PTM at other sites will remain undetected (ie the antibody will be specific for only that single modification).

Mass spectrometric detection of specific PTMs is based on mass changes. Depending on the type of modification, a specific change in the mass of the modified amino acid or peptide occurs. Subsequently, the change in mass is detected by the mass spectrometer to identify the presence of a PTM in a peptide sample. Using tandem mass spectrometric methods, identification of the specific site of PTM can be achieved by subsequent fragmentation and sequencing of the relevant peptide. Yet, technical challenges hamper MS-based investigation of biologically important PTMs, such as ADP-ribosylation, one of the key signaling molecules that regulate DNA repair, a critical process in maintaining genome stability that is compromised in cancer and aging.

Data increasingly implicate PTMs not only during aging and/or under pathological conditions but in the normal functioning of the cell. In turn, PTMs are increasingly studied for their role in health and disease. For example, the precise and accurate measurement of distinct PTM-containing moieties offers potential biomarker utility to aid early diagnosis, prognosis, monitoring response to therapy and decisions regarding inclusion in clinical trials as new medicines are developed. However, technical difficulties limit these studies, leaving many unanswered questions. The identification of unknown/unexpected PTMs by proteomic data reanalysis is an emerging subfield of proteomics recently boosted by the increased availability of raw data shared in public repositories. Notably, though, a sampling of the proteome in a given organism or cell provides only a snapshot of a highly dynamic process, confounding the analytical problem and ultimately arguing for time-resolved inventories. Thus, while many tools are currently available for the study of PTMs, new methods are needed to further advance the study of these modifications.

6.6.1.2: Examples of PTMs

Protein PTMs involve the covalent addition of some chemical group by enzymatic catalysis. Typically, an electrophilic fragment of a co-substrate is added to an electron-rich protein side chain, which acts as a nucleophile in the transfer. Common covalent protein PTMs include phosphorylation, acylation, methylation, sumoylation, ubiquitination, glycosylation, lipidation, oxidation and disulfide bond formation (either internal within a single protein or linking two protein/peptide chains together). Examples of common PTMs are provided in Figure \(\PageIndex{1}\) below.

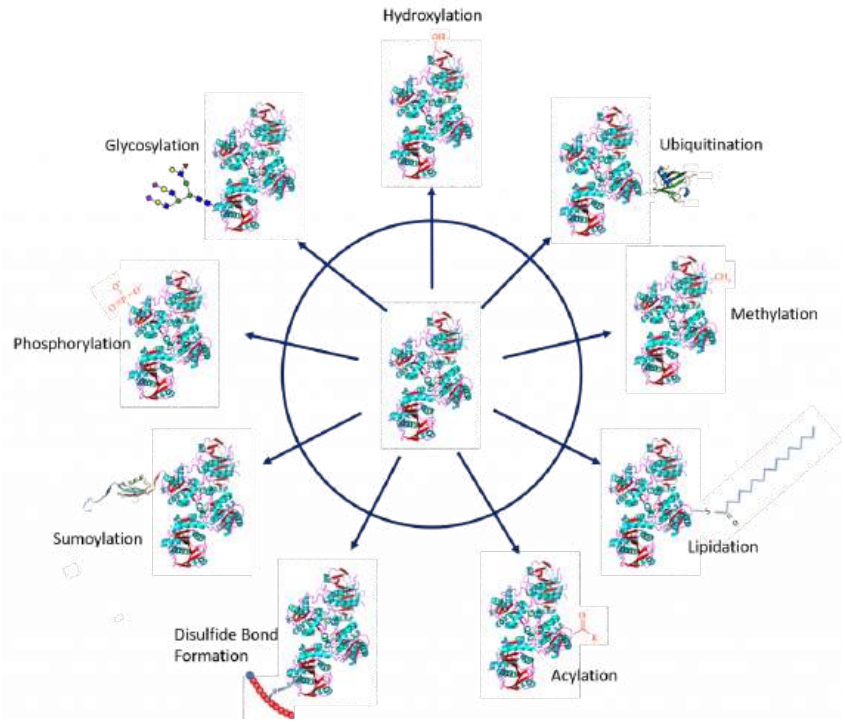


Figure 6.6.1: **Major types of covalent additions to protein side chains.** Common modifications include oxidation, ubiquitination, methylation, lipidation, acylation, disulfide bond formation, sumoylation, phosphorylation, and glycosylation. *Figure modified from: Santos, A.L, and Lindner, A.B. (2017) Oxidative Medicine and Longevity, Article ID: 5716409, Goettig, P. (2016) Int. J. Mol. Sci. 17(12), 1969, Rogerdodd, Jakob Suckale, Akinlade, A., et al (2014) Int. Archives of Med 7(50):28, WilsonNR, and Sivart13*

Protein Phosphorylation

One of the most common posttranslational modifications, protein phosphorylation, is the reversible addition of a phosphoryl group from adenosine triphosphate (ATP) to amino acid side chains such as serine, threonine, and tyrosine residues as shown in Figure 6.6.2. This modification causes conformational changes that either (1) affect the catalytic activity to activate or inactivate the protein and/or cause the tendency of a protein to misfold and aggregate or (2) recruit other proteins to bind; both responses result in altered protein function and cell signaling. Phosphorylated proteins have critical and well-known functions in diverse cellular processes across eukaryotes, but phosphorylation also occurs in prokaryotic cells. In humans, about one-third of proteins are estimated to be substrates for phosphorylation. Indeed, phosphorylated proteins are now identified and characterized by high-throughput phosphoproteomics studies. The reversibility of protein phosphorylation is attributed to the actions of ***kinases*** and ***phosphatases***, which phosphorylate and dephosphorylate substrates, respectively. The temporal and spatial balance of kinase and phosphatase concentrations within a cell mediates the size of its phosphoproteome.

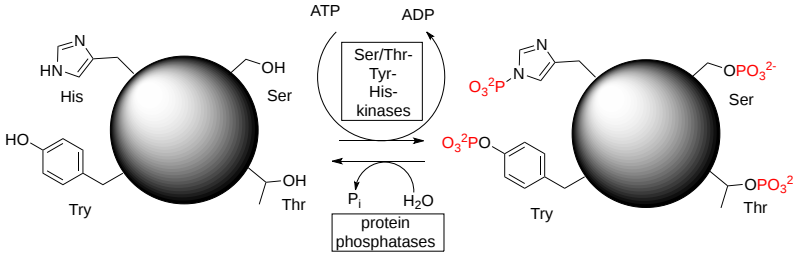


Figure 6.6.2: Amino acid phosphorylation in proteins

Protein Acylation

The simplest form of acylation is protein N-Acetylation. This occurs at the amino terminus amine and the ε-amino group of the lysine side chains through the action of acetylases. The acetylation of lysine side chains can be reversed through the actions of deacetylases (similar to the combined actions of protein kinases and phosphatases). The general reactions of protein N-acetylation are shown in Figure 6.6.3. Interestingly, 80–90% of eukaryotic proteins are acetylated, yet the underlying biological significance remains unclear.

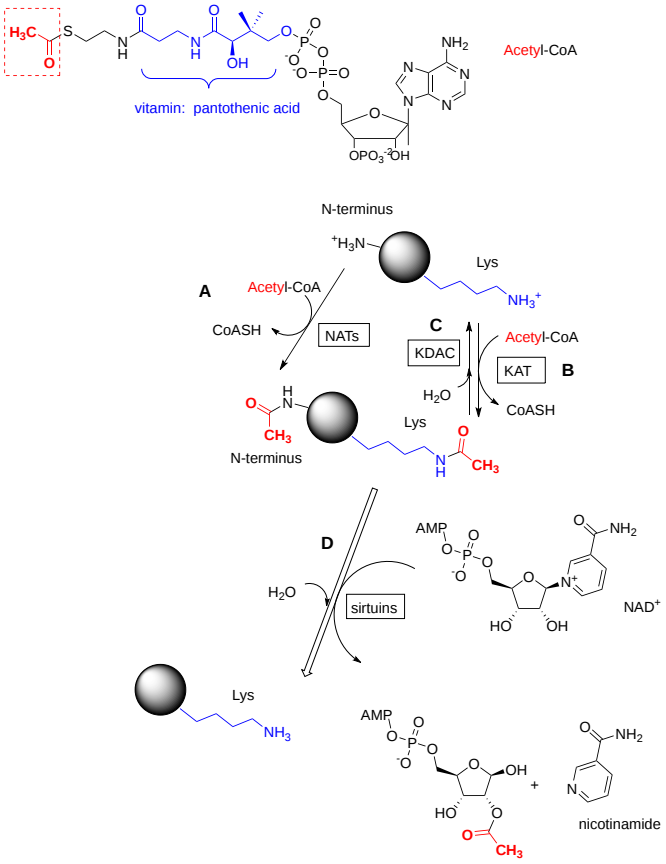


Figure 6.6.3: Protein N-acetylation and deacetylation reactions

The acetyl donor is acetyl-CoA. The N-terminus (reaction A) is acetylated by **N-terminal acetyltransferases (NATs)**. The ε-amino group of lysine (reaction B) is acetylated by **lysine acetyltransferases (KATs).** The latter modification is reversed through two different enzymes. One is a simple hydrolysis (reaction C) catalyzed by lysine deacetylases (KDACs). The other (reaction D) requires NAD^+ and is catalyzed by enzymes called **sirtuins** (silent information regulator). Since sirtuins use NAD^+, one of the main oxidizing agents in the biological world, they link the protein **acetylome** with metabolic status and promote metabolic health.

Histones, positively charged proteins found in eukaryotic cell nuclei, pack and order the DNA into structural units called nucleosomes as shown in Figure 6.6.4.

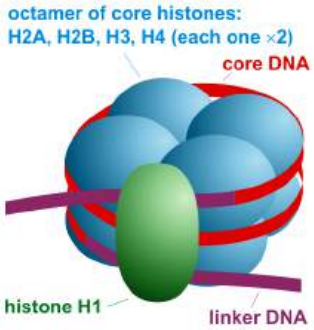


Figure 6.6.4: Nucleosome organization. https://commons.wikimedia.org/wiki/F...ganization.png. Creative Commons Attribution-Share Alike 3.0 Unported license.

The histones have multiple lysine and arginine residues that interact with the negatively charged phosphate groups of the DNA backbone. The nucleosome core is formed of two H2A-H2B dimers and an H3-H4 tetramer, forming two nearly symmetrical halves by tertiary structure (C2 symmetry; one macromolecule is the mirror image of the other). The 4 'core' histones (H2A, H2B, H3 and H4) are relatively similar in structure and highly conserved through evolution, all featuring a 'helix-turn-helix' motif (a DNA-binding protein motif that recognizes specific DNA sequence). They also share the feature of long 'tails' on one end of the amino acid structure, which is the location of post-translational modification, specifically N-acetylation.

Histone acetylation typically results in transcriptional activation; deacetylation typically results in transcriptional suppression. Acetylation occurs via ***histone acetyltransferases (HATs)*** and is reversible via the action of ***histone deacetylases (HDACs)***. One group of histone deacetylases are the **sirtuins** (silent information regulator), which maintain gene silencing via hypoacetylation. Sirtuins have been reported to aid in maintaining genomic stability. Figure 6.6.5 shows the effect of acetylation on histone:DNA packing.

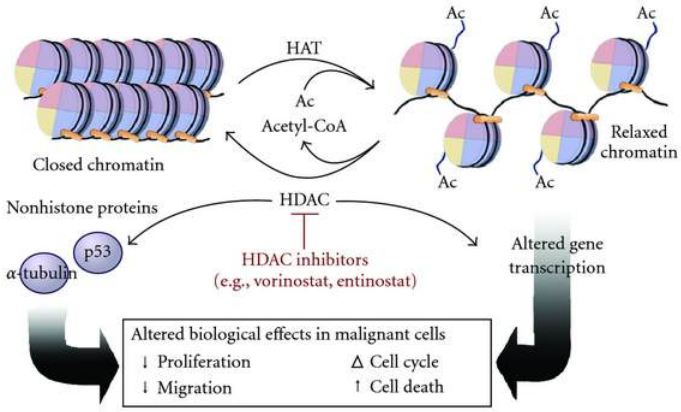


Figure 6.6.5: **Histone octamers form the backbone for chromosomal structure.** Acetylation and methylation affect histone-DNA interactions. DNA methylation increases histone-DNA affinity and blocks transcriptional activation. Acetylation of the histone tails disrupts histone-DNA interactions and facilitates gene expression. https://commons.wikimedia.org/wiki/F...cetylation.jpg. Creative Commons Attribution 3.0 Unported

Although first described in histones, acetylation is also observed in cytoplasmic proteins. Acetylated proteins can also be modulated by cross-talk with other posttranslational modifications, including phosphorylation, ubiquitination, and methylation. Therefore, acetylation may contribute to cell biology beyond transcriptional regulation.

Protein Glycosylation

Protein glycosylation involves the addition of a diverse set of sugar moieties to the protein core. Glycosylation has significant implications for protein folding, conformation, distribution, stability, and activity. Glycosylated proteins can have additions of simple monosaccharides. For example, many nuclear transcription factors are modified in this way. Alternatively, some proteins are modified with highly complex branched polysaccharides, such as those seen on cell surface protein receptors.

More than half of all mammalian proteins are believed to be glycosylated. While proteins exhibit improved stability and trafficking after glycosylation *in vivo*, glycan structures can alter protein functions or activities. Glycan structures often are modified by glycan-processing enzymes working within a cell at any given time. However, the structures are sometimes protein-specific, depending on protein trafficking properties and interactions with other cellular factors.

There are three types of protein glycosylation in higher eukaryotes: N-linked, O-linked, and C-linked. These types reflect their glycosidic linkages to amino acid side chains. In N-linked glycosylation, β-N-acetylglucosamine (GlcNAc) is attached through an amide linkage to the side chain of Asn within a consensus sequence of AsnXaaSer/Thr (Figure 8.8). N-linked glycans have multiple functions. While they act as ligands for glycan-binding proteins in cell-cell communication, they also can regulate glycoprotein aggregation in the plasma membrane and affect the half-life of antibodies, cytokines, and hormones in serum.

O-linked glycosylation in higher eukaryotes occurs through several different mechanisms. The most abundant type of O-linked glycosylation is the ***mucin-type***, which involve the attachment of an α-N-acetylgalactosamine (GalNAc) to the hydroxyl group of Ser/Thr side chains. ***Mucins*** are a family of high molecular weight, heavily glycosylated proteins (glycoconjugates) produced by epithelial tissues in most animals. A key characteristic of mucin proteins is their ability to form gels; therefore they are a key component in most gel-like secretions, serving functions from lubrication to cell signaling to forming chemical barriers. Aberrant expression of mucin-type O-linked glycans occurs in cancer cells and may provide targets for anticancer vaccines.

O-linked glycosylation occurring with the addition of α-O-mannose is the only form of O-linked glycosylation in yeast, but also occurs in the brains of higher eukaryotes. Higher eukaryotes also have an α-O-fucose modification of Ser/Thr residues. This type of glycosylation modulates signaling pathways during eukaryotic development. Another modification, β-O-galactosylation, may contribute to rheumatoid arthritis.

Finally, C-linked glycosylation involves the addition of α-mannose (Man) to the 2-position of the indole side chain of tryptophan residues. While first identified on ribonuclease 2, it also occurs on other proteins, including some of the mucin proteins, thrombospondin, and the Ebola virus soluble glycoprotein. Figure 6.6.6 shows examples of glycan adducts.

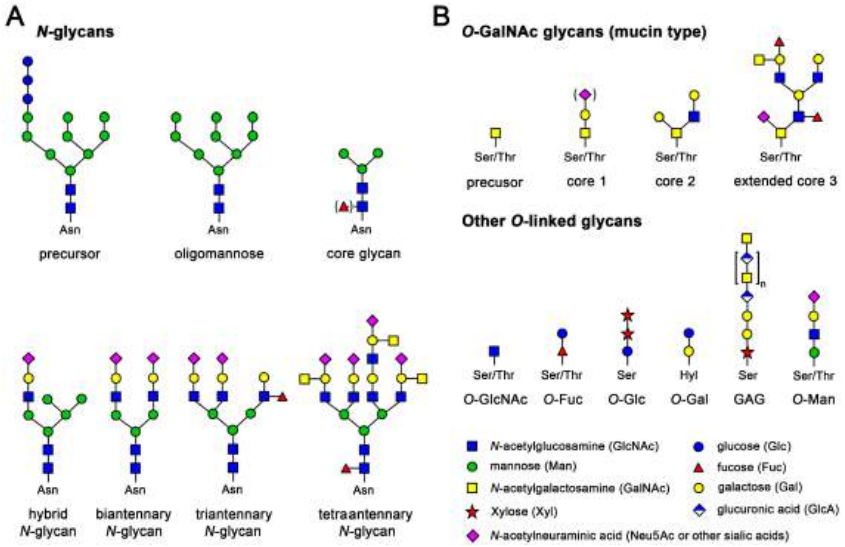


Figure 6.6.6: Glycan adducts to proteins Goettig, P. (2016) Int. J. Mol. Sci. 17(12), 1969

We will discuss glycoprotein in great detail in the next chapter.

Protein Ubiquitination and Sumoylation

Ubiquitination (or ubiquinitylation) is the addition of an 8 kDa polypeptide, called ubiquitin, to lysine residues of target proteins via the C-terminal glycine residues of ubiquitin. The addition of one ubiquitin can lead to the addition of more to form ubiquitin chains on the target protein. The activity of the target protein can be modified by the addition of one ubiquitin. This can regulate the activity of a protein, its location within the cell, and its interaction with other proteins. The addition of a single ubiquitin to lysine

120 (K120) in humans and to K123 in yeast alters gene transcription through remodeling of chromatin. However, if multiple ubiquitins are added to the target protein forming a polyubiquitinated protein, the protein is tagged for degradation by the 26S proteasome. Figure 6.6.7 shows different patterns of ubiquitylation for proteins.

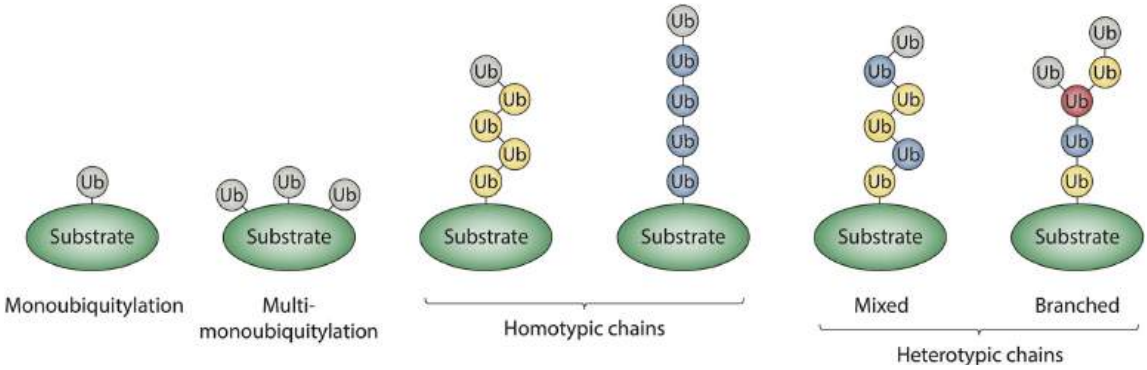


Figure 6.6.7: Classification of ubiquitin modifications. Ubiquitins modified on one acceptor site are colored in blue or yellow; the branch point ubiquitin is colored in red; unmodified or "terminal" ubiquitins are colored gray. French et al. Cell Discovery (2021) 7:6 Cell Discovery. https://doi.org/10.1038/s41421-020-00237-y. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

An elaborate system of enzymes is involved in attaching a ubiquitin through its C-terminal Gly 76, to an internal lysine (or cysteine) in a target protein, and then adding more ubiquitins to form the structures shown above. Figure 6.6.8 shows the first step in the process which requires three enzymes, E1, E2, and E3.

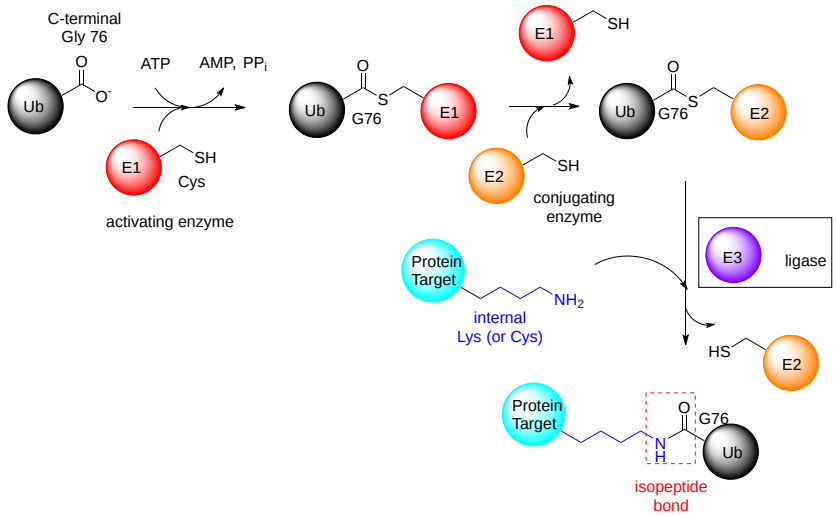


Figure 6.6.8: shows the first step in the process which requires three enzymes, E1 (Ub-activating), E2 (Ub-conjugating) and E3 (Ub-ligating) enzymes

Once the first (or proximal) ubiquitin is added, more can be added to form chains. The steps involved in adding a second ubiquitin to the proximal one are shown in Figure 6.6.9.

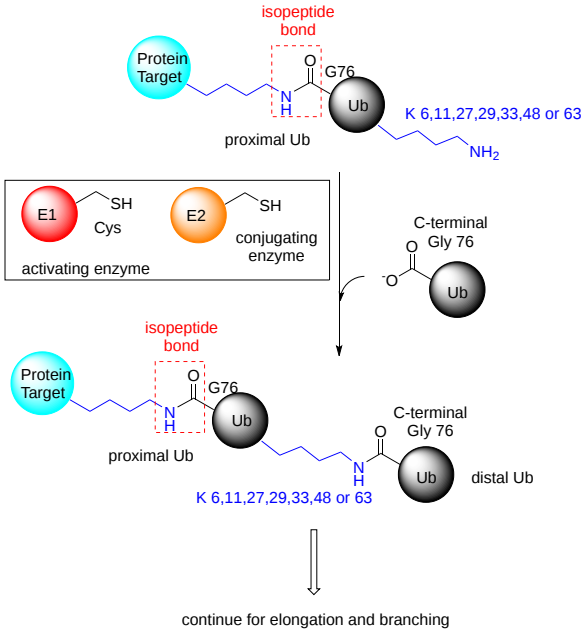


Figure 6.6.19: Addition of additional ubiquitin to form chains

Similarly, protein sumoylation is a reversible posttranslational modification whereby a **s**mall **u**biquitin-like **mo**difier (**SUMO**) protein is covalently attached to a target. Like ubiquitin, SUMO is linked to a lysine side chain of target proteins and is removed by SUMO-specific isopeptidases. Sumoylation controls many aspects of nuclear function. Protein sumoylation is involved in many extranuclear neuronal processes and potentially in a wide range of neuropathological conditions.

Figure 6.6.10 shows ubiquitin (1UBQ, magenta) and SUMO-1 (1A5R, cyan) superimposed. Note the similarity in their structures even though they share only 18% sequence identity. Key amino acids are shown in spacefill.

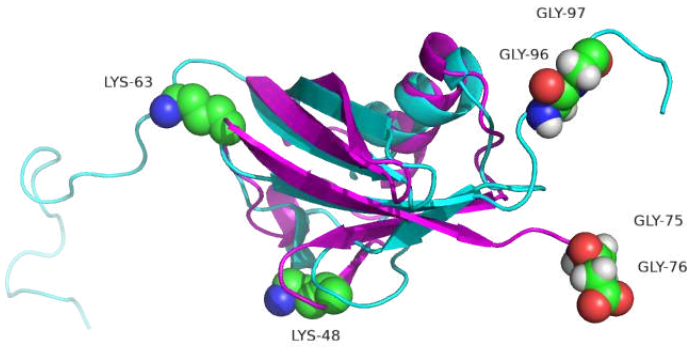


Figure 6.6.10: Ubiquitin (1UBQ) and SUMO-1 (1A5R) superimposed

SUMO-1 doesn't target proteins for degradation. One of its roles is in the transport of proteins between the nucleus and cytoplasm, and it is involved in the modulation of protein-protein interactions. Both ubiquitin and SUMO-1 have C-terminal glycine involved in isopeptide bond formation between ubiquitin or SUMO and the target protein. A major difference between ubiquitin and SUMO-1 is along disordered N-terminal end in SUMO. Also lysines 48 and 63, found in ubiquitin, are replaced by other amino acids. This may account for the observation that SUMO-1 does not form polymers.

The interplay between ubiquitylation and SUMOlaytion is complex. If the ubiquitylation of proteins in inhibited, SUMOylation of newly synthesized proteins increases. These can end up in phase-separated condensates called Promyelocytic Leukemia Nuclear Bodies (PML nuclear bodies), 0.1–1 μm condensates without a membrane found in most mammalian cell nuclei. These are somehow involved in gene expression in the nucleus.

Protein Oxidation

The reaction of proteins with a variety of free radicals and reactive oxygen species (ROS) leads to oxidative protein modifications such as the formation of protein hydroperoxides, hydroxylation of aromatic groups and aliphatic amino acid side chains, oxidation

of sulfhydryl groups, oxidation of methionine residues, conversion of some amino acid residues into carbonyl groups, cleavage of the polypeptide chain, and formation of cross-linking bonds. Aromatic and sulfur-containing amino acid residues are particularly susceptible to oxidative modification.

Unless repaired or removed from cells, oxidized proteins are often toxic and can impair cellular viability, since oxidatively modified proteins can form large aggregates. Oxidatively damaged proteins undergo selective proteolysis, primarily by the 26S proteasome in a ubiquitin- and ATP-independent way. Upon extensive protein oxidation, these aggregates can become progressively resistant to proteolytic digestion and actually bind the 20S proteasome and irreversibly inhibit its activity.

Protein carbonylation is defined as an irreversible posttranslational modification (PTM) whereby a reactive carbonyl moiety, such as an aldehyde, ketone, or lactam, is introduced into a protein. The first identified source of protein-bound carbonyls was metal-catalyzed oxidation (MCO). MCO results from the Fenton reaction when transition metal ions are reduced in the presence of hydrogen peroxide, generating highly reactive hydroxyl radicals in the process. These hydroxyl radicals can oxidize amino acid side chains or cleave the protein backbone, leading to numerous modifications including reactive carbonyls. For example, oxidation of proline and arginine results in the production of glutamic semialdehyde, while lysine is oxidized to aminoadipic semialdehyde and threonine to 2-amino-3-ketobutyric acid. Direct oxidation of other amino acid residues can also lead to protein-bound carbonyls. Tryptophan oxidation by ROS produces at least seven oxidation products. Among them are kynurenine and N-formyl kynurenine, as well as their hydroxylated analogs, which contain aldehyde or keto groups formed by oxidative cleavage of the indole ring.

Another important source of protein-bound carbonyls is reactive lipid peroxidation products, which are produced during oxidation of polyunsaturated fatty acids. Protein carbonylation can also occur via glycoxidation. Reactive α-carbonyls formed during glycoxidation, such as glyoxal, methylglyoxal, and 3-deoxyglucosone, can then modify the basic residues Lys and Arg to generate, for example, pyrralines and imidazolones. Glycation (i.e., the reaction of reducing sugars such as glucose or fructose with the side chains of lysine and arginine residues) forms Amadori and/or Hynes products. These glycated residues can be further changed by ROS into advanced glycation end products (AGE) carrying carbonylated moieties.

Some oxidative modifications are made enzymatically and have key regulatory or structural functions within the modified proteins. For example, proline can be converted to hydroxyproline and lysine to hydroxylysine, as shown in Figure 6.6.11. 4-Hydroxyproline makes up about 13.5% of the residues within the mammalian collagen family of proteins. Recall that collagen is the main protein of the connective tissue and represents about one-fourth of the total protein content in many animals. Hydroxyproline contributes to the stability of the triple helix and also aids in cross-linking collagen fibers to form larger macromolecular complexes.

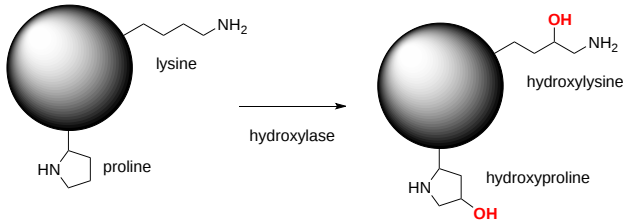


Figure 6.6.11: Hydroxlation of lysine and proline. Xu, Y., et. al. (2014) Int J Mol Sci 15(5):7594-7610.

Protein Methylation

Alkyl substituents are also a common posttranslational modification. The introduction of such alkyl groups results in the alteration of the hydrophobicity of the modified protein. The most common type of protein alkylation is protein methylation, which is mediated by methyltransferase enzymes. One-carbon methyl groups are added to nitrogen or oxygen (N- and O-methylation, resp.) on amino acid side chains, increasing protein hydrophobicity or neutralizing a negative charge when bound to carboxylic acids. While N-methylation is typically irreversible, O-methylation is potentially reversible. Methylation occurs so often that its primary methyl donor, S-adenosyl methionine (SAM), is one of the most-used enzymatic substrates after ATP. The methylation of a histone protein is shown in Figure 6.6.12

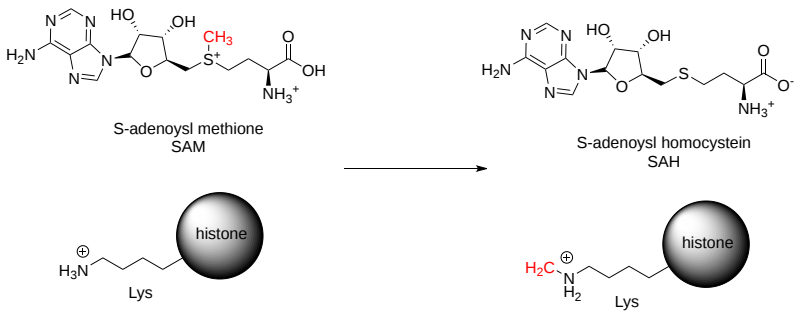


Figure 6.6.12: Methylation of a histone protein by SAM

This modification (not unlike phosphorylation reactions) plays a role in regulation of protein-protein interactions. For instance, the arginine methylation of proteins can either inhibit or promote protein-protein interactions depending on the type of methylation. Protein methylation is also a common modification found in ***histone proteins.*** The transfer of methyl groups from S-adenosyl methionine to histones is catalyzed by enzymes known as ***histone methyltransferases***. The N-terminal tails of histones H3 and H4 receive methyl groups on specific lysines. Methylation then determines if gene transcription is activated or repressed, thus leading to different biological outcomes.

Histone methylation was traditionally thought to be irreversible. However, ***histone demethylases*** demonstrate the reversibility of this PTM. The simultaneous removal of one histone methylation and the addition of another can enable transcriptional fine-tuning.

In addition to methylation, histone acetylation, deacetylation, and mono-ubiquitination are also essential parts of gene regulation, as shown in Figure \(\PageIndex{13}\) below. Acetylation removes the positive charge on the histones, thereby decreasing the interaction of the N termini of histones with the negatively charged phosphate groups of DNA. As a consequence, condensed chromatin is transformed into a more relaxed structure associated with greater levels of gene transcription.

Nonhistone proteins also exhibit methylation as a common PTM which regulates signal transduction pathways. Furthermore, methylation works in concert with other types of PTMs, to exert influence on not only chromatin remodeling but also gene transcription, protein synthesis, and DNA repair.

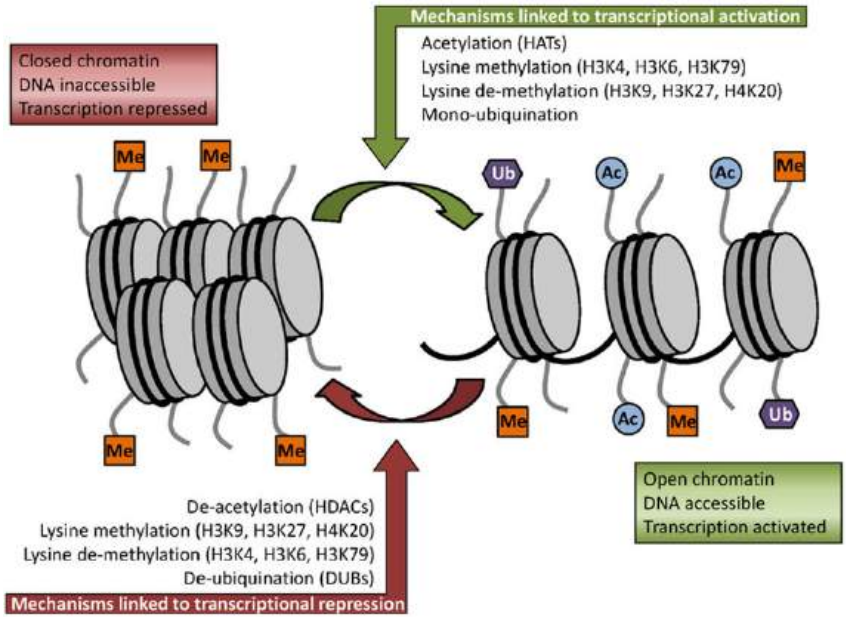


Figure \(\PageIndex{13}\): Some of the key histone modifications influencing gene expression (Me: methylation, Ub: ubiquination, Ac: acetylation). Figure from: Ullah, M.F. (2015) Medicines 2:141-156.

6.6.2: Allosteric Regulation

Allosteric regulation fine-tunes most biological processes, including signal transduction, enzyme activity, metabolism and transport. ***Allostery***, an intrinsic property of a protein, is referred to as the regulation of activity at one site (also known as an ***orthosteric site***) in a protein by a topographically and spatially distant site; the latter is designated as an ***allosteric site. Allosteric***

regulation occurs through the binding of a modulator (e.g., small molecule or protein) at an allosteric site to engender a conformational change that affects function at the orthosteric site. This effect may cause the re-distribution of the conformational ensemble by either stabilizing an active conformation (***allosteric activation***) or destabilizing an inactive conformation (***allosteric inhibition***) in response to allosteric perturbations (Figure 8.13). Traditionally, the repertoire of allostery was primarily confined to determining the allosteric effects or mechanisms in individual multi-subunit or monomer proteins by conformational transitions. Recently, increasing evidence has indicated that allosteric signals can propagate across several or numerous proteins to sculpt allosteric networks.

Allosteric regulation is particularly important in the cell's ability to adjust enzyme activity based on the surrounding environmental conditions. Feedback control loops, such as feedback inhibition from downstream products or feedforward from upstream substrates are common allosteric regulatory mechanisms found in nature. Another example of allostery includes oxygen binding to one of the subunits of hemoglobin that prompts cooperative binding to other subunits.

We mentioned before that PTMs can lead to conformational changes in proteins. Other methods to induce such allosteric changes include the binding of small molecules, the binding of proteins, changes in the redox environment of a protein, which affects disulfide bonds, and general changes in protein flexibility and dynamics. Let's look at some examples of allostery brought about by these changes. Many of these examples are presented by Laskowski et al (https://doi.org/10.1016/j.febslet.2009.03.019).

Allosterism by small molecules: opening and closing active sites

Small molecule binding at an allosteric site can lead to small and large conformation changes in a distal active site. An often observed motion is a partial or full hinge-clamping conformational change that brings critical catalytic groups into a more organized and effective active site, which excludes water as a competing hydrolytic substrate (for example). One example of a small hinge-clamping change is seen in phosphoglycerate dehydrogenase (PGDH, which catalyzes the first step (oxidation by NAD^+) in the synthesis of serine. It converts 3-phosphoglycerate into 3-phosphohydropyruvate. The end product of the pathway, serine, binds to an allosteric site on PGDH and inhibits it, a classic example of feedback inhibition of the first reaction of a pathway by the end product. PGDH has a regulatory binding domain (RBD) that binds serine, a substrate binding domain (SBD) and an NAD^+/NADH binding domain (NBD) which is where the allosteric inhibitor binds, plus binding domains for substrate (SBD) and the NAD nucleotide (NBD). Figure 6.6.14 shows the small hinge-bending conformational change on bind conversion of monomeric *apo*-PGDH (1psd) to *holo*-PGDH with serine (not shown) bound in the RBD allosteric domain. The RBD-SBD domains, relative to the NBD domain, undergoes a 15° rotation.

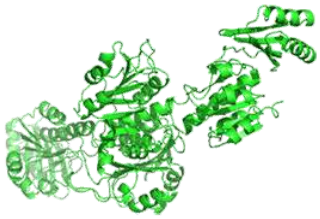


Figure 6.6.14: Small hinge-bending conformational changes on bind conversion of monomeric *apo*-PGDH (1psd) to *holo*-PGDH on binding the allosteric end-product inhibitor serine (not shown)

The functional enzyme is a tetramer (only the monomer is shown in the above figure). The rotation in the full tetramer aligns catalytic residues and closes off the active site.

Conformational changes are ubiquitous on the binding of ligands to protein. These changes also occur on substrate binding at the active site. It is important to recognize substrate-induced changes in the active site are **not** allosteric changes since the conformational changes, no matter how big, are caused by binding at the active site, not at an allosteric site. A classic example of a huge conformational change on binding of substrate is elicited by the binding of glucose or glucose analogs to the active site of hexokinase, the first enzyme in the glycolytic pathway. Figure 6.6.15 shows the large conformational change on binding a glucose analog to hexokinase (1hkg, no glucose; 2yhx, with glucose analog). The background is shown in black to remind readers that this is **NOT** an example of allosterism, although the changes facilitate catalysis by excluding water, a competing substrate) from the active site after glucose and ATP (not shown) bind. On binding the glucose analog in the active, the active site becomes sequestered ("jaw" clamping down on binding).

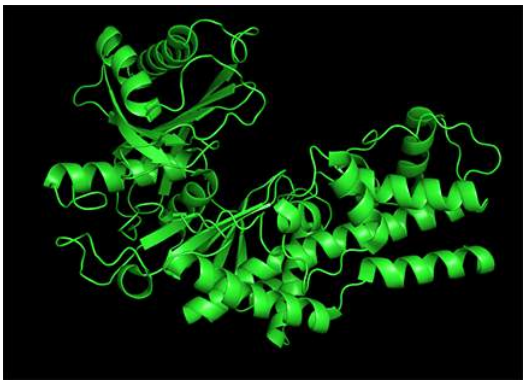


Figure 6.6.15: Large, substrate-induced (non-allosteric) conformational change on binding a glucose analog to (1hkg, no glucose to 2yhx, with a glucose analog.

Allosterism by small molecules: Subtile electrostatic effects

Often no large conformation change is evident in the protein. In those cases, subtle rearrangements of key residues in the active site or near it (which might promote access) may be the result of allosteric effector binding. A simple change in the electrostatic environment might account for the effect. Such appears to be the case with chorismate mutase, an enzyme in the bacterial, fungal, and plant pathways for the synthesis of aromatic amino acids tyrosine and phenylalanine. The enzyme catalyzes the conversion of chorismate to prephenate, which proceeds on to the aromatic amino acids. An offshoot pathway takes chorismate to the last aromatic amino acid tryptophan. The enzyme chorismate mutase then is key in the metabolic decision pathway to make tyrosine/phenylalanine or tryptophan. The enzyme is activated by tryptophan and inhibited by tyrosine. The aligned sequences of two crystal structures of the enzyme with the bound allosteric activator tryptophan (Trp 502 in pdb structure 1csm showing a magenta glutamic acid 23) and with bound inhibitor (Tyr 300 in pdb structure 2csmA showing cyan Glu 23) are shown in Figure 6.6.16.

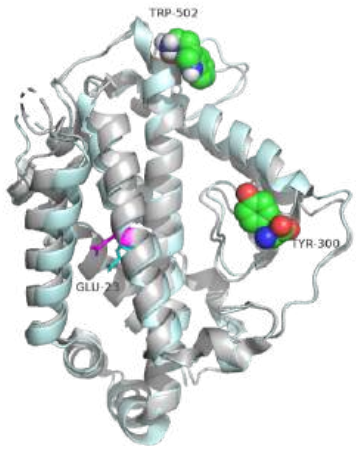


Figure 6.6.16: Aligned structures of chorismate mutase with bound allosteric activator tryptophan (Trp 502 , pdbID 1csm , magenta glutamic acid 23) and with bound inhibitor (Tyr 300 pdbID 2csmA, cyan Glu 23)

There are no significant changes in the overall structure of the protein or in its active site. However, the alignment of glutamic acid 23 is different, and this may be the basis for the observed allosteric effects. When active, glutamic acid 23 is buried in the active site pocket, but in the inhibited site, it swings into the binding site. Since the chorismate is a charged dicarboxylic acid, Glu 23 probably repels its binding, inhibiting the enzyme's activity.

Allosterism by phosphorylation

We'll focus on one example of a huge conformational change on phosphorylation of a serine side chain in glycogen phosphorylase (GP), an enzyme that catalyzes the phosphorolysis of an acetal link between a terminal glucose on a glucose polymer, glycogen, and the next glucose in the chain. This reaction is not a hydrolysis, in which water acts as a nucleophile to cleave acetal bonds. This is a key reaction in metabolism since it cleaves a major energy storage molecule, glycogen.

The enzyme is a multimer that exists in two major states, a T-state and R-state glycogen phosphorylase (GP). We've explored T and R states and allosterism in their interconversion when we discussed the binding of oxygen to hemoglobin in Chapter 5.3. Crystal

structures of the T and R-state bound to allosteric activators IMP and AMP are known. High AMP concentrations imply low ATP concentrations and the need to mobilize glycogen reserves. Conversion to the R state is also promoted by the phosphorylation of Seriner14, which activates the enzyme. The allosteric activators AMP and IMP bind to a disordered loop (313-326) and change it to a form promoted by nucleotide binding. which is disordered in the free structure, and adopts a conformation dictated mainly by the type of nucleotide that binds at this site. Figure 6.6.17 shows a monomer of the European rabbit glycogen phosphorylase in the R (active) state bound to AMP (3e3n).

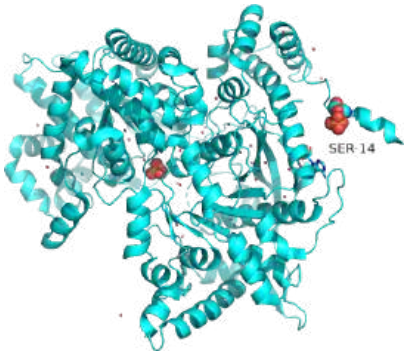


Figure 6.6.17: Monomer of the European rabbit glycogen phosphorylase in the R (active) state bound to AMP (3e3n).

Only one chain of the active homodimer is shown. The helix containing the phospho-Ser 14 swings away from the rest of the protein.

Figure 6.6.18 shows the difference between the T (inactive) state of rabbit muscle glycogen phosphorylase (7P7D) and the R (active) state (3e3n, neither phospho-Ser 14 or AMP shown). Note the huge conformational state elicited on phosphorylation of Ser 13.

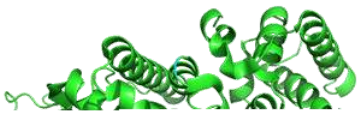


Figure 6.6.18: Conformational changes going from T (inactive) state of rabbit muscle glycogen phosphorylase (7P7D) to the R (active) state (3e3n) on phosphorylation of Ser 13.

Allosterism by small molecules: control of quaternary structure

The binding of small molecules at allosteric sites may promote or inhibit the formation of the correct functional multimeric structure of a protein. One example is ATP phosphoribosyltransferase (APRT), which catalyzes the first step in the synthesis of histidine. As we saw with phosphoglycerate dehydrogenase (PGDH), APRT has three domains. I and II comprise the active site that is located between them, and a regulatory domain, III (the functional protein is a dimer). As with PGDH, APRT is allosterically inhibited by the end product of the pathway, histidine. On histidine binding to the inactive form (1nh7), a large conformational change results (1nh8), as shown in Figure 6.6.19.

Figure 6.6.19: Conformational changes on pathway end product histidine binding to the inactive form (1nh7) form of ATP phosphoribosyltransferase (APRT) inactivating the enzyme (1nh8)

This change drives the formation of a hexamer - $(dimer)_3$, which closes off the active site and inactivates the enzyme. The enzyme ribonuclease reductase behaves similarly.

Allosteric by protein binding

Another protein can bind to an enzyme and activate it by promoting an allosteric change. An example is the binding of a regulatory protein CyclinA to cyclin-dependent kinase 2 (CDK2). On binding of cyclin to one side of CDK2's catalytic cleft, a large conformational change occurs, which activates the enzyme by altering active site geometry and making the active site more accessible. CDK enzymes are involved in cell-cycle progression. For more control of the cell cycle, both the binding of cyclin and the phosphorylation of CDK2 are required for activity. On binding cyclin, the active site is more available for substrate (ATP) binding. Next, the amino acid that needs to be phosphorylated for activation, Thr 16, is made accessible. Figure 6.6.20 shows the conformational change in apo-CDK2 (1hc1) on cyclin binding, which allows ATP binding (1fin).

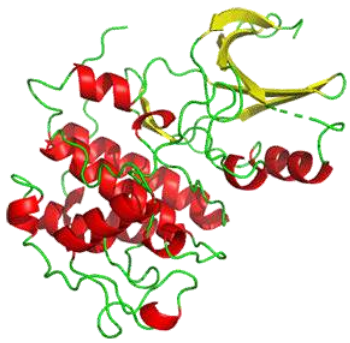


Figure 6.6.20 show the conformational change in apo-CDK2 (1hc1) on cyclin binding which allows ATP binding (1fin).

Cyclin (the regulatory protein) is shown in gray. ATP is shown in spacefill. CDK2 is shown as a colored cartoon.

Allosterism by disulfide bonds

It should make sense that cleaving an intrachain or interchain disulfide, whose presence constraint a specific protein conformation, should significantly alter protein structure and function. One example is the protein botulinum neurotoxin type A. It is neurotoxin and Zn protease that cleaves proteins like synaptobrevins, syntaxin and SNAP-25 in the neurosynapase that are required for neurotransmitter release.

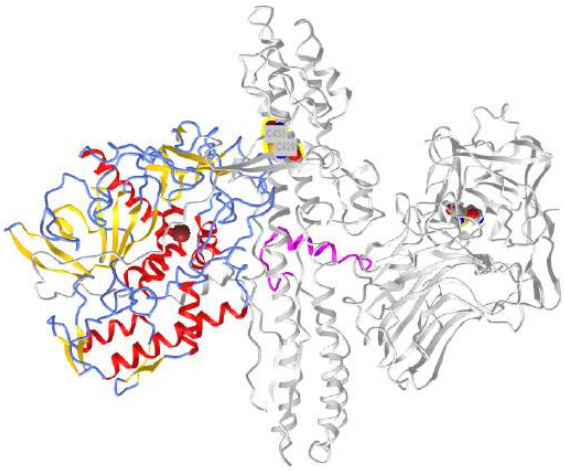
It has a catalytic (Zn peptidase/protease) and a translocation domain in a single protein chain that contains two disulfide bonds when synthesized protein (PDB code 3bta). On cleavage at a select peptide bond, it is split into A and B chains, which remain connected by a single disulfide bond, with the A chain containing the catalytic domain. The B chain contains the translocation domain, which effectively still blocks the active site of the peptidase/protease domain. The disulfides are cleaved when the protein enters an endosome, which contains a more reducing environment. The separate A chain now separates for the inhibiting B domain and expresses protease activity. The active protein then enters the cytoplasm.

Figure 6.6.21 shows the domain structure and disulfide bonds of the unprocessed form. The N-terminal Zn proteinase domain (Peptidase_M27, amino acids 1-409) is followed by an intrachain disulfide between C429 and C 453.

Figure 6.6.21: The domain structure and disulfide bonds of the unprocessed form of botulinum neurotoxin type A

The mature form is proteolyzed between amino acids 447 and 448 (within the sequences connected by the disulfide (429-453). The protein remains inactive until the disulfide bond is cleaved.

Figure 6.6.22 shows an interactive iCn3D model of botulinum neurotoxin type A (3bta).



NCBI iCn3D Figure 6.6.22: Botulinum neurotoxin type A (3bta) (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...CfSsBH94jcPyb8

The N-terminal Zn peptidase/protease domain is shown in color and the other regions of the protein are shown in gray. The two disulfide bonds are shown in spacefill, with CPK color. On cleavage of the disulfide bond between C429 and C453, the Zn protease domain, which is inhibited in the presence of the constraining disulfides, is released and activated.

Allosterism by change in protein flexibility and population shifts

Sometimes crystal structure show little difference between the apo- and holo-forms of the protein containing a separate allosteric binding site. One example is the enzyme dihydrodipicolinate synthase (DHDPS) which catalyzes the first step in the bacterial and plant pathways for lysine synthesis. DHDPS catalyzes the synthesis of 4-hydroxy-2,3,4,5-tetrahydro-2-dipicolinic acid from aspartate-β-semialdehyde and pyruvate. One explanation for allosterism in the protein is that it occurs on time scales too short to be seen in crystal or NMR structures. Rather some transient conformation might cause allosterism. Such changes are detectable in molecular dynamics simulation of the protein in the presence and absence of the allosteric effector molecule.

This possibility suggests that the allosteric effector might change the distribution of possible transient conformational states within a population of such states. If an allosteric effector bound one of the transient states in the population, that state would be "removed" from the ensemble of states, shifting the dynamic equilibrium among the populated states to produce more of the specific conformation that binds the allosteric effector. This sounds remarkably like the T to R state conversion in the MWC model oxygen binding to hemoglobin and the allosterism discussed above with the enzyme glycogen phosphorylase only a much shorter time scale. This model is consistent with a conformational selection model in which the allosteric molecule binds preferentially to a different preexisting conformational state that exists transiently.

6.6.3: Zymogen Activation

A ***zymogen*** also called a ***proenzyme***, is an inactive precursor of an enzyme. A zymogen requires a biochemical change (such as a hydrolysis reaction revealing the active site, or changing the configuration to reveal the active site) for it to become an active enzyme.

Protease enzymes secreted by the pancreas are initially synthesized as zymogens. The pancreas secretes zymogens to help prevent the enzymes from inappropriately digesting proteins in the pancreatic cells in which they are synthesized. Enzymes like Trypsin are synthesized as proenzymes. For trypsin, trypsinogen is an inactive precursor that is translated in the rough endoplasmic reticulum and transported to the Golgi apparatus for sorting. Trypsinogen is always co-synthesized and packed with a pancreatic secretory trypsin inhibitor (PSTI) that inhibits its premature activation. Thus, there are two mechanisms in place to maintain the inactivity of the protease within the pancreas: (1) synthesis of the zymogen or proenzyme form, and (2) co-expression of a trypsin inhibitor protein that will bind and inhibit any prematurely cleaved trypsin until it has reached the small intestine. An animation showing structural differences between bovine trypsinogen (magenta) with just amino acids 10-15 of the presequence (spacefill orange) and mature trypsin (cyan) is shown in Figure 6.6.23

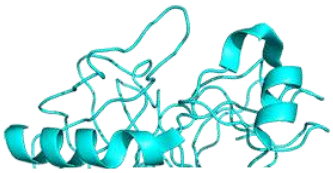


Figure 6.6.23: Animation showing structural differences between bovine trypsinogen (magenta) with just amino acids 10-15 of the presequence (spacefill orange) and mature trypsin (cyan)

During packaging within the Golgi system, the trypsinogen and other digestive enzymes condense into core particles and are packed in zymogen granules. The condensed enzymes are stable and minimal activation happens within the zymogen granules. Once the pancreatic cells receive secretory stimuli, these zymogen granules are released in to the lumen of the pancreatic duct, which carries the digestive enzymes into the duodenum. Once in the duodenum, enteropeptidase activates trypsinogen by removing the 7-10 amino acids trypsinogen activation peptide (TAP) from the N-terminal region (see above). Removal of TAP induces conformational change resulting in active trypsin. TAP is immunologically distinct from the same sequence within trypsinogen, thereby allowing the detection of trypsinogen activation *in situ*.

Once activated, trypsin will cleave and activate other zymogen proteases and lipases present in the duodenum. These include the activation of elastase, chymotrypsin, carboxypeptidase, and lipase as shown in Figure 6.6.24 Zymogens are also found in other cellular processes as well. For example, intracellular proteases known as caspases, are activated in a similar manner during the process of cellular apoptosis or programmed cell death. The process of blood clotting also involves the activation of zymogens.

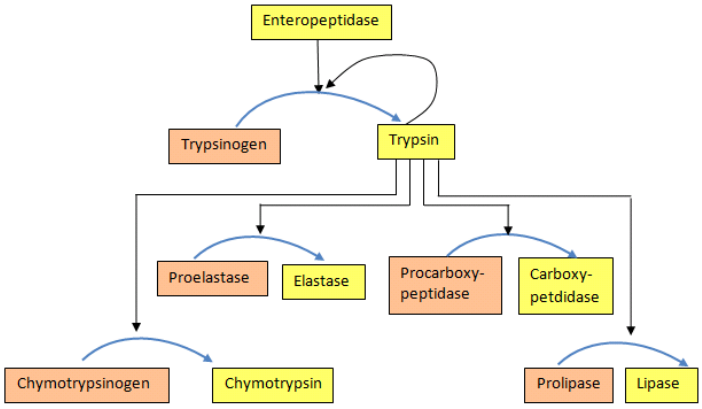


Figure 6.6.24: Activation of proenzymes

6.6.4: Isozymes

Isozymes (also known as **isoenzymes**) are enzymes that differ in amino acid sequence but catalyze the same chemical reaction. These enzymes usually display different kinetic parameters (e.g. different K_M or K_{cat} values), or different regulatory properties. The existence of isozymes permits the fine-tuning of metabolism to meet the particular needs of a given tissue or developmental stage. In many cases, isozymes are coded for by homologous genes that have been duplicated within the genome and then diverged over

time. *Isozymes* should not be confused with **allozymes,** which are allelic variants of the same gene locus that are found within a population. Allozymes represent enzymes from different alleles of the same gene. Isozymes represent enzymes from different genes that process or catalyze the same reaction, or from the same gene that produces, through differential splicing of primary RNA transcripts, sequence of similar but different sequences. We will focus on isozymes within this section. Part of the regulation derived from the production of different isozymes can arise from the differential expression of isozymes.

Isozymes are usually the result of gene duplication. Over evolutionary time, if the function of the new variant remains *identical* to the original, then it is likely that one or the other will be lost as mutations accumulate, resulting in a pseudogene. However, if the mutations do not immediately prevent the enzyme from functioning, but instead modify either its function or its pattern of expression, then the two variants may both be favored by natural selection and become specialized for different functions. For example, they may be expressed at different stages of development or in different tissues. Some isozymes may also arise from convergent evolution and may not share a high degree of sequence homology or common ancestry.

6.6.4.1: The Cyclooxygenase I and II (Cox-1 and Cox-2) Isozymes

As an example of isozymes, we will discuss **cyclooxygenases COX-1** and **COX-2**, which are also called **Prostaglandin Synthases.** They regulate a key step in prostaglandin and thromboxane synthesis and are the targets of **nonsteroidal antiinflammatory drugs (NSAIDs)**, such as aspirin, ibuprofen, naproxen and celebrix (Figure 6.6.25). The prostaglandins (PG) are a group of physiologically active lipid compounds called eicosanoids having diverse hormone-like effects in animals. Prostaglandins have been found in almost every tissue in humans and other animals. They are derived enzymatically from the fatty acid arachidonic acid. Every prostaglandin contains 20 carbon atoms, including a 5-carbon ring. They are a subclass of eicosanoids and of the prostanoid class of fatty acid derivatives.

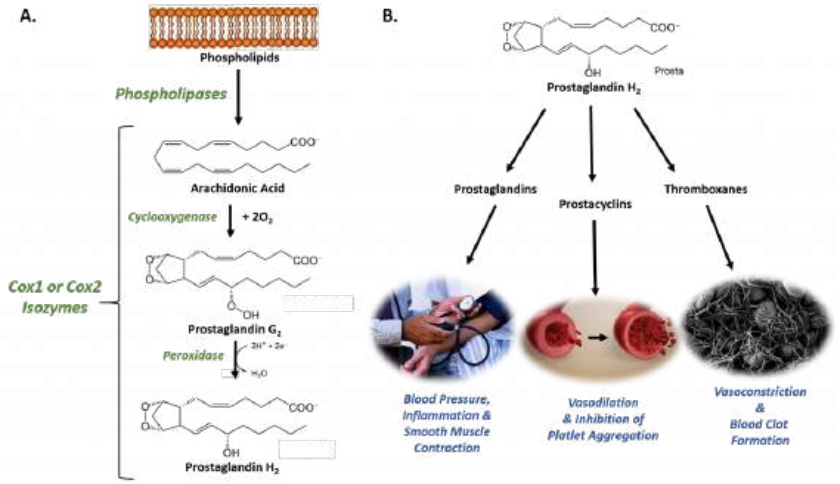


Figure 6.6.25: The Role of COX-1 and COX-2 in Prostaglandin Biosynthesis. (A) COX-1 and COX-2 are bifunctional enzymes that mediate the cyclooxygenase and peroxidase reactions that convert Arachidonic Acid to Prostaglandin H_2. (B) Some of the physiological effects that prostaglandins, prostacyclins, and thromboxanes have on biological processes.Images from: Администрация Волгоградской области , scientific animations, and Fuzis

Cycloxygenases are in a class of enzymes called dioxygenases, that incorporate both atoms of O_2 into a substrate. We will explore the mechanism of cyclooxygenase in more detail in Chapter 13.5: Biological Oxidation and Reduction Reactions.

The structural differences between prostaglandins account for their different biological activities. A given prostaglandin may have different and even opposite effects in different tissues. The ability of the same prostaglandin to stimulate a reaction in one tissue and inhibit the same reaction in another tissue is determined by the type of receptor to which the prostaglandin binds. They act as autocrine or paracrine factors with their target cells present in the immediate vicinity of the site of their secretion. Prostaglandins differ from endocrine hormones in that they are not produced at a specific site but in many places throughout the human body and tend to act locally once secreted. Prostaglandins are implicated in various physiological processes such as gastrointestinal cytoprotection, hemostasis and thrombosis, as well as renal hemodynamics.

Through their role in vasodilation, prostaglandins are also involved in inflammation and can trigger the onset of a fever or the sensation of pain. They are synthesized in the walls of blood vessels. They prevent needless clot formation and regulate the contraction of smooth muscle tissue. The prostacyclins, a special class of prostaglandins, are powerful, locally-acting vasodilators

and inhibit the aggregation of blood platelets. Conversely, thromboxanes (produced by platelet cells) are vasoconstrictors and facilitate platelet aggregation. Their name comes from their role in clot formation or thrombosis.

The cyclooxygenases COX-1 and COX-2 regulate the first two steps in prostaglandin and are bifunctional enzymes containing two active sites. The first active site performs the *bis*-oxygenation and cyclization of arachidonic acid, whereas the second active site mediates a peroxidase (reduction) reaction to form PGH_2. The enzyme is an example of a dioxygenase, which uses O_2 as a substrate. The enzyme contains a heme cofactor. The functional enzyme of both COX-1 and COX-2 are homodimers of 70 kDa subunits, each having an N-terminal epidermal growth factor domain, a membrane binding domain and a C-terminal catalytic domain. The cyclooxygenase active site in on the opposite side of the peroxidase active site in the catalytic domain.

Figure 6.6.26 shows an interactive iCn3D model of arachidonic acid bound to V349I murine COX-2 (6OFY). Just one chain of the dimer is shown.

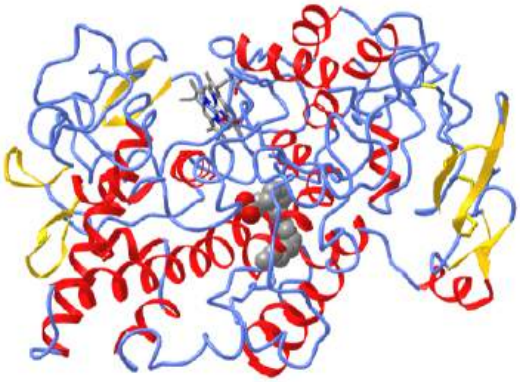


Figure 6.6.26: Arachidonic acid and heme bound to V349I murine COX-2 (6OFY) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?gUtX3uAtwoozBvJUA

The heme is shown in stick while the arachidonic acid is shown in spacefill. Note that arachidonic acid is very kinked and not extended due to its four cis double bonds. Three additional amino acids are highlighted that play key roles in NSAID inhibition of Cox. These include Ser 530 (in model 531), which is covalently acetylated by aspirin, and Arg 120 (in model 121) and Tyr 355 (in model 356). These play key role in binding of NSAIDs and access to the arachidonic acid binding site.

The COX-1 enzyme is widely distributed in many tissues where it is constitutively expressed. Expression of the COX-2 isoform (shown in Figure 6.6.2), on the other hand, is normally undetectable in most tissues (except for the central nervous system, kidneys, and seminal vesicles). COX-2 is induced by various inflammatory and mitogenic stimuli. More recently, a third isoform named COX-3 was identified as a COX-1 splicing variant. This new variant may play a role in processes such as fever and pain. Additionally, a high level of COX-2 expression is found in several forms of cancer. For example, COX-2 overexpression is related to poor prognosis in certain breast cancers and endometrial adenocarcinomas.

COX-2, unlike COX-1, is induced in inflammatory cells and activated by various inflammatory and mitogenic stimuli. Under these conditions, COX-2 activity leads to the production of prostanoid mediators that trigger important inflammatory processes. Although inflammation is initially a necessary process to fight infection, if it is maintained or remains uncontrolled, it can provoke chronic pathologies and tissue damage. This is why the inhibition of COX proteins are key targets for anti-inflammatory and pain-management. Their actions are illustrated in Figure 6.6.27.

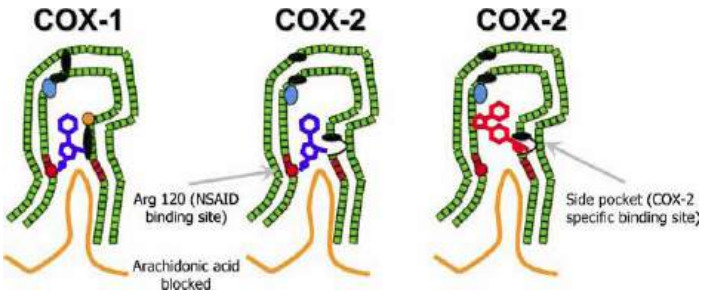


Figure 6.6.27: **NSAID Inhibition of COX-1 and COX-2 Enzymes.** The schematic representation of the COX-1 (large green figure) active site being inhibited by a nonselective NSAID (central blue figure). The entrance channel to COX-1 is blocked by the NSAID. Binding and transformation of arachidonic acid (bottom yellow figure) within COX-1 is prevented. *Middle Lower Panel* shows the inhibition of COX-2 by a nonselective NSAID (central blue figure). Nonselective binding uses an amino acid residue, Arg120, that is conserved in both enzymes. The *right panel* shows the inhibition of COX-2 by COX-2 selective NSAIDs (central red figure). The COX-2 side pocket allows specific binding of the COX-2 selective NSAID's. The entrance channel to COX-2 is blocked. The bulkier COX-2-selective NSAID will not fit into the narrower COX-1 entrance channel, allowing COX-1 to remain active. Upper Figure by Saiz, M., Gonzalez, R., & Garcia, E., (2019) Protopedia and Lower Figure by Meek, I.L., et al. (2010) Pharmaceuticals 3(7):2146-2162.

6.6.4.2: Nonsteroidal Antiinflammatory Drugs (NSAIDs)

Clinically available NSAIDs can be separated into 3 different classes based upon their mechanism of action:

- **ASPIRIN:** - Acts to irreversibly inhibit COX 1 & COX-2 by covalent acetylation of Ser 530 in the active site. Most notably, low doses of aspirin can suppress platelet COX-1 activity by 95% or more, an effect that is permanent for the lifetime of the platelet, since platelets lack DNA and cannot synthesize new enzymes. Due to aspirin's antithrombotic properties at low doses, this treatment has been found to have cardioprotective effects and is often prescribed for patients at high risk of myocardial infarction. All other NSAIDs interact with COX isoforms reversibly and produce variable COX inhibition (ranging from 50% to 95%) in a time-dependent fashion based on how quickly they are metabolized in the body.
- **NON-SELECTIVE COX INHIBITORS:** Different non-selective NSAIDs have varying inhibitory effects against COX-1 & COX-2 (Figure 8.3). The two most commonly used over-the-counter drugs in this group (ibuprofen & naproxen) produce reversible platelet inhibition ranging from 50 to 95% in a reversible, time-dependent manner (more on this below). These NSAIDs may be insufficient to provide cardio-protection throughout a commonly used dosing interval and are not commonly used for this purpose. Ketorolac (Toradol ®), an NSAID most commonly used in a hospital setting to treat moderately severe pain, is classified as a non-selective NSAID, although it is arguably, a very selective for COX-1 inhibitor (Figure 8.3). Inhibition of COX-1 can result in unwanted side effects, such as gastrointestinal discomfort and in severe cases, ulceration.
- **COXIBS:** Selective COX-2 inhibitors were designed and marketed to treat pain and inflammation, while avoiding the GI side effects. However soon after they were introduced into the market, their use led to the first reported incidence of increased cardiovascular events (myocardial infarction and stroke) in 2004. Rofecoxib (Vioxx ®), one of the most selective COX-2 inhibitors was removed from the market because of mounting evidence for significant cardiovascular toxicity (Drazen, 2005). Celecoxib (Celebrex ®) is currently the only FDA approved coxib available in the US, and it has been given a black box warning indicating the potential risk of cardiovascular toxicity. It has a 10-20 fold selectivity for COX-2 over COX-1. Etoricoxib (Arcoxia ®) is a second coxib with ~10^6 fold selectivity for COX-2 over COX-1 that is available outside of the United States.

Figure 6.6.28shows Selectivity and Treatment Efficacy of Nonsteroidal Anti-inflammatory Drugs (NSAIDs).

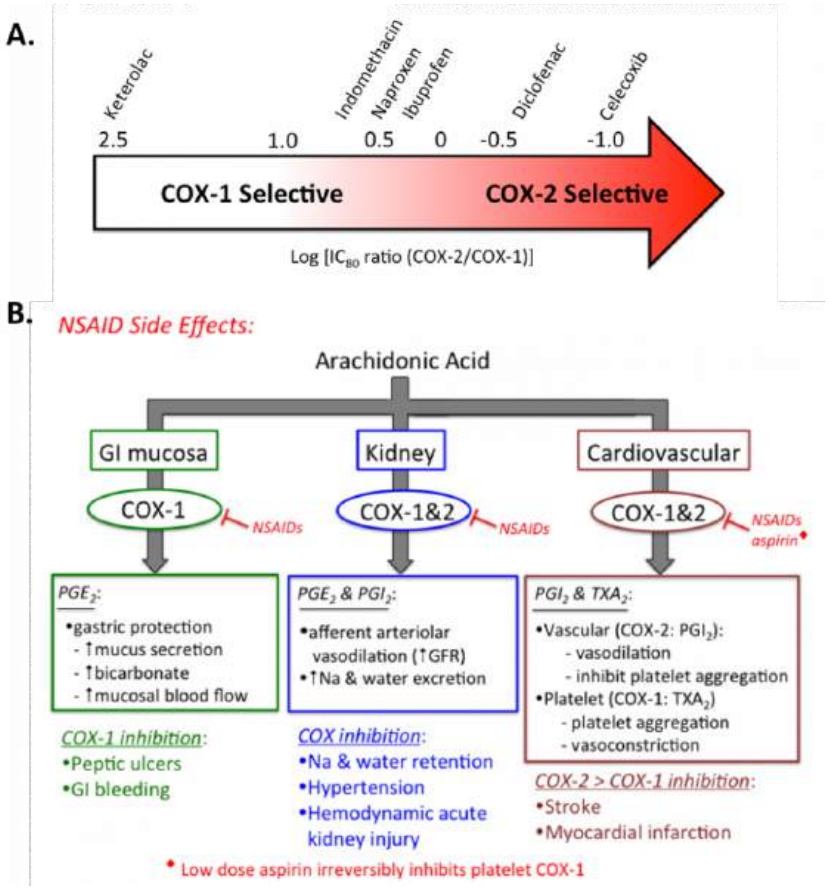


Figure 6.6.28: Selectivity and Treatment Efficacy of Nonsteroidal Anti-inflammatory Drugs (NSAIDs). (A) Relative COX-1 & COX-2 selectivity for commonly used non-aspirin NSAIDs. Celecoxib (Celebrex ®) is the only COX-2 selective NSAID on the market in the US. Adapted from Danelich et al (2015). (B) Major physiological roles for COX-1 & COX-2, and mechanisms underlying drug-induced side effects. PGI_2: prostacyclin, TXA_2: thromboxane.*Figures from: Clarkson, C.W. (2018) TUSOM Pharmwiki*

The structures of ibuprofen (Advil), naproxen, aspirin, celebrix and acetaminophen (Tylenol, which reduces fever and pain but not inflammation so its not a NSAID) are shown in Figure 6.6.29

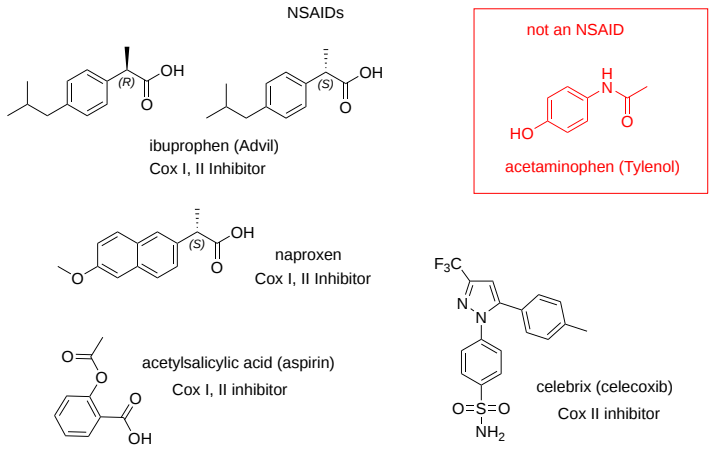


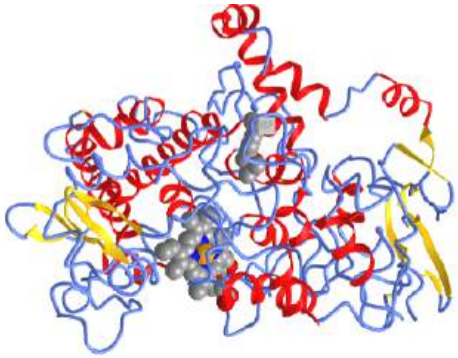
Figure 6.6.29: Structures of ibuprofen, naproxen, celebrix (NSAIDs) and acetaminophen (not an NSAID)

Acetominophen - Tylenol

Although this extremely popular drug relieves pain and reduces fever, it does **not** reduce peripheral inflammation and is hence **not classified as an NSAID**. It does have effects central nervous system affects. It does **not** appear to bind to the active site of COX-1 or COX-2 (no PDB structures available), but can reduce their activities probably either upstream or downstream of the

pathways that these enzymes are part of. It does decrease prostaglandin production in the CNS, which reduces CNS pain and fever. It might act as a reducing substrate in the peroxidase site. Many proteins bind this drug. which is understandable given its relatively simple structure shown in Figure 6.6.5. A metabolite of it, p-aminophenol, can be esterified to arachidonic acid to produce the fatty acid amide, which can work through cannabinoid receptors to reduce pain.

Figure 6.6.30 shows an interactive iCn3D model of the NSAID ibuprofen bound to cyclooxygenase-2 (4PH9). Just one chain of the functional dimer is shown.



NCBI iCn3D Figure 6.6.30: Ibuprofen bound to cyclooxygenase-2 (4PH9). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...syYYjSoeZRGNU8

Ibuprofen is sold as the racemic mixture of R and S enantiomers but it was thought that only the S isomer's inhibition of Cox II leads to its antinflammatory effects while the other isomer had no effects and no side effects. The crystal structures of IBP bound to both Cox I and II are known. At least in the mouse Cox2, the S-isomer binds more tightly than the R-isomer. Arg-120 and Tyr-355 at the entrance of the cyclooxygenase channel are important in the action of IBP in Cox2 (which are labeled in in Figure 6.6.6.

As mentioned above, aspirin inhibits COXs through covalent acetylation of a key serine side chain. The other NSAIDs appear to inhibit prostaglandin H(2) synthase through blockage of the channel for arachidonic acid binding. This should make you think that they act as competitive inhibitors, which is true for ibuprofen and a derivative, methyl flurbiprofen. However, the effect of two others NSAIDs, aclofenac and flurbiprofen, seems to have an added time component not found in simple competitive inhibition model. They are called slow tight-binding inhibitors and a simple model describing it is shown in Figure 6.6.31.

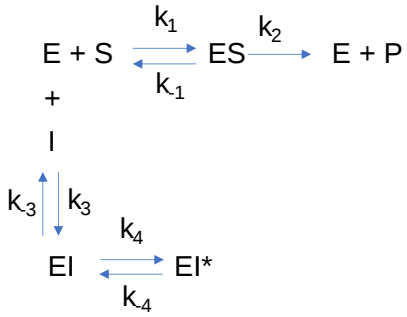


Figure 6.6.31: Model for slow binding inhibition

A time step suggests a conformational change after initial binding. Yet crystal studies of the four inhibitors mentioned above show that the NSAID-occupied active sites are essentially the same, suggesting no major global conformational change. It may be that the rate involves slow hydrogen bonding around two key residues, Arg 120 and Tyr 355.

Slow binding and slow tight-binding inhibition

All the enzyme inhibition models we discussed in Chapter 6.4 are based on reversible rapid equilibrium binding of an inhibitor to E (competitive), ES (uncompetitive), and E and ES (mixed/noncompetitive). Equations can be derived for **slow binding** inhibition and for **slow tight-binding**. In these models, the degree of inhibition at a fixed concentration of the inhibitor changes with time to establish an "equilibrium" among all species. The model above works for both. Both types of inhibitors (slow and slow-tight) inhibit enzymes in a time-depended fashion. In slow binding, it takes longer to establish equilibrium among the

three species, E, EI and EI*. In slow binding, the steps determining the concentration of EI (k_3I and k_{-3}) are assumed to be fast compared to those involved in determining the concentration of EI* (k_4 and k_{-4}). When [I] >> [E], the following dissociation constants for pure competitive ($K_{IS} = K_C$) and for the slow binding inhibition (K_I), which accounts for mass balance, hold.

$$K_{IS} = K_C = \frac{[E][I]}{[EI]} \tag{6.6.1}$$

$$K_I^* = \frac{[E][I]}{[EI] + [EI^*]} \tag{6.6.2}$$

In initial rate Michaelis-Menten kinetics, v_0 is measured as a function of [S]. The initial rate v_0 is determined by measuring [P] vs t in the beginning of the reaction when the substrate is not depleted. Valid v_0 values require that [P] vs t curves are linear. The slope of that line is the initial velocity, v_0. This is true for uninhibited and reversibly inhibited (competitive, uncompetitive, and mixed) reactions. However, with slow binding inhibition, the [P] vs t curves bend with time as the slower formation of EI* occurs. It's a bit like the pre-steady state assumption in enzyme kinetics. At very short times (msec), the P vs t curves bend as a steady state emerges. It's the same with slow binding inhibition. This is illustrated in Figure 6.6.32

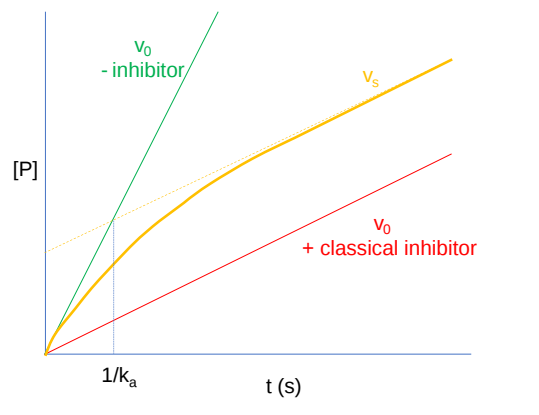


Figure 6.6.32: P vs t curves for uninhibited (green), classical inhibition (red) and slow binding inhibition (gold)

A progress curve equation showing [P] vs t for slow binding in inhibition is shown below without inhibition.

$$[P] = v_s t + \frac{(v_o - v_s)\left(1 - e^{-k_a t}\right)}{k_a} + C \tag{6.6.3}$$

In this equation, k_a is the apparent first-order rate constant to the development of the steady state at a given substrate and inhibitor concentration, v_0 is the very first initial rate and v_s is the steady-state rate.

Slow tight-binding inhibition occurs when the rate constants for net production of EI* are fast compared to the step for the formation of EI. In this case, even small amounts of I will produce EI* as the reaction is pulled to EI*, and more complicated equations must be used to determine product vs time equations.

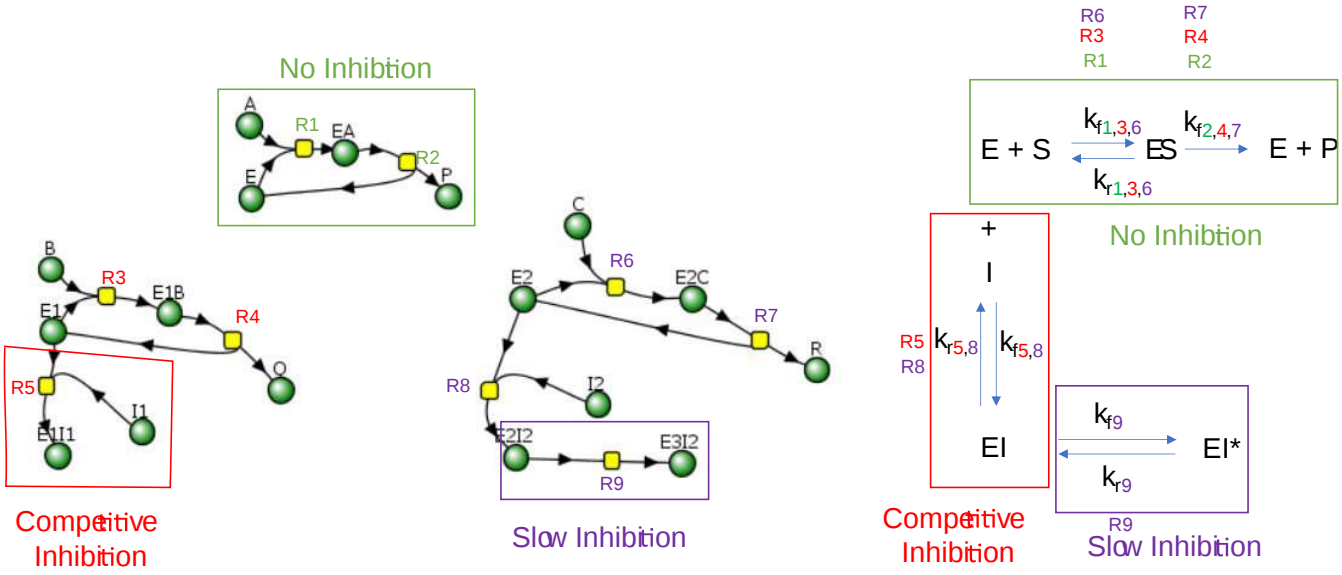
Other chemical models could also account for slow binding inhibition. These include a slow binding rate constant k_4, a slow isomerization of EI after fast binding of I, or a slow binding to a specific, low population conformation of enzyme in a process called conformational selection. Slow-binding inhibitors of enzymes (such as neural acetylcholine esterase) are known.

Here is a Vcell simulation showing progress curves for an uninhibited reactions, one in the presence of a competitive inhibitor and one in the presence of a slow (competitive) inhibitor.

VCell MODEL

Comparison No Inhibition, Competitive Inhibition, and Slow (Competitive) Inhibition

Vcell reaction diagram with all equations based on mass action (not Michael-Menten kinetics)



Concentrations

- A, B and C (substrates), t_0 = 1 uM
- I_1 and I_2 (inhibitors) = 1 uM
- E, E_1 and E_2 (enzymes) at t_0 = 0.1 uM
- P, Q and R (products) at t_0 = 0 uM

Rate Constants

All forward (k_f) and reverse (k_r) rate constants are set initially to 1 except for

- k_{r2} for all reactions = 0 and k_{f4} = 1.

Select Load [model name] below

Load SlowInhibition

Select **Start** to begin the simulation.

Only plots of P, Q and R (products) are initially shown. Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

6.6.4.3: Coxibs and the Thromboxane/Prostacyclin Imbalance Hypothesis

Previous research indicates that in the cardiovascular system, a greater inhibition of COX-2 vs COX-1 (as produced by COX-2 selective “coxibs”) can tip the normal balance between the effects produced by prostacyclin & thromboxane, resulting in an increased likelihood for platelet aggregation and vasoconstriction. These effects can help to explain the higher incidence of myocardial infarction and stroke observed when these drugs have been used clinically. The mechanisms involved are illustrated in Figure 6.6.33.

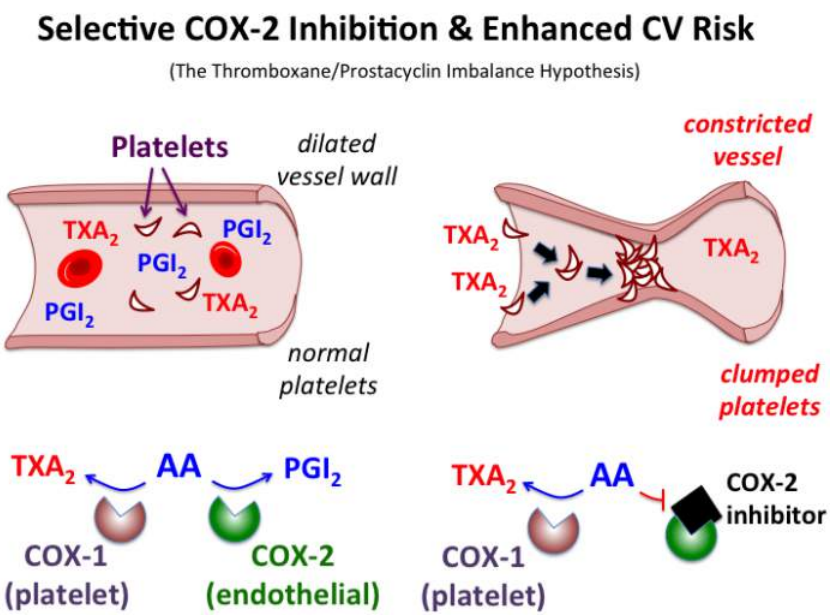


Figure 6.6.33: **COX-2 Inhibitors & Cardiovascular Risk.** *Figure from: Clarkson, C.W. (2018) TUSOM Pharmwiki*

The left graphic illustrates the normal balanced effect between prostacyclin (PGI_2) and Thromboxane (TXA_2). PGI_2 is produced primarily by COX-2 activity in the endothelial cell wall of blood vessels. PGI_2 produces vasodilation, and inhibits platelet activation. In contrast, TXA_2 is produced primarily by COX-1 activity inside platelets, and produces vasoconstriction and enhanced platelet aggregation. When there is a balanced effect of both PGI_2 & TXA_2, normal vascular homeostasis is maintained. However, when the balance is tipped in favor of TXA_2 formation after selective inhibition of COX-2 (right graphic), vasoconstriction and platelet clumping are more likely to occur, potentially causing an increased risk for cardiovascular events such as myocardial infarction and stroke. Overall, the COX-1/COX-2 isozyme example sheds light on the complexity of biological systems and the ability for slight adjustments in gene expression to create varied and tissue-specific responses.

6.6.5: Some References

1. Shen, Q., Wang, G., Li, S., Liu, X., Lu, S., Chen, Z., Song, K., Yan, J., Geng, L, Huang, Z., Huang, W., Chen, G., and Zhang, J. (2016) ASDv3.0: Unraveling allosteric regulation with structural mechanisms and biological networks. Nucleic Acids Research 44(D1):D527-D535. Available at: https://academic.oup.com/nar/article/44/D1/D527/2503129
2. Wikipedia contributors. (2020, January 19). Isozyme. In *Wikipedia, The Free Encyclopedia*. Retrieved 03:42, May 19, 2020, from https://en.Wikipedia.org/w/index.php?title=Isozyme&oldid=936548836
3. Wikipedia contributors. (2020, April 30). COX-2 inhibitor. In *Wikipedia, The Free Encyclopedia*. Retrieved 07:00, May 22, 2020, from en.Wikipedia.org/w/index.php?title=COX-2_inhibitor&oldid=954080651
4. Clarkson, C.W. (2018) Major Side Effects of NSAIDs and COX-2 Selective Inhibitors. *TUSOM Pharmwiki*. Available at: http://tmedweb.tulane.edu/pharmwiki/doku.php/nsaid_side_effects?do=
5. Santos, A.L. and Lindner, A.B. (2017) Protein Porsttranslational Modifications: Roles in Aging and Age-Related Disease. *Oxidative Medicine and Cellular Longevity*, Article ID: 5716409. Available at: https://www.hindawi.com/journals/omcl/2017/5716409/#copyright
6. Wikipedia contributors. (2020, May 12). Protein phosphorylation. In *Wikipedia, The Free Encyclopedia*. Retrieved 17:38, May 25, 2020, from en.Wikipedia.org/w/index.php?title=Protein_phosphorylation&oldid=956228409
7. Szylveszter, K.P., Németh, T., and Mócsai, A. (2019) Tyrosine Kinases in Autoimmune and Inflammatory Skin Diseases. Front. Immunol. 10.3389(2019.01862). Available at: https://www.frontiersin.org/articles/10.3389/fimmu.2019.01862/full
8. Wikipedia contributors. (2020, May 7). Histone. In *Wikipedia, The Free Encyclopedia*. Retrieved 21:12, May 25, 2020, from en.Wikipedia.org/w/index.php?title=Histone&oldid=955458038
9. Wikipedia contributors. (2020, May 13). Mucin. In *Wikipedia, The Free Encyclopedia*. Retrieved 01:10, June 7, 2020, from en.Wikipedia.org/w/index.php?title=Mucin&oldid=956387296
10. Wikipedia contributors. (2020, March 15). Allosteric regulation. In *Wikipedia, The Free Encyclopedia*. Retrieved 04:07, June 7, 2020, from en.Wikipedia.org/w/index.php?title=Allosteric_regulation&oldid=945637073

11. Dixit, Ajay. Dawra, Rajinder K. Dudeja, Vikas. Saluja, Ashok K. (2016). Role of trypsinogen activation in genesis of pancreatitis.
Pancreapedia: Exocrine Pancreas Knowledge Base, DOI: 10.3998/panc.2016.25
12. Coll-Martinez, B., and Crosas, B. (2019) How the 26S Proteasome Degrades Ubiquitinated Proteins in teh Cell. *Biomolecules* 9(9):395. Retrieved from: https://www.mdpi.com/2218-273X/9/9/395
13. Liu, W., Tang, X., Qi, X., Ghimire, S., Ma, R., Li, S., Zhang, N., and Si H. (2020) The Ubiquitin Conjugating Enzyme: An Important Ubiquitin Transfer Platform in Ubiquitin-Proteosome System. *Int J Mol Sci* 21(8):2894. Retrieved from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7215765/

6.7: Ribozymes - RNA Enzymes

6.7.1: Ribozymes

Any molecule that displays any of the catalytic motifs seen in the earlier chapters (general acid/base catalysis, electrostatic catalysis, nucleophilic catalysis, intramolecular catalysis, and transition state stabilization) can be a catalyst. So far we have examined only protein catalysts. These can fold to form unique 3D structures, which can have active sites with appropriate functional groups or nonprotein "cofactors" (metal ions, vitamin derivatives) that participate in catalysis. There is nothing special about the ability of proteins to do this. RNA also can form secondary and tertiary structures as we will see in Chapter 8. RNA molecules that act as enzymes are called **ribozymes**.

We are presenting the section before Chapter 8 for a few reasons. Most readers have encountered the structures of RNA and DNA before. They most likely know about three different types of RNA, ribosomal RNA (rRNA), transfer RNA (tRNA) and messenger RNA (mRNA). Likewise, they know from introductory biology classes the essential dogma of biology: (DNA, the holder of the genetic code) is transcribed into RNA which is translated into a protein sequence. Finally, most have studied (even at the high school level) that DNA of many species has exons and introns (intervening sequences), the latter that are spliced out of RNA transcripts to form mature RNA. In this section, we will discuss the catalytic properties of ribozymes, so the introductory background we just mentioned, although important, takes a "second" seat to the chemistry of catalysis, which is the main topic of Chapter 6.

There are 12 classes of ribozymes

- small self-cleaving RNAs (9 classes)
- Group I introns
- Group II introns
- Ribonuclease P

The large ribonucleoprotein nanoparticles, the spliceosome and ribosome, are also functionally ribozymes as well.

The term ribozyme is used for RNA that can act as an enzyme. Ribozymes are mainly found in selected viruses, bacteria, plant organelles, and lower eukaryotes. Ribozymes were first discovered in 1982 when Tom Cech’s laboratory observed Group I introns acting as enzymes. This was shortly followed by the discovery of another ribozyme, Ribonuclease P, by Sid Altman’s laboratory. Both Cech and Altman received the Nobel Prize in chemistry in 1989 for their work on ribozymes.

Ribozymes can be categorized based on size. Small ones, which usually don't require metal ions for activity, vary from 30-150 nucleotides while large ones can be a few thousand nucleotides in length. This translates into approximate molecule weights (using this formula for single-stranded RNA: (# nucleotides x 320.5) + 159.0) into 9800 for a 30-mer and 640,000 for a 2000 mer, typical of small and very large proteins/protein complexes, respectively. into two groups depending upon their size – small and large. Large ribozymes, which required metal ions for activity, can vary in size from a few hundred to several thousand nucleotides. Examples of small ones include hammerhead, viroid, hairpin, and riboswitch ribozymes. Examples of large ones include type I and II self-spicing introns, bacterial ribonuclease P, as well as the RNA in spliceosomes and ribosomes. Many also are not really true enzymes since they catalyze their own cleavage, although some can cleave presented RNA substrates. Large ones act as true catalysts.

Since RNAs can carry genetic information and act as enzymes, they probably evolved before proteins, which require nucleic acids for their synthesis. In addition, DNA required a special enzyme (ribonucleotide reductase) encoded by DNA to reduce the 2'OH to a 2'H. Artificial ribozymes have been made to catalyze many reactions that require protein enzymes. We will explore some ribozymes in several classes.

6.7.2: Small self-cleaving RNAs

We'll consider four small self-cleaving RNAs- hammerhead, viroid and hairpin ribozymes as well as the glucosamine-6-phosphate riboswitch (glmS). All catalyze the cleavage of an internal phosphodiester bond (cis catalysis) or in a presented substrate (trans catalysis) by a transesterification reaction. Internal cis catalysis reactions cleave the ribozyme into two fragments, which inactivates their catalytic activity. In that sense, they don't act as true catalysts since they engage in only one cycle of cleavage. Trans catalysis in which a substrate RNA binds to and is cleaved by the ribosome would be considered true catalysis.

The cleavage reaction in an internal cis cleavage is an SN2 trans-esterification reaction as shown in Figure 6.7.1.

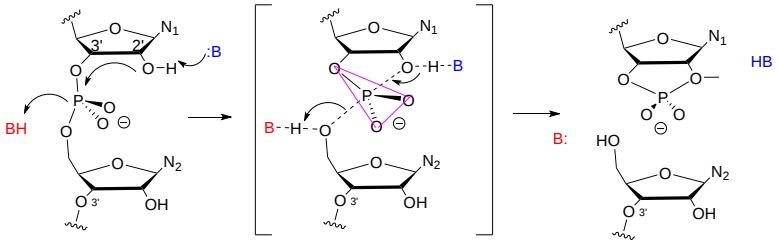


Figure 6.7.1: SN2 trans-esterification cleavage of an internal phosphodiester bond in small RNA ribozymes

In the reaction, an adjacent general base (:B) abstracts a proton of the C'2 OH of nucleotide N_1. The resulting 2' O^- acts as a nucleophile in a SN2 reaction and attacks the δ^+ phosphorous in the phosphodiesterase bond, forming a pentavalent, trigonal pyramidal sp^3d hybridized intermediate/transition state, which collapses breaking the phosphodiesterase bond between nucleotide N_1 and N_2. This is an inline mechanism in that the incoming nucleophilic O in C2' and the exiting one on the O of 5' CH_2OH of nucleotide 2 are axial to each other separated by 180^0. A general acid, BH, facilitates the departure of the exiting nucleophile by its protonation. Bound metal ions may facilitate the reaction and can be considered cofactors but might be more involved in maintaining a catalytically-active structure. Self-cleaving small RNAs are also found in humans and may be part of long noncoding RNAs.

a. Hammerhead RNA

The hammerhead ribozyme is a small RNA ribozyme with a conserved core with three helical stems. It has a structure similar to the head of a hammerhead shark. One predicted secondary structure of a hammerhead ribozyme is shown below in Figure 6.7.2

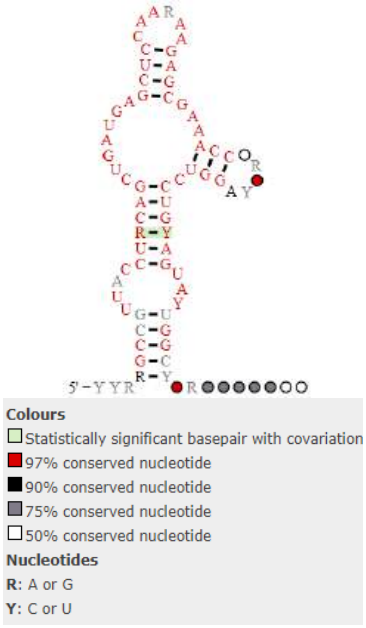


Figure 6.7.2: Predicted secondary structure and sequence conservation of the HH9 ribozyme found conserved from lizard to human genomes.

A possible trigonal pyramidal intermediate/transition state in the Hammerhead ribosome is shown in Figure 6.7.3.

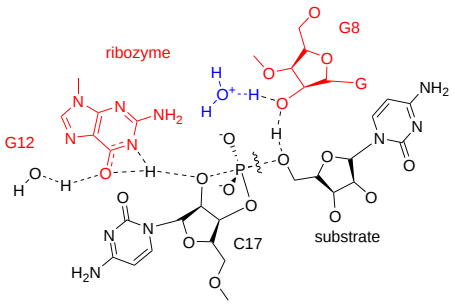
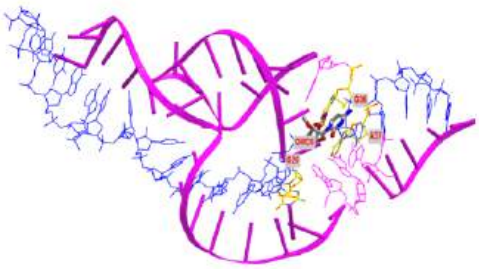


Figure 6.7.3: Role of G8 and G12 in active site of full length Schistosoma mansoni hammerhead ribozyme. After Martick and Scott, Cell. 126, 2006 DOI:https://doi.org/10.1016/j.cell.2006.06.036

A deprotonated G-12 in the ribozyme probably acts as a general base that activates the 2'-OH to form the incoming nucleophile that attacks the trans-substrate RNA. The 2′-OH of G-8 in the ribozyme appears to hydrogen bond to the 5′-O of the departing nucleophile in the substrate where bond scission occurs. The ribozyme increases the rate by 1000-fold.

Figure 6.7.4 shows an interactive iCn3D model of the full-length Schistosoma manson catalytically active hammerhead ribozyme (3dz5). This is an example of a ribozyme that acts in trans as the cleaved phosphodiester bond is a bound RNA single-stranded substrate.



NCBI iCn3D Figure 6.7.4: Full-length Schistosoma manson catalytically active hammerhead ribozyme (3dz5). (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?qNBw2MY1RZ7xKogH7

The numbering system is a bit different in the iCn3D model above. PDB requires sequential numbering whereas ribozymes sequences are numbered in discontinuous ways. Core residues that are conserved are given common numbers. However, the different projecting double-stranded RNA regions, which vary among ribozymes, are numbered differently. The G8 and G12 in Figure 6.7.3 are numbered G20 and G36, respectively.

b. ***Viroids***

Viroids are small single-stranded circular RNAs that infect plant cells. They are not packaged with viral capsid protein. Some enter cells along with viruses and are called **virusoids** or **viroidlike satellite RNAs**. Intrastrand pairing occurs and they are synthesized as tandem repeats containing multiple adjacent copies of the viroid. These repeats are cut and ligated to form the individual mature viroid by internal ribozymes sequences. One example is the hepatitis delta virus (HDV), a satellite of the hepatitis B virus. A possible mechanism of catalysis of the hepatitis delta virus ribozyme involving general acid/base catalysis is shown in Figure 6.7.5.

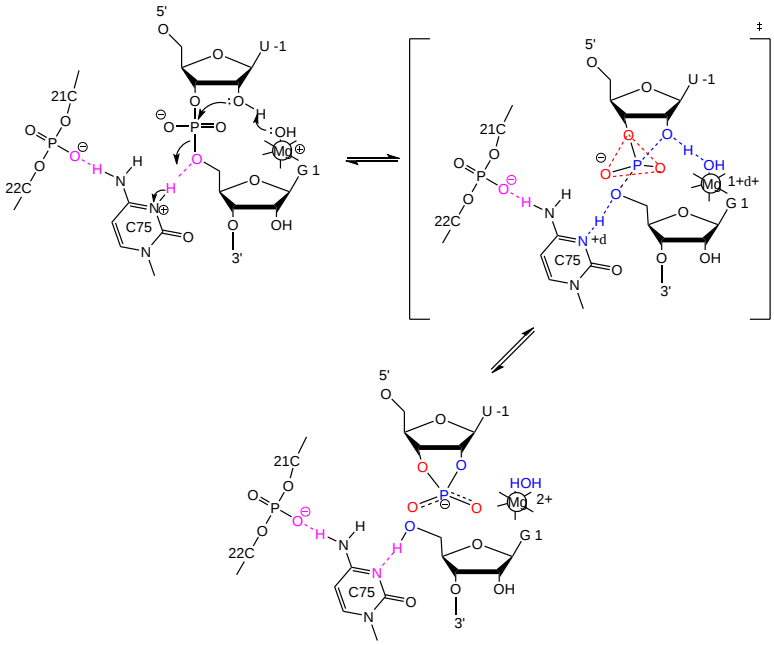


Figure 6.7.5: General acid/base catalysis by Hepatitis delta virus (HDV) ribozyme

c. ***Hairpin ribozyme***

Hairpin ribozymes are encoded by the satellite RNA of plant viruses. They are about 50 nucleotides long, and can cleave itself internally, or, in a truncated form, can cleave other RNA strands in a transesterification reaction. The structure consists of two domains, stem A required for binding (self or other RNA molecules) and stem B, required for catalysis. Self-cleavage in the hairpin ribozyme occurs in stem A between an A and G bases (which are splayed apart) when the 2' OH on the A attacks the phosphorous in the phosphodiester bond connecting A and G to form a pentavalent intermediate.

Rupert et al solved the crystal structure of a hairpin ribozyme with a non-cleavable substrate analog containing a 2'-O-CH3 group on the ribose. This acts as a nucleophile in the transesterification cleavage of RNA.

A38 in Stem B appears to be able to interact with the products (the cleaved A now in the form of a cyclic phosphodiester with itself) and the departing G, and with a transition state pentavalent analog of the sessile A-G bond in which the phosphodiester linking A and G in the substrate is replaced with a pentavalent vanadate bridge between A and G. This is Illustrated in Figure 6.7.6.

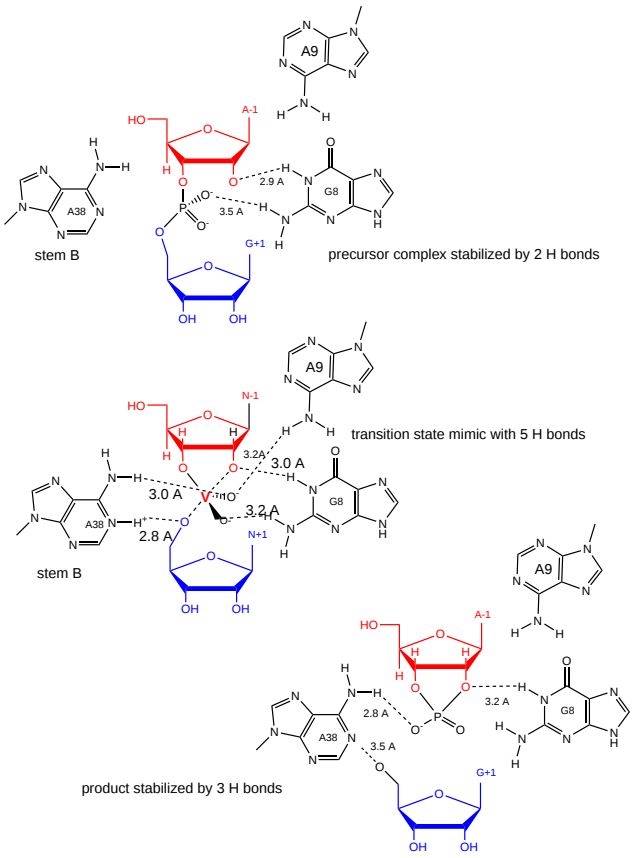


Figure 6.7.6: Transition state binding stabilization of hairpin ribozyme (adapted from Rupert et al, Science, 298 pg 1423 2002)

However, A 38 does not appear to react with the sessile A -G groups in the normal substrate, indicating that the main mechanism used by this ribozyme is transition state binding. Since RNA molecules have fewer groups available for acid/base and electrostatic catalysis (compared to protein enzymes), ribozymes, presumably the earliest type of biological catalyst, probably make more use of transition state binding as their predominant mode of catalytic activity.

Figure 6.7.7 shows an interactive iCn3D model of the hairpin ribozyme in the catalytically-active conformation (1M5K).

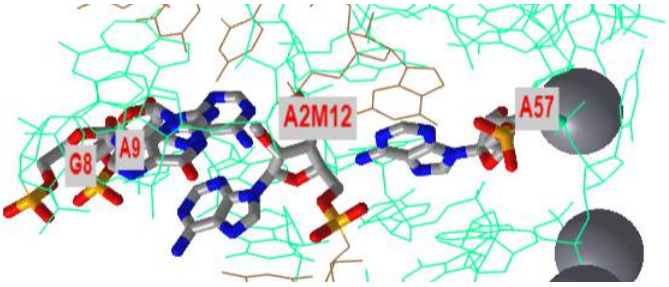


Figure 6.7.7: Hairpin ribozyme (human) in the catalytically-active conformation (1M5K). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...goBWNHUbLU7my8

The hairpin ribozyme is shown in cyan sticks and the inhibitor substrate in brown sticks. The inhibitor contains a 2'-O-methyl adenosine (A2M12), so it can not be cleaved and instead acts as an inhibitor. The A38 shown in the catalytic mechanism is labeled A57 in the iCn3D.

d. ***Glucosamine-6-phosphate riboswitch (glmS)***:

A novel use of ribozymes was recently reported by Winkler et. al. They discovered that the 5' end of the mRNA of the gene glmS (from Gram-positive bacteria) is a ribozyme. The GlmS gene encodes glucosamine-6-phosphate synthetase (GlmS), which catalyzes the reaction of fructose-6-phosphate and glutamine to glucosamine-6-phosphate (GlcN6P) and glutamate. This is the first committed step in bacterial cell wall synthesis. Glucosamine-6-phosphate binds to the ribozyme (3' end of the mRNA) and acts as a cofactor leading to self-cleavage of the ribozyme. What an amazing mechanism for pathway inhibition. At high GlcN6P

concentrations, it binds to the ribozyme, inhibiting its own synthesis. G40 in the active site appears to act as a general base. Figure 6.7.8 shows an interactive iCn3D model of GlmS Ribozyme Bound to Its Catalytic Cofactor, glucosamine 6 phosphate (GlcN6P) (**2NZ4).**

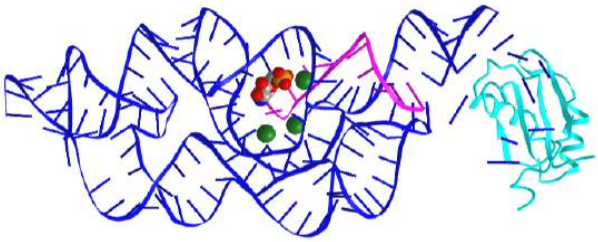


Figure 6.7.8: GlmS ribozyme bound to its catalytic cofactor glucosamine 6 phosphate (GlcN6P) (2NZ4) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...J19zdfwEujH4c8

Riboswitches are discussed in greater detail in Chapter 28.1: Regulation of Gene Expression in Bacteria.

6.7.3: Group I and Group II Introns

Introns present in RNA molecules must be removed to form mature RNA. In humans, about 80% of introns are less than 200 nucleotides long, but some can be 10,000 nucleotides or longer in length. Before we discuss introns, we'll provide a quick background on RNA splicing. There are two major types of self-splicing introns, Groups I and II. Other introns are removed by a ribonucleoprotein called the spliceosome. Some call these Group III introns. A simple two-step mechanism for the self-splicing Group I and II introns is shown in Figure 6.7.9.

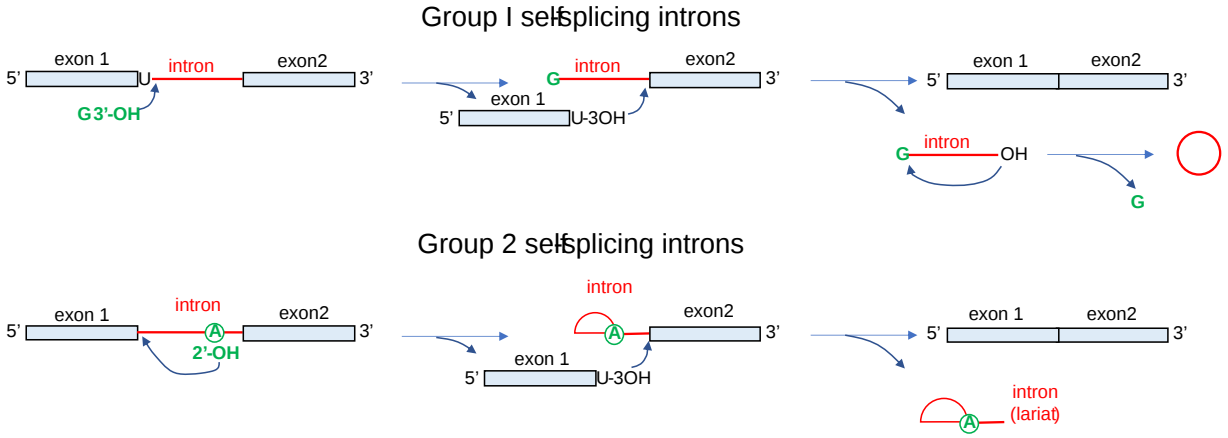


Figure 6.7.9: Self-splicing of Group I and II introns. (after N.K. Tanner /FEMS Microbiology Reviews 23 (1999)

Both required first a scission of the RNA strand followed by ligation of the two exons to form the mature RNA. Note that Group I introns require an external guanosine nucleophile and the removed intron forms a circular RNA when removed. In contrast, in Group II introns, an internal A residue acts as the first nucleophile in the scission reaction and the intron on removal forms a branched lariat structure.

A simple two-step mechanism for Group II introns which are spliced out from pre-mRNA in the nucleus by a ribonucleoprotein complex called the **splicesome** is shown in Figure 6.7.10.

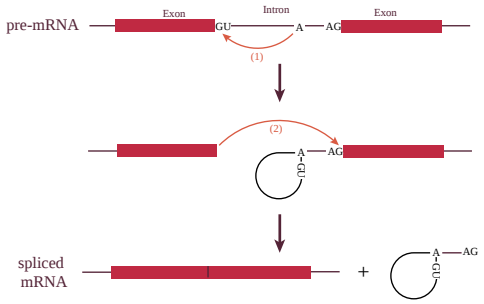


Figure 6.7.10: Two steps of canonical RNA processing, from pre-mRNA to spliced RNA and the branched lariat intron. https://commons.wikimedia.org/wiki/F...g_reaction.svg. Creative Commons Attribution-Share Alike 3.0 Unported

Note that the mechanism is extremely similar to the auto-removal of Group II introns, suggesting an evolutionary relationship between the two.

Figure 6.7.11 below shows a more detailed view of the catalytic cycle of the spliceosome. Five small ribonucleoproteins (U1, U2, U4/U6 and U5 snRNPs) assemble on the nuclear pre-mRNA and facilitate the removal of the intro,n but the main mechanism involves ribozyme activity.

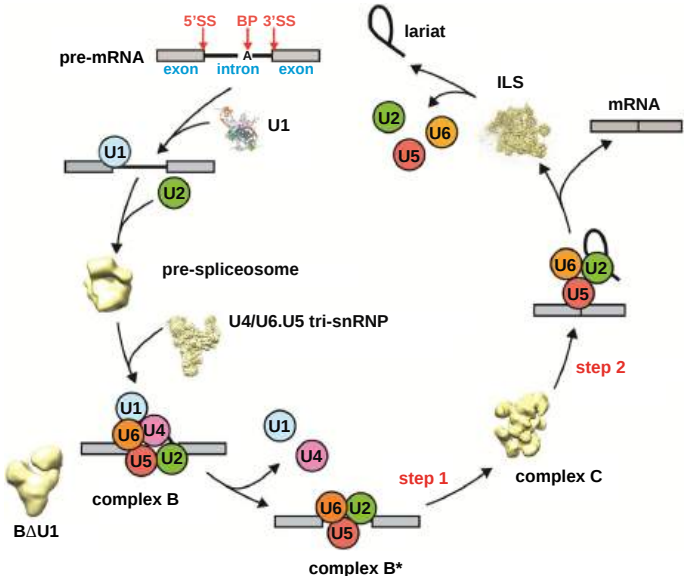


Figure 6.7.11: Step-wise spliceosome assembly from its U-snRNP components. Nguyen et al. Current Opinion in Structural Biology 2016, 36:48–57. http://dx.doi.org/10.1016/j.sbi.2015.12.005. CC BY license (http://creativecommons.org/licenses/by/4.0/).

6.7.3.1: Group I introns

These are found in bacteria, lower eukaryotes (including mitochondrial and chloroplast RNA) and higher plants and are in ribosomal RNA (rRNA), mRNA and tRNA. They are also found in Gram-positive bacteria bacteriophages (viruses that attack bacteria). As shown in Figure 6.7.9, they require guanosine as a cofactor and have a single active site for both scission and ligation to produce the mature mRNA, tRNA or rRNA. Mg^2 is required not for catalysis per se but to maintain the correct tertiary structure of the ribozyme with the correct secondary structure. In Group I introns, the splicing reaction is initiated by a **guanosine cofactor**. They have one active site that catalyzes the initial cleavage of the phosphodiester bond and the final religation after cleavage.

The Group I catalytic core from Tetrahymena thermophila has two domains. A cleft is formed between them when they pack which can bind the short helix with the 5' splice site. and the guanosine cofactor. This "active" site is preformed without substrates similar to the active sites of protein enzymes. Figure 6.7.12 shows the secondary structure and the reaction of the group I intron ribozyme from *Tetrahymena*.

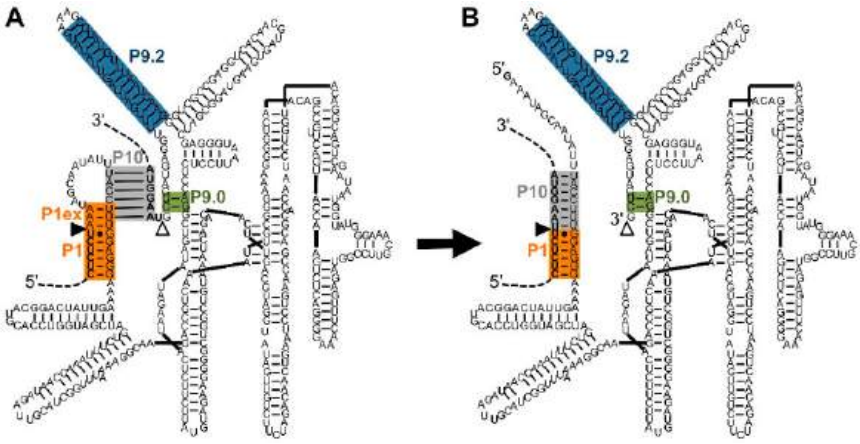
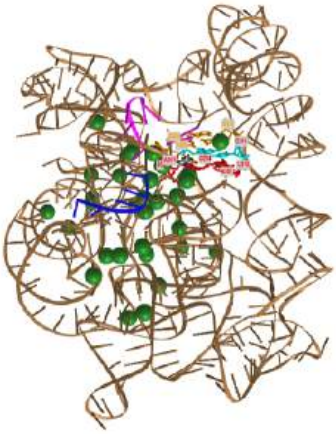


Figure 6.7.12: Secondary structure and reaction of the natural, *cis*-splicing group I intron ribozyme from *Tetrahymena*. Müller. *Molecules* **2017**, *22*(1), 75; https://doi.org/10.3390/molecules22010075. Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).

The intron runs between the two triangles, which show the 5' (filled triangle) and 3' (open triangle) splice sites. The orange P1 helices contain the 5′-splice site (G:U). The P10 helix (grey), P9.0 helix (green), and the P9.2 helix (blue) are structured to present the appropriate 3′-splice site. During splicing (from (**A**) to (**B**)), the P1 helix extension (orange) is opened to expose the 3′-hydroxyl group of the terminal uridine at the 5′-splice site. The P10 helix then facilitates a conformational change, in which the 3′-exon (upper dashed line) is positioned adjacent to the 5′-exon (lower dashed line), allowing the nucleophilic attack of the 3′-uridine the 3′-splice site, joining 5′-exon and 3′-exon.

Figure 6.7.13 shows an interactive iCn3D model from cryoEM of the full-length holo L-16 ScaI Tetrahymena ribozyme (7EZ2).



NCBI iCn3D Figure 6.7.13: Holo L-16 ScaI Tetrahymena ribozyme (7EZ2) (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...PhK3BSbFLDnDt8

Three sets of coplanar bases are found in the active site. These include C262, A263 and G312 (top layer, brown), G264, C311 and the ωG - labeled G3 in the iCn3D model- (cyan layer) and A261, A265 and U310 (bottom layer red). The ωG - labeled G3 - is the nucleophile.

The structure is nearly identical to the apo-form of the ribozyme with just an internal guide RNA sequence undergoing a large change and the guanosine binding site undergoing a small shift on binding RNA substrates.

6.7.3.2: Group 2 Introns

Group II introns are found in mRNA of bacteria and some Archaea, in rRNA, tRNA, and mRNA of chloroplasts and mitochondria, and in fungi, plants, and protists. No Class 2 introns appear to be found in eukaryotic genomes. Some of these introns are in gene-encoding proteins, but most are in bacterial noncoding sequences. In Group II introns, the splicing reaction is initiated by an adenosine cofactor, as shown in Figure 6.7.9.

What's especially interesting about group II introns is that they can reinsert in DNA so that can be considered to be mobile genetic elements. A **maturase/reverse transcriptase** enzyme (562 amino acids) is associated with the intron, which helps stabilize the active site of the intron for the **reversible** session and ligation of the intron. Reintegration of the excised branched lariat intron into DNA is called retrotransposition (copy/paste). Mitochondrial and chloroplast Group 2 intron have lost their mobility and act as classic introns.

The structures of a group II intron from Thermosynechococcus vestitus (a cyanobacteria) before and after integration have been determined. A branch-site domain VI helix swings 90°, enabling DNA integration. The maturase/revere transcriptase protein assists excision of the intron through the interaction of domain VI of that intron that positions the key adenosine for branched lariat formation during forward splicing, as shown in Figure 6.7.9. The changes in the structure of the group II intron retroelement before (6ME0) and after DNA integration (6MEC) are shown in Figure 6.7.14

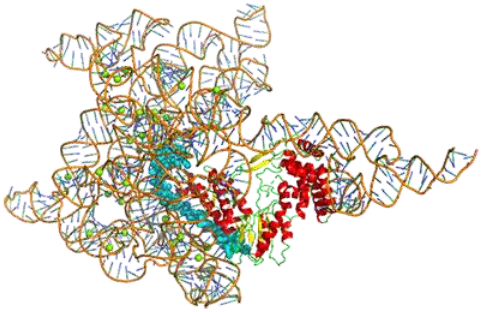
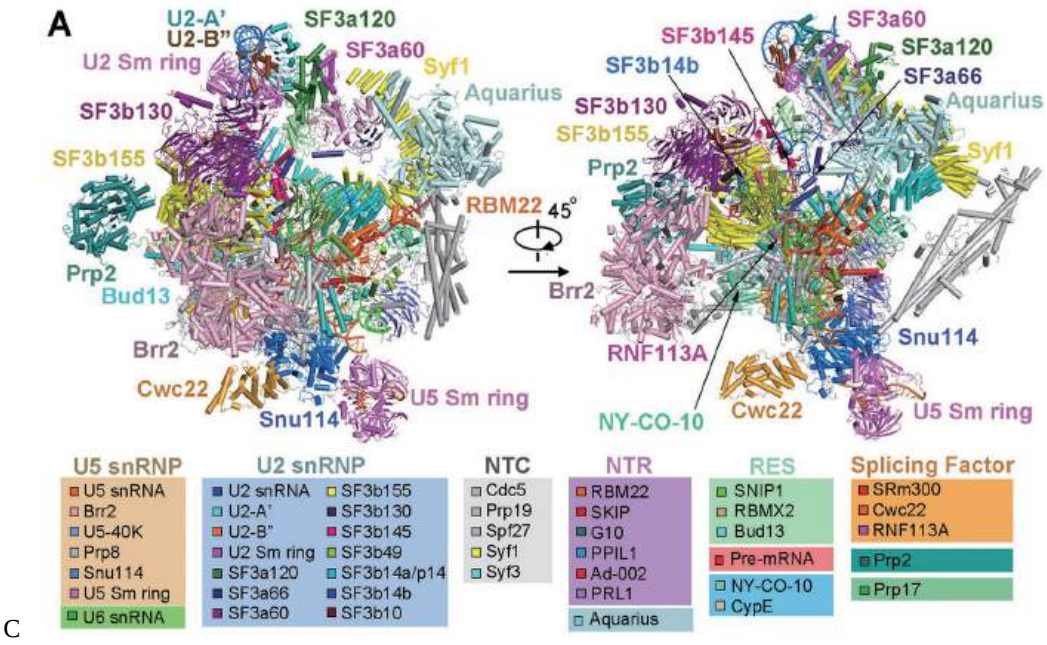


Figure 6.7.14: Changes in the structure of the group II intron retroelement before (6ME0) and after DNA integration (6MEC)

The target DNA before integration is shown in cyan spacefill and in orange spacefill after integration. The protein is shown as a colored cartoon. The Group II intron is 867 nucleotides and the sense target DNA is 47 nucleotides in length.

6.7.4: Spliceosomal Introns

A comparison of Figure 6.7.9 through Figure 6.7.11 show the similarities between spliceosomal introns and group II self-splicing introns. Spliceosomal and group II self-splicing introns are structurally and mechanistically homologous, right down to the stereochemistry of the splicing reaction. In eukaryotes, introns in pre-mRNA are removed by splicing and subsequent exon ligation, releasing the intron as a branched lariat molecule. This reaction is performed by the spliceosome, a large nuclear ribonucleoprotein (RNP) complex. Spliceosomes remove introns and splice the exons of most **nuclear genes**. They are composed of 5 kinds of small nuclear RNA (**snRNA**) molecules and over 100 different protein molecules. It is the RNA — not the protein — that catalyzes the splicing reactions. The molecular details of the reactions are similar to those of Group II introns, and this has led to speculation that this splicing machinery evolved from them. Figure 6.7.15 shows two views of the cryo-EM structure of the human-activated spliceosome (the Bact complex)



C

Figure 6.7.15: Cryo-EM structure of the human-activated spliceosome (the Bact complex). Zhang et al. https://www.nature.com/articles/cr201814.pdf. Creative Commons Attribution 4.0 Unported License. http://creativecommons.org/licenses/by/4.0/

There are 52 proteins, 3 small nuclear RNAs (snRNA), and one pre-mRNA. The total molecular mass is 1.8M. U2, U5, and U6 snRNAs are colored marine, orange, and green, respectively. Pre-mRNA is colored red. Figure 6.7.16 shows just the structural changes in RNA and protein components that occur between the early Bact complex (left panel) and the mature Bact complex (right panel).

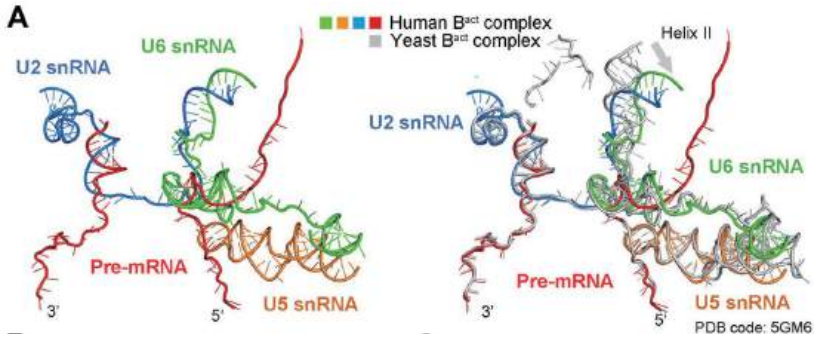


Figure 6.7.16: Structural changes in RNA and protein components that occur between the early Bact complex (left panel) and the mature Bact complex of the human spliceosome.

6.7.5: Ribonuclease P

This enzyme is a ribonucleoprotein that cleaves RNA through the catalytic action of one essential RNA subunit that displays ribozyme activity. It's found in most organisms. As with many ribozymes, the activity is increased 2-3 fold with bound proteins that stabilize the folded ribozyme and help bind the preferred substrate, which is pre-tRNA. Figure 6.7.17 shows bacterial RNase P ribozyme in complex with tRNA (3q1r).

Figure 6.7.17: Bacterial RNase P ribozyme in complex with tRNA (3q1r).

The tRNA is shown as a magenta cartoon/surface and the protein in cartoon form. The enzyme cleaves the 5' head end of the precursors of transfer RNA (tRNA) molecules. In bacteria, the enzyme is a heterodimer with one RNA and protein subunit.

Figure 6.7.18 shows a possible transition state with key residues involve in binding two Mg^{2+} ions in the active site. These ions are essential for catalysis as they stabilize the pentavalent intermediate/transition state.

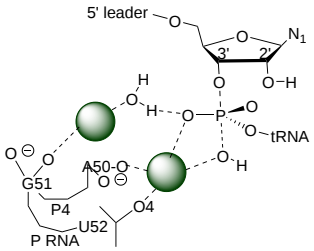


Figure 6.7.18: Transition state of RNase P with key RNA active site

Figure 6.7.19 shows an interactive iCn3D model of the active site with key residues labeled of bacterial RNase P ribozyme in complex with tRNA (3q1r).

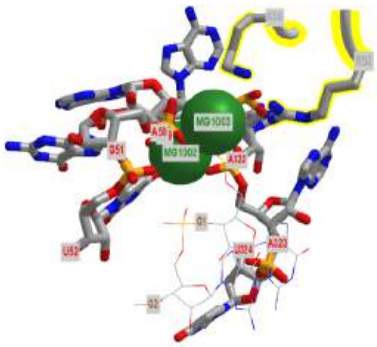


Figure 6.7.19: Bacterial RNase P holoenzyme in complex with tRNA (3q1r) (Copyright; author via source).
Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...6SjVMa8g6K47x5

6.7.6: RNA Polymerase Ribozyme

If primordial RNA acted as both a metabolic enzyme catalyst as well as the holder of the genetic information, it would need to have RNA polymerase activity. Artificial ribozymes with class I RNA ligase activity have been made. Figure 6.7.20 shows an interactive iCn3D model of the active site region of the Class I ligase ribozyme-substrate preligation complex, C47U mutant, Mg^{2+} bound (3R1L).

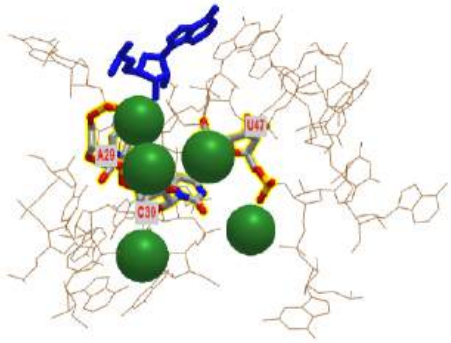


Figure 6.7.20: Active site of the Class I ligase ribozyme-substrate preligation complex, C47U mutant, Mg^{2+} bound (3R1L). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...EEQV3Mzc8Jzdv6

The blue stick represents just the 3' terminal adenosine end of the target substrate (5'UCCAGUA3') to which a new nucleoside would be added. The brown represents the active site region of the ribozyme. Three catalytic residues A29, C30 and C47 have been identified in the actual ribozyme. In iCn3D model is of a mutant, C47U, which has no catalytic activity. The green sphere represents Mg^{2+} ions.

Figure 6.7.21 show the active site residues and how they might facilitate stabilization of the pentavalent intermediate/transition state and the similarity of the active site to a protein RNA polymerase.

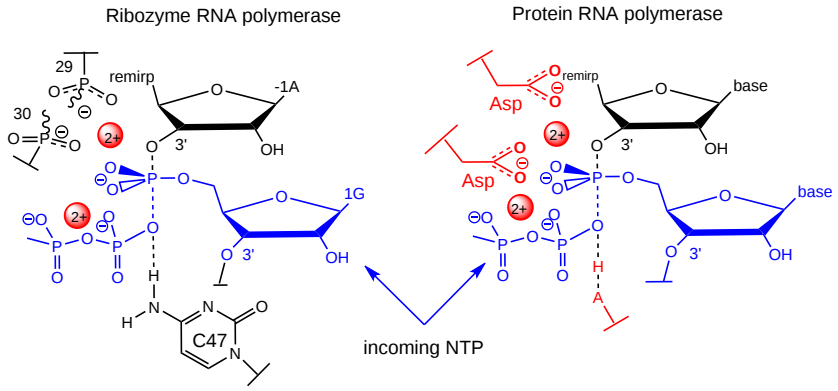


Figure 6.7.21: Comparison of Transition State Models of Ribozyme RNA ligase and a protein RNA Polymerases (after Schechner et al)

A divalent Mg^{2+} in the active site of the ribozyme enhances the nucleophilicity of the 3-OH on the primer, which attached the terminal phosphate of the G(1)TP substrate to form a pentavalent intermediate. The Mg cation is stabilized by oxygens on P 29 and 30 of the ribozyme. The Mg ion also stabilizes the developing charge in the transition state and in the charge in the intermediate. Stabilization of analogous divalent cations in the protein polymerase occurs through Asp side changes in the protein.

6.7.7: Ribosome

Protein synthesis from mRNA templates occurs on a ribosome, a nanomachine composed of proteins and ribosomal RNAs (rRNA). The ribosome is composed of two very large structural units. The smaller unit (termed 30S and 40S in bacteria and eukaryotes, respectively) coordinates the correct base pairing of the triplet codon on the mRNA with another small adapter RNA, transfer or tRNA, that brings a covalently connected amino acid to the site. Peptide bond formation occurs when another tRNA-amino acid molecule binds to an adjacent codon on mRNA. The tRNA has a cloverleaf tertiary structure with some intrastrand H-bonded secondary structure. The last three nucleotides at the 3' end of the tRNA are CpCpA. The amino acid is esterified to the terminal 3'OH of the terminal A by a protein enzyme, aminoacyl-tRNA synthetase.

Covalent amide bond formation between the second amino acid to the first, forming a dipeptide, occurs at the peptidyl transferase center, located on the larger ribosomal subunit (50S and 60S in bacteria and eukaryotes, respectively). The ribosome ratchets down

the mRNA so the dipeptide-tRNA is now at the **P** or **Peptide site**, awaiting a new tRNA-amino acid at the **A** or **Amino site**. The figure below shows a schematic of the ribosome with bound mRNA on the 30S subunit and tRNAs covalently attached to amino acid (or the growing peptide) at the A and P site, respectively. Figure 6.7.22 shows a cartoon model of the prokaryotic ribosome with bound mRNA, tRNAs and the P and A sites.

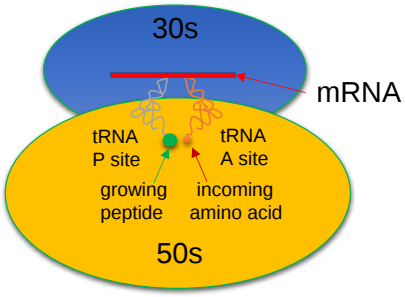


Figure 6.7.22: Cartoon model of the prokaryotic ribosome with bound mRNA, tRNAs and the P and A sites.

We present another more detailed model of the ribosome complex illustrating protein synthesis in Figure 6.7.23

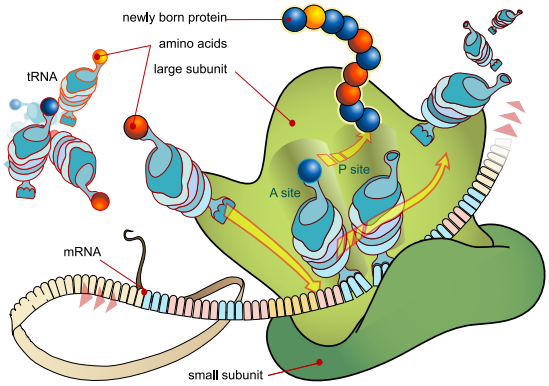


Figure 6.7.23: Ribosome mRNA translation. https://en.Wikipedia.org/wiki/File:R...slation_en.svg. public domain by its author, LadyofHats

A likely mechanism (derived from crystal structures with bound substrates and transition state analogs) for the formation of the amide bond between a growing peptide on the P-site tRNA and the amino acid on the A-site tRNA is shown below. Catalysis does not involve any of the ribosomal proteins (not shown) since none is close enough to the peptidyl transferase center to provide amino acids that could participate in general acid/base catalysis, for example. Hence the rRNA must act as the enzyme (i.e. it is a ribozyme). Initially, it was thought that a proximal adenosine with a perturbed pKa could, at physiological pH, be protonated/deprotonated and hence act as a general acid/base in the reaction. However, none was found. The most likely mechanism to stabilize the oxyanion transition state at the electrophilic carbon attack site is precisely located water, which is positioned at the oxyanion hole by H-bonds to uracil 2584 on the rRNA. The cleavage mechanism involves the concerted proton shuffle shown below. In this mechanism, the substrate (Peptide-tRNA) assists its own cleavage in that the 2'OH is in position to initiate the protein shuttle mechanism. (A similar mechanism might occur to facilitate hydrolysis of the fully elongated protein from the P-site tRNA.) Of course, all of this requires perfect positioning of the substrates and isn't that what enzymes do best? The main mechanisms for catalysis of peptide bond formation by the ribosome (as a ribozyme) are intramolecular catalysis and transition state stabilization by the appropriately positioned water molecule. These processes are illustrated in Figure 6.7.24

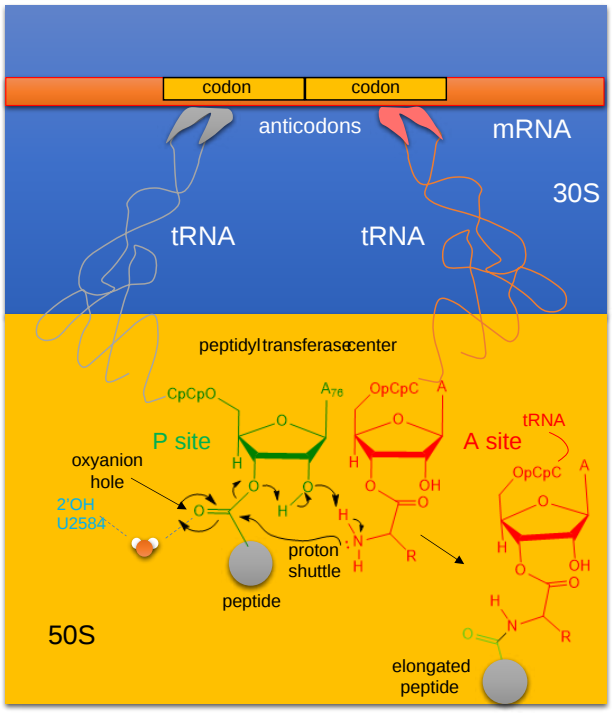


Figure 6.7.24: Mechanism Peptide Bond Formation by the Ribosome

The crystal structure of the eukaryotic ribosome has recently been published (Ben-Shem et al). It is significantly larger (40%) with a mass of around $3x10^6$ Daltons. The 40S subunit has one rRNA chain (18) and 33 associated proteins, while the larger 60S subunit has 3 rRNA chains (25S, 5.8S and 5S) and 46 associated proteins. The larger size of the eukaryotic ribosome facilitates more interactions with cellular proteins and greater regulation of cellular events. Figure 6.7.25 shows the two copies of the 80S yeast ribosome (4v88), presented to humble readers and authors alike.Figure 6.7.25: The 80S yeast ribosome (4v88). Each subunit is given a different color.

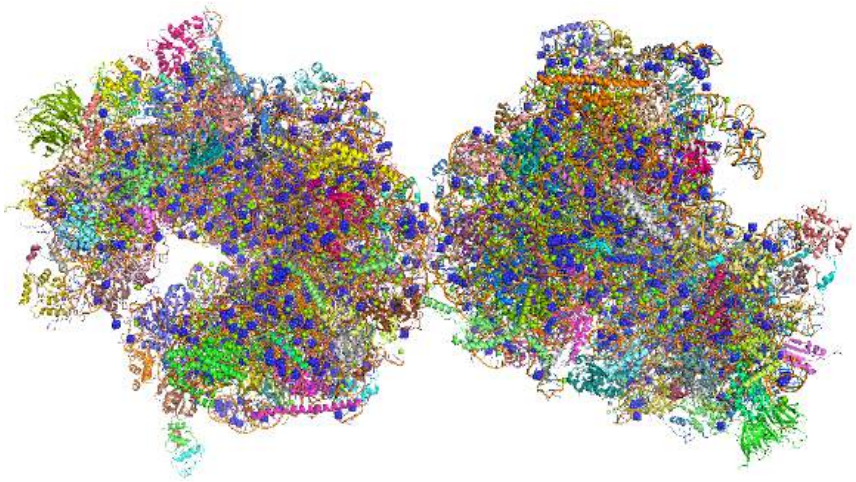


Figure 6.7.25: Structure of the yeast 80S ribosome.

6.7.8: Ribozyme methyltransferase

The ribozymes described above and generally found in nature catalyze phosphoryl transfer reactions and with the ribosome, peptide bond formation. In vitro evolution can be used to drive new enzymatic functionalities, which would have been required in a RNA-only world that preceded the use of proteins as catalysts. RNA ribozymes are limited in having only 4 bases that can be employed in binding and catalytic steps, compared to 20 amino acids which can serve the same function in proteins. However, as in the case of protein, small molecule cofactors that bind to a potential ribozyme might facilitate greater catalytic efficiency and an expanded

repertoire of reaction types. Indeed, we have seen above how small molecules can bind to riboswitches. Figure 6.7.26 shows the reaction and structure of a methyltransferase 1 ribozyme (MTR1) that acts as a methyltransferase. The small ligand, O^6-methylguanine, binds to the ribozymes and acts as a cofactor in the methylation of adenine 63 in the RNA.

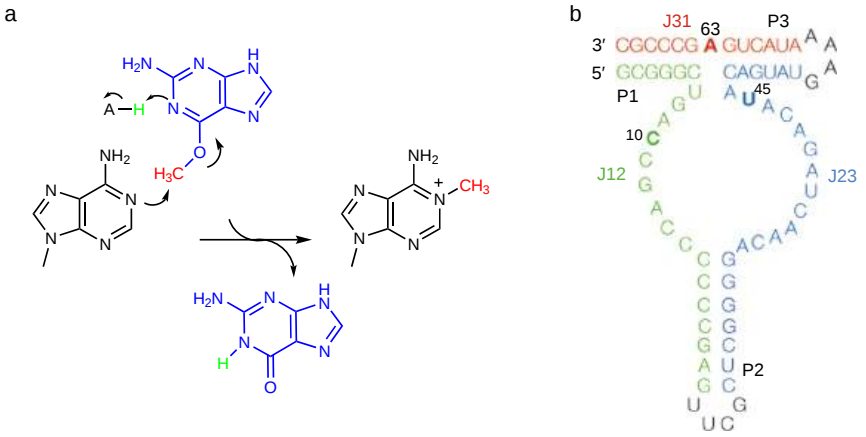


Figure 6.7.26: Deng, J., Wilson, T.J., Wang, J. *et al.* Structure and mechanism of a methyltransferase ribozyme. *Nat Chem Biol* **18**, 556–564 (2022). https://doi.org/10.1038/s41589-022-00982-z. Creative Commons Attribution 4.0 International License. Creative Commons Attribution 4.0 International License, http://creativecommons.org/licenses/by/4.0/

Panel a shows the chemical reaction in which the methyl group of the small ligand O^6-methylguanine is transferred to N1 of Adenine 63 in the RNA.

Panel b shows the sequence of the MTR1 ribozyme as crystallized for the experiments. The RNA is a three-way junction composed of three arms P1, P2 and P3. A GNRA tetraloop has been added to the end of the P3 helix so that the entire ribozyme comprises a single RNA strand. Subsections of the strands are named J12 (colored green), J23 (colored blue) and J31 (colored red).

X-ray crystal structures of the ribozyme in the presence of the cofactor O^6-methylguanine were determined. The final structure contained guanine and an A63 methylated adenosine (1MA) implying the methyl group of the O^6-methylguanine had transferred to A63, leaving guanine bound in the active site. The structure of the guanine bound to the ribozyme is shown in Figure 6.7.27.

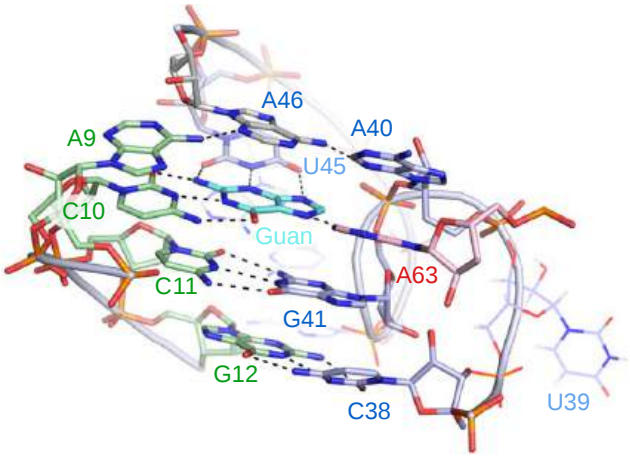


Figure 6.7.27: The four planes of nucleobase interactions in the core of the ribozyme. The four planes are composed of (from bottom to top) G12•C38, C11•G41, exogenous guanine hydrogen bonded to C10, U45 and A63 and the triple interaction A9•A46•A40. Deng, J et al, ibid.

A hypothetical reaction mechanism is shown in Figure 6.7.28

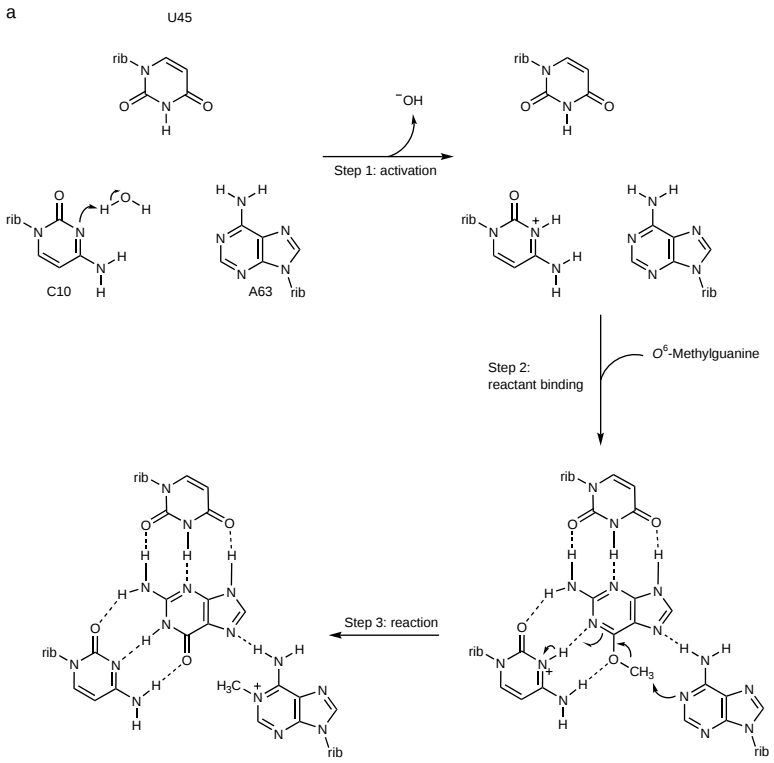


Figure 6.7.28: Proposed catalytic mechanism for the ribozyme methyltransferase. Deng, J et al, ibid.

In step 1, the nucleobase of C10 becomes protonated, and in step 2, the O^6-methylguanine becomes bound. However, it is likely that these two steps are coordinated, as the binding will raise the pK_a of the cytosine.

The methyl transfer reaction occurs in step 3 by the nucleophilic attack of A63 N1 on the methyl group of O^6-methylguanine and the coordinated breakage of the guanine O6–C bond. This involves a train of electron transfers, the movement of the proton from C10 N3 to guanine N1 and a concomitant shift of the positive charge from C10 to the N^1-methyladenine at position 63. In principle, the guanine can now be released as product, although there is no evidence that this occurs with the present form of the ribozyme. Regeneration of active ribozyme would also require an exchange of the substrate strand to place unmethylated adenine at position 63.

The proposed mechanism is fully consistent with the structure of the MTR1 riboswitch and the effect of the substitutions at C10 and U45 on activity. The complete loss of methylation activity of the C10U variant is fully consistent with the proposed role as a general acid in addition to ligand binding.

Figure 6.7.29 shows an interactive iCn3D model of a methyltransferase ribozyme (7V9E).

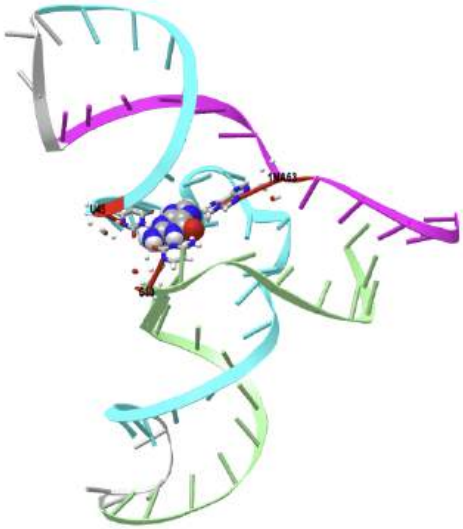


Figure 6.7.19: A methyltransferase ribozyme (7V9E). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...LtyVP4DBnkEgH6

The parts of the ribozyme are named J12 (light green), J23 (cyan) and J31 (magenta). The active site residues, C10, U45 and 1MA63 (1-methyladenosine) are shown in CPK-colored sticks and labeled (disregard small separated spheres). The Ba^{2+} ion in the crystal structure is not displayed. The guanine is shown in spacefill.

6.8: Cofactors and Catalysis - A Little Help From My Friends

6.8.1: Cofactors and Electron Pushing: Sources and Sinks

To make and break bonds, electrons have to be moved. In drawing reaction mechanisms, we showed how electrons move from "sources" to "sinks". In many enzyme-catalyzed reactions, vitamin derivatives are used as substrates or "cofactors" or "coenzymes" to facilitate the flow of electrons in bond-making and breaking. The section focuses on cofactors, which facilitate the flow of electrons from the substrate to product. We will see these enzymes in more detail in specific chapter sections.

Cofactors are molecules that bind to enzymes and are required for catalytic activity. They can be divided into two major categories: ***metals*** and ***coenzymes. Metal cofactors*** commonly found in human enzymes include iron, magnesium, manganese, cobalt, copper, zinc, and molybdenum. ***Coenzymes*** are small organic molecules that are often derived from vitamins. ***Coenzymes*** can bind loosely with the enzyme and release from the active site. As such, they are also considered substrates for the reaction. Alternatively, they may be tight binding and cannot dissociate easily from the enzyme. In this case, after their initial participation in an enzyme-catalyzed reaction, the enzyme would no longer be able to use the cofactor in another round of catalysis until the initial state of the cofactor is reformed, which takes another chemical reaction and often an additional substrate.

Tight-binding coenzymes are referred to as ***prosthetic groups***. Enzymes not yet associated with a required cofactor are called ***apoenzymes***, whereas enzymes bound with their required cofactors are called ***holoenzymes.*** Sometimes organic molecules and metals combine to form coenzymes, such as in the case of the heme cofactor (Figure 7.15). Coordination of heme cofactors with their enzyme counterparts often involves interactions with histidine residues, as shown in the succinate dehydrogenase enzyme shown in Figure 6.8.1.

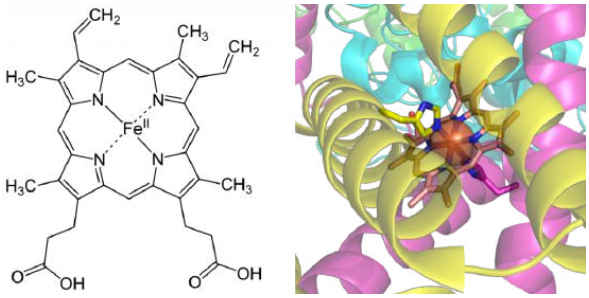
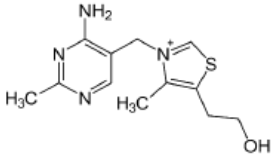
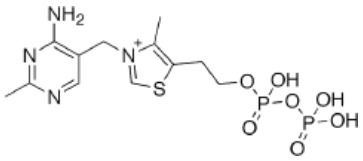
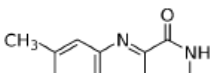
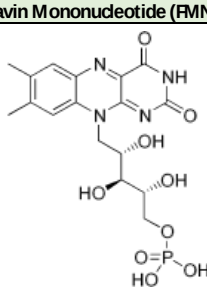
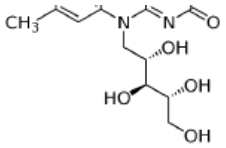
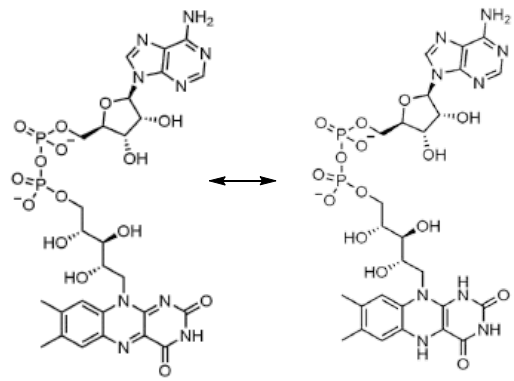
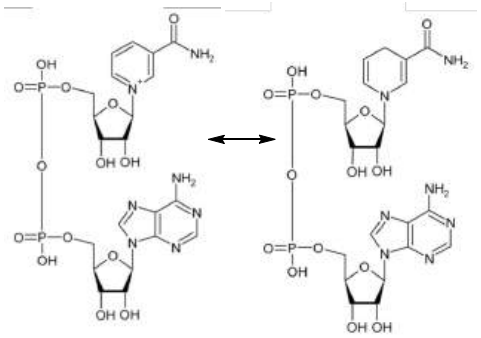
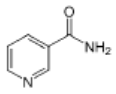
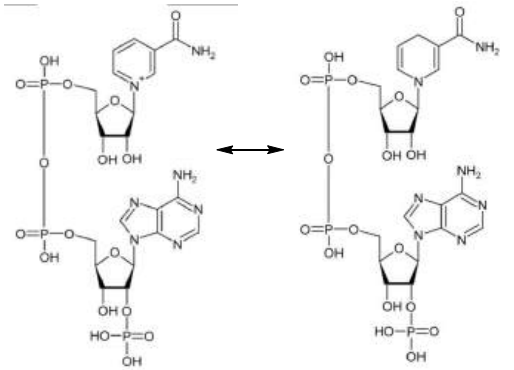
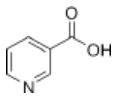
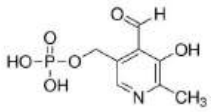
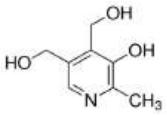
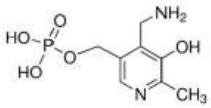
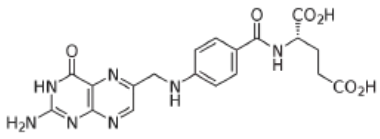
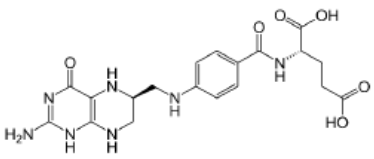


Figure 6.8.1: The Heme Cofactor. The family of heme cofactors contains an iron metal coordinated with a porphyrin ring structure, as shown in the *left-hand panel* within the structure of Heme B. In the right-hand panel, Heme B is shown complexed with the succinate dehydrogenase enzyme from the Kreb Cycle. The structure of Heme B shown in the left-hand panel is from Yikrazuul and the crystal structure of Succinate Dehydrogenase complexed with Heme B is from Richard Wheeler.

Many biological cofactors are vitamin B derivatives, as shown in Table 6.8.1 below. Many vitamin deficiencies cause disease states due to the inactivity of ***apoenzymes*** that can not function without the correctly bound ***coenzyme.***

B Vitamins	Modified Enzyme Cofactors
Vitamin B1 (Thiamine)	Thiamine Diphosphate (TPP)
	
Vitamin B2 (Riboflavin)	Flavin Mononucleotide (FMN)
	

	Flavin Adenine Dinucleotide (FAD <-> $FADH_2$) 
Vitamin B3 (Niacinamide) - amide form	Nicotinamide adenine dinucleotide (NAD <-> NADH) 
	
Vitamin B3 (Niacin) - carboxylic acid form	Nicotinamide dinucleotide phosphate (NADP <-> NADPH) 
	
Vitamin B6 (Pyridoxine)	Pyridoxal5'-phosphate 
	Pyridoxamine 5'-phosphate 
Vitamin B9 (Folic Acid)	Tetrahydrofolate
	
Vitamin B12 (Cyanocobalamin)	Adenosylcobalamin

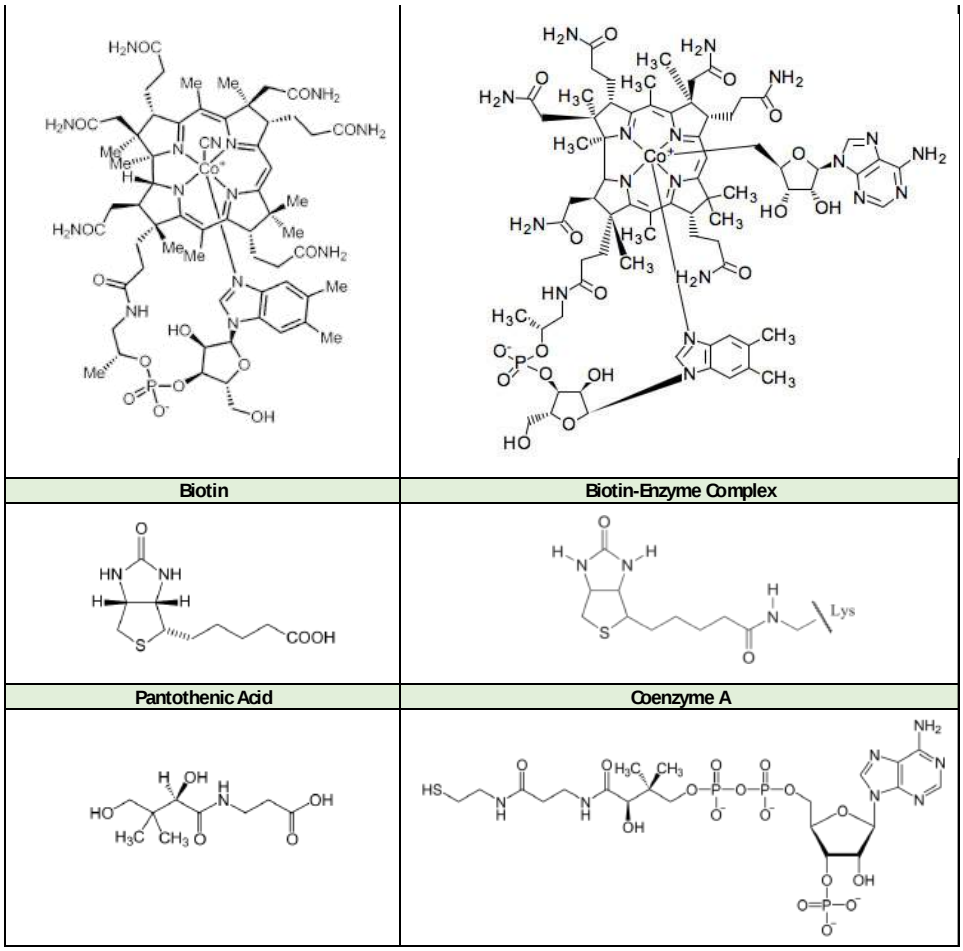


Table 6.8.1: Essential B-Vitamins and their Modified Enzyme Cofactors

Cofactors can help to mediate enzymatic reactions through the use of any of the different catalytic strategies listed above. They can serve as nucleophiles, mediate covalent catalysis, form electrostatic interactions with the substrate, and stabilize the transition state. They can also cause strain distortion or facilitate acid-base catalysis. Metal-aided catalysis can often use homolytic reaction mechanisms that involve radical intermediates. This can be important in reactions such as those occurring in the electron transport chain that requires the safe movement of single electrons.

We present plausible mechanisms for prototypical reactions using some of the cofactors shown in Table 6.8.1 above. Each shows the flow of electrons from a source to a sink. The source is often a pair of electrons on an anion, formed by the prior removal of a proton from the atom by a general base. A sink could be a carbonyl O, which receives a pair of electrons from one of the C=O bonds of the carbonyl. As a bond is made to the carbonyl, one of the double bonds must break with the electrons going (temporarily if the reaction is a nucleophilic substitution reaction) to the carbonyl O, an excellent sink since it is so electronegative. An even better sink is a positive N of an iminium ion; examples are shown below. Just the "business parts" of the cofactors are shown below.

To appreciate the mechanism used by cofactors, and show a clear example of an electron source/sink, let's look at a reaction that doesn't require a cofactor, the spontaneous decarboxylation of a β-keto acid as shown in Figure 6.8.1.

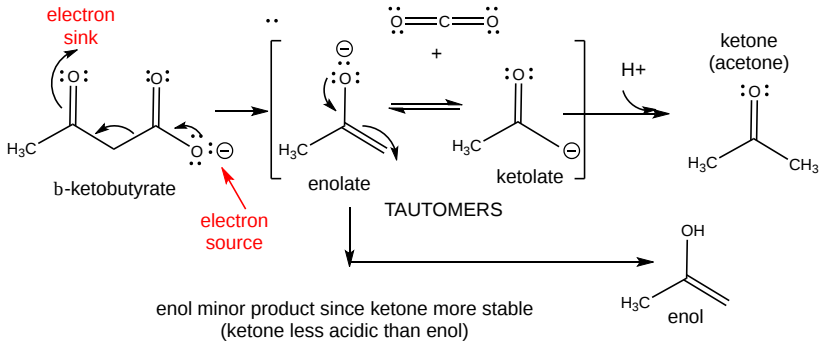


Figure 6.8.1:spontaneous decarboxylation of a β-keto acid

Even though no cofactor is required, nucleophilic catalysis by an amine through Schiff Base formation would speed up the reaction (as we will see below). Now let's look at how some of the cofactors listed in Table 6.8.1 above facilitate electron flow in reactions.

6.8.2: Thiamine pyrophosphate - decarboxylation of α-keto acids

Thiamine pyrophosphate (TPP) facilitates the decarboxylation of α-keto acids. TPP is a derivative of thiamine, vitamin B1, whose deficiency causes beriberi. TPP is covalently attached to the enzyme, such as in pyruvate dehydrogenase and alpha-ketoglutarate dehydrogenase, two enzymes that catalyze the decarboxylation of α-keto acids. The structure and "business" end of TPP and its catalytic activity are shown in Figure 6.8.2.

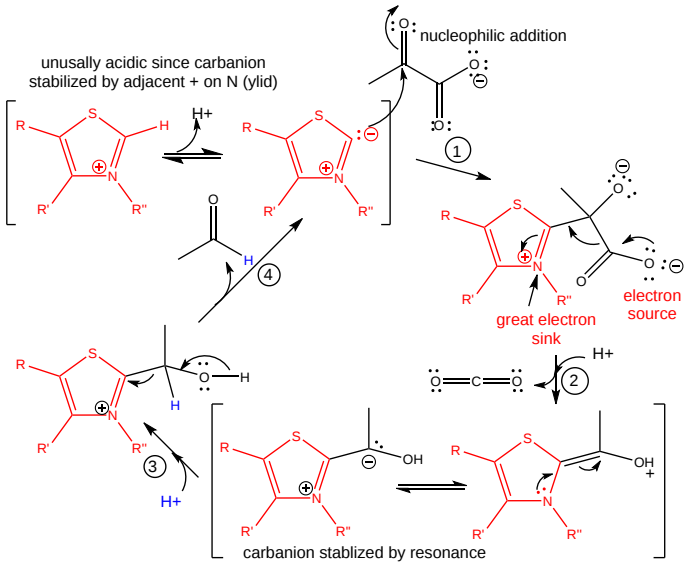


Figure 6.8.2: Role of TPP in the decarboxylation of pyruvate (step 1) and release of acetaldehyde (step 4)

The number of arrows leading to the product does not reflect the actual number of steps.

Figure 6.8.3 shows an interactive iCn3D model of the thiamin diphosphate-dependent enzyme pyruvate decarboxylase from the yeast Saccharomyces cerevisiae (1pvd).

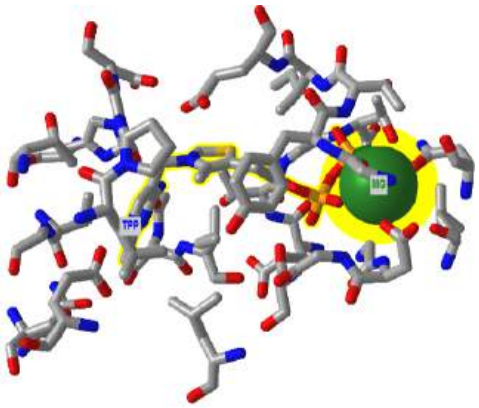


Figure 6.8.3: Thiamin diphosphate-dependent enzyme pyruvate decarboxylase from the yeast Saccharomyces cerevisiae (1pvd) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...cCq9EJrquwggj8

6.8.3: Flavin Adenine Dinucleotide (FAD) - hydride transfer

FAD and its reduced form, $FADH_2$, are tightly or covalently attached to an enzyme, so FAD must be regenerated in each catalytic cycle. Figure 6.8.4 shows an example of how this cofactor facilitates the transfer of a $:H^-$ hydride ion to the "business end" of FAD. In contrast to a transfer of protons (H^+), an acid/base reaction, hydride transfer removes 2 electrons from the substrate (in this case, succinate) along with a proton in an oxidation reaction as FAD is reduced.

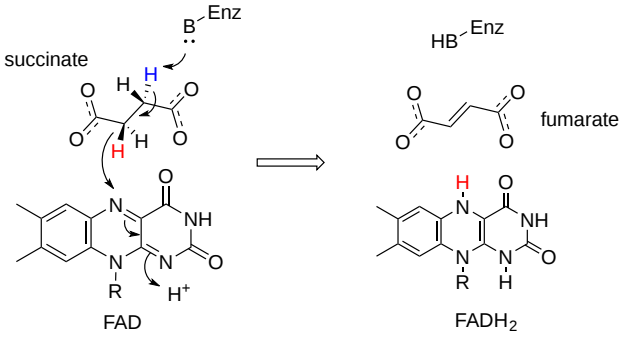


Figure 6.8.4: Oxidation of succinate by FAD

Figure 6.8.5 shows an interactive iCn3D model of the FAD-binding domain of cytochrome P450 BM3 from Priestia megaterium in complex with $NADP^+$ (4DQL)

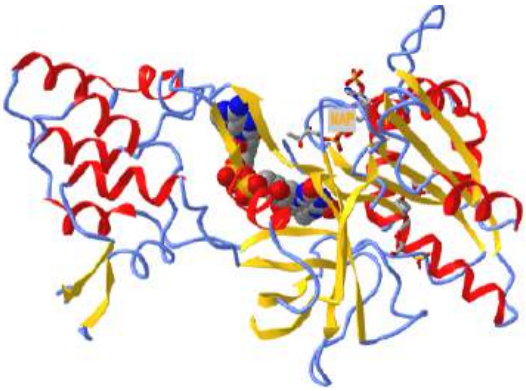


Figure 6.8.5: FAD binding domain of cytochrome P450 BM3 in complex with $NADP^+$ (4DQL). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?hoT1WDCUv1wZFMyRA

FAD is shown in spacefill. $NADP^+$, which reoxidizes the reduced $FADH_2$ back to FAD, is shown in sticks and labeled NAP.

6.8.4: Nicotinamide Adenine Dinucleotide (FAD) reactions

NAD^+ and a phosphorylated form, $NADP^+$, are one of nature's most widely used oxidizing agents and are used as dissociable substrates/cofactors for many different types of enzyme-catalyzed oxidation reactions. Since it binds (as a substrate) and dissociates (as a product) after each catalytic cycle, the free enzyme is continually active. The biological synthesis of NAD^+ requires the vitamin nicotinic acid, also called niacin (nicotinic acid), an absence of which causes pellagra.

Oxidation of an alcohol to an aldehyde: The oxidation of ethanol to acetaldehyde by NAD^+, catalyzed by the enzyme alcohol dehydrogenase, is shown in Figure 6.8.6.

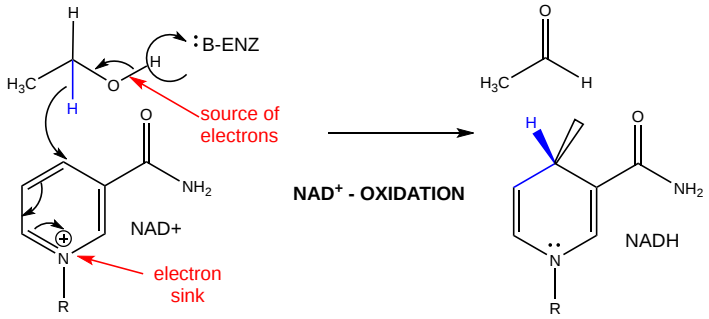


Figure 6.8.6: Oxidation of ethanol by NAD^+

The product acetaldehyde contributes to hangovers after ethanol consumption. Note that this reaction is a hydride transfer, which would not be expected to occur in the aqueous environment of a cell, given the extreme reactivity and basicity of a $:H^-$ hydride ion. This transfer happens in the active site of the enzyme, which is anhydrous after binding substrates.

Oxidative decarboxylation of an alcohol: A two-step mechanism for this reaction is shown in Figure 6.8.7

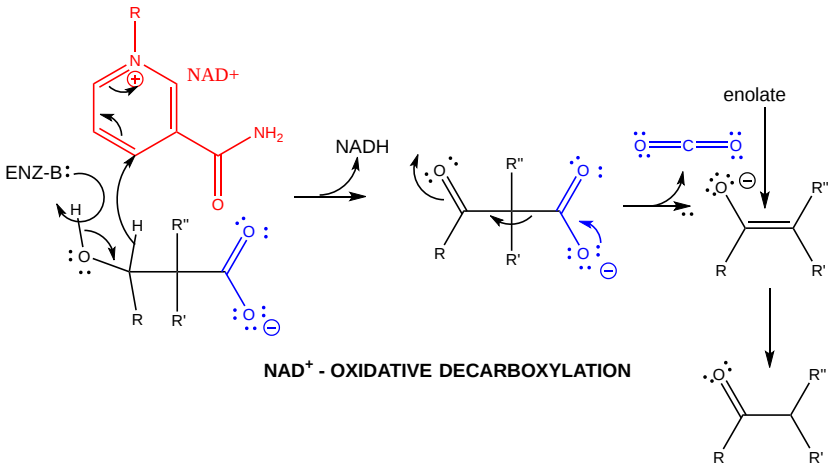


Figure 6.8.7: Oxidative decarboxylation of an alcohol

After the first step, an electron sink (the oxygen of the carbonyl) is present at the β-carbon, facilitating the decarboxylation step.

Oxidative deamination of an amine: A two-step reaction, a hydride transfer to form a Schiff base, followed by hydrolysis of the Schiff base, is shown in Figure 6.8.8.

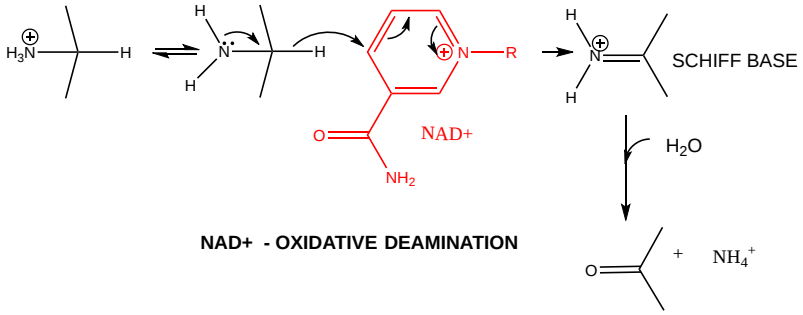


Figure 6.8.8: **Oxidative deamination of an amine**

We will discuss Schiff base chemistry in more detail below.

6.8.5: Pyridoxal Phosphate Enzymes

Pyridoxal phosphate (PLP) is a derivative of vitamin B6 or pyridoxal. Deficiencies cause convulsions, chronic anemia, and neuropathy. It assists in many reactions (catalyzed by PLP-dependent enzymes). The PLP is bound covalently to lysine residues in a Schiff base linkage (aldimine). This form reacts with many free amino acids (as substrates) to replace the Schiff base to Lys of the enzyme with a Schiff base to the amino acid substrate. First, we will review of Schiff base (an imine) formation by the reaction of an aldehyde or ketone with an amine as shown in Figure 6.8.9.

Schiff base (imine) formation

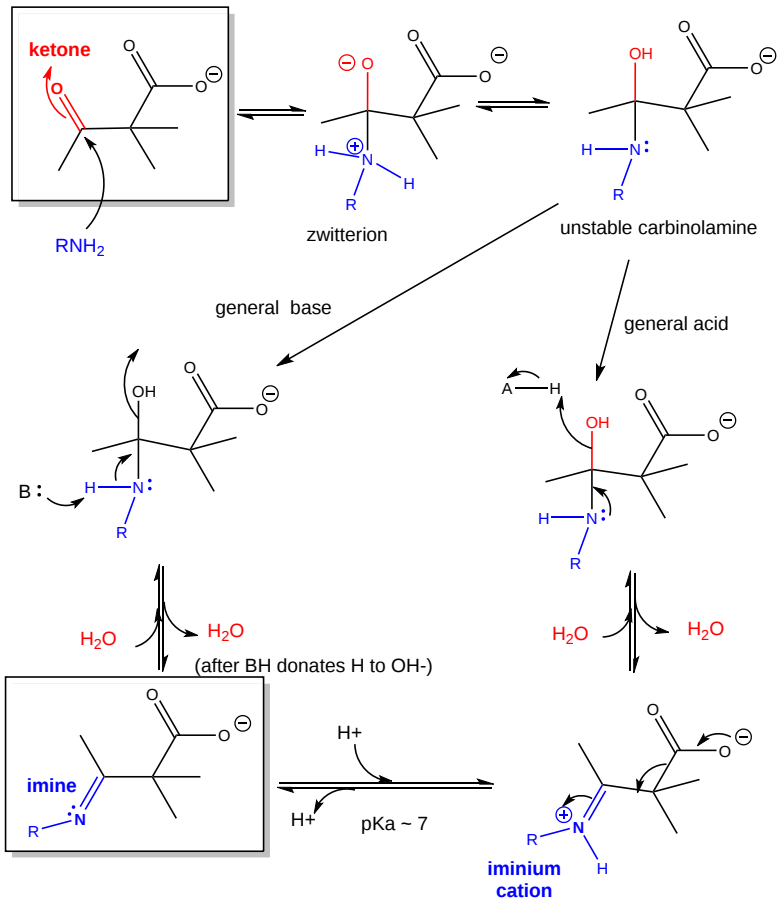


Figure 6.8.9

The reaction is essentially a nucleophilic attack of a carbonyl carbon of an aldehyde or ketone by an amine, followed by a dehydration step. Note that the net effect is to replace one electron sink, a carbonyl (C=O), with an **imine** (C=NH ↔ $C=NH_2^+$), with pKa around 7.0. Hence at neutral pH, 50% of the imine is protonated to form the **iminium cation**, a much better electron sink than the starting carbonyl!

The structure of pyridox**al** phosphate, which contains a reactive **al**dehyde, is converted to an imine by reaction with the ε-amino side chain of a lysine in the active site of a PLP-dependent enzyme, is shown in Figure 6.8.10.

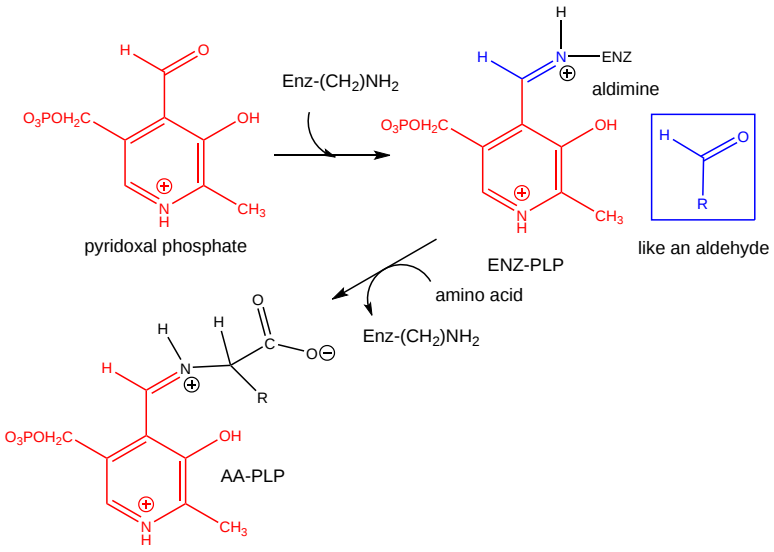


Figure 6.8.10

The figure also shows the replacement of the enzymes' lysine ε-NH_2-PLP bond to that of a free amino as an incoming substrate, a process which should proceed with a ΔG^0 of approximately 0. This occurs in PLP-dependent enzymes with free amino acids as substrates (we will discuss several examples below).

Figure 6.8.11 shows an interactive iCn3D model of the E. Coli Aspartate aminotransferase, W140H mutant, maleate complex (1ARI).

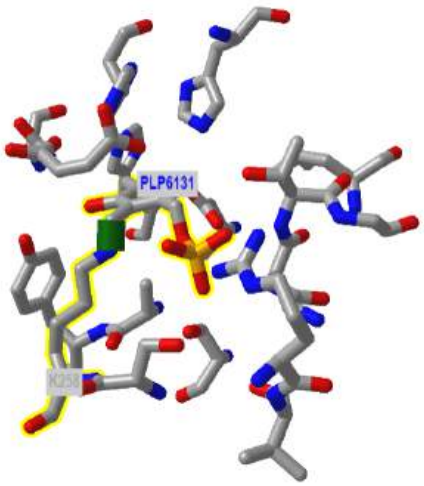


Figure 6.8.11: **Aspartate aminotransferase, W140H mutant, maleate complex** (1ARI). (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...ZQixXhCwX5EEQ9

Note that the PLP is in Schiff base linkage with the ε-NH_2 group of a lysine in the enzyme's active site.

PLP is quite impressive!

From a chemistry perspective, PLP is an ideal molecule to facilitate electron flow in biochemical reactions. William Jencks noted this in his classic text, Catalysis in Chemistry, in which he wrote this elegant description of its properties:

"It has been said that God created an organism especially adapted to help the biologist find an answer to every question about the physiology of living systems; if this is so, it must be concluded that pyridoxal phosphate was created to provide satisfaction and enlightenment to those enzymologists and chemists who enjoy pushing electrons, for no other coenzyme is involved in such a wide variety of reactions, in both enzyme and model systems, which can be reasonably interpreted in terms of the

chemical properties of the coenzyme. Most of these reactions are made possible by a common structural feature. That is, electron withdrawal toward the cationic nitrogen atom of the imine and into the electron sink of the pyridoxal ring from the alpha carbon atom of the attached amino acid activates all three of the substituents of this carbon for reactions which require electron withdrawal from this atom."

We'll present three examples of the reaction of an amino acid with a PLP-dependent enzyme. In each case, a different bond to the α-carbon of the amino acid substrate is broken.

α-decarboxylation of an amino acid: Figure 6.8.12 shows a plausible reaction mechanism.

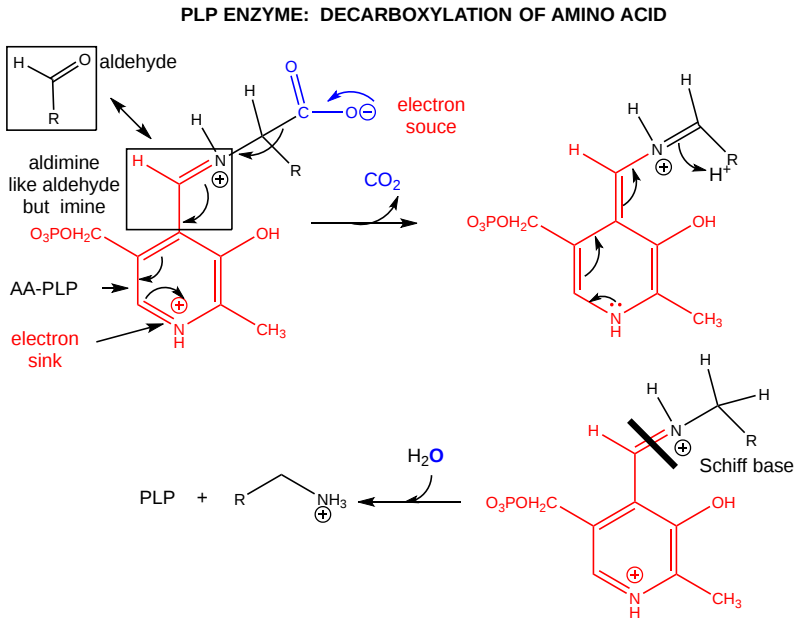


Figure 6.8.12: PLP-dependent decarboxylation of an amino acid.

β-elimination from serine: The enzyme serine dehydratase catalyzes the reaction shown in Figure 6.8.13

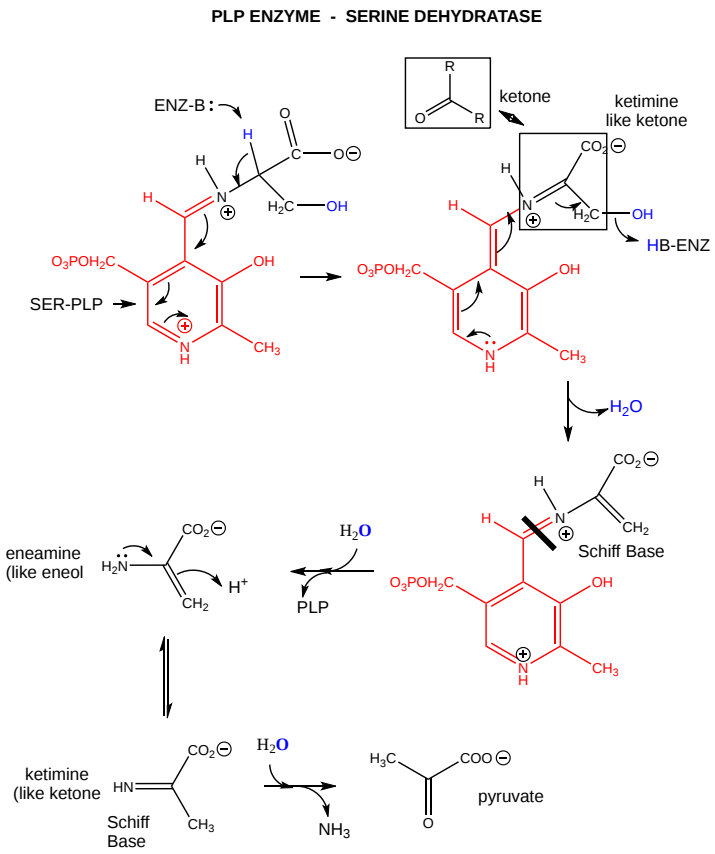


Figure 6.8.13: PLP-dependent β-elimination from serine

Racemization of amino acids: Amino acid racemases use PLP as a cofactor using a mechanism shown in Figure 6.8.14

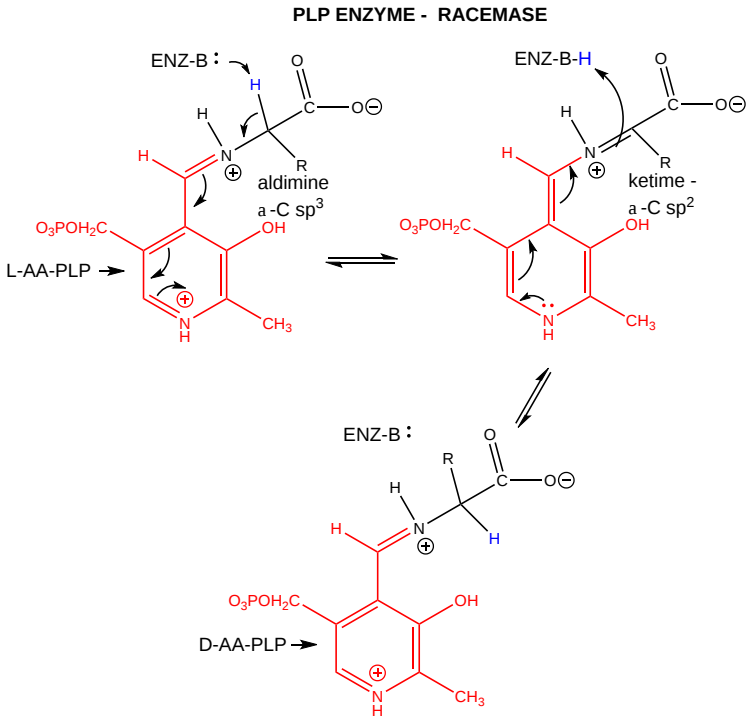


Figure 6.8.14: PLP-dependent racemization of amino acids

Why do racemases exist since the biological world consists of only L-amino acids? There are two possible reasons. Some D-amino acids are found, such as in bacterial cell walls. In addition, amino acids spontaneously racemize on their own, albeit at a slow rate.

Racemases that have oxygen atoms in the beta-carbon racemize at a much higher rate since they can stabilize the carbanion intermediate formed when the alpha proton is removed in the process of racemization. The concentration of D-Asp and D-Asn can be used in dating biological material as well.

Transamination reactions: PLP enzymes also catalyze the transamination reaction, which is shown in Figure 6.8.15

Amino Acid 1 + α-keto acid 1 ↔ α-keto acid 2 + Amino Acid 2 For example: Figure 6.8.x

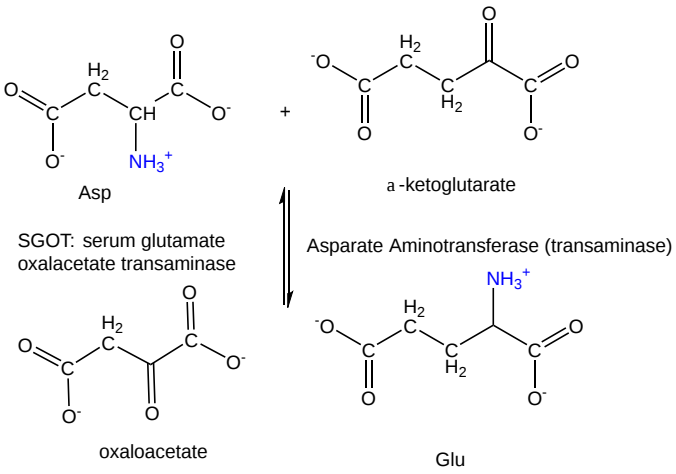


Figure 6.8.15: PLP-dependent transamination reaction

First, Asp, bound to PLP through a Schiff base link, loses the α-H and forms a ketimine through a tautomerization reaction, which ultimately hydrolyzes to form the released oxaloacetate and pyridoxamine. The pyridoxamine reacts with α-ketoglutarate in the reverse of the first three reactions to form Glu.

We will explore other cofactors in future chapters.

CHAPTER OVERVIEW

7: Carbohydrates and Glycobiology

7.1: Monosaccharides and Disaccharides
7.2: Polysaccharides
7.3: Glycoconjugates - Proteoglycans, Glycoproteins, Glycolipids and Cell Walls
7.4: The Sugar Code and Lectin Decoding
7.5: Working with Carbohydrates
7.6: Chapter 7 Problems - Answer Key
7.7: Chapter 7 Problems

7.1: Monosaccharides and Disaccharides

7.1.1: Introduction

Carbohydrate or glycan biochemistry is very complex and challenging owing to the stereochemical complexity of simple sugars, a large number of positions on the sugars used to form linkages between other sugars to create polymers, the large number of chemical modifications to base sugars, and the lack of a genetic template to instruct glycan polymer formation. It is no wonder that our understanding of complex glycans has developed after that of the chemically simpler polymers like nucleic acids and proteins.

In addition, the terminology used to describe them varies as well. We use these general descriptions of them:

Sugar: usually refers to low molecular weight carbohydrates like glucose, lactose, and sucrose, but it can also refer broadly to any carbohydrate.

Carbohydrate: a general term that applies to simple sugars to complex sugar polymers like glycogen, starch, and cellulose. The name derives from the formula for simple sugars like glucose ($C_6H_{12}O_6$), which can be written as $C_6(H_2O)_6$ - a carbo (C) - hydrate (H_2O).

Glycan: A general term for molecules containing simple sugars and sugar derivatives linked in a polymer, either standalone molecules or attached to other molecules like proteins.

7.1.2: Monosaccharides Structures

The above definition of sugar needs some further nuance. From a chemical perspective, sugars can be defined as polyhydroxy-aldehydes or ketones. The simplest sugars contain at least three carbon atoms, and the most common are the aldo- and keto-trioses, tetroses, pentoses, and hexoses. The 3C sugars are glyceraldehyde and dihydroxyacetone, as shown in Figure 7.1.1.

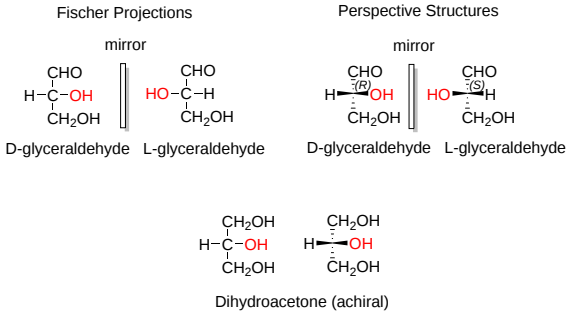


Figure 7.1.1: Three-carbon sugars

Glucose, an aldohexose, is a central sugar in metabolism. It and other 5C and 6C sugars can cyclize through intramolecular nucleophilic attack of one of the free hydroxyl groups on the carbonyl carbon of the aldehyde or ketone. Such intramolecular reactions occur if stable 5- or 6-member rings can form. The resulting rings are labeled furanose (5-member) or pyranose (6-member) based on their similarity to furan and pyran. On nucleophilic attack to form the ring, the carbonyl O becomes an OH that points either below (α anomer) or above (β anomer) the ring.

Figure 7.1.2 shows different representations of the linear and cyclic forms of the sugars D-glucose, D-ribose, and D-fructose

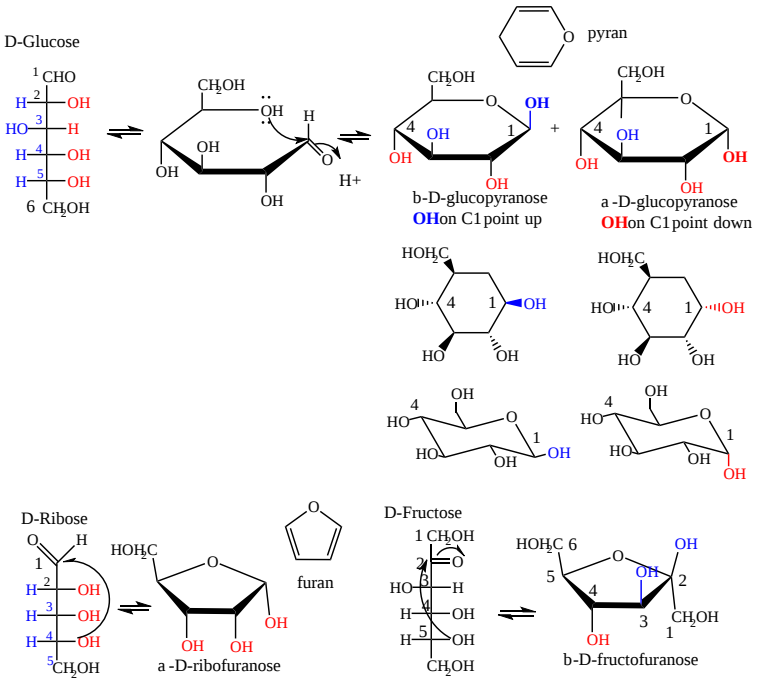


Figure 7.1.2: Linear and cyclic forms of D-glucose, D-ribose and D-fructose

Monosaccharides in solution exist as equilibrium mixtures of the straight and cyclic forms. In solution, glucose (Glc) is mostly in the pyranose form, fructose is 67% pyranose and 33% furanose, and ribose is 75% furanose and 25% pyranose. However, in polysaccharides, Glc is exclusively pyranose and fructose and ribose are furanoses.

Sugars can be drawn in the straight chain form as either Fischer projections or perspective structural formulas.

In the Fisher projection, the vertical bonds point down into the plane of the paper. That's easy to visualize for 3C sugars but more complicated for larger ones. For those, draw a wedge and dash line drawing of the molecule. When determining the orientation of the OHs on each C, orient the wedge and dash drawing in your mind so that the C atoms adjacent to the one of interest are pointing down. Sighting towards the carbonyl C, if the OH is pointing to the right in the Fisher project, it should be pointing to the right in the wedge and dash drawing, as shown below for D-erythrose and D-glucose. Figure 7.1.3 shows how to convert Fisher projections to wedge dash representations.

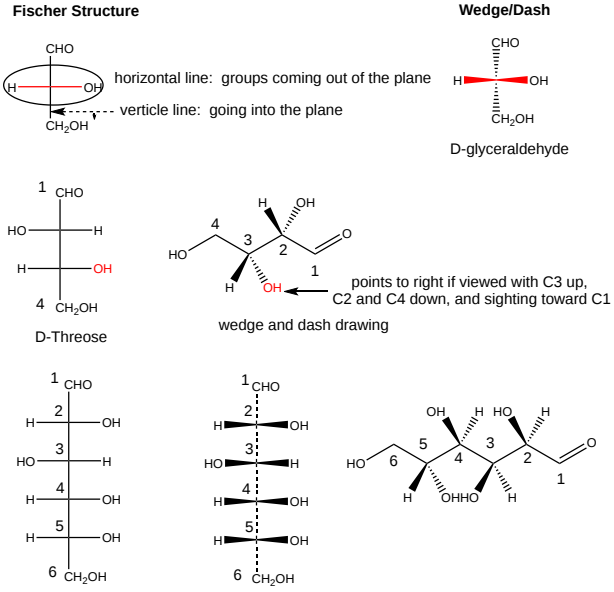


Figure 7.1.3: Converting Fisher projections to wedge dash representations.

Figure 7.1.4 shows an interactive iCn3D model of D-glucose in a linear form.

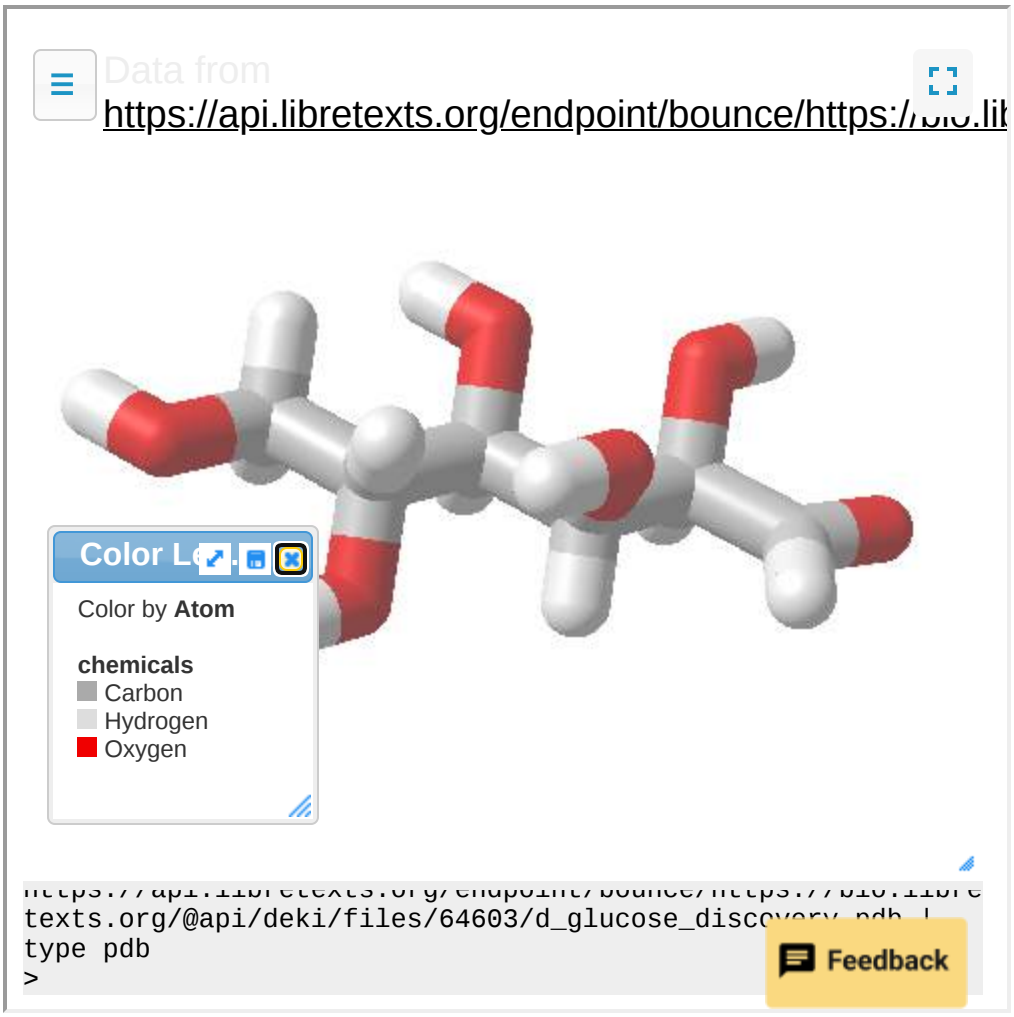


Figure 7.1.4: D-glucose (Copyright; author via source).

Orient the molecule as shown in Figure 7.1.5 below, with the carbonyl oxygen pointed to the far right, and compare it to the orientation shown in Figure 7.1.5 to reinforce your understanding of Fisher and wedge/dash projections.

Figure 7.1.5:

Cyclic forms can be drawn either as the Haworth projections, which show the molecule as cyclic and planar with substituents above or below the ring) or the more plausible bent forms (showing glucose in the chair or boat conformations, for example). β-D-glucopyranose is the only aldohexose that can be drawn with all its bulky substituents (OH and CH_2OH) in equatorial positions, which probably accounts for its widespread prevalence in nature. Figure 7.1.6 shows four different representations of glucose.

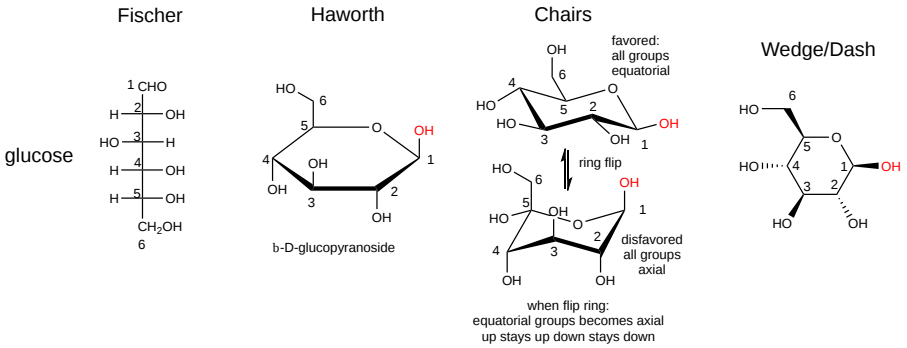


Figure 7.1.6: Fisher and cyclic Haworth, chair and wedge/dash representations of glucose

Haworth projections are more realistic than the Fisher projections, but you should be able to draw both structures. Generally, if a substituent points to the right in the Fisher structure, it points down in the Haworth. If it points left, it points up. Generally, the OH on the **α**-anomer points down (**α**nts down) while on the **β**-anomer, it points up (**β**utterflies up) as illustrated in Figure 7.1.7

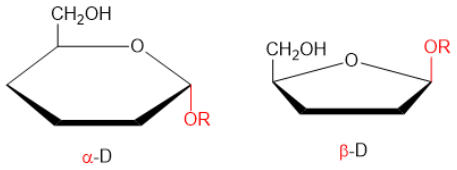


Figure 7.1.7: Alpha and beta Haworth representations of sugars

In the Haworth projections, the bulky R group of the next carbon after the carbon whose OH group was the nucleophile for ring formation is pointed up if the OH engaged in the attach was on the right-hand side in the straight chain Fisher diagram (as in α-D-glucopyranose above when the CH_2OH group is up). It is pointed down if the OH engaged in the attack was on the left-hand side in the straight chain Fisher diagram (as in α-D-galactofuranose above when the $(CHOH)CH_2OH$ group is down). The rest of the OH groups still follow the simple rule that if they point to the right in the Fisher straight chain form, they point down in the Haworth form.

The Fisher structures of most common monosaccharides (other than glyceraldehyde and dihydroxyacetone), which you will encounter most frequently, are shown in Figure 7.1.8.

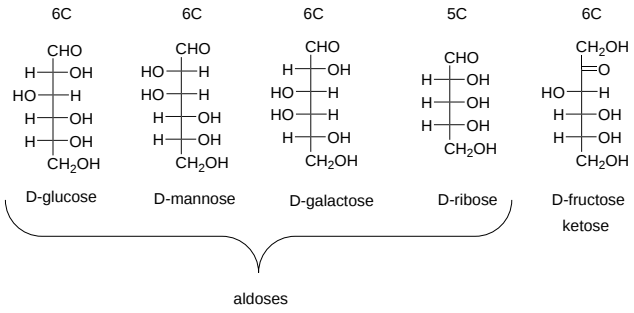


Figure 7.1.8: Most common monosaccharides discussed in this book

The mirror image of D-Glc is L-Glc. The D- and L- designations refer to the center of asymmetry most remote from the aldehyde or ketone. By convention, all chiral centers are related to D-glyceraldehyde, so sugar isomers related to D-glyceraldehyde at their last asymmetric center are D sugars.

Figure 7.1.9 shows multiple renderings of common hexoses.

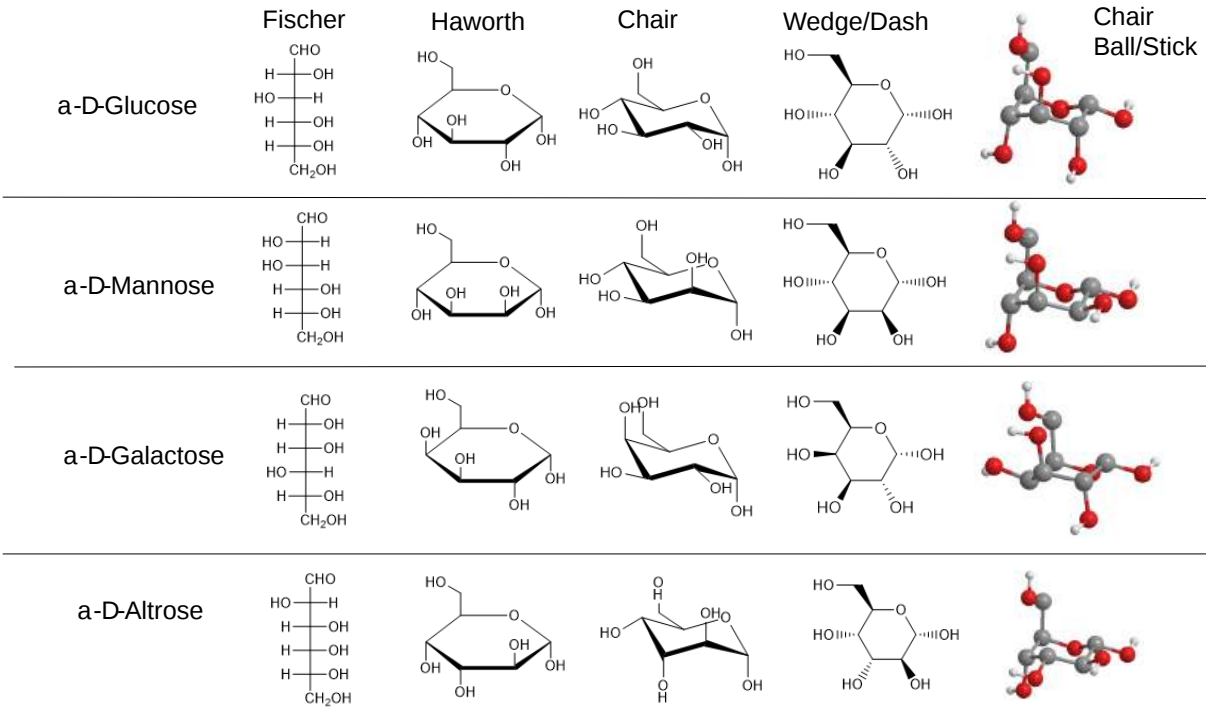


Figure 7.1.9: Multiple renderings of common hexoses

7.1.3: Isomers

Sugars can be configurational (interconverted only by breaking covalent bonds) or conformational isomers. Figure 7.1.10 reviews different configurational isomers.

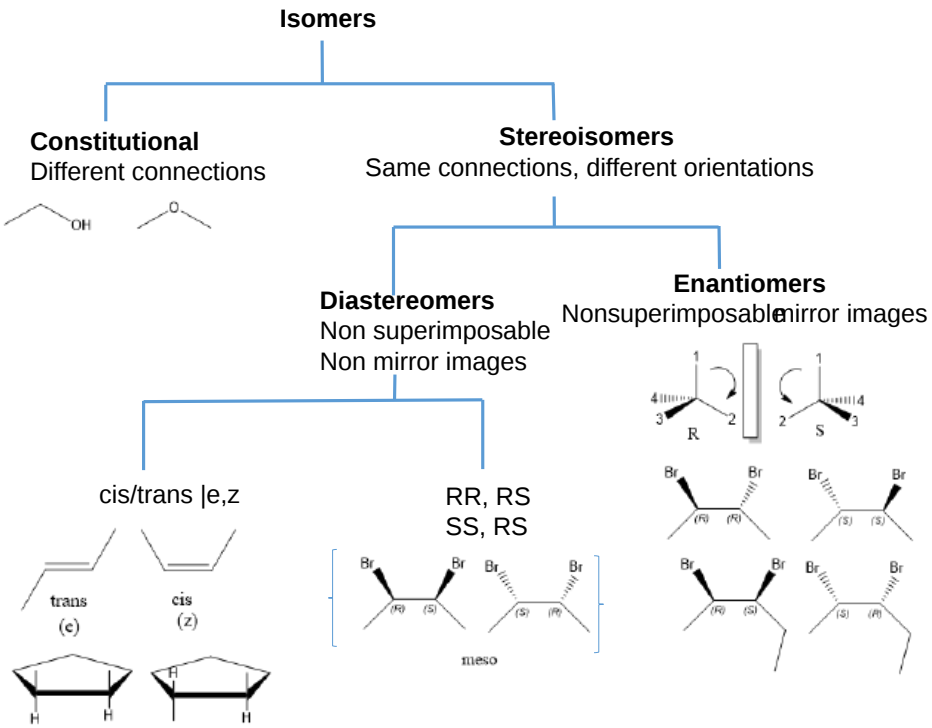


Figure 7.1.10: Types of isomers

The configurational isomers include enantiomers (stereoisomers that are mirror images of each other), diastereomers (stereoisomers that are not mirror images), **epimers** (diastereomers that differ at one stereocenter), and **anomers** (a particular form of stereoisomer, diastereomer, and epimer) shows enantiomers, diastereomers, epimers and anomers of 6 carbon sugars.

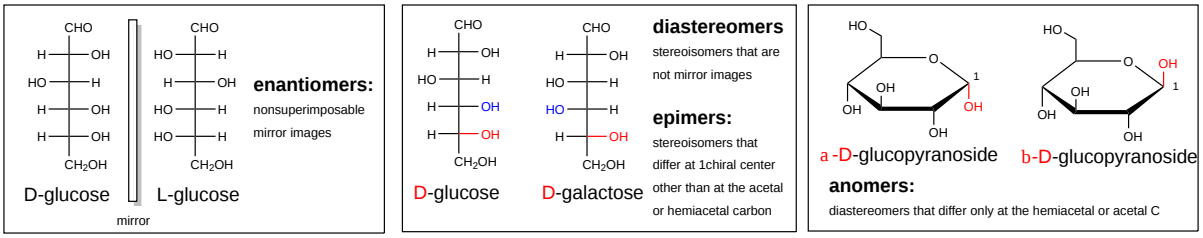


Figure 7.1.11: Enantiomers, diastereomers, epimers and anomers of 6 carbon sugars.

Sugars can also exist as conformational isomers, which interchange without breaking covalent bonds. These include chair and boat conformations of the cyclic sugars as shown in Figure 7.1.12

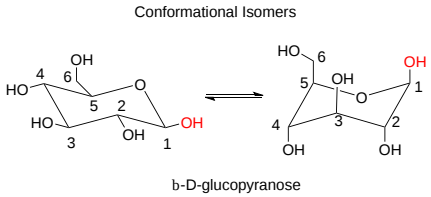


Figure 7.1.12: Conformational isomers of b-D-glucopyranose

7.1.4: Monosaccharide Derivatives

Many derivatives of monosaccharides are found in nature. These include

- oxidized forms in which the aldehyde and/or alcohol functional groups are oxidized to carboxylic acids
- phosphorylated forms in which phosphates are transferred from ATP to form phosphoester derivatives
- amine derivatives such as glucosamine or galactosamine
- acetylated amine derivatives such as N-Acetyl-GlcNAc (GlcNAc) or GalNAc
- lactone forms (intramolecular esters) in which an OH group attacks a carbonyl C that was previously oxidized to a carboxylic acid
- condensation products of sugar derivatives with lactate ($CH_3CHOHCO_2^-$) and pyruvate ($CH_3COCO_2^-$), both from the glycolytic pathway, to form muramic acid and neuraminic acids (also called sialic acids), respectively.

Figure 7.1.13some simple monosaccharide derivatives.

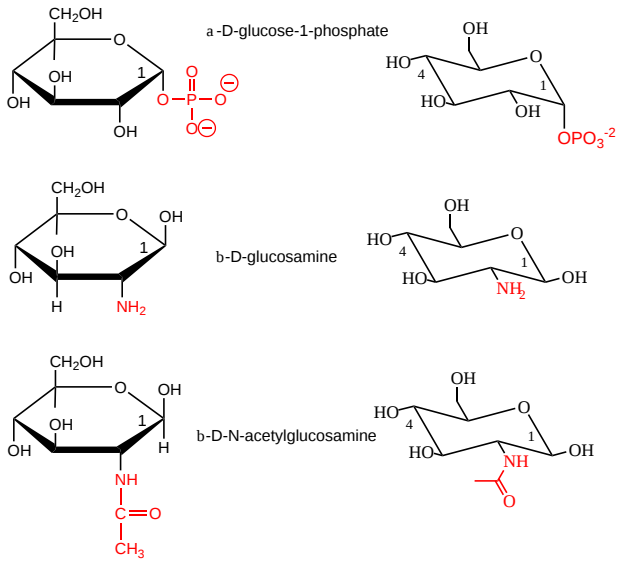


Figure 7.1.13: Monosaccharide derivatives

Figure 7.1.14 shows some additional oxidative derivatives of glucose shown in Fischer projections.

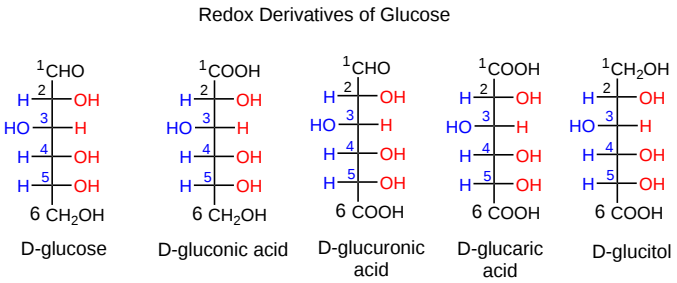


Figure 7.1.14: Redox derivatives of glucose

Other important derivatives of monosaccharides are **sialic acids**. **N-acetylmuramic acid**, found in bacterial cell walls, consists of GlcNAc in ether link at C3 with lactate, while **N-acetylneuraminic acid** results from an intramolecular cyclization of a condensation product of ManNAc and pyruvate. These sialic acids are shown in Figure 7.1.15.

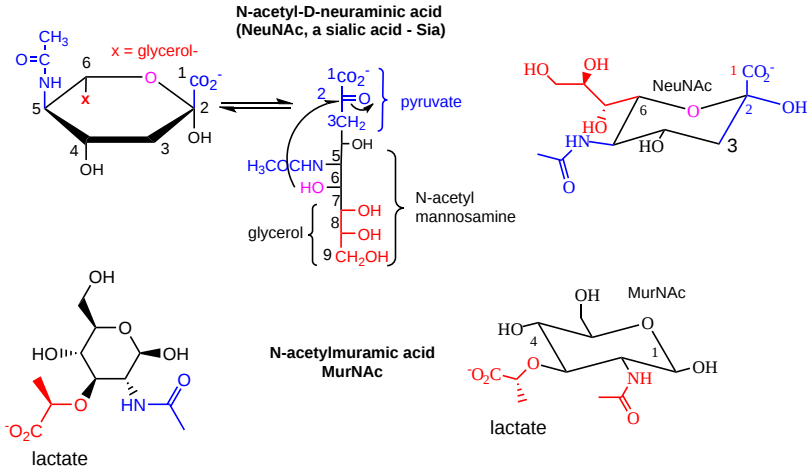


Figure 7.1.15: Sialic Acids

Sugars are very complicated as the linkages and substituents are so diverse. Figure 7.1.16 show differences in sialic acids between humans and chimps.

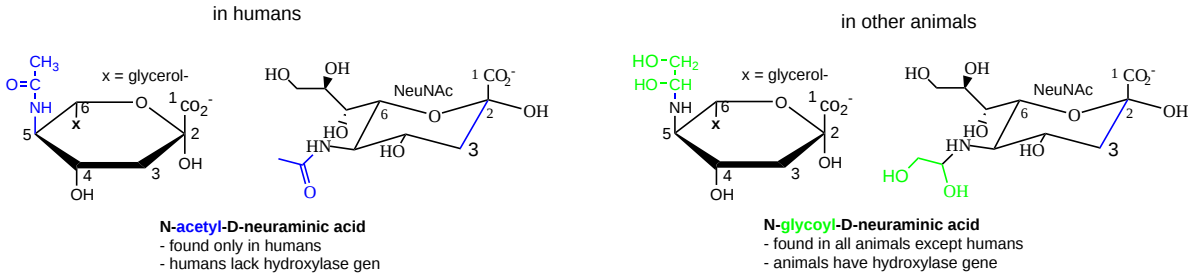


Figure 7.1.16: Sialic acids between humans and chimps.

What happens when non-vegan humans eat animal products (meat, milk) with N-glycoyl neuraminic acids (Neu5Gc)? Some get incorporated into human membrane glycans. Sialic acids on surface proteins can serve as "receptors" that allow binding of self-cells as well as foreign cells or proteins that have evolved to bind them. A toxin, SubAB, secreted by E. Coli 0157, can bind Neu5Gc. Hence eating meat products can make us more susceptible to bacteria that recognize Neu5Gc.

7.1.5: Formation of Hemiacetals, Acetals, and Disaccharides

Monosaccharides that contain aldehydes can cyclize through an intramolecular nucleophilic attack of an OH at the carbonyl carbon in an addition reaction to form a **hemiacetal.** In the past the group was called a **hemiketal** if the attack was on a ketone but now they are also called hemiacetals. On the addition of acid (which protonates the anomeric OH, forming water as a potential leaving group), another alcohol can add, forming an **acetal** with water leaving. These reactions are shown in Figure 7.1.17.

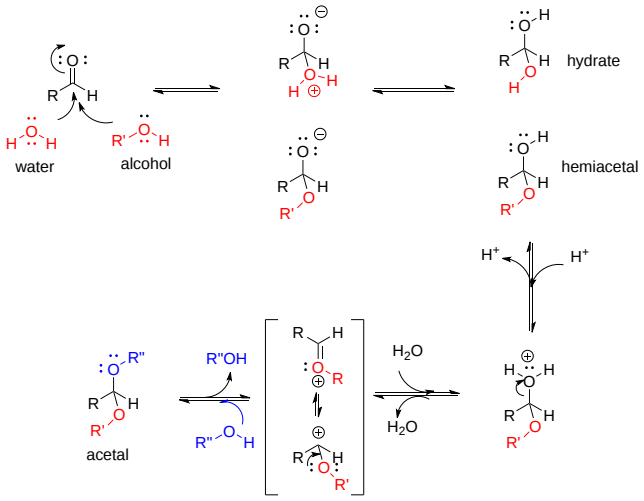


Figure 7.1.17: Hemiacetal and acetal formation

If the other alcohol is a second monosaccharide, a disaccharide results. The acetal link bonding to the two monosaccharides is called a **glycosidic link**. If the link between the two sugars involves an anomeric carbon, the newly formed OH group at the link can be designated either as α (if the O on the carbon involved in the glycosidic link is pointing down) or β (if the O is pointing up). For a 2-2 link between hexoses (i.e. between two non-anomeric carbons, the α/β designation is not used. Since sugars contain so many OH groups, which can act as the "second" alcohol in acetal formation, links between sugars can be quite diverse. These include α and β forms of 1-2, 1-3, 1-4, 1-5, 1-6, etc. links. Here are examples of disaccharides:

- maltose: Glc(α 1,4)Glc, which can be considered a disaccharide hydrolysis product of the polysaccharide glycogen or starch (discussed in the section)
- cellobiose: Glc(Glc(α 1,4)Glc 1,4)Glc, which can be considered a disaccharide hydrolysis product of cellulose.
- lactose: Gal(β 1,4)Glc, also known as milk sugar.
- sucrose: Glc(α 1,2)Fru. Since fructose is attached through the anomeric OH of this ketose, the fructose is not in equilibrium with its straight-chain keto form; hence, sucrose is a nonreducing sugar. Note also since the anomeric C-OH or each sugar is used, the α/β designation in the dissacharide is used. Hence sucrose would be abbreviated as Glc(α1,2β)Fru

The differences between lactose and sucrose are illustrated in Figure 7.1.18. Note that the β-D-fructofuranose ring is flipped (left to right as in turning one of your hands over) compared to Figure 7.1.16.

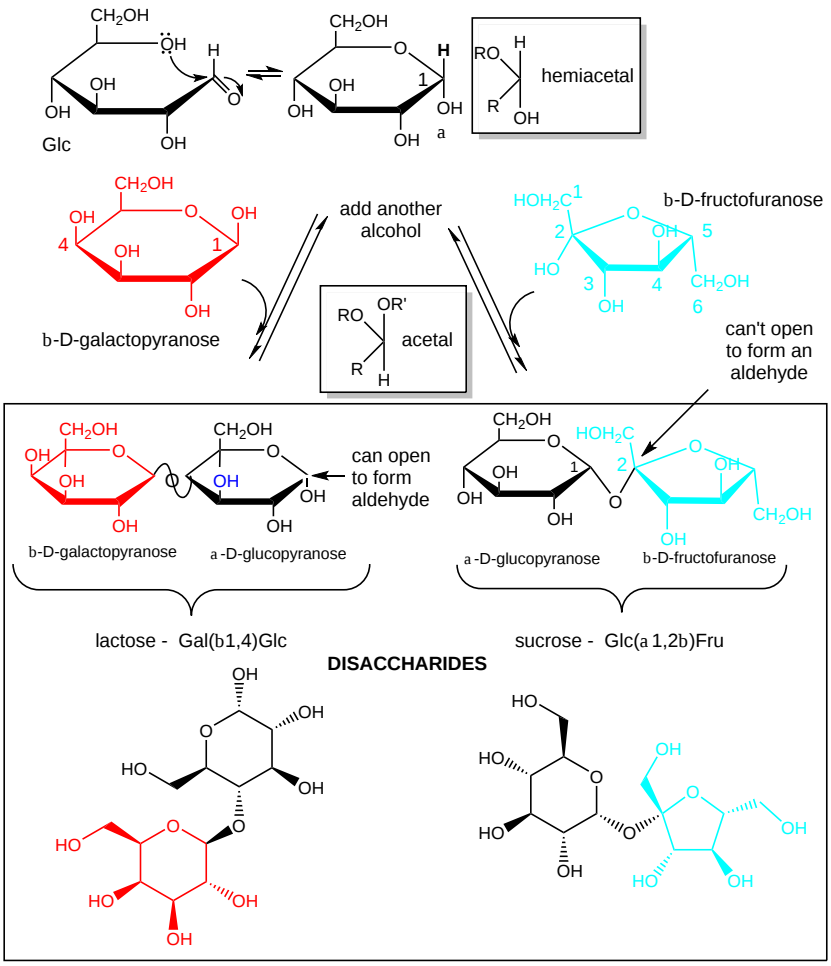


Figure 7.1.18: Structures of lactose and fructose

Acetal links between sugars in glycans can be hydrolyzed by water (catalyzed by H^+), just as with the other key biological polymers, proteins and nucleic acids.

The disaccharides described above that are linked through a 1,4 linkage are called reducing sugars since they can act as reducing agents in reactions in which they get oxidized. For example, in lactose, since galactose is attached to glucose through the OH on C4, the anomeric glucose carbon, C1, could revert to the noncyclic aldehyde form. This aldehyde is susceptible to oxidation by reagents (Benedict's Solution - with citrate, Fehling's reagent - with tartrate) as these reagents are subsequently reduced. In both reagents, reducing sugars reduce a basic blue solution of $CuSO_4$ (Cu^{2+}) to a brick-red precipitate of Cu_2O (Cu+). Sugars (monosaccharides, disaccharides and polysaccharides) that have a potentially open aldehyde at C1 or have an α-hydroxymethyl ketone group which can isomerize to an aldehyde under basic conditions (such as fructose) are called **reducing sugars**. These oxidizing agents are mild and react with aldehydes and not ketones.

If a monosaccharide, disaccharide, or even polysaccharide has a least one hemiacetal link (for instance the second sugar in lactose), it is a reducing sugar, as the monomer with the cyclic hemiacetal can reversibly open to form an aldehyde. However, if the only links in sugars are full acetals (such as in sucrose when the link is between the two anomeric carbons), the sugar is not reducing.

Alpha-gal syndrome

Alpha-gal syndrome (AGS) is a relatively newly discovered disease caused by the bite of a tick. Tick saliva contains the disaccharide galactose-α-1,3-galactose (alpha-gal). After a tick bite, people develop an immune response to the disaccharide in the form of IgE antibodies. Further bites could cause a mild rash up to an anaphylactic response.

What makes AGS worse is that red meat also contains the disaccharide but is not found in fish, birds, or people. Hence people who mount a strong IgE response to the disaccharide can also elicit the same response when they eat red meat or even if they drink cow's milk, for example. Estimates that up to 450,000 people in the US may develop serious and even life-threatening symptoms after eating red meat.

The structure of Gal(α1,3)Gal is shown in Figure 7.1.19 below.

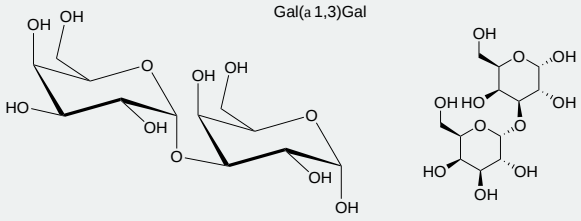


Figure 7.1.19: The structure of the disaccharide Gal(α1,3)Gal

7.2: Polysaccharides

Polysaccharides contain many monosaccharides in glycosidic links and may have many branches. They serve as either structural components or energy storage molecules. Polysaccharides consisting of single monosaccharides are homopolymers. The most common are starch, glycogen, dextran, cellulose, and chitin. We'll discuss based on whether the acetal link is alpha or beta.

7.2.1: α 1,4 main chain links

Starch and **Glycogen**: These polysaccharides are polymers of glucose linked in α 1,4 links with α 1,6 branches. **Starch**, found in plants, is subdivided into **amylose**, which has no branches, and **amylopectin,** which does. Starch granules consist of about 20% amylose and 80% amylopectin. **Glycogen,** the main CHO storage in animals, is found in muscle and liver, and consists of glucose residues in α 1,4 links with lots of α 1,6 branches (many more branches than in starch).

Here are various ways to render in 2D the chemical structure of a branched glycogen and starch fragment, as shown in Figure 7.2.1.

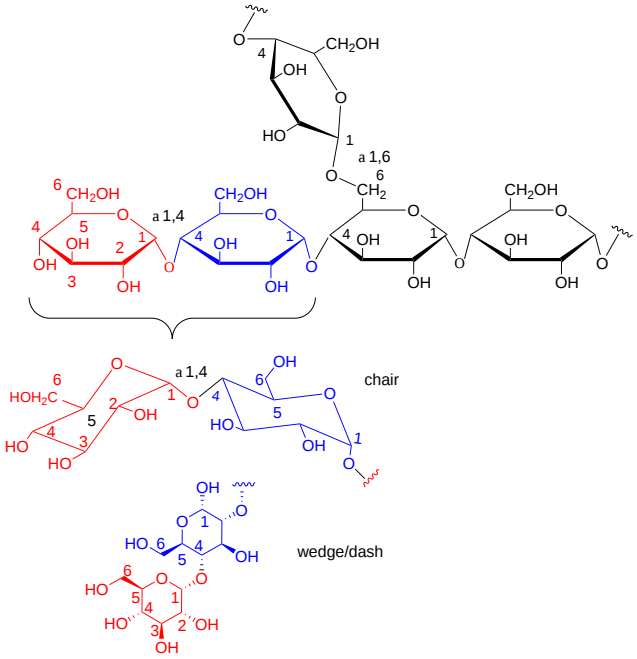


Figure 7.2.1: Structure of a branched glycogen and starch fragment

The top part of the figure shows the Haworth structure. The bottom part shows two glucose units in red and blue in the more structurally clear chair and wedge/dash representations.

Figure 7.2.2 shows an interactive iCn3D model of 10 glucose monosaccharides in an α-(1,4) linkage with five glucose units with α-(1,4) linkages attached to the main chain through α-(1,6) branch at glucose 6 of the main chain. The type of substructure would be found in starch (amylopectin) and glycogen.

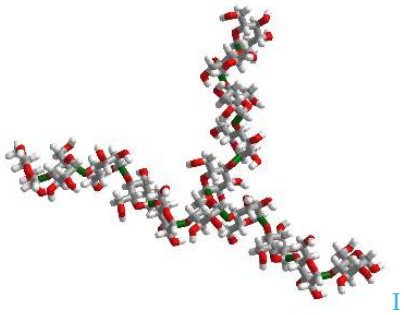


Figure 7.2.2: α-(1,4) linked glucose with an α-(1,6) branch (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...TXXy589xPNgmG9

Figure 7.2.3 the structure in the iCn3D model in a diagrammatic fashion in which glucose is represented as a blue circle with the acetal/glycosidic/glucosidic linkages between the monosaccharides written between the circles. The 14A label shows that the acetal linkage is an α-(1,4) link with a single α-(1,6) branch.

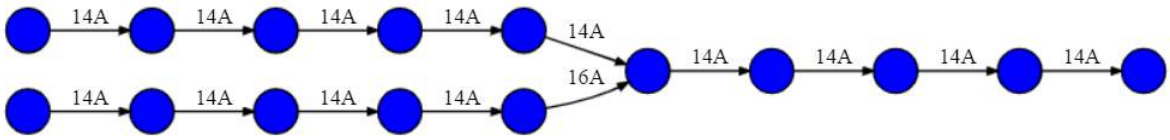


Figure 7.2.3: Symbolic Nomenclature For Glycans (SNFG) for an α-(1,4) linked glucose polymer with an α-(1,6) branch.

The linkages are written in a variety of conventions. These include 14A, 14α, 4A and 4α. The between many sugars is often a 1,x link where x is 2,3, 4, 5 or 6. In those cases the 1 can be omitted. The program used to generate images in this text uses both numbers and A or B.

What makes carbohydrates so complex is their 3D structures. Like proteins and nucleic acids, they can adopt a myriad of conformations. As the monomeric units are so homogeneous, especially in homopolymers, it isn't easy to get crystal structures for them so computer models are often used.

Studies have shown that the simple starch fraction amylose α 1,4 polymer of glucose, often envisioned as a straight chain, can adopt three main conformations. They are double-helical A- (found chiefly in cereals), double helical B-(found primarily on tubers) amyloses and single-helical V-amylose (or simply A, B and V structures). The A and B do **NOT** represent alpha or beta in this classification system. The A and B forms consist of double helices aligned in a parallel fashion with about 6 glucoses per turn. The helices appear to be left or right-handed, and this ambiguity might arise from a lack of crystal structures.

In contrast, a well-defined structure of the V helix is known. It folds into a left-handed helix with 6 glucoses per turn and a pitch of about 8Å. Unlike alpha helices of proteins and the double-stranded helix of DNA, the center of the helix is **NOT** packed tightly and can accommodate small molecules. One is iodine (actually triiodide, I_3^-), which, when bound in amylose with a sample of starch, exhibits a dark blue color. This is the basis for starch indicators that you may have used in titration reactions in biology and chemistry courses.

In proteins, alpha helices might self-associate during folding to form a 4-helix bundle. Likewise, the helices in V-amylose can associate into bundles. Figure 7.2.4 shows an interactive iCn3D model of the actual structure of a V-amylose, cycloamylose 26 (1C58). It consists of a linear cycloamylose strand of 26 glucose monomers, which has collapsed to form secondary structure with 6-residue helices packed together into a tertiary structure of 4-helix bundle. The blue sphere "cartoon" color coding of each glucose residue corresponds to the blue circles in the diagrammatic representation above.

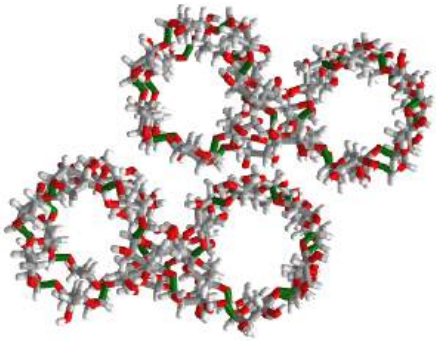


Figure 7.2.4: V-amylose, cycloamylose 26 (1C58). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...6tghM5HMTypm88

Rotate the model to explore it. Trace the chains by following the blue sphere symbolic representation for glucose as you trace the main chain. Rotate it to view down the helix axes to see the 4 holes that can each accommodate a I_3^-. In the menu button (☰), choose Style, Chemicals, Sphere to see a spacefilling model that shows the holes within each helix. Remember, NO unoccupied holes exist within either a protein alpha helix or a double-stranded DNA molecule.

The well-known macrocyclic compound cyclodextrins (for example α-cyclodextrin) are structures equivalent to one turn of V-amylose. The V-amylose helix is partially stabilized by hydrogen bonds from donors and acceptors within the helix from the OH3 on the i^{th} glucose and the OH2 on the i^{th}+1 glucose as well as from the OH6 on i^{th} glucose and OH2 on the i^{th} + 6 glucose.

In vivo, glycogen is synthesized by the attachment of glucose monomers to a core protein called glycogenin. Figure 7.2.5 shows a model of a glycogen particle with glycogenin at its core.

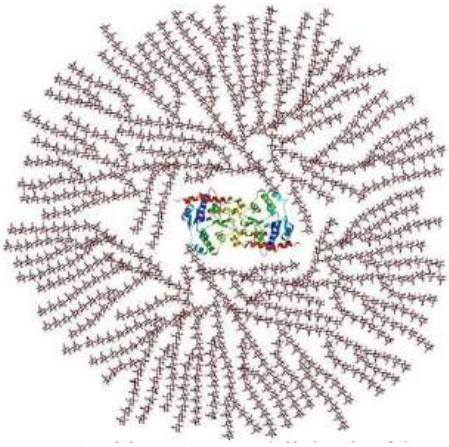


Figure 7.2.5: Glycogen particle with glycogenin at its core

The dimeric protein glycogenin is an enzyme that autoglucosylates itself in a stepwise fashion. The first glucose is added at Tyr 195. At some point, the active site must get buried and the protein can no longer add more monomers.

It makes chemical sense to store glucose residues as either glycogen or starch, one large molecule. A review of colligative properties would inform you that if glucose was stored as the monosaccharide, a great osmotic pressure difference would be found between the outside and inside of the cell. Glycogen, with its many branchs, is a single molecule. When glucose is needed, it is cleaved one residue at a time from all the branches (at the nonreducing ends) of glycogen, producing a large amount of free glucose quickly.

Phi/Psi angles can also be described for the starch/glycogen main chain (around the acetal O) in a fashion comparable to that for proteins (around the alpha carbon). The phi torsion angle describes rotation around the C1-O bond of the acetal link, and the psi angle describes rotation around the O-C4 bond of the same acetal link, with the glucopyranose ring considered as a rigid rotator (just as the 6 atoms in the planar peptide bond unit). The most extended form of a glucose polymer occurs when the glycosidic link is β 1,4 (as in cellulose), which forms linear chains. This would be analogous to the more extended parallel beta strand (phi/psi angles of -119^0, -113^0) and antiparallel beta strands (phi/psi angles of -139^0, $+135^0$) of proteins. The α 1,4 linked main chain of glycogen and starch causes the chain to turn and form a large helix. Iodine (or I_3^-) can fit into the helix, which turns a solution/suspension of starch blue, which turns starch purple. The less extended structure is analogous to the less extended protein alpha helix, which has phis/psi angles of -57^0,-47^0.

Figure 7.2.6 shows phi/psi angles for acetal/glycosidic linkage in maltose, a dissacharide of glucose, is shown below.

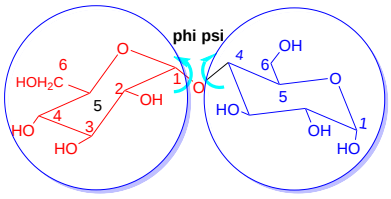


Figure 7.2.6: Phi/psi angles for acetal/glycosidic linkage in the glucoside dissacharide maltose

7.2.2: α 1,6 main chain links

Dextran is a branched polymer of glucose in α 1,6 links with α 1,2, α 1,3, or α 1,4 linked side chain. This polymer is used in some chromatography resins. Figure 7.2.7 shows chair structures (A) and wedge/dash structures (B) for dextran showing the main chain α 1,6 link with one α 1,3 branch.

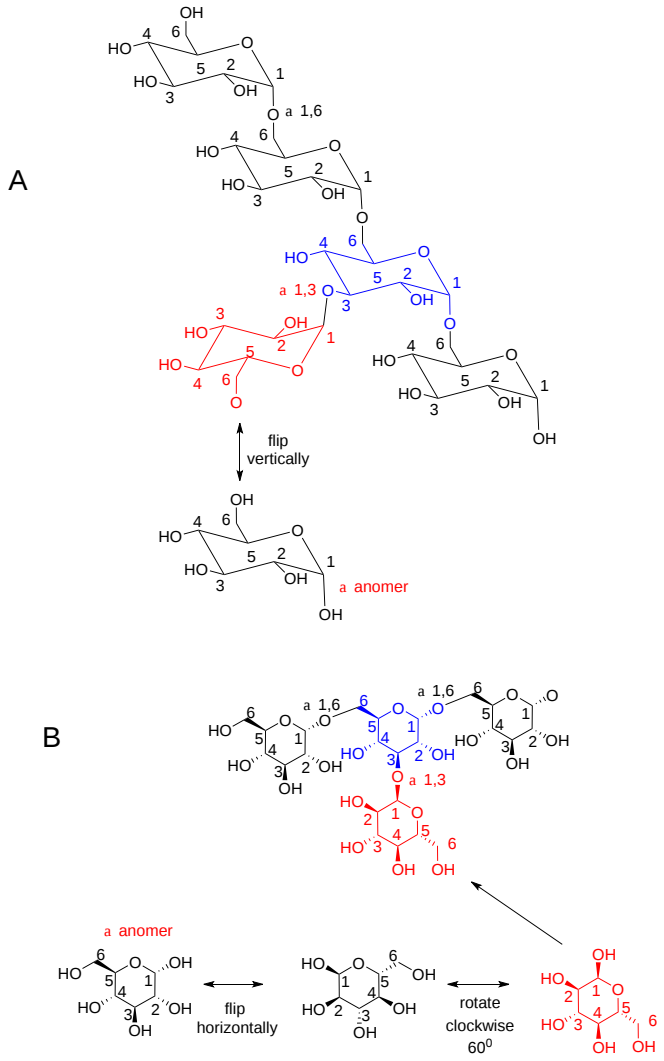


Figure 7.2.7: Chair structures (A) and wedge/dash structures (B) for dextran showing the main chain α 1,6 link with one α 1,3 branch

Depending on its molecular weight, it is soluble in water (forming viscous solutions) and organic solvents. It is also used as a food thickener and stabilizer. It is synthesized by lactic acid-forming bacteria using sucrose as an energy source. Most uses are commercial.

7.2.3: β 1,4 links

Cellulose, a structural homopolymer of glucose in plants, has of β 1,4 main chain links without branching. Multiple chains are held together by intra and inter-chain H-bonds. It is the most abundant biological molecule in nature. Various rendering of 4 glucose residues in cellulose are shown in Figure 7.2.8. Haworth structures are not shown. Instead, more chemically informative chair and wedge/dash structures are used. It's important to see the structures displayed in many ways, since different representations of carbohydrate structurse can be found in different sources.

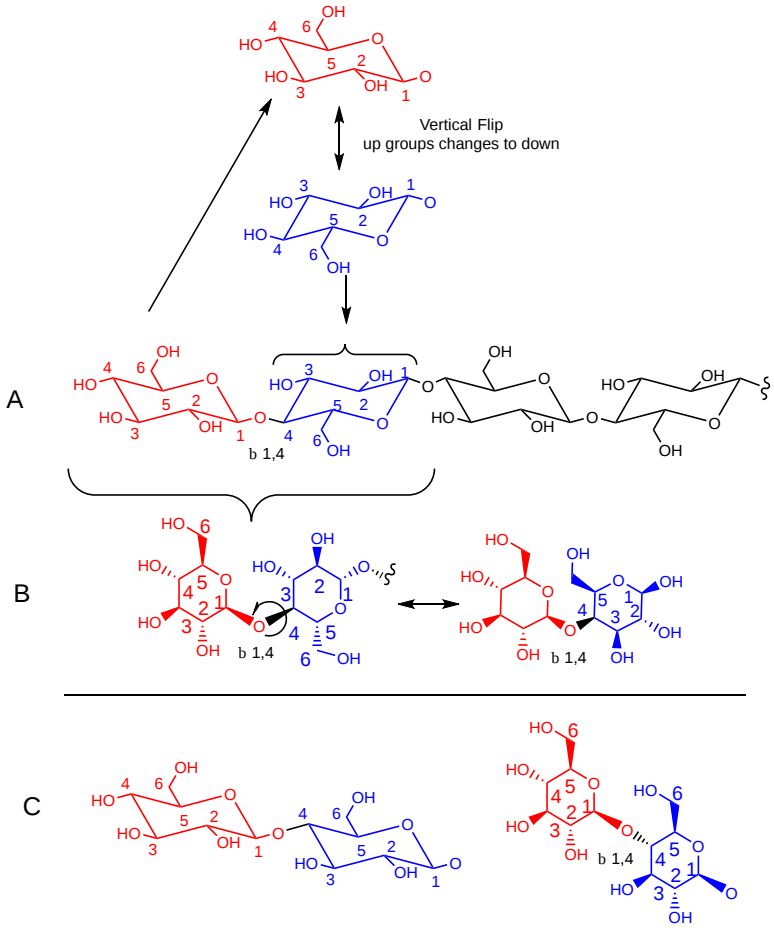


Figure 7.2.8: Rendering of 4 glucose residues in cellulose

In A, the most common chair representation, the 2nd and 4th residues from the right-hand end are flipped versions of residues 1 and 3. Residues 1 and 2 are colored red and blue for clarity. This unit is repeated to generate the full chain. The top part of A show a simplified version of the flip of the red ring to produce the blue ring to help you see that they are indeed identical structures.

The same structure as in A is shown in the left part of B in wedge/dash from (looking down on the ring). The right-hand side of B shows a variant of the left-hand side of B that is generated by simple 180^0 rotation around the bond indicated in the left of B.

In C, the simple repeat is shown without the chain flips in A and B. The acetal/glycosidic/glucosidic bond seems to be shown in a straight line in the chair structures (a bit confusing and structurally deceptive) but is shown more clearly in the adjacent wedge/dash structure.

All of the structures are correct, but the one shown in A is most often used.

One long chain of starch can interact with other chains in a structure stabilized by intrachain and interchain hydrogen bonds. Different sources display different hydrogen bonds. Some common ones are shown below. These chains align in parallel and twist to form larger cellulose fibers. Figure 7.2.9 shows an interactive iCn3D model of cellulose chains.

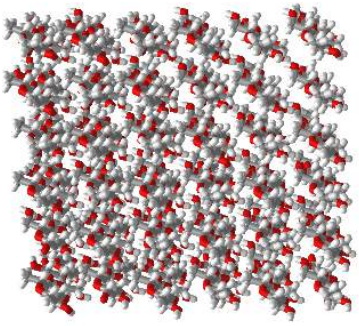


Figure 7.2.9: Cellulose chains (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...TUHKki2VN4jeTA

Chitin

The glycan is the major component in the exoskeletons of anthropoids and mollusks. It is a β 1,4 linked polymer of N-acetylglucose (GlcNAc). Compare this to cellulose which is a β 1,4 linked polymer of glucose. What a difference an N-acetyl substituent makes!

The basic chemical structure of chitin is shown in chair form in Figure 7.2.10 along with the symbolic nomenclature for glycans (SNFG).

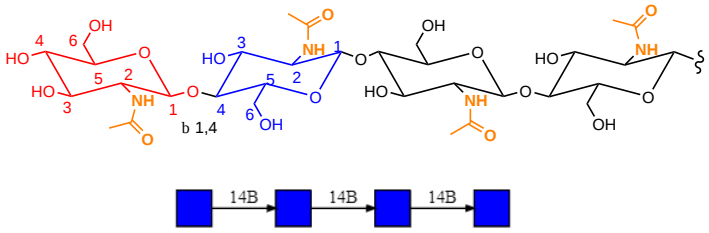


Figure 7.2.10: Chitin - Chain and symbolic nomenclature for glycans SNFG)

7.2.4: Symbolic nomenclature for glycans (SNFG) -

Before we go further into the complexities of glycan structure, let's explore the symbolic nomenclature for glycan structures. The Consortium for Functional Glycomics (2005) proposed a scheme based on specific colored geometric shapes for each, as shown for the example glycan shown in Figure 7.2.11 for a complex glycan.

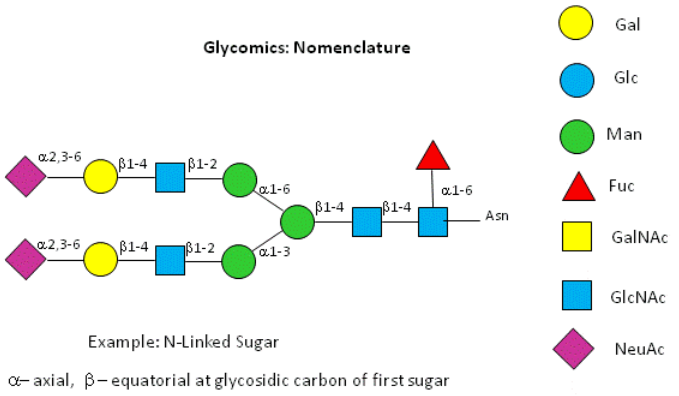


Figure 7.2.11: SNFG representation of a complex glycan

This nomenclature has recently been updated in Appendix 1B of Essentials of Glycobiology, 3rd Edition (Glycobiology 25(12): 1323-1324, 2015. doi: 10.1093/glycob/cwv091 (PMID 26543186) and is summarized in the Figure 7.2.12

SHAPE	White (Generic)	Blue	Green	Yellow	Orange	Pink	Purple	Light Blue	Brown	Red
Filled Circle	Hexose	Glc	Man	Gal	Gul	Alt	All	Tal	Ido	
Filled Square	HexNAc	GlcNAc	ManNAc	GalNAc	GulNAc	AltNAc	AllNAc	TalNAc	IdoNAc	
Crossed Square	Hexosamine	GlcN	ManN	GalN	GulN	AltN	AllN	TalN	IdoN	
Divided Diamond	Hexuronate	GlcA	ManA	GalA	GulA	AltA	AllA	TalA	IdoA	
Filled Triangle	Deoxyhexose	Qui	Rha			6dAlt		6dTal		Fuc
Divided Triangle	DeoxyhexNAc	QuiNAc	RhaNAc							FucNAc
Flat Rectangle	Di-deoxyhexose	Oli	Tyv		Abe	Par	Dig	Col		
Filled Star	Pentose		Ara	Lyx	Xyl	Rib				
Filled Diamond	Nonulosonate		Kdn				Neu5Ac	Neu5Gc	Neu	
Flat Hexagon	Unknown	Bac	LDManHep	Kdo	Dha	DDManHep	MurNAc	MurNGc	Mur	
Pentagon	Assigned	Api	Fru	Tag	Sor	Psi				

Figure 7.2.12: Symbolic Nomenclature for Glycans (SNFG) for the most common monosaccharides

7.2.5: Glycosaminoglycans - Heteropolysaccharides with dissacharide repeating units

Many polysaccharides consist of repeating disaccharide units. A major class of polysaccharides with disaccharide repeats include the **glycosaminoglycans** (**GAGs**), all which contain one amino sugar in the repeat and in which one or both of the sugars contain negatively charged sulfate and/or carboxyl groups. The extent and position of sulfation vary widely between and within GAGs. GAGs are found in the vitreous humor of the eye and synovial fluid of joints, and in connective tissue like tendons, cartilage, etc, as well as skin. They are found in the extracellular matrix and are often covalently attached to proteins to form proteoglycans. From a bird's eye view, they are all elongated polyanions.

They and their structures are very complicated and exceedingly diverse. This makes them difficult to understand for those who want clear and unambiguous structures. From a biological perspective, they present in their local environment an incredibly diverse array of potential binding sites for ligand (both small and large). Because of these they also have functions in cell signaling. In addition, some GAGs are free-standing, others are covalently attached to proteins (a bit like glycogen is attached to glycogenin). These large molecules are called **proteoglycans**. We will discuss this later in Chapter 7.4 when we discuss the "carbohydrate code"

Here are the ring structures and descriptions of important GAGs. The common disaccharide repeat unit is shown twice for each structure, with the knowledge that sulfation patterns may differ for the disaccharide repeats in the actual chains. Note also that the first member of each disaccharide repeat shows the ring flipped vertically (top to bottom) as was shown in the structures for other beta-linked glycans (cellulose, chitin) above.

In a long chain, selecting which is the repeating disaccharide unit is a bit relative, as shown in Figure 7.2.13 for the repeating disaccharide sequence of N-acetylglucosamine (blue square) and N-acetylgalactosamine (yellow square).

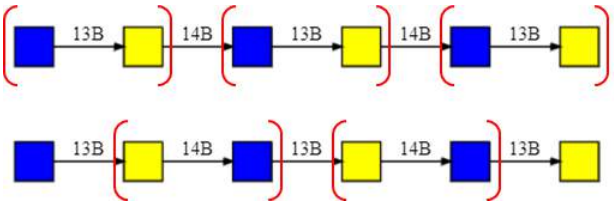


Figure 7.2.13: SNFG representation of a disaccharide sequence of N-acetylglucosamine (blue square) and N-acetylgalactosamine (yellow square)

In the top, the repeating units (blue-yellow) are connected to each other through beta 1,4 links while in the bottom, the connection of the repeating unit (yellow-blue) is beta 1,3. Without knowing the full chain, the best choice of annotating the repeating unit is illusive. What's most important however is to note the alternating acetal/glycosidic links throughout the whole sequence. In the figures below different disaccharide repeats are highlighted.

Hyaluronic acid

This is a polymer of glucuronate (β 1,3) GlcNAc. It offers a backbone for theattachment of protein and other GAGs. It's the only GAG without sulfate. Figure 7.2.14 shows a tetrasaccharide fragment with two disaccharide repeats. The internal acetal/glycosidic link of the illustrated disaccharide repeat is β 1,3 while the connection between the disaccharides is β 1,4. For one last time, the vertical flip of the glucuronic acid is shown to allow a better understanding of its flipped presentation in the actual GAG.

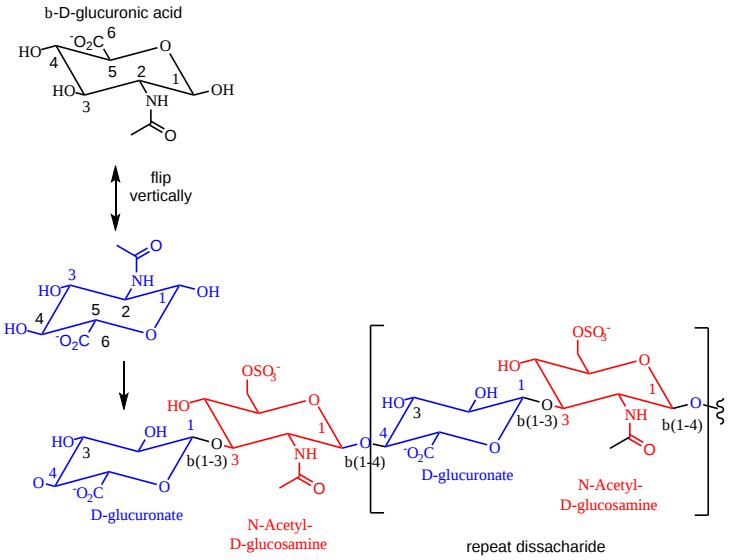


Figure 7.2.14: Hyaluronic acid - tetrasaccharide fragment with two disaccharide repeats

Hyaluronic acids are found in a variety of locations including synovial fluid, the extracellular matrix and skin, where it helps control skin moisture. It is water soluble and displays twin antiparallel left-oriented helices. Covalent conjugates of the chemotherapeutic drugs doxorubicin and camptothecin linked to hyaluronic acid, whose overall structure is similar to "worm-like micelles", have been used successfully to treat skin cancers.

Keratan sulfate

This GAG contains repeats of N-acetyl-D-glucosamine-6-phosphate in β 1,3 link to D-galactose or D-galactose-6-sulfate. The link between Gal and the modified glucosamine is β 1,4. Keratin sulfate is highest abundant in the cornea of the eye but is also found in other connective tissues such as bone, cartilage and tendon, as well as in the
central and peripheral nervous system.

Figure 7.2.15 shows a tetrasaccharide containing two repeating disaccharides.

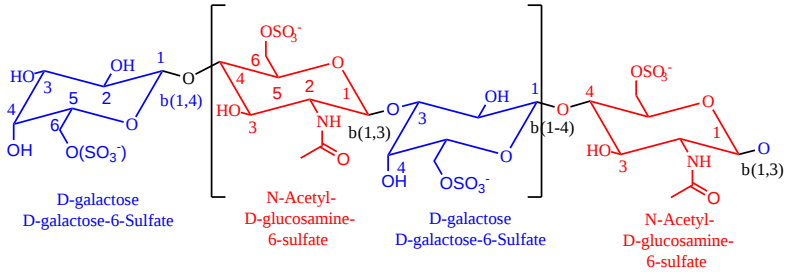


Figure 7.2.15: Keratan sulfate disaccharide repeats

Chondrotin sulfate

Th repeat dissacharide unit is D-glucuronate β(1,4) GalNAc-4 or 6-sulfate. It's found in connective tissue matrix as well as the cell surface (in the form of proteoglycans), in basement membranes, as well as intracellular granules. A tetrasacccharide showing two disaccharide repeats is shown in Figure 7.2.16.

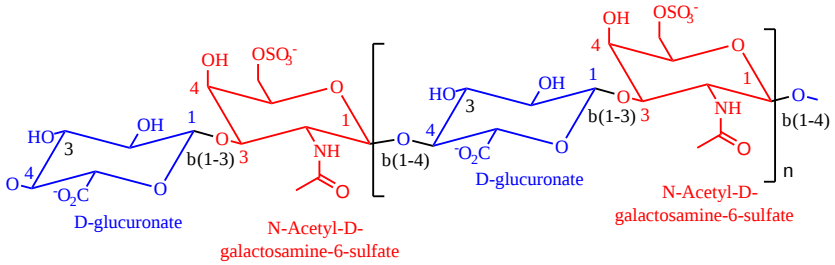


Figure 7.2.16: Chondroitin sulfate disaccharide repeats

Dermatan sulfate

This glycosoaminoglycan is similar to chondroitin sulfate. It is first made as a polymer of the disaccharide unit of D-gluconic and N-acetyl-D-galactosamine. The gluconic acid is epimerized to L-iduronic acid, followed by sulfation. Its structure is shown in Figure 7.2.17.

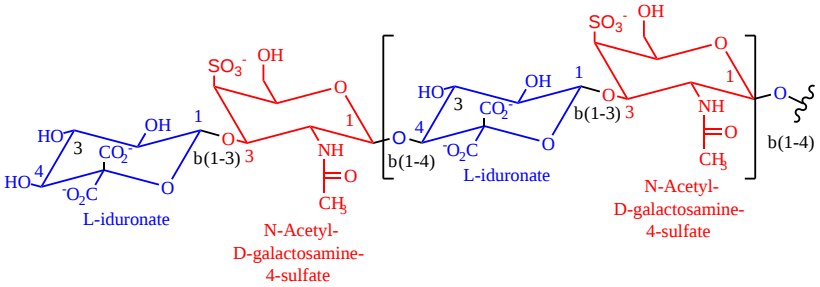


Figure 7.2.17: Dermatan sulfate disaccharide repeats

Heparin

This GAG contains a highly trisulfated disaccharide repeat as shown in Figure 7.2.18. Note that the molecule can contain glucuronate or iduronate, and the degree of sulfation of the chains varies. (Remember, there is no genetic code that specifies the actual sequence or sulfation pattern in these polymers.)

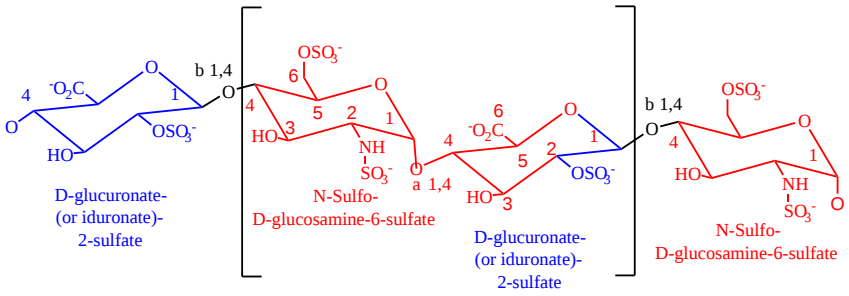


Figure 7.2.18: Heparin disaccharide repeats

Most people are familiar with the anti-clotting properties of heparin administered as a drug. Heparin acts as a "catalyst" to accelerate the inhibition of the enzyme thrombin, which cleaves fibrinogen and activates platelets to form clots, by the blood protein antithrombin III. Heparin works in two ways to facilitate thrombin inactivation. It has a specific binding site for antithrombin III which causes a conformation in the protein, making it a more effective inhibitor. Thrombin, a positively-charged serine protease, can bind the heparin, a polyanion, nonspecifically. When it does, it diffuses along the heparin chain, where it can find bound antithrombin III much more quickly than if the inhibitor was free in the blood. Heparin effectively changes the search path of thrombin from a 3D to a 1D search.

Figure 7.2.19 shows an interactive iCn3D model of the amino acids in antithrombin III within seven angstroms of a bound heparin 5mer (1NQ9). Dotted lines represent hydrogen bonds and salt bridges between the two. Heparin is highlighted in yellow

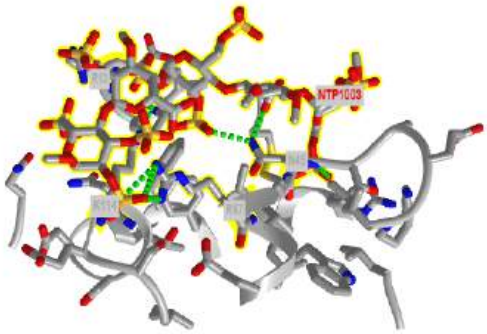


Figure 7.2.19: Heparin 5mer - antithrombin III complex (1NQ9. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...fnvt6YfbJovat8

7.2.6: Agarose

Agarose is the main polysaccharide component derived from red algae. Agarose is a polymer of a disaccharide repeat of (1,3)-β-D-galactopyranose-(1,4)-3,6-anhydro-α-L-galactopyranose, is often used for a gelable solid phase for electrophoresis of nucleic acid and as a component of chromatography beads. As with starch, which is present as mixtures of amylose and amylopectin, agarose is often found with agaropectin, which is a sulfated galactan. A tetrasaccharide fragment with two dissacharide repeats is shown in Figure 7.2.20

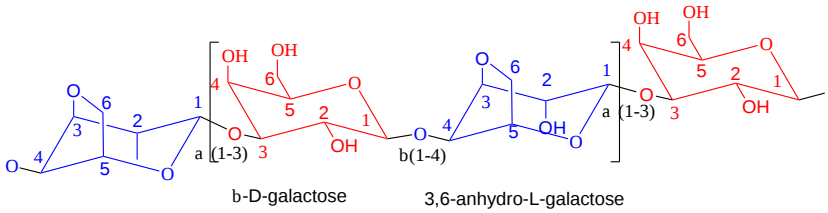


Figure 7.2.20: Agarose disaccharide repeats

7.3: Glycoconjugates - Proteoglycans, Glycoproteins, Glycolipids and Cell Walls

Many proteins, especially those destined for secretion or insertion into membranes, are post-translationally modified by the attachment of carbohydrates. They are usually attached through either Asn or Ser side chains. Carbohydrate modifications on the protein appear to be involved in the recognition of other binding molecules, prevention of aggregation during protein folding, protection from proteolysis, and increases the half-life of the proteins. In contrast to a protein sequence determined by a DNA template, sugars are attached to proteins by enzymes that recognize appropriate sites on proteins and attach the sugars. Since there are many sugars with many functional groups that can serve as potential attachment sites, the structures of the oligosaccharides attached to proteins are enormously varied, complex, and hence "information rich" compared to linear or folded polymers like DNA and proteins.

7.3.1: N-linked Glycoproteins

These contain carbohydrates attached through either a GlcNAc or GalNAc to an Asn in a X-Asn-X-Thr sequence of the protein. There are three types of N-linked glycoproteins, high mannose, complex, and hybrid. They all contain the same core oligosaccharide - (Man)3(GlcNAc)2 attached to Asn as shown in Figure 7.3.1.

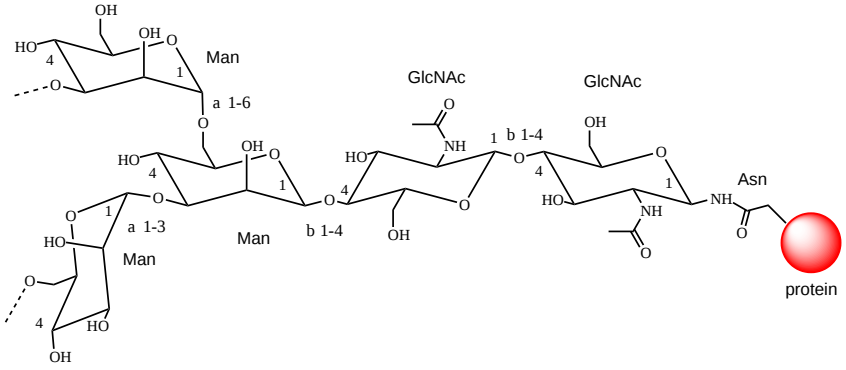
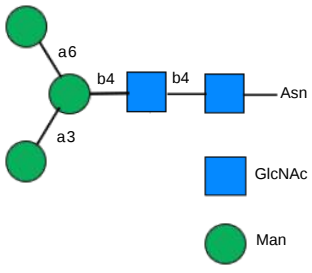
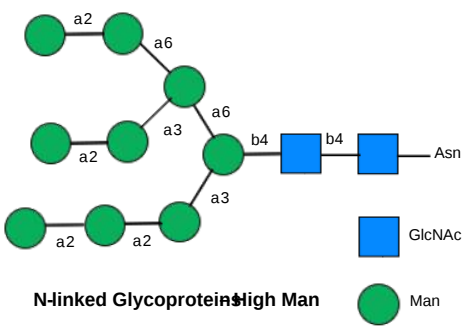


Figure 7.3.1: Core oligosaccharide - (Man)3(GlcNAc)2 attached to Asn in N-linked glycoproteins

Table 7.3.1 below shows the SNFG representation for the main core and variant glycans in **N-linked glycoproteins**. Note that the designation of α2 implies an α(1→2) linkage. Unless otherwise stated the linkage is presumed to start from carbon 1.

Core	 **N-linked Glycoprotein Core**
High mannose	 **N-linked Glycoprotein High Man**

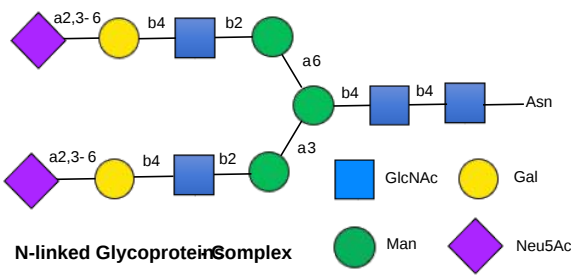
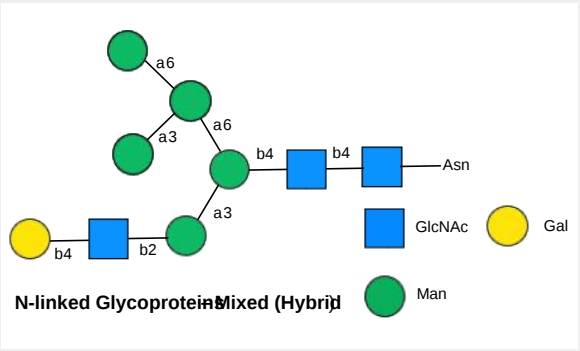
Complex	
Mixed hybrid	

Table 7.3.1: SNFG representation for the main core and variant glycans in **N-linked glycoproteins**

Complex N-linked glycans don't contain mannose outside of the core glycan and have GlcNAc attached to the branching mannoses in the core structure. The complex glycan shown above has a Gal(β1,4)GlcNAc sequence which could be named as the disaccharide lactosamine. Often, lactosamine repeats in the sequence.

Hybrid glycans have both unsubstituted terminal mannoses (as in the high-mannose type) and substituted mannoses with an N-acetylglucosamine attached (as in the complex type. GlcNAc residues added to the core in the hybrid and complex N-glycoproteins are called antennae. Figure 7.3.2 shows an example of a biantennary N-linked glycan with two GlcNAc branches linked to the core. The core is outlined in red and the two GlcNAcs are labeled 1 and 2.

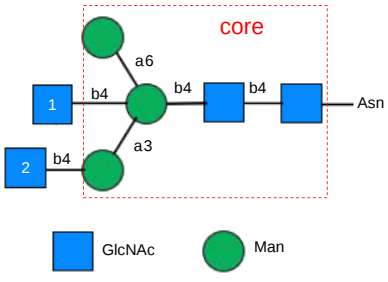


Figure 7.3.2: Biantennary N-linked glycan with two GlcNAc branches linked to the core

Complex glycans also have bi-, tri- and tetraantennary forms and comprise most of N-glycans. As shown in Table 7.3.1 above, complex N-linked glycans usually end with sialic acid residues. About 50% of the surface area of the COVID Sars-2 spike protein is covered with glycans as shown in the model structure in Figure 7.3.3. The protein surface is gray and the glycans (biantennary LacNAc N-glycans) in spacefill CPK with carbon in cyan.

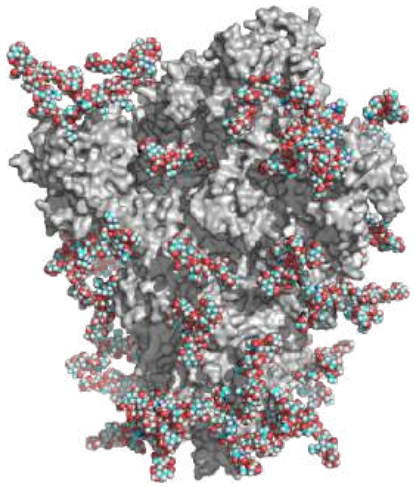


Figure 7.3.3: COVID Sars-2 spike protein is covered with glycans (colored spacefill). (PDB file coordinates 5.Swiss.3.M3F1.CYX.TER from Analysis of the SARS-CoV-2 spike protein glycan shield: implications for immune recognition. Oliver C. Grant, David Montgomery, Keigo Ito, Robert J. Woods. **doi:** https://doi.org/10.1101/2020.04.07.030445)

In the hybrid oligosaccharide shown above, one terminus contains Gal(β1,4)GlcNAc. However, in all other mammals except man, apes, and old-world monkeys, an additional Gal is often connected in an α1,3 link to the Gal to give a terminus of Gal(α1,3)Gal(β1,4)GlcNAc. These animals have an additional enzyme, an α1,3 Gal transferase. Bacteria also have this enzyme, and since we have been exposed to this link through bacterial infection, we mount an immune response against it. Why is this important? Pig hearts turn out to be similar to human hearts, so they might be good candidates for transplantation into humans (xenotransplants). However, the Gal-α1,3-Gal link is recognized as foreign, and we mount a significant immune response against it. Several biotech firms are trying to delete the pig α1,3 Gal transferase which would prevent the addition of the terminal Gal, and make them good donors for transplanted hearts.

Figure 7.3.4 shows an interactive iCn3D model of an N-linked glycoprotein, human beta-2-glycoprotein-I (Apolipoprotein-H) (1C1Z).

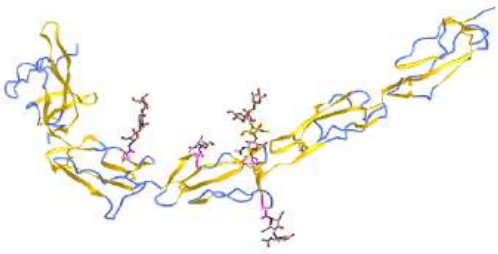
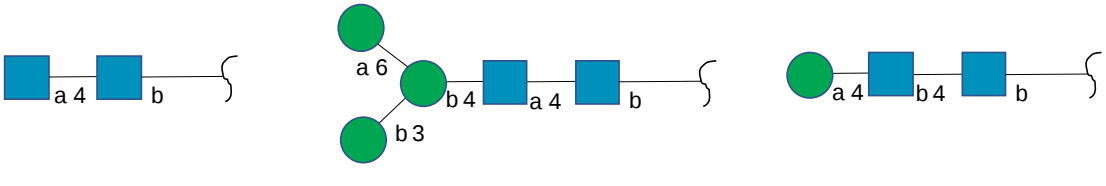


Figure 7.3.4: N-linked glycoprotein, human beta-2-glycoprotein-I (Apolipoprotein-H) (1C1Z). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...GkZSVFdciXiBT7

? Exercise 7.3.1

The glycan structures for the beta-2-glycoprotein-I are shown below. Identify the monosaccharides in each and specific to which asparagine they are linked.



Answer

add answer here.

Figure 7.3.5 shows an interactive iCn3D model of the GP120 HIV protein that contains a high mannose, complex and hybrid N-linked glycans. Most glycoproteins in the Protein Data Bank do not contain attached glycans. The glycans here were added with the program GlyProt at 3 of 17 possible Asn residues that would presumably have attached glycans. Use your mouse or key paid to hover over the monomers in the attached glycans. Abbreviations for the given residues in the model are as follows: adm = alpha-D-Man, bdg= beta-D-Glc or Gal, adn = alpha-D-neuramindase.

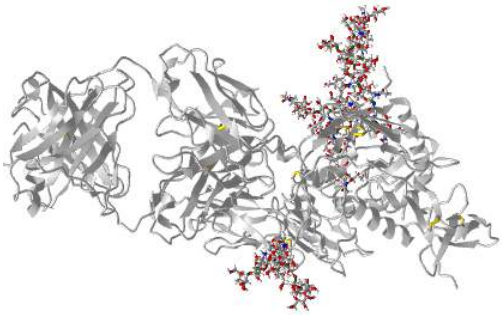


Figure 7.3.5: GP120 HIV protein that contains a high mannose, complex and hybrid N-linked glycans. (Copyright; author via source). No external link is available for this structure.

? Exercise 7.3.1

Which Asn chains contain the high Man, complex, and hybrid glycans?

Answer

Add answer here

The coronavirus pandemic of 2020-23 has been deadly (over 1.1 million deaths in the USA alone as of February 8, 2023 and 6.8 million around the world). However, the 1918 influenza pandemic was far worse with an estimated 650,000 deaths in the USA and 50 million around the world in a population since the population was less than 1/3 of the present. However, the number of USA deaths during the Covid-19 pandemic is greater than during the 1918 Flu pandemic. An additional 3 million deaths in the USA were probably prevented by vaccination as of December 2022. A large number of deaths in the developing world would have been prevented if wealthy nations allocated more resources to produce and distribute the vaccines. The evolution of the virus in nonvaccinated areas might come back to haunt wealthy countries if present vaccines become ineffective against the mutants. A worse pandemic might await us. An avian version of the influenza virus (H5N1), presently endemic in wild birds and now found in mink populations, has infected 240 people as of January 2023 and killed 56% of them. A quick note: the 1918 pandemic affected youth the most.

Influenza and the Avian Flu

The influenza virus, a simple yet deadly virus, is shown below in Figure 7.3.6. It interacts with human cells through a surface protein, hemagglutinin (HA).

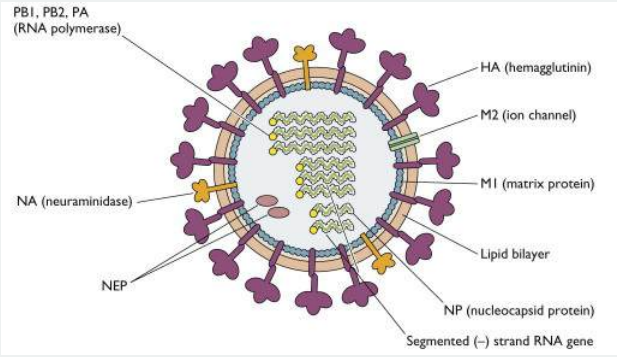


Figure 7.3.6: Influenza virus. Vincent Racaniello. https://www.virology.ws/2009/04/30/s...fluenza-virus/. Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

The virus binds to host cells through the interaction of HA with cell surface carbohydrates. Once bound, the virus internalizes, ultimately leading to the release of the RNA genome of the virus into the host cell.

The hemagglutinin protein is the most abundant protein on the viral surface. 15 avian and mammalian variants have been identified (based on antibody studies). Only 3 adapted to humans in the last 100 yr, giving pandemic strains H1 (1918), H2 (957) and H3 (1968). Three recent avian variants (H5, H7, and H9) jump directly to humans recently but have low human-to-human transmissibility.

The influenza hemagglutinin protein has the following characteristics:

- the mature form is a homotrimer (3 identical protein subunits), MW 220,000 with multiple sites for covalent attachment of sugars. Hemagglutinin is a glycoprotein.
- each monomer is synthesized as a single polypeptide chain precursor (HA0), which is cleaved into HA1 and HA2 subunits by the protease trypsin in epithelial cells of the lung.
- structure known for human (H3), swine (H9), avian (H5) subtypes.

Hemagglutinin binds to sialic acid (Sia) covalently attached to many cell membrane glycoproteins. The sialic acid is usually connected through an α(2,3) or α(2,6) link to galactose on N-linked glycoproteins. The subtypes found in avian (and equine) influenza isolates bind preferentially to Sia (α2,3) Gal which predominates in the avian GI tract where viruses replicate. Human influenza isolates prefer Sia α(2,6)

Sia (α2,3) Gal which predominates in the avian GI tract where viruses replicate. Human influenza isolates prefer Sia α(2,6) Gal. The human virus of H1, H2, and H3 subtypes (causes of the 1918, 1957, and 1968 pandemics) recognize Sia α(2,6) Gal, the major form in the human respiratory tract. The swine influenza HA binds to Sia α(2,6) Gal and some Sia (α2,3) Gal both of which are found in swine. The structures of the Sia-Gal disaccharide are shown in Table 7.3.2 below.

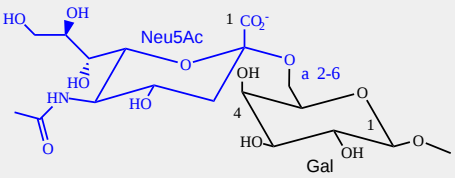
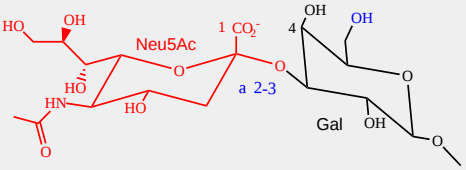
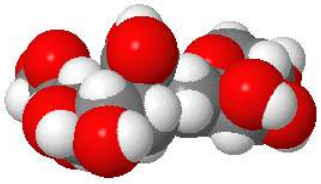
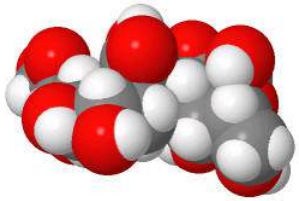
Sia α(2,6) Gal (Human)	Sia α(2,3) Gal (Avian and some Swine)
	
 (made with Sweet, with an OH, not AcNH on sialic acid on C5)	 (made with Sweet, with an OH, not AcNH on sialic acid on C5)

Table 7.3.2: Structures of Sia α(2,6) Gal (human) and Sia α(2,3) Gal (avian/swine)

The H5N1 avian flu H5N1) virus is deadly but presently lacks human-to-human transmissibility. Why? One reason is that it appears to bind deep in the lungs and is not released easily on coughing or sneezing. It appears that cell surface glycoproteins deeper in the respiratory tract have Sia (α2,3) Gal which accounts for this pathology.

Before it leaves the cell, the virus forms a bud on the intracellular side of the cell with the HA and NA in the cell membrane of the host cell. The virus in this state would not leave the cell since its HA molecules would interact with sialic acid residues in the host cell membrane, holding the virus in the membrane. Neuraminidase hydrolyzes sialic acid from cell surface glycoproteins, allowing the virus to complete the budding process and be released from the cell as new viruses. The drugs Oseltamivir (Tamiflu) and zanamivir (Relenza) bind to and inhibit neuraminidase, whose activity is necessary for viral release from infected cells. Tamiflu appears to work against N1 of the present H5N1 avian influenza viruses. Governments across the world are hopefully stockpiling this drug in case of a pandemic caused by the avian virus jumping directly to humans and becoming transmissible from human to human.

7.3.2: O-linked Glycoproteins

The CHOs are usually attached from a Gal (β 1,3) GalNAc to a Ser or Thr of a protein as shown in Figure 7.3.7.

Figure: O-linked Glycoproteins

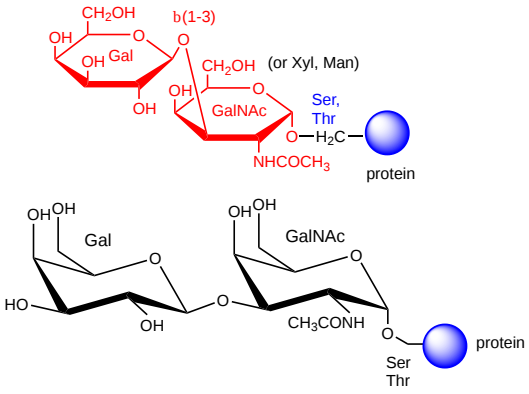


Figure 7.3.7: O-linked Glycoproteins

The blood group antigens (CHOs on cells attached to either proteins or lipids) are examples. The sugars shown as chairs (in contrast to structures found in many texts) in Figure 7.3.8 are the blood group antigens. They are attached to a core heterosaccharide (shown as a red ellipse below), which is connected to either a membrane glycoprotein or glycolipid.

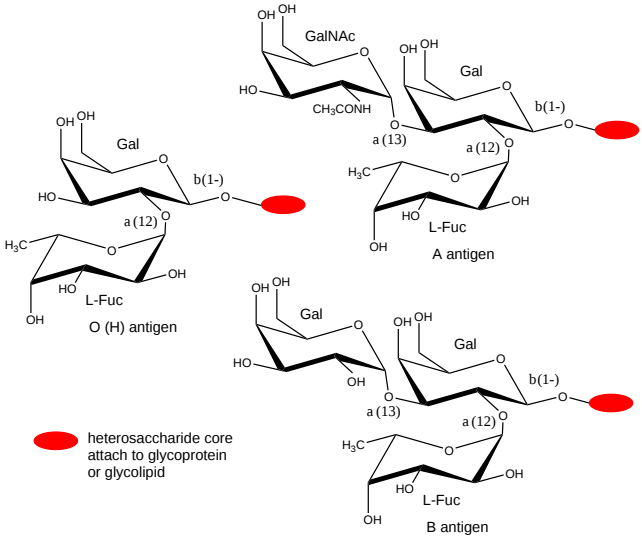


Figure 7.3.8 : Blood group glycans

Figure 7.3.9 shows the SNFG representation for the A antigen in the glycolipid form.

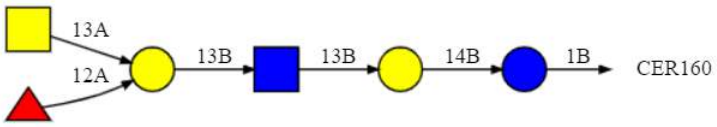


Figure 7.3.9: SNFG representation for the A antigen in the glycolipid form

The trimeric branched residues on the left-hand side represent the A antigen shown above. The red triangle is L-fucose. Yellow represents galactose or GalNac, while blue is glucose or GlcNAc.

7.3.3: Proteoglycans

Some proteins are so modified with CHOs that they contain more CHOs than amino acids. Proteins linked to glycosaminoglycans are together called proteoglycans (PGs). The consists of a core protein linked to one or more glycosaminoglycans. GAGs are linear sulfated glycans which we described earlier. The structures of a few proteoglycans are known. The GAGs are O-linked to the protein, typically to a Ser of a Ser-Gly dipeptide often repeated in the protein. Some of the proteoglycans also contained N-linked oligosaccharide groups. Figure 7.3.10 shows a representation of proteoglycan structure.

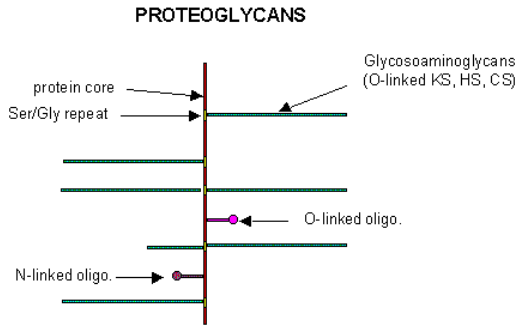


Figure 7.3.10: Representation of proteoglycan structure.

PGs can be soluble and are found in the extracellular matrix, or as integral membrane proteins. There are about 43 genes for proteoglycans. Differential splicing of the RNA transcripts give rise to soluble and transmembrane forms. Given the diversity of sugars and the varying extent of sulfation, the CHO part of PGs provides an incredible variety of binding structures at or near the cell surface. Figure 7.3.11 shows the variety of proteoglycans found in mammalian cells. PGs help form the extracellular matrix, which provides a rich binding environment between cells.

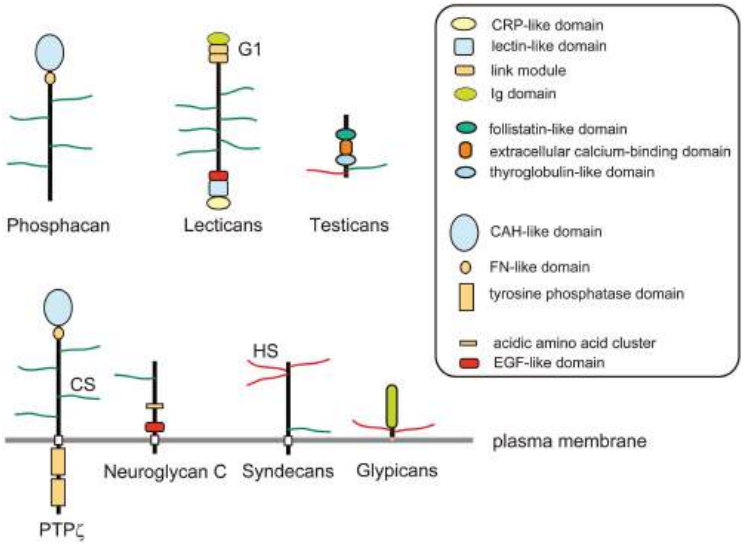


Figure 7.3.11: Proteoglycans found in mammalian cells

One PG, syndecan, binds through its intracellular domain to the internal cytoskeleton of the cell, while interacting with another protein - fibronectin - in the extracellular matrix. Fibronectin also binds other molecules which can regulate cellular growth and other interactions. PGs act like glue in connecting the extracellular and intracellular functions of the cell. There are four different core syndecan proteins (SDCs 1–4), with SDC4 lacking the cytoplasmic and transmembranes and thus is a soluble form in the intracellular matrix. The glycan components of syndecans are mostly heparan sulfate while SDC 1 and 3 also have two chondroitin sulfate chains.

Most proteins bind PGs through a PG binding motif of BBXB or BBBXXB where B is a basic amino acid. Some proteins bind to specific sequences in specific GAGs. For instance, antithrombin 3, an inhibitor of blood clotting, binds specifically to heparin. This enhances its interaction with clotting proteins such as thrombin and Factor Xa. Figure 7.3.12 shows an interactive iCn3D model of a 5 residue fragment of heparin interacting with the key amino acids side chains of Factor Xa (2gd4).

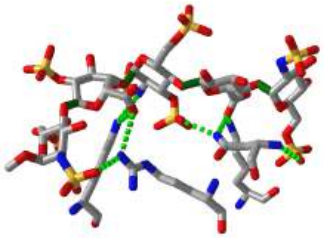


Figure 7.3.12: Heparin-5mer interactions with antithrombin III in a complex with Factor Xa. (2gd4) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...5nZ99mUhULUi59

For those more chemically oriented, the extracellular matrix (ECM) might appear to be a nondescript mess, since chemists are used to well-defined structures. Figure 7.3.13 shows a cartoon of the ECM and may clarify the components. Few structure files exist for them given the inherent flexibility of the glycan components.

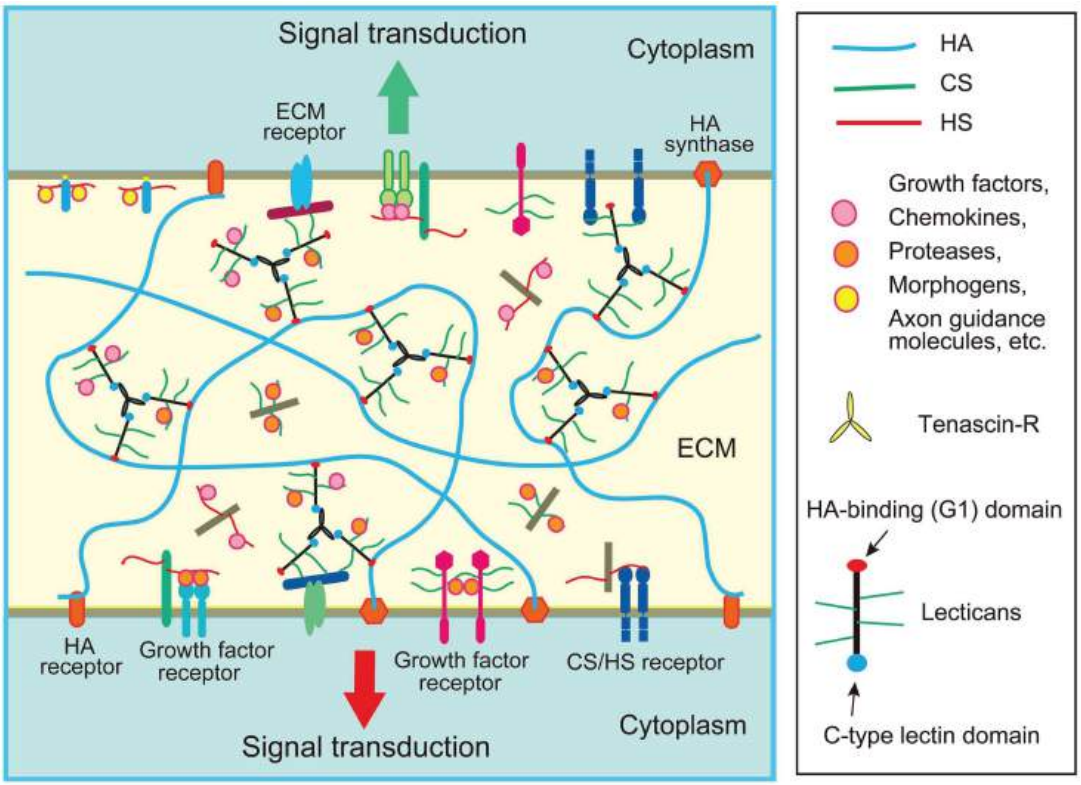


Figure 7.3.13: Cartoon representation of the extracellular matrix. Frontiers in Neuroscience 9(50):98 (2015). DOI: 10.3389/fnins.2015.00098. CC BY 4.0

7.3.4: Cell Walls and Glycolipids

In contrast to eukaryotic cells, bacteria and plant cells have a cell wall in addition to a lipid bilayer membrane. These are essentially carbohydrate polymers that determine cell shape, affording protection from exterior pathogens, hypotonic conditions and high internal osmotic pressures, preventing swelling and bursting of the cells. This is especially important in plants, which need strength and rigidity against the "turgor" pressure of the aqueous cytoplasm against the cell membrane. This prevents wilting in plants. The cell wall in plants and probably bacteria are involved in cell signaling across the cell membrane.

7.3.4.1: Bacteria Cell Walls

Two types of cell walls occur in Nature.

7.3.4.1.1: a. Gram positive bacteria-

These bacteria can be stained with Gram stain. The wall consists of a GlcNAc (β 1,4) MurNAc repeat. (GlcNAc is often abbreviated as NAG while MurNAc is abbreviated as NAM.) This is similar to the GlcNAc (β 1,4) GlcNAc homopolymer chitin, except that every other GlcNAc contains a lactate molecule covalently attached in an ether-linkage to the C3 hydroxyl to form the monomer N-Acetylmuramic acid. A pentapeptide (Ala-D-isoGlu-Lys-D-Ala-D-Ala) is attached through an amide link to the carboxyl group of the lactate in MurNAc. The GlcNAc (β 1,4) MurNAc strands are covalently connected by a pentaglycine bridge through the epsilon amino group of the pentapeptide Lys on one strand and the terminal D-Ala of a pentapeptide on another strand. A small part of the structure of a gram-positive bacterial cell wall is shown in Figure 7.3.14 It shows one repeating GlcNAc-MurNAc disaccharide unit in front (darker) and one in the back (lighter) connected through the peptides shown.

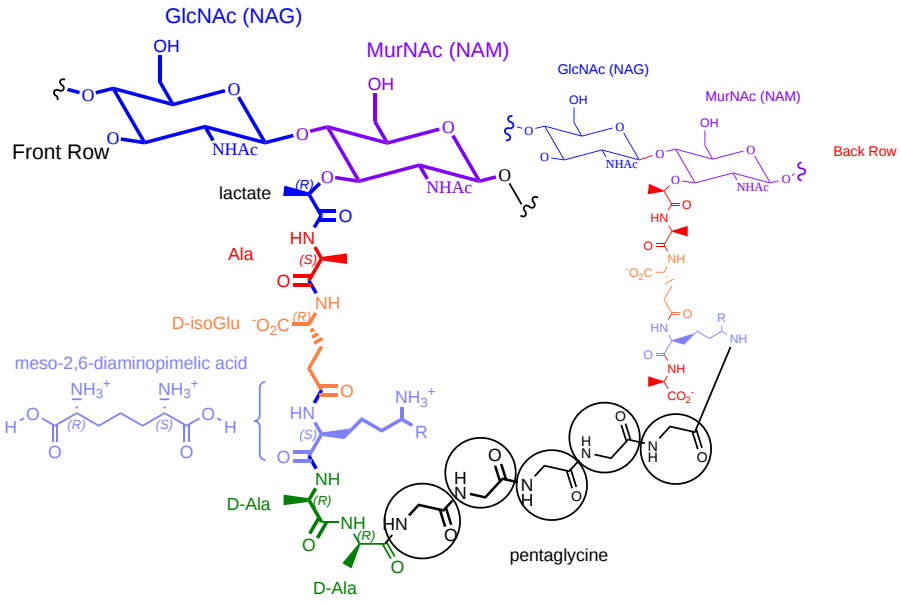


Figure 7.3.14: Part of the structure of a gram-positive bacterial cell wall

The SNFG representation of a larger section of the gram-positive cell well is shown in Figure 7.3.15

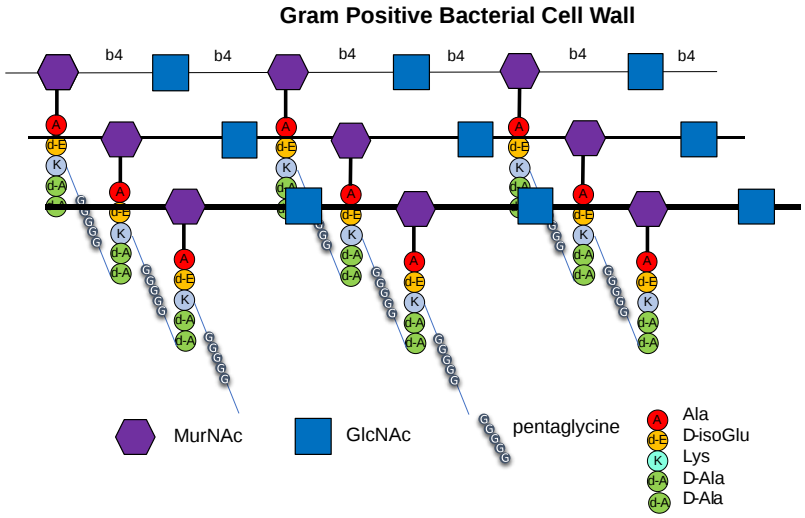


Figure 7.3.15: SNFG representation of a larger section of the gram-positive cell wall.

One final structure is found in Gram + peptidoglycan cell walls. Techioic acids are often attached to the carbon 6 of MurNAc. Teichoic acid is a polymer of glycerol or ribitol to which alternative GlcNAc and D-Ala are linked to the middle C of the glycerol. Multiple glycerols are linked through phosphodiester bonds. These teichoic acids often make up 50% of the dry weight of the cell

wall and present a foreign (or antigenic) surface to infected hosts. These often serve as receptors for viruses that infect bacteria (called bacteriophages). Its structure is illustrated in Figure 7.3.16.

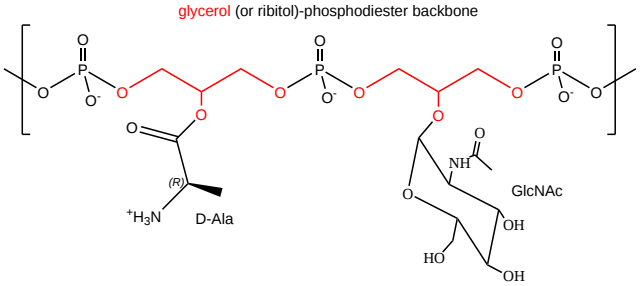


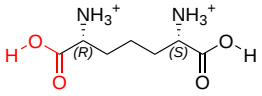
Figure 7.3.16 : Teichoic acid

Notice that all monomeric units of peptidoglycan and attached teichoic acid derivatives are covalently attached on form one large molecule comprising the entire cell wall! This structure, along with the Gram-negative cell wall structures, is the largest single macromolecule in nature.

7.3.4.1.2: *b. Gram-negative bacteria*

These bacteria can NOT be stained with Gram stain. The wall consists of the same structure as in Gram-positive bacteria but the GlcNAc (β 1,4) MurNAc strands are covalently connected through a direct amide bond between a derivative of Lys, meso-diaminopimelic acid (m-A2pm), on one peptide strand and to the last D-Ala of a pentapeptide on another strand. (i.e. there is no pentaGly spacer). The connector peptide is Ala-D-isoGlu-m-A2pm-D-Ala-D-Ala

m-A2pm replaces Lys 3 of the peptide in most Gram-negative species and in Gram-positive bacteria of the genus *Bacillus* and *mycobacteria*. The stereochemistries at each chiral center are different (R and S), but because the molecule has a plane of symmetry, it is an example of a meso-compound, a diastereoisomer of a molecule, which does not have a different enantiomeric version. The structure is shown in Figure 7.3.17.



meso diaminopimelic acid

Figure 7.3.17: Meso-diaminopimelic acid

A small part of the structure of a Gram-negative bacterial cell wall is shown in Figure 7.3.18.

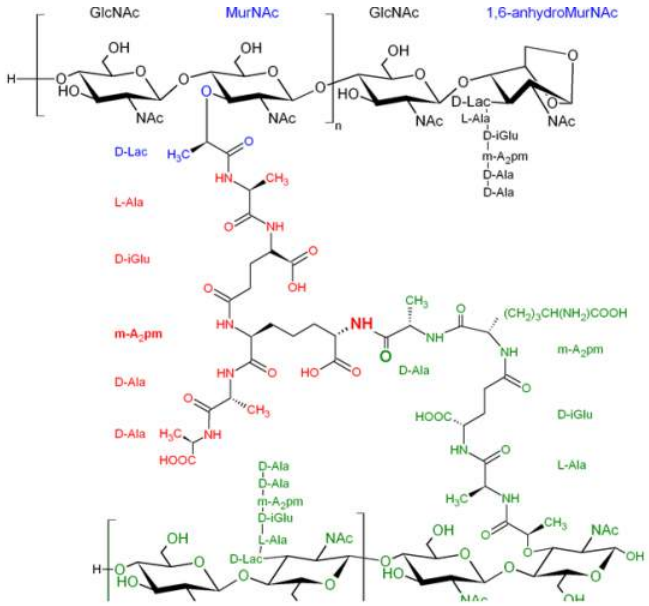


Figure 7.3.18: Part of the structure of a gram negative bacterial cell wall. http://www.glycopedia.eu/e-chapters/...ular-structure. https://gagdb.glycopedia.eu/license

Figure 7.3.19 shows an interactive iCn3D model (actual computed model, not a crystal or NMR structure) of the Gram-negative peptidoglycan of E. Coli. The PDB coordinates were kindly provided by Jame Gumbart. The peptide part of the peptidoglycan is represented in spacefill. The repeating $(GlcNac\text{-}MurNac)_n$ and pentapeptide

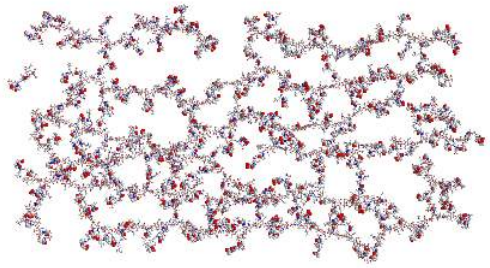


Figure 7.3.19: Gram-negative peptidoglycan of E. Coli. PDB coordinates kindly provided by Jame Gumbart. No external link is available for this structure.

In addition, Gram-negative bacterial don't have teichoic acid polymers. Rather they have a **second, outer lipid bilayer**. The cell wall peptidoglycan (PG) is sandwiched between the inner and outer bilayers. The space between the lipid bilayers is called the periplasm. The outer membrane is coated with a lipopolysaccharide (LPS) of varying composition. The LPS determines the antigenicity of the bacteria. The different LPS are called the O-antigens. Figure 7.3.20 shows the structure of the Gram-negative bacterial membrane organization. (PS is LPS, PG is peptidoglycan)

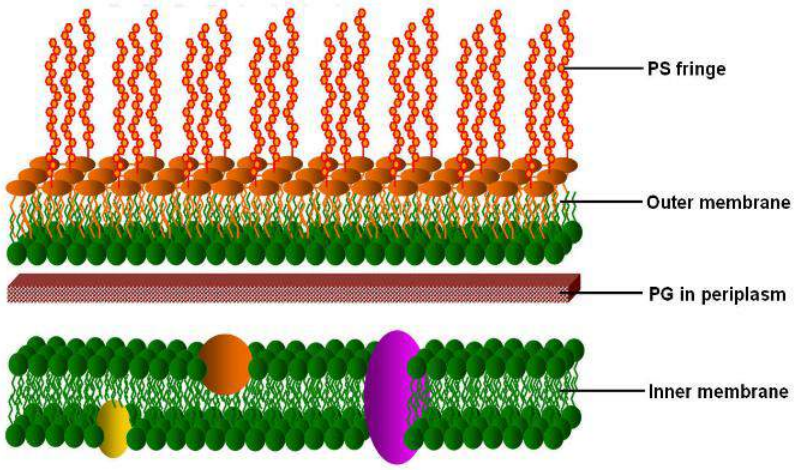


Figure 7.3.20: Overall structure of the Gram-negative bacterial membrane organization. https://cronodon.com/BioTech/Bacteria_envelope.html.

A detailed view of the structure of the lipopolysaccharide (LPS) from *Salmonella tryphimurium* is shown in Figure 7.3.21 below.

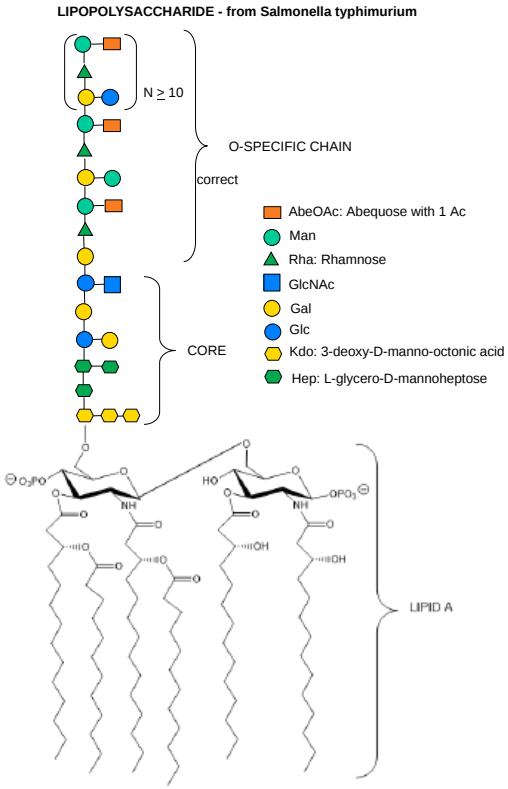


Figure 7.3.21: Lipopolysaccharide (LPS) from *Salmonella tryphimurium*

? Exercise 7.3.1

questions on this:diff btw g+ and g- cells

https://microbiologyinfo.com/differe...tive-bacteria/.

Answer

Add texts here. Do not delete this text first.

7.3.4.1.4: c. Archaeal Cell Membranes and Walls

We have already discussed that the lipids in Archaeal cell membranes contain L (instead of D) glycerol derivatives and that ether links (more stable in reactive environments) replace ester links with isoprenoid (sometimes branched) chains replacing fatty acid chains. The cell wall is quite different as well, and some don't have one. The type of cell wall is depending on the environmental need for stability. They don't contain peptidoglycans. Figure 7.3.22 shows four different types.

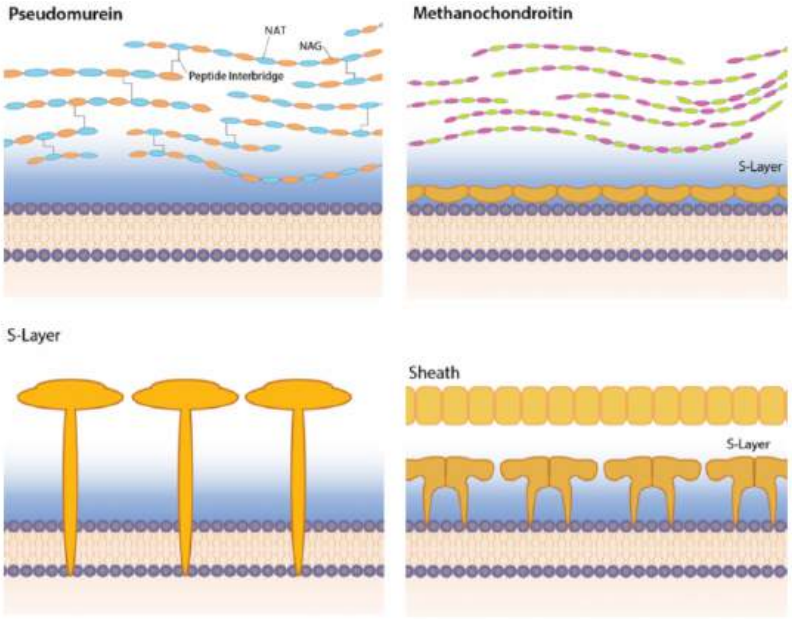


Figure 7.3.22:Archaeal cell walls https://open.oregonstate.education/g...apter/archaea/. https://creativecommons.org/licenses/by-nc/4.0/

Some differences include the presence of

- pseudomurein - This is the closest to the peptidoglycans presented above. Instead of repeating disaccharide units of (NAM-NAG)$_n$, they have a repeating disaccharide unit of N-acetylalosaminuronic acid (NAT)-NAG. The structure of NAT is shown in Figure 7.3.23

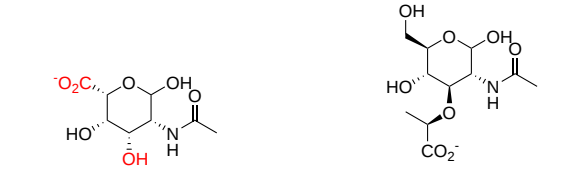


Figure 7.3.23: N-acetylalosaminuronic acid (NAT)compared to N-acetylmuramic acid (NAM)

- methanochondroitin - This is similar to the glycosaminoglycan chondrotin sulfate
- S-Layer
- Sheath/S-Layer

7.3.4.1.5: d. Plant Cell Wall

If you thought bacterial cell walls were complicated, wait until you see plant cell walls! There are about 35 different types of plant cells, and each may have a different cell wall depending on the local needs of a given cell. Cells synthesize thin cell wall that extends and stay thin as the cell grows.

Figure 7.3.24 shows the primary cell wall of plants. The primary cell wall contains cellulose microfibrils (no surprise) and two other polymers, pectin and hemicellulose. The middle lamella consisting of pectins is somewhat analogous to the extracellular matrix discussed above.

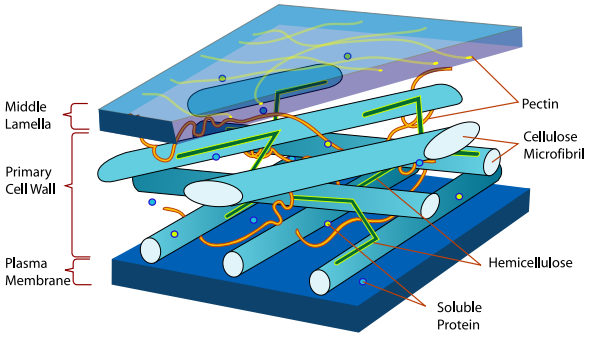


Figure 7.3.24: Primary cell wall of plants. https://commons.wikimedia.org/wiki/F...diagram-en.svg

After cell growth, the cell often synthesizes a secondary cell wall thicker than the first for extra rigidity. The enzymatic machinery for its synthesis is in the cytoplasm and the cell membrane. It is deposited between the cell membrane and the primary cell wall, as shown in the animated image shown in Figure 7.3.25.

Figure 7.3.25: Primary and secondary cell wall of plants

Figure 7.3.26 shows a structural representation of both the primary and secondary cell wall.

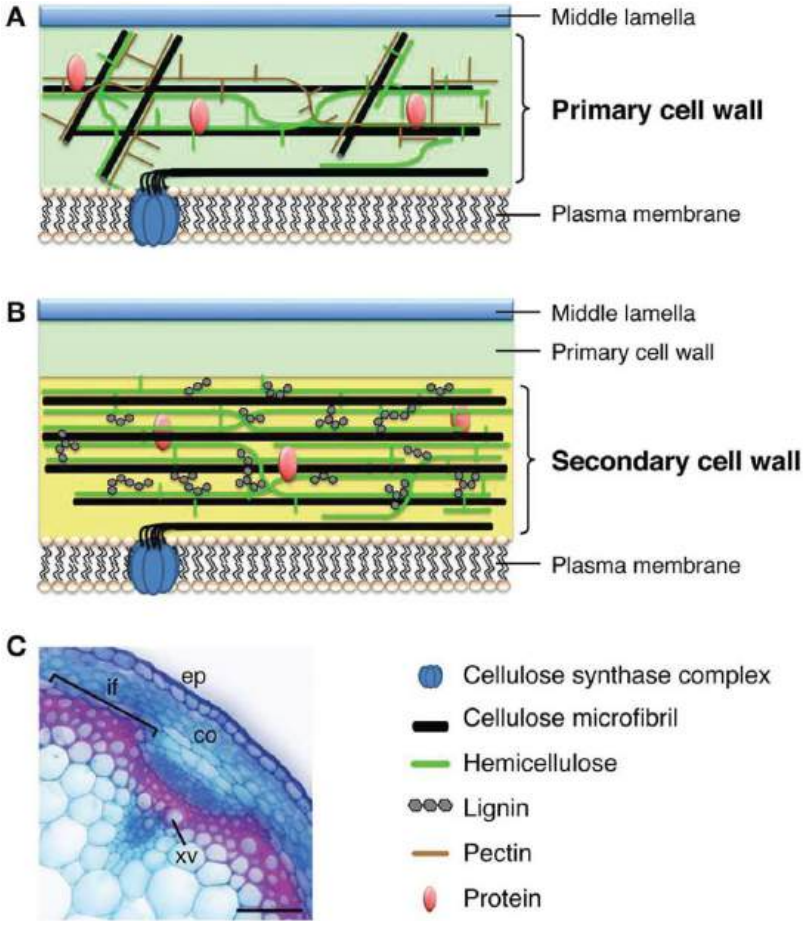


Figure 7.3.26: A structural representation of both the primary and secondary cell walls of plants. Nakano Yoshimi et al. Frontiers in Plant Science (6), 288 (2015) https://www.frontiersin.org/article/...015.00288/full. Creative Commons AttributionLicense (CC BY).

The middle lamella, which contains pectins, lignins and some proteins, helps "glue together" the primary cell walls of surrounding plants.

Primary Cell Wall:

The main component of the primary plant wall is the homopolymer cellulose (40% -60% mass) in which the glucose monomers are linked $\beta(1\rightarrow4)$-linked into strands that collect into microfibrils through hydrogen bond interactions. Two other groups of polymers,

hemicellulose and **pectin** make up the plant cell wall.

Hemicellulose can make up to 20-40% by the mass These polymers have β(1,4) backbones of glucose, mannose, or xylose (called xyloglucans, xylans, mannans, galactomannans, glucomannans, and galactoglucomanannans along with some β(1,3 and 1,4)-glucans. The most abundant hemicellulose in higher plants are the xyloglucans which have a cellulose backbone linked at O6 to α-D-xylose. **Pectin** consists of linked galacturonic acids forming homogalacturonans, rhamnogalacturonans, and rhamnogalacturonans II (RGII) [12] [13]. Homogalacturonans (α1→4) linked D-GalA making up more than 50% of the pectin

Figure 7.3.27 shows some variants of the cell wall components of a plant.

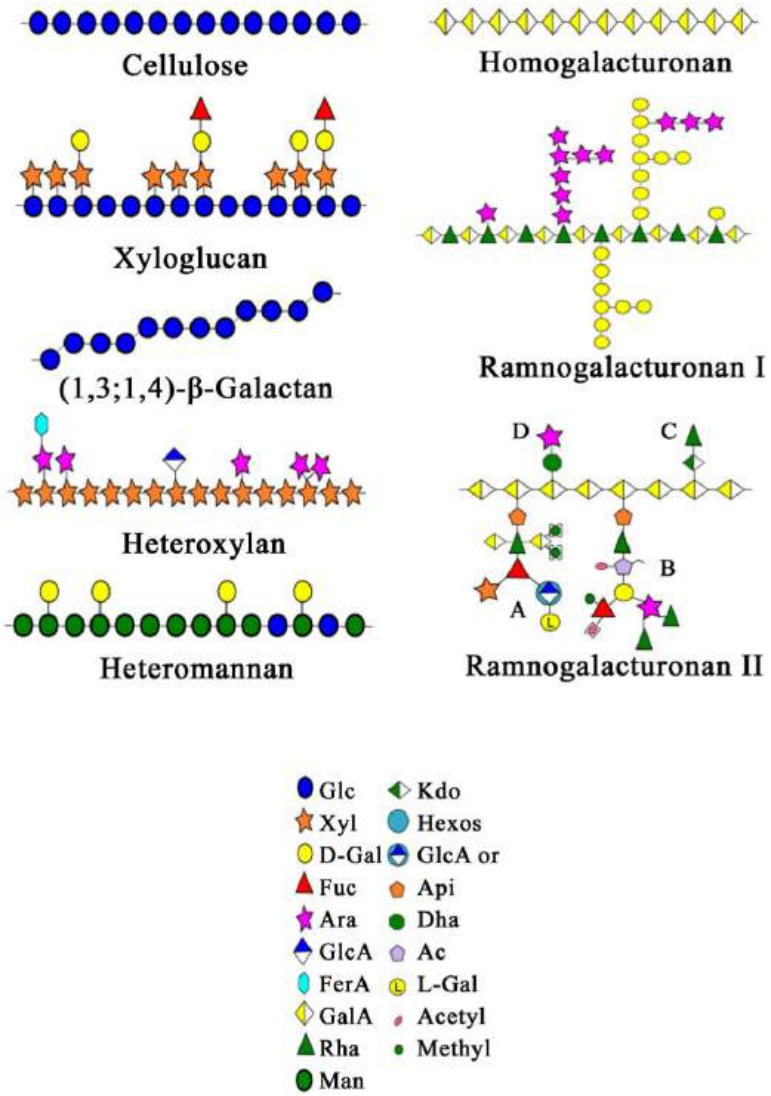


Figure 7.3.27: Variant of the cell wall components of a plant. Costa and Plazanet. Advances in Biological Chemistry 06(03):70-105. DOI: 10.4236/abc.2016.63008. License CC BY 4.0

Secondary Cell Wall

The structure of the secondary cell wall depends on the function and environment of the cell. It contains cellulose fibers, hemicellulose and in addition a new polymer, **lignin**. It is abundant in xylem vessels and fiber cells of woody plants. It gives the plant extra stability and new functions, including the transport of fluids within the plant through channels.

Lignins, which can make up to 25% of the biomass weight, are made from derivatives of phenylalanine, but more directly from cinnamic acid. This derives from is made from phenylalanine which is hydroxylated and converted through other steps to hydroxycinnamyl alcohols called monolignols as shown in Figure 7.3.28. Three common monomer (M) derivatives, p-coumaryl, coniferyl, and sinapyl alcohols can polymerize into lignins, with the units in the polymer (P) names hydroxyphenyl, guaiacyl and syringly, respectively.

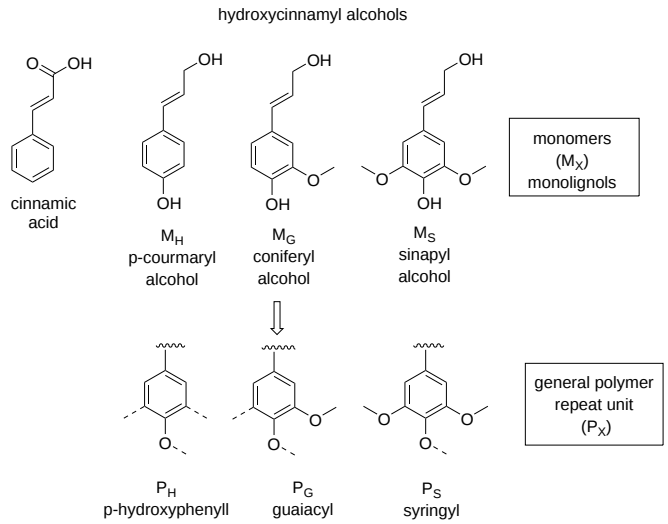


Figure 7.3.28: Monolignols and their polymers

Lignols are activated phenolic compounds, which form phenoxide free radicals (catalyzed by enzymes called peroxidases), which can attack other lignols to form covalent dimers. Reaction mechanisms for the dimerization of the M_S sinapyl alcohol free radical are shown as an example in Figure 7.3.29.

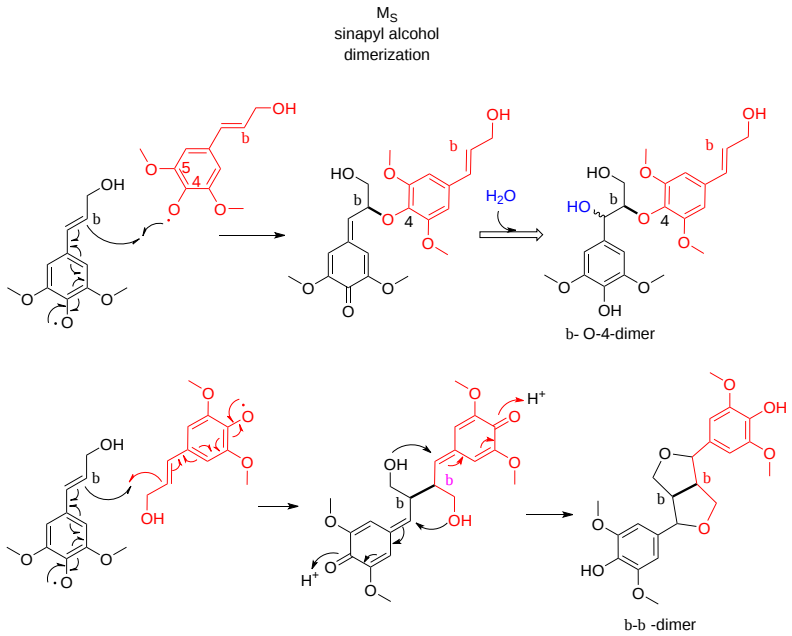


Figure 7.3.29: Dimers of lignols

Now imagine this polymerization continuing through the formation of additional phenolic free radicals and coupling at a myriad of sites to form a huge covalent lignin polymer. Figure 7.3.30 shows one example of a larger lignin.

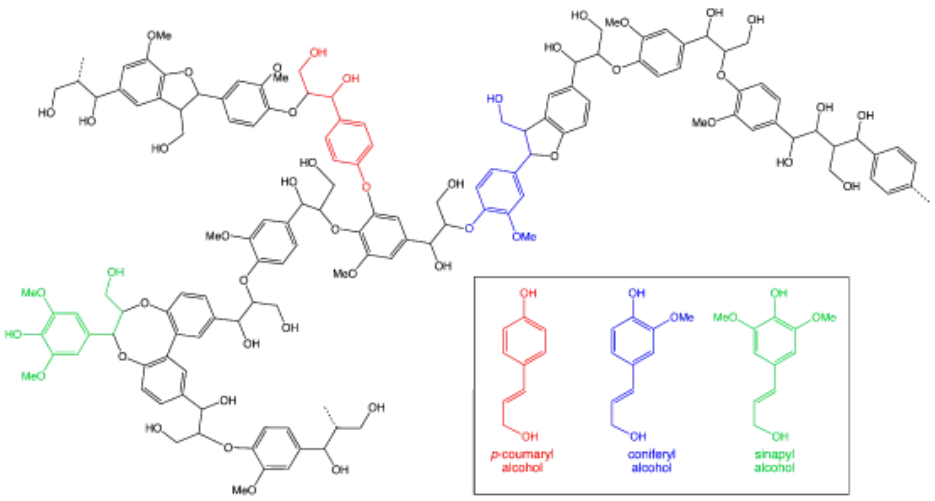


Figure 7.3.30: A larger lignin. https://commons.wikimedia.org/wiki/C...ile:Lignin.png . By Smokefoot - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=16022799

Finally, Figure 7.3.31 shows an image of a poplar tree cell wall, made using surface Raman scattering, showing lignin, cellulose, and lipids in secondary xylem cell walls.

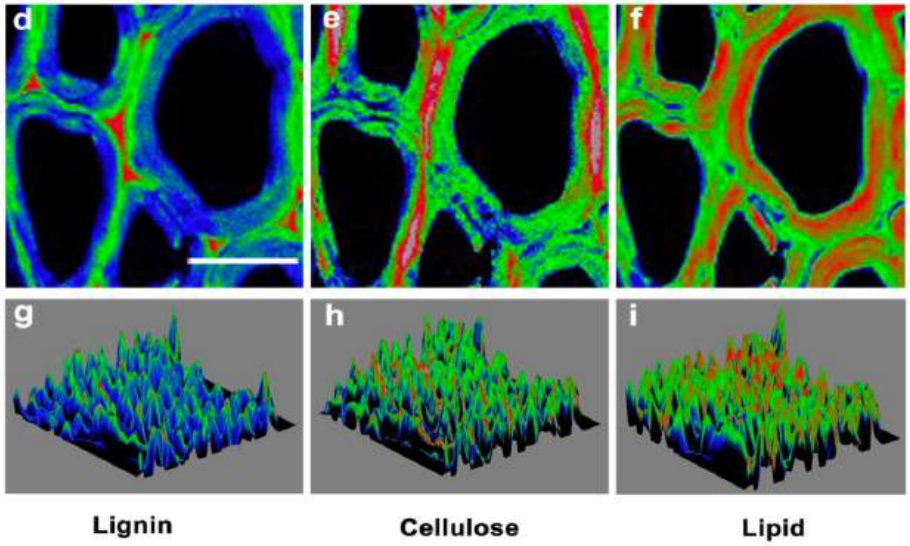


Figure 7.3.31: Surface Raman scattering (SRS) images of lignin, cellulose, and lipids in xylem cell walls. **d-e are** SRS images of lignin, cellulose, and lipids in the secondary xylem cell walls of poplar, respectively. The 3D surface plots are shown in **g–i**. Xu, H., *et al.* A label-free, fast and high-specificity technique for plant cell wall imaging and composition analysis. *Plant Methods* **17,** 29 (2021). https://doi.org/10.1186/s13007-021-00730-9. http://creativecommons.org/licenses/by/4.0/.

7.3.5: The Extracellular Matrix (ECM) and Basement Membranes

We won't formally discuss cell membranes until Chapter 11, but since anyone reading this book has previously seen biological membranes (including the Gram-negative and positive bilayers discussed above), let's explore a term that most chemistry students, but perhaps not biology students, will find very confusing. That topic is the basement membrane. The basement membrane is encountered often so often, that we will explore its overall structure here even though it is not a lipid bilayer. It fits well here since it is a complex structure consisting of proteins and proteoglycans. It's very amorphous which makes its structure difficult to those hoping for crystal structures or even complex bilayers. It is somewhat similar to the cell wall in functionality. We will offer a cursory explanation. For a great overall introduction, please visit Introduction to Extracellular Matrix and Cell Adhesion in BioLibre texts. Some of the images (when noted) below come from that Cell Biology book chapter.

The extracellular matrix (ECM) is a general term for the large protein and polysaccharide network formed on secretion by some cells in a multicellular organism. They act as connective material to hold cells in a defined space. Cell density can vary greatly

between different tissues of an animal, from tightly-packed muscle cells with many direct cell-to-cell contacts to liver tissue, in which some of the cells are only loosely organized, suspended in a web of extracellular matrix, shown in Figure 7.3.32

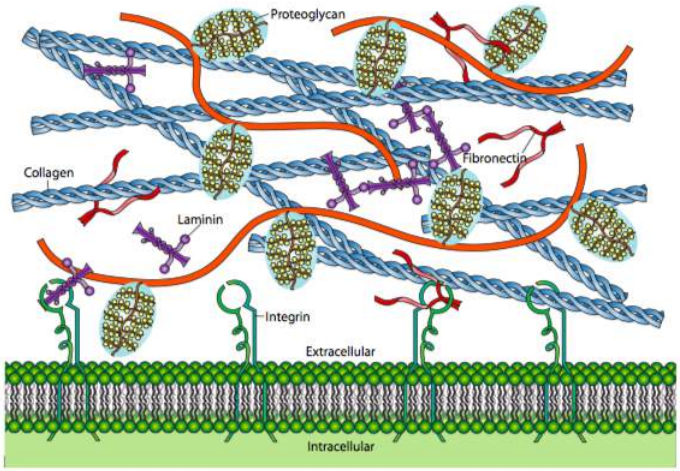


Figure 7.3.32: Extracellular matrix (ECM). Typical components include collagen, proteoglycans (with hydration shell depicted around sugars), bronectin, and laminin. The cellular receptors for a number of these ECM components are integrins, although the exact integrin αβ *pair may differ.*

The ECM is a generic term encompassing mixtures of polysaccharides and proteins, including collagens, bronectins, laminins, and proteoglycans, all secreted by the cell. The proportions of these components can vary greatly depending on tissue type. Two, quite different, examples of ECM are the basement membrane underlying the epidermis of the skin, a thin, almost two-dimensional layer that helps to organize the skin cells into a nearly-impenetrable barrier to most simple biological insults, and the massive three-dimensional matrix surrounding each chondrocyte in cartilaginous tissue. The ability of the cartilage in your knee to withstand the repeated shock of your footsteps is due to the ECM proteins in which the cells are embedded, not to the cells that are actually rather few in number and sparsely distributed. Although both types of ECM share some components in common, they are distinguishable not just in function or appearance, but in the proportions and identity of the constituent molecules

Figure 7.3.33 shows a general structure of the basement membrane. Think of it as an amorphous polymer mixture (somewhat similar to a polyacrylamide gel).

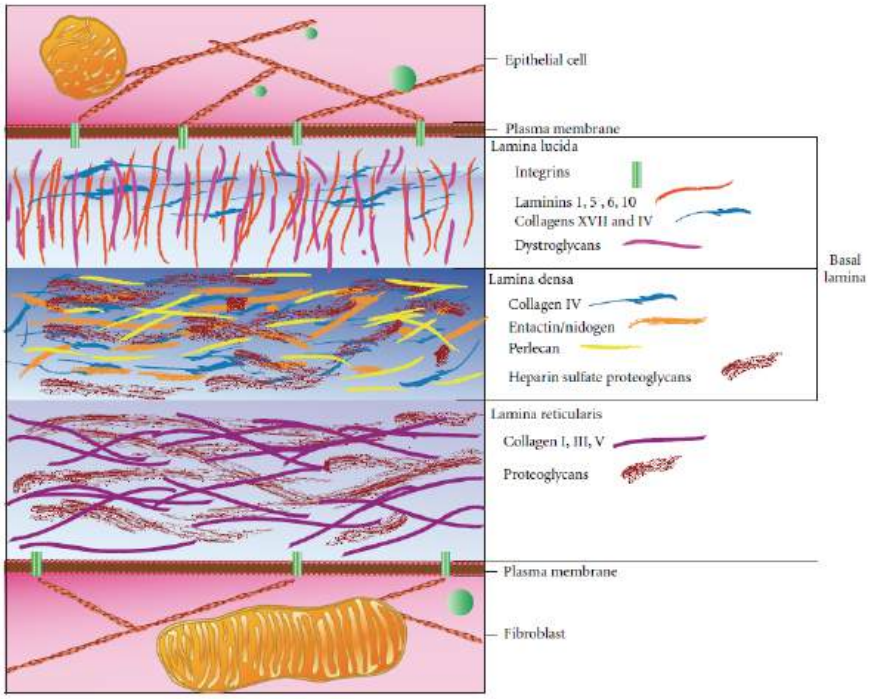


Figure 7.3.33: General structure of the basement membrane. Mentor and DuBois. Journal of Cell Biology · February 2012. DOI: 10.1155/2012/723419 · Source: PubMedCreative Commons License. https://creativecommons.org/licenses/by/3.0/

7.4: The Sugar Code and Lectin Decoding

7.4.1: Introduction

By now, you should be convinced that the structures of glycans are extraordinarily complex and, in many ways, much more complicated than proteins and nucleic acids. Their structure diversity is staggering, given the number of different sugar monomers, stereocenters, linkages, lengths, conformers, dynamic flexibility and chemical modifications. Yet evolution has allowed this astronomical diversity, which must serve more than just simple functions such as protection of proteins from degradation, to pick one example. Much of the diversity derives from the lack of an equivalent genetic code for glycan synthesis.

Since all events in biology start with a binding interaction, let's ponder the binding interaction of glycans with partner "ligands" such as proteins, lipids, nucleic acids, etc. A binding site on a glycan could be a single monosaccharide to a much larger and more complex interface. Figure 7.4.1 shows an interactive iCn3D model of one of the few glycoproteins with pdb coordinates, the unliganded simian immunodeficiency virus (SIV) gp120 core glycoprotein (3fus).

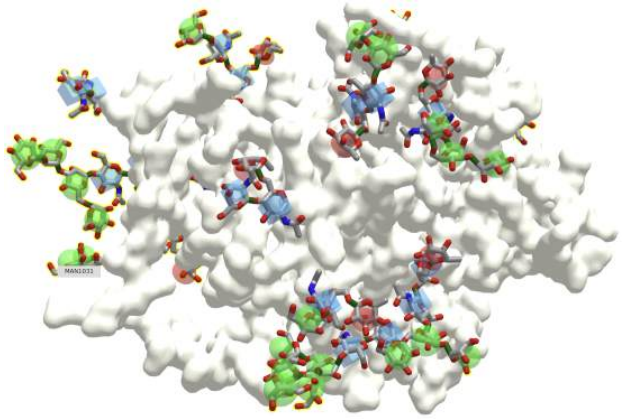


Figure 7.4.1: Simian immunodeficiency virus (SIV) gp120 core glycoprotein (3fus). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?DzZYhLp8JL1vPYYw6

The protein surface is shown in ivory and the glycans are shown in color sticks with the correct symbolic color-coded spheres or cubes around them.

Now let's convert in our imagination an image file showing one face of the protein to a black and white QR code as shown in Figure 7.4.2.

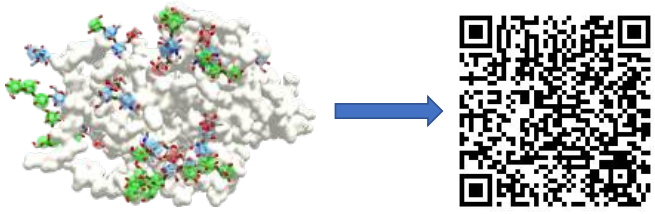


Figure 7.4.2: An imaginary QR code for the surface of a glycoprotein.

Computers can recognize information encoded in the QR codes and decode them into another form of information, such as a menu at a restaurant. Likewise, organisms have evolved "readers" to decode the **glycan code** written by enzymes (glycan synthases, hydrolases, and modifying enzymes). The glycan code is written onto the 3D-surfaces of polysaccharides, glycoproteins, glycolipids, and proteoglycans. It should be no surprise that the biological readers of the glycan code are mostly proteins, which locate and bind to the correct "QR" code displayed on the surface of glycan.

Luckily, the QR code metaphor for the glycan code is a bit exaggerated since the readers of the glycan code, glycan-binding proteins, seem to recognize just small sections of a glycan. They can be compared to protein antibodies that bind to foreign molecules such as proteins. The binding site on a foreign protein recognized by an antibody is called an **epitope**. Epitopes can be continuous (linear) stretches of the foreign protein sequence, or discontinuous (conformational), made of some continuous stretches of amino acid and some further away in the sequence but close in the 3D folded protein. The average continuous epitope is often 5-6 amino acids long. Yet that might be an underestimation since an analysis of all contact residues (within a conservative 4 A°

distance) for target proteins and their bound antibodies found in the Protein Data Bank is around 18-19 amino each (Stave and Lindpaintner). Glycan-binding proteins presumably also bind a mixture of continuous and discontinuous glycan sequences. Linear one would be much easier to determine and study.

Now, let's explore the family of these glycan-binding proteins (GBP), the readers of the glycan code.

7.4.2: Glycan-Binding Proteins (GBPs)

There appear to be nine types of glycan-binding proteins (GBPs). These include nonenveloped capsid virus GBPs,enveloped-virus GBPs (ex. influenza and coronaviruses), eukaryotic microbial GBPs (ex yeast), and bacterial toxin GBPs (ex botulinum toxin). Bacterial **adhesins** (parts of organelles like flagella), **lectins** (soluble proteins) and lectin domain-containing proteins are also examples of **glycan-binding proteins (GBP).** We will discuss in more detail three other types, **C-type lectins, galectins, and siglecs**.

In the broadest sense, if a **lectin** is a protein that binds a specific carbohydrate motif (i.e a glycan code) without modifying the motif, then any glycan-binding protein could be called a lectin. This would exclude enzymes that synthesize, degrade or modify glycans, as well as antibodies that would recognize foreign or self glycan sequences. Table 7.4.1 below shows some lectins and their target glycan ligand from plants, animals, viruses and bacteria.

Lectin Family/Lectin	Abbreviation	Ligand(s)
Plants		
Concanavalin A	ConA	Man α1- OCH3
Griffonia simplicifolia lectin 4	GS4	Lewis b (Leb) tetrasaccharide
Wheat germ agglutinin	WGA	Ner5Ac(α 2,3)Gal(β 1,4)GlcGlcNAc(β1,4)GlcNAc
Ricin		Gal(β 1,4)Glc
Animals		
Galectin-1		Gal(β 1,4)Glc
Mannose-binding protein	MBP-A	High Mannose Octasaccahride
Viral		
Influenza Virus hemagglutinin	HA	Neu5Ac(α 2,6)Gal(β1,4)Glc
Polyoma virus protein 1	VP1	Neu5Ac(α 2,3)Gal(β1,4)Glc
Bacterial		
Enterotoxin	LT	Gal
Cholera toxin	CT	GM1 pentasaccharide

Table 7.4.1: Lectin families and their ligands.

In animals, lectins facilitate cell-cell interactions by forming multiple, but weak interactions (also called **multivalent interactions**) between the protein and many sugars on the ligand to which it binds.

Now let's consider the other three classes of glycan-binding proteins (or lectins), **C-type lectins, galectins,** and **siglecs** in more detail. Focus on the very different structures of the carbohydrate-binding domains.

7.4.3: C-Type Lectins

C-Type lectins comprise the largest number of glycan-binding proteins. These proteins have a glycan or carbohydrate recognition domain that depends on the Ca^{2+} ion. They bind self glycans as well as those on pathogens, which can target viruses to specific cells. Many are on the surface of immune cells. They have an N-terminal glycan-binding domain, also called a C-lectin (CLECT) domain or a carbohydrate recognition domain (CRD). However, some proteins with the domain do not appear to bind either Ca^{2+} or glycans. They serve as adhesion molecules and are also involved in cell signaling. Some residues in the lectin binding domain appear critical for binding lectins. These include an EPN motif, which interacts with Man, GlcNAc, Fuc, and Glc) and a WND motif involved in binding to Gal and GalNAc.

Let's look now at one example of a C-type lectin, the **selectins**.

P-Selectins

These are involved in the interaction of immune cells in the blood with endothelial cells that line the blood vessel wall. Think of the challenges an immune cell faces as it moves from the blood into a tissue where an infection might occur! Blood flow in vessels at a rate inversely proportional to the total cross-sectional area of the blood vessel. That rate is about 5-20 cm/sec in arteries, 1.5-7 cm/sec in veins and about 1 mm/sec (1000 μm)/sec) in capillaries. Assuming an average diameter of 10 μm for a lymphocyte, the cell would move about 100 cell lengths per second. An equivalent speed for a human with an arm span of 6 feet (approximately fitting into a circle of diameter = 6 feet as drawn by Leonardo da Vinci) would be around 600 feet/second. The cell must go from its typical circulating speeds to a stop before it can move through the blood vessel wall into tissues. Nature has solved this by providing a way to slow down the moving cell until its final capture. The cells roll along the endothelial cells, making transient low-affinity interactions, which slow it down enough until high-affinity ones effectively stop it (unless it dissociates first).

Also, you wouldn't want immune cells to stop and move into tissue in the absence of an infection signaled by mediator molecules. Another problem solved! P-selections are stored in the intracellular granules of platelets (the source of the name P-selectin) and endothelial cells so moving immune cells are not spuriously captured in the absence of some signal. In the presence of the right chemical signal, endothelial and platelets get active and P-selectin is transported rapidly to the cell surface where "capture" occurs before the cell can move into the underlying tissue. P-selectin mediates the first transient interactions and subsequent rolling of immune cells on activated platelets and endothelial cells.

Figure 7.4.3 is a video animating the rolling and "capture" of a lymphocyte by endothelial cells. (See the video for the reference.) Note that cancer cells also can move through the endothelial cells of blood vessels in the process of forming metastases.

Figure 7.4.3: Video animation of a lymphyocyte rolling and being captured by endothelial cells lining blood vessel walls

P-selectins hence are receptors for molecules on immune cells. They bind Ca^{2+} ions, which helps create an active conformation. Their binding ligands are glycan codes and nearby sections of protein connected to the glycan. The glycan ligand on the surface of a circulating immune cell is the **sialyl-Lewis X (SiaLew^X)** glycan or a derivative of it. One of the immune membrane proteins with SiaLew^X is the **P-selectin glycoprotein ligand 1** (**PSGL-1**, the gene name), also called **SELPLG**. It mediates rapid rolling of leukocyte rolling over vascular surfaces during the initial steps in inflammation through interaction with SELPLG"

P-selectin is a mediator of cell adhesion (to other cells). As such it could also be classified as an adhesion protein. The three main types of selectins:

1. L-selectins: found on leukocytes ("white" blood cells that are circulating immune cells).
2. P-selectins: found on activated platelets (which can aggregate to form a type of blood clot) and activated endothelial cells. Activation occurs during the inflammatory response which can lead to the quick movement of pre-formed selectins stored within the cytoplasm to the membrane. In addition, their expression can be induced.

3. E-selectins: found on activated endothelial cells only after the cells have been induced to form them by certain immune hormones called cytokines released by immune cells during an inflammatory response.

Figure 7.4.4 shows the domain structure of human P-selectin.

Figure 7.4.4: Domain structure of human P-selectin (from https://smart.embl.de/)

It contains an N-terminal C-Lectin (CLECT) domain, which is also called the carbohydrate-recognition domain (CRDs) or the C-type lectin domain (CTLD). In addition, it has an epidermal growth factor domain (EGF), 9 complement control protein (CCP) domains and the blue transmembrane domain.

Figure 7.4.5 shows the structure of the $SLew^x$ glycan along with its symbol nomenclature for glycans (SNFG) representation.

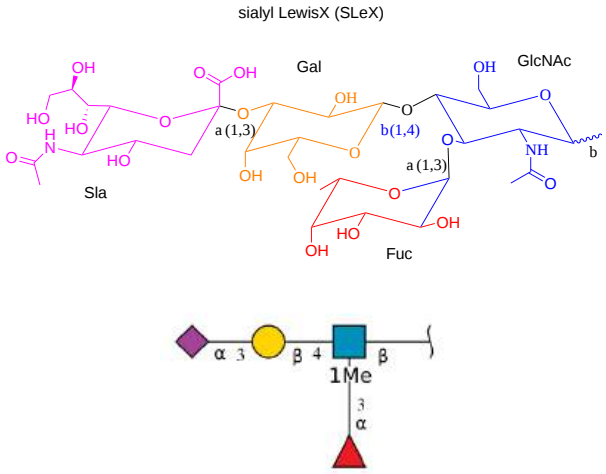


Figure 7.4.5: Structure of the $SLew^x$ glycan

Figure 7.4.6 shows an interactive iCn3D model of the crystal structure of P-selectin lectin/EGF domains complexed with SLeX (1g1r) to which P-selectin binds with weak affinity. Fucose interacts with the Ca^{2+} ion. The glycan interacts with the CLECT domain.

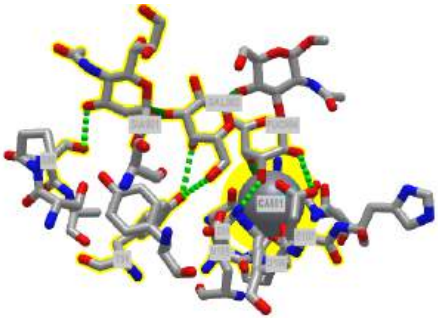


Figure 7.4.6: P-selectin lectin/EGF domains complexed with SLeX (1g1r). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i..s1GYbAfXfnuSz9

$SLew^x$ is not present in isolation but rather attached to a membrane protein on an immune cell, which serves as a ligand for the P-selectin on activated endothelial cells or platelet. (The $SLew^x$ can also be part of a glycolipid.) Now let's contrast the interactions of the P-selectin LE domain with "naked" $SLew^x$ with those present between P-selectin LE with a higher affinity natural binding ligand, human P-selectin glycoprotein ligand 1 (PSGL-1), an immune cell integral membrane protein. PSGL-1 is expressed on neutrophils, monocytes and most lymphocytes. The P-selectin:PSGL-1 complex has a much lower K_D (higher affinity compared to binding of the unmodified SLe^X (1g1s). PSGL-1 is a disulfide-linked homodimer. When sulfated on a specific Tyr (48) the protein displays high affinity for P-selectin. In contrast, when sulfated on a different Tyr (51), it displays high affinity for L-selectin instead

The $SLex^X$ type glycan O-linked to the peptide is a bit more complicated than the simple $SLex^X$ ligand as it is connected to a protein through an O-linked bond at a threonine. The SNFG is shown in Figure 7.4.7.

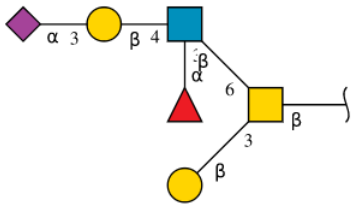


Figure 7.4.7: SNFG if $SLex^X$ type glycan O-linked to a peptide

The crystal structure of a trisulfated, $SLew^x$-modified peptide from the N terminal region of PSGL-1 (1G1S) bound to P-selectin lectin and EGF domains (P-LE) has been solved. Figure 7.4.8 shows an interactive iCn3D model which shows some of the interactions between the PSGL-1 peptide (green backbone) and P-LE (magenta backbone).

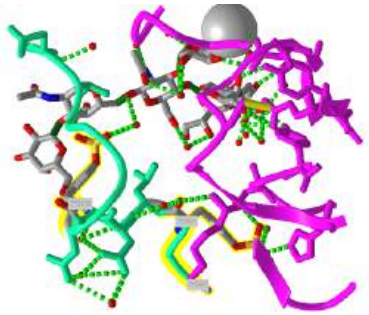


Figure 7.4.8: P-selectin Lectin/EGF domains and bound PSGL-1 peptide (1G1S). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i..E2mVULZc8WqVR9

In the crystal structure, the peptide from the P-selectin ligand (which again is a membrane protein) contains 3 sulfated tyrosine residues (605, 607, and 610) which correspond to amino acids 5, 7, and 10 in the peptide). No electron density for the side chain of Tyr 605 was seen. Tyr 607 binds through multiple interactions to the P-selectin LE domain, and is most likely responsible for the high-affinity interaction of P-selectin with the P-selectin glycoprotein ligand (again represented by the green chain). In contrast, Tyr 610 interacts through an intermediary water molecule with the glycan $SLex^X$ of the peptide.

Figure 7.4.9 shows the electrostatic surface potential map of one of the dimers of the P-select LE domains. Blue represents positive potential and red negative. The backbone of the P-selectin ligand peptide is shown in green with all of the negatively charged side chains (Tyr, Asp and Glu) shown in stick with CPK colors. Note that these amino acids are all bound in blue (positive) regions of P-selection. The glycan portion attached to the peptide (stick, CPK color) is positioned mostly over negative potential, allowing hydrogen bonding between the hydrogen bond donors of sugar OHs with the protein.

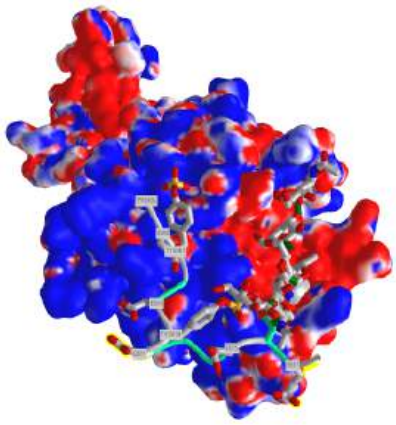


Figure 7.4.9: Electrostatic surface potential map of one monomer of the P-selectin LE domains with bound P-selectin ligand peptide.

You could surmise that the blue region of positive potential could also bind other strongly negatively charged ligands (such as heparin and other glycosaminoglycans) which could inhibit the function of this protein as it would prevent binding of the PSGL-1.

Figure 7.4.10 shows an interactive iCn3D model which shows the surface electrostatic potential of the P-selectin Lectin/EGF domains and bound PSGL-1 peptide

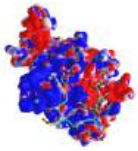


Figure 7.4.10: Electrostatic potential of the P-selectin Lectin/EGF domains and bound PSGL-1 peptide (1G1S). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i..X2CSQQzgbFn2VA

The blue represents positive potential and the red negative. The backbone of the P-selectin ligand peptide is shown in green with all of the negatively charged side chains (Tys, Asp and Glu) shown in stick with CPK colors.

There are also nonpolar interactions not shown in the figure and model above. The aromatic ring of Tyr 607 (7) interacts with the nonpolar parts of a Ser ($-CH_2$) and Lys $(-CH_2)_4$ side chains and the ring of Tys 610 (10) interacts with two leucine side chains.

The selectins are also part of a class of molecules called **adhesion** molecules. As described for the selectins, adhesion molecules contain

- an extracellular CHO binding domain (the lectin domain), which mediates binding to adjacent cells or to the extracellular matrix;
- a transmembrane domain;
- and a cytoplasmic domain which often interacts with the cytoskeleton within the cell.

This initial binding mediated by selectin-CHO interactions activates the expression of another adhesion molecule on the leukocyte, **integrin**, a heterodimer with an a and b chain. These cause strong leukocyte-endothelial cell interactions, leading to the movement of the leukocytes through the vessel wall. Other classes of adhesion molecules (in addition to selectins and integrins) are **cadherins** (calcium-dependent adhesion molecules), and the immunoglobulin-like superfamily (ICAM1, ICAM2, VCAM). VCAM (Vascular Adhesion Molecule) binds to integrin expressed on activated lymphocytes, leading to the passage of the lymphocyte from the lumen of the vessel into the tissues. Integrins appear to bind proteins in the extracellular matrix through RGD (Arg-Gly-Asp) and also through LDV (Leu-Asp-Val) motifs on the proteins, including fibronectin (RGD), thrombospondin (RGD & LDV), fibrinogen (RGD & LDV), van Willebrand Factor (RGD), vitronectin (RGD). They also bind other matrix proteins with an "alpha domain" including collagen and laminin. Integrin/Adhesion molecule interactions involve protein/protein interactions.

A fertilized egg (in the blastocyst stage which is ready for implantation in the uterine cell wall) express L-selectin which allows a low affinity (rolling-type) interaction of the fertilized egg with the uterine **epithelial** cells. These cells expressed the CHO ligands on their surface which bind to the L-selectin on the blastocyst. The CHO ligands are only transiently expressed on the surface of the epithelial cells of the uterus, presumably only when the uterus is primed for implantation. After the initial interaction of the blastocyst and epithelial cells, further expression of integrins on the blastocyst surface might result. Problems in any of these molecular steps could result in infertility. Figure 7.4.11 shows. endothelial cell/leukocyte Interactions mediated through selectins, integrins, and ICAMs.

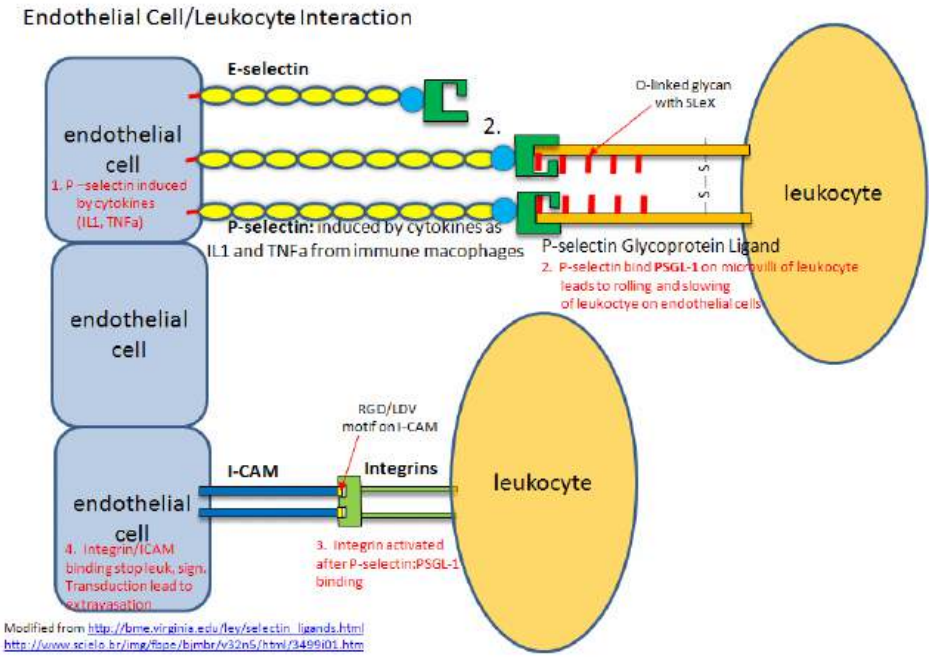


Figure 7.4.11: Endothelial cell/leukocyte Interactions mediated through selectins, integrins, and ICAMs

Post-translational modifications of protein modification (like glycosylation) can confer new binding and biological functions to a protein. Site-directed mutagenesis can be used to replace surface amino acids with cysteine or methionine with nonnatural amino acid analogs that contain azide or alkyne groups. These modified groups could then direct the location of chemical modifying reagents (such as sugars) to these sites. A protein completely unrelated to PSGL-1 has been selectively modified using this approach to contain covalently attached glycans and sulfated tyrosine side chain. The unrelated protein bound to P-selectin.

7.4.4: Mannose Receptor.

What do you do with a protein that no longer has the correct structure(s) to perform its designed function(s). Proteins, as with any molecule, undergo chemical changes during their biological lifetime. They must be recognized as aberrant and then removed from "service", ultimately being degraded into component amino acids for reuse. There are no repair enzymes for proteins as for DNA. One modification that changes glycoproteins and signals the need for their removal is the removal of terminal sialic acid residue, forming asialoglycoproteins, whose glycans end in galactose, as you can envision from Figure 7.4.12 which shows a typical structure of a N-linked glycoprotein.

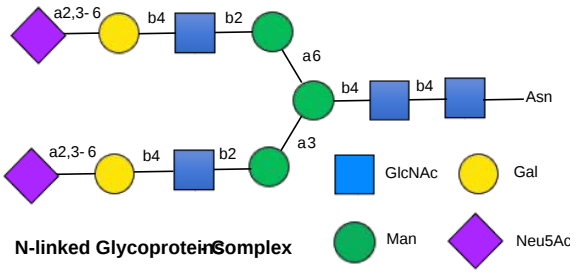


Figure 7.4.12: Typical structure of a N-linked glycoprotein

The **asialoglycoprotein receptor**, a member of the C-Type lectin family, is a transmembrane protein, which binds terminal galactose and N-acetylgalactosamine sugars on the end of circulating asialoglycoproteins, leading to their endocytosis into the cell. It is expressed on the surface of hepatocytes (liver cells). Receptors of this type are also called scavenger receptors as they remove proteins from circulation.

The **mannose receptor** (also called CD206), also expressed in liver endothelial cells, is another C-Type lectin involved in binding and removal of glycoproteins from the circulation. It binds both sulfated and non-sulfated glycans. It also is a receptor that allows binding and phagocytosis of bacterial and fungal pathogens by a type of immune cells called macrophages and dendritic cells. Unfortunately, tumor cells can use the same process for uptake into macrophages, leading to the promotion of tumor cell growth. The protein binds and scavenges sulfated glycoprotein hormones, mannose-bearing glycoproteins released during inflammation, lysosomal enzymes released from cells on injury and fragments of collagen.

Figure 7.4.13 shows the domain structure of the human mannose receptor.

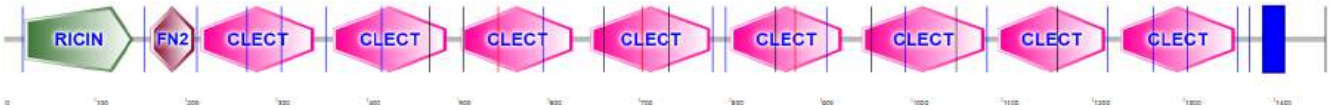


Figure 7.4.13: Domain structure of the human mannose receptor.

Given the large number of CLECT domains, you might surmise that this protein could bind a number of different target glycans from both self and pathogens. What is different about the domain structure compared to P-selectin is the presence of an N-terminal Ricin and a Fibronectin type 2 (FN2) domain. The FN2 domain has two cystines from the 4 conserved cysteines involved in disulfide bonds. What's so interesting about the mannose receptor is that it binds glycans both in the CLECT domains and in the FN2 domain.

Glycan binding at the CLECT domain: The CLECT domain binds targets containing mannose, fucose and N-acetylglucosamine with a preference for Man(α1,2)Man or fucose. Figure 7.4.14 shows an interactive iCn3D model. of the **CLECT** 4 domain of the mannose receptor complexed with Man(α1,2)Man (7jue). Interactions of fucose lin ligands such as Lewis-a-trisaccharide strengthen the binding.

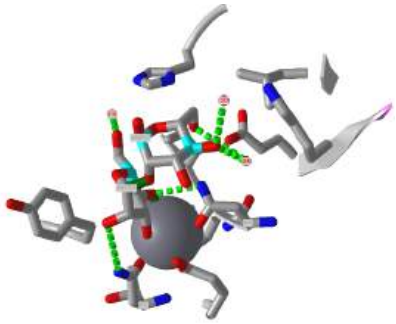


Figure 7.4.14: **CLECT** 4 domain of the mannose receptor complexed with Man(α1,2)Man (7jue). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i..jhn5eMLMrG6mX7

The receptor can bind a variety of glycans. Both mannose and N-acetylglucosamine interact with bound Ca^{2+} through equatorial OHs on carbon 3 and 4 of the ring while fucose uses OHs on carbon 2 and 3, or 3 and 4.

The interaction with fungal pathogens is obviously medically important. Fungi like yeast have an outer structure composed of a membrane bilayer and a mixture of glycans, which deploys an incredibly complex "glycan code" to host infected by them, as illustrated in Figure 7.4.15.

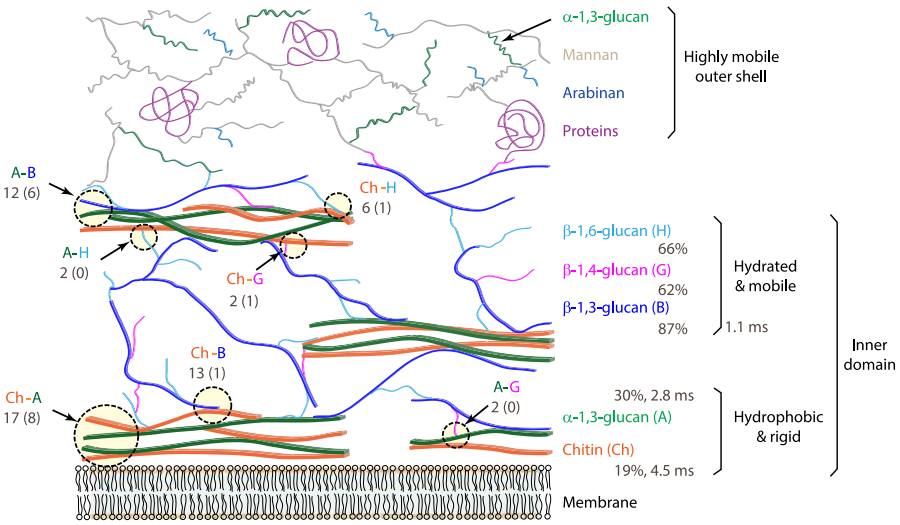


Figure 7.4.15: Structure of fungal cell wall. Kang, X., Kirui, A., Muszyński, A. *et al.* Molecular architecture of fungal cell walls revealed by solid-state NMR. *Nat Commun* **9,** 2747 (2018). https://doi.org/10.1038/s41467-018-05199-0. Creative Commons Attribution 4.0 International License: http://creativecommons.org/licenses/by/4.0/

Mannans, polymers of just mannose, differ widely in structure. Their main backbone can be DMan(α-1,6)DMan or DMan(β-1,4)DMan with many branches.

Glcyan binding at the FN2 (Cysteine-Rich) Domain (1FWU): The mannose receptor can also bind non-mannose sulfated glycans, such as 3-SO4-LEWIS(X), for which the SNFG representation is shown in Figure 7.4.16.

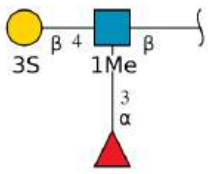


Figure 7.4.16: 3-SO4-LEWIS(X) non-mannose sulfated glycans for mannose receptor

The mannose receptor binds this glycan, which does not even contain mannose, through the FN2 domain (which contains four disulfide bonds) and not through the CLECT calcium-dependent carbohydrate-binding domain. Hence the protein can bind both sulfated and nonsulfated glycans.

Figure 7.4.17 shows an interactive iCn3D model of the complex of the FN2 domain of the mannose receptor with the non-mannose containing 3-SO4-LEWIS(X) glycan (1fwu).

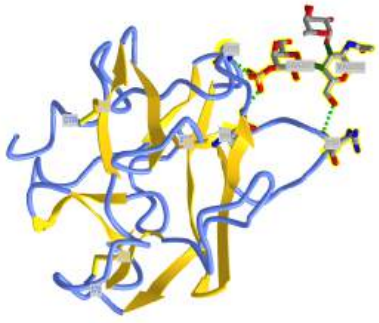


Figure 7.4.17: FN2 domain of the mannose receptor with the non-mannose containing 3-SO4-LEWIS(X) glycan (1fwu). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i..kXn2Rz8qMzebo9

Look at the number of CLECT domains in the domain structure diagram for the mannose receptor above. Along with interactions of sulfated glycans at the FN2 domain, these would enable the binding of widely diverse glycan structures. Reported ligands for the

mannose receptor include those with high mannose content released during inflammation (lysosomal hydrolases, collagen peptides, and tissue plasminogen activator), and sulfated ones (including the pituitary hormones lutropin and thyrotropin.

7.4.5: Galectins

This family of glycan-binding proteins contains a common carbohydrate recognition domain (CRD) of about 130 amino acids, which bind Galβ1,3GlcNAc or Galβ1,4GlcNAc disaccharides (hence the name galectins) as well as other glycan motifs . They are expressed in almost all cells and multicelluar organisms. There are 15 different types, grouped together in how the CRD is functionally expressed (as dimers, tandem repeats, or chimeras), as illustrated in Figure 7.4.18. The figure also shows their role in cancer biology.

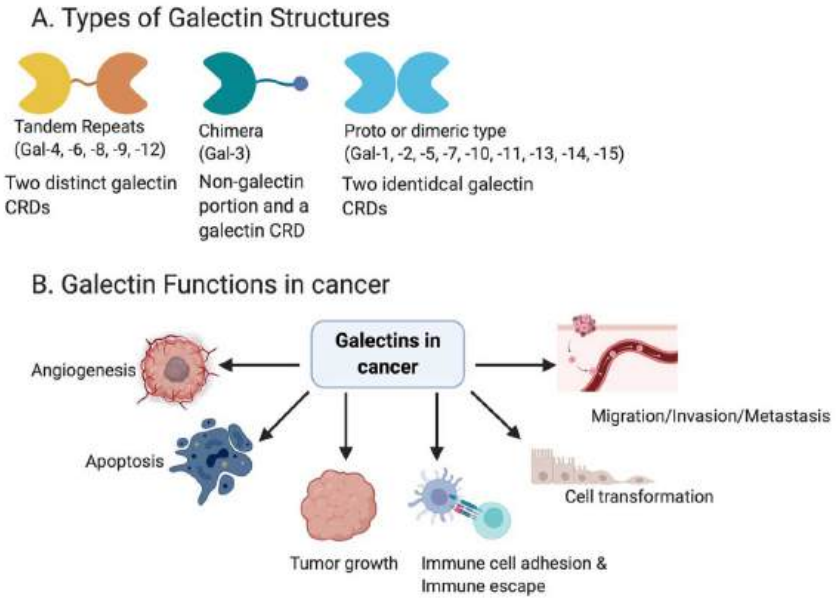


Figure 7.4.18: Galnectins structure and function Shimada, C.; Xu, R.; Al-Alem, L.; Stasenko, M.; Spriggs, D.R.; Rueda, B.R. Galectins and Ovarian Cancer. *Cancers* **2020**, *12*, 1421. https://doi.org/10.3390/cancers12061421. **Creative Commons Attribution License**

The carbohydrate-binding domain of the galactins has a jellyroll-like protein architecture with two anti-parallel β-sheets forming a β-sandwich.

Galectin I

This protein is secreted and is found in the extracellular matrix, as well as in the cytoplasm. It induces apoptosis in T-cells. It binds beta-galactosides as well as other glycans. The main ligand of galectin-1 has a Galβ1-4GlcNAc (or LacNAc) structure. Figure 7.4.19 shows an interactive iCn3D model of Human Galectin-1 in Complex with Type 1 N-acetyllactosamine (Gal(β1,3)GlcNAc), which binds less tightly than Galβ1,4GlcNAc (Type 2)(4XBL)

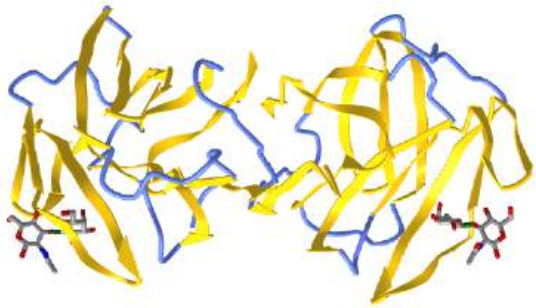


Figure 7.4.19: Human Galectin-1 in Complex with Type 1 N-acetyllactosamine (Gal(β1,3)GlcNAc) (4XBL). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i..MTyV7QvPKqqibA

A comparison of the crystal structures shows different phi/psi angles for bound Type I (135°) versus the more tightly bound Type 2 (-108°), which shows the nuance in binding conformations in the interactions of glycans with glycan-binding proteins.

7.4.6: Siglecs

The proteins are sialic acid-binding immunoglobulin (Ig)-like lectins found on immune cells like basophils, macrophages, mast cells and eosinophils. One type (Siglec-4)is found in myelinated structures in the central and peripheral nervous systems. They all have an N-terminal extracellular immunoglobulin domain (abbreviated as IG or V-Set) and a differing number of IG-like domains, also called C2-set Ig domains. The glycan binding epitope recognized by Siglecs are sialylated oligosaccharides on a section of the protein containing a conserved arginine. Figure 7.4.20compares the domain structures of the human Siglec family.

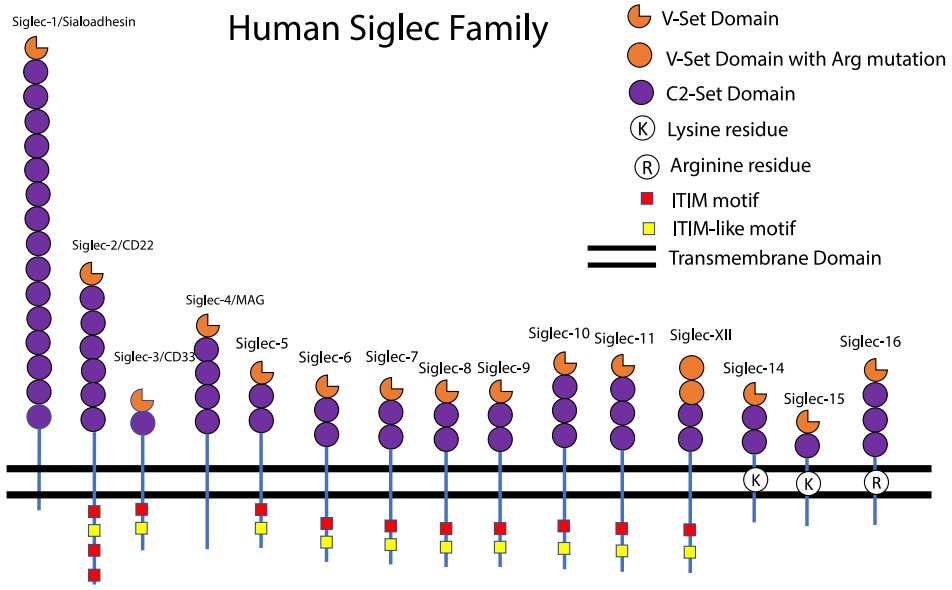


Figure 7.4.20: Domain structures of the human Siglec family Siddiqui, S.S.; Matar, R.; Merheb, M.; Hodeify, R.; Vazhappilly, C.G.; Marton, J.; Shamsuddin, S.A.; Al Zouabi, H. Siglecs in Brain Function and Neurological Disorders. *Cells* **2019**, *8*, 1125. https://doi.org/10.3390/cells8101125. Open access article distributed under the Creative Commons Attribution License

Here is one example of a Siglec.

Siglec-8

This protein is expressed on immune cells like basophil, mast cells and eosinophils. When activated by infection and prolonged inflammation, they release the contents of intracellular granules which have potent physiological effects that can lead to allergic and asthmatic responses. On infection and other inflammatory states, immune cytokines are released that in a signaling process lead to the release of sialoglycans that act as ligands, binding to the Siglec-8 on the surface of the immune cells. One type of sialoglycan released is mucins, which are very large glycoproteins with many 6′S sLe^x glycans attached. These "multivalent" glycan epitopes can bind to Siglec-8 lead to signaling in the cells and ultimate inhibition of cell function (including by death or apoptosis). The mucins in mucus (cross-linked mucins), which cover epithelial cells or airways, also act as a first line of defense as they can bind viruses through mulitple-contact (multivalent) binding sites, effectively trapping the viruses. The glycan structure recognized by Siglec-8 is sialic acid and sulfate (NeuAcα2-3[6S]Galβ1-4G[Fucα1-3]GlcNAc-). Given their role in inhibiting and inducing apoptosis in immune cell, the family of siglecs are likely involved as checkpoints, which are important in cancer and inflammatory conditions.

Figure 7.4.21shows the domain structure of Siglec-8

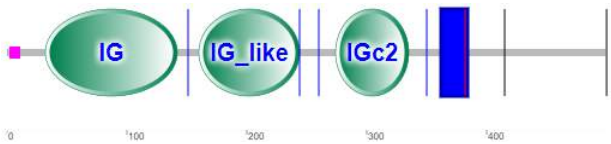


Figure 7.4.20: Domain structure of Siglec-8

Note that there is no CLECT domain, but rather immunoglobulin- (IG) or IG-like domains, which seems logical given their role in binding glycan "epitopes". The IG domain is also called the immunoglobulin V-set domain (V-Set). The blue rectangle represents the transmembrane domain (single helix). The cytoplasm contains a tyrosine-inhibitory motif (ITIM) involved in transducing the signal on binding 6′S sLe^x glycans to the IG domains.

As discussed above, humans lack a hydrolase gene necessary for the hydroxylation of Neu5Ac to Neu5Gc, which is found in chimps who possess the enzyme. Chimp's immune systems seem to confer protection from acquiring simian versiosn of AIDS, cirrhosis, and other diseases which humans acquire when they are infected with the human versions of the HIV virus, hepatitis B or C, or other viruses. These diseases and others associated with overactive T cells (rheumatoid arthritis, asthma, type-I diabetes) are not common in chimps. It turns out that there is a link between the type of sialic acid and the expression of siglecs that influences the difference for our disease propensity. Varki et al have shown that chimps and gorillas show much higher levels of expression of siglecs on T cells, which are critical regulatory and effector cells in the immune system. When siglecs on T cells are activated, T-cell responses are down-regulated. Although HIV virus ultimately kills T helper cells, the virus initially activates them on infection, leading to their proliferation and production of a larger number of cells for the virus to infect.

Figure 7.4.21 shows an interactive iCn3D model of human Siglec-8 lectin domain in complex with 6'sulfo sialyl Lewisx (2N7B)

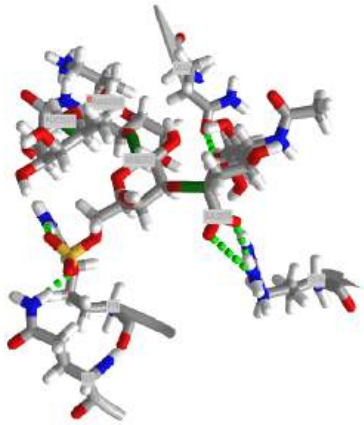


Figure 7.4.19: Human Siglec-8 lectin domain in complex with 6'sulfo sialyl Lewisx (2N7B). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i..Gaa93cu86jMPA9

7.5: Working with Carbohydrates

The material in this chapter is derived from the open access article referenced below and used under the following Creative Common's License.

Shirakawa, A.; Manabe, Y.; Fukase, K. Recent Advances in the Chemical Biology of *N*-Glycans. *Molecules* **2021**, *26*, 1040. https://doi.org/10.3390/molecules26041040. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (**http://creativecommons.org/licenses/by/4.0/**). It deals with the analysis and synthesize of N-glycans, but the same principles would be used for study of O-glycans, proteoglycans, etc. Text in boxes have been added to offer addition simplifying information when necessary for undergraduate students. References can be found in the original paper linked above.

7.5.1: Introduction

Glycosylation is the most common post-translational modification of proteins. Over 60% of proteins are linked to glycans. Asparagine-linked oligosaccharides (*N*-glycans) have a core pentasaccharide composed of mannose and glucosamine and are classified into three types: high-mannose, hybrid, and complex as shown in **Figure 7.5.1** below. In the biosynthesis of *N*-glycan-modified proteins, the high-mannose type *N*-glycan consisting of 14 residues ($Glc_3Man_9GlcNAc_2$) is first attached to proteins in the endoplasmic reticulum (ER). The initial high-mannose *N*-glycans play an important role in protein folding in the ER. Glycoproteins then migrate to the Golgi apparatus and are subsequently converted into complex-type *N*-glycans. Complex *N*-glycans have diverse structures due to differences in their associated synthesizing enzymes, resulting in different functions for each structure. For example, polylactosamine, consisting of a repeating structure of galactose and glucosamine, is involved in cancer metastasis and immune response. Sialic acids, in contrast, control immunity via recognition by Siglecs expressed in immune cells. Core fucose, which is a fucose linked to the glucosamine 6 position at the reducing end, and bisecting glucosamine, which is a glucosamine linked to the branched mannose 4 position, also play various roles and are closely related to many diseases. Hybrid *N*-glycans have both high-mannose and complex-type structures. Thus, *N*-glycans have diverse structures and are involved in a variety of biological phenomena. However, the molecular bases of their modes of action are yet to be fully elucidated.

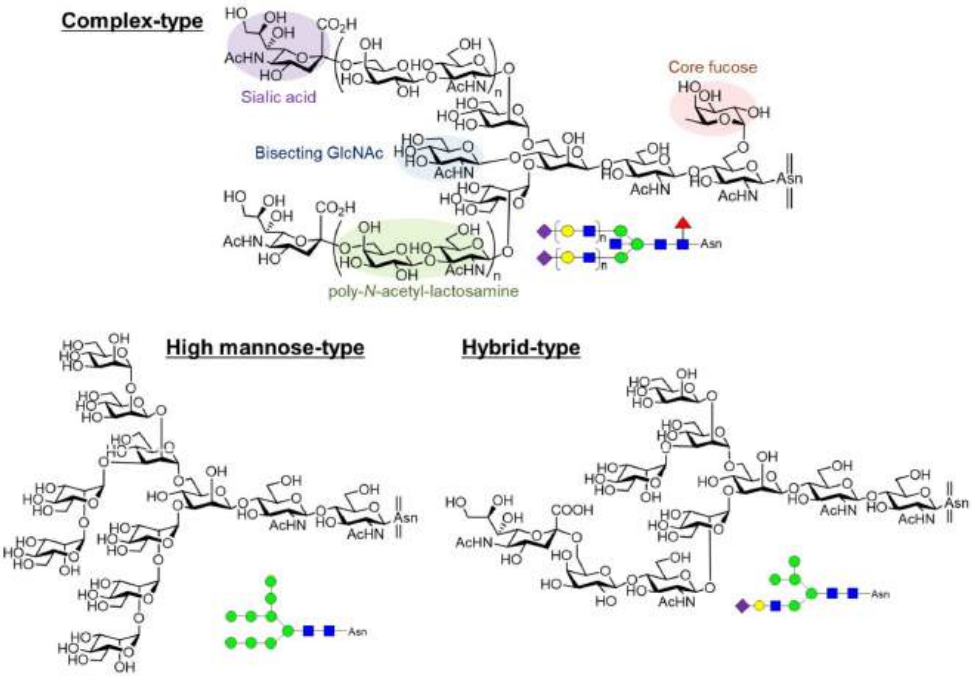


Figure 7.5.1: Structures of *N*-glycans. Complex-type *N*-glycans have diverse structures with/without sialic acid, poly-*N*-acetyl-lactosamine, bisecting GlcNAc, core fucose and so on. High-mannose-type *N*-glycan is composed of 14 residues ($Glc_3Man_9GlcNAc_2$) containing 3 glucoses, 9 mannoses, and 2 GlcNAc. Hybrid-type *N*-glycans have both high-mannose and complex-type structures.

Chemical synthesis, enzymatic synthesis, and isolation of diverse, pure *N*-glycans have been vigorously investigated for analyzing *N*-glycan functions at the molecular level. Chemical synthesis is an extremely potent approach that allows the de novo construction

of glycan structures. Any desired glycan structures can be constructed, including partial and artificial structures. Danishefsky et al. successfully synthesized various *N*-glycans with multiantennary structures. Unverzagt et al. achieved convergent synthesis of complex-type *N*-glycans with bisecting glucosamine and/or core fucose. We have also reported the synthesis of *N*-glycans. In addition, Ito et al, Boons et al., Wang et al., Wong et al., and Schmidt et al. have achieved *N*-glycan synthesis. Since the chemical synthesis of *N*-glycans with complex structures is a challenging process that requires multiple steps, the isolation of *N*-glycans from natural resources has also been explored. Kajihara et al. established an efficient method for the isolation of *N*-glycans from egg yolk, which has become a standard method for *N*-glycan preparation. In recent years, the preparation of *N*-glycans using enzymatic reactions has also been extensively investigated. Ito et al., Boons et al., Wang et al., and Wong et al. successfully constructed a wide range of *N*-glycan libraries via enzymatic synthesis using isolated or chemically synthesized glycans as substrates. Thus, over the years, the technical basis for a sufficient supply of various *N*-glycans has been established.

Owing to the increased availability of pure *N*-glycans, their functional elucidation has advanced considerably in recent years as shown on **Figure 7.5.2** below. *N*-Glycan functions have mainly been analyzed using molecular biological techniques, including knockout of biosynthetic enzymes. However, it is difficult to determine the precise structure–activity relationship using these methods. Although interaction analysis of lectins using relatively small glycan fragments, such as disaccharides and trisaccharides, has been used to study function, it is not possible to estimate the conformational effect or multivalent interactions of the complex structure of *N*-glycans. Recent interaction analysis of lectins using various *N*-glycans has elucidated the significance of such complex structures. The increased availability of *N*-glycans also allows one to prepare glycoproteins with homogeneous glycoforms, enabling the elucidation of *N*-glycan function on the distinct protein. Furthermore, *N*-glycans have the potential to be used in the development of novel drugs. This review provides an overview of the recent chemical biology study of *N*-glycans.

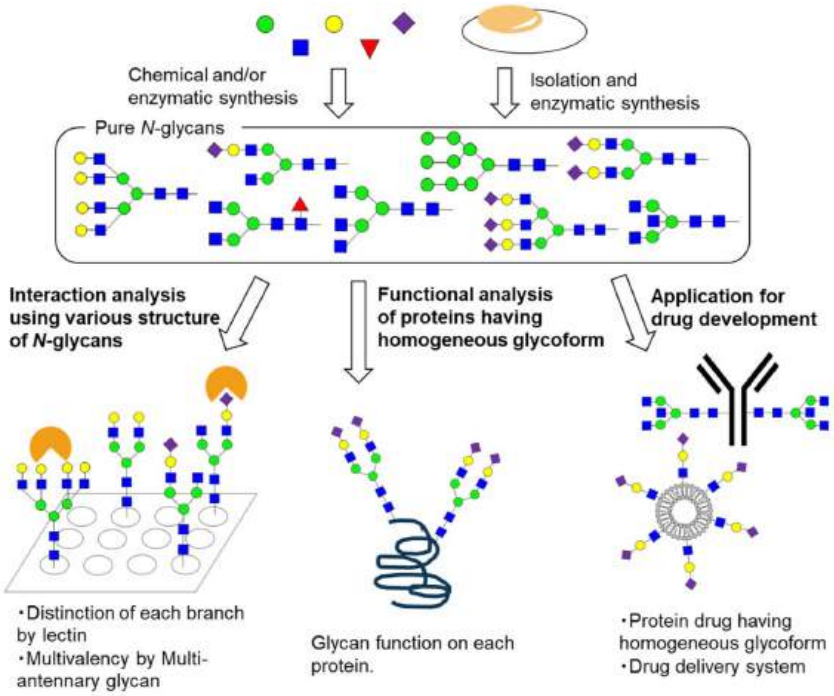


Figure 7.5.2: The chemical biology study of homogeneous *N*-glycans. Chemical synthesis, enzymatic synthesis, and isolation of diverse, pure *N*-glycans enable their functional analysis at the molecular level. Interaction analysis using various *N*-glycans revealed the significance of complex *N*-glycan structures—for example, distinction of each branch and multivalent interaction in lectin recognition. Functional analysis using proteins with homogeneous glycoforms clarifies glycan function on each protein. Application of *N*-glycans for drug development is also investigated

7.5.2: Elucidation of the Molecular Basis of *N*-Glycan Recognition by Lectins

Glycans control various biological phenomena through their recognition by lectins. Thus, interaction analysis between *N*-glycans and lectins is essential to elucidate *N*-glycan functions .

7.5.2.1: Methods for the Glycan–Lectin Interaction Analysis

The glycan–lectin interaction analysis methods include analyses using glycan arrays, nuclear magnetic resonance (NMR), isothermal titration calorimetry (ITC), surface plasmon resonance (SPR), fluorescent polarization (FP), and X-ray crystallography.

ITC gives thermodynamic parameters, whereas SPR provides kinetic parameters. FP realizes a simple and easy assay system. X-ray crystallography provides precise structural information. In this review, we focus on studies using glycan arrays and NMR, which are effective methods for elucidating the interaction between glycans and lectins.

Glycan arrays are used to detect the binding of lectins to immobilized glycans. An advantage of a glycan array is that a large number (dozens and hundreds) of samples can be examined in a high-throughput manner using a small amount of glycans. Interaction analysis using various structures of glycans provides insights into precise structure–activity relationships. Glycan immobilization methods are divided into two categories: noncovalent and covalent. Noncovalent immobilization utilizes hydrophobic interactions, charge interactions, and biotin–streptavidin interactions among others. Covalent immobilization methods typically use coupling of an amino group introduced at the reducing end of a glycan to the plate surface activated with *N*-hydroxysuccinimide. Many other methods, including thiol-maleimide coupling and alkyne-azide click reactions, have been reported. As for the detection, fluorescence is usually used to realize high-throughput analysis.

NMR can be used to analyze interactions at the atomic level. Saturation transfer difference (STD) NMR is a particularly powerful method for the analysis of glycan–lectin interactions. In this method, saturation transfer from the protein to the ligand is observed as STD signals after the saturation of protein by radio frequency as shown in **Figure 7.5.3** below. The closer the protons are to the protein, the stronger the STD signals observed. STD-NMR was originally developed as a method for screening ligands from mixture systems but is now widely used for the analysis of protein–ligand binding modes. This method works when the affinity is not high (K_D is 10^{-3} to 10^{-8} M), because STD signals are measured when the protein and ligand are in an equilibrium state of binding and dissociation. Since glycan–lectin interactions are usually weak (K_D in mM-μM), STD-NMR is highly effective. NMR is also a powerful tool for the conformational analysis of glycans. Importantly, this method does not require labeled proteins. In addition, a small amount of receptor is necessary (typically micromolar range). However, an excess of the ligands is used (typically the millimolar range), thus, low solubility of the ligand causes a problem. While conformation is an important factor for glycan recognition, the flexibility of glycans makes conformational analysis difficult. In addition to analysis based on coupling constants and the nuclear Overhauser effect (NOE), an analysis using pseudocontact shift (PCS) by paramagnetic metals has recently been developed, and its efficiency has been demonstrated.

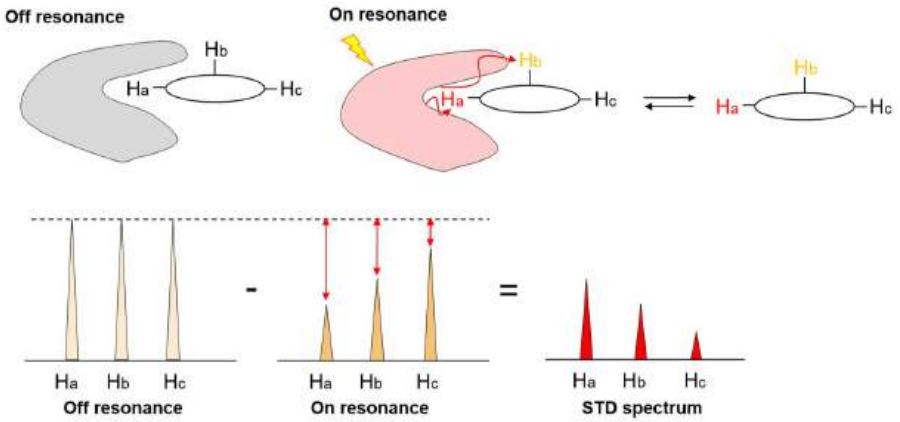


Figure 7.5.3: Mechanism of Saturation transfer difference nuclear magnetic resonance (STD-NMR). “Off resonance” experiment gives a reference spectrum. Under “on resonance” conditions, the saturation is transferred from protein to ligand by spin diffusion through intermolecular nuclear Overhauser effects (NOEs). The closer the protons are to the protein, the stronger the STD signals that are observed.

Examples of the analysis of glycan–lectin interactions using glycan arrays and NMR are introduced below.

7.5.2.2: Analysis of Sugar–Lectin Interactions Using Glycan Arrays

Glycan arrays are excellent tools for the comprehensive analysis of glycan–lectin interactions. Previous interaction analysis using small fragments, such as disaccharides and trisaccharides, revealed the minimum structure (epitope) required for recognition of individual lectins. Meanwhile, recent advances in the preparation of the whole structure of various *N*-glycans have allowed the full realization of structure–activity relationships, elucidating the significance of complexity of *N*-glycan structures as shown in **Figure 7.5.4** below. For example, these advances have provided insights into the differences in the lectin recognition of each branch, the improvement of affinity due to the inclusion of multiple recognition units (multivalent effect), the influence of chain length on affinity, and remote (heterovalent) recognition.

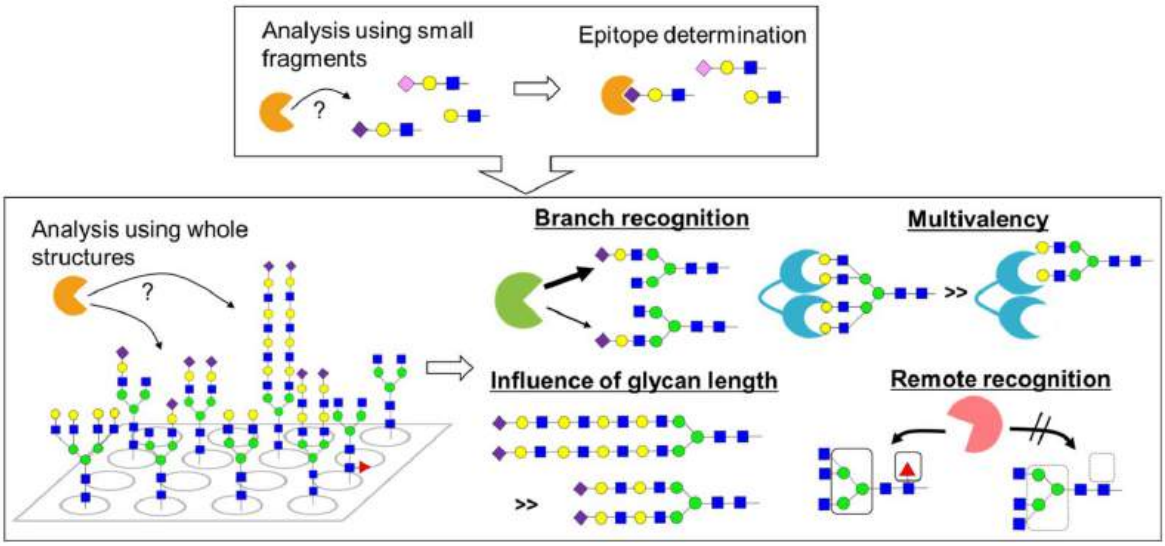


Figure 7.5.4: Analysis of glycan–lectin interaction using glycan arrays. Interaction analysis using small fragments, such as disaccharides and trisaccharides, revealed the epitope required for lectin recognition, whereas interaction analysis using whole structures of *N*-glycans revealed the significance of complexity of *N*-glycan structures; these analyses provided the insights into the differences in the lectin recognition of each branch, the improvement of affinity due to the inclusion of multiple recognition units (multivalent effect), the influence of chain length on affinity, and remote (heterovalent) recognition

Wang et al. demonstrated the differences of each *N*-glycan branch in lectin recognition by comprehensive interaction analysis of various *N*-glycans with several lectins using a glycan array. Plant-derived *Sambucus nigra* lectin (SNA), which recognizes sialic acid, recognized the sialic acid on the α1,3-branched chain more strongly than the sialic acid on the α1,6-branched chain. Meanwhile, plant-derived *Maackia amurensis* lectin (MAL-I) and virus-derived lectin hemagglutinin (HA) strongly bind to sialic acid on the α1,6-branched chain. MAL-I also interacts with terminal galactose; in this case, MAL-I strongly recognizes galactose on the α1,3-branched chain, suggesting that MAL-I has two distinct glycan recognition domains. In addition, *Erythrina cristagalli* lectin (ECL), which recognizes the lactosamine structure, has a higher affinity to lactosamine on the α1,3-branched chain than on the α1,6-branched chain. *Phaseolus vulgaris* erythroagglutinin (PHA-E) prefers terminal galactose on the α1,6-branched chain and terminal glucosamine on the α1,3-branched chain, whereas wheat germ agglutinin (WGA) strongly interacts with glucosamine on the α1,3-branched chain.

Branch selective binding of C-type lectins and monoclonal antibodies was also revealed by using glycan array including *N*-glycan positional isomers prepared by chemo-enzymatic method. DC-SIGN, C-type lectin recognizing glycan on bacteria and viruses, showed strong binding to hybrid- and complex-type glycans and *N*-glycans presenting Lex epitopes. DC-SIGN showed preferential binding to the biantennary glycans with terminal galactose or *N*-acetylgalactosamine on the α1,6-branched chain, whereas DC-SIGNR showed the opposite binding behavior. L-SECtin showed the preference to GlcNAc1,2-Man residues on the 3-arm of the complex and hybrid *N*-glycans.

N-glycans have symmetric structures on the nonreducing end side, and these studies indicate that glycan structures on each branched chain have distinct functions.

The structural redundancy of *N*-glycans plays an important role in enhancing their affinity to lectins because of their multivalency. Interaction analysis of Siglec-1, -2, -9, and -10 with sialic acid-containing *N*-glycans using a glycan array showed a higher affinity for four-branching *N*-glycans than for two-branching *N*-glycans. Multivalent effects were also confirmed in ECL, which recognizes the lactosamine structure, and *Ricinus communis* agglutinin (RCA120), which recognizes terminal galactose. Similarly, interactions with galectin were also enhanced as the number of recognition units increased.

The structure at the remote positions of the lectin recognition unit can affect its interaction. *Lens culinaris* agglutinin (LCA), which binds to core fucose, only recognizes core fucosylated biantennary and triantennary *N*-glycans with particular branching patterns, but did not recognize triantennary *N*-glycans with other branching patterns or tetraantennary *N*-glycans. These results indicate that the branching structure away from the core fucose affected recognition by LCA, although its recognition site is core fucose. On the other hand, the affinity between HA from H3N3 and sialic acids at the nonreducing end was increased by chain elongation; the insertion of a polylactosamine repeating structure enhanced the affinity. These studies revealed that both the epitope and the whole glycan structure are important for the recognition of *N*-glycans.

Glycan arrays, comprising *N*-glycans along with glycolipids and *O*-glycans, have been used to investigate the host–pathogen interactions in diagnostic and therapeutic applications. For example, the inhibition of human anti-N9 antibodies to influenza neuraminidases was analyzed by glycan array. The binding study of the H3N2 influenza viruses using glycan microarrays demonstrated the changes in virus hemagglutinin that affect the receptor binding properties of the viruses.

Glycan arrays can also be used to explore artificial glycoligands as new drug candidates that target lectins. High-affinity ligands for Siglecs or several C-type lectins, which are involved in immune regulation, are expected to be lead compounds for drug development. However, glycan–lectin interactions are usually weak, which is a major issue in the utilization of glycans as bioactive molecules. Thus, the synthesis of glycans and derivatization of artificial molecules, followed by high-throughput screening using glycan arrays, is expected to be a powerful approach to address this limitation.

7.5.2.3: Analysis Using NMR

NMR analysis can provide insights into glycan–lectin interactions at the atomic level. STD-NMR can be used for high-resolution epitope mapping. Similar to the results obtained from glycan arrays, NMR analysis also reveals that not only epitopes but the whole structure of *N*-glycans plays an important role in glycan–lectin interactions. The conformational analysis of *N*-glycans using NMR is a powerful approach that provides a rational explanation for the molecular basis of the recognition of complexity of *N*-glycan structures.

STD-NMR allows for a detailed analysis of glycan–lectin interactions. Many researchers have analyzed the interactions between sialic acid containing *N*-glycans and Siglecs . Silipo et al. analyzed the interaction between Siglec-2 and sialyl *N*-glycans using STD-NMR and molecular dynamics (MD) simulations. The Siglec-2 epitope was clearly shown by STD-NMR, and the conformation of sialyl *N*-glycans was predicted by NMR analysis and MD simulations. When biantennary sialyl *N*-glycans were recognized, Siglec-2 only interacted with the sialyl disaccharide at the nonreducing end, and the other part was expected to protrude from the protein surface. These results suggest that multiantennary *N*-glycans with multiple sialic acids can interact with several Siglec-2 and induce the formation of Siglec-2 oligomers on B cells.

STD-NMR analysis using the whole structure of *N*-glycans has demonstrated that lectins not only recognize small units, such as disaccharides and trisaccharides, but also interact with *N*-glycans in a more complex manner. In *N*-glycan recognition by *Pisum sativum* agglutinin (PSA), a mannose-recognition lectin, core fucose was shown to alter its binding mode. When biantennary *N*-glycans without core fucose were used for the interaction analysis with PSA, the mannose on each branch gave comparable STD signals. While for the core fucose containing *N*-glycans, the STD signals of mannose on the α1,6-branched chain were weakened, and instead, an interaction with the methyl group of the core fucose was observed. STD-NMR using a fluorine derivative (2D STD-TOCSYreF) indicated that the mannose on the α1,3-branched chain was more strongly recognized by PSA than the mannose on the α1,6-branched chain. In addition, dectin-1, which recognizes fungal β-glucan, was found to recognize core fucose on immunoglobulin (IgG) . STD-NMR analysis indicated that dectin-1 interacted not only with core fucose but also with an Fmoc group attached to the amino group of asparagine introduced at the reducing end. These results suggest that dectin-1 recognizes amino acids with aromatic side chains, such as phenylalanine and tyrosine, together with core fucose. On the other hand, STD-NMR is also effective for the analysis of substrate recognition by glycosyltransferases. The STD-NMR analysis of FUT8, a fucosyltransferase that builds core fucose structure, revealed the precise interaction between FUT8 and *N*-glycan . FUT8 recognizes not only glucosamine at the reducing end (reaction point) but also the whole glycan structure. In particular, FUT8 strongly interacted with the α1,3-branched chain at the nonreducing end.

Advanced STD-NMR methods have been developed. Saturation transfer double difference (STDD)-NMR is useful for the direct observation of ligands binding on the surfaces of living cells. Clean-STD can avoid accidental saturation to give improved detection of ligand–protein interactions at low concentration of protein. Second dimension STD-NMR, i.e., STD-TOCSY, STD-HSQC, STD-NOESY, can overcome the problems of proton overlapping typical of glycan NMR analysis.

Conformation analysis of glycans using NMR provides important insights into complex glycan–lectin interactions. The *N*-glycan conformation can be predicted by combining PCS-based NMR analysis and MD simulations as shown in **Figure 7.5.5** below. Kato et al. analyzed the conformation of high-mannose glycans by PCS-based NMR analysis using ^{13}C-labeled compounds and Tm^{3+} as a paramagnetic metal ion tag. They elucidated the conformational change caused by mannose trimming during the *N*-glycan biosynthetic process. Unverzagt and Barbero et al. distinguished each branch of tetraantennary *N*-glycan based on the PCS method and analyzed the differences in the recognition of each branch by lectins . *Datura stramonium* seed lectin (DSL), which recognizes the lactosamine structure, interacts more strongly with the lactosamine on the α1,6-branched chain than with that on the α1,3-branched chain. On the other hand, no differences in the strengths of STD signals of each branch were observed with *Ricinus communis* agglutinin (RCA120), which recognizes terminal galactose, indicating that RCA120 recognizes all branches without

distinction. In a similar analysis between sialic acid containing biantennary *N*-glycans and HA, STD signals from both sialic acids were observed, suggesting the contribution of two sialic acids in a multivalent effect. Furthermore, interesting results have been reported showing that the *N*-glycan conformation directly affects lectin recognition. *N*-Glycans have three back-fold conformations and two extended conformations, in which the α1,6-branched chain is folded toward the reducing end or extended, respectively. The addition of core fucose or bisecting glucosamine significantly changes their conformational equilibria and reduces the number of major conformations from five to four and five to two, respectively. Crystal structure analysis and transferred NOE (TrNOE) analysis revealed that *Calystegia sepium*-derived calsepa and *Phaseolus vulgaris*-derived phytohemagglutinin (PHA-E), which recognize bisecting glucosamine containing *N*-glycans, recognize *N*-glycans in the back-fold conformation induced by bisecting glucosamine addition

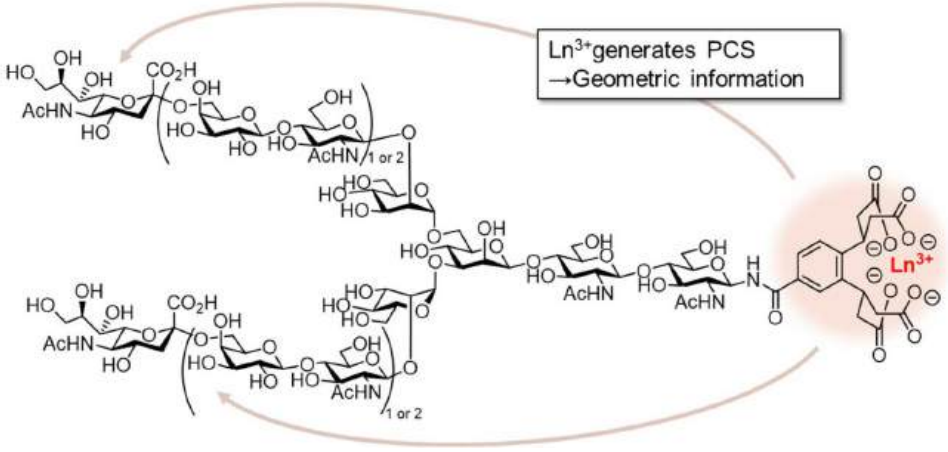


Figure 7.5.6*a*: Conformation analysis of *N*-glycan using pseudocontact shift (PCS). Chelation with paramagnetic metals can induce PCS to give geometric information of *N*-glycan.

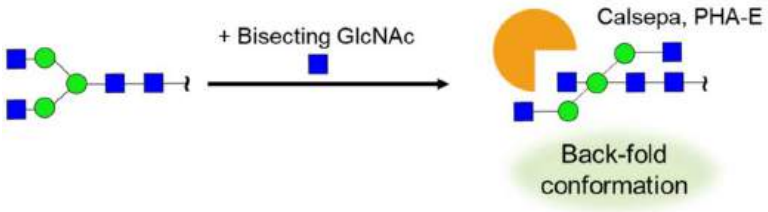


Figure 7.5.6*b*: Recognition of bisecting GlcNAc containing *N*-glycan by Calsepa and *Phaseolus vulgaris* erythroagglutinin (PHA-E). Attachment of bisecting GlcNAc enhances back-fold conformation, which is recognized by Calsepa and PHA-E.

7.5.3: Functional Analysis of *N*-glycans on Glycoproteins

Analysis of *N*-glycan functions on glycoproteins needs to be considered with proteins. In recent years, improvements in the techniques for the synthesis of peptides and proteins, as well as glycans, have enabled the preparation of glycoproteins with homogeneous glycans. *N*-Glycans on glycoproteins can be modified by Endo-β-*N*-acetylglucosaminidases (ENGases). Synthesized glycoproteins with homogeneous glycan structures have helped elucidate precise glycan functions.

A series of synthetic studies of glycoproteins and glycoprotein mimics by Ito and Kajihara et al. revealed the precise function of *N*-glycans in a quality-control mechanism for glycoproteins in the endoplasmic reticulum (ER). ER has a quality control system that promotes the correct folding of ribosome-produced proteins. In the case of *N*-glycosylated proteins, high-mannose *N*-glycans work as tags for protein folding. A common dolichol-linked oligosaccharide precursor containing terminal glucose trisaccharide is first synthesized in the ER and is transferred to proteins by the oligosaccharyltransferase (OST). The folding process then starts. The first glycosidase (GCSI) cleaves the terminal glucose and the second glycosidase (CGSII) further cleaves glucose residues to afford monoglucosylated or nonglucosylated glycoproteins. The folded nonlucosylated glycoproteins are then transferred to the glycan modification process. The UDP-glucose:glycoprotein glucosyltransferase (UGGT) complex distinguishes misfolded glycoproteins and transfers glucose to the nonreducing end of the high-mannose glycan. This monoglucosylation serves as a marker for misfolded glycoproteins and the chaperone proteins calnexin/calreticulin (CNT/CRT) promotes folding. CGSII then cleaves glucose residue to transfer the glycoproteins for the glycan modification process. This cyle is illustrated in **Figure 7.5.7** below.

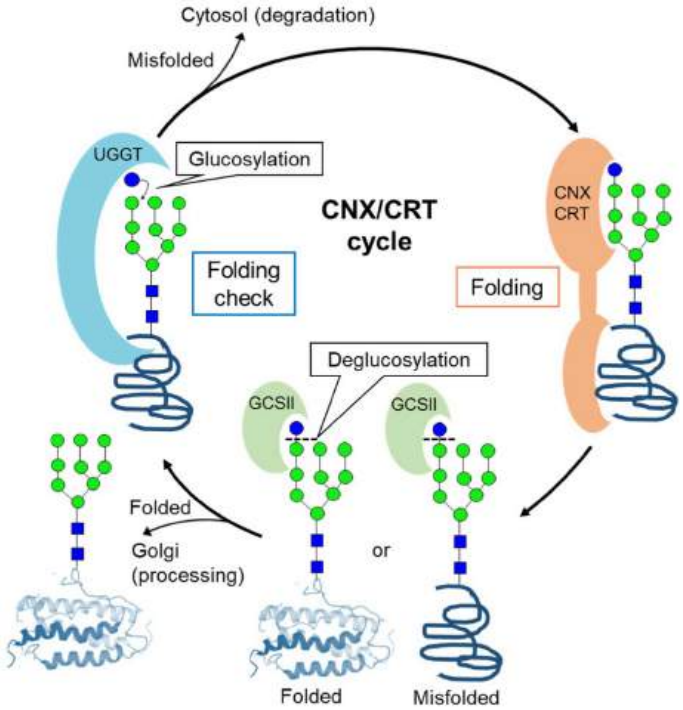


Figure 7.5.7: Protein quality control utilizing high-mannose-type *N*-glycan as a tag. UDP-glucose:glycoprotein glucosyltransferase (UGGT) complex distinguishes misfolded glycoproteins to transfer glucose to the nonreducing end of the high-mannose glycan. This monoglycosylation serves as a marker for misfolded glycoproteins and the chaperone proteins calnexin/calreticulin (CNT/CRT) promotes folding.

The defects in this process cause congenital disorders of glycosylation (CDGs), which are severe genetic diseases. CDG is classified into Type I and Type II. In Type I, the enzymes are mutated in synthesis and transfer a common dolichol-linked oligosaccharide precursor and enzyme substrates. Type II defects the modification process of *N*-glycans in the ER and Golgi. Lack of GCS1 causes CDG-IIb. Unfolded proteins lead to ER stress and cause CDGs .

Ito et al. introduced methotrexate (MTX) at the reducing end of high-mannose *N*-glycans and prepared a complex with dihydrofolate reductase (DHFR), which recognizes MTX. Such glycoprotein mimics were used to analyze the interaction with UGGT. They also investigated various aglycone structures as substrates of UGGT. In addition, chemically synthesized glycoproteins were used for the analysis of substrate recognition by UGGT. UGGT showed higher enzymatic activity against high-mannose *N*-glycans on misfolded interleukin-8 (IL-8) than against those on the folded one. Furthermore, they synthesized several glycoproteins and isotope-labeled glycopeptides and revealed that UGGT recognizes hydrophobic patches on misfolded proteins. As shown above, they elucidated the molecular basis of the quality-control mechanism based on high-mannose *N*-glycans using glycoprotein mimics and chemically synthesized glycoproteins.

Maintaining the appropriate folding is also critical for in the degradation process. Mutations in human *N*-glycanase 1 (NGLY1) cause the congenital disorder of deglycosylation (CDDG). Suzuki revealed that *N*-GlcNAc proteins are accumulated by the action of Endo-β-*N*-acetylglucosaminidase (ENGase) in Ngly1-defective cells . During ER-associated degradation, *N*-GlcNAc proteins form aggregates that seem to be toxic. Suzuki also revealed that lethality of Ngly1-KO mice is partially rescued by the additional deletion of the Engase gene, suggesting that ENGase inhibitors are targets for CDDG.

In recent years, the influence of *N*-glycan modifications on the bioactivity of proteins has been gradually elucidated using synthetic glycoproteins. Hematopoietic hormone erythropoietin (EPO), which is used to treat renal anemia, has three *N*-glycan-modification sites. EPO with various glycoforms is used as a drug. Several groups have reported the synthesis of EPO with homogeneous glycoforms, and the effect of *N*-glycans on their biological activities has been investigated. In addition, various neoglycoprotein analogues of EPO have been reported. Kajihara et al. synthesized five types of EPO, which is introduced sialic acid containing *N*-glycans into three *N*-glycosylation sites with different patterns, and showed the relationship between glycosylation sites and hematopoietic activities. Increasing the number of sialic acids containing *N*-glycans on EPO improved the stability in blood, leading to an improvement in hematopoietic activity. Moreover, the metabolic stability of EPO was highly correlated with hydrophobicity, suggesting that glycan modifications enhance the in vivo stability by covering hydrophobic sites on the protein

surface. Kajihara et al. also synthesized two types of interferon-β (IFN-β) with sialic acid-containing and noncontaining (asialo) *N*-glycans, and their activities were evaluated. IFN-β modified with sialic acid-containing *N*-glycans exhibited higher activity than that modified with asialo *N*-glycans, suggesting that sialic acid extended the in vivo half-life of IFN-β. Thus, *N*-glycans are closely related to the stability of glycoproteins in vivo. Indeed, Tanaka et al. demonstrated the effect of *N*-glycans on protein metabolic stability by positron emission tomography (PET) imaging using glycodendrimers as pseudoglycoproteins. On the other hand, *N*-glycosylation can also affect binding affinity to a receptor. Okamoto et al. synthesized two types of chemokine CCL1 with and without *N*-glycan, in which *N*-glycosylation reduced the activity of CCL1, suggesting that CCL1 biological activity can be regulated by *N*-glycan modification. Thus, it should be noted that the role of *N*-glycan modifications can be different between proteins. We reported that dectin-1 specifically recognized core fucosylated IgG and did not interact with other core fucosylated proteins, suggesting that core fucose on IgG has specific physiological functions. The role of *N*-glycans on distinct proteins is an important topic for future work.

7.5.4: Use of *N*-glycans for Drug Development

The increased supply of *N*-glycans has led to an increase in the use of *N*-glycans for drug development. Because *N*-glycans are endogenous molecules, they are unlikely to be toxic or immunogenic and, thus, are expected to have high safety profiles.

7.5.4.1: Next-Generation Protein/Peptide Drugs Modified with Homogeneous N-Glycans

Controlling the glycan structure is an important issue in the preparation of glycoprotein and glycopeptide drugs. Biopharmaceuticals, including antibodies, are common pharmaceuticals. Although many proteins utilized in biopharmaceuticals are glycoproteins, their actual glycan structures are often neglected or ignored. However, the significance of the role of glycans on the function of glycoproteins has recently been illuminated, and the importance of the glycan structure has been highlighted. The preparation of glycoproteins with homogeneous glycans is also important from the viewpoint of quality control.

IgG antibodies have *N*-glycans at Asn297 in the Fc region of the heavy chain, and their structures affect activity, dynamics, and safety (**Figure 7.5.8**). The importance of core fucose on these *N*-glycans is well known. The removal of the core fucose from IgG antibodies dramatically enhances antibody-dependent cellular cytotoxicity (ADCC) activity . Mogamulizumab, the antibody without core fucose, is actually in current use. Bisecting glucosamine and terminal galactose have been reported to affect ADCC and complement-dependent cytotoxicity (CDC) activities. Therefore, modifications of *N*-glycans on IgG antibodies have been extensively investigated. ENGase provides a powerful tool. *N*-Glycans on antibodies can be trimmed, and other *N*-glycans can be introduced by ENGase. Antibody–drug conjugates (ADCs) have also been prepared using this method in which *N*-glycans were changed into a structure with a tag for subsequent reactions, and small molecular drugs were introduced via bio-orthogonal reactions. This approach allows for the introduction of drugs into the Fc region without affecting antigen recognition. Furthermore, the *N*-glycan structure can be made homogeneous. Wong et al. introduced *N*-glycans with 3-position fluorinated sialic acids into antibodies. Because this fluorinated *N*-glycan was not degraded by sialidase and modification with sialic acid containing *N*-glycan can enhance the metabolic stability of proteins, this antibody is expected to show a significant improvement in pharmacokinetics.

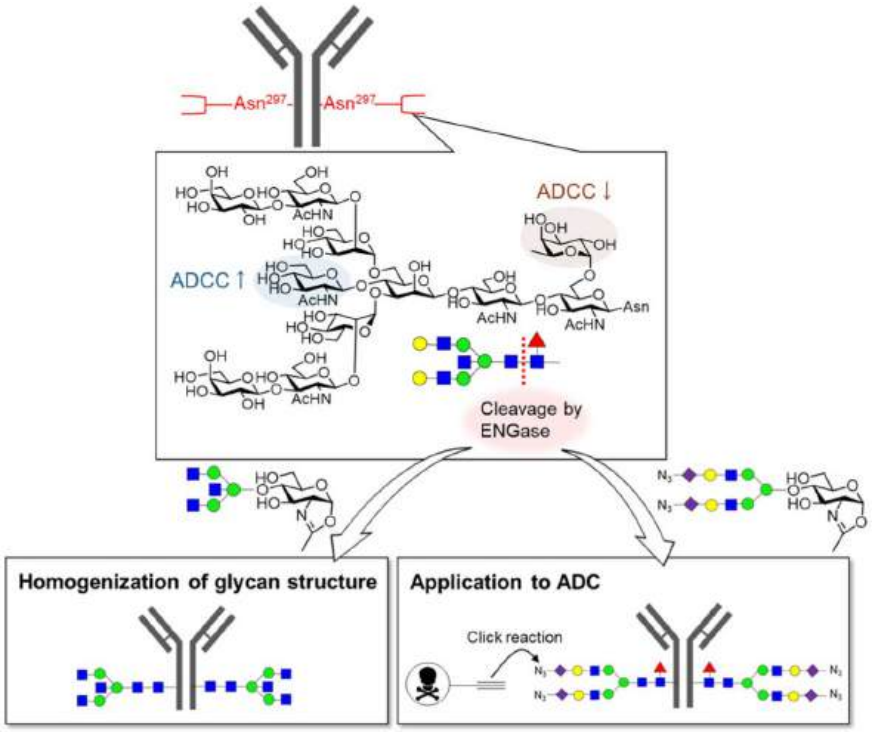


Figure 7.5.8: Glycan editing of immunoglobulin (IgG) antibodies. *N*-Glycans at Asn297 of IgG affect their activity. Core fucose reduces the antibody-dependent cellular cytotoxicity (ADCC) activity, whereas bisecting GlcNAc enhances the ADCC activity. *N*-Glycan editing using Endo-β-*N*-acetylglucosaminidase (ENGase) can give IgG as a homogeneous glycoform or can be applied for the preparation of Antibody–drug conjugates (ADCs).

N-Glycans play important roles not only in antibodies, but also in many other glycoproteins. As described above, the structure–activity relationship study of *N*-glycans on EPO demonstrates the importance of the *N*-glycan structure on the bioavailability and bioactivity of proteins. Hossain and Wade et al. reported that the physical properties of insulin can be improved by adding *N*-glycan to insulin, which originally has no glycans. Introduction of sialic acid containing *N*-glycans to insulin successfully inhibited problematic fibril formation. In addition, *N*-glycan-modified insulin bound to its receptor with almost the same affinity as the natural form, and further improvements in its metabolic stability were observed. Currently, PEGylation has been generally used to enhance the bioavailability of proteins; however, PEG is not without adverse effects. Considering that *N*-glycans are endogenous glycans and are expected to be extremely safe, "*N*-glycan modification" has the potential to become a common strategy for improving the protein/peptide bioactivity.

The structure of *N*-glycans is also important for vaccine development. Viruses use host biosynthetic systems to synthesize proteins. Consequently, viral proteins are subjected to glycan modification. Therefore, glycoproteins and glycopeptides are candidate antigens for vaccine development, and their glycan structures influence their functions. HIV vaccine candidates containing *N*-glycans have been designed and synthesized . Wang et al. reported that the glycan structure on the antigen was critical for the neutralization activity of antibodies, clearly demonstrating the importance of the glycan structure in vaccine design. In addition, Wang showed the importance of glycan structures in the development of influenza HA-based vaccines. For the development of vaccines against COVID-19, the spike protein is a promising antigen candidate. This protein is heavily glycosylated, but *N*-glycan modifications of spike proteins have been reported to reduce their antigenicity. However, *N*-glycan-modified antigens may induce antibodies against endogenous *N*-glycans, which should be carefully examined. Overall, glycans are likely to be important for developing highly efficient and safe vaccines.

7.5.4.2: Drug Delivery Systems (DDSs) Using N-Glycans

N-Glycans interact with various biomolecules, including many lectins, and thus show distinct dynamics in vivo. Therefore, DDSs using *N*-glycans have been investigated. Because glycan–lectin interactions are weak, multivalent materials, including polymers, dendrimers, and liposomes, are usually utilized to enhance their interactions.

We synthesized dendrimers of sialic acid containing *N*-glycans and evaluated their dynamics in vivo using PET imaging. We revealed that the structure of *N*-glycans affected the uptake of dendrimers into specific organs. In addition, Tanaka et al. developed

an *N*-glycan-based DDS using albumin as a multivalent scaffold (**Figure 7.5.9**). The albumins modified with *N*-glycans were used as carriers of metal catalysts to realize chemical reactions at the desired organ in vivo. It should be noted that they achieved metal-catalyzed reactions in vivo by utilizing the hydrophobic pocket of albumin.

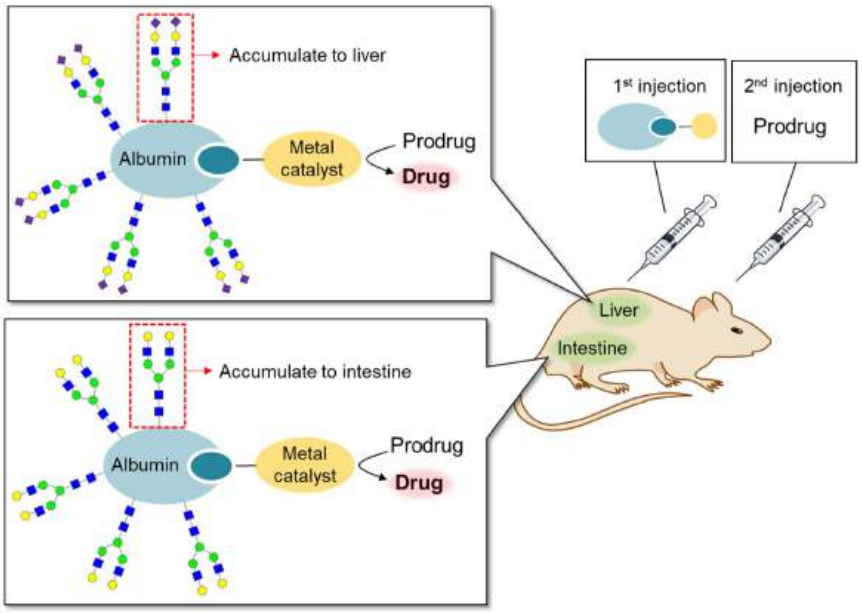


Figure 7.5.9: In vivo reaction using artificial glycosylated albumin metalloenzymes. Specific *N*-glycan conjugated albumin is specifically uptaken into the specific organs. Thus, the albumins conjugated with *N*-glycans were used as carriers of metal catalysts to realize chemical reactions for the activation of prodrug at the desired organ.

Siglecs, which recognize sialic acid, are expressed on immune cells and are involved in immune regulation. Immune cells can be targeted by utilizing sialyl glycan–Siglec interactions. Paulson et al. developed high-affinity Siglec ligands by the derivatization of sialic acid. They synthesized *N*-glycans containing these artificial structures, which exhibited a high affinity for Siglec-2. They achieved B cell targeting using liposomes displaying this *N*-glycan. Utilization of different sialyl glycans enables the targeting of various immune cells. In addition to Siglecs, DDSs targeting galectins, which recognize galactose, have also been investigated.

7.5.5: Future Perspectives

Glycans exist as polysaccharides in nature and are involved in multivalent interactions for pattern recognition. Conformational control via the formation of polysaccharides also plays an important role in glycan functions. In addition, many glycans function only when they are linked to proteins or lipids. Such emergent glycan functions can only be revealed by analysis using the whole glycan structure or glycoconjugates. As described herein, the increased availability of various *N*-glycans has led to the elucidation of the significance of complex *N*-glycan structures. The influence of *N*-glycan modification on some protein functions was also discussed. On the other hand, the molecular basis of glycan functions on membrane proteins remains to be elucidated, although glycans are attached to almost all membrane proteins and have diverse functions.

Recent advances in the engineering of cell-surface glycans are expected to provide a powerful approach to tackle this challenging issue. Bertozzi et al. developed metabolic labeling of cell surface glycan by incorporating unnatural sugar analogs using "**click chemistry**" having the reaction tag followed by the **bioorthogonal reactions**. We'll discuss these techniques further below.

In addition to glycan function analysis, the therapeutic application of metabolic glycan labeling is being vigorously investigated . Glycan engineering by chemical and chemoenzymatic methods has also been investigated. In addition, de novo glycans on cell surfaces have also been reported, such as the direct introduction of defined glycan structures into plasma membranes by lipid insertion, liposomal fusion, and tag technology. Such glycan editing technique enables glycan functions to be explored on membrane proteins on living cell surfaces.

A major feature of glycans is their heterogeneity. Glycans attached to the same site on the same protein can have diverse structures. In addition, many proteins have multiple glycosylation sites to which various glycans can be added. Although studies using pure *N*-glycans have revealed the functions of individual *N*-glycans, little is known about their function in combination with each other. Kurbangalieva and Tanaka et al. prepared albumins labeled with several *N*-glycans and observed their dynamics in vivo. Interestingly, their dynamics were altered depending on the *N*-glycosylation pattern. These results suggest that the simultaneous

interaction of multiple *N*-glycans may result in the expression of functions different from those of individual *N*-glycans. Little is known about whether the interactions of glycans with multiple lectins work collaboratively or competitively. A bottom-up approach to the construction of controlled glycoforms is expected to be a powerful strategy to address this difficult issue.

Glycans are considered to be the third most important life chain and have attracted increasing attention in recent years. However, unlike nucleic acids and proteins, their functional analysis and regulation have been delayed due to the lack of simple preparation methods. Recent advances in the preparation of *N*-glycans are expected to accelerate functional studies.

7.5.6: Click Chemistry and Bioorthogonal Reactions

Click chemistry is a powerful way to covalently connect two molecules - hence the name click chemistry. Sharpless, Meldal and Bertozzi were awarded the Nobel Prize in Chemistry in 2022 for its development and application. It has been use to synthesize active site inhibitors for enzymes as well as to label glycan and other biomolecules in vitro and in vivo.

In click chemistry, two molecules are "stitched" together to form a new molecule, much like two activated amino acids condense to form a dipeptide. Click chemistry was developed to emulate the simple solutions found in nature to produce polymers. Click chemistry reactions should not be sensitive to water and oxygen, easy to purify products, and proceed with a favorable ΔG (< - 20 kcal/mol, -84 kJ/mol). Azides and acetylenes were the first used but the addition of Cu^{1+} as a catalyst made the reaction very fast. The reaction can take place easily in blood and even urine.

Drugs that inhibit enzymes typically bind to the active site of the enzyme where catalysis occurs. Binding of an inhibitor precludes binding of the normal reactants (substrates) for the enzyme, inhibiting its activity. Using click chemistry, two small reactive molecules selected to bind independently in the active site can covalently react with each other to form a new drug with very high specificity and very high binding affinity (low K_D). This has been used to synthesize noncovalent inhibitors of the enzyme acetylcholinesterase (Barry Sharpless Lab, Scripps Lab). The reactive groups chosen in the example below are azide and acetylene derivatives, which when held in close approximation in the binding site of the enzyme undergo a cycloaddition reaction to form a triazole.

Figure 7.5.10 shows the reaction of an azide and acetylene in solution (without a "binding template"), which leads to equal amounts of the syn and anti products.

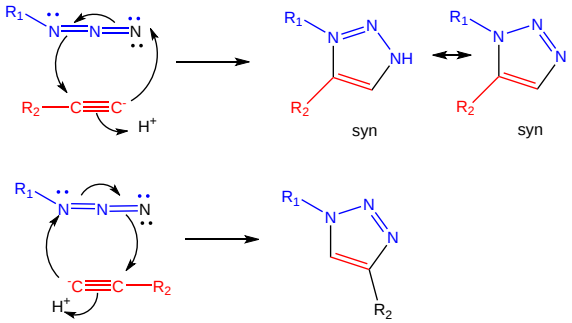


Figure 7.5.10: Click chemistry reaction between azide and acetylide without a directing template

The actual mechanism (not the simplified version shown) requires catalysis by copper ions (Cu^{1+}), which forms a complex with the acetylide (deprotonated acetylene). This decreases the pKa of the acetylene functional group, making it a better nucleophile. A dicopper intermediate is suggested in which the azide interacts with the second copper. Subsequent rearrangement lead to the triazole products.

The reaction of the an azide and acetylide in an extended active site binding site leads to the production of only the syn product as shown for a click inhibitor of the enzyme acetylcholinesterase in Figure 7.5.11. The product would be a potent inhibitor of the enzyme.

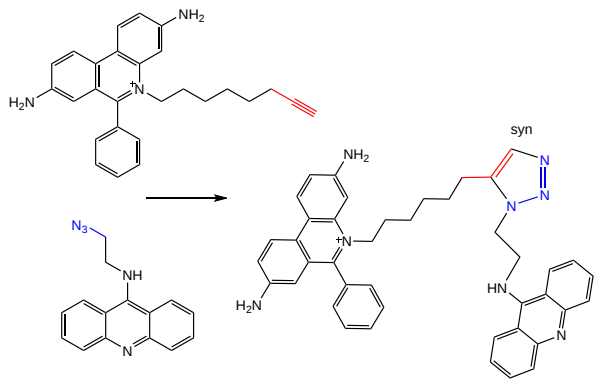
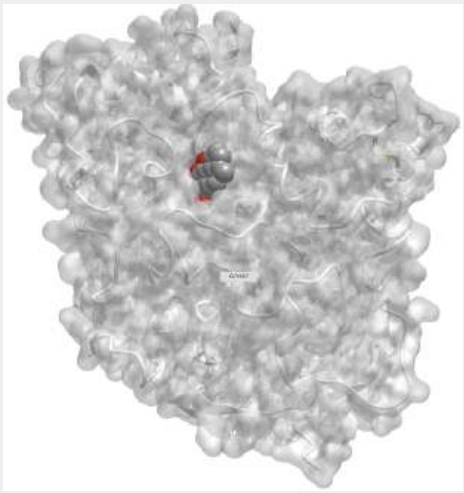
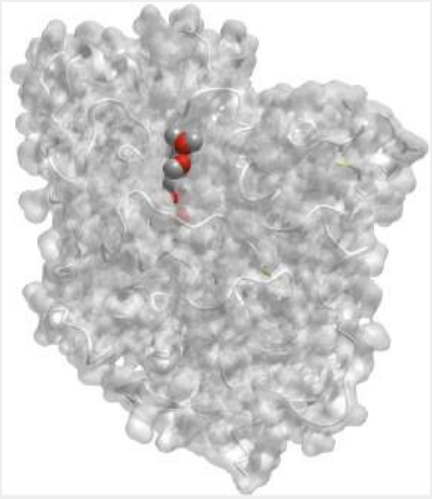


Figure 7.5.11: Click chemistry reaction between azide and acetylide with a directing template

This target-guided synthesis creates a bivalent inhibitor (one that binds at two different regions of an extended binding site). It would have a lower K_D than either of the separate inhibitors.

The enzyme has a catalytic site at the end of very deep (20 Å) and narrow pocket. It also has a peripheral site near the surface of the extended binding pocket. Hence it makes a great potential target for click-chemistry inhibitors.

Figure 7.5.12 shows interactive iCn3D models of human acetylcholinesterase in complex with peripheral site inhibitor dihydrotanshinone I (4M0E) and the protein in complex with peripheral and active site-spanning inhibitor territrem B (4M0F) to illustrate the structural features of the enzyme that make it ideal for click chemistry inhibitors.

Human acetylcholinesterase in complex with the peripheral site inhibitor dihydrotanshinone I (4M0E)	Human acetylcholinesterase in complex with the peripheral and active site-spanning inhibitor territrem B (4M0F)
	
(Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...y8btPa8WMYNj4A	(Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...veBZM1DLSxzKX6

Click chemistry can be used in vivo in **bioorthogonal reactions**. These are specific reactions that take place in potentially reactive biological environments but without interfering with normal biological functions and activities. The reaction is different than when it occurs in a test tube in that a "bioorthogonal reporter such as a fluorphore" (that can be used for instance to track a target molecule in vivo when linked to it) covalently links to a target biomolecule (such as a surface glycan) without influencing its activity. Instead of using Cu ions as catalyst as in *in vitro* click chemistry, normal catalytic mechanisms of the cell are used.

Some more classical condensation reaction can be considered early examples of bioorthogonal reactions. Azides are reactive and are rare biologically, but biomolecules (lipids, proteins, nucleic acids, etc) can be easily labeled with azides or alkynes, facilitating their reaction with click chemistry. Copper ions however can generate ROS, which limits its potential for in vivo click reactions. Here are some classical and more modern bioorthogonal condensing reactions.

- carbonyls with hydrazines and alkoxylamines to form oximes and hydrazones.

- carbonyls and amines to form Schiff bases
- triarylphosphines and organic azides (Staudinger ligation)
- copper(I)-catalyzed azide–alkyne cycloaddition (CuAAC) described above
- strained cyclooctynes and azides in strain-promoted azide–alkyne cycloadditions (SPAAC) which occur fast enough and don't require a copper ion catalyst
- 1,2,4,5-tetrazine and an alkene or alkyne dienophile ([4 + 2]/retro [4 + 2]) cycloaddition to form a dihydropyridazine or pyridazine conjugate in a inverse electron-demand Diels–Alder reaction (IEDDA)

Figure 7.5.13 illustrates newer generation click reactions for connecting two molecules

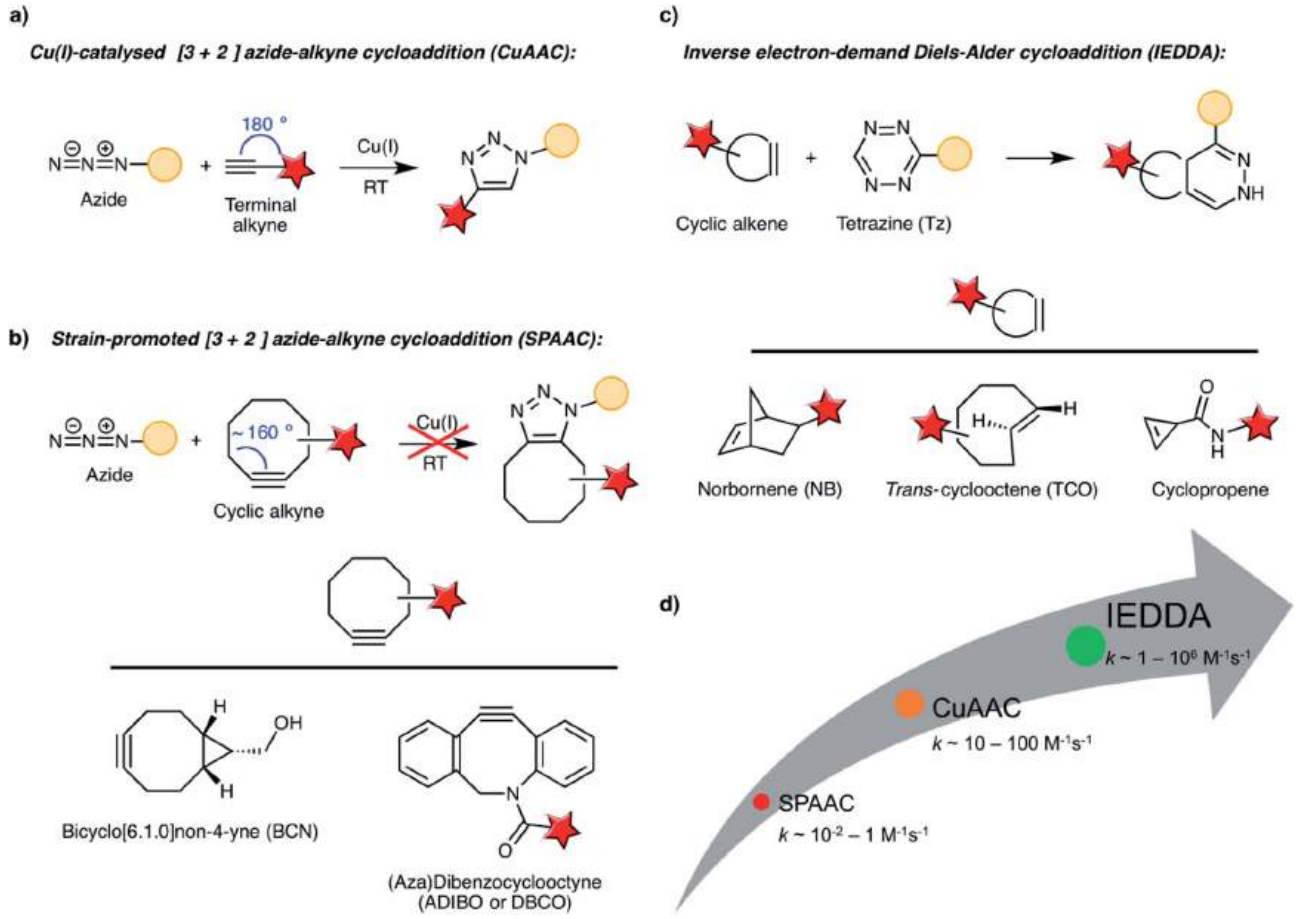


Figure 7.5.13: Newer generation click/bioorthogonal reactions Idiago-L'opez et al., Nanoscale Adv., 2021, 3, 1261. DOI: 10.1039/d0na00873g. Creative Commons Attribution-NonCommercial 3.0 Unported

Panels (a–c) show the main click chemistry reaction use in biochemical labeling reactions. Panel (d) shows a comparison of reaction kinetics of CuAAC, SPAAC and IEDDA.Figure

Figure 7.5.14 shows some applications of bioorthogonal reactions

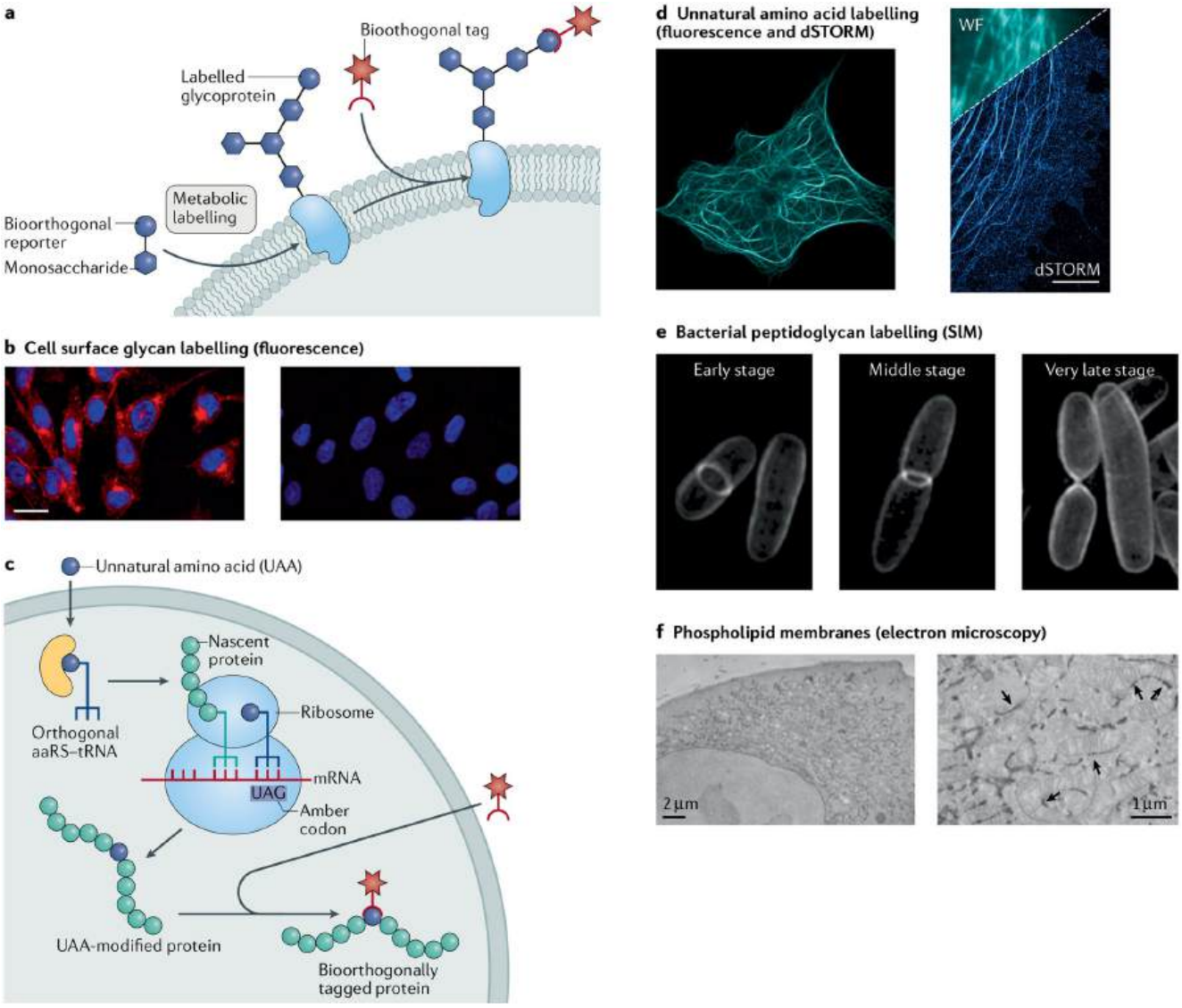


Figure 7.5.14: Applications for labeling different molecule types in cells. Scinto et al. Nat Rev Methods Primers. 2021 ; 1: . doi:10.1038/s43586-021-00028-z. Creative Commons (https://creativecommons.org/licenses/by/4.0/).

Panel a shows a model for metabolic engineering for cell labeling and imaging.

Panel b shows fluorescence microscopy of CHO cells incubated in the presence (left) or absence (right) of peracetylated N-azidoacetylmannosamine (Ac4ManNAz) and labeled with a fluorophore by the Staudinger ligation.

Panel c shows a model for genetic code expansion as a strategy for cell labeling and imaging.

Panel d shows fluorescence and direct stochastic optical reconstruction microscopy (dSTORM) super-resolution images of COS-7 cells where microtubule- microtubule-associated protein was encoded with an unnatural trans-cyclooctene (TCO) amino acid and tetrazine ligation was used to attach a microscopy dye.

Panel e shows structured illumination microscopy (SIM) images of Escherichia coli, where N-azidoacetyl-muramic acid (NAM) was metabolically incorporated into the bacterial peptidoglycan and fluorophore-labeled by copper(I)-catalyzed azide–alkyne cycloaddition (CuAAC).

Panel f shows electron microscopy images of HeLa cells, where azido-choline was metabolically incorporated, and cyclooctyne/azide click chemistry was used to conjugate electron microscopy imaging agents. The arrows indicate sites of endoplasmic reticulum–mitochondria contacts. aaRS, aminoacyl-
aminoacyl-tRNA synthetase.

Figure 7.5.15 shows the application of click chemistry to label surface glycoproteins called integrins, which we will explore in greater detail in Chapter 12.10.

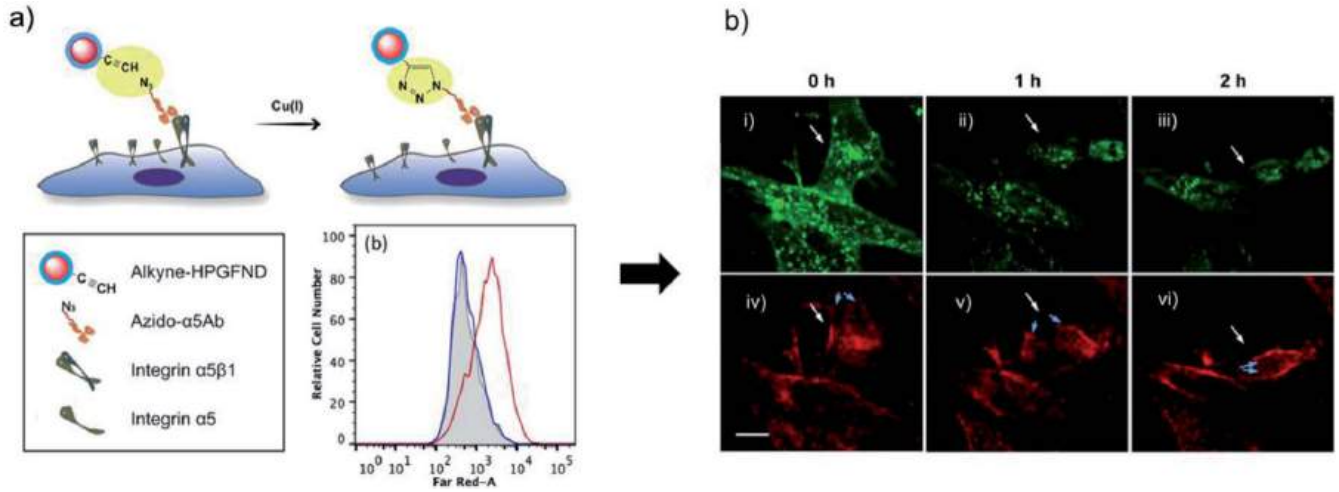


Figure 7.5.15: Bioorthogonal labeling of integrin a5 membrane proteins using azide-modified antibodies and alkyne-HPGFNDs: (a) flow cytometry analysis of fluorescence signals from HFW cells preincubated with Azido-5aAb (red) and control cells (blue). (b) HFW cells labeled with Alexa
Fluor 488-conjugated wheat germ agglutinin (i–iii) and 100 nm alkyne-HPGFNDs (iv–vi). White arrows indicate the cell migration route and blue
arrows show the migration of integrin a5 on cells filopodia. Scale bars: 20 mm. Idiago-L'opez et al., ibid

7.6: Chapter 7 Problems - Answer Key

7.6.1: Chapter 7: Carbohydrates and Glycobiology - Answer Key for Problems

Note: Tables, graphs, diagrams, and figures can be copied from the research papers, the papers must have Creative Commons licenses. Science Advances, Nature Communications, JBC and PLOS journals usually have to right kind of CC licenses.

1. jdkfjdkf

2. djkfjdkjf

3. kjfkdjfkjdk

7.7: Chapter 7 Problems

Chapter 7: Carbohydrates and Glycobiology - Problems

Note: Tables, graphs, diagrams, and figures can be copied from the research papers, the papers must have Creative Commons licenses. Science Advances, Nature Communications, JBC and PLOS journals usually have to right kind of CC licenses.

1. jdkfjdkf

2. djkfjdkjf

3. kjfkdjfkjdk

CHAPTER OVERVIEW

8: Nucleotides and Nucleic Acids

8.1: Nucleic Acids - Structure and Function
8.2: Nucleic Acids - RNA Structure and Function
8.3: Nucleic Acids - Comparison of DNA and RNA
8.4: Chromosomes and Chromatin
8.5: References
8.6: Enzymes for Genetic modifications

8.1: Nucleic Acids - Structure and Function

8.1.1: Introduction to Nucleic Acids

Alongside proteins, lipids, and complex carbohydrates (polysaccharides), *nucleic acids* are one of the four major types of macromolecules that are essential for all known forms of life. The nucleic acids consist of two major macromolecules, *Deoxyribonucleic acid (DNA)* and *ribonucleic acid (RNA)* that carry the genetic instructions for the development, functioning, growth, and reproduction of all known organisms and viruses. Both consist of polymers of a sugar-phosphate-sugar backbone with organic heterocyclic bases attached to the sugars. The sugar in DNA is deoxyribose while in RNA it is ribose. DNA contains four bases, cytosine and thymine (pyrimidine bases) and guanine and adenine (purine bases). DNA in vivo consists of two antiparallel strands intertwined to form the iconic DNA double-stranded helix. RNA is single-stranded but may adopt many secondary and tertiary conformations not unlike that of a protein. Figure 8.1.1 shows a low-resolution comparison of the structure of DNA and RNA.

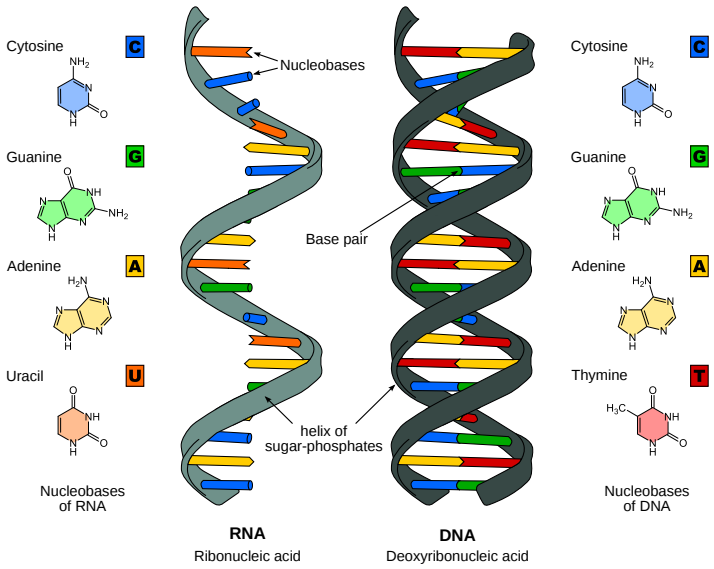


Figure 8.1.1: Low-resolution comparison of the structure of DNA and RNA. https://commons.wikimedia.org/wiki/F...DNA_RNA-EN.svg. Creative Commons Attribution-Share Alike 3.0 Unported license.

The biological function of DNA is quite simple, to carry and protect the genetic code. Its structure serves that purpose well. In the next section, we will study the functions of RNA, which are much more numerous and complicated. The structure of RNA has evolved to serve those added functions.

The core structure of a nucleic acid monomer is the *nucleoside*, which consists of a sugar residue + a nitrogenous base that is attached to the sugar residue at the 1′ position as shown in Figure 8.1.2. The sugar utilized for RNA monomers is ribose, whereas DNA monomers utilize deoxyribose that has lost the hydroxyl functional group at the 2′ position of ribose. For the DNA molecule, four nitrogenous bases are incorporated into the standard DNA structure. These include the *Purines:* Adenine (A) and Guanine (G), and the *Pyrimidines:* Cytosine (C) and Thymine (T). RNA uses the same nitrogenous bases as DNA, except for Thymine. Thymine is replaced with Uracil (U) in the RNA structure.

When one or more phosphate groups are attached to a nucleoside at the 5′ position of the sugar residue, it is called a *nucleotide*. *Nucleotides* come in three flavors depending on how many phosphates are included: the incorporation of one phosphate forms a *nucleoside monophosphate*, the incorporation of two phosphates forms a *nucleoside diphosphate*, and the incorporation of three phosphates forms a nucleoside triphosphate as shown in Figure 8.1.2.

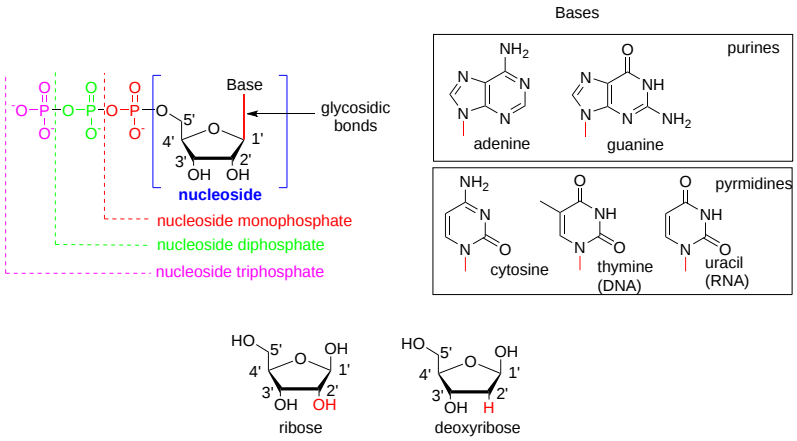


Figure 8.1.2: The Monomer Building Blocks of Nucleic Acids. The site of the nitrogenous base attached to the sugar residue (glycosidic bond) is shown in red.

8.1.2: DNA and RNA Hydrogen-bonded structures

Figure 8.1.3 below shows a "flattened" structure of double-stranded B-DNA that best shows the backbone and hydrogen-bonded base pairs between two antiparallel strands of the DNA. Unlike the protein α-helix, where the R-groups of the amino acids are positioned to the outside of the helix, in the DNA double-stranded helix, the nitrogenous bases are positioned inward and face each other. The backbone of the DNA is made up of repeating sugar-phosphate-sugar-phosphate residues. Bases fit in the double-helical model if pyrimidine on one strand is always paired with purine on the other. From Chargaff's rules, the two strands will pair A with T and G with C. This pairs a keto base with an amino base, and a purine with a pyrimidine. Two H-bonds can form between A and T, and three can form between G and C. This third H-bond in the G:C base pair is between the additional exocyclic amino group on G and the C2 keto group on C. The pyrimidine C2 keto group is not involved in hydrogen bonding in the A:T base pair.

Furthermore, the orientation of the sugar molecule within the strand determines the directionality of the strands. The phosphate group that makes up part of the nucleotide monomer is always attached to the 5′ position of the deoxyribose sugar residue. The free end that can accept a new incoming nucleotide is the 3′ hydroxyl position of the deoxyribose sugar. Thus, DNA is directional and is always synthesized in the 5′ to 3′ direction. Interestingly, the two strands of the DNA double helix lie in opposite directions or have a head-to-tail orientation.

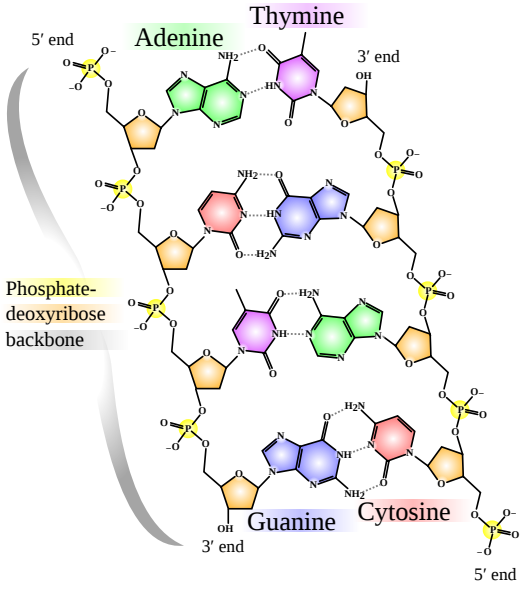


Figure 8.1.3: "Flattened" Structure of DNA Madeleine Price Ball. https://en.Wikipedia.org/wiki/File:D..._structure.svg. Wikimedia Commons

By analogy to proteins, DNA and RNA can be loosely thought to have primary and secondary structures. For a single strand, the primary sequence is just the base sequence read from the 5' to 3' end of the strand, with the bases thought of as "side chains" as illustrated in Figure 8.1.4 for an RNA strand which contains U instead of T.

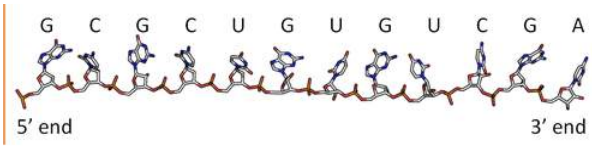


Figure 8.1.4: "Primary" sequence of a single RNA strand. https://en.Wikipedia.org/wiki/Nucleic_acid_sequence. Creative Commons Attribution-ShareAlike License

Since it is found partnered with another molecule (strand) of DNA, the double-stranded DNA, which consists of two molecules held together by hydrogen bonds, might be considered to have secondary structure (analogous to alpha and beta structure in proteins). Of course, the hydrogen bonds are not between backbone atoms but between side chain bases in double-stranded DNA.

Figure 8.1.5 shows an interactive iCn3D model of the iconic structure of a short oligomer of double-stranded DNA (1BNA).

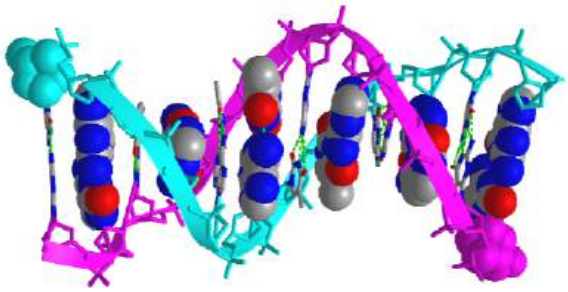


Figure 8.1.5: Iconic structure of a short oligomer of double-stranded DNA (1BNA). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...b5HUbmuQrCobg8

The backbones of the antiparallel strands are magenta (chain A) and cyan (chain B). The 5' sugar-phosphate end of each chain is shown in spacefill and colored magenta (chain A) and cyan (chain B). The hydrogen-bonded interstrand base pairs are shown alternatively in spacefill and sticks to illustrate how the bases stack on top of each other.

Figure 8.1.6 shows types of "secondary (flat representations) and their 3D or tertiary representations found in nucleic acids.

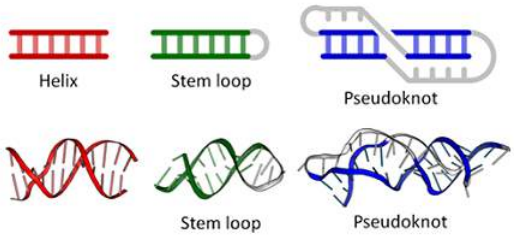


Figure 8.1.6: https://en.Wikipedia.org/wiki/Nucleic_acid_sequence. Creative Commons Attribution-ShareAlike License

Figure 8.1.7 shows an interactive iCn3D model of the tertiary structure of the T4 hairpin loop on a Z-DNA stem (1D16).

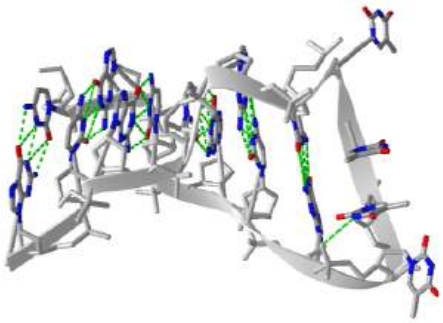


Figure 8.1.7: T4 hairpin loop on a Z-DNA stem (1D16). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...8C7qBqgh8ZTJH9

The hairpin shown is from a synthetic DNA oligomer C-G-C-G-C-G-T-T-T-T-C-G-C-G-C-G which adopts an alternative Z-DNA conformation (which we will explore below) with a loop at one end. The thymine bases 7, 8, and 9 are generally perpendicular to one another and stack together, along with the ribose of T7.

Figure 8.1.8 shows an interactive iCn3D model of pseudoknot in RNA (437D).

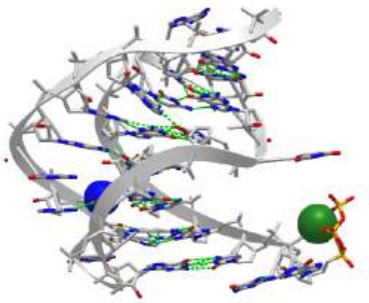


Figure 8.1.8: Minor groove RNA triplex in the crystal structure of a ribosomal frameshifting viral pseudoknot (437D). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...ZtdeJqQXvjCKfA

The pseudoknot has two stems that form a "helix" and two loops. The knot consists of a hairpin in the nucleic acid structure with the loop between the helices paired to another part of the nucleic acid. Pseudoknots can be found in mRNA and ribosomal RNA and affect the translation of the RNA (decoding to instruct the synthesis of a protein sequence). RNA viruses have pseudoknots which likewise affect protein synthesis as well as RNA replication. Pseudoknots also occur in DNA.

8.1.3: Synthesis and structure of DNA

The nucleotide that is required as the monomer for the synthesis of both DNA and RNA is *nucleoside triphosphate*. During the incorporation of the nucleotide into the polymeric structure, two phosphate groups, (P_i-P_i , called *pyrophosphate*) from each triphosphate are cleaved from the incoming nucleotide and further hydrolyzed during the reaction, leaving a *nucleoside monophosphate* that is incorporated into the growing RNA or DNA chain as shown in **Figure 8.1.9** below. Incorporation of the incoming nucleoside triphosphate is mediated by the nucleophilic attack of the 3′-OH of the growing DNA polymer. Thus, DNA synthesis is directional, only occurring at the 3′-end of the molecule.

The further hydrolysis of the pyrophosphate (P_i-P_i) releases a large amount of energy ensuring that the overall reaction has a negative ΔG. Hydrolysis of P_i-P_i $\leftrightarrow$ $2P_i$ has a ΔG = -7 kcal/mol (-29 kJ/mol) and is essential to provide the overall negative ΔG (-6.5 kcal/mol, -27 kJ/mol) of the DNA synthesis reaction. Hydrolysis of the pyrophosphate also ensures that the reverse reaction, pyrophosphorolysis, will not take place removing the newly incorporated nucleotide from the growing DNA chain.

This reaction is mediated in DNA by a family of enzymes known as DNA polymerases. Similarly, RNA polymerases are required for RNA synthesis. A more detailed description of polymerase reaction mechanisms will be covered in Chapters X and Y, covering DNA Replication and Repair, and DNA Transcription.

Figure 8.1.9: Nucleic Acid Synthesis: In nucleic acid synthesis, the 3' OH of a growing chain of nucleotides attacks the α-phosphate on the next NTP to be incorporated (blue), resulting in a phosphodiester linkage and the release of pyrophosphate (PPi). The DNA polymerase further mediates the hydrolysis of the pyrophosphate preventing the reverse reaction from occurring and releasing enough energy to drive the reaction forward. The synthesis of DNA is shown in this diagram. *Image modified from Michal Sobkowski*

DNA was first isolated by Friedrich Miescher in 1869. The double-helix model of DNA structure was first published in the journal *Nature* by James Watson and Francis Crick in 1953 based upon the crucial X-ray diffraction image of DNA from Rosalind Franklin in 1952, followed by her more clarified DNA image with Raymond Gosling, Maurice Wilkins, Alexander Stokes, and Herbert Wilson, and base-pairing chemical and biochemical information by Erwin Chargaff. The prior model was triple-stranded DNA.

The realization that the structure of DNA is that of a double-helix elucidated the mechanism of base pairing by which genetic information is stored and copied in living organisms and is widely considered one of the most important scientific discoveries of the 20th century. Crick, Wilkins, and Watson each received one-third of the 1962 Nobel Prize in Physiology or Medicine for their contributions to the discovery. (Franklin, whose breakthrough X-ray diffraction data was used to formulate the DNA structure, died in 1958, and thus was ineligible to be nominated for a Nobel Prize.)

Watson and Crick proposed two strands of DNA – each in a right-hand helix – wound around the same axis. The two strands are held together by H-bonding between the complementary base pairs (A pairs with T and G pairs with C) as shown in **Figure 8.1.10** below. Note that when looking from the top view, down on a DNA base pair, that the position where the base pairs attach to the DNA backbone is not equidistant, but that attachment favors one side over the other. This creates unequal gaps or spaces in the DNA known as the *major groove* for the larger gap, and the *minor groove* for the smaller gap (Figure 4.5). Based on the DNA sequence within the region, the hydrogen-bond potential created by the nitrogen and oxygen atoms present in the nitrogenous base pairs causes unique recognition features within the major and minor grooves, allowing for specific protein recognition sites to be created.

Figure 8.1.10: The Major and Minor Grooves of DNA. Top view of an (A) A-T base pair and a (B) G-C base pair showing the formation of the major and minor groove sides of the DNA. (C) Side view of the DNA double helix with the major and minor grooves indicated. The DNA backbone is shown in green, potential nitrogen hydrogen-bonding locations are indicated in blue, and oxygen hydrogen-bonding locations are in red. Figure C modified from *dullhunk*

Figure 8.1.1 shows a schematic representation of available hydrogen bond donors and acceptors in the major and minor grove for TA and CG base pairs.

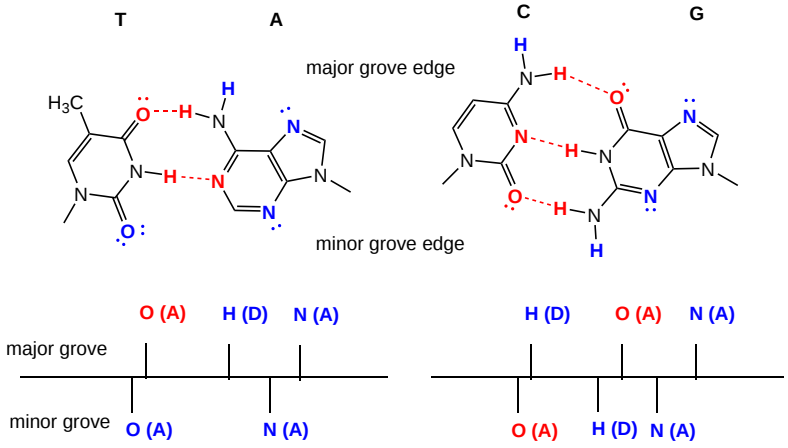


Figure 8.1.11: Available hydrogen bond donors and acceptors in the major and minor groove for TA and CG base pairs

Figure 8.1.12 shows an interactive iCn3D model of DNA showing the major and minor grooves.

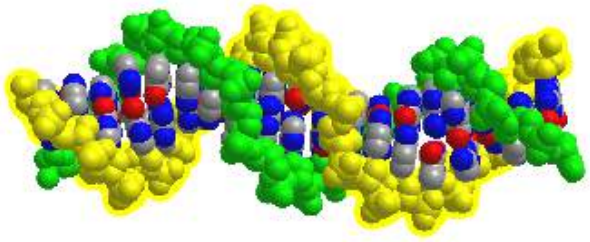


Figure 8.1.12: Major and minor groves of ds-DNA. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?WQWkYi1FrM4DSY2i7

The two sugar-phosphate backbones are shown in green and yellow. Some of the red (oxygen) and blue (nitrogen) atoms in the major grove (and to a much less extent in the minor groove) are not involved in inter-strand G-C and A-T base pairing and so would be available to hydrogen bond donors with specific binding proteins that would display complementary shape and hydrogen bonds acceptors and donors.

Figure 8.1.13 shows an interactive iCn3D model of the N-terminal fragment of the yeast transcriptional activator GAL4 bound to DNA (1D66).

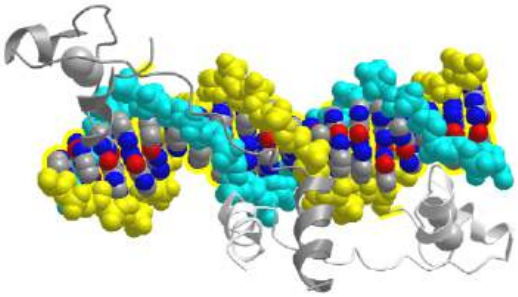


Figure 8.1.13: N-terminal fragment of the yeast transcriptional activator GAL4 bound to DNA (1D66). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...5kLYSSfG7rsmS9

The N-terminal fragment binds to conserved CCG triplets found at both ends of the DNA in the major grove. The protein shown is a dimer held together by a short coiled-coil interaction domain so the site has 2-fold symmetry. A small Zn^{2+}-containing secondary structure motif in each member of the dimer interacts with the major grove. An extended chain connects the DNA binding and interaction domains of each protein.

In addition to the major and minor grooves providing variation within the double helix structure, the axis alignment of the helix along with other influencing factors such as the degree of solvation can give rise to three forms of the double helix, the *A-form (A-DNA)*, the *B-form (B-DNA)*, and the *Z-form (Z-DNA)* as shown in Figure 8.1.14

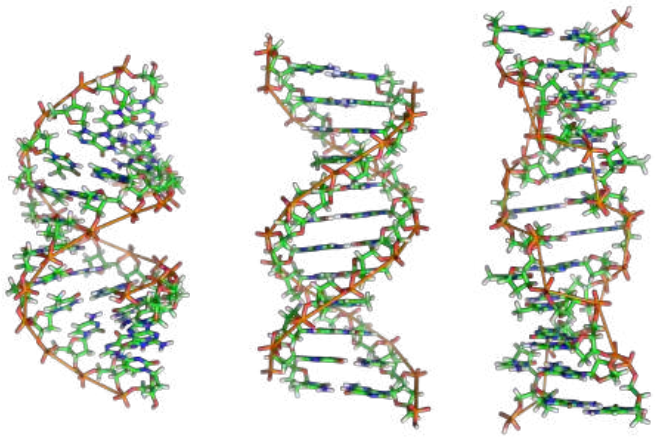


Figure 8.1.14: Structures of A-DNA (left), B-DNA (middle), and Z-DNA (right). https://en.Wikipedia.org/wiki/File:A..._and_Z-DNA.png. Creative Commons Attribution-Share Alike 3.0 Unported

Both the A- and B-forms of the double helix are right-handed spirals, with the B-form being the predominant form found *in vivo*. The A-form helix arises when conditions of dehydration below 75% of normal occur and have mainly been observed *in vitro* during X-ray crystallography experiments when the DNA helix has become desiccated. However, the A-form of the double helix can occur *in vivo* when RNA adopts a double-stranded conformation, or when RNA-DNA complexes form. The 2′-OH group of the ribose sugar backbone in the RNA molecule prevents the RNA-DNA hybrid from adopting the B-conformation due to steric hindrance.

The third type of double helix formed is a left-handed helical structure known as the Z-form or Z-DNA. Within this structural motif, the phosphates within the backbone appear to zigzag, providing the name Z-DNA. *In vitro*, the Z-form of DNA is adopted in short sequences that alternate pyrimidine and purines when high salinity is present. However, the Z-form has been identified *in vivo*, within short regions of the DNA, showing that DNA is quite flexible and can adopt a variety of conformations. A comparison of features between A-, B-, and Z-form DNA is shown in Table 4.1.

Table 4.1 Comparisons of B-DNA, A-DNAand Z-DNA

	B-DNA	A-DNA	Z-DNA
helix sense	Right Handed	Right Handed	Left Handed
base pairs per turn	10	11	12
vertical rise per bp	3.4 Å	2.56 Å	19 Å
rotation per bp	+36°	+33°	-30°
helical diameter	19 Å	19 Å	19 Å

The double-stranded helix of DNA is not always stable. This is because the stair step links between the strands are noncovalent, reversible interactions. Depending on the DNA sequence, denaturation (melting) can be local or widespread and enables various crucial cellular processes to take place, including DNA replication, transcription, and repair.

Both sequence specificity and interaction (whether covalent or not) with a small compound or a protein can induce tilt, roll, and twist effects that rotate the base pairs in the x, y, or z axis, respectively as seen in **Figure 8.1.15**, and can therefore change the helix's overall organization. Furthermore, slide or flip effects can also modify the geometrical orientation of the helix. Hence the flip effects, and (to a lesser extent) the other above-defined movements modulate the double-strand stability within the helix or at its ends. Indeed, under physiological conditions, local DNA 'breathing' has been evidenced at both ends of the DNA helix and B- to Z-DNA structural transitions have been observed in internal DNA regions. These types of locally open DNA structures are good substrates for specific proteins which can also induce the opening of a 'closed' helix. The processes of DNA replication and repair will be discussed in more detail in Chapter 28.

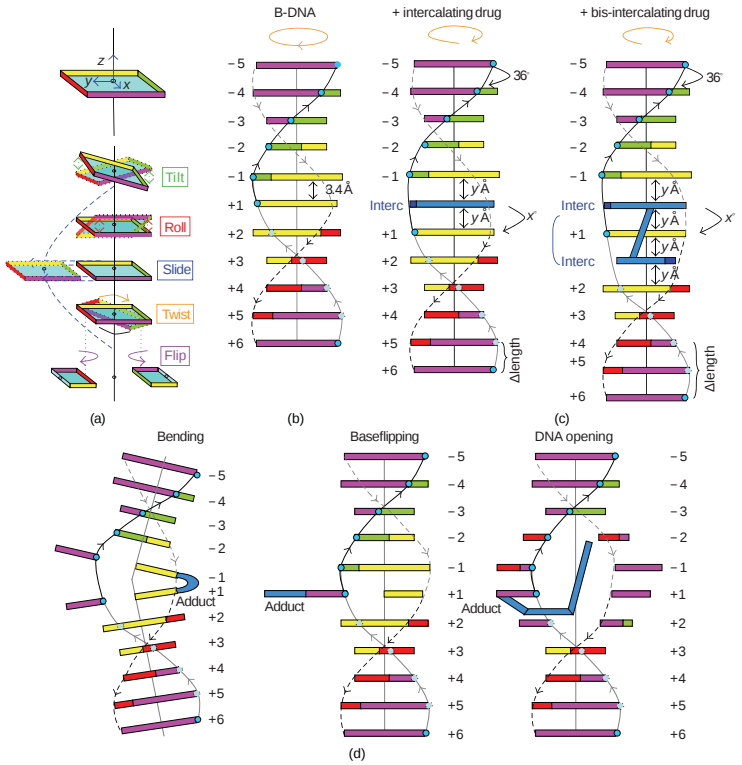
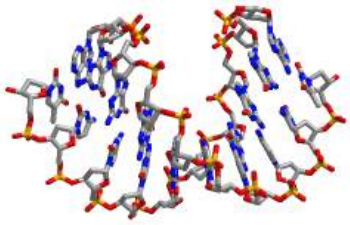
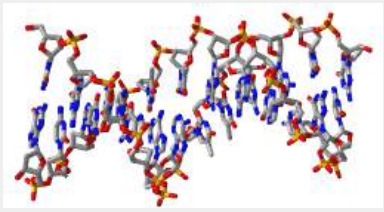
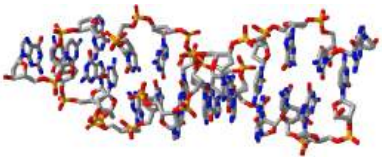


Figure 8.1.15: Localized Structural Modification of the DNA Double Helix. (a) Base pair orientations with the x, y, and z axes result in different kinds of rotation (tilt, roll, or twist) or slipping of the bases (slide, flip) regarding the central helix axis. (b) Matove B-DNA with nearly 11 base pairs within one helical turn. (c) Mono- or bis-intercalation of a small molecule (shown in blue) between adjacent base pairs resulting in an unwinding of the DNA helix (orange arrow on the top) and a lengthening of the DNA helix (ΔLength) depending on the X and y Å values that are specific for a defined DNA intercalating compound. (d) Representation of the DNA bending, base flipping, or double strand opening induced by some DNA destabilizing alkylating agents (adducts shown in blue). Adapted from Calladine and Drew's schematic box representation. Lenglet and David-Cordonnier (2010) Journal of Nucleic Acids, http://dx.doi.org/10.4061/2010/290935. Creative Commons Attribution License,

Figure 8.1.16 shows interactive iCn3D models of A-DNA (top), B-DNA (center), and Z-DNA (bottom). (Copyright; author via source). Click the image for a popup or use the external links in column 1.

A-DNA (440D)	
B-DNA (1BNA)	
Z-DNA (4OCB)	

NCBI iCn3D Figure 8.1.16: A, B and Z-DNA. Click the image for a popup or use the links in column 1

We studied the structure of proteins in-depth, discussing resonance in the peptide backbone, allowed backbone angles φ, ψ, and ω, side chain rotamers, Ramachandran plots, and different structural motifs. We also explored them dynamically using molecular dynamic simulations. We also discussed the thermodynamics of protein stability, and how stability could be altered by changing environmental factors such as solution composition and temperature.

In contrast, our understanding of the structural parameters and the dynamics of nucleic acids is less advanced. This may seem paradoxical, especially given the apparent simplicity of the iconic structure of DNA presented in textbooks. Yet look at the types of secondary structures of nucleic acid presented and then the complicated tertiary and quaternary structures of RNA.

The backbone of nucleic acid has a 5-membered sugar ring, which adds rigidity to the backbone, linked to another sugar ring by $CH_2O(PO_3)O$- connectors, which add some additional conformational freedom. We'll explore the effects of the pentose ring geometry in RNA and DNA in chapter section 8.3. To illustrate a yet unexplored complexity of nucleic acid structure, consider just the orientation of rings in double-stranded DNA and in regions of RNA where double-stranded structures form. The variants in the orientation of the hydrogen-bonded base pairs and the corresponding parameters that define them are shown in Figure 8.1.17.

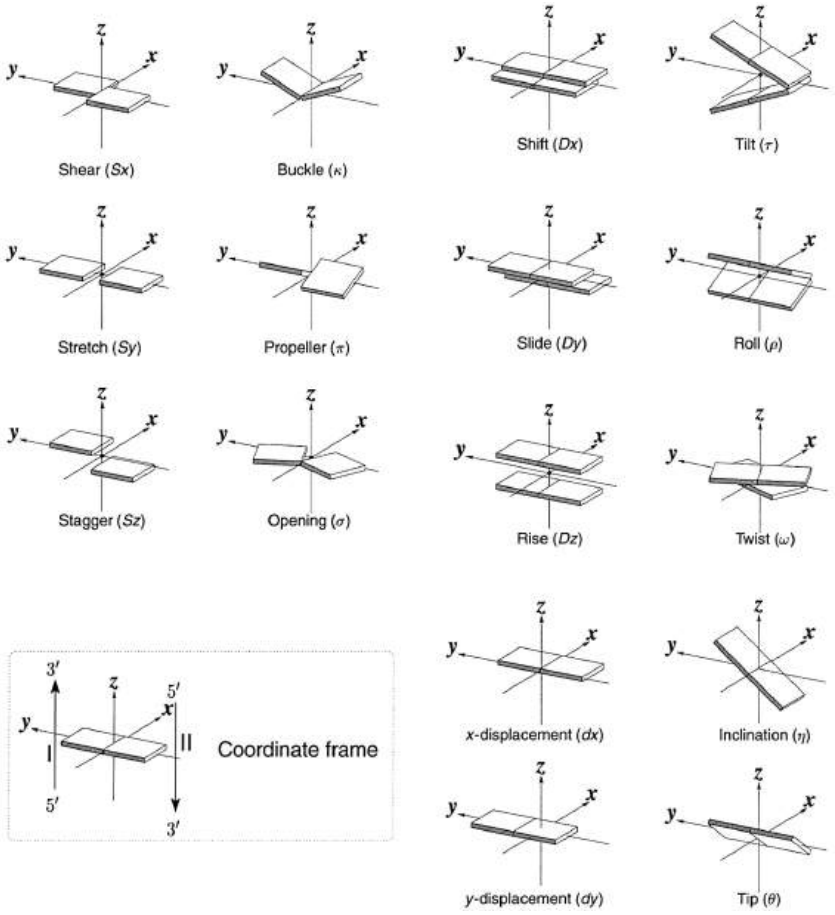


Figure 8.1.17: Base pair orientation and corresponding parameters in nucleic acids. http://x3dna.org/highlights/schemati...air-parameters (with permission). 2008 3DNA *Nature Protocols* paper (NP08), the initial 3DNA *Nucleic Acids Research* paper .

Consider just two of these, the **propellor** and **twist** angles. If you examine the iCn3D models of nucleic acids presented above, you will see that the base pairs are not perfectly flat but are twisted. Larger propeller angles are associated with increased rigidity. The propellor angles for A, B, and Z DNA are +18º, + 16 +/-7 º, and about 0º, respectively. The twist angles A, B, and Z DNA are +33º, +36 º, and -30º, respectively. The lower the twist angle, the higher the number of base pairs per turn. This of course affects the pitch of the helix (the length of one complete turn). All of these terms should be minimized to computationally determine the lowest energy state for a given double-stranded nucleic acid.

8.1.4: Alternative Base Pairing in DNA and RNA

A first glance at a DNA or RNA structure reveals a myriad of possible hydrogen bond donors and acceptors in the bases of the nucleic acid. Hence it should come as no surprise that a variety of alternative or **noncanonical** (not in the canon or dogma) intermolecular hydrogen bonds can form between and among bases, leading to alternatives to the classical Watson-Crick base pairing. There are 28 possible base pairs with two hydrogen bonds between them. As structure determines function and activity, these alternative structures also influence DNA/RNA function. We will consider four different types of noncanonical base pairing: reverse Watson Crick, wobble, Hoogsteen, and reverse Hoogsteen base pairs.

In DNA, these types of noncanonical base pairs can occur when bases become mismatched in double-stranded regions. In RNA, which we will explore more fully in Chapter 8.2, double-stranded molecules form by separate RNA molecules aren't common. Instead, the molecule folds on itself in 3D space to form a complex tertiary structure containing regions of helical secondary structure. RNAs also form quaternary structures when bound to other nucleic acids and proteins. Larger RNAs have loops with complex secondary and tertiary structures which often require noncanonical base pairing, stabilizesbilize the alternative structures. The noncanonical structures are also important for RNA-protein interactions in the RNA region which binds proteins. If one considers RNA and protein binding as a coupled equilibrium, it should be clear that protein binding to RNA might also induce conformation changes, specifically noncanonical base pairs, in the RNA. For example, the HIV Rev peptide binds to a target site in the envelop gene of HIV (which has an RNA genome) and leads to the formation of an RNA loop with hydrogen bonding between two purines.

Figure 8.1.18 shows an interactive iCn3D model of the REV Response element RNA complexed with REV peptide (1ETF).

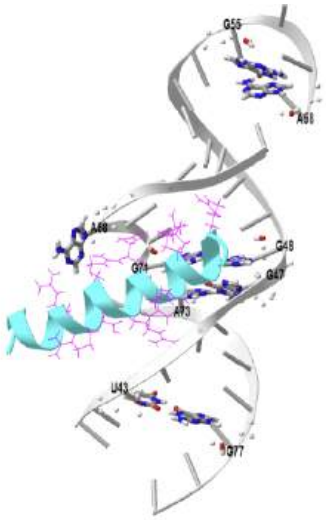


Figure 8.1.18: REV Response element RNA complexed with REV peptide (1ETF). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...T8CJ3pCe986Vx9

The peptide is shown in cyan and its arginine side chains are shown as cyan lines. There are an extraordinary number of arginines that form ion-ion interactions with the negatively charged phosphates in the major grove of this double-stranded A-RNA. The noncanonical base pairs are shown in CPK-colored sticks. A wobble base, U43-G77, see below, is shown as well as three homopurine base pairs, G47-A73, G55-A58, and G48-G71. The solitary A68 base is shown projecting away from the RNA.

Figure 8.1.19 shows the Watson Crick and the first set of alternative non-canonical base pairs.

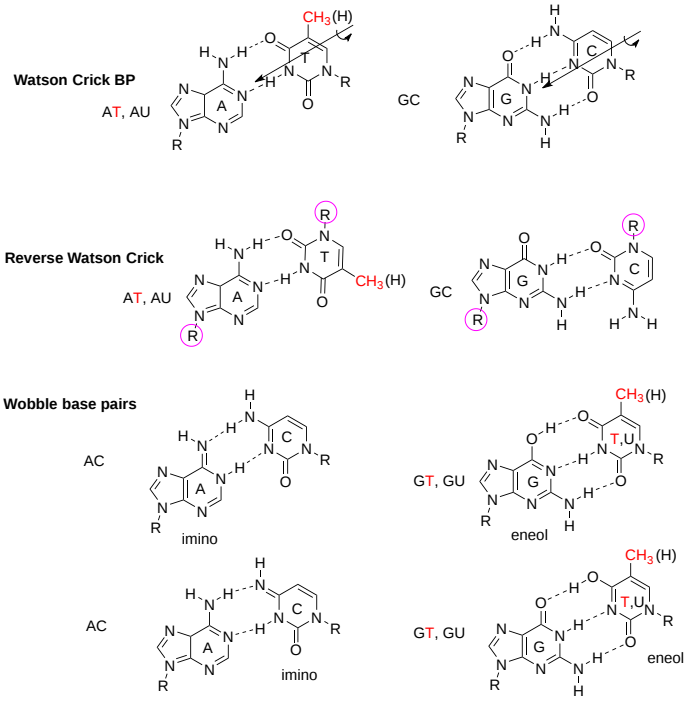


Figure 8.1.19: Some noncanonical base nucleic acid base pairs

Let's look at them in more detail.

Reverse Watson Crick: The reverse Watson-Crick AT (AU) and GC pairs can sometimes be found at the end of DNA strands and also in RNA. In forming the reverse base pairs, the pyrimidine can rotate 180^{o} along the axis shown and then rotate in the plane to align the hydrogen bond donors and acceptors as shown in the top part of the figure. The glycosidic bond between the N in the base and the sugar (the circled R group) is now in an "antiparallel" arrangement in the reverse base pair.

Wobble Base Pairs

The bases in nucleic acids can undergo tautomerization to produce forms that can base pair noncanonically. They are termed **wobble** base pairs and include G-T(U) base pairs from keto–enol tautomerism and A-C base pairs from amino–imino tautomerism, as illustrated in Figure 18 above.

Figure 8.1.20 shows an interactive iCn3D model of the GT Wobble Base-Pairing in Z-DNA form of d(CGCGTG) (1VTT). Two such GT pairs are found in the structure.

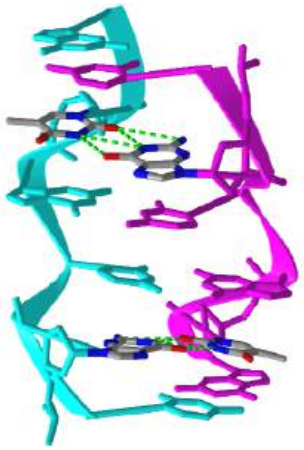


Figure 8.1.20: GT Wobble Base-Pairing in Z-DNA form of d(CGCGTG) (1VTT). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...LtwfzyeqDCaPEA

The water around the wobble base pairs can form hydrogen bonds and stabilize the pair if a hydrogen bond is missing.

Figure 8.1.21 shows an interactive iCn3D model of dsRNA with G-U wobble base pairs (6L0Y).

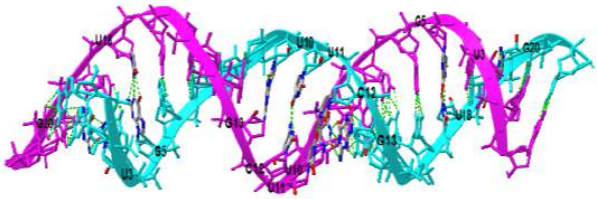


Figure 8.1.21: dsRNA with G-U wobble base pairs (6L0Y). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...reyeD6JQM1djq6

The structure contains many GU wobble base pairs as well as two CU base pairs between two pyrimidine bases.

Inosine, a variant of the base adenine, can be found in RNA. It is formed by the deamination of adenosine by the enzyme adenosine deaminase. A nucleotide having inosine is named hypoxanthine. Hypoxanthine can form the wobble base pairs I-U, I-A, and I-C when incorporated into RNA, as illustrated in Figure 8.1.22

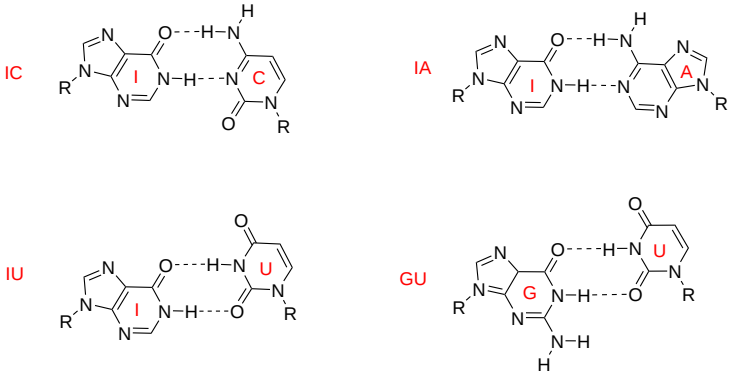


Figure 8.1.22 Wobble bases pairs using hypoxanthine with the base inosine

Wobble base pair interactions are especially important in translation when a protein sequence is made from a messenger RNA template (which will be discussed in Unit III). For that decoding process to occur, two RNA molecules, messenger RNA (mRNA) and a transfer RNA (t-RNA) covalently attached to a specific amino acid like glutamic acid, must bind to each other through a 3 base pair interaction. The 3 bases on the mRNA are called the codon, and the 3 complementary bases on the tRNA are called the anticodon. The triplet base pair are antiparallel to each other. The interaction between mRNA and tRNA are illustrated in Figure 8.1.23

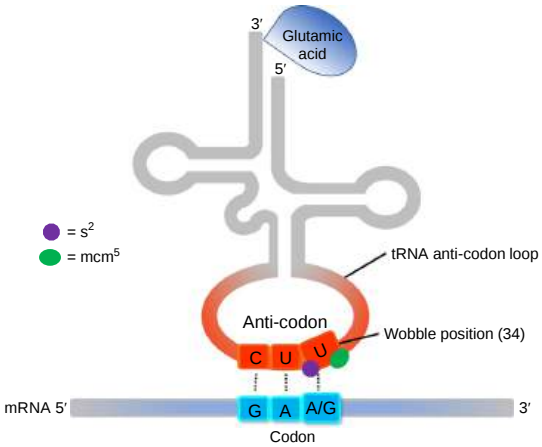


Figure 8.1.23 The wobble uridine (U34) of tRNA molecules that recognize both AAand AG-ending codons for Lys, Gln, and Glu, is modified by the addition of both a thiol (s2) and a methoxy-carbonyl-methyl (mcm5). This double modification enhances the translational efficiency of AA-ending codons. Goffena, J et al. *Nat Commun* **9,** 889 (2018). https://doi.org/10.1038/s41467-018-03221-z. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

The third 3' base on the mRNA is less restricted and can form noncanonical, specifically, wobble base pairs, with the 5' base in the anti-codon triplet of tRNA. The term wobble arises from the subtle conformational changes used to optimize the pairing of the triplets. Wobble bases occur much more in tRNA than other nucleic acids.

Hoogsteen base pairing

Flexibility in DNA allows rotation around the C1'-N glycosidic bond connecting the deoxyribose and base in DNA, allowing different orientations of AT and GC base pairs with each other. The normal "anti" orientation allows "Watson-Crick" (WC) base pairing between AT and GC base pairs while the altered rotation allows "Hoogsteen" base pairs. The different orientations for an AT base pair are shown in Figure 8.1.24

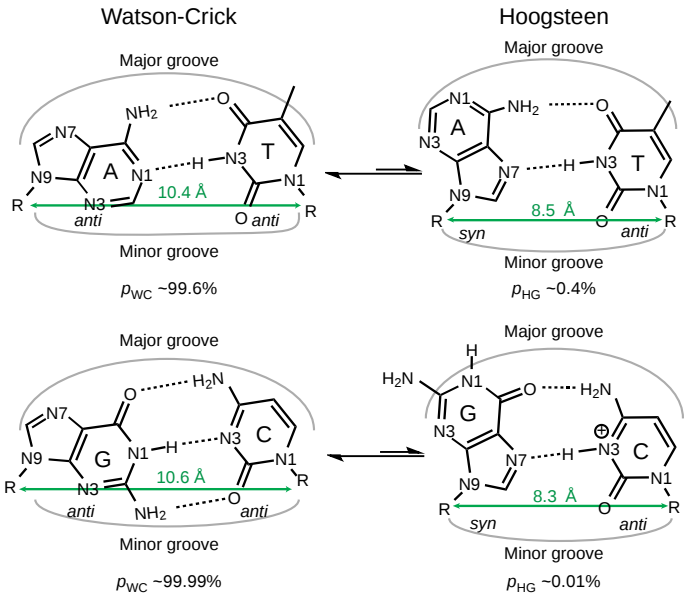


Figure 8.1.24: Xu, Y., McSally, J., Andricioaei, I. *et al.* Modulation of Figure 8.1.*x*aHoogsteen dynamics on DNA recognition. *Nat Commun* **9,** 1473 (2018). https://doi.org/10.1038/s41467-018-03516-1Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Hoogsteen base pairing is usually seen when DNA is distorted through interactions with bound proteins and drugs that intercalate between base pairs. Figure 8.1.25 shows an interactive iCn3D model of a Hoogsteen base pair embedded in undistorted B-DNA - MATAlpha2 homeodomain bound to DNA (1K61).

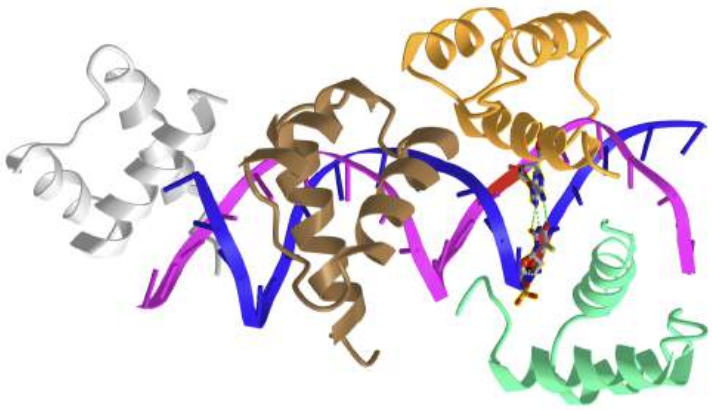


Figure 8.1.25: A Hoogsteen base pair embedded in undistorted B-DNA - MATAlpha2 Homeodomain bound to DNA (1K61). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...SLLRv1m8HQXKcA

The same DNA without bound protein has no Hoogsteen base pairs. To form Hoogsteen base pairs, a rotation around the glycosidic-base bond must occur. Hoogsteen base pairs between G and C can also occur on rotation but in addition, the N3 of

cytosine is protonated, as shown in Figure 14 above.

Evidence suggests that Hoogsteen base pairing may be important in DNA replication, binding, damage, or repair. They can induce kinking of the DNA near the major grove.

There are also examples of reverse Hoogsteen base pairing, as shown in Figure 8.1.26

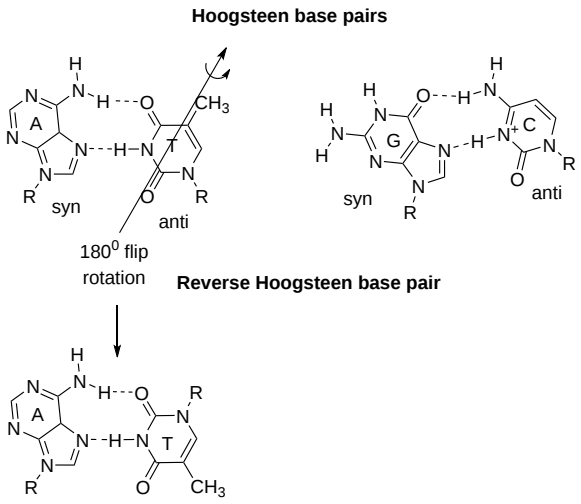


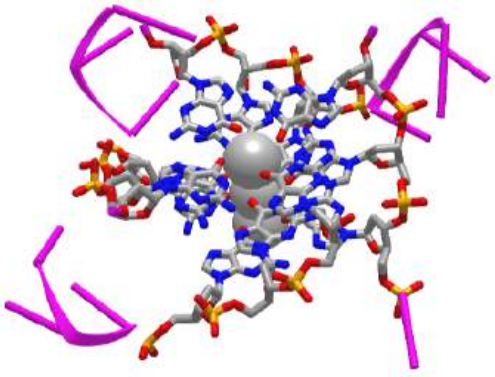
Figure 8.1.26: The reverse Hoogsteen AT base pair

8.1.5: Additional Alternative Structures: Quadruplexes, Triple Helices, and 4-Way Junctions

Quadruplexes

These can be formed in DNA and RNA from G-rich sequences involving tetrads of guanine bases that are hydrogen bonded. They are a bit hard to describe in words so let's first examine one particular structure.

In human cells, telomeres (the ends of chromosomes) contain 300-8000 repeats of a simple **TTAGGG** sequence. The repetitive **TTAGGG** sequences in telomeric DNA can form quadruplexes. Figure 8.1.27 shows an interactive iCn3D model of parallel quadruplexes from human telomeric DNA (1KF1). The structure contains a single DNA strand **(**5'-AGGGTTAGGGTTAGGGTTAGGG-3') which contains four TTAGGG repeats.



NCBI iCn3D Figure 8.1.27: parallel quadruplexes from human telomeric DNA (1KF1). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...y5joFHDgWJQsQ6

Rotate the model to see 3 parallel layers of quadruplexes. In each layer, **4** noncontiguous guanine bases interact with a K^+ ion. Hover over the guanine bases in one layer and you will find that one layer consists of guanines 4, 10, 16, and 22, which derive from the last G in each of the repeats in the sequence of the oligomer used (5'-AGGGTTAGGGTTAGGGTTAGGG-3'). These quadruplexes certainly serve in recognition and as binding site for telomerase proteins. The guanine-rich telomere sequences which can form quadruplex may also function to stabilize chromosome ends

A Quadruplex can be formed in 1 strand of nucleic acid (as in the above model) or from 2 or 4 separate strands. They also must have at least 2 stacked triads. As in the example above, single-stranded sections can form intramolecular G-quadruplex from a $G_mX_nG_mX_oG_mX_pG_m$ sequence, where m is the number of Gs in each short segment (3 in the structure above). If a segment is longer than others, a G might be in a loop.

Triple Helices

These structures can occur in DNA (and also RNA) that contain homopurine and homopyrimidine sequences that have a mirror repeat symmetry. Hence they can occur naturally. A mirror repeat contains a center of symmetry on a single strand. Here is an example: 5'-GCATGGTACG-3'.

They can also occur when a third single-strand DNA (called a triplex-forming oligonucleotide or TFO) binds to a double-stranded DNA. The TFOs bind through Hoogsteen base pairing in the major grove of the ds-DNA. They can bind tightly and specifically in a parallel or antiparallel fashion. Specific and locally higher concentrations of divalent cations or positively charged polyamines like spermine act to stabilize the extra negative charge density from the binding of a third polyanionic DNA strand.

An example of a triple helix system that has been studied in vitro is shown in Figure 8.1.28

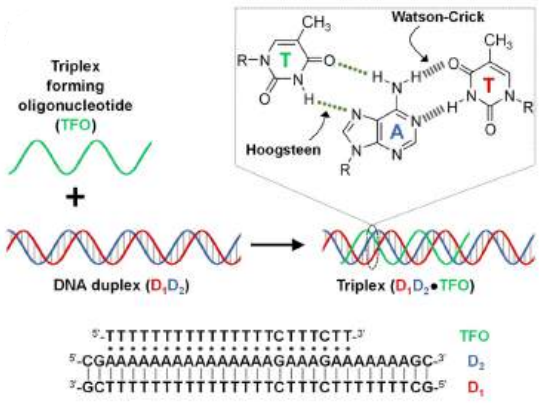


Figure 8.1.28 Intermolecular triplex formation and their oligonucleotide sequences (where "•" and "-" indicates Hoogsteen and Watson–Crick base pairings, respectively). Inset: chemical structure of a parallel T•AT triplet. Guerrini, L. and Alvarez-Puebla, R.A. Nanomaterials 2021, 11, 326. https://doi.org/10.3390/nano11020326. Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/)

The double-stranded canonical helix (D_1D_2) consists of 31 base pairs in which strand D_1 is pyrimidine-rich and D_2 is purine-rich strand (D_2). A 22-nucleotide Triple helix forming oligonucleotide (TFO) that is rich in pyrimidines binds the 19 AT and 2 C-GC base triplets. The TFO binds along the major grove of the D_2 strand which is purine rich.

If the binding of the third strand in the major groove occurs at the site where RNA polymerase binds to a gene, then the third strand can inhibit gene transcription. Binding can also lead to a mutation or recombination at the site.

Figure 8.1.29 shows the base pairing of purine and pyrimidines of the third strand to the canonical AT dn GC base pairs of the original double-stranded DNA.

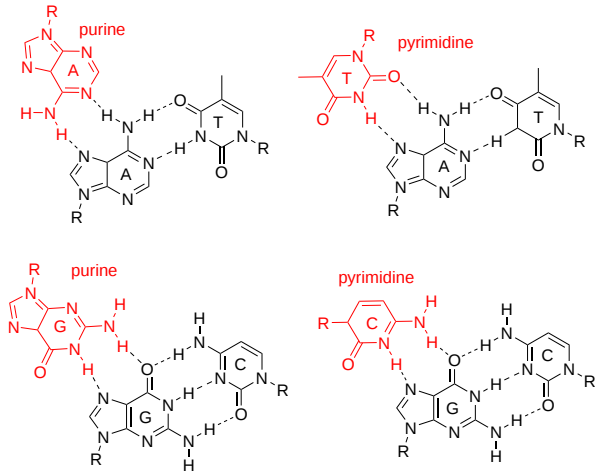


Figure 8.1.29: Base pairing in triple helix motifs. (after Jain et al. Biochimie. 2008. doi: 10.1016/j.biochi.2008.02.011

Figure 8.1.30 shows an interactive iCn3D model of a solution conformation of a parallel DNA triple helix (1BWG).

Figure 8.1.30: Solution conformation of a parallel DNA triple helix (1BWG). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...5JU813eNjND8E7

Triple helix formation can also occur within a single strand of DNA. The resulting structure is called **H-DNA**. An example is shown below. Note that the central blue, black, and red sequences are all mirror image repeats (around a central nucleotide). During processes that unravel DNA (replication, transcription, repair), the self-association of individual mirror repeats can form a locally stable triple helix, as shown in Figure 8.1.31.

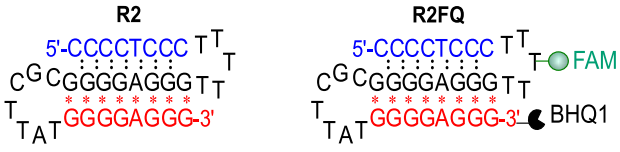


Figure 8.1.31: Schematic illustrations of (A) the H-DNA or intramolecular triplex structure used in this study; del Mundo et al. (2019) Nucleic acids research. 47. e73. 10.1093/nar/gkz237. Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/)

The * between in the G*G and A*A denote Hoogsteen hydrogen bonding (purine motifs) in this intramolecular triple helix. Reverse Hoogsteen hydrogen bonds can also occur.

Triple helices can form when single-stranded DNA formed during replication, transcription or DNA repair with half of the required mirror symmetry folds back into the adjacent major grove and base pairs using Hoogsteen/reverse Hoogsteen bonding, which can be stabilized by Mg^{2+}.

Recent Updates: **Four-Way Junctions**

Nonhelical sections of DNA can, as we will see in the next section on RNA, can bind small target molecules through noncovalent interactions. (RNA examples that we will see in the next chapter section include aptamers, ribozymes and riboswitches.) One example is "lettuce" single-stranded DNA that can bind small fluorophores modeled after the intrinsic fluorophore of the green fluorescent protein. When bound to the lettuce DNA, the fluorescence of the fluorophore is dramatically enhanced. Figure 8.1.32 below shows the structure of extrinsic DNA fluorophores based on GFP that bind to the single-stranded "lettuce" DNA.

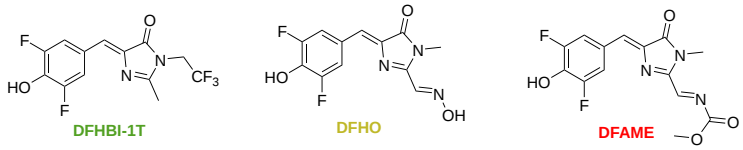


Figure 8.1.32 Structure of extrinsic DNA fluorophores based on GFP that bind to DNA. The font color of the names indicates the color of the emitted fluorescence.

Figure 8.1.33 shows an interactive iCn3D model of a solution conformation of a ssDNA:DFAME fluorophore complex (8FI0). The blue dotted lines shown π-π stacking interactions and the green dotted line a hydrogen bond.

Figure 8.1.33: Solution conformation of a ss-53 mer DNA:DFAME fluorophore complex (8FI0). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...97eLqwNWmaTNC9

The DFAME ligand as shown in sticks.

Figure 8.1.34 shows a closeup of DFAME (colored spacefill) bound to the lettuce DNA.

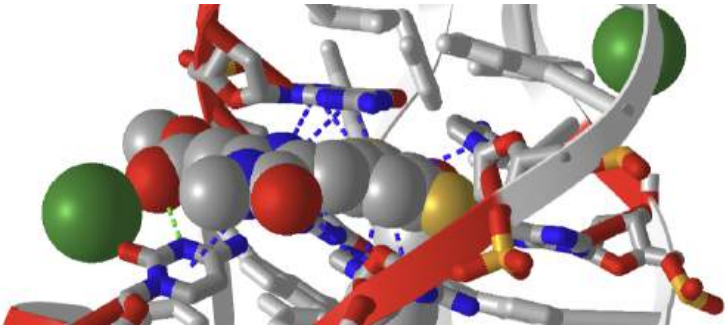


Figure 8.1.34: Pi stacking interactions (blue dotted lines) between the extrinsic fluorophore DFAME (spacefill) and ssDNA.

The DNA fold is characterized as a four-way junction (also seen in RNA but they are more L or H-shaped). On either end are B-DNA duplexes and the ssDNA between them forms stem-loops with odd base pairing in the stems. The overall structure is like a cloverleaf. Two coaxial stacks of nucleotides form what is called a G-quadruplex where the fluorophore binds. Pi base stacking between diagonally packed bases, along with the binding of Mg^{2+} and K^{+}, stabilize the structure.

8.1.6: Stability of nucleic acids

After looking at the myriad of structures showing the nearly parallel hydrogen-bonded base pairs, and from ideas from most textbooks and classes you have taken, you probably think that double-stranded DNA is held together and stabilized by hydrogen bonds between the bases. It is well known that the greater the percentage of GC compared to AT, the greater the stability of the dsDNA, which translates into a higher "melting temperature (T_M)", the temperature at which the dsDNA is converted to ssDNA.

There is a linear relationship between GC content and T_M. The figures above show that GC base pairs have 3 inter-base hydrogen bonds compared to 2 in AT base pairs. These observations support the simple notion that inter-base hydrogen bonds are the source of dsDNA stability.

You would be in general correct in this belief, but you'd be missing the more important contributor to ds-DNA stability, base (π) stacking, and the noncovalent interactions associated with the stacking. The main contributors to stability are hydrophobic interactions in the anhydrous hydrogen-bonded base pairs in the helix. Given that the hydrogen bond donors and acceptors that contribute to base pairing exist in the absence of competing water, the donors and acceptors are free to fully engage in bonding. The hydrogen bond interaction energy is hence more favorable in the stack. The stacking energy is similar for an AT-AT stack and a GC-GC stack (about -9.8 kcal/mol, 41 kJ/mol). Hence AT and GC base pairs contribute equally to stability. The excess stability of dsDNA enriched in GC base pairs can still be explained by the extra stabilization for an additional hydrogen bond per GC base pair

Proteins are stabilized by a myriad of interactions, but the folded state is marginally more stable than the ensemble of the unfolded state. Marginal stability is important as protein conformation often must be perturbed on binding and ensuing function. The same must be true of double-stranded DNA, which must "unfold' or separate on replication, transcription, and repair. It is well known that dsDNA structure is sensitive to hydration (see the section on A, B, and Z DNA). Small molecules like urea, as we saw with proteins, can also denature DNA into single strands.

DNA must be stable enough to be the carrier of genetic information but dynamic enough to allow events that required partial unfolding. Other water-soluble molecules like ethylene glycol ethers (polyethylene glycol-400) and diglyme (dimethyl ether of diethylene glycol), which are more hydrophobic than water, appear to reduce base stacking interactions while maintaining them, and at the same time allow a longitudinal extension or breathing of the helix. This dynamic extension may be required for transitions of B-DNA to Z-DNA, for example. The extensions also allow transient "hole" to appear between base pairs which might assist in the binding of intercalating agents like some transition metal complexes. The extension caused by these ethers and natural extensions would decrease base stacking but appear at the same time to strengthen the hydrogen bonding between bases.

Longitudinal helical extensions might be important when homologous gene recombine. In that process the homologous DNA strand but exchange with a paired homolog. This processing is associated with strand extension and disruption of base pair at every third base. Recombination also must allow chain extension as it maintains base-pairing fidelity.

DNA structures get obviously more complicated as it packs into the nucleus of a cell and form chromosomes, as shown in Figure 8.1.35. We will study the packing of DNA in other sections.

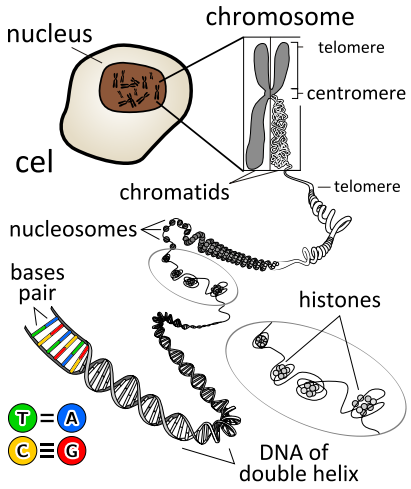


Figure 8.1.35 : Packing of DNA into the chromosome. https://commons.wikimedia.org/wiki/F...omosome_en.svg .

8.1.7: References

Börner, R., Kowerko, D., Miserachs, H.G., Shaffer, M., and Sigel, R.K.O. (2016) Metal ion induced heterogeneity in RNA folding studied by smFRET. *Coordination Chemistry Reviews* 327 DOI: 10.1016/j.ccr.2016.06.002 Available at: https://www.researchgate.net/publication/303846502_Metal_ion_induced_heterogeneity_in_RNA_folding_studied_by_smFRET

Hardison, R. (2019) B-Form, A-Form, and Z-Form of DNA. Chapter in: R. Hardison's *Working with Molecular Genetics*. Published by LibreTexts. Available at: https://bio.libretexts.org/Bookshelves/Genetics/Book%3A_Working_with_Molecular_Genetics_(Hardison)/Unit_I%3A_Genes%2C_Nucleic_Acids%2C_Genomes_and_Chromosomes/2%3A_Structures_of_Nucleic_Acids/2.5%3A_B-Form%2C_A-Form%2C_and_Z-Form_of_DNA

Lenglet, G., David-Cordonnier, M-H., (2010) DNA-destabilizing agents as an alternative approach for targeting DNA: Mechanisms of action and cellular consequences. *Journal of Nucleic Acids* 2010, Article ID: 290935, DOI: 10.4061/2010/290935 Available at: https://www.hindawi.com/journals/jna/2010/290935/

Mechanobiology Institute (2018) What are chromosomes and chromosome territories? *Produced by the National University of Singapore*. Available at: https://www.mechanobio.info/genome-regulation/what-are-chromosomes-and-chromosome-territories/

National Human Genome Research Institute (2019) The Human Genome Project. *National Institutes of Health*. Available at: https://www.genome.gov/human-genome-project

Wikipedia contributors. (2019, July 8). DNA. In *Wikipedia, The Free Encyclopedia*. Retrieved 02:41, July 22, 2019, from https://en.Wikipedia.org/w/index.php?title=DNA&oldid=905364161

Wikipedia contributors. (2019, July 22). Chromosome. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:18, July 23, 2019, from en.Wikipedia.org/w/index.php?title=Chromosome&oldid=907355235

Wikilectures. Prokaryotic Chromosomes (2017) In MediaWiki, Available at: https://www.wikilectures.eu/w/Prokaryotic_Chromosomes

Wikipedia contributors. (2019, May 15). DNA supercoil. In *Wikipedia, The Free Encyclopedia*. Retrieved 19:40, July 25, 2019, from en.Wikipedia.org/w/index.php?title=DNA_supercoil&oldid=897160342

Wikipedia contributors. (2019, July 23). Histone. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:19, July 26, 2019, from en.Wikipedia.org/w/index.php?title=Histone&oldid=907472227

Wikipedia contributors. (2019, July 17). Nucleosome. In *Wikipedia, The Free Encyclopedia*. Retrieved 17:17, July 26, 2019, from en.Wikipedia.org/w/index.php?title=Nucleosome&oldid=906654745

Wikipedia contributors. (2019, July 26). Human genome. In *Wikipedia, The Free Encyclopedia*. Retrieved 06:12, July 27, 2019, from en.Wikipedia.org/w/index.php?title=Human_genome&oldid=908031878

Wikipedia contributors. (2019, July 19). Gene structure. In *Wikipedia, The Free Encyclopedia*. Retrieved 06:16, July 27, 2019, from en.Wikipedia.org/w/index.php?title=Gene_structure&oldid=906938498

8.2: Nucleic Acids - RNA Structure and Function

8.2.1: RNA: Structure and Function

Ribonucleic acids are very similar in chemical structure to DNA except they contain ribose instead of deoxyribose. They also have the pyrimidine base uracil instead of thymine, as shown in Figures 1 and 2 above. These two small changes (but mostly the first) confer on it a very different set of biological functions than DNA. This should not surprise us and the basis of all chemistry and biochemistry is that chemical structure determines chemical and biochemical functions and activities. We discussed in the previous section how RNA can adopt complex tertiary structures, which requires the presence of more noncanonical base pairs and chemical modification of bases. In this section we wi,ll explore the plethora of different types of RNA structures and their functions.

The sequence of RNA is made from DNA through a process called transcription (converting the information of DNA, a nucleic acid, into RNA, another nucleic acid). RNA can form double-stranded helices but typically these are viral in origin. DsRNA is a pathogen-associated molecular pattern (PAMP) that binds Toll-like receptor 3 (TLR3) as we saw in Chapter 5.5. If both strands of DNA are transcribed, the resulting strands can anneal form dsRNA. In addition, a single strand of RNA can fold on itself if the 5' and 3' ends are complementary to form a stem-hairpin loop. Figure 8.2.1 shows a stem-loop from a messenger RNA (4QOZ) when it is bound to a specific RNA binding protein (not shown).

Figure 8.2.1 A stem-loop from a messenger RNA (4QOZ) when it is bound to a specific RNA binding protein (not shown).

Larger ss-RNA can form tertiary structures with many regions of intrastrand hydrogen bonds forming secondary structures, as shown in Figure 8.2.2 for one type of RNA called a transfer RNA

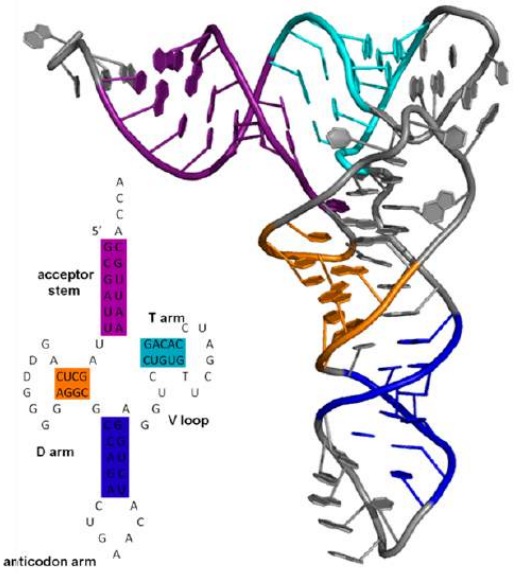


Figure 8.2.2: Secondary and tertiary structure of tRNA showing unpaired (gray) and hydrogen-bonded secondary structure regions (color). Kyle Schneider https://commons.wikimedia.org/w/inde...curid=12309266

Figure 8.2.3 a computed model for secondary structure within a much larger single RNA molecule S11 that is part of the ribosome. The figure shows color-coded differences in accessibility when the S11 RNA is free (blue) and bound to the protein NSP2 (red), which induces structural rearrangements.

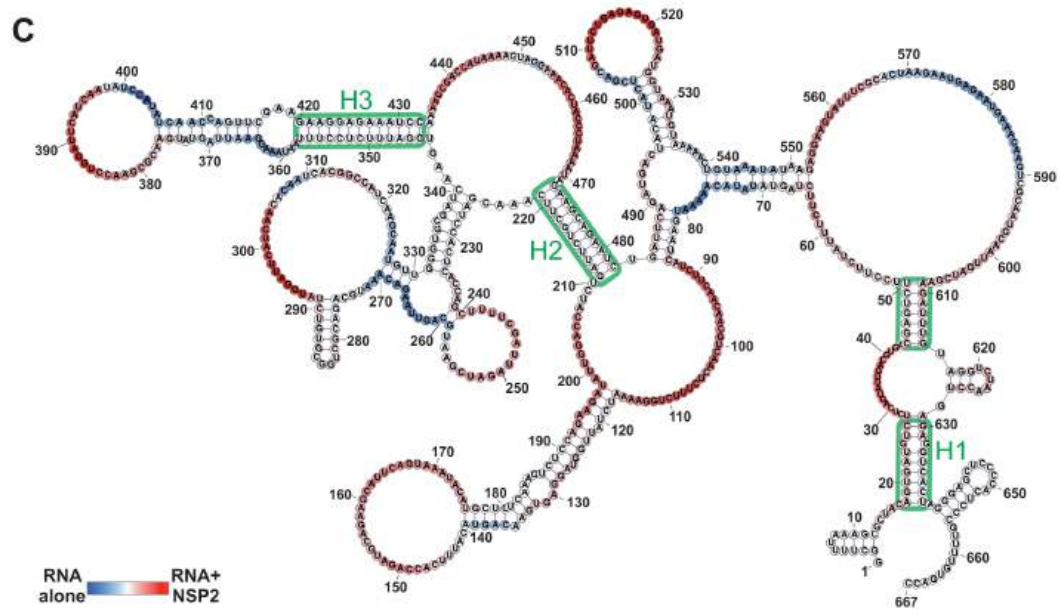


Figure 8.2.3: Model showing secondary structure on ribosomal S11 RNA and accessibility changes on binding the protein NSP2.

You can imagine that a different set of intrachain H-bonded double-stranded regions could easily form, with the most likely determined by sequence, local environment, and protein binding partners. Each RNA molecule would have a thermodynamic folding landscape similar to protein. Programs are available to determine secondary structures from RNA sequences. Each RNA molecule would have a thermodynamic folding landscape similar to protein. Also, the structures are dynamic as we saw with proteins.

Another feature that makes RNAs complicated is that many different types of RNA are made from DNA using RNA polymerases. They are loosely divided into two types of RNA. One is **coding RNA**, which contains the sequence information that will be translated into a protein sequence. The other type is called **noncoding RNA**. These RNAs regulate a myriad of cellular processes including transcription to produce the coding RNA.

8.2.1.1: Coding RNA

The DNA template from which the coding sequence of a translatable RNA is produced is called a **gene.** The coding RNA which has the exact sequence that is translated into protein is called **messenger RNA (mRNA)**. The exact sequence of RNA in mRNA that encodes a protein is derived from a longer contiguous DNA sequence in the nucleus from which sections called **intervening sequences** or **introns** have been removed. The coding sequences of DNA which are separated by introns are called **exons**. When coding DNA is transcribed, a long contiguous sequence containing both exons and introns is transcribed into one long primary transcript called **heteronuclear RNA**. The introns in the heteronuclear RNA are removed, in a splicing reaction catalyzed by a large complex called the **spliceosome** to form mRNA. The process is illustrated in Figure 8.2.4. The first RNA sequence made is the heteronuclear RNA.

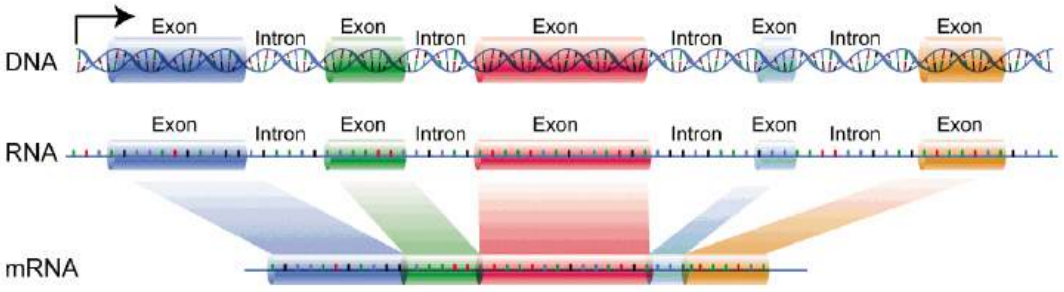


Figure 8.2.4: https://commons.wikimedia.org/wiki/F...ns_introns.gif. This work is in the **public domain** in the United States

The long single stranded mRNA molecule binds to ribosomes, nanomachines which orchestrate the translation of the mRNA sequence into a protein sequence. There are around 20,000 human genes that produce an even large number of mRNA that arise

from differential splicing of the primary transcript.

8.2.1.2: Noncoding RNA (ncRNA)

Not long ago, few thought about possible RNA transcripts from non protein-coding regions of the genome, except for two types of RNA required for the translation of mRNA. These two are ribosomal RNAs (rRNA) which are found in ribosomes and transfer RNAs, to which amino acids are esterified and transferred to a growing protein chain of the ribosome. Many more classes have been discovered and given names that are quite confusing to someone more vind with protein structures. One way to classify noncoding RNAs (ncRNAs), which implies non-protein coding RNAs is based on size.

- short noncoding RNAs (sncRNAs) are <200 nucleotides
- long noncoding RNAs (lncRNAs)are >200 nucleotides

These function to regulate gene expression at both the transcription and post-transcriptional levels. Some have catalytic functions. Some affect chromosome structure and chemical modification.

8.2.1.2.1: Long Noncoding RNAs (lncRNAs)

There many be between 16,000 to over 100,000 human lncRNAs encoded into the genome which adds much complexity to our understanding of the function of RNA transcripts. An online lncipedia is a database of searchable lncRNA sequences. There are many types of lncRNAs. The first we will consider is ribosomal RNA.

a. Ribosomal RNA (rRNA):

These RNAs fit the simple definition of lnRNAs (>200 nucleotides and are not protein- ing), but most would not think of them as lncRNAs since they have always been in their own category of a nonprotein-coding gene. rRNAs vary in length from between 1500 and 3000 nucleotides long in bacteria and about 1800 and 5000 nucleotides long in humans and are the core structure of ribosomes, the nanomachines which translate bound mRNA into a protein sequence.

Figure 8.2.5 shows an interactive iCn3D model of the structure of 23S rRNA of the large ribosomal subunit from Deinococcus radiodurans (2O44) (long load time).

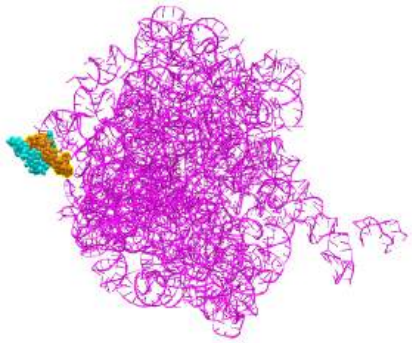


Figure 8.2.5: Structure of 23S rRNA of the large ribosomal subunit from Deinococcus radiodurans (2O44). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?Tb2BVwSTe6LviPou5

The red (highlighted yellow) spacefill is the 5' start of the rRNA. The chain has a complex tertiary structure, much like a protein sequence, and ends at the cyan spacefilling 3' end. It has 2880 nucleotides.

Let's focus on more classical examples of long noncoding RNAs (i.e not rRNA). One way to categorize them is based on the position in the genome that encodes them. The different types include long **intergenic** noncoding RNAs (lincRNA), intronic lncRNAs, antisense RNAs (as lncRNAs) and other variants. These are illustrated in Figure 8.2.6, where the lncRNA is shown in pink.

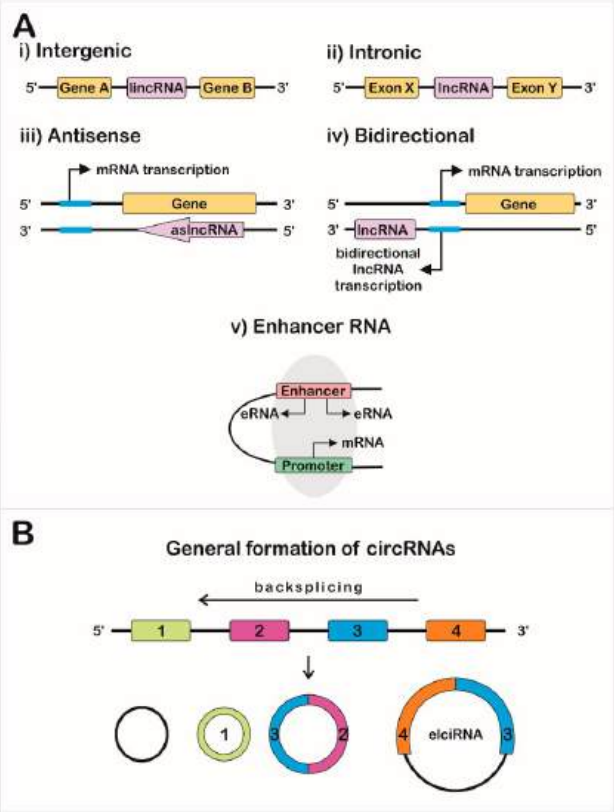


Figure \(\PageIndex{6}\): https://en.Wikipedia.org/wiki/Long_non-coding_RNA. Creative Commons Attribution-Share Alike 3.0 Unported

Another variant is exon and intron-containing circRNAs (EIciRNAs) as illustrated in panel B in Figure 8.2.13. These are presumably produced from pre-mRNA for a given mRNA, and appear to regulate gene expression through RNA-RNA interactions with U1 snRNA, which starts the assembly of the spliceosome on pre-mRNA when it binds to the 5′ pre-mRNA splice site.

Now let's consider more typical examples of long non-coding RNAs (lncRNAs), which are often bound to target proteins.

b. mamRNA (a lnc RNA)

The lncRNA named mamRNA (**M**mi1 **a**nd **M**ei2-associated RNA) binds two proteins, Mmi1 and Mei2 in *Schizosaccharomyces pombe* that control the balane between meisois and mitosis in yeast. (*Schizosaccharomyces pombe* is a "fission" yeast that divides by fission and not budding.) The MamRNA has two variants of length 550 and 700 nucleotides. Binding of mamRNA leads to the ubiquitinylation of the Mei2 in the complex. Mmi1 R is an RNA-binding protein that binds to a modified version of adenosine that has been methylated at N6 and is found internally in mRNA. Mei2 (meiosis protein 2) is necessary of meiosis. The binding of mamRNA leads to the ubiquitinylation of the Mei2 in the complex, targeting it for proteolysis. Mei2 concentrations relatively increase, shifting yeast from mitosis to meiosis. Figure 8.2.7 shows a cartoon depicting these interactions.

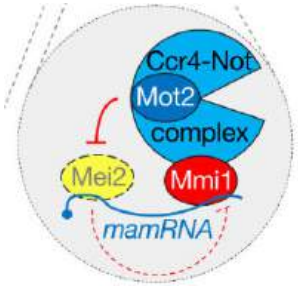


Figure 8.2.7: Binding of the lncRNA mam2 to yeast proteins Mei2 and Mmi1. Andric, V. et al. Yeast. Non-coding RNA 2021, 7, 34. https://doi.org/10.3390/ncrna7020034. Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

Figure 8.2.8 shows an interactive iCn3D model of the S. pombe Mei2 RRM3 protein domain bound to the Mei2 binding "domain" of mamRNA (6YYM) which in this structure is only 8 nucleotides long (not the full length this lncRNA which is 550 and 700 nucleotides long).

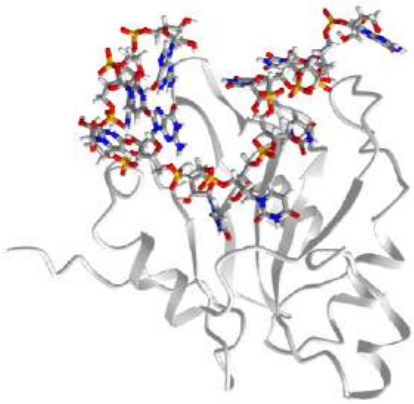


Figure 8.2.8: S. pombe Mei2 RRM3 protein domain bound to an 8 nucleotide section of the yeast lncRNA mam (6YYM). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...5xmQjwAmQL9on7

c. ToxI - a lnRNA inhibitor of the endonuclease ToxN

Those with a more chemistry-centric background might be surprised to know that viruses also "infect" bacteria. These viruses are called **bacteriophages**. It is estimated that there are over 10^{30} in nature. Some covalently incorporate into genomes where they reside and are incorporated permanently into the genome. They are a main driver of bacterial genome evolution as they shape the bacteria's immune response and adaptation.

One very interesting example is the type III toxin-antitoxin (TA) system in E. Coli. It consists of a toxin, ToxN, which is a nuclease that cleaves internally after the second A in a AAA sequence. It acts on mRNA but especially pre-mRNA sequences. It is inhibited by the binding of a lncRNA called ToxI (toxin inhibitor). The RNA sequence of the ToxI inhibitor has 36 "domain" repeats of a pseudoknot, one of which is sufficient to inhibit the ToxN. The ToxN endonuclease cleaves the ToxI lncRNA as it assembles the complex. It also cleaves its mRNA. Figure 8.2.9 shows an interactive iCn3D model of the protein toxin (ToxN):lncRNA (ToxI) which is a shortened version of 29 nucleotide section from *Pectobacterium atrosepticum* (**2xdb**).

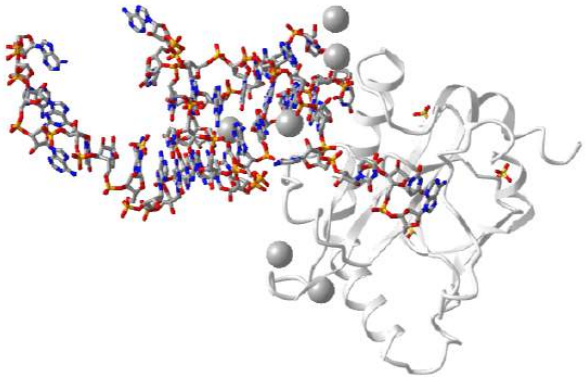


Figure 8.2.9: Protein toxin (ToxN) - 29 NT fragment of lncRNA antitoxin ToxI complex (**2xdb**). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...KjNPGghiUyAXr9

8.2.1.2.2: Short Noncoding RNA

Short noncoding RNAs (sncRNAs) are less than 200 nucleotides in length. By definition, this would include **transfer RNAs (tRNAs)** which bring to the ribosome amino acids covalently attached to their 3' end for incorporation into a growing protein chain during the translation of mRNA. As with rRNA for lncRNAs, these are really in a class of their own. Others include **small nuclear RNAs (snRNAs)** involved in splicing, **small nucleolar RNAs (snoRNAs)** involved in the modification of rRNAs, and **microRNAs (miRNAs)**, involved in the inhibition of translation and transcription, **PIWI-interacting RNAs (piRNAs)**, and endogenous **small interfering RNAs (siRNAs).** It is difficult to remember the subtle difference among these, which makes them a bit difficult to understand. We will tell their stories with a few targeted examples.

a. Transfer RNA:

Transfer RNAs act as adapter molecules between transcription and translation. They are between 76 and 90 nucleotides long and have a cloverleaf shape. An enzyme, aminoacyl-tRNA synthase, covalently attaches a select amino acid at its 3' end. Another end of the tRNA hydrogen bonds through 3 nucleotides (the anticodon) to a triplet nucleotide (the codon) on the mRNA that encodes a specific amino acid at that triplet position. Figure 8.2.10 shows an interactive iCn3D model of the structure of yeast phenylalanine tRNA (1EHZ)

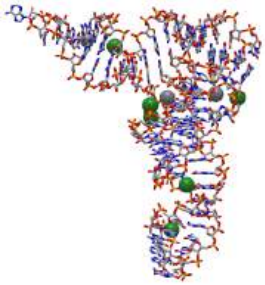


Figure 8.2.10: Yeast phenylalanine tRNA (1EHZ) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?iAWpKknDWbQm5hHL6

b. Small nuclear RNA (snRNA):

The spliceosome is a nanoparticle that catalyzes the removal of introns from pre-mRNA in eukaryotes (prokaryotes appear devoid of introns). The yeast spliceosome has a molecular weight of 1.3 million and contains 5 small ribonucleoproteins (RNPs) with many other associated proteins. Each of the 5 RNPs has a small nuclear RNA (U1, U2, U4, U5 and U,6) which is enriched in uracils. U6 is highly conserved and is directly involved in catalysis. Figure 8.2.11 shows an interactive iCn3D model of the core structure of the U6 small nuclear ribonucleoprotein complex with most of the U6 RNA bound.

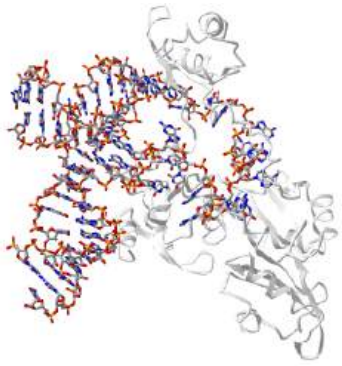


Figure 8.2.11: Core structure of the U6 small nuclear RNA - protein (ribonucleoprotein) complex (4N0T). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...P3SLVRAKhP7W57

c. MicroRNAs (miRNAs) and small inhibitory RNAs (siRNAs)

MicroRNAs (miRNAs) control the expression of thousands of genes in plants and animals. They are single-stranded but fold on themselves to form a stem-hairpin. The miRBase is a microRNA database containing almost 40,000 miRNA sequences. miRNAs are highly conserved and are found in animals, plants, and some unicellular eukaryotes. They interact with the 3′ untranslated regions of mRNAs and inhibit or prevent their translation. Several key proteins, **RNA polymerase II, Drosha and Dicer** are involved in the canonical pathway while the others appear to be independent of **Drosha**, which is a ribonuclease III double-stranded (ds) RNA endoribonuclease.

Dicer is a dsRNA) endoribonuclease which cleaves long dsRNAs and short hairpin pre-microRNAs (miRNA) into fragments of either 21-23 nucleotides (short interfering RNA) or 19-25 nucleotides (microRNAs). Each has two nucleotides that are unpaired at the 3' end. These bind to the enzyme complex RISC (RNA-induced silencing complex) whicn then targets the to mRNA complementary to the siRNA/miRNA (RISC) causing cleavage of the mRNA and hence inhibiting translation.

Small inhibitory RNAs (siRNAs) are very similar to miRNA (to the point that differentiating between them is somewhat arbitrary). They both engage in RNA interference (RNAi) of mRNA translation. Here some reported differences:

- The substrate for dicer cleavage is dsRNA (that could be added exogenously) of length 30-100+ for siRNA but the actual pre-miRNA of length 7-100 nucleotides that may contain hairpins with some mismatches (i.e. not a perfect stem and hairpin)
- The final RNA after dicer processing is double-stranded for both and 21-23 nucleotides long for siRNA and 19-25 for miRNA
- siRNAs are perfectly complementary to the target mRNAs while miRNAs, which are not necessarily perfectly complementary, bind typically to the 3' untranslated end of the mRNA
- Because of the perfect complementarity to target mRNA, siRNA interact with only one mRNA while miRNAs, given that they are not perfectly complementary to their target sequences, can bind different mRNAs
- Given their higher affinity binding, the siRNA leads to dicer endonuclease cleavage of the target mRNA while inhibition of mRNA translation by miRNAs arises from binding of the miRNA to the mRNA or, if the match between the miRNA and mRNA is high enough, endonuclease cleavage of the mRNA.

Figure 8.2.12 shows a canonical and several alternative pathways for their transcription and processing from the noncoding miRNA genes.

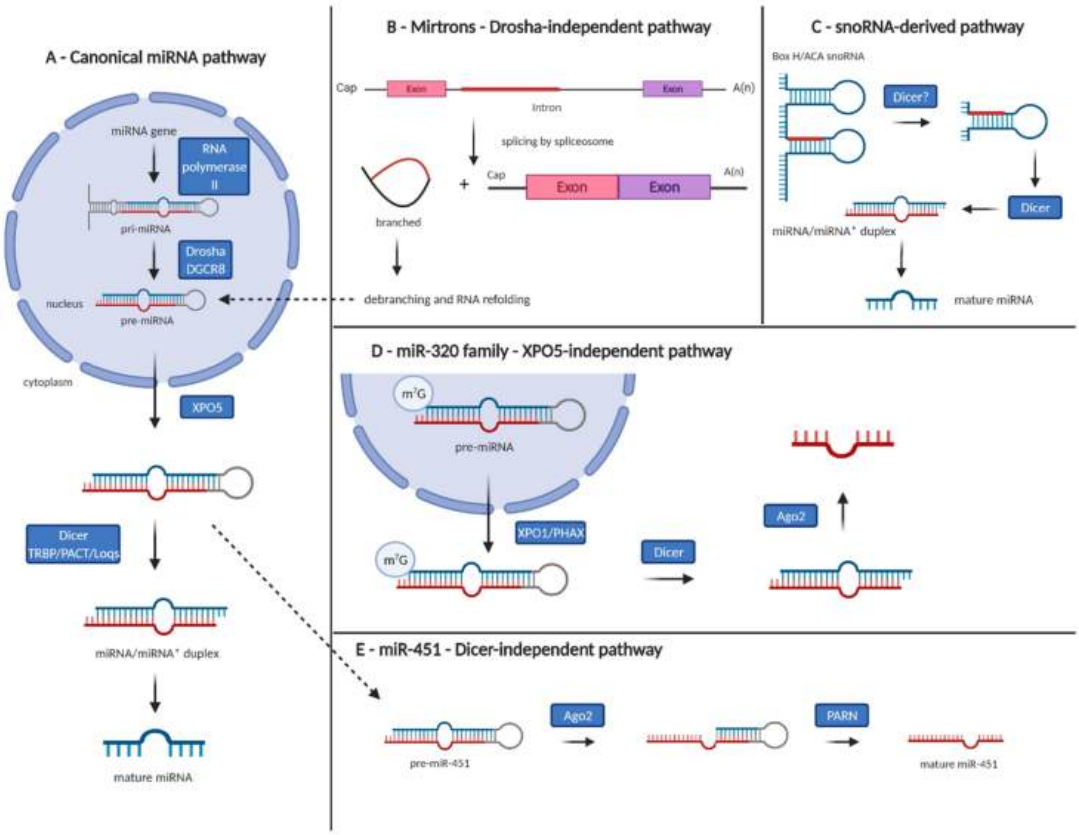


Figure 8.2.12:

Figure 1. Canonical and non-canonical pathways of microRNA biogenesis. (A) Canonical pathway—microRNA gene is transcribed by RNA polymerase II into primary microRNA (pri-miRNA), cleaved by microprocessor complex Drosha/DGCR8, and precursor microRNA (pre-miRNA) is exported from the nucleus to the cytoplasm by Exportin 5 (XPO5) and further processed by Dicer and its partners into 18–25 nucleotide long microRNA duplex with 2-nucleotide 30 overhangs. The guide strand is subsequently bound by the Argonaute proteins 1-4 (AGO1-4) and retained in the microRNA-induced silencing complex to target mRNAs for post-transcriptional silencing. (B) Mirtrons—generated through mRNA splicing independently of Drosha-mediated processing step. (C) Small nucleolar RNA-derived microRNAs—Drosha-independent pathway. (D) Exportin 5-independent transport of pre-miRNAs from the nucleus to the cytoplasm has been described in the case of miR-320 family. (E) Dicer-independent processing of miR-451—pre-miR-451 is directly loaded into AGO2, cleaved and trimmed by poly(A)-specific ribonuclease PARN to produce mature miR-451. Gregorova et al. Cancers 2021, 13, 1333. https://doi.org/10.3390/cancers13061333. Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/)

siRNA that are perfect matches to specific mRNA can be easily designed and purchased for translation inhibition and through that gene silencing studies. Both miRNAs and siRNA are potentially therapeutic as it is much simpler to design a drug that targets a mRNA sequence (a 1D sequence target) than a protein active site (a 3D target). Additionally, they can be used to inhibit protein synthesis of target proteins that don't have a "druggable" active site.

The protein Argonaute is involved in miRNA and siRNA silencing of genes through their mRNAs in the RISC (RNA-induced silencing complex). RISC contains the protein argonaute 2 (AGO2) bound to a "**guide**" RNA which is either microRNA (miRNA) or short interfering RNA (siRNA). It is the miRNA or siRNA that directly interacts with the "target" - the mRNA.

Example miRNA:

Figure 8.2.13 shows an interactive iCn3D model of human Argonaute2 Bound to a Guide (miRNA) and Target RNA (4W5O). The two RNA sequences are 5' UUCACAUUGCCCAAGUCUUU 3' and 5' CAAUGUGAAA 3'.

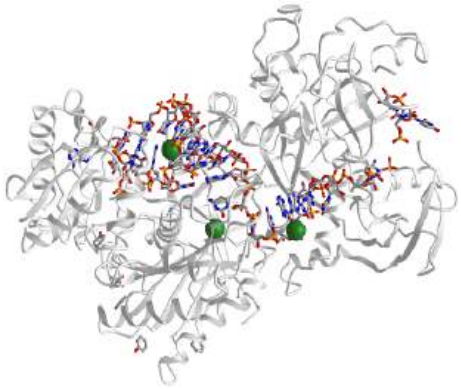


Figure 8.2.13: human Argonaute2 Bound to a Guide and Target RNA (miRNA) (4W5O). (Copyright; author via source). Click the image for a popup or use this external link: **https://structure.ncbi.nlm.nih.gov/icn3d/share.html?9fvyQmPkk4Kctei59**

Example: siRNA (Small interfering RNA)

Virus genomes are ultimately decoded into new viruses by the host replication, transcription and translation machinery. Host cells have evolved ways to silence mRNAs from viruses. Unfortunately, viruses, in response, evolve ways to suppress host RNA silencing. Many viral proteins are used to suppress silencing by the host. One is the viral p19 protein, which preferentially binds to host short interfering RNAs (siRNAs) than to microRNAs (miRNAs). A single mutation in the viral p19 proteins changes selectivity which allows it to bind to a specific human miRNA called miR-122. This shows the subtle complexities of protein:RNA interactions. Figure 8.2.14 shows an interactive iCn3D model of the viral suppressor of RNA silencing protein and a 21 residue small interfering RNAs (6BJV)

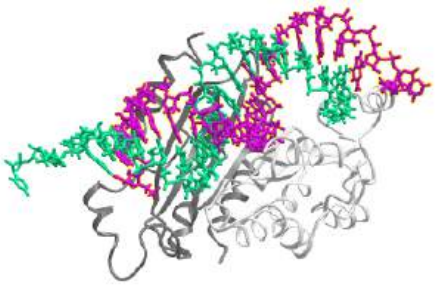


Figure 8.2.14: Viral suppressor of RNA silencing protein p19 mutants and a double-stranded 21 nucleotide small interfering RNAs (6BJV). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...VXzSgezTEPcPt9

Example: piRNA (a specific miRNA)

Piwi proteins are RNA-binding proteins in plants and animals and are similar in structure to argonaute. They bind a guide RNA called piwi-interacting RNAs (piRNAs) and lead to the silencing of sequences in the genome called transposable elements that can move around the genome. piWi has endonuclease activity and can cleave mRNA. The piRNAs are just one type of miRNA. Figure 8.2.15 shows an interactive iCn3D model of Ephydatia fluviatilis (a sponge) PiwiA with a guide (piRNA) and-target RNA(7KX9)

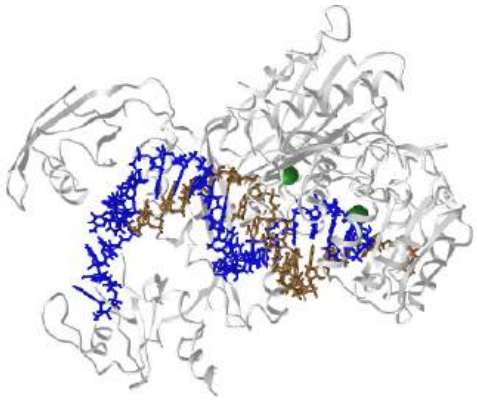


Figure 8.2.15: Ephydatia fluviatilis PiwiA-piRNA-target complex (7KX9). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?j9ikApUgsVvqsiM5A

lncRNAs and miRNAs in the brain

After the decoding of the human genome, many have been struggling to understand how the complexity of the human brain (large size, greater connectivity among neurons) arises given that we appear to have only around 20,000 proteins genes encoded by the genome (not counting small proteins of less than 100 amino acids). Long noncoding RNAs (lncRNAs) and miRNAs appear to be significant pieces of this puzzle. Their ability to regulate transcription during development may hold the key. Additional roles on these RNAs outside of transcriptional regulation are being discovered. Some are transported away from the nucleus to serve other functions in axons, dendrites, etc. For example, the lncRNA Gm38257 binds to proteins that structure the synapse.
something rather unexpected: Instead of simply regulating gene expression, it binds to proteins

The repertoire of miRNA appears to be significantly increased in "intelligent" organisms such as humans and octopi. For example, a large increase (179) in miRNAs occurs in proceeding in the evolutionary scale from mice (which have about 24,000 protein-encoding genes) to humans (around 20,000). miRNAs and lncRNAs may be involved (causative or correlative?) with brain disease. An example is the miR-124 is significantly elevated (3.5X) times higher in hippocampal cells from mouse models of Alzheimer's compared to normal mice. Altered expression of the lncRNA named Gomafu, RNCR2 or MIAT) appears to affect certain psychiatric diseases.

f. small nucleolar RNA

The nucleolus is a small structure in the nucleus that helps assemble the ribosomal RNAs that are synthesized in the nucleus. They then are transported through the nuclear membrane into the cytoplasm where they combine with proteins translated from mRNA in the cytoplasm to form complete ribosomes. As we will describe below, rRNA is chemically modified by enzymes (much like the post-translational modification of proteins). One such modification is 2'-O-methylation in archaea and eukaryotes. A class of small nucleolar RNAs (snoRNAs) which vary from 10-21 base pairs are called C/D RNAs, and they "guide" the modification. Hence they are also called "guide" RNAs. These snoRNAs bind to 3-4 proteins into ribonucleoproteins. Figure 8.2.16 shows an interactive iCn3D model of the box C/D ribonucleoprotein 40 nt snoRNA "guide" and a 10 nucleotide RNA target substrates. It appears that the maximal duplex RNA formed (from the guide and target) is 10 base pairs long.

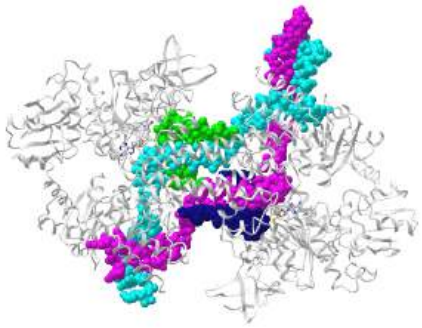


Figure 8.2.16: Box C/D ribonucleoprotein with a small nucleolar RNA (snoRNA) 12 nucleotide "guide" and 13 nucleotide target RNA for 2'-O-methylation (5GIO). (Copyright; author via source). Click the image for a popup or use this external link: **https://structure.ncbi.nlm.nih.gov/icn3d/share.html?q45tA8LEko87mYTw6**

8.2.2: References

Börner, R., Kowerko, D., Miserachs, H.G., Shaffer, M., and Sigel, R.K.O. (2016) Metal ion induced heterogeneity in RNA folding studied by smFRET. *Coordination Chemistry Reviews* 327 DOI: 10.1016/j.ccr.2016.06.002 Available at: https://www.researchgate.net/publication/303846502_Metal_ion_induced_heterogeneity_in_RNA_folding_studied_by_smFRET

Hardison, R. (2019) B-Form, A-Form, and Z-Form of DNA. Chapter in: R. Hardison's *Working with Molecular Genetics*. Published by LibreTexts. Available at: https://bio.libretexts.org/Bookshelves/Genetics/Book%3A_Working_with_Molecular_Genetics_(Hardison)/Unit_I%3A_Genes%2C_Nucleic_Acids%2C_Genomes_and_Chromosomes/2%3A_Structures_of_Nucleic_Acids/2.5%3A_B-Form%2C_A-Form%2C_and_Z-Form_of_DNA

Lenglet, G., David-Cordonnier, M-H., (2010) DNA-destabilizing agents as an alternative approach for targeting DNA: Mechanisms of action and cellular consequences. *Journal of Nucleic Acids* 2010, Article ID: 290935, DOI: 10.4061/2010/290935 Available at: https://www.hindawi.com/journals/jna/2010/290935/

Mechanobiology Institute (2018) What are chromosomes and chromosome territories? *Produced by the National University of Singapore.* Available at: https://www.mechanobio.info/genome-regulation/what-are-chromosomes-and-chromosome-territories/

National Human Genome Research Institute (2019) The Human Genome Project. *National Institutes of Health*. Available at: https://www.genome.gov/human-genome-project

Wikipedia contributors. (2019, July 8). DNA. In *Wikipedia, The Free Encyclopedia*. Retrieved 02:41, July 22, 2019, from https://en.Wikipedia.org/w/index.php?title=DNA&oldid=905364161

Wikipedia contributors. (2019, July 22). Chromosome. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:18, July 23, 2019, from en.Wikipedia.org/w/index.php?title=Chromosome&oldid=907355235

Wikilectures. Prokaryotic Chromosomes (2017) In MediaWiki, Available at: https://www.wikilectures.eu/w/Prokaryotic_Chromosomes

Wikipedia contributors. (2019, May 15). DNA supercoil. In *Wikipedia, The Free Encyclopedia*. Retrieved 19:40, July 25, 2019, from en.Wikipedia.org/w/index.php?title=DNA_supercoil&oldid=897160342

Wikipedia contributors. (2019, July 23). Histone. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:19, July 26, 2019, from en.Wikipedia.org/w/index.php?title=Histone&oldid=907472227

Wikipedia contributors. (2019, July 17). Nucleosome. In *Wikipedia, The Free Encyclopedia*. Retrieved 17:17, July 26, 2019, from en.Wikipedia.org/w/index.php?title=Nucleosome&oldid=906654745

Wikipedia contributors. (2019, July 26). Human genome. In *Wikipedia, The Free Encyclopedia*. Retrieved 06:12, July 27, 2019, from en.Wikipedia.org/w/index.php?title=Human_genome&oldid=908031878

Wikipedia contributors. (2019, July 19). Gene structure. In *Wikipedia, The Free Encyclopedia*. Retrieved 06:16, July 27, 2019, from en.Wikipedia.org/w/index.php?title=Gene_structure&oldid=906938498

8.3: Nucleic Acids - Comparison of DNA and RNA

Now that we have an understanding of the structures of DNA and the structures and various functions of RNA, we can now more fully explore how their chemical similarities and difference contribute to different functions.

8.3.1: Chemical modifications of DNA and RNA

Post-translation modifications of proteins alter their structural/functional properties. Likewise, intentional chemical modifications of nucleic acid bases alter both their structures and potentially their transcriptional and translational status. Figure 8.3.1 shows common modifications of bases in DNA.

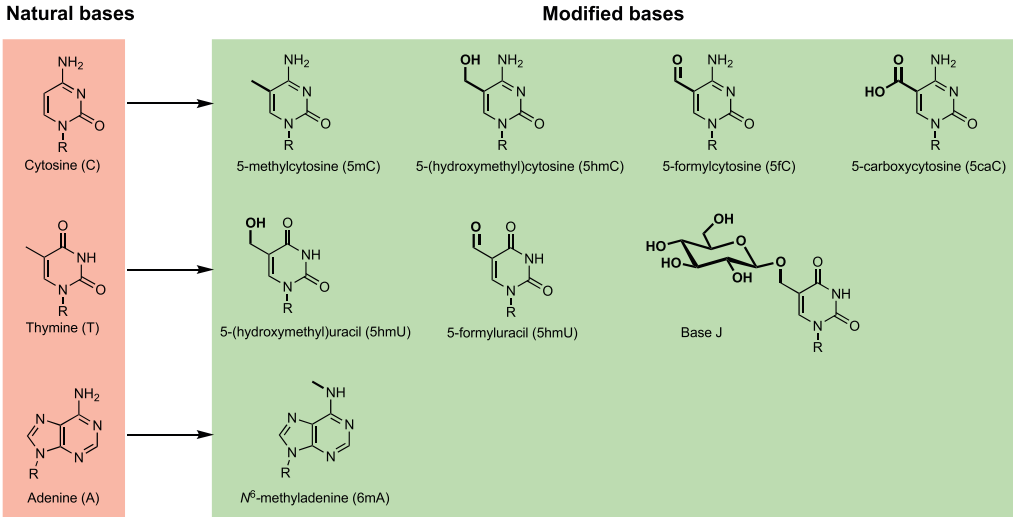


Figure 8.3.1: Common modifications of bases in DNA. Matthew K.Bilyard et al. Current Opinion in Chemical Biology. Volume 57, August 2020, Pages 1-7. https://doi.org/10.1016/j.cbpa.2020.01.014. Under a Creative Commons license

Likewise, RNA is chemically modified. Figure 8.3.2 shows common modifications of bases in RNA. Methylation and subsequent hydroxylation to hydroxymethyl are common to both DNA and RNA. Methylation of DNA often represses the transcription of the DNA into RNA. Hence it has huge potential to alter gene transcription. Such changes to the DNA are called **epigenetic** modifications. These changes can be passed down to future generations as well and affect the phenotype of a cell. Histone proteins involved in DNA packing into nucleosomes can also be methylated and acetylated, altering the interaction of the DNA with the nucleosome core and further packing, again affecting transcription.

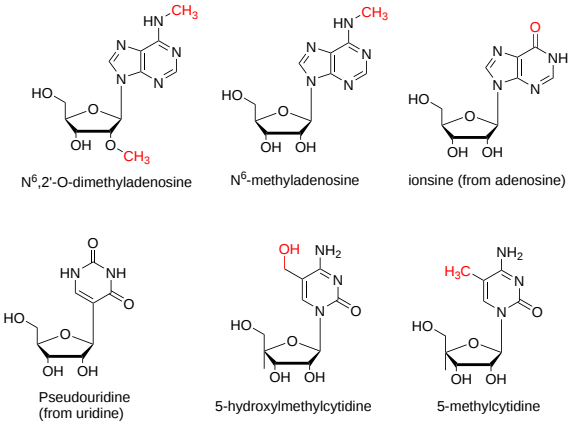


Figure 8.3.2: Common modifications of bases in RNA

The chemical modification to RNA also can change the reading out of the genome. The **epitranscriptome** refers to the collective chemical modifications to RNA, and its understanding is part of new field called **epitranscriptomics**.

8.3.2: Mutations

Mutation can arise from the chemical modification of bases. Uracil in RNA is a demethylated form of thymine in DNA. In RNA, AU base pairs replace AT base pairs. Why the need for uracil in RNA? The question could be rephrased as to why the need for thymine, with its extra methyl group, in DNA. It's useful to think about the consequence of replacing a single H in a molecule with a -CH_3. Take HOH, water, as an example. Our bodies are over 60% water. We drink liters of water of concentration 55 M each day. Yet if we drink 0.07 L of methanol, CH_3OH, half of us would die! Let's probe some consequences of the U (no -CH_3) and T (with -CH_3) changes in DNA. It can get confusing but just remember that the normal base pairs in DNA are AT, but AU base pairs also form (they norm in RNA). The -CH_3 substituent on thymine does not affect its base pairing.

a. Spontaneous deamination of cytosine in DNA

Why are we now discussing cytosine in DNA? One reason is that the most common mutation in DNA is a C to T replacement. One way that happens is through the spontaneous hydrolytic deamination of cytosine in DNA to uracil, which we have presumed to be found only in RNA. The mechanism for this deamination and subsequence conversion of a GC to an AT base pair is shown in Figure 8.3.3. The inset box shows a simplified mechanism for spontaneous deamination.

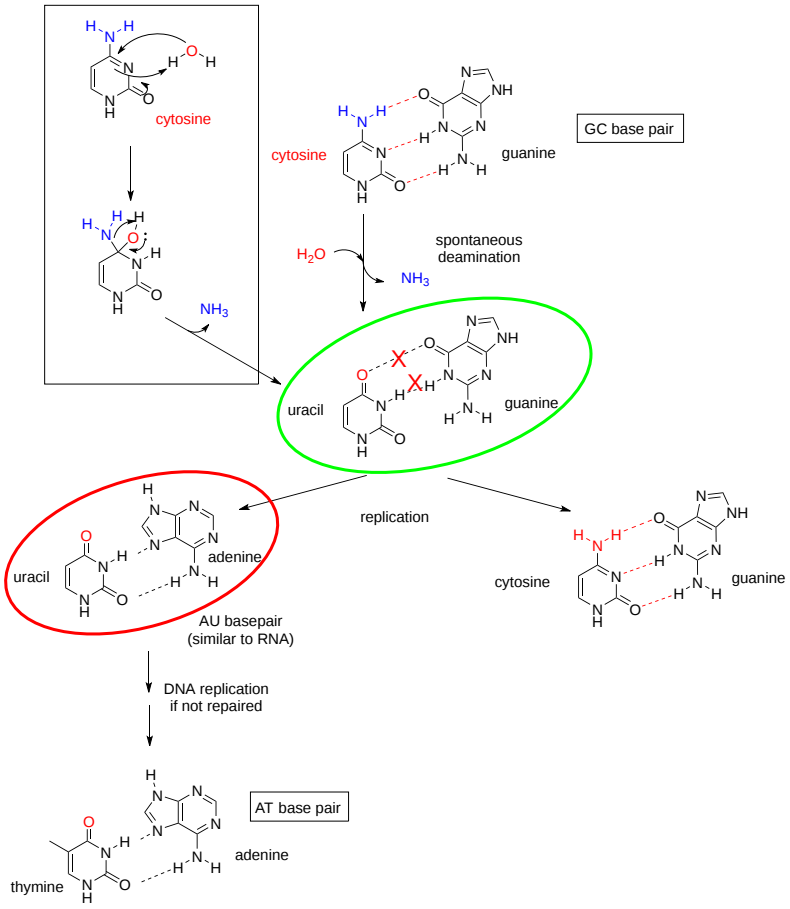


Figure 8.3.3: GC to AT base pair mutation on spontaneous hydrolytic deamination of cytosine in DNA.

Hence a possible consequence of the deamination reaction is a GC to AT base pair mutation if the uracil in DNA is not removed before DNA replication. Fortunately, the enzyme uracil-DNA glycosylases can remove any uracils found in DNA, leaving an abasic site, which can be fixed with DNA repair enzymes.

We can now ask the question, why T and not U in DNA? Pretend you are a DNA repair enzyme and you see a UA base pair in DNA. How can you tell if the UA base pair is correct and intended to be there or if it should be a CG base pair that underwent deamination? The most common uracil-DNA glycosylases remove the uracil whether it is across from guanine, the correct base but which can not hydrogen bond with uracil (in the **green oval** in Figure 8.3.3), or if is across from adenine, the wrong base (in **red oval**), which is present after a round of replication. Evolution has addressed this problem by adding a methyl group to uracil to

form thymine and using that base, which forms a base pair with adenine. Now no decision on which base across from a uracil (guanine if the uracil arose from deamination) or across from a "uracil-like" thymine (adenine) is correct.

b. Other mutations

Since we are considering chemical modifications to DNA and mutations, it is appropriate to give a more expanded background on them. In addition to mutations caused by spontaneous hydrolytic deamination of cytosine, mutations can also arise through the addition of a wrong base during DNA replication, by chemical damages caused by radiation or chemical modifying agents. How many mistakes in replication are made? If you received a 99% on an examination, you would be ecstatic. That's not good enough for DNA replication. In Cell Biology by the Numbers, they calculate it this way. Assume the replication /repair is so good that it takes 10^8 replications to make a mistake (an error rate of 10^{-8}/BP). Assume also there are 3 x 10^9 base pairs in the human genome. This leads to a mutation rate 10-100 mutations/genome/generation or about 0.1-1 mutations/genome/replication. Not bad!

Figure 8.3.4 shows how common point mutations might arise just randomly.

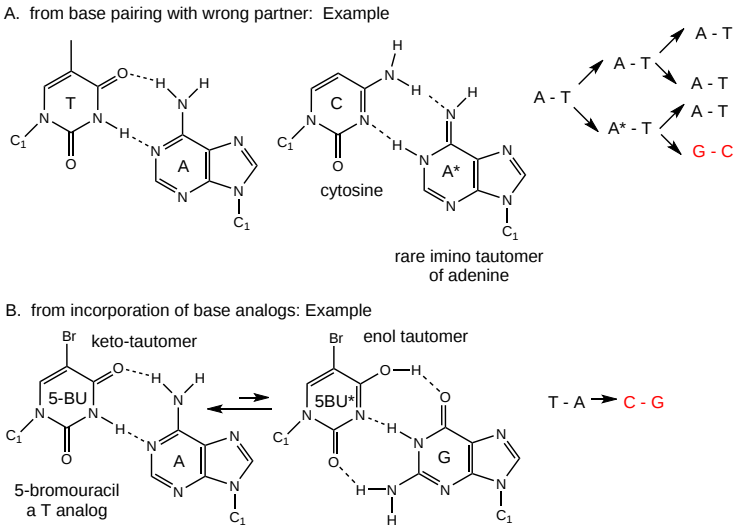


Figure 8.3.4: How common point mutations might arise randomly

Chemical agents also can cause point mutations. Figure 8.3.5 shows point mutations arising from oxidative deaminations (not hydrolytic) by nitrous acid/nitrosamines and from alkylating agents.

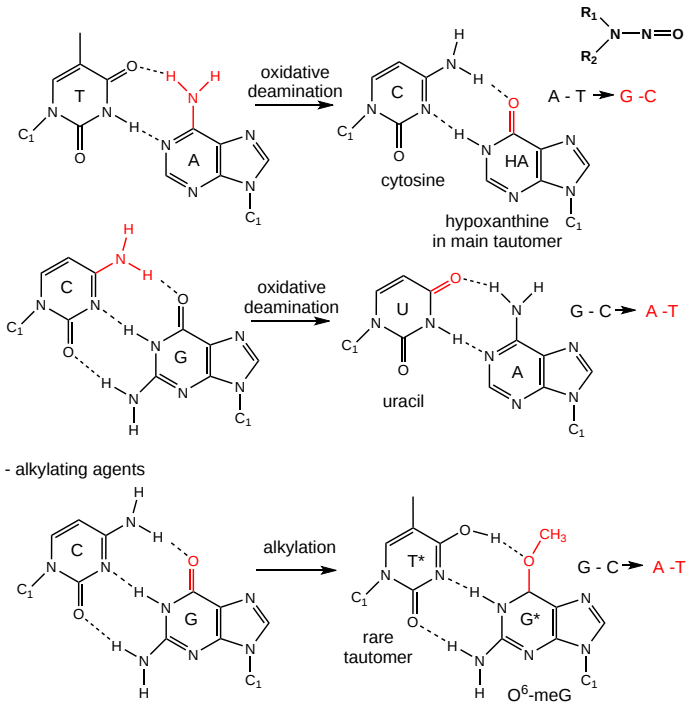


Figure 8.3.5: Nitrous acid/nitrosamines and alkylating agent point mutations

Figure 8.3.6 shows a variety of alkylating agents with mutagenic potential.

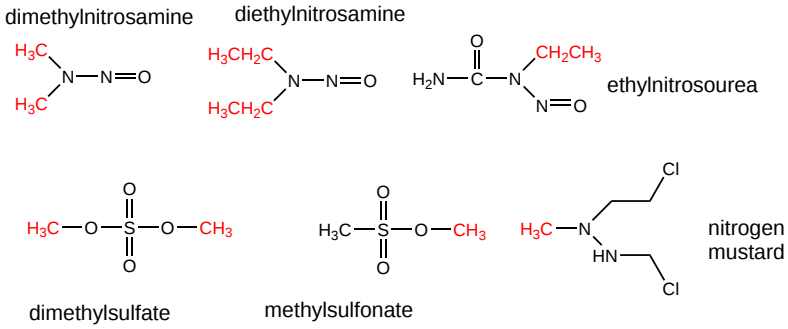


Figure 8.3.6: Alkylating agents with mutagenic potential.

Finally, large-scale changes in chromosome structure can also occur as shown in Figure 8.3.7, usually with profound consequences.

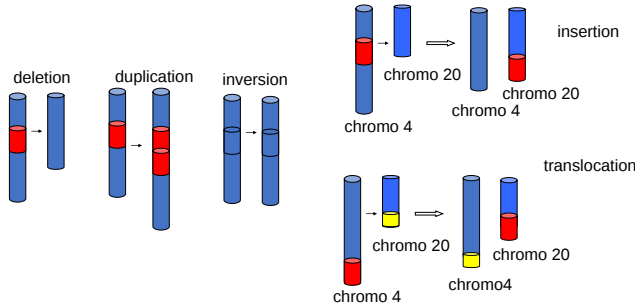


Figure 8.3.7: Large scale structural rearrangements in DNA

8.3.3: Why DNA and RNA - A chemical perspective

Asking a "why" question (like above) in the sciences is really not appropriate as such teleological questions are more philosophical or religious. Yet we will in this section, in part, to be in the company of Alexander Rich, who wrote a very cool article entitled "Why RNA and DNA have different Structures".

Given that RNA expresses catalytic activities and can carry genetic information (some viruses have ds and ss RNA as their genome), it has been suggested that early life might have been based on RNA. DNA would evolve later as a more secure carrier of genetic information. An inspection of the chemical properties of DNA, RNA, and proteins shows them to have attributes needed for their expressed function. Let's examine each for structural features that might be important for function.

a. Why does DNA lack a 2' OH group (found in RNA), which has been replaced with hydrogen? This required the evolutionary creation of a new enzyme, ribonucleotide reductase, to catalyze the replacement of the OH in a ribonucleotide monomer to form the deoxyribonucleotide form. One possible explanation is offered in the figure below. DNA, the main carrier of genetic information, must be an extremely stable molecule. An OH present on C'2 could act as a nucleophile and attack the proximal P in the phosphodiester bond, leading to a nucleophilic substitution reaction and potential cleavage of the link. RNA, an intermediary molecule, whose concentration (at least as mRNA) should rise and fall based on the need for a potential transcript, should be more labile to such hydrolysis. Figure 8.3.8 shows a possible reaction diagram for the internal cleavage of RNA. (The reaction would probably proceed with no actual intermediate, but just a transition state.

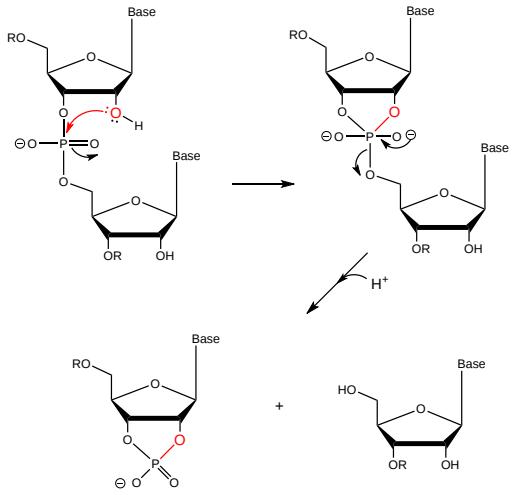


Figure 8.3.8: Internal cleave of RNA using the C'2-OH as an intramolecular nucleophile

b. Why do both DNA and RNA contain a phosphodiester link between adjacent monomers instead of more "traditional" links such as carboxylic acid esters, amides, or anhydrides? One possible explanation is given below. Nucleophilic attack on the sp^3 hybridized P in a phosphodiester is much more difficult than for a more open sp^2 hybridized carboxylic acid derivative. In addition, the negative charge on the O in the phosphodiester link would decrease the likelihood of a nucleophilic attack. The negative charges on both strands in ds-DNA probably help keep the strands separated allowing the traditional base pairing and double-stranded helical structure observed. The cleavage of the phosphodiester link in DNA and a hypothetic ester link is shown in Figure 8.3.9. Again, the reaction of the phosphodiester shows a pentavalent intermediate, but most like the reaction proceeds directly from the transition state.

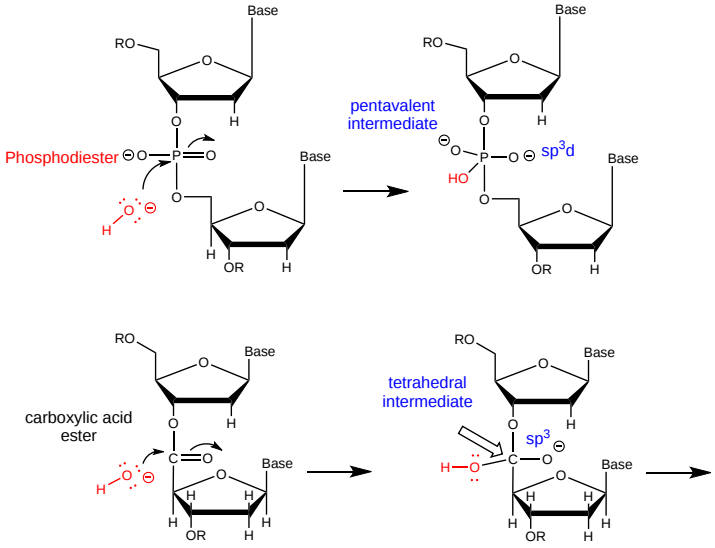


Figure 8.3.9: cleavage of the phosphodiester link in DNA and a hypothetic ester link

c. Why is DNA found as a repetitive double-stranded helix but RNA is usually found as a single-stranded molecule that can form complicated tertiary structures with some ds-RNA motifs?

Another reason for the absence of the 2' OH in DNA is that it allows the deoxyribose ring in DNA to pucker in just the right way to sterically allow extended ds-DNA helices (B type). The pucker in deoxyribose and ribose can be visualized by visualizing a single plane in the sugar ring defined by the ring atoms C1', O, and C4'. If a ring atom is pointing in the same direction as the C4'-C5' bond, the ring atom is defined as **endo**. If it is pointing in the opposite direction, it is defined as **exo**. In the most common form of double-stranded DNA, B-DNA, which is the iconic extended double helix you know so well, C2' is in the endo form. It can also adopt the C3' endo form, leading to the formation of another less common helix, a more open ds-A helix. In contrast, steric interference prevents ribose in RNA from adopting the 2'endo conformation, and allows only the 3'endo form, precluding the occurrences of extended ds-B-RNA helices but allowing more open, A-type helices.

Figure 8.3.10 shows another comparison between the A-RNA and B-DNA double helices and the C'3 and C'2 endo forms of the ribose

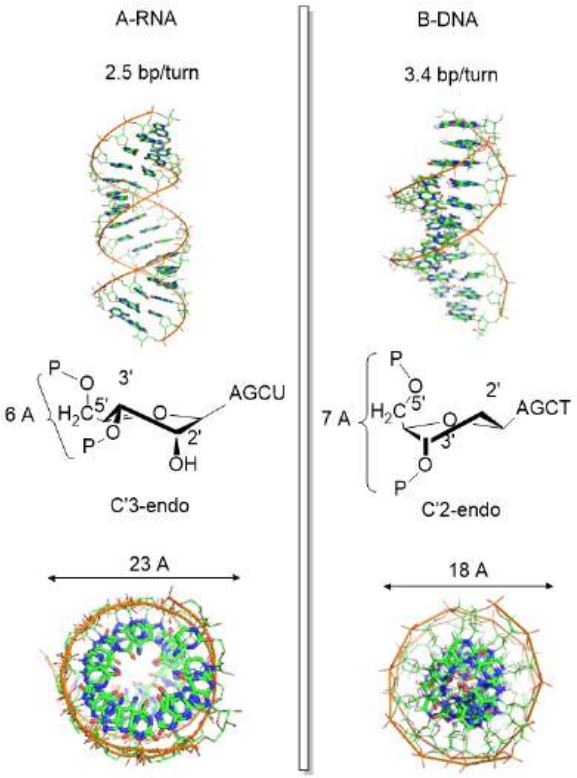
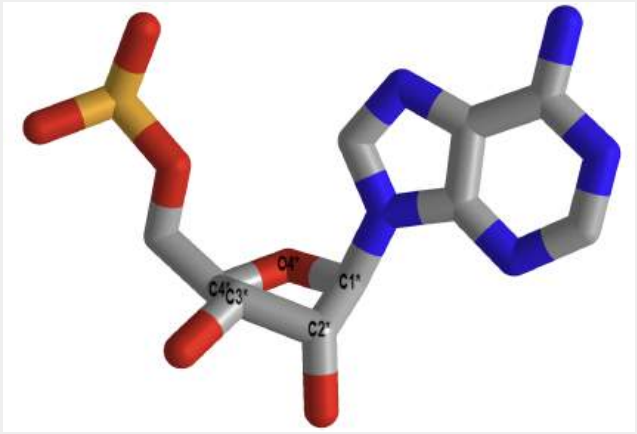
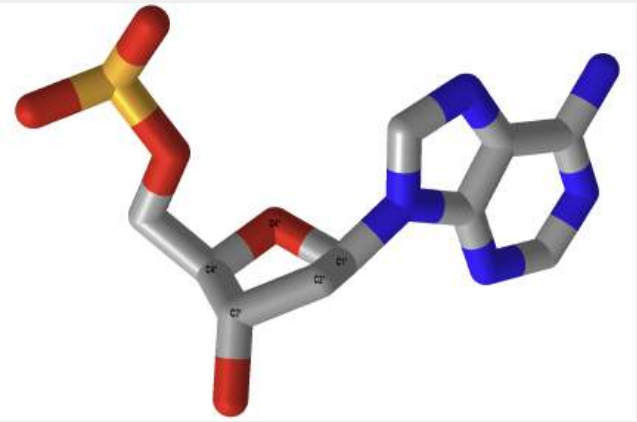


Figure 8.3.10: after Zhou et al Nature Structural and Molecular Biology. doi:10.1038/nsmb.3270

Figure 8.3.5 shows interactive iCn3D models of the pentoses in a strand of A-RNA (413D), double-stranded, left, and B-DNA (1BNA), double-stranded, right.

C'3-endo ribose, A-RNA (413D, double stranded)	C'2 endo ribose, B-DNA (1BNA, double stranded)
	
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...KPueqrBADczh26	Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...BEn5nqsCQG2JH6

d. What about the molecular dynamics of A-RNA and B-DNA?

The information above suggests that the sugar ring of DNA is conformationally more flexible than the ribose ring of RNA. This can be inferred from the observation that dsDNA can adopt B and A forms, which requires a switch from the 2' endo in the B form to the 3'endo form in the A form. The smaller H on the 2'C would offer less steric interference with such flexibility. The rigidity in ribose is associated with a smaller 5'O to 3'O distance in RNA leading to a compression of the nucleotides into a helix with a smaller number of base pairs/turn.

The increased flexibility in DNA allows rotation around the C1'-N glycosidic bond connecting the deoxyribose and base in DNA, allowing different orientations of AT and GC base pairs with each other. The normal "anti" orientation allows "Watson-Crick"

(WC) base pairing between AT and GC base pairs while the altered rotation allows "Hoogsteen" (Hoog) base pairs. The different orientations for an AT base pair are shown in Figure 8.3.11.

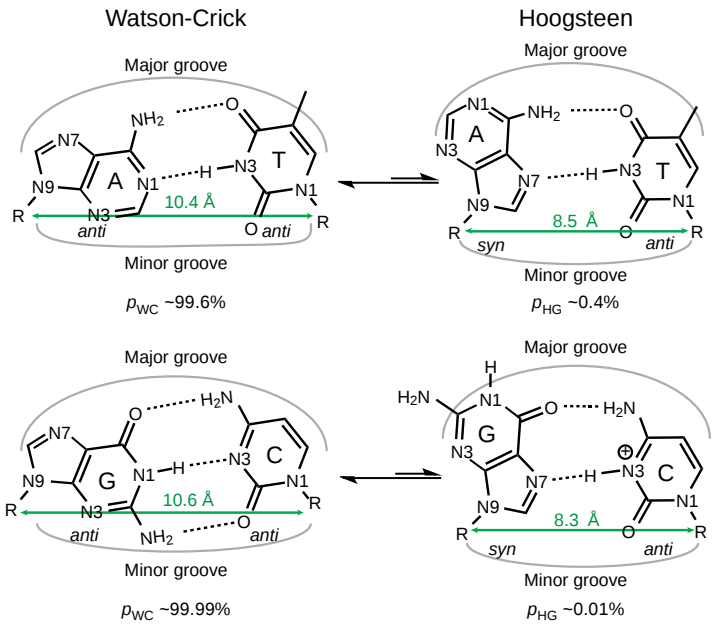


Figure 8.3.11: Xu, Y., McSally, J., Andricioaei, I. *et al.* Modulation of Hoogsteen dynamics on DNA recognition. *Nat Commun* **9,** 1473 (2018). https://doi.org/10.1038/s41467-018-03516-1Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

The Watson-Crick (WC) and Hoogsteen (HG) base pairs in B-DNA are in a dynamic equilibrium with the equilibrium greatly favoring the WC form as indicated by the arrows in the figure above. In a DNA:protein complex, the WC ↔ HG equilibrium can favor the WG form for AT and GC^+ forms (in the latter, the C is protonated) when those base pairs are also involved in protein recognition. They can also occur more frequently in damaged DNA. In contrast, molecular dynamic studies show that the HG base pairs A-U and GC^+ are strongly disfavored in ds A-RNA.

One type of DNA damage is methylation on N1-adenosine and N1-guanosine. This modification prevents normal Watson-Crick base pairing but for DNA, these modified bases can still engage in Hoogsteen base pairing, preserving the overall structure of dsDNA and its ability to stably carry genetic information. This same methylation occurs normally in post-transcriptional modified RNA. Hence, N1 adenosine and N1 guanosine methylation prevent any type of base pairing in the modified RNA. These properties make DNA a better carrier of molecular information and offer another way to regulate the structural and functional properties of RNA.

Hoogsteen base pairs can be found in distorted dsDNA structures (caused by protein:DNA interactions) but also in normal B-DNA.

Figure 8.3.12 shows a Hoogsteen base pair between dA7 and dT37 in the MAT α 2 homeodomain:DNA complex (pdb 1K61). Note that the dA base in the Hoogsteen base pair is rotated syn (with respect to the deoxyribose ring) instead of the usual anti, allowing the Hoogsteen base pair.

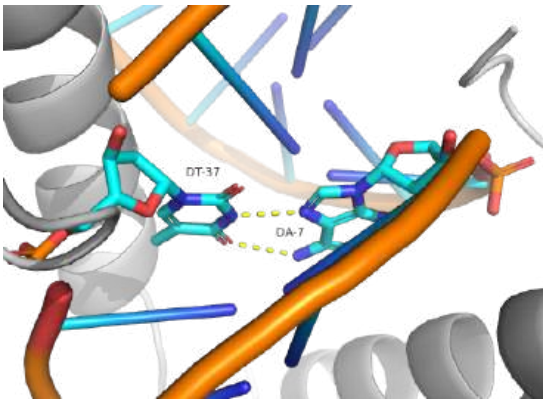


Figure 8.3.12: Hoogsteen base pair between dA7 and dT37 in the MAT α 2 homeodomain:DNA complex (pdb 1K61)

8.3.4: A Structural Comparison

Now let's review the kinds of structures adopted by the 3 major macromolecules, DNA, RNA, and proteins. DNA predominately adopts the classic ds-BDNA structure, although this structure is wound around nucleosomes and "supercoiled" in cells since it must be packed into the nucleus. This extended helical form arises in part from the significant electrostatic repulsions of two strands of this polyanion (even in the presence of counter-ions). Given its high charge density, it is not surprising that it forms complexes with positive proteins and does not adopt complex tertiary structures. RNA, on the other hand, can not form long B-type double-stranded helices (due to steric constraints of the 2'OH and the resulting 3'endo ribose pucker). Rather it can adopt complex tertiary conformations (albeit with significant counter-ion binding to stabilize the structure) and in doing so can form regions of secondary structure (ds-A RNA) in the form of stem/hairpin forms. Proteins, with their combination of polar charged, polar uncharged, and nonpolar side chains have little electrostatic hindrance in the adoption of secondary and tertiary structures. RNA and proteins can both adopt tertiary structures with potential binding and catalytic sites, making them ideal catalysts for chemical reactions. RNA, given its 4 nucleotide alphabet, can also carry genetic information, making it an ideal candidate for the first evolved macromolecules enabling the development of life. Proteins with a great abundance of organic functionalities would eventually supplant RNA as a better choice for life's catalyst. DNA, with its greater stability, would supplant RNA as the choice for the main carrier of genetic information (Figure 8.3.13):

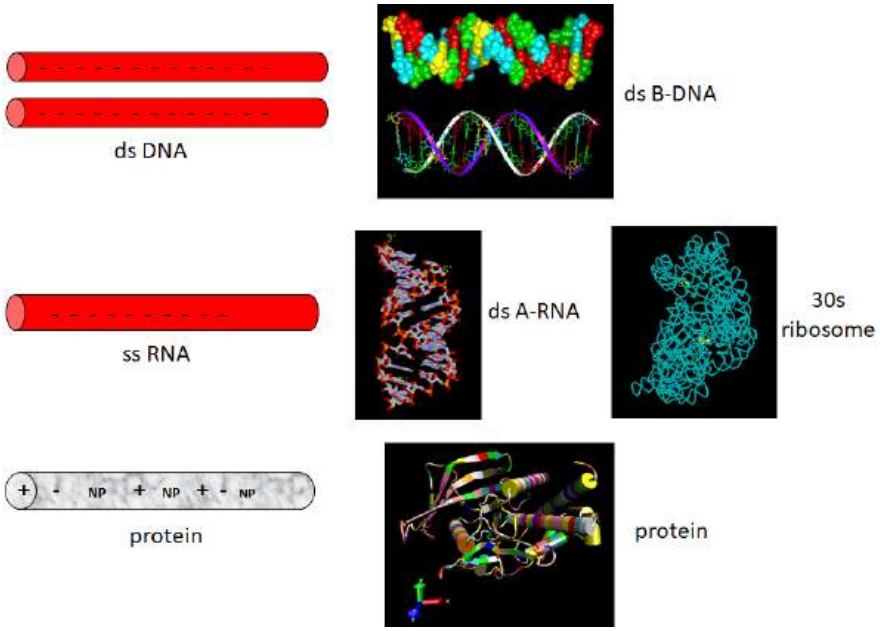


Figure 8.3.13: Summary comparison of DNA, RNA, and Protein structures

A final note on the simplicity of the dsDNA structure. A mutation causing a single base pair change in DNA does **not** change the iconic ds-stranded DNA structure. If it did, DNA would not be a reliable molecule to store and read out the genetic blueprint. In contrast, a single mutation in the DNA leading to a single amino acid substitution may lead to a protein with altered structure and

function. On one hand that could be deleterious or even fatal to the organism. On the other hand, the new protein structure might have new functionalities that allow adaptation to new environments or allow new types of reactions. Evolution would favor the latter.

8.3.5: References

Börner, R., Kowerko, D., Miserachs, H.G., Shaffer, M., and Sigel, R.K.O. (2016) Metal ion induced heterogeneity in RNA folding studied by smFRET. *Coordination Chemistry Reviews* 327 DOI: 10.1016/j.ccr.2016.06.002 Available at: https://www.researchgate.net/publication/303846502_Metal_ion_induced_heterogeneity_in_RNA_folding_studied_by_smFRET

Hardison, R. (2019) B-Form, A-Form, and Z-Form of DNA. Chapter in: R. Hardison's *Working with Molecular Genetics*. Published by LibreTexts. Available at: https://bio.libretexts.org/Bookshelves/Genetics/Book%3A_Working_with_Molecular_Genetics_(Hardison)/Unit_I%3A_Genes%2C_Nucleic_Acids%2C_Genomes_and_Chromosomes/2%3A_Structures_of_Nucleic_Acids/2.5%3A_B-Form%2C_A-Form%2C_and_Z-Form_of_DNA

Lenglet, G., David-Cordonnier, M-H., (2010) DNA-destabilizing agents as an alternative approach for targeting DNA: Mechanisms of action and cellular consequences. *Journal of Nucleic Acids* 2010, Article ID: 290935, DOI: 10.4061/2010/290935 Available at: https://www.hindawi.com/journals/jna/2010/290935/

Mechanobiology Institute (2018) What are chromosomes and chromosome territories? *Produced by the National University of Singapore*. Available at: https://www.mechanobio.info/genome-regulation/what-are-chromosomes-and-chromosome-territories/

National Human Genome Research Institute (2019) The Human Genome Project. *National Institutes of Health*. Available at: https://www.genome.gov/human-genome-project

Wikipedia contributors. (2019, July 8). DNA. In *Wikipedia, The Free Encyclopedia*. Retrieved 02:41, July 22, 2019, from https://en.Wikipedia.org/w/index.php?title=DNA&oldid=905364161

Wikipedia contributors. (2019, July 22). Chromosome. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:18, July 23, 2019, from en.Wikipedia.org/w/index.php?title=Chromosome&oldid=907355235

Wikilectures. Prokaryotic Chromosomes (2017) In MediaWiki, Available at: https://www.wikilectures.eu/w/Prokaryotic_Chromosomes

Wikipedia contributors. (2019, May 15). DNA supercoil. In *Wikipedia, The Free Encyclopedia*. Retrieved 19:40, July 25, 2019, from en.Wikipedia.org/w/index.php?title=DNA_supercoil&oldid=897160342

Wikipedia contributors. (2019, July 23). Histone. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:19, July 26, 2019, from en.Wikipedia.org/w/index.php?title=Histone&oldid=907472227

Wikipedia contributors. (2019, July 17). Nucleosome. In *Wikipedia, The Free Encyclopedia*. Retrieved 17:17, July 26, 2019, from en.Wikipedia.org/w/index.php?title=Nucleosome&oldid=906654745

Wikipedia contributors. (2019, July 26). Human genome. In *Wikipedia, The Free Encyclopedia*. Retrieved 06:12, July 27, 2019, from en.Wikipedia.org/w/index.php?title=Human_genome&oldid=908031878

Wikipedia contributors. (2019, July 19). Gene structure. In *Wikipedia, The Free Encyclopedia*. Retrieved 06:16, July 27, 2019, from en.Wikipedia.org/w/index.php?title=Gene_structure&oldid=906938498

8.4: Chromosomes and Chromatin

8.4.1: Chromatin

When stained and viewed in a microscope, eukaryotic nuclear DNA in nondividing cells is observed in two different states, **heterochromatin** (dark areas) and euchromatin (light areas), as shown in Figure 8.4.1.

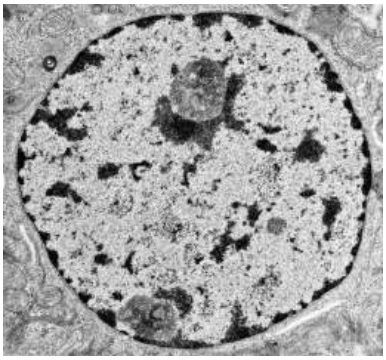


Figure 8.4.1: Nucleus showing heterochromatin (dark) and euchromatin (light). http://medcell.med.yale.edu/histolog...ochromatin.php

The heterochromatin is darkly stained and found along the inner side of the nuclear envelope as well as inside the nucleus. Euchromatin doesn't stain well since it is more dispersed, and is found throughout the nucleus. The image above is of a cell in interphase, a part of the cell cycle in which the cell is in-between cell divisions. Here are some differences between heterochromatin and euchromatin:

Heterochromatin:

- contains genes that are transcriptionally inactive and not expressed;
- is not found in prokaryotic cells;
- in humans, one of the X chromosomes is silenced in heterochromatin while the other is transcribable and in euchromatin;
- protects the tightly packed in it from nucleases.

Euchromatin:

- contains about 90% of genome DNA;
- contains chromosomes that are unfolded into nucleosomal DNA resembling beads on a string and is characterized by looser association with packaging histone proteins;
- is the only form of DNA in bacteria;
- is transcriptionally active so it is open to the binding of RNA polymerase and transcription factors;
- has some genes which can be transcriptionally silenced by moving them into heterochromatin.

Look as long as you wish but you won't see the iconic pictures of chromosomes in Figure 8.4.1. They are there but dispersed. They are only discretely visible at certain times in the cell cycle when cells are preparing to divide.

Cell Cycle

For those with a more chemistry-centric background, a short review of the cell cycle would be helpful. Somatic (non-germ cells) spend only part of their lives preparing for and dividing into two cells in a process called mitosis. Figure 8.4.2 shows a simple model for the cell cycle.

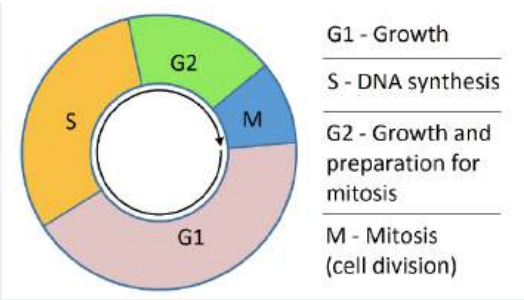
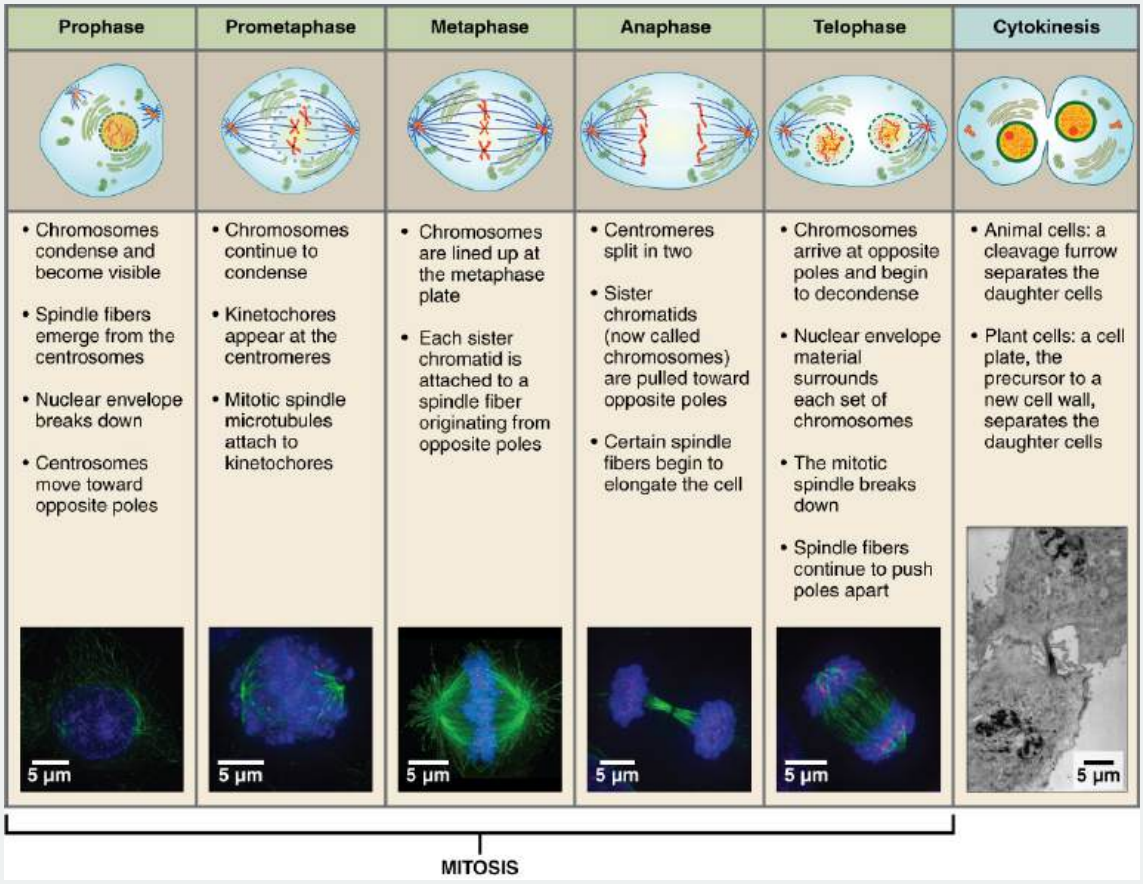


Figure 8.4.2: The Cell Cycle. https://commons.wikimedia.org/wiki/F...cle_simple.png. Creative Commons Attribution-Share Alike 3.0 Unported

Before a cell gets ready to divide, it is in **interphase** (between mitotic phases), which encompasses the **G1, S, and G2** phases of the cell cycle. In interphase, cells grow and replicate their DNA **Mitosis** or cell division occurs after G2 and consists of a series of new discrete phases, including **prophase, metaphase, anaphase, and telophase**, after which the cell divides in a process called cytokinesis. Table 8.4.1 below shows the stages of mitosis and cytokinesis which occurs after interphase.



Let's start with a discussion of the structure of chromosomes as classically observed in mitotic cells. Then we will look more closely at the seemingly amorphous and more complicated structures of chromatin.

8.4.2: Chromosomes

Within eukaryotic cells, DNA is organized into long linear structures called **chromosomes.** A **chromosome** is a "thread-like" structure in the nucleus of animal and plant that consists of a single but long molecule of double-stranded DNA (so it's two ss-DNAs) with a myriad of bound proteins. The proteins bind to and condense the DNA molecule to prevent it from becoming an unmanageable tangle. Before typical cell division (mitosis), these chromosomes are replicated by DNA polymerase to make two identical chromosomes, one for each future daughter cell. The two identical chromosomes, called **sister chromatids**, bind to each

other at a common structure called the **centromere**. A replicated chromosome (sister chromatids) bound to each other at the centromere, is shown in Figure 8.4.3. Each chromosome in the sister chromatid structure represents one chromatid.

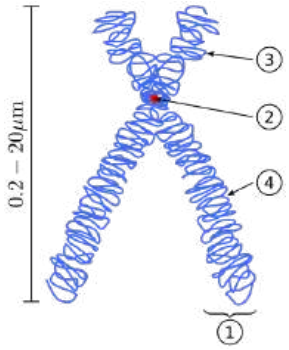


Figure 8.4.3: **Diagram of Replicated and Condensed Eukaryotic Chromosome (sister chromatids).** (1) Chromatid – one of the two identical parts of the chromosome after S phase. (2) Centromere – the point where the two chromatids are joined together. (3) Short arm is termed *p*; Long arm is termed *q*. *Image* by: Magnus Manske, Dietzel65, and Tryphon

In humans, each cell normally contains 23 pairs of chromosomes, for a total of 46. In twenty-two of these pairs, called *autosomes*, each member of the pair is similar. The 23rd pair, the sex chromosomes, differ between males and females. Females have two copies of the X chromosome, while males have one X and one Y chromosome. Figure 8.4.4 shows a DNA **karyotype** with 22 pairs of autosomes and one pair of sex chromosomes. Karyotypes are prepared using standardized staining procedures that reveal characteristic structural features for each chromosome, usually from white blood cells. The karyotype of human cells shown in Figure 4 below shows an extra copy of chromosome 21 (trisomy 21), which causes Down's Syndrome.

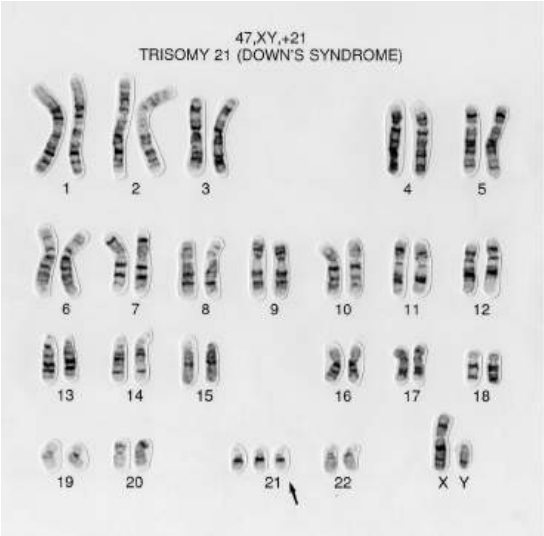


Figure 8.4.4: Karyotype of the human chromosome from a human male with Down's Syndrome. https://wellcomecollection.org/works/wmcdanw6. **Creative Commons Attribution 4.0 International License**

Each species has a unique chromosomal complement. For example, chickens have 39 pairs of chromosomes (78 total) with 38 autosomal pairs and one pair of sex chromosomes (Z and W). In this species, ZW chickens are female and ZZ chickens are male. The total length of the human genome is over 3 billion base pairs. The total length of the human genome is over 3 billion base pairs. The genome also includes mitochondrial DNA.

Eukaryotic organisms (animals, plants, fungi, and protists) store most of their DNA inside the cell nucleus as linear nuclear DNA. However, DNA in the mitochondria and chloroplast are circular (with no ends). Bacteria and Archaea cells do not have organelle structures and thus, store their DNA only in a region of the cytoplasm known as the *nucleoid region*. Prokaryotic chromosomes consist of *double–stranded circular DNA*.

8.4.3: Packing of DNA: Supercoiling

The genome of a cell is often significantly larger than the cell itself. For example, if the DNA from a human cell containing 46 chromosomes were stretched out in a line, it would extend more than 6 feet (2 meters)! How is it possible that the genetic

information not only fits into the cell but fits into the cell nucleus? Eukaryotes solve this problem by a combination of supercoiling and packaging DNA around the histone family of proteins (described below). Prokaryotes do not contain histones (with a few exceptions). Prokaryotes tend to compress their DNA using **nucleoid-associated-proteins (NAPs)** and **supercoiling,** as shown in Figure 8.4.5.

DNA supercoiling refers to the over- or under-winding of a DNA strand and is an expression of the strain on that strand. Supercoiling is important in a number of biological processes, such as compacting DNA and regulating access to the genetic code. DNA supercoiling strongly affects DNA metabolism and possibly gene expression. Additionally, certain enzymes such as **topoisomerases** can change DNA topology to facilitate functions such as DNA replication or transcription.

In a "relaxed" double-helical segment of B-DNA, the two strands twist around the helical axis once every 10.4–10.5 base pairs of sequence. Adding or subtracting twists, as some enzymes can do, impose strain. If a DNA segment under twist strain were closed into a circle by joining its two ends and then allowed to move freely, the circular DNA would contort into a new shape, such as a simple figure-eight, as shown in Figure 8.4.5. Such a contortion is a *supercoil*. The noun form "supercoil" is often used in the context of DNA topology.

Figure 8.4.5: DNA Supercoiling. The supercoiled structure of linear DNA molecules with constrained ends. The helical nature of the DNA duplex is omitted for clarity. *mage by: Richard Wheeler*

Positively supercoiled (overwound) DNA is transiently generated during DNA replication and transcription, and, if not promptly relaxed, inhibits (regulates) these processes. The simple figure eight is the simplest supercoil and is the shape a circular DNA assumes to accommodate one too many or one too few helical twists. The two lobes of the figure eight will appear rotated either clockwise or counterclockwise with respect to one another, depending on whether the helix is over- or underwound. For each additional helical twist being accommodated, the lobes will show one more rotation about their axis. As a general rule, the DNA of most organisms is negatively supercoiled.

Lobal contortions of a circular DNA, such as the rotation of the figure-eight lobes above, are referred to as *writhe*. The above example illustrates that twisting and writhing are interconvertible. Supercoiling can be represented mathematically by the sum of *twist* and *writhe*, The *twist* is the number of helical turns in the DNA and the writhe is the number of times the double helix crosses over on itself (these are the supercoils). Extra helical twists are positive and lead to positive supercoiling, while subtractive twisting causes negative supercoiling. Many topoisomerase enzymes sense supercoiling and either generate or dissipate it as they change DNA topology.

In part, because chromosomes may be very large, segments in the middle may act as if their ends are anchored. As a result, they may be unable to distribute excess twist to the rest of the chromosome or to absorb twist to recover from underwinding—the

segments may become *supercoiled*, in other words. In response to supercoiling, they will assume an amount of writhe, just as if their ends were joined.

Supercoiled circular DNA forms two major structures; a *plectoneme* or a *toroid*, or a combination of both. A negatively supercoiled DNA molecule will produce either a one-start left-handed helix, *the toroid*, or a two-start right-handed helix with terminal loops, the *plectoneme*. *Plectonemes* are typically more common in nature, and this is the shape most bacterial plasmids will take. For larger molecules, it is common for hybrid structures to form – a loop on a toroid can extend into a plectoneme as shown in Figure 8.4.6. DNA supercoiling is important for DNA packaging within all cells and seems to also play a role in gene expression.

Figure 8.4.6: Bacterial DNA Supercoiling. Atomic force microscopy (AFM) visualization of torsionally relaxed (A), and negatively supercoiled (B) bacterial plasmids pBR322. (C) Electron microscopy image of the E. coli chromosomal DNA displaying a hybrid toroidal-plectoneme structure. Image A and B from Witz, G. and Stasiak, A. (2009) Nucleic Acids Research 38(7):2119-2133. Image C from Prokaryotic Chromosomes

In addition to forming supercoiled structures, circular chromosomes from bacteria have been shown to undergo the processes of *catenation* and *knotting* upon the inhibition of topoisomerase enzymes. *Catenation* is the process by which two circular DNA strands are linked together like chain links, whereas *DNA knotting* is the interlooping structures occurring within a single circular DNA structure. These are illustrated in Figure 8.4.7. *In vivo*, the action of topoisomerase enzymes is critical to keep knots and catenoids from tangling the DNA structure.

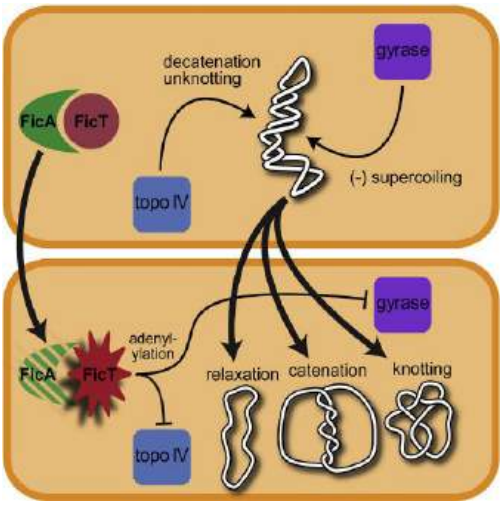


Figure 8.4.7: DNA Catenation and Knotting. The upper structure shows the negatively supercoiled form of bacterial DNA. The inhibition of topoisomerase enzyme activity leads to the relaxation, catenation, and knotting of the chromosomal structure. *Harms, A. et al. (2015) Cell Reports 12(9):1497-1507*.

Mitochondrial and Chloroplast DNA

Mitochondrial and Chloroplast DNA are circular suggesting a bacterial origin for both of these organelle structures. Sequence alignments further lend support for the *endosymbiotic theory*, which proposes that bacteria were engulfed by early eukaryotic organisms and subsequently became symbiotic to their eukaryotic counterpart, rather than being digested.

In the cells of eukaryotic organisms, the vast majority of the proteins present in the mitochondria (numbering approximately 1500 different types in mammals) are coded for by nuclear DNA. However, sequencing of the human mitochondrial genome has revealed 16,569 base pairs encoding 13 proteins, as shown in Figure 8.4.8. Many of the mitochondrially produced proteins are required for electron transport during the production of ATP.

Figure 8.4.8: Mitochondrial Genome. Mitochondria are organelle structures containing a double membrane, thought to have originated as an independent prokaryotic organism that was originally engulfed by a eukaryotic organism, where it became a symbiotic counterpart. Mitochondria contain circular chromosomal DNA that shares high sequence similarity with alphaprotobacteria. The human mitochondrial genome contains 16,569 base pairs encoding 13 proteins and ribosomal RNA (rRNA) components. *Images adapted from The National Human Genome Research Institute and Shanel, Knopfkind, and JHC*

8.4.4: Histones and Nucleosomes

Within eukaryotic chromosomes, chromatin proteins, known as **histones**, compact and organize DNA. These compacting structures guide the interactions between DNA and other proteins, helping control which parts of the DNA are transcribed.

Histones are highly basic proteins found in eukaryotic cell nuclei that package and order the DNA into structural units called **nucleosomes**. They are the chief protein components of chromatin, acting as spools around which DNA winds, and playing a role in gene regulation. Without histones, the unwound DNA in chromosomes would be very long (a length-to-width ratio of more than 10 million to 1 in human DNA). For example, each human diploid cell (containing 23 pairs of chromosomes) has about 1.8 meters of DNA wound on the histones, and the diploid cell has about 90 micrometers (0.09 mm) of chromatin.

There are five major families of histones, H1/H5, H2A, H2B, H3, and H4. Histones H2A, H2B, H3, and H4 are known as the core histones, while histones H1/H5 are known as the linker histones.

The core histones all exist as dimers, which are similar in that they all possess the histone fold domain: three alpha helices linked by two loops (Figure 4.13). It is this helical structure that allows for interaction between distinct dimers, particularly in a head-tail fashion (also called the handshake motif). The resulting four distinct dimers then come together to form one octameric nucleosome core, approximately 63 Angstroms in diameter. Around 146 base pairs (bp) of DNA wrap around this core particle 1.65 times in a left-handed super-helical turn to give a particle of around 100 Angstroms across, called a nucleosome as illustrated in Figure 8.4.9.

Figure 8.4.9: Nucleosome Core Structure. Histones H2A and H2B dimerize, and Histones H3 and H4 dimerize. Two dimers of each join to form a histone core octamer. The DNA double helix winds 1.65 times around the octamer core forming the nucleosome structure. *image adapted from Nucleosome Structure*

The linker histone H1 binds the *nucleosome* at the entry and exit sites of the DNA, thus locking the DNA into place and allowing the formation of higher order structure (Figure 4.14). The most basic such formation is the 10 nm fiber or beads on a string conformation. This involves the wrapping of DNA around nucleosomes with approximately 50 base pairs of DNA separating each pair of nucleosomes (also referred to as linker DNA).

The nucleosome contains over 120 direct protein-DNA interactions and several hundred water-mediated ones. Direct protein – DNA interactions are not spread evenly about the octamer surface but rather located at discrete sites. These are due to the formation of two types of DNA binding sites within the octamer; the α1α1 site, which uses the α1 helix from two adjacent histones, and the L1L2 site formed by the L1 and L2 loops. Salt links and hydrogen bonding between both side-chain basic and hydroxyl groups and main-chain amides with the DNA backbone phosphates form the bulk of interactions with the DNA. This is important, given that the ubiquitous distribution of nucleosomes along genomes requires it to be a non-sequence-specific DNA-binding factor. Although nucleosomes tend to prefer some DNA sequences over others, they are capable of binding practically to any sequence, which is thought to be due to the flexibility in the formation of these water-mediated interactions. In addition, non-polar interactions are made between protein side-chains and the deoxyribose groups, and an arginine side-chain intercalates into the DNA minor groove at all 14 sites where it faces the octamer surface. The spatial distribution and strength of DNA-binding sites about the octamer surface distort the DNA within the nucleosome core. The DNA is non-uniformly bent and also contains twist defects. The twist of free B-form DNA in solution is 10.5 bp per turn. However, the overall twist of nucleosomal DNA is only 10.2 bp per turn, varying from a value of 9.4 to 10.9 bp per turn.

The histone tail extensions constitute up to 30% by mass of histones, but are not visible in the crystal structures of nucleosomes due to their high intrinsic flexibility, and have been thought to be largely unstructured (Figure 4.14). The N-terminal tails of histones H3 and H2B pass through a channel formed by the minor grooves of the two DNA strands, protruding from the DNA every 20 bp. The N-terminal tail of histone H4, on the other hand, has a region of highly basic amino acids (16-25), which, in the crystal structure, forms an interaction with the highly acidic surface region of an H2A-H2B dimer of another nucleosome, being potentially relevant for the higher-order structure of nucleosomes. This interaction is thought to occur under physiological conditions also and suggests that acetylation of the H4 tail distorts the higher-order structure of chromatin.

Figure 8.4.10 shows an interactive iCn3D model of the human nucleosome (3afa). One member of each pair of histones is shown in cartoon rendering, while the other member of the pair is shown in the same color but in spacefill rendering. The structure of a human nucleosome (3afa) is shown below (H2A is shown in cyan, H2B in blue, H3 in magenta, and H4 in purple). Each strand of DNA is shown in a different shade of gray.

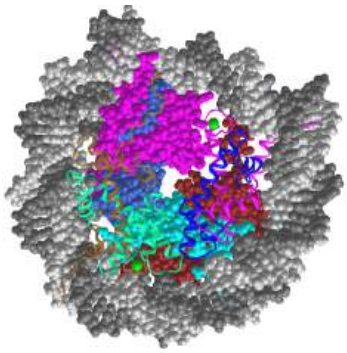


Figure 8.4.10: Human nucleosome (3afa). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...B2SwQHYDLj4BJ6

The packing of DNA from dsDNA to the metaphase chromosomes is schematically shown in Figure 8.4.11. The formation of the DNA double helix represents the first-order packaging of the chromosome structure. The formation of nucleosomes represents the second level of packaging for eukaryotic chromosomes. *In vitro* data suggests that nucleosomes are then arranged into either a solenoid structure which consists of 6 nucleosomes linked together by the Histone H1 linker proteins or a zigzag structure that is similar to the solenoid construct. Both the solenoid and zigzag structures are approximately 30 nm in diameter. The solenoid and zigzag structures reported from *in vitro* data have not yet been confirmed to occur *in vivo.*

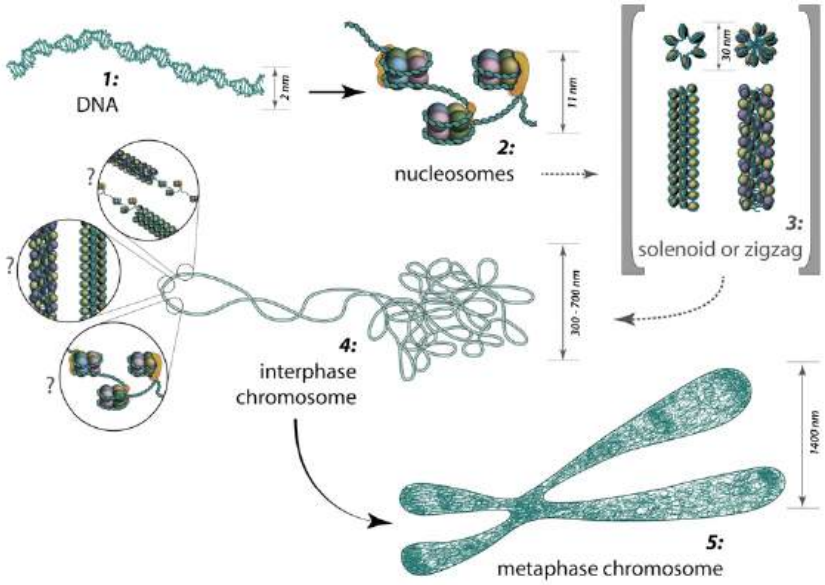


Figure 8.4.11: Chromosome Structure. (1) DNA double helix is approximately 2 nm in diameter. (2) The nucleosome core structure is approximately 11 nm in diameter. (3) The solenoid/zigzag structure is approximately 30 nm in diameter and is proposed to form chromosome loops (4) during cellular interphase and more condensed chromosome territories (5) during mitosis. *Image by MBInfo*

8.4.5: Telomeres

At the ends of the linear eukaryotic chromosomes are specialized regions of DNA called telomeres. The main function of these regions is to allow the cell to replicate chromosome ends using the enzyme telomerase, as the enzymes that normally replicate DNA cannot copy the extreme 3′ ends of chromosomes. A cartoon showing telomeres and their extension is shown in Figure 8.4.12

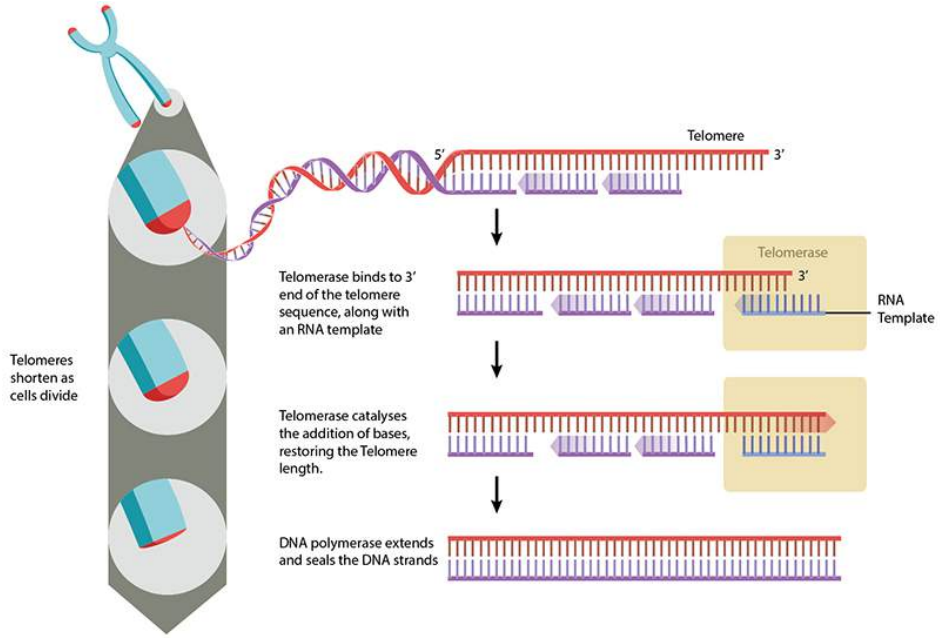
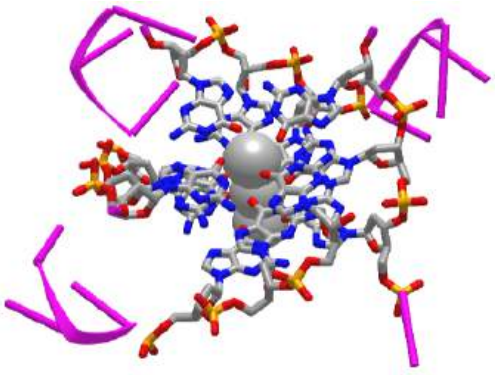


Figure 8.4.12: Telomer structure and extension. https://www.mechanobio.info/genome-r...are-telomeres/. Creative Commons Attribution-NonCommercial 4.0 International License.

These specialized chromosome caps also help protect the DNA ends, and stop the DNA repair systems in the cell from treating them as damage to be corrected.

In human cells, telomeres contain 300-8000 repeats of a simple **TTAGGG** sequence. The repetitive **TTAGGG** sequences in telomeric DNA can form unique higher-order structures called **quadruplexes**. Figure 8.4.13 shows an interactive iCn3D model of parallel quadruplexes from human telomeric DNA (1KF1). The structure contains a single DNA strand **(**5'-AGGGTTAGGGTTAGGGTTAGGG-3') which contains four TTAGGG repeats.



NCBI iCn3D Figure 8.4.13: A buried phenylalanine in low molecular weight protein tyrosyl phosphatase (1xww) (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...y5joFHDgWJQsQ6

Rotate the model to see 3 parallel layers of quadruplexes. In each layer, **4** noncontiguous guanine bases interact with a K^+ ion. Hover over the guanine bases in one layer and you will find that one layer consists of guanines 4, 10, 16, and 22, which derive from the last G in each of the repeats in the sequence of the oligomer used (5'-AGGGTTAGGGTTAGGGTTAGGG-3'). These quadruplexes certainly serve as recognition and binding sites for telomerase proteins. The guanine-rich telomere sequences which can form quadruplex may also function to stabilize chromosome ends

During DNA replication, the double-stranded DNA is unwound and DNA polymerase synthesizes new strands. However, as DNA polymerase moves in a unidirectional manner (from 5' to 3'), only the **leading strand** can be replicated continuously. For the complementary **lagging strand**, DNA replication is discontinuous. In humans, small RNA primers attach to the lagging strand

DNA, and the DNA is synthesized in small 5'-3' stretches of about 100-200 nucleotides, which are termed **Okazaki fragments.** The RNA primers are removed and replaced with DNA and the Okazaki DNA fragments are ligated together. At the end of the lagging strand, it is impossible to attach an RNA primer, meaning that there will be a small amount of DNA lost each time the cell divides. This ‘end replication problem’ has serious consequences for the cell as it means the DNA sequence cannot be replicated correctly, with the loss of genetic information. Hence most telomeres have 3´ overhangs. Bacteria DNA, which is circular, does not have the problem.

To prevent this, telomeres are repeated hundreds to thousands of times at the end of the chromosomes. Each time cell division occurs, a small section of telomeric sequences is lost to the end replication problem, thereby protecting the genetic information. At some point, the telomeres become critically short. This decreases leads to cell senescence, where the cell is unable to divide, or apoptotic cell death. Telomeres are the basis for the **Hayflick** limit, the number of times a cell can divide before reaching senescence.

Telomeres can be restored by the enzyme **telomerase**, which extends the telomere's length. Telomerase activity is found in cells that undergo regular division, such as stem cells and lymphocyte cells of the immune system. The enzyme has two major subunits. One is the catalytic enzyme named **telomerase reverse transcriptase (TERT)**, which is an RNA-dependent DNA polymerase. Almost all DNA polymerases use DNA as a template for replication. Some RNA viruses that use RNA as their genetic information like HIV encode their own **reverse transcriptase**, which directs the polymerization of a DNA copy of the viral RNA genome as part of its life cycle.)The virus that causes the Covid-19 pandemic is called severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). It is also an RNA virus but in contrast to HIV it uses and encodes an RNA-dependent RNA polymerase.)

The second subunit of telomerase is the **telomerase RNA (TR)**, which contains a template from which the new telomer is made. The enzyme makes many DNA copies from this to create a multitude of DNA repeat in the telomeres. Figure 8.4.14 shows the structure of the telomerase RNA used to build new telomers. Ovals show proteins that form part of the complex. 'For this discussion, the most important is TERT, telomerase reverse transcriptase, the RNA-dependent DNA polymerase which synthesizes new telomeric DNA from the template sequence of the RNA.

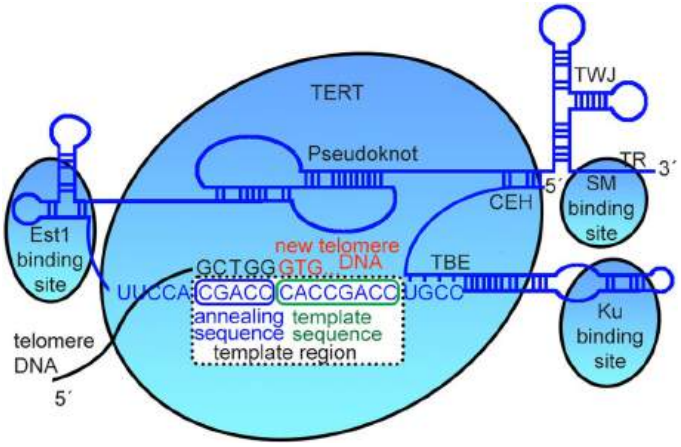


Figure 8.4.14: Telomerase RNA organization for the yeast Lachancea sp. The most relevant binding sites for protein-RNA interactions are indicated by blue ellipses (Est1, SM, and KU), and the bigger one is the contact region with the telomerase reverse transcriptase (TERT). TBE template boundary element, TWJ three-way junction, CEH core enclosed helix. Adapted from Waldl et al.36 Te template region is located between the EST1 binding site and Ku binding hairpins. Peska, V *et al. Sci Rep* **11,** 12784 (2021).Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Trace the black single-stranded telomeric DNA strand from the 5' end (bottom left) to its 3' end containing the terminal sequence **5'GCTGG3'**. Now start tracing the telomerase RNA starting with its 5' end, which is between the CEH and SM binding site in Figure 8.4.14. It bends sharply near TBE, and continues to the right through the Ku binding site, where it forms stem-loops. It then continues into the template region containing a template sequence (3'CACCGACC5'), from which new telomeric DNA is made (5'GTGGCTGG3' - a slightly different repeat than the human repeat), and an annealing sequence (3'CGACC5'), which is complementary to the 3' end of the existing telomeric DNA (5'GCTGG3'). The RNA continues in the 3' direction through the Est1 bind site region, through the pseudoknot, and ends at the 3' end past the SM binding site.

Telomeres can also be extended through the Alternative Lengthening of Telomeres (ALT) pathway. In this case, rather than being extended, telomeres are switched between chromosomes by homologous recombination. As a result of the telomere swap, one set of daughter cells will have shorter telomeres, and the other set will have longer telomeres

A downside to telomere extension is the potential for uncontrolled cell division and cancer. Abnormally high telomerase activity has been found in the majority of cancer cells, and non-telomerase tumors often exhibit ALT pathway activation. As well as the potential for losing genetic information, cells with short telomeres are at a higher risk for improper chromosome recombination, which can lead to genetic instability and aneuploidy (an abnormal number of chromosomes).

8.4.6: Chromatin Structure

During interphase, distinct chromosomes as shown in Figure 8.4.4 are not observed. Rather each chromosome occupies a spatially limited, roughly elliptical domain which is known as a **chromosome territory (CT)**. Each chromosome territory is comprised of higher-order chromatin units of ~1 Mb each. These units are likely built up from smaller loop domains that contain the solenoid/zigzag structural motifs. On the other hand, 1Mb domains can themselves serve as smaller units in higher-order chromatin structures. With the development of high-throughput biochemical techniques, such as 3C (chromosome conformation capture) and 4C (chromosome conformation capture-on-chip and circular chromosome conformation capture), numerous spatial interactions between neighboring chromatin territories have been described as shown in Figure 8.4.15.

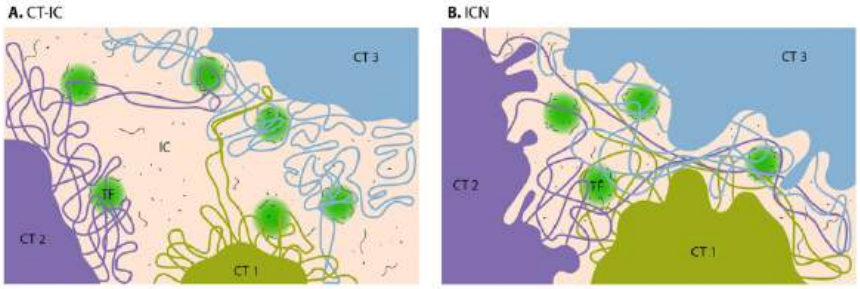


Figure 8.4.15: Computer Models of Chromosome Territory (CT) Structure. In the CT-IC model, the space between discrete CTs can be visualized in light and electron microscope and is called an interchromatin compartment (IC). Transcription factories (TF, green color) are localized predominantly in the perichromatin region. In the ICN model, the interchromatin compartment is not apparent. Instead, the space between CTs is occupied by intermingling decondensed chromatin loops, which often share the same transcription factories. *Image by MBInfo*

Chromosome territories are known to be arranged radially around the nucleus. This arrangement is both cell and tissue-type specific and is also evolutionarily conserved. The radial organization of chromosome territories was shown to correlate with their gene density and size. In this case, the gene-rich chromosomes occupy interior positions, whereas larger, gene-poor chromosomes, tend to be located around the periphery. Chromosome territories are also dynamic structures, with genes able to relocate from the periphery towards the interior once they have been 'switched on'. In other cases, genes may move in the opposite direction or simply maintain their position.

Figure 8.4.16 shows nano- to more micro-folding structures of chromosomes in the nucleus. The top left shows the most zoomed-in view where DNA is wound around histone complexes (nucleosomes), which condense to form 10 nm fibers. These condense further into **Topologically Associated Domains (TAD)** which further separate into **Compartments A and B**. These then pack into discrete **territories**. Individual chromosomes occupy their own chromosome territories in the nucleus.

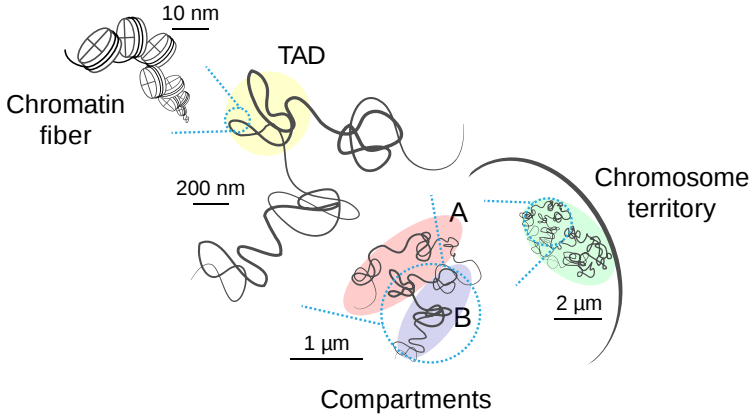


Figure 8.4.16: Schematic view of chromosome folding inside the nucleus Szabo Q et al. Sci Adv. 2019 Apr 10;5(4):eaaw1668. doi: 10.1126/sciadv.aaw1668. PMID: 30989119; PMCID: PMC6457944.Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC)

This may seem very complex and it is, but it is somewhat analogous to protein folding, which starts with a linear primary sequence and moves into more complicated secondary structures, secondary structure motifs, domains, tertiary structures, and quaternary structures, which display varying degrees of symmetry.

Because these different types of folding and compaction of chromosomes are difficult to visualize, we will present several different representations of these nano- to microstructures to help your understanding. Figure 8.4.17 offers one that shows scaling factors and differential structures of chromosomes and chromatin.

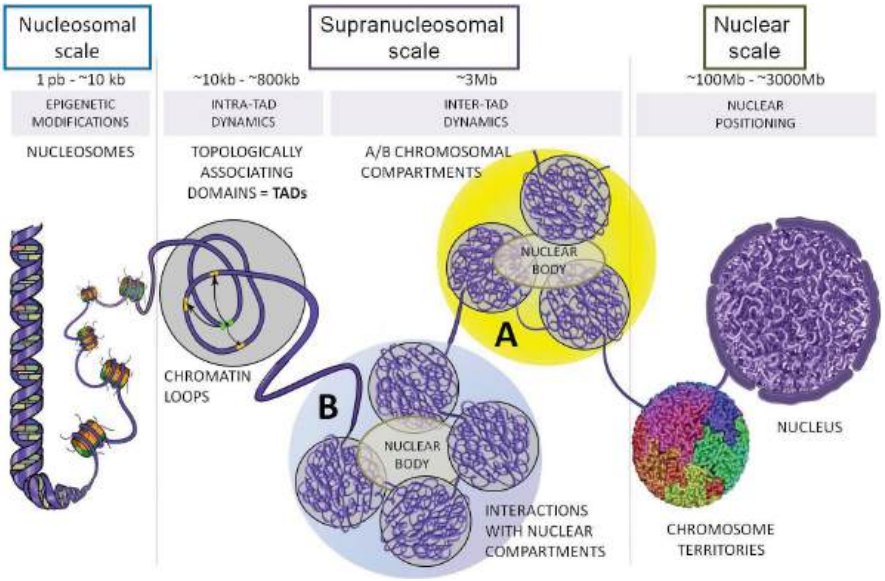


Figure 8.4.17: Schematic representation of genome organization in mammals. a, V.; Baudement, M.-O.; Lesne, A.; Forné, T. Contribution of Topological Domains and Loop Formation to 3D Chromatin Organization. *Genes* **2015**, *6*, 734-750. https://doi.org/10.3390/genes6030734. Creative Commons Attribution license (**http://creativecommons.org/licenses/by/4.0/**).

The large-scale A compartment is gene-rich and actively transcribed so it best represents euchromatin. In contrast, B compartments are gene-poor and best represent heterochromatin. At the subscale level, boundaries between TADs are transcriptionally rich and can be separated from other TADS by heterochromatin "islands".

Another representation of chromatin organization that emphasizes TADs is shown in Figure 8.4.18

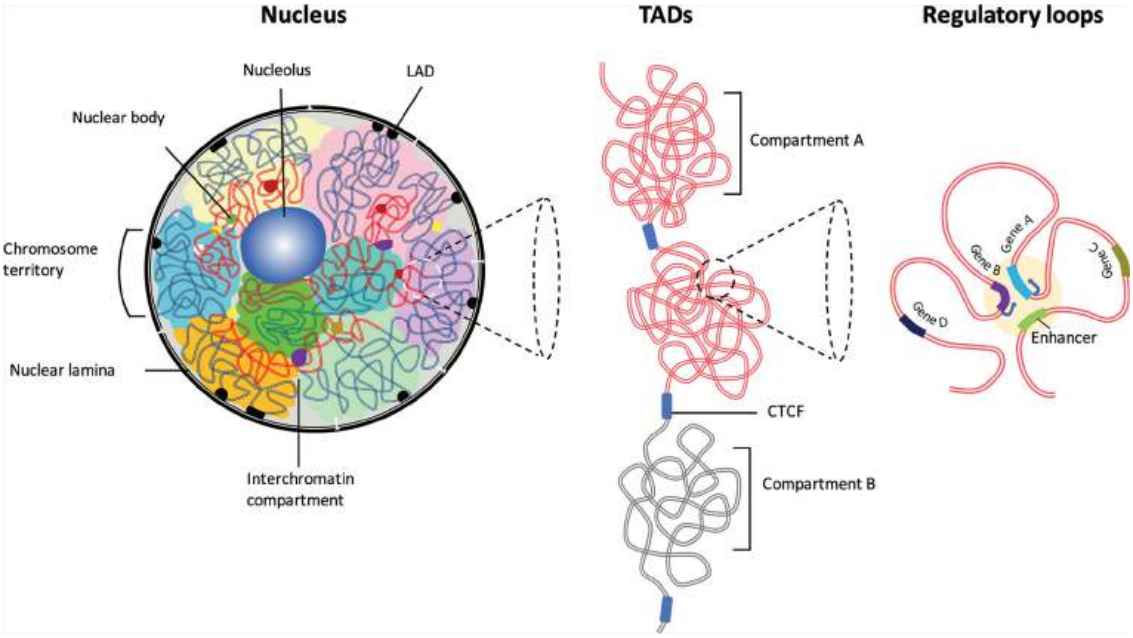


Figure 8.4.18: The hierarchical organization of the 3D chromatin inside the interphase nucleus. Chromatin interactions occur predominantly between compartments with similar biochemical or functional properties. The majority of the chromatin interactions are intra-chromosomal. Preferential self-interactions within the heterochromatin and euchromatin (A and B compartments) regions result in the formation of topologically associating domains (TADs), demarcated by boundary elements enriched with CTCF/Cohesin proteins. Jagan M. R. Pongubala and Cornelis Murre. Front. Immunol., 29 March 2021 | https://doi.org/10.3389/fimmu.2021.633825. Creative Commons Attribution License (CC BY).

The TAD boundaries in Figure 8.4.18 show a blue CTCF in-between compartments, with each TAD consisting of many interacting loops. A more detailed view is shown in Figure 8.4.19.

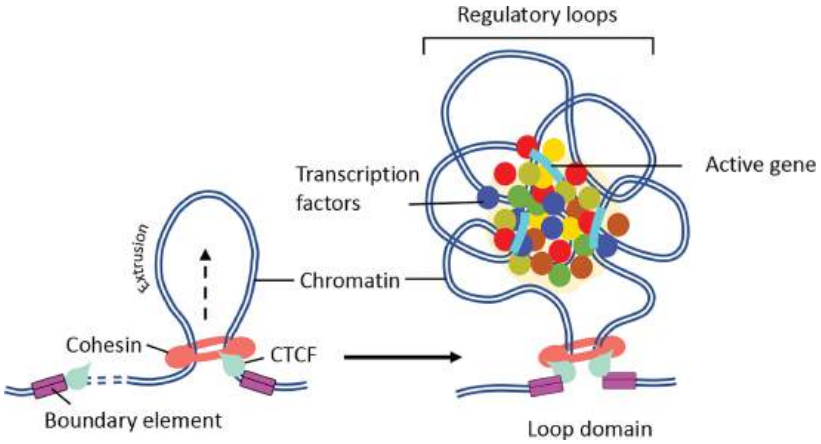


Figure 8.4.19: Jagan M. R. Pongubala and Cornelis Murre. Front. Immunol., 29 March 2021 | https://doi.org/10.3389/fimmu.2021.633825. Creative Commons Attribution License (CC BY).

A protein complex containing cohesin (which forms a ring) and CTCF is found at sites where loops of relationally transcribable DNA are extruded through the ring. The extruded loops interact with others to form a cluster of regulatory loops containing genes with similar potential for transcription (activated if they end up in Compartment A or repressed if in Compartment B). Cohesin is a functional complex that forms a ring that traps sister chromatids. During anaphase, the complex is cleaved and the sister chromatids separate. CTCF is a chromatin-binding factor with many associated activities.

"Mechanism of chromatin loop formation. TADs contain varying numbers of chromatin loops generated through loop extrusion by CTCF/cohesin
complexes. (Right panel) In the presence of NIPBL and MAU2, the cohesin complex loaded onto the DNA. Then, cohesin extrudes chromatin until a pair of convergent CTCF binding sites is reached. (Right panel) The N-terminus of CTCF and convergent

positioning of the CTCF-DNA complex stabilizes cohesin binding and stall chromatin extrusion leading to the establishment of higher-order chromatin organization. The intervening DNA between two convergent CTCF sites leads to the formation of a loop domain, which adopts a variety of complex shapes comprised of multiple regulatory loops. The internal structure of the loop domain is likely determined by polymer chromatin-chromatin self-interactions, which may be further stabilized by phase separation. The contacts within the loop domains facilitate the targeting of enhancers to specific genes (104). The black arrow depicts the direction of loop extrusion."

Special elements in the DNA called **enhancers** and **silencers** of gene transcription have long been known to influence gene transcription. These are sequences that are cis (i.e. on the same molecule of DNA, not trans factors like separate proteins that bind to promoters, for example) and can be quite distant from the proximal promoter sequence which controls the transcription of a target gene. How might they work from such a large distance from the promoter? One obvious answer is the DNA folds in 3D space so the enhancers and silencers are close to each in 3D space. Chromatin loop formation facilitates such promoter and enhancer/silencer interactions. Evidence suggests that specific enhancers and silencers are housed in loops in specific TADs so their effects are limited to a subset of genes. Enhancers and promoters seem to interact only within TADs, which suggests that TADS are a fundamental folding "domain" based on gene regulation. If boundaries between different TADS were removed, then promoters and enhancers in one TAD might affect transcription in other TADS, leading to aberrant gene expression. Chromatin loop formation facilitates interactions between promoter and enhancer/silencer elements. Figure 8.4.20 shows how enhancers and silencers can be brought close in space to promoters and their genes through TAD formation.

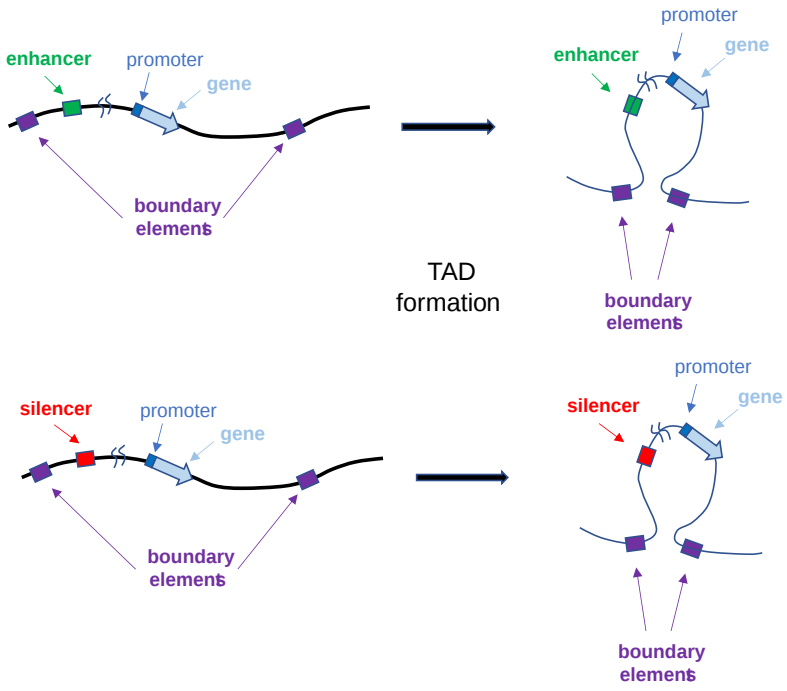


Figure 8.4.20: TADs bring enhancers and silencers close to promoters and their genes

TADs appear highly conserved in mammals and comprise most (90%) of the genome. The median size is about 880 Kb. The Boundaries between TADs have CCCTC-binding factor (CTCF) and the structural maintenance of chromosomes (SMC) cohesin complex. In Drosophila, TADs are organized by epigenetic state (methylation, condensation). For example, some TADS are transcriptionally active with epigenetic histone modifications (for example trimethylation of histone H3 lysines 4 and 36 or H3Kme3 and H3K36me3) that activate transcription). Others are transcriptionally repressed (enriched in H3K27me3 and containing Polycomb group (PcG) proteins) and some are more classically representative of heterochromatin.

8.4.7: Mechanisms of Separation of Euchromatin and Heterochromatin

From a chemistry focus, what are the interactions which stabilize TADs, compartments, and ultimately heterochromatin and euchromatin? It appears that the easiest way to conceptualize their separation in the nucleus is to use the idea of phase separations. Figure 8.4.21 shows a cartoon representation of the separation of heterochromatin and euchromatin in the zebrafish embryo. The

euchromatin is shown as more centrally located and dispersed with red dots indicating transcriptionally active sites. In the late blastula stage, gene expression is dramatically increased and the nucleolus and heterochromatin are not seen.

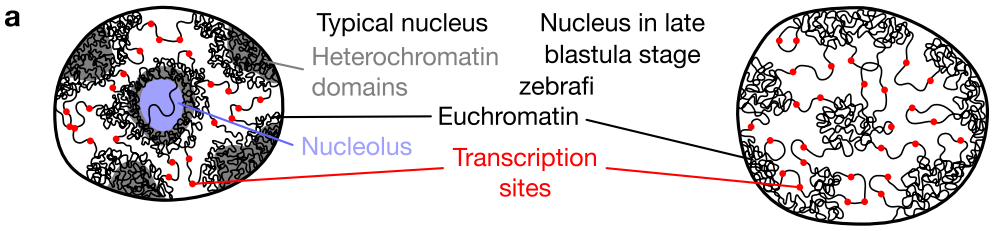


Figure 8.4.21: Sketch of nuclear compartmentalization in a typical nucleus and the nucleus of a late blastula (sphere) stage zebrafish embryo. Hilbert, L., Sato, Y., Kuznetsova, K. *et al.* Transcription organizes euchromatin via microphase separation. *Nat Commun* **12,** 1360 (2021). https://doi.org/10.1038/s41467-021-21589-3. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Analyses show that transcription forms regions enriched in RNA, RNA binding proteins, and accordingly transcriptionally active chromatin that are separated from transcriptionally inactive, heterochromatin. Given the dispersion of the DNA required for transcription, the regions enriched in RNA are also depleted in "stainable" DNA. Micrographs showing the depletion of DNA in the actively transcribed regions are shown in Figure 8.4.22 Note the clear lack of DNA in the white rectangles, which contain high levels of RNA and RNA polymerase II (as a proxy for protein).

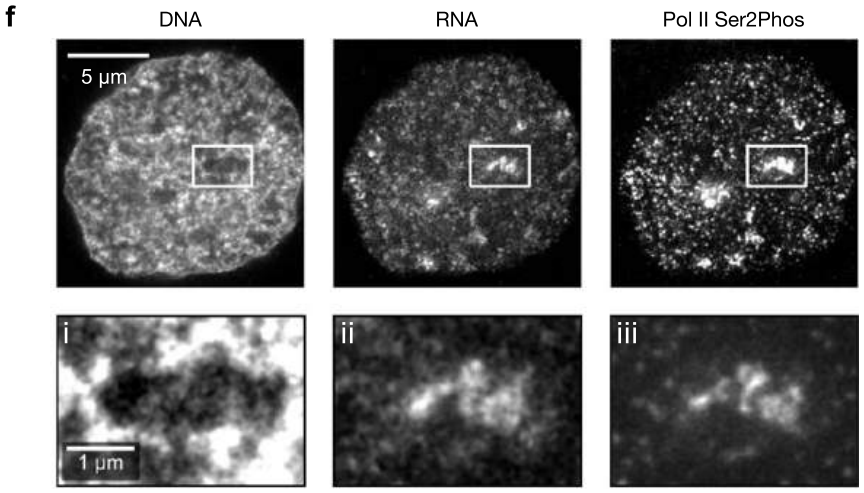


Figure 8.4.22: (F) Representative three-color micrographs showing DNA, RNA, and transcriptional activity (Pol II Ser2Phos) in a nuclear mid-section after transcription onset

It appears that the presence of a high concentration of RNA drives "phase" separation of heterochromatin and euchromatin. Hence euchromatin might act as an "oil in water" type microemulsion, with the euchromatin core stabilized by "tethered" RNA acting as an amphiphile. Ribonucleoproteins (RNPs) can be modeled as phase-separated droplets or condensates Figure 8.4.23 shows how increasing RNA leads to the formation of euchromatin "domains" which stay dispersed in the presence of continuing RNA formation.

Figure 8.4.23: (a) Cartoon representation of conventional phase separation and a microemulsion. The right panel focuses on the amphiphile in the microemulsion. Hilbert, L., Sato, Y., Kuznetsova, K. *et al.* Transcription organizes euchromatin via microphase separation. *Nat Commun* **12,** 1360 (2021). https://doi.org/10.1038/s41467-021-21589-3. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

A final summary cartoon that describes the formation and separation of chromosomal DNA into euchromatin and heterochromatin in Zebra fish is shown in Figure 8.4.24

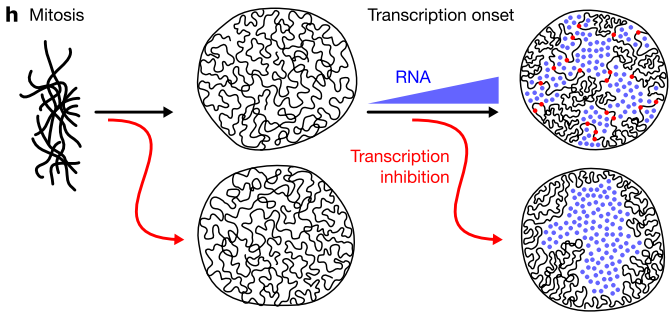


Figure 8.4.24: (h) Sketch summarizing the experimental observations to this point. Hilbert, L., Sato, Y., Kuznetsova, K. *et al.* Transcription organizes euchromatin via microphase separation. *Nat Commun* **12,** 1360 (2021). https://doi.org/10.1038/s41467-021-21589-3. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

8.5: References

This page is a draft and is under active development.

4.4 References

Börner, R., Kowerko, D., Miserachs, H.G., Shaffer, M., and Sigel, R.K.O. (2016) Metal ion induced heterogeneity in RNA folding studied by smFRET. *Coordination Chemistry Reviews* 327 DOI: 10.1016/j.ccr.2016.06.002 Available at: https://www.researchgate.net/publication/303846502_Metal_ion_induced_heterogeneity_in_RNA_folding_studied_by_smFRET

Hardison, R. (2019) B-Form, A-Form, and Z-Form of DNA. Chapter in: R. Hardison's *Working with Molecular Genetics.* Published by LibreTexts. Available at: https://bio.libretexts.org/Bookshelves/Genetics/Book%3A_Working_with_Molecular_Genetics_(Hardison)/Unit_I%3A_Genes%2C_Nucleic_Acids%2C_Genomes_and_Chromosomes/2%3A_Structures_of_Nucleic_Acids/2.5%3A_B-Form%2C_A-Form%2C_and_Z-Form_of_DNA

Lenglet, G., David-Cordonnier, M-H., (2010) DNA-destabilizing agents as an alternative approach for targeting DNA: Mechanisms of action and cellular consequences. *Journal of Nucleic Acids* 2010, Article ID: 290935, DOI: 10.4061/2010/290935 Available at: https://www.hindawi.com/journals/jna/2010/290935/

Mechanobiology Institute (2018) What are chromosomes and chromosome territories? *Produced by the National University of Singapore.* Available at: https://www.mechanobio.info/genome-regulation/what-are-chromosomes-and-chromosome-territories/

National Human Genome Research Institute (2019) The Human Genome Project. *National Institutes of Health*. Available at: https://www.genome.gov/human-genome-project

Wikipedia contributors. (2019, July 8). DNA. In *Wikipedia, The Free Encyclopedia*. Retrieved 02:41, July 22, 2019, from https://en.Wikipedia.org/w/index.php?title=DNA&oldid=905364161

Wikipedia contributors. (2019, July 22). Chromosome. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:18, July 23, 2019, from https://en.Wikipedia.org/w/index.php?title=Chromosome&oldid=907355235

Wikilectures. Prokaryotic Chromosomes (2017) In MediaWiki, Available at: https://www.wikilectures.eu/w/Prokaryotic_Chromosomes

Wikipedia contributors. (2019, May 15). DNA supercoil. In *Wikipedia, The Free Encyclopedia*. Retrieved 19:40, July 25, 2019, from https://en.Wikipedia.org/w/index.php?title=DNA_supercoil&oldid=897160342

Wikipedia contributors. (2019, July 23). Histone. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:19, July 26, 2019, from https://en.Wikipedia.org/w/index.php?title=Histone&oldid=907472227

Wikipedia contributors. (2019, July 17). Nucleosome. In *Wikipedia, The Free Encyclopedia*. Retrieved 17:17, July 26, 2019, from https://en.Wikipedia.org/w/index.php?title=Nucleosome&oldid=906654745

Wikipedia contributors. (2019, July 26). Human genome. In *Wikipedia, The Free Encyclopedia*. Retrieved 06:12, July 27, 2019, from https://en.Wikipedia.org/w/index.php?title=Human_genome&oldid=908031878

Wikipedia contributors. (2019, July 19). Gene structure. In *Wikipedia, The Free Encyclopedia*. Retrieved 06:16, July 27, 2019, from https://en.Wikipedia.org/w/index.php?title=Gene_structure&oldid=906938498

8.6: Enzymes for Genetic modifications

It is difficult to read newspapers and newsmagazines without encountering the CRISPR-Cas9 gene editing system that has the potential to make gene editing routine in disease diagnosis, treatment, and cure, as well as in genetic modification of organisms to improve their quality and quantity for food and natural product production. In this chapter section, we will explore the mechanism of restriction enzymes that made gene cloning possible as well as the CRISPR-Cas gene editing system.

8.6.1: Restriction Endonucleases

A ***restriction enzyme***, ***restriction endonuclease***, or ***restrictase*** is an enzyme that cleaves DNA into fragments at or near specific recognition sites within molecules known as restriction sites. Restriction enzymes are one class of the broader endonuclease group of enzymes. Restriction enzymes are commonly classified into five types, which differ in their structure and whether they cut their DNA substrate at their recognition site, or if the recognition and cleavage sites are separate from one another. To cut DNA, all restriction enzymes make two incisions, once through each sugar-phosphate backbone (i.e. each strand) of the DNA double helix. Here we will focus on the Type II restriction enzymes that are routinely used in molecular biology and biotechnology applications.

As with other classes of restriction enzymes, Type II Restriction Enzymes occur exclusively in unicellular microbial life forms—mainly bacteria and archaea (prokaryotes)—and are thought to function primarily to protect these cells from viruses and other infectious DNA molecules. Inside a prokaryote, the restriction enzymes selectively cut up *foreign* DNA in a process called ***restriction digestion*****;** meanwhile, host DNA is protected by a modification enzyme (a methyltransferase) that modifies the prokaryotic DNA and blocks cleavage. Together, these two processes form the ***restriction-modification system***.

The first Type II Restriction Enzyme discovered was HindII from the bacterium *Haemophilus influenzae* Rd. The event was described by Hamilton Smith (Figure 7.23) in his Nobel lecture, delivered on 8 December 1978:

‘"In one such experiment we happened to use labeled DNA from phage P22, a bacterial virus I had worked with for several years before coming to Hopkins. To our surprise, we could not recover the foreign DNA from the cells. With Meselson’s recent report in our minds, we immediately suspected that it might be undergoing restriction, and our experience with viscometry told us that this would be a good assay for such an activity. The following day, two viscometers were set up, one containing P22 DNA and the other Haemophilus *DNA. Cell extract was added to each and we began quickly taking measurements. As the experiment progressed, we became increasingly excited as the viscosity of the* Haemophilus *DNA held steady while the P22 DNA viscosity fell. We were confident that we had discovered a new and highly active restriction enzyme. Furthermore, it appeared to require only* Mg^{2+} *as a cofactor, suggesting that it would prove to be a simpler enzyme than that from* E. coli *K or B.*

After several false starts and many tedious hours with our laborious, but sensitive viscometer assay, Wilcox and I succeeded in obtaining a purified preparation of the restriction enzyme. We next used sucrose gradient centrifugation to show that the purified enzyme selectively degraded duplex, but not single-stranded, P22 DNA to fragments averaging around 100 bp in length, while Haemophilus *DNA present in the same reaction mixture was untouched. No free nucleotides were released during the reaction, nor could we detect any nicks in the DNA products. Thus, the enzyme was clearly an endonuclease that produced double-strand breaks and was specific for foreign DNA. Since the final (limit) digestion products of foreign DNA remained large, it seemed to us that cleavage must be site-specific. This proved to be case and we were able to demonstrate it directly by sequencing the termini of the cleavage fragments.’"*

Figure (\PageIndex{35}\): Hamilton Smith and Daniel Nathans at the Nobel Prize press conference, 12 October 1978 (reproduced with permission from Susie Fitzhugh). Original Repository: Alan Mason Chesney Medical Archives, Daniel Nathans Collection. Image from: Pingoud, A., Wilson, G.G., and Wende, W. (2014) Nuc Acids Res 42(12):7489-7527.

Restriction enzymes are named according to the taxonomy of the organism in which they were discovered. The first letter of the enzyme refers to the genus of the organism and the second and third to the species. This is followed by letters and/or numbers identifying the isolate. Roman numerals are used to specify different enzymes from the same organism. For example, the enzyme 'HindIII' was discovered in *Haemophilus influenzae*, serotype d, and is distinct from the HindI and HindII endonucleases also present within this bacterium. The DNA-methyltransferases (MTases) that accompany restriction enzymes are named in the same way, and given the prefix 'M.'. When there is more than one MTase, they are prefixed 'M1.', 'M2.', etc, if they are separate proteins or 'M1~M2.' when they are joined.

Restriction Enzymes that recognize the same DNA sequence, regardless of where they cut, are termed ***'isoschizomers'*** (iso = equal; skhizo = split). ***Isoschizomers*** that cut the same sequence at different positions are further termed ***'neoschizomers'*** (neo = new). ***Isoschizomers*** that cut at the same position are frequently, but not always, evolutionarily drifted versions of the same enzyme (e.g. BamHI and OkrAI). ***Neoschizomers,*** on the other hand, are often evolutionarily unrelated enzymes (e.g.EcoRII and MvaI).

Type II Restriction Enzymes are a conglomeration of many different proteins that, by definition, have the common ability to cleave duplex DNA at a fixed position within, or close to, their recognition sequence. This cleavage generates reproducible DNA fragments, and predictable gel electrophoresis patterns, properties that have made these enzymes invaluable reagents for laboratory DNA manipulation and investigation. Almost all Type II Restriction Enzymes require divalent cations, usually Mg^{2+}, as essential components of their catalytic sites. Ca^{2+}, on the other hand, often acts as an inhibitor of Type II Restriction Enzymes.

The recognition sequences of Type II Restriction Enzymes are ***palindromic***, with two possible types of palindromic sequences. The ***mirror-like palindrome*** is similar to those found in ordinary text, in which a sequence reads the same forward and backward on a single strand of DNA, as in GTAATG. The ***inverted repeat palindrome*** is also a sequence that reads the same forward and backward, but the forward and backward sequences are found in complementary DNA strands (i.e., of double-stranded DNA), as in GTATAC (GTATAC being complementary to CATATG). ***Inverted repeat palindromes*** are more common and have greater biological importance than ***mirror-like palindromes.*** The position of cleavage within the palindromic sequence can vary depending on the enzyme and can produce either single-stranded overhanging sequences (sticky ends) or blunt-ended DNA products. Examples of staggers and blunt end cuts by restriction enzymes are shown in Table 8.6.8 below.

EcoR1	G|AATTC CTTAA|G
Sma1	CCC|GGG GGG|CCC

Table 8.6.8: Staggered and blunt end cut sequences by EcoR1 and Sma1

Methylation can be used by the host to protect its own genome from cleavage. For example, the methylation of the EcoRI recognition sequence by the M.EcoRI methyltransferase (MTase), changes the sequence from GAATTC to GAm6ATTC (m6A = N6-methyladenine). This modification completely protects the sequence from cleavage by EcoRI.

Type II Restriction Enzymes initially bind non-specifically with the DNA and proceed to slide down the DNA scanning for recognition sequences as shown in Figure (\PageIndex{36}\):. Upon binding to the correct palindromic sequence the enzyme associates with the metal cofactor and mediates catalytic cleavage of the DNA using the mechanism of ***strain distortion*** and ***catalysis by approximation***.

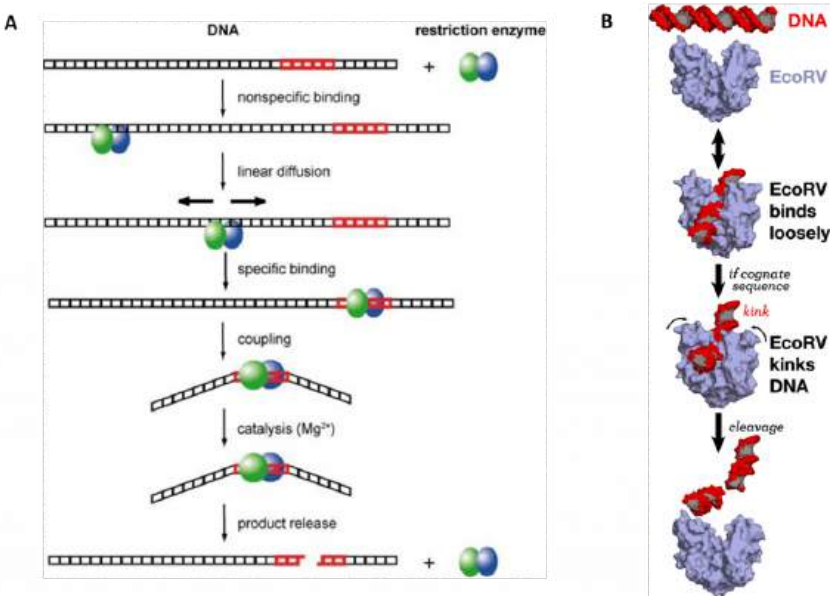


Figure (\PageIndex{36}\): DNA Recognition and Cleavage by Type II Restriction Endonucleases. (A) Pictorial view of an EcoRV dimer scanning nonspecifically along the DNA until a specific binding site is recognized. This causes coupling with the metal cofactor and strain distortion of the DNA. Hydrolysis of the phosphodiester bond is mediated and the DNA cleavage products are released from the enzyme. (B) shows a space-filling model of EcoRV DNA recognition and cleavage. Figure (A) from Pingoud, A., Wilson, G.G., and Wende, W. (2014) Nuc Acids Res 42(12):7489-7527. and Figure (B) from Thomas Splettstoesser

One of the most important questions regarding the catalytic mechanism of a hydrolase is whether hydrolysis involves a covalent intermediate, as is typical for the proteases described previously. This can be decided by analyzing the stereochemical course of the reaction. This was done first for EcoRI, and later for EcoRV. Both enzymes were found to cleave the phosphodiester bond with inversion of the chiral center at the phosphorus, which argues against the formation of a covalent enzyme–DNA intermediate. Thus, it is proposed that cleavage involves the direct nucleophilic attack of the substrate by a water molecule, as shown in Figure (\PageIndex{37}\) below.

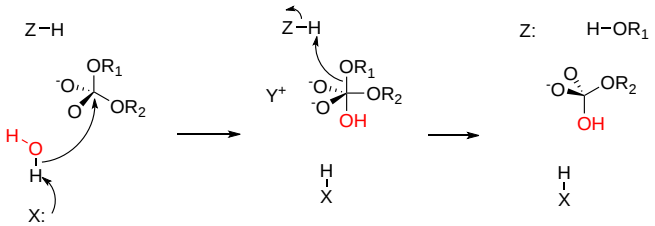


Figure (\PageIndex{37}\): A General Mechanism for DNA Cleavage by EcoRI and EcoRV. An activated water molecule attacks the phosphorous in-line with the phosphodiester bond to be cleaved, which proceeds with an inversion of configuration. X, Y, and Z are a general base, a Lewis acid, and a general acid, respectively. Figure adapted from:

Pingoud, A., Wilson, G.G., and Wende, W. (2014) Nuc Acids Res 42(12):7489-7527.

Type II restriction enzymes typically form a homodimer when binding with DNA, as shown in the crystal structure of ***Bgl*II** in Figure 7.26B. ***Bgl*II** catalyzes phosphodiester bond cleavage at the DNA backbone through a phosphoryl transfer to water. Studies on the mechanism of restriction enzymes have revealed several general features that seem to be true in almost all cases, although the actual mechanism for each enzyme is most likely some variation of this general mechanism (Figure 7.25). This mechanism requires a base to generate the hydroxide ion from water, which will act as the nucleophile and attack the phosphorus in the phosphodiester bond. Also required is a Lewis acid to stabilize the extra negative charge of the pentacoordinate transition state phosphorus, as well as a general acid or metal ion that stabilizes the leaving group (3'-O^-). In some Type II Restriction Enzymes, two divalent metal cofactors are required (such as in EcoRV and BamHI), whereas other enzymes only require one divalent metal cofactor (such as in EcoRI and BglII).

Structural studies of endonucleases have revealed a similar architecture for the active site with the residues following the weak consensus sequence Glu/Asp-$(X)_{9-20}$-Glu/Asp/Ser-X-Lys/Glu. ***Bgl*II's** active site is similar to other endonucleases', following the sequence Asp-$(X)_9$-Glu-X-Gln. In its active site, there sits a divalent metal cation, most likely Mg^{2+}, that interacts with Asp-84, Val-94, a phosphoryl oxygen, and three water molecules. One of these water molecules is able to act as a nucleophile because of its proximity to the scissile phosphoryl group (Figure 7.26A). The nucleophilic water molecule is positioned for attack onto the phosphoryl group by a hydrogen bond with the side chain amide oxygen of Gln-95 and its contact with the metal cation. Interaction with the metal cation effectively lowers its pK_a, promoting the water's nucleophilicity as shown in Panel A of Figure (\PageIndex{38}\) below (from Pingoud, A., Wilson, G.G., and Wende, W. (2014) Nuc Acids Res 42(12):7489-7527). During hydrolysis, the divalent cation can stabilize the 3'-O^- leaving group and coordinate proton abstraction from one of the coordinated water molecules

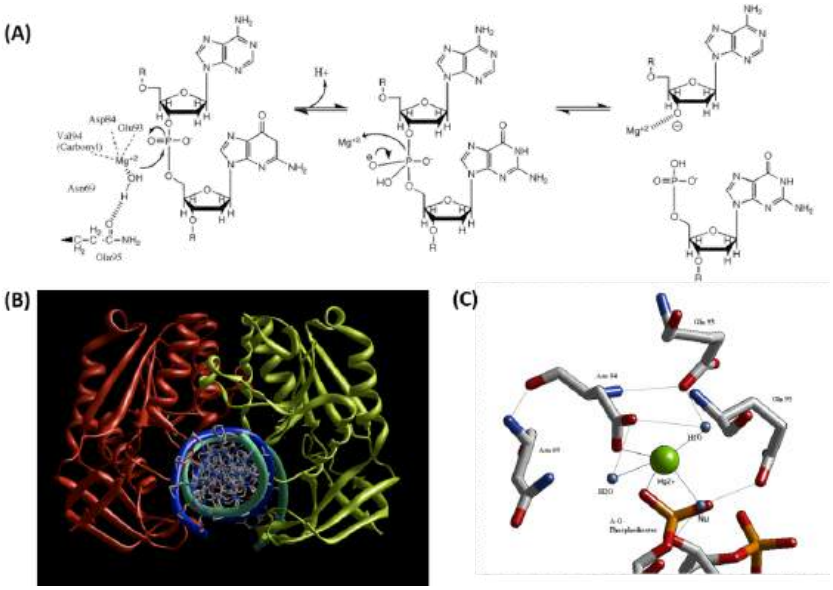


Figure (\PageIndex{38}\): Proposed Reaction Mechanism for the Type II Restriction Endonuclease, *Bgl*II. (A) Schematic diagram of the catalytic mechanism demonstrating the utility of Mg2+ ions and polar amino acid residues within the active site to activate and position a water molecule for nucleophilic attack on the phosphodiester bond of the DNA substrate. (B) Crystal structure of the BglII dimer with double-stranded DNA and (C) Coordination of the Mg2+ cofactor within the active site of the BglII enzyme.

Figures from G Williams

8.6.2: CRISPR-Cas 9

The **CRISPR** (clustered regularly interspaced short palindromic repeats) operon was initially discovered as part of the adaptive immune system of bacteria and archaea, which must defend themselves against viruses (bacteriophages) and unwanted plasmids transferred from both bacteria. It would be ideal for bacteria to recognize previous exposure to viruses and their nucleic acids as the basis of their immunological memory system. Given the tendency of viral DNA to integrate into the host genome (which allows later transcription and translations of the viral genes in the process of new virus production), immunological memory could be based on that viral integrated DNA. Without going into detail, viral DNA can be integrated between two direct repeats in the bacterial genome. DNA from different viruses from previous exposures is also incorporated in the same fashion. One site of

integration is the CRISPR operon. The DNA of the CRISPR operon contains both protein-coding and noncoding regions which are transcribed and processed to form at least three RNA molecules, as shown in Figure (\PageIndex{24}\) below.

- a coding Cas 9 mRNA this is translated to produce the **Cas** 9 (**C**RISPR-**as**sociated protein);
- a noncoding cr-RNA (CRISPR RNA)
- a noncoding tracr-RNA (trans-activating CRISPR RNA)

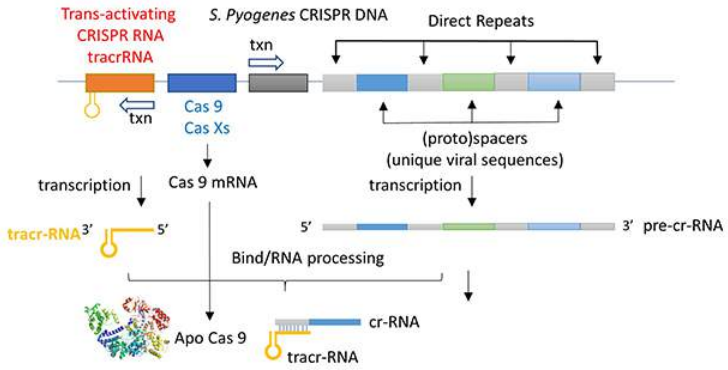


Figure (\PageIndex{24}\): DNA of the CRISPR operon

The two mature noncoding RNAs eventually associate to form a binary complex. When using CRISPR-Cas 9 in eukaryotic gene editing applications, the two noncoding RNAs are covalently combined into one large **s**ynthetic **g**uide RNA (**sg-RNA**), described later in this section. The Cas 9 protein is an endonuclease that cleaves both strands of bound target dsDNA in a blunt-end fashion at specific sequences. This occurs after the DNA binds to two arginines (1333 and 1335) in Cas9 through a short (3-5+ bases) recognition **protospacer adjacent motif** (**PAM**) located three base pairs from the cleavage site. The DNA must also bind in a complementary and specific fashion to the **protein-bound** noncoding cr+tracr-RNAs (or a single sg-RNA molecule for gene editing applications). Binding and cleavage of target DNA would render DNA from an invading bacteriophage inactive.

Basic research into the bacterial CRISPR system has led to revolutionary and explosive applications of this gene editing system in eukaryotes. The hope is that CRISPR technology will give us a precise and incredibly cheap way to do gene therapy in diseased cells and organisms. Given its role in transforming our ability to edit the genome and potentially cure genetically-based diseases, we will offer a detailed explanation of its mechanism.

We have discussed the structure and function of many proteins. Protein enzymes are key to life as they catalyze almost all biological reactions. Most key enzymes are regulated. The activity of Cas 9 must be carefully controlled. Think of the consequences if the enzyme were to cleave promiscuously at off-site targets! This section will help you understand several critical features of this enzyme:

1. How does the enzyme find its correct target site, a 20 nucleotide DNA sequence, and a proximal PAM site, among all the possible alternative sites? Think of how many PAM sequences there must be in the host DNA genome!
2. How can the enzyme be "turned" on when it finds its target site and remain off when free, but more importantly when it is bound off-site?

First, we will discuss the apo- form of the enzyme without bound substrate and RNA.

Apo- and Holo-Cas 9

This section will focus on the Type II-A Cas9 from *Streptococcus pyogenes* (SpyCas9 or SpCas9). Cas 9 is an endonuclease that cleaves both strands of DNA 3 base pairs from a DNA motif, NCC/NGG, called PAM. It has two distinct lobes. The nuclease lobe (NUC), amino acids 1-56 and 718-1368, has two different nuclease domains for the two cleavages. The recognition or receptor lobe (REC), amino acids 94-717, interacts with the RNA molecules. There is also an arginine-rich bridge helix (57-93).

The enzyme has two catalytic nuclease domains:

- HNH-like nuclease domain cleaves the "target" DNA strand, which is complementary to the RNA the confers specificity to the enzyme. The key catalytic residues are His 840 and Asn 854. It also contains a Mg ion;
- Ruv-like domain that cleaves the complementary "non-target" strand with key active site residues Asp 10, Glu 762, Asp 986, and His 983. It also contains a bound Mn ion. The two lobes are separated by two linkers, amino acids 712-717, and an arginine-rich bridge (basic helix - BH), amino acids 628-658.

The overall structure of the apoenyzme (without bound RNA and DNA,pdb id 4cmp) is shown in Figure (\PageIndex{25}\) below, which shows the NUC domain (light blue) with the two catalytic domains (HNH and Ruv), the REC domain (orange) and the BH helix (red).

Figure (\PageIndex{25}\): Apoenyzme Cas9 (without bound RNA and DNA (4cmp)

A close up view showing the two catalytic sites is shown in Figure (\PageIndex{26}\) below.

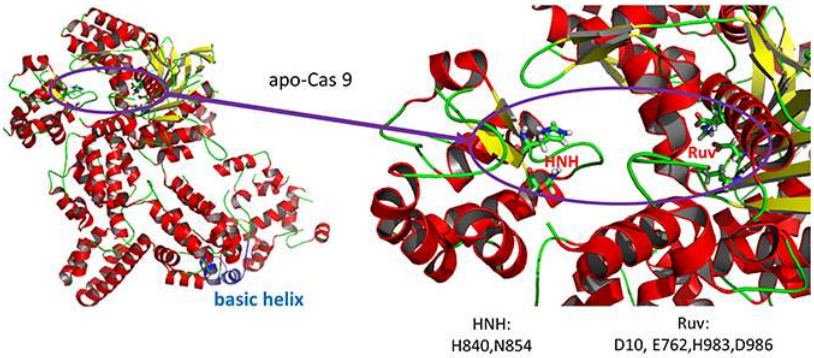


Figure (\PageIndex{26}\): Two catalytic sites in Cas9

Figure 8.6.27 shows an interactive iCn3D model of Streptococcus pyogenes Cas9 in complex with guide RNA and target DNA (4OO8) (long load time). The Cas9 enzyme is shown as a gray transparent surface with an underlying cartoon rendering. The DNA is shown as colored sticks. The RNA is shown as a cyan cartoon.

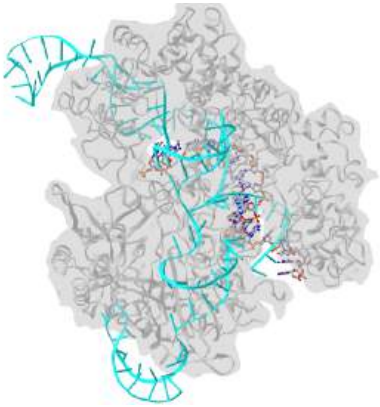


Figure 8.6.27: **Streptococcus pyogenes Cas9 in complex with guide RNA and target DNA (4OO8)**. (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...RjzJBFVt5qRjS7 (long load time)

A comparison of the crystal structure of the apo-Cas 9 and the ternary Cas 9: sgRNA:DNA target strand complex shows a significant conformational change on binding nucleic acids. The structure of the holoenzyme (ternary complex) is shown in Figure (\PageIndex{28}\) below.

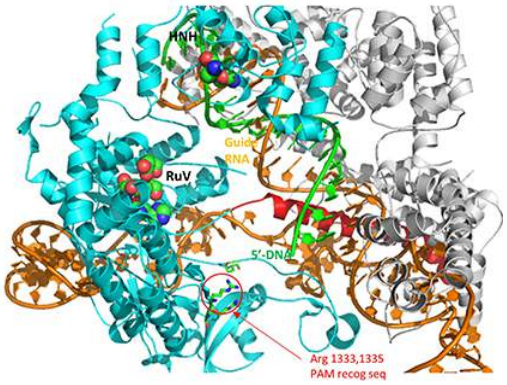


Figure (\PageIndex{28}\): Structure of the holoenzyme Cas9 (ternary complex)with bound guide RNA and DNA

The extent of the conformation change between apo- and holo-Cas 9 enzymes can be seen by examining the distance between D435 and E 944/945 in Figure (\PageIndex{29}\) below. The importance of this change will be described later.

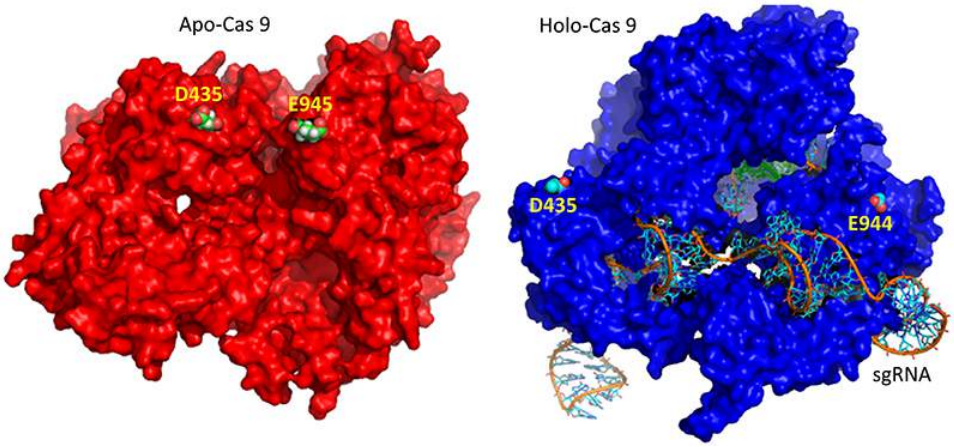


Figure (\PageIndex{29}\): Distance between D435 and E 944/945 going from the apo-Cas9 (left) to the holo-Cas 9 (right) enzyme

Figure (\PageIndex{30}\) below shows the pathway from the transcription of the relevant CRISPR genes (coding and noncoding) to the assembly of the ternary complex and the blunt end cut of the target DNA strand three nucleotides from the PAM sequence.

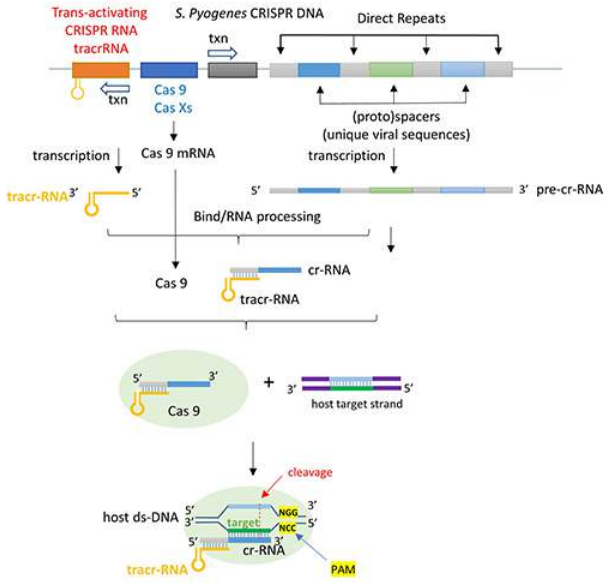


Figure (\PageIndex{30}\): Pathway from the transcription of the relevant CRISPR genes (coding and noncoding) to the assembly of the ternary complex and the blunt end cut of the target DNA strand three nucleotides from the PAM sequence

Figure (\PageIndex{31}\) below shows an expanded view of the ternary complex.

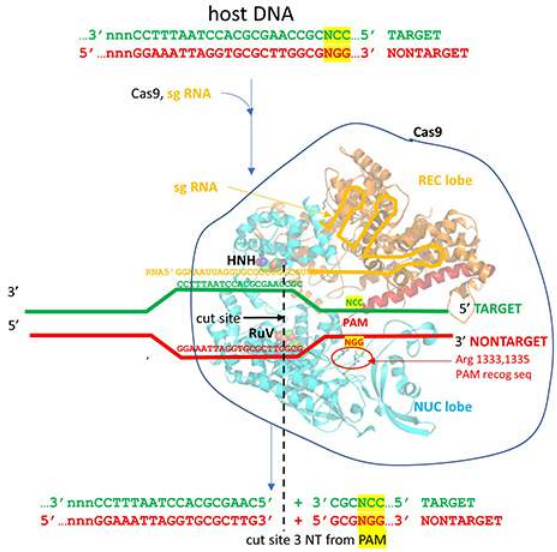


Figure (\PageIndex{31}\): Expanded view of the ternary complex of Cas9 with guide RNA and DNA

Mechanism of DNA binding and cleavage

The above figures do not speak to the mechanism of the binding processes that form the ternary complex. Kinetic and structural studies have been conducted to elucidate the mechanism of binding and cleavage and address the following questions:

- which binds first, the RNA or DNA?
- What are the consequences of the profound conformational changes on the formation of the ternary complex?

The **specificity** of target DNA binding depends both on enzyme:PAM DNA and enzyme:sgRNA (or tracr- and crRNA) interactions. It should seem improbable that the trinucleotide PAM DNA sequence (NGG in *S. pyogenes*), which interacts with a pair of arginines (R 1333, R 1335) through H-bonding, as shown in the images above, and other local sites in Cas 9 would provide the sole or even the majority of the binding interactions. Figure (\PageIndex{32}\) below shows the Args:PAM interaction (pdb code 4un3)

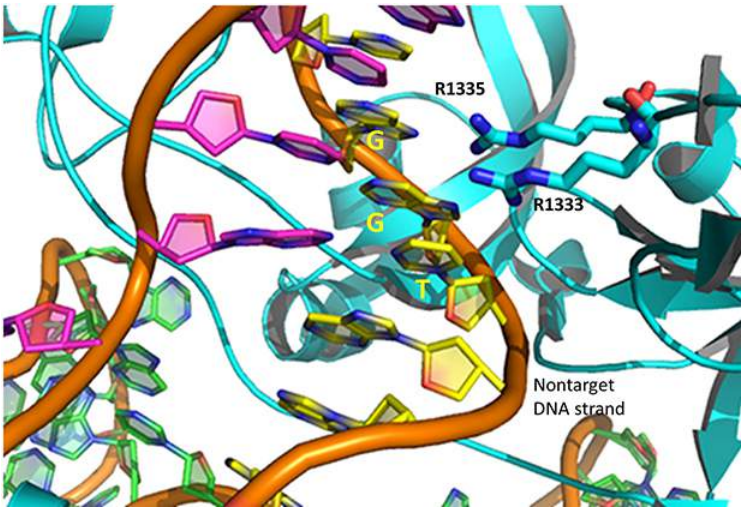


Figure (\PageIndex{32}\): Args:PAM interaction in holo-Cas9 (4un3)

Hence it is most likely that RNA binds first. Indeed, it does with the tracRNA implicated in the recruitment of Cas and the crRNA providing specificity for target DNA binding. The resulting Cas9:RNA binary complex could then search the relevant DNA genome. That would include the DNA of the bacteriophage in viral infection or eukaryotic DNA if the CRISPR DNA operon with the genes for Cas 9 and a sg-RNA was transfected into the eukaryotic cell. After RNA binding, the enzyme would change conformation and allow loose DNA binding through Cas 9: PAM interactions.

Studies have shown that the apo form can also bind DNA, but it does so loosely and indiscriminately. It dissociates quickly and binding is affected by generic polyanions such as the glycosaminoglycan heparin, which indicates its nonspecific nature. Once bound, both off-target and target DNAs would then be surveyed. If a target DNA contained a PAM sequence, the complex would undergo another conformational change to position the HNH and Ruv nuclease catalytic residues and locally unwind the duplex DNA to make the blunt-end cuts.

Cas 9 binding to the PAM site would promote better interaction of the unwound DNA and the bound RNA. If no PAM was present, no catalytically-effective Cas 9:target DNA would form. This prevents off-site cleavage. These allosteric changes and controls are vital to the function of the endonuclease. Here are some findings that support this proposed mechanism:

- the conformation of apo Cas 9 is catalytically inactive;
- on binding RNA to form a binary complex, Cas 9 undergoes a dramatic conformational change, mostly in the REC lobe. However, on binding DNA in a nonspecific fashion, the conformational changes are much smaller. This suggests that most changes in conformation occur before DNA binding. In a way, RNA acts as an allosteric activator of the enzyme (as well as the major source of binding specificity to target DNA). Conformational changes can be determined directly by comparison of crystal structures or spectral techniques such as fluorescence resonance energy transfer (FRET) between two different attached fluorophores.
- Cas 9: RNA interactions lead to ordering of the region of the RNA that interacts with the DNA PAM sequence and adjacent deoxynucleotides (a "seed sequence"), allowing the Cas 9:RNA complex to scan and interact with potential DNA targets with PAM sequences;
- Once a PAM site is found, conformational changes lead to unwinding of the dsDNA, which allows heteroduplex formation between the crRNA and the target DNA strand;
- since Cas 9 recognizes a variety of DNA target sequences (but of course only a specific PAM sequence), the binding of the target sequence depends on the geometry, not the sequence, of the target DNA;
- since binding of off-target DNA to the Cas 9:RNA complex occurs but with very infrequent cleavage, binding and cleavage are very distinct steps;
- on specific DNA binding, the HNH catalytic site moves near to the sessile DNA bond site. Crystal structures show that the active site His is not sufficiently close to facilitate cleavage, suggesting that binding of a second metal ion (see below) may be necessary. Molecular dynamics studies show that the HNH domain is "remarkably plastic".

Figure (\PageIndex{33}\) below show an animation that illustrates the relative conformational changes going from the apo Cas 9 to the binary Cas 9:sgRNA complex to the ternary Cas 9: sgRNA: target DNA complex. The NUC catalytic domain is shown in light blue, the REC (receptor or RNA binding domain) in orange, sgRNA in red, and the target DNA in green. Note again that on

binding RNA to form a binary complex, Cas 9 undergoes a dramatic conformational change, mostly in the REC lobe. The pdb protein sequences shown were aligned using pdbEfold.

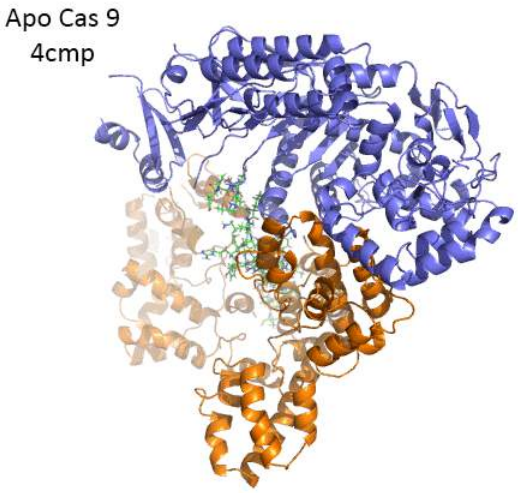


Figure (\PageIndex{33}\): Animation illustrating conformational changes going from the apo Cas 9 to the binary Cas 9:sgRNA complex to the ternary Cas 9: sgRNA: target DNA complex

A potential abbreviated catalytic mechanism for the Ruv nuclease domain is shown in Figure (\PageIndex{34}\) below. The red arrows indicate the second set of electron movements. His 983 acts as a general base to abstract a proton from the water making it a more potent nucleophile. An intermediate trigonal bipyramidal phospho-intermediate is formed, which, along with the preceding transition state, is stabilized by the proximal Mg^{2+} ion (an example of electrostatic or metal ion catalysis). The magnesium is positioned through its interaction with negatively charged carboxyl groups of Asp 10, Glu 762, and Asp 986.

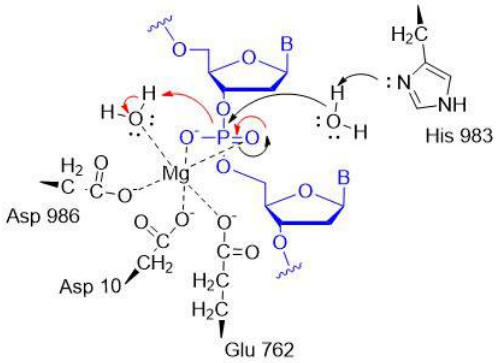


Figure (\PageIndex{33}\): Abbreviated catalytic mechanism for the Ruv nuclease domain of Cas9

A second metal ion might be recruited to the Ruv site to further facilitate the cleavage of the DNA. The HNH catalytic site has a structure (beta-beta-alpha) and conserved His in common with a class of nucleases that require one metal ion. In contrast, the Ruv catalytic site does not have this common secondary structural motif and has a critical histidine, both common features found in endonucleases that use two metal ions.

CRISPR and Eukaryotic Gene Editing

How could blunt-end cutting of both strands of DNA by Cas 9 lead to the holy grail of specific eukaryotic gene editing with no off-site effects? Cutting the DNA genome seems like a bad idea. It is potentially so bad that a myriad of DNA repair mechanisms has evolved to fix the cut. These include homologous recombination. If corrective DNA is supplied as well as the components of the CRISPR system, a cell could effectively add the corrective DNA after the double-stranded cut and repair a deleterious mutation. Consult a molecular biology textbook for more insight into homologous recombination.

Mutations in the PAM sequence prevent Cas9 nuclease activity. Hence the NGG PAM sequence is vital for the interactions and activities described above. This would seem to limit the utility of CRISPR-Cas 9 in eukaryotic gene editing until one realizes that

the GG dinucleotide has a 5.2% frequency of occurrence in the human genome, which corresponds to over 160 million occurrences. Even then it might not occur in a desired gene target. Cas 9 nuclease from other bacteria extends the range of activity of the CRISPR/Cas system as they interact with other PAM sequences (NNAGAA and NGGNG for *S. thermophilus* and NGGNG for *N. meningidtis*). Likewise, mutations in the *S. pyogenes* PAM (NGG) have been made as well. A D1135E mutation retains but increases the specificity for the normal NGG PAM site. D1135V, R1335Q, and T1337R mutations alter the optimal PAM recognition site to NGAN or NGNG.

CRISPR editing can be easily used to knock out specific genes. In addition, if cells are transfected with a plasmid with many target sequences, the system can be used to edit multiple genes in one experiment. This would be very useful in studies of diseases linked to multiple genes. Since the cost of CRISPR reagents (plasmids, RNAs) is so inexpensive, and the specificity of editing is so high, the great excitement about CRISPR use for gene editing in human disease and for modification of plant and fungal genomes is warranted.

Other systems have been developed to specifically bind to a target DNA sequence and then cleave it. They typically contain a protein that binds to a specific DNA target and an associated endonuclease that cleaves within the target DNA site. Typical prokaryotic restriction enzymes bind to and cut at a specific nucleotide sequence (for example Eco R1 cleaves at G/AATTC palindromic sequences) to form sticky ends. The protein itself binds to this DNA recognition site. Other examples are based on the structure of known transcription factors. Libraries of genetically engineered proteins with Zn finger DNA binding domains (designed for specific DNA target sequences) fused to endonucleases have been created for this purpose. Other examples are proteins called TALENs (transcription activator-like effector nucleases). These are fusion proteins containing a TAL effector DNA-binding domain and a nuclease. In each of these cases, a 3D-folded protein is the specific target DNA recognition molecule. Think how much easier it is to make in effect a 1D-DNA recognition element, a simple linear RNA sequence, which would adopt the correct 3D structure on the binding of its complementary target.

One major problem in the use of CRISPR for gene editing must be solved: how to get the CRISPR components in the correct cells in an organism. In effect, it's the same problem faced by small drug designers only the components are much larger. *Ex vivo* applications, when diseased cells are removed from the body, repaired by CRISPR, and then reinjected, are likely to have more success. In these cases, electroporation would allow the uptake of Cas 9 and the sg-RNA. I*n vivo* therapy has included the use of adeno-associated viruses in which genes for Cas 9 and sg RNA could be encapsulated. This technique, used for other gene delivery systems, has the advantage of being tolerated immunologically. However, this system allows for continual gene expression which is undesirable for gene editing. After an initial "fix" of a mutant gene, continued expression of the CRISPR-Cas 9 genes would increase the chances for off-target cutting. A more recent approach is to deliver the mRNA in artificial lipid nanoparticles that can be taken into cells. Once free and translated into protein and sg RNA inside the cell, gene editing has a chance to occur before the RNA and protein are degraded.

8.6.2.0.1: 7.4 References:

1. Wikipedia contributors. (2020, April 21). Nucleophile. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:39, April 26, 2020, from en.Wikipedia.org/w/index.php?title=Nucleophile&oldid=952368939
2. Oregon Institute of Technology (2019) Organic Chemistry II (Lund). In Libretexts. Retrieved 10:58 am, April 27, 2020 from: https://chem.libretexts.org/Courses/Oregon_Institute_of_Technology/OIT%3A_CHE_332_--_Organic_Chemistry_II_(Lund)
3. Wikipedia contributors. (2020, April 12). Bond cleavage. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:15, April 27, 2020, from en.Wikipedia.org/w/index.php?title=Bond_cleavage&oldid=950494652
4. Wikipedia contributors. (2020, February 24). Arrow pushing. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:25, April 27, 2020, from en.Wikipedia.org/w/index.php?title=Arrow_pushing&oldid=942438883
5. Wikipedia contributors. (2020, April 16). Acid dissociation constant. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:48, April 27, 2020, from en.Wikipedia.org/w/index.phptitle=Acid_dissociation_constant&oldid=951313744
6. Farmer, S., Reusch, W., Alexander, E., and Rahim, A. (2016) Organic Chemistry. Libretexts. Available at: https://chem.libretexts.org/Core/Organic_Chemistry
7. Ball, et al. (2016) MAP: The Basics of GOB Chemistry. Libretexts. Available at:https://chem.libretexts.org/Textbook_Maps/Introductory_Chemistry_Textbook_Maps/Map%3A_The_Basics_of_GOB_Chemistry_(Ball_et_al.)/14%3A_Organic_Compounds_of_Oxygen/14.10%3A_Properties_of_Aldehydes_and_Ketones
8. McMurray (2017) MAP: Organic Chemistry. Libretexts. Available at:https://chem.libretexts.org/Textbook_Maps/Organic_Chemistry_Textbook_Maps/Map%3A_Organic_Chemistry_(McMurry)
9. Soderburg (2015) Map: Organic Chemistry with a Biological Emphasis. Libretexts. Available at:https://chem.libretexts.org/Textbook_Maps/Organic_Chemistry_Textbook_Maps/Map%3A_Organic_Chemistry_With_a_Bio

logical_Emphasis_(Soderberg)
10. Ophardt, C. (2013) *Biological Chemistry.* Libretexts. Available at:https://chem.libretexts.org/Core/Biological_Chemistry/Proteins/Case_Studies%3A_Proteins/Permanent_Hair_Wave
11. Soderberg, T. (2016) *Organic Chemistry with a Biological Emphasis*. Libretexts. Available at:https://chem.libretexts.org/Textbook_Maps/Organic_Chemistry_Textbook_Maps/Map%3A_Organic_Chemistry_with_a_Biological_Emphasis_(Soderberg)
12. Ball, et al. (2016) MAP: The Basics of General, Organic, and Biological Chemistry. Libretexts. Available at:https://chem.libretexts.org/Textbook_Maps/Introductory_Chemistry_Textbook_Maps/Map%3A_The_Basics_of_GOB_Chemistry_(Ball_et_al.)
13. Clark, J. (2017) Organic Chemistry. Libretexts. Available at:https://chem.libretexts.org/Core/Organic_Chemistry/Amides/Reactivity_of_Amides/Polyamides
14. Wikipedia contributors. (2018, December 28). Metabolism. In *Wikipedia, The Free Encyclopedia*. Retrieved 19:28, December 29, 2018, fromen.Wikipedia.org/w/index.php?title=Metabolism&oldid=875751739
15. Ball, Hill, and Scott. (2012) Enzyme Activity, section 18.7 from the book Introduction to Chemistry: General, Organic and Biological (v1.0) retrieved on Dec 31, 2018 fromhttps://2012books.lardbucket.org/books/introduction-to-chemistry-general-organic-and-biological/s21-07-enzyme-activity.html
16. Wikipedia contributors. (2018, November 29). Mechanism of action. In *Wikipedia, The Free Encyclopedia*. Retrieved 05:00, January 1, 2019, fromen.Wikipedia.org/w/index.php?title=Mechanism_of_action&oldid=871201209
17. Mótyán, J.A., Tóth, F., and Tőzsér, J. (2013) Research Applications of Proteolytic Enzymes in Molecular Biology. *Biomolecules 3*(4), 923-942; https://doi.org/10.3390/biom3040923
18. Wikipedia contributors. (2020, April 11). Adenylate kinase. In *Wikipedia, The Free Encyclopedia*. Retrieved 19:28, May 4, 2020, from en.Wikipedia.org/w/index.php?title=Adenylate_kinase&oldid=950311736
19. Ahern, K., Rajagopal, I., and Tan, T. (2019) Biochemistry Free and Easy. Available at Oregon State University (http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy) and Libretexts (https://bio.libretexts.org/Bookshelves/Biochemistry/Book%3A_Biochemistry_Free_For_All_(Ahern%2C_Rajagopal%2C_and_Tan)/04%3A_Catalysis/4.03%3A_Mechanisms_of_Catalysis)
20. Wikipedia contributors. (2020, April 16). Serine protease. In *Wikipedia, The Free Encyclopedia*. Retrieved 14:32, May 6, 2020, from en.Wikipedia.org/w/index.php?title=Serine_protease&oldid=951309456
21. Wikipedia contributors. (2020, April 16). Restriction enzyme. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:12, May 16, 2020, from en.Wikipedia.org/w/index.php?title=Restriction_enzyme&oldid=951351229
22. Pingoud, A., Wilson, G.G., and Wende, W. (2014) Type II restriction endonucleases - a historical perspective and more. *Nuc Acids Res* 42(12)7489-7527. Retrieved from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4081073/pdf/gku447.pdf
23. Wikipedia contributors. (2019, July 25). BglII. In *Wikipedia, The Free Encyclopedia*. Retrieved 20:48, May 16, 2020, from en.Wikipedia.org/w/index.php?title=BglII&oldid=907885716
24. De la Peña, M, GarcÍa-Robles, I., and Cervera, A. (2017) The Hammerhead Ribozyme: A Long History for a Short RNA. *Molecules* 22(1):78. Retrieved from: https://www.mdpi.com/1420-3049/22/1/78/htm
25. Jakubowski, H. (2019) Biochemistry Online. *Libretexts.* Available at: https://bio.libretexts.org/Bookshelves/Biochemistry/Book%3A_Biochemistry_Online_(Jakubowski)

CHAPTER OVERVIEW

9: Investigating DNA

9.1: DNA Isolation, Sequencing, and Synthesis
9.2: Bioinformatics
9.3: Cloning and Recombinant Expression
9.4: DNA Microarrays
9.5: In Situ Hybridization
9.6: References

9.1: DNA Isolation, Sequencing, and Synthesis

9.1.1: Genomic and complementary DNA

The ability to sequence the DNA of an organism has revolutionized our understanding of biology and evolution. DNA can be isolated directly from living, dead, and even extinct species and the "primary" sequence of A, G, C, and T bases in the molecule determine. We can read (sequence), write (synthesize), and edit (mutate) DNA at will. Before we explore how to purify, sequence (read) and synthesize (write) DNA, it's important to differentiate two different types of DNAs, genomic and **complementary DNA** (cDNA), which is made by reverse transcribing messenger RNA into DNA. Since mRNA has no nucleotides encoded by introns, c-DNA provides just the coding sequences for protein.

Genomic deoxyribonucleic acid (gDNA) is chromosomal DNA, which does not include the extra-chromosomal DNA found in the mitochondria of eukaryotes or plasmids in bacteria (plasmids will be discussed in more detail in section 5.3 during the discussion of gene cloning and expression). Most organisms have the same genomic DNA in every cell (one exception is the genomic DNA for antibodies in B cells and T cells receptors in T-cells that have been altered as the cells become more terminally differentiated. It is also important to remember that only certain genes are active (expressed) in each cell. The subset of expressed genes is specific to a given differentiated cell and allows for the expression of specific cell functions. Liver cells, for example, don't express the gene for the protein opsin, which is expressed in retinal cells and is required for vision.

The genome of an organism (encoded by the genomic DNA) is the (biological) information of heredity that is passed from one generation of organism to the next. That genome is transcribed to produce various RNAs, which are necessary for the function of the organism. Precursor mRNA (pre-mRNA) is transcribed by RNA polymerase II in the nucleus. pre-mRNA is then processed by splicing to remove introns, leaving the exons in the mature messenger RNA (mRNA). Additional processing includes the addition of a 5' cap and a poly(A) tail to the pre-mRNA. The mature mRNA may then be transported to the cytosol and translated by the ribosome into a protein. Other types of RNA include ribosomal RNA (rRNA) and transfer RNA (tRNA). These types are transcribed by RNA polymerase II and RNA polymerase III, respectively, and are essential for protein synthesis. However, 5s rRNA is the only rRNA that is transcribed by RNA Polymerase III. ***cDNA is derived from mRNA, so it contains only exons but no introns.***

Figure 9.1.1 shows the flow of information stored in eukaryotic DNA and its eventual expression in mRNA.

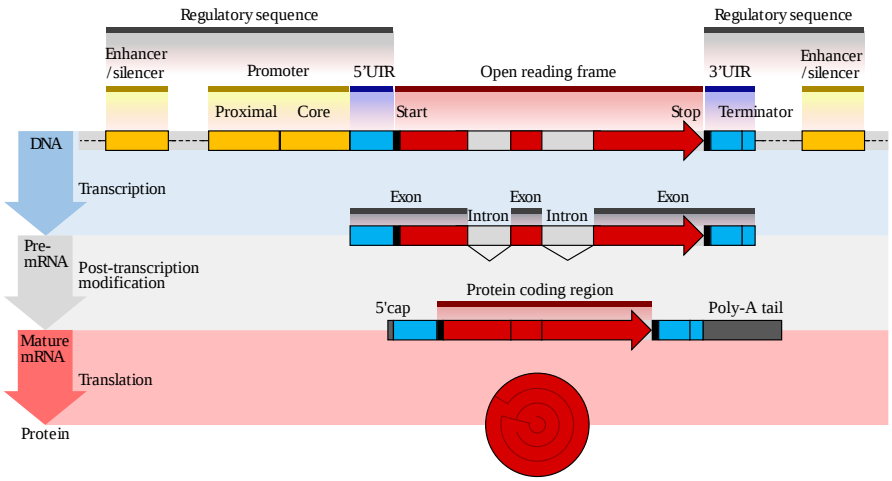


Figure 9.1.1: Eukaryotic gene structure. Shafee T, Lowe R (2017). *WikiJournal of Medicine* **4** (1). DOI:10.15347/wjm/2017.002. ISSN 20024436.Wikimedia Commons

Red indicates coding exons, which are separated by gray introns. At the beginning and end of a gene sequence (encoded by 3 exons in the figure below) are 5' and 3' **untranslated regions (UTRs)** which are also transcribed and are represented in the mature mRNA. Also, there are potential regulatory sequences, which are not transcribed (yellow) on both sides of the transcribable part of the gene. The 5'-end **promoter** is where transcription factors and RNA polymerase assemble before the start of transcriptions. In addition, there are regulatory **enhancers** and **silencers** more distal from the gene sequences. An open reading frame (ORF) is a region of the DNA that can be decoded into a mRNA and doesn't have a stop signal (codon) in it that would prematurely terminate transcription.

In contrast, **complementary DNA** (**cDNA**) is synthesized from a single-stranded RNA (e.g., messenger RNA (mRNA) or microRNA) template in a reaction catalyzed by the enzyme reverse transcriptase. Reverse transcriptase is an enzyme found in retroviruses such as HIV that have RNA as their core genetic material. Upon entering the host cell the RNA is reverse-transcribed to produce a copy of cDNA that can then integrate into the host's genomic DNA. In biotechnology, reverse transcriptase is often used to create cDNA from the mRNA expressed in specific cells or tissues. In this way, the eukaryotic genes can be cloned without any introns housed in the structure. This is especially useful if the goal is to express the protein in a prokaryotic (bacterial) host. Recall that bacterial DNA does not contain any intron sequences within its chromosomal DNA. Thus, if you are using a prokaryotic system to express eukaryotic proteins, you must use cDNA, as the prokaryotic system will not be able to remove intron sequences following gene transcription. The term *cDNA* is also used, typically in a bioinformatics context, to refer to an mRNA transcript's sequence expressed as DNA bases (GCAT) rather than RNA bases (GCAU).

The gene organization of prokaryotes is different in that they don't have introns. In addition, some genes for a common pathway for example are continuous in the DNA. These stretches of DNA are called operons. Transcription from an operon produces a polycistronic RNA transcript. The words cis and are used in chemistry to describe R groups on the same size (cis) or opposite sides (trans) of a double bond. In DNA, cis-elements are in a single DNA section while trans-elements usually refer to proteins (away from the gene) binding to the DNA. Hence the term polycistronic for bacterial operons (with multiple genes sequentially arranged in the DNA sequence). Figure 9.1.2 shows the organization of prokaryotic gene structure.

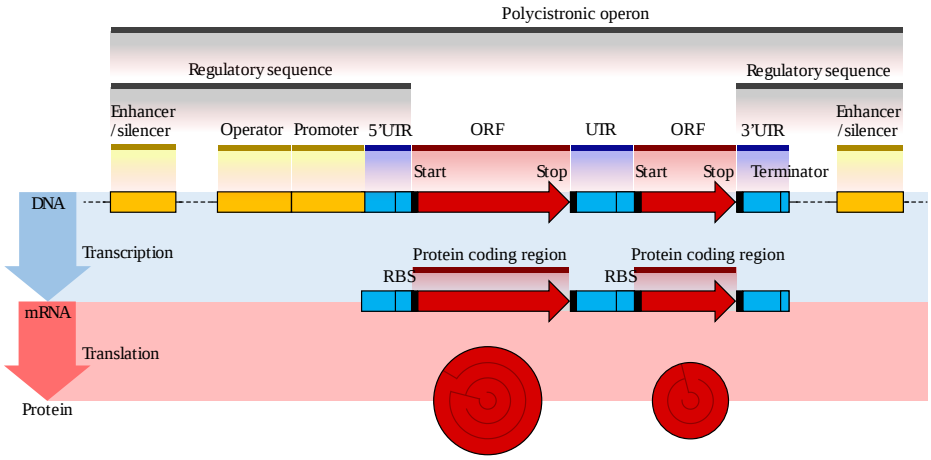


Figure 9.1.2: Prokaryotic gene structure. Shafee T, Lowe R (2017). "Eukaryotic and prokaryotic gene structure". *WikiJournal of Medicine* **4** (1). DOI:10.15347/wjm/2017.002. ISSN 20024436.Wikimedia Commons

9.1.2: DNA Extraction/Purification

The first isolation of DNA was done in 1869 by Friedrich Miescher. Now purification kits are available from multiple manufacturers.

DNA can be isolated from whole tissue or cell cultures. Let's consider just DNA extraction from cells grown in the lab. Cells are first collected by centrifugation and then treated with detergents like sodium dodecyl sulfate to lyse the cell membranes. Proteases and DNAase-free RNAase can be added to digest proteins and RNA, respectively.

Methods involving phenol/chloroform extractions:

In older methods, a mixture of phenol and chloroform or phenol/chloroform/isoamyl alcohol is used to extract DNA from the solution. Students who have performed liquid/liquid extractions in chemistry lab courses should recognize that this will form a biphasic mixture with water. Nonpolar substances like lipids and cellular debris will partition into the nonpolar phase (chloroform/phenol) or in the interface between them (as suspended insoluble material). Chloroform is very dense as it contains a chlorine atom). Phenol is somewhat soluble in water (8 g/100 g water) but very soluble in chloroform. In mixing during the extraction, the dissolved phenol alters the properties of water sufficiently enough to push the delicate equilibrium of proteins from the native to the denatured states, which aggregate and precipitate. On settling, DNA will remain in the aqueous phase. The use of chloroform/phenol in DNA extractions has a potential problem. Phenol (hydroxylated benzene) can lose one electron from the oxygen atom forming a free radical, which can be stabilized by resonance with pi electrons in the aromatic ring. Free radicals can damage DNA, so most new methods of purification do **not** use phenol/chloroform extractions.

Most methods involve precipitating the extracted DNA at some point in the purification process using cold ice-cold ethanol or isopropanol. DNA is to a first approximation a long polyanion so it would be very difficult to purify "naked" DNA from solution since the extensive negative charges on the DNA would prevent aggregate and precipitate formation. This is not a problem if the ionic strength of the medium is sufficiently high so bound positively charged counter-ions can shield the negative charges from each other, allowing precipitation.

Methods involving adsorption chromatography using silica gel: Nucleic acids bind or adsorb to a solid phase (silica or other) depending on the pH and the salt concentration of the buffer. If small amounts are needed (such as for isolating recombinant plasmid from bacteria), small spin columns are employed. This method relies on the fact that nucleic acid will bind to the solid phase of silica gel under certain conditions and then be released when those conditions are altered. These features are illustrated in Figure 9.1.3.

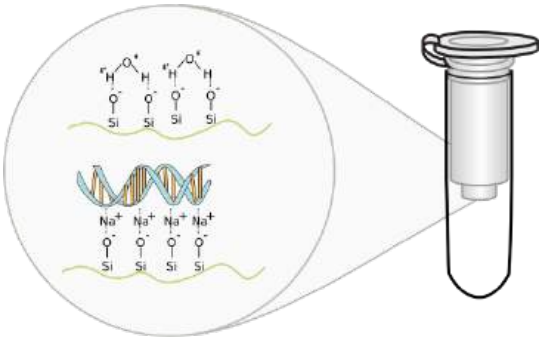


Figure 9.1.3: DNA-Silica gel interactions. *Image by Squidonius*

The solution containing DNA is applied to a small spin column containing silica gel and placed in a mini-centrifuge. On spinning, the nucleic acid will bind to the silica gel membrane as the solution passes through. After multiple "spin" washes to remove non-specific cellular components from the column, the DNA is eluted with a low salt elution buffer (or simply water). Unlike RNA which degrades very quickly, DNA is quite stable and can be stored for long periods at -20^{o}C.

Many student readers have used silica gel chromatography or silica gel thin-layer chromatography to separate and analyze organic mixtures. These techniques are usually performed in a mixture of organic solvents (like hexane/ethyl acetate for example). The use of silica gel to purify DNA from an aqueous solution might seem strange so we will briefly explore how DNA binds.

In silica gel, each silicon atom is tetrahedrally bonded (sp^3 hybridization) to four oxygen atoms, with each oxygen atom covalently linked to two silicon atoms. At the surface of the particle, the oxygen atoms are capped with H atoms, so the entire surface contains a "sea" of OHs, and therefore hydrogen bond donors and acceptors. At lower pHs, some could ionize to form O^- ions. If salt concentrations are high enough, the percentage of O^- ions increases, since they are stabilized by the cations in the salt (shifting the equilibrium to the ionized state).

DNA, a long negatively charged anion, can bind to the silica surface using two types of noncovalent interaction. It can form hydrogen bonds with the silica gel surface OHs. In addition, it can interact with the surface through ion-ion interactions mediated by bridging cations (like Na^+ from the high-concentration salt solution) as illustrated in the figure above. The binding solutions used in the spin column adsorption steps have high concentrations of chaotropic salts that disrupt water structure and hydrogen bonding. The salts also denature proteins and in effect dehydrate the DNA. Some chaotropic salts used include sodium iodide, sodium perchlorate, guanidinium thiocyanate, and guanidinium chloride. The sodium or guanidinium acts as bridging cations, allowing the adsorption of the negative DNA to negative charges on the silica gel surface. Sodium acetate and Tris-HCl are included to buffer pH from 6-7. Now it becomes easy to understand how pure water or low salt concentration solutions elute the bound DNA after extensive column washing since pure water or low salt solutions would strip the bound intermediary cations from the silica column.

After isolation, the DNA is dissolved in a slightly alkaline buffer, usually in a Tris-EDTA buffer, or in ultra-pure water. EDTA binds divalent cations like Ca^{2+}, which activate nucleases. Modifications made to these standard techniques are often done if the tissue being used is difficult to break down, if contaminants persist in the lysis solution that inhibit further reactions, or if the sample is extremely minimal, as is often the case in forensic investigations. In addition, different commercial kits will be tailored for the isolation of larger genomic DNA or smaller plasmid DNA.

The purity of a DNA preparation is usually determined by measuring the absorbance of the solution at 230, 260 (peak absorbance nucleic acids), and 280 nm (peak absorbance proteins), often using an instrument that requires a tiny droplet of solution. Figure (\PageIndex{4}\) below shows the relative absorbance spectra of proteins and nucleic acids.

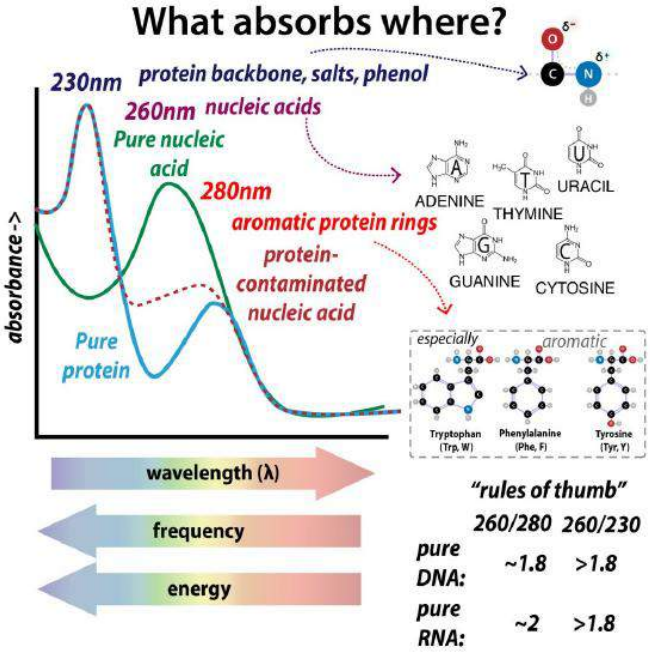


Figure (\PageIndex{4}\) relative absorbance spectra of proteins and nucleic acids. Brianna Bibel. The Bumbling Chemist https://thebumblingbiochemist.com/36...purity-ratios/

The A_{260}/A_{280} gives a measure of protein contamination, with a value around 1.8 indicating "pure" DNA. The A_{260}/A_{230} gives information on protein and other solution contamination. A pure solution of DNA that has an A_{260} = 1.0 has a concentration of 50 µg/mL = 50 ng/µL.

When the temperature of a dsDNA solution is increased, in some range of temperature the A_{260} increases by about 37%. This is called the hyperchromic effect which occurs when the bases in DNA unstack on denaturation of dsDNA to single-stranded DNA as the intrastrand hydrogen bonds break. Figure (\PageIndex{5}\) below shows a graph of A_{260} vs temperature. The midpoint of the cooperative change in the absorbance signifies that the population of DNA molecules is 50% denatured.

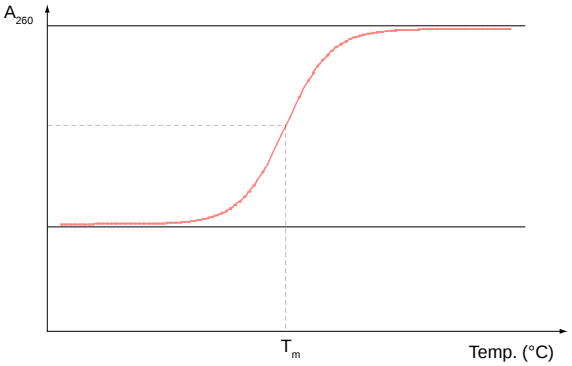


Figure (\PageIndex{5}\): Thermal denaturation curve of dsDNA. Fdardel - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/inde...?curid=5949679

Given the cooperative of unfolding, it is less likely that the population consists of individual molecules that are 50% denatured (i.e. a single molecule being 50% double-stranded and 50% single-stranded). Accordingly, a lower concentration (33 µg/mL) of fully single-stranded DNA gives an A_{260} = 1.

9.1.2.1: DNA Sequencing Techniques

DNA sequencing is the process of determining the nucleic acid sequence – the order of nucleotides in DNA. It includes any method or technology that is used to determine the order of the four bases: adenine, guanine, cytosine, and thymine. The advent of rapid DNA sequencing methods has greatly accelerated biological and medical research and discovery.

Knowledge of DNA sequences has become indispensable for basic biological research and in numerous applied fields such as medical diagnosis, biotechnology, forensic biology, virology, and biological systematics. Comparing healthy and mutated DNA sequences can diagnose different diseases including various cancers, characterize antibody repertoire, and can be used to guide patient treatment. Having a quick way to sequence DNA allows for faster and more individualized medical care to be administered, and for more organisms to be identified and cataloged.

The rapid speed of sequencing attained with modern DNA sequencing technology has been instrumental in the sequencing of complete DNA sequences, or genomes, of numerous types and species of life, including the human genome and other complete DNA sequences of many animal, plant, and microbial species.

The first DNA sequences were obtained in the early 1970s by academic researchers using laborious methods based on two-dimensional chromatography. Following the development of fluorescence-based sequencing methods with a DNA sequencer, DNA sequencing has become easier and orders of magnitude faster.

The canonical structure of DNA has four bases: thymine (T), adenine (A), cytosine (C), and guanine (G). DNA sequencing is the determination of the physical order of these bases in a molecule of DNA. However, DNA bases are often modified by epigenetic processes to control gene expression. Thus, many other modified bases may be present in a DNA molecule than the standard four bases. For example, in some viruses (specifically, bacteriophages), cytosine may be replaced by hydroxymethyl- or hydroxymethylglucose cytosine. In eukaryotic DNA, variant bases with methyl groups or phosphosulfate may be found as shown in Figure (\PageIndex{6}\) below. Depending on the sequencing technique, a particular modification, e.g., the 5mC (5 - methylcytosine) common in humans, may or may not be detected.

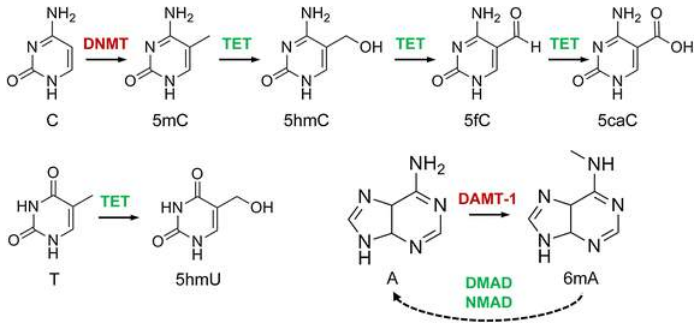


Figure (\PageIndex{6}\): **DNA modifications with epigenetic regulatory functions and their interdependencies.** Cytosine (C) is methylated to 5-methylcytosine (5mC) by DNA methyltransferases (DNMT) and then further oxidized to 5hmC, 5fC, and 5caC by Tet dioxygenases. 5-Hydroxyuracil (5hmU) is produced by Tet-catalysed oxidation of thymine (T). N6-methyladenine (6mA) is likely catalyzed by DNA N6 adenine methyltransferases (DAMT-1 in *C. elegans*), even though the biochemical activity of these enzymes remains to be characterized. The Tet-like ALKB enzymes NMAD (N6-methyl adenine demethylase 1) and DMAD (DNA 6mA demethylase) are involved in 6mA demethylation in *C. elegans* and *Drosophila*, respectively, possibly by using a conserved dioxygenase mechanism. *Image by Breiling and Lyko (2015) Epigenetics and Chromatin 8:24*

9.1.2.1.1: Early DNA sequencing methods

The first method for determining DNA sequences involved a location-specific primer extension strategy established by Ray Wu at Cornell University in 1970. DNA polymerase catalysis and specific nucleotide labeling, both of which figure prominently in current sequencing schemes, were used to sequence the cohesive ends of lambda phage DNA. Between 1970 and 1973, Wu, R Padmanabhan, and colleagues demonstrated that this method can be employed to determine any DNA sequence using synthetic location-specific primers. Frederick Sanger then adopted this primer-extension strategy to develop more rapid DNA sequencing methods at the MRC Centre, Cambridge, UK, and published a method for "DNA sequencing with chain-terminating inhibitors" in 1977. Walter Gilbert and Allan Maxam at Harvard also developed sequencing methods, including one for "DNA sequencing by chemical degradation". In 1973, Gilbert and Maxam reported the sequence of 24 base pairs using a method known as wandering-spot analysis. Advancements in sequencing were aided by the concurrent development of recombinant DNA technology, allowing DNA samples to be isolated from sources other than viruses.

Maxam-Gilbert sequencing requires radioactive labeling at one 5' end of the DNA and purification of the DNA fragment to be sequenced. Chemical treatment then generates breaks at a small proportion of one or two of the four nucleotide bases in each of the four reactions (G, A+G, C, C+T). The concentration of the modifying chemicals is controlled to introduce on average one modification per DNA molecule. Thus a series of labeled fragments is generated, from the radiolabeled end to the first "cut" site in each molecule. The fragments in the four reactions are electrophoresed side by side in denaturing acrylamide gels for size separation. To visualize the fragments, the gel is exposed to X-ray film for autoradiography, yielding a series of dark bands each corresponding to a radiolabeled DNA fragment, from which the sequence may be inferred.

The technical aspects of Maxam-Gilbert sequencing caused it to go out of favor once the Sanger sequencing method had been well established, as described below.

9.1.2.1.2: Sanger Sequencing Method

The chain-termination method developed by Frederick Sanger and coworkers in 1977 soon became the method of choice, owing to its relative ease and reliability. When invented, the chain-terminator method used fewer toxic chemicals and lower amounts of radioactivity than the Maxam-Gilbert method. Because of its comparative ease, the Sanger method was soon automated and was the method used in the first generation of DNA sequencers.

The classical chain-termination method requires a single-stranded DNA template, a DNA primer, a DNA polymerase, normal deoxynucleotide triphosphates (dNTPs), and modified di-deoxynucleotide triphosphates (ddNTPs), the latter of which terminate DNA strand elongation. These chain-terminating nucleotides lack a 3'-OH group required for the formation of a phosphodiester bond between two nucleotides, causing DNA polymerase to cease the extension of DNA when a modified ddNTP is incorporated. The ddNTPs may be radioactively or fluorescently labeled for detection in automated sequencing machines.

The DNA sample is divided into four separate sequencing reactions, containing all four of the standard deoxynucleotides (dATP, dGTP, dCTP, and dTTP) and the DNA polymerase. To each reaction is added only one of the four dideoxynucleotides (ddATP, ddGTP, ddCTP, or ddTTP), while the other added nucleotides are ordinary ones as shown in (\PageIndex{7}\) below.

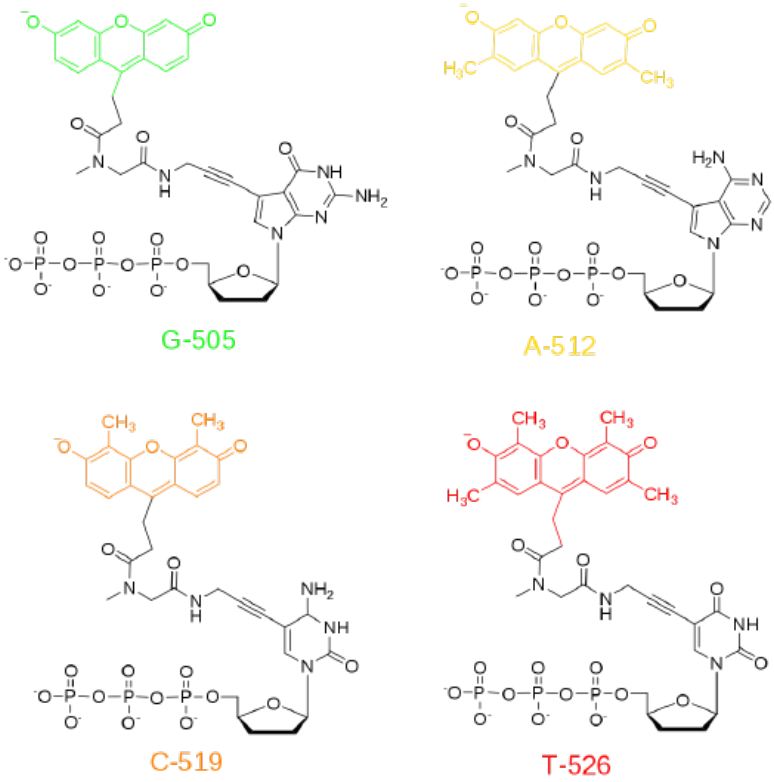


Figure (\PageIndex{7}\): **Fluorescent ddNTPs for Sanger Sequencing.** Dideoxynucleotides are utilized for sequencing as they cannot be extended further once they are incorporated into the nascent DNA. *Image by Fibonachi*

The dideoxynucleotide concentration should be approximately 100-fold lower than that of the corresponding deoxynucleotide (e.g. 0.005mM ddTTP : 0.5mM dTTP) to allow enough fragments to be produced while still transcribing the complete sequence. In total, four separate reactions are needed in this process to test all four ddNTPs. This is illustrated in Figure (\PageIndex{8}\) below.

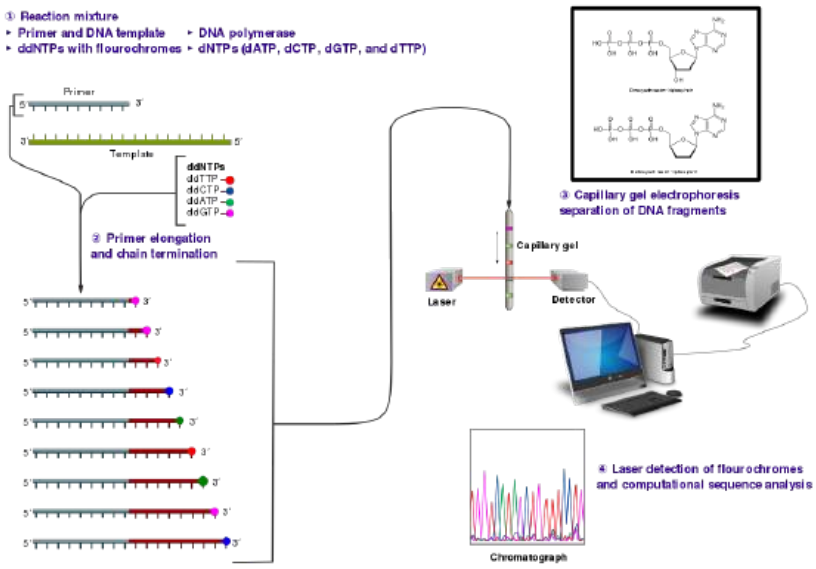


Figure (\PageIndex{8}\): **The Sanger (chain-termination) method for DNA sequencing.** (1) A primer is annealed to a sequence, (2) Reagents are added to the primer and template, including DNA polymerase, dNTPs, and a small amount of all four dideoxynucleotides (ddNTPs) labeled with fluorophores. During primer elongation, the random insertion of a ddNTP instead of a dNTP terminates the synthesis of the chain because DNA polymerase cannot react with the missing hydroxyl. This produces all possible lengths of chains. (3) The products are separated on a single-lane capillary gel, where the resulting bands are read by an imaging system. (4) This produces several hundred thousand nucleotides a day, data that require storage and subsequent computational analysis. *Image by Estevezj*

Following rounds of template DNA extension from the bound primer, the resulting DNA fragments are heat denatured and separated by size using gel electrophoresis. This technique was frequently performed using a denaturing polyacrylamide-urea gel with each of the four reactions run in one of four individual lanes (lanes A, T, G, C). The DNA bands may then be visualized by autoradiography or UV light and the DNA sequence can be directly read off the X-ray film or gel image, as shown in Figure (\PageIndex{9}\) below.

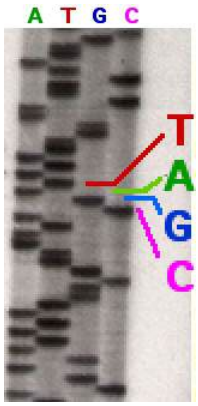


Figure (\PageIndex{9}\): **Traditional Sanger Sequencing Gel.** Sequence visualized by autoradiography. Each lane contains a single reaction that has all four regular nucleotides and a small amount of one of the dideoxynucleotides (ddNTPs). Over time, the ddNTPs will be incorporated at each position containing that specific nucleotide. The gel can then be read from the bottom to the top, as the smallest fragments (those fragments terminated the closest to the primer at the 5'-end) will run the farthest distance in the gel. The sequence of this fragment is 5'-TACGAGATATATGGCGTTAATACGATATATTGGAACTTCTATTGC-3'. *Image by John Schmidt*

Automation of the Sanger sequencing method was made possible when the shift from radioactively tagged nucleotides to fluorescently tagged nucleotides was made. Within the automated sequencers, capillary gel electrophoresis is performed rather than separating the samples using gel electrophoresis. The output from capillary electrophoresis are fluorescent peak trace chromatograms, as shown in **Figure** (\PageIndex{10}\) below. Automated DNA-sequencing instruments (DNA sequencers) can sequence up to 384 DNA samples in a single batch. Batch runs may occur up to 24 times a day greatly enhancing the speed with

which samples may be sequenced and analyzed. Common challenges of DNA sequencing with the Sanger method include poor quality in the first 15-40 bases of the sequence due to primer binding and deteriorating quality of sequencing traces after 400-500 bases.

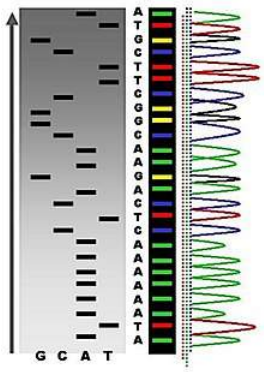


Figure (\PageIndex{10}\): **Side by Side Comparison of Gel Electrophoresis and Capillary Electrophoresis.** *Lefthand Diagram* shows the traditional autoradiogram of Sanger sequencing samples. The *Righthand Diagram* shows the same reactions using fluorescently tagged ddNTPs separated by capillary electrophoresis. The chromatogram output is shown on the far right. *Image by Abizar*

Sanger sequencing is the method that prevailed from the 1980s until ~2005. Over that period, great advances were made in the technique, such as fluorescent labeling, capillary electrophoresis, and general automation. These developments allowed much more efficient sequencing, leading to lower costs. The Sanger method, in mass production form, is the technology that produced the first human genome in 2001, ushering in the age of genomics.

9.1.2.1.3: Microfluidic Sanger Sequencing

Microfluidic Sanger sequencing is a ***lab-on-a-chip application*** for DNA sequencing, in which the Sanger sequencing steps (thermal cycling, sample purification, and capillary electrophoresis) are integrated on a wafer-scale chip using nanoliter-scale sample volumes (Figure (\PageIndex{11}\)). This technology generates long and accurate sequence reads while obviating many of the significant shortcomings of the conventional Sanger method (e.g. high consumption of expensive reagents, reliance on expensive equipment, personnel-intensive manipulations, etc.) by integrating and automating the Sanger sequencing steps.

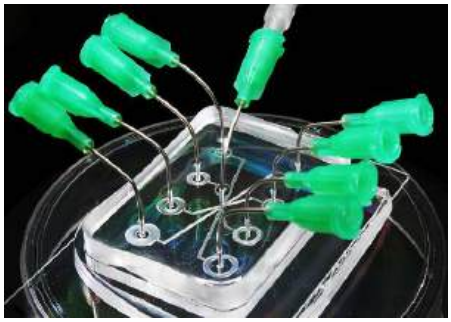


Figure (\PageIndex{11}\): **Lab-On-A-Chip Technologies.** Example of a microfluidic lab-on-a-chip device sitting on a polystyrene dish. Stainless steel needles inserted into the device serve as access points for fluids into small channels within the device, which are about the size of a human hair. *Image by National Institute of Standards and Technology (NIST)*

9.1.2.1.4: Next-generation sequencing (NGS)

Next-generation sequencing (NGS), also known as high-throughput sequencing, is the catch-all term used to describe several different modern sequencing technologies. These technologies allow for the sequencing of DNA and RNA much more quickly and cheaply than the previously used Sanger sequencing, and as such revolutionized the study of genomics and molecular biology. We present information from specific companies that have developed these new technologies without endorsements. Such technologies include:

Illumina Sequencing - In NGS, vast numbers of short reads are sequenced in a single stroke using the lab-on-a-chip technology described above. To do this, the input sample must be cleaved into short sections. In Illumina sequencing, 100-150bp reads are used. Somewhat longer fragments are ligated to generic adaptors and annealed to a slide using the adaptors. PCR is carried out to

amplify each read, creating a spot with many copies of the same read. They are then separated into single-stranded DNA to be sequenced as shown in Figure (\PageIndex{12}\) below.

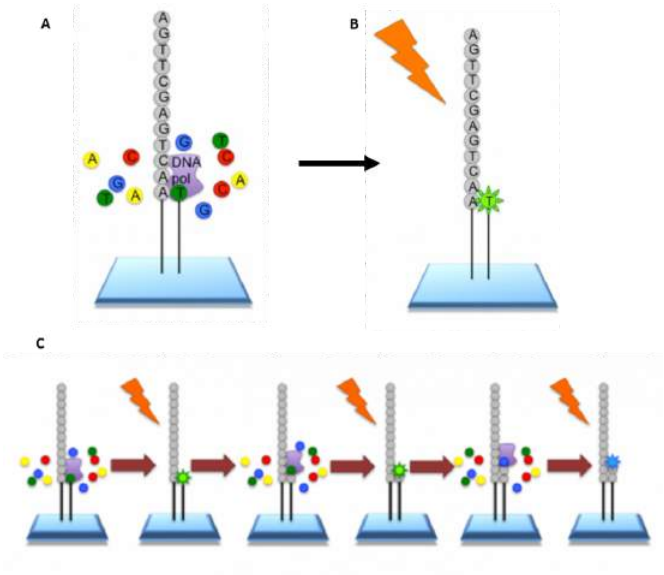


Figure (\PageIndex{12}\): **Procedure for Illumina Sequencing.** (A) The slide with PCR-amplified fragments of DNA is flooded with nucleotides and DNA polymerase. These nucleotides are fluorescently labeled with each color corresponding to a specific base. The reactions also have a terminator present, so that only one base is added at a time. (B) An image is taken of the slide. In each reaction location, there will be a fluorescent signal indicating that a specific base has been added. (C) The data is recorded and the slide is then prepared for the next cycle. In preparation, the terminators are removed, which will allow the next base to be added, and the fluorescent signal is cleaved, preventing the fluorescent signal from contaminating the next image. The process is repeated, adding one nucleotide at a time (G, A, T, or C) and imaging in between. All of the sequence reads will be the same length as single bases are added at each cycle. *Image modified from* EMBL-EBI

Roche 454-Sequencing is similar to the Illumina process but can sequence much longer reads. Like Illumina, it does this by sequencing multiple reads at once by reading optical signals as bases are added. As in Illumina, the DNA or RNA is fragmented into shorter reads, in this case, up to 1kb (1,000bp). Generic adaptors are added to the ends and these are annealed to beads, one DNA fragment per bead. The fragments are then amplified by PCR using adaptor-specific primers. Each bead is then placed in a single well of a slide with each well containing a single bead, covered in many PCR copies of a single sequence. The wells also contain DNA polymerase and sequencing buffers (Figure (\PageIndex{13}\).

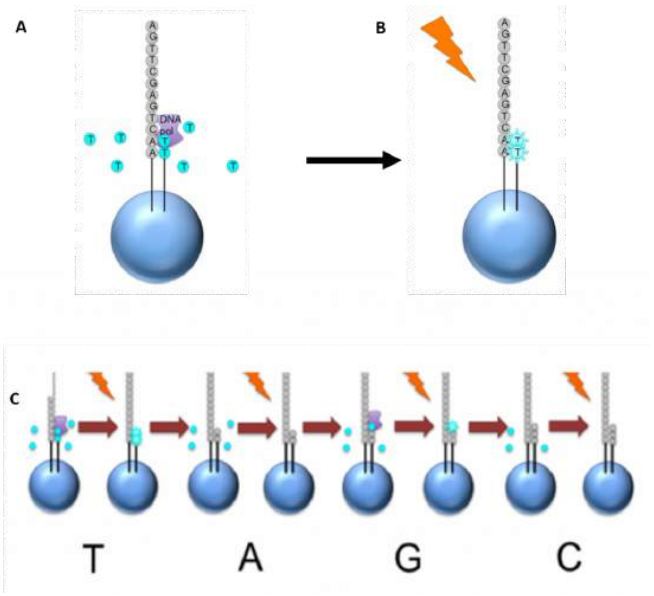


Figure (\PageIndex{13}\): **Procedure for Roche 454 Sequencing.** (A) Once the PCR product is attached to the bead, the slide is flooded with one of the four NTP species. Where this nucleotide is next in the sequence, it is added to the sequence read. If that single base repeats, then more will be added. So if we flood with Guanine bases, and the next in the sequence is G, one G will be added, however, if the next part of the sequence is GGGG, then four Gs will be added. (B) The addition of each nucleotide releases a light signal. These locations of signals are detected and used to determine which beads the nucleotides are added to. (C) The NTP mix is washed away. The next NTP mix is now added and the process is repeated, cycling through the four NTPs. All of the sequence reads from 454 sequencings will be different lengths, because different numbers of bases will be added with each cycle. *Image modified by EMBL-EBI*

Newer technologies such as the ***Ion Torrent Technology*** detect sequence data using electrical signals on a semiconductor chip, rather than optically reading dye-labeled nucleotides. This is possible as the addition of a dNTP to the DNA polymer causes the release of an H^+ ion (Figure (\PageIndex{14}\)). As in other kinds of NGS, the input DNA or RNA is fragmented, this time ~200bp. Adaptors are added and one molecule is placed onto a bead. The molecules are amplified on the bead by emulsion PCR. Each bead is placed into a single well of a slide. This is illustrated in Figure (\PageIndex{14}\) below.

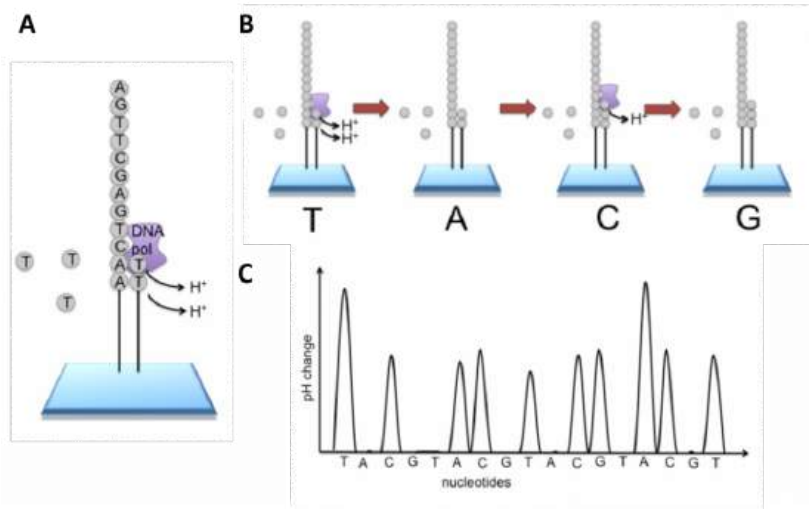


Figure (\PageIndex{14}\): **Ion Torrent Sequencing Technology.** (A) Similar to 454 sequencings, the slide is flooded with a single species of dNTP, along with buffers and polymerase. The pH is monitored in each of the wells after the addition of the specific dNTP. The pH will decrease when the dNTP is incorporated into the polymer causing the release of a proton (H^+). The changes in pH allow the determination of different dNTP species. (C) The pH change, if any, is used to determine how many bases (if any) were added with each cycle. *Image modified from EMBL-EBI*

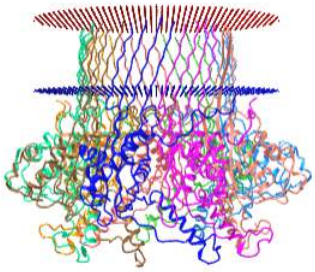
Nanopore Sequencing

In this technique, flow cells are constructed which contain nanopores in a nonlipid membrane that is resistant to current flow. When a voltage is applied across a membrane separating two salt solutions, the current through each nanopore channel can be detected by

a sensor chip. When a larger molecule moves through the pore, a disruption in basal current occurs. Computer algorithms have been developed that detect base-specific changes in the current as the base (even chemical-modified ones) moves through the membrane. The sequence is then decoded in real-time.

Single-stranded DNA or RNA can be driven through the pore by a transmembrane potential in a process similar to electrophoresis. The enzyme DNA helicase, a motor protein, can be attached to the outer part of the pore protein. This enzyme binds single-stranded DNA and moves along the DNA in a process that requires ATP. If the helicase is attached to the pore protein, the single DNA would then move, allowing control of its movement through the pore.

The nanopores are made of real membrane proteins (which we discuss in Chapter 11). One example is α-hemolysin, a heptamer that has an inner pore diameter of 1 nm. When embedded in real cells, it can allow the flow of K^+ (diameter around 250 pm = 0.25 nm) and other ions across the cell membrane, changing the osmotic balance and lysing the cell. The pore size of the proteins used in nanopore sequencing allows single-stranded DNA to flow through it. An interactive iCn3D model of protein used for nanopore sequencing, Curli transport lipoprotein CsgG (4uv3), is shown in the membrane bilayer in Figure 9.1.15.



NCBI iCn3D Figure 9.1.15: Curli transport lipoprotein CsgG (4uv3) for DNA nanopore sequencing. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...LC5K3vSjYjtHN7

Nanopore technologies have enabled the production of small hand-held DNA sequencing devices that can be plugged into the USB drive on a laptop computer and utilized in the field under real-time collection conditions. Future modifications might replace protein pores with synthetic solid-state nanopores. For instance, the membrane might be made of graphene with pores of a specific size made in it.

Figure 9.1.16 shows an animation of a single-stranded DNA moving through a protein pore (blue) assisted by a motor protein (magenta)

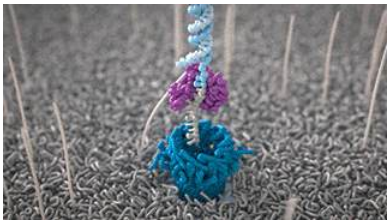


Figure 9.1.16: Single-stranded DNA moving through a protein pore (blue) assisted by a motor protein (magenta) Oxford Nanopore Tech. https://nanoporetech.com/how-it-works

Single Molecule, Real-Time (SMRT) sequencing

In this technique, either RNA or DNA is converted to dsDNA. Deoxynucleotide "adapters" are added to connect the 5'-end of strand 1 to the 3'-end of strand 2 and another adapter to connect the 3'-end of strand 1 to the 5'-end of strand 2 resulting in a "circular" ss-DNA molecule. This single molecule is then drawn into a nanophotonic nanowell made in a thin metal film deposited on glass. The dimensions of each well allow only a single circular ss-DNA molecule. Hundred of different circular ss-DNA molecules entering individual wells are shown in Figure 9.1.17. The blue stretches represent the adapters.

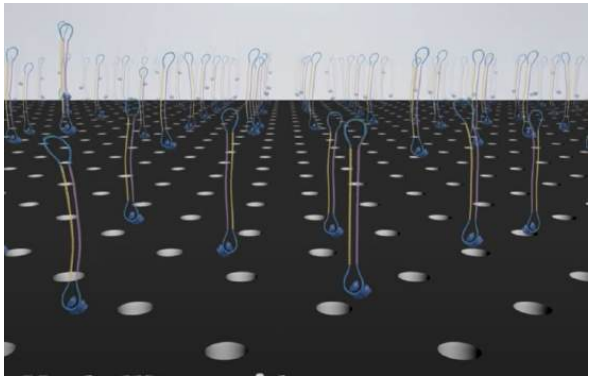


Figure 9.1.17. Nanowells for SMRT sequencing. https://www.youtube.com/embed/_lD8JyAbwEo

The wells are approximately 100 nm in diameter. DNA polymerase and dNTPs can be added to the nanowells, which contained a single molecule of immobilized circular ssDNA. The DNA is immobilized by its biotinylated or attachment to magnetic beads which interact with streptavidin-coated wells. When confined to the wells the apparent concentration of reactants for the polymerization can be quite high, allowing robust DNA polymerase activity.

DNA sequencing using real-time fluorescence monitoring can be done in a massively parallel fashion. The fluorophore is connected to the terminal phosphate of the dNTP. When a phosphodiester bond is made by the DNA polymerase, the fluorophore is released as a leaving group, and the result is a natural, unmodified DNA that continues to grow. The YouTube video below shows the entire process of a single molecule, real-time sequencing.

The four main advantages of Next Generation Sequencing (NGS) over classical sequencing are described below.

Sample size

NGS is significantly cheaper, quicker, needs significantly less DNA, and is more accurate and reliable than Sanger sequencing. Let us look at this more closely. For Sanger sequencing, a large amount of template DNA is needed for each read. Several strands of template DNA are needed for each base being sequenced (i.e. for a 100bp sequence you'd need many hundreds of copies, for a 1000 bp sequence you'd need many thousands of copies), as a strand that terminates on each base is needed to construct a full sequence. In NGS, a sequence can be obtained from a single strand. In both kinds of sequencing multiple staggered copies are taken for contig construction and sequence validation.

Speed

NGS is quicker than Sanger sequencing in two ways. Firstly, the chemical reaction may be combined with the signal detection in some versions of NGS, whereas in Sanger sequencing these are two separate processes. Secondly and more significantly, only one read (maximum ~1kb) can be taken at a time in Sanger sequencing, whereas NGS is massively parallel, allowing 300Gb of DNA to be read on a single run on a single chip.

Cost

The reduced time, people power, and reagents in NGS mean that the costs are much lower. The first human genome sequence cost in the region of $2.7 billion in 2003. Using modern Sanger sequencing methods, aided by data from the known sequence, a full human genome still cost $300,000 in 2006. Sequencing a human genome with NGS today costs roughly $1,000.

Accuracy

Repeats are intrinsic to NGS, as each read is amplified before sequencing, and because it relies on many short overlapping reads, each section of DNA or RNA is sequenced multiple times. Also, because it is so much quicker and cheaper, it is possible to do more repeats than with Sanger sequencing. More repeats mean greater coverage, which leads to a more accurate and reliable sequence, even if individual reads are less accurate for NGS.

The nanopore and single molecule, real-time (SMRT) sequencing were recently employed to complete the sequence for the full human genome (2022). Previous genomic sequences were missing regions with highly repetitive sequences at the centromeres and telomeres. The "Telomere-to-Telomere (T2T) Consortium performed the analysis which added 200 megabases of new sequence information that was missing from the previous best sequence.

9.1.2.2: DNA Synthesis Techniques

DNA synthesis is the natural or artificial creation of deoxyribonucleic acid (DNA) molecules. The term DNA synthesis can refer to ***DNA replication*** (which will be covered in more detail in Chapter XX), ***polymerase chain reaction (PCR),*** or ***gene synthesis*** (physically creating artificial gene sequences).

9.1.2.2.1: Polymerase Chain Reaction (PCR)

The Polymerase chain reaction (PCR) refers to a technique employed widely in the basic and biomedical sciences. PCR is a laboratory technique utilized to amplify specific segments of DNA for a wide range of laboratory and/or clinical applications. Building on the work of Panet and Khorana's successful amplification of DNA *in-vitro*, Kary Mullis and coworkers developed PCR in the early 1980s, having been met with a Nobel prize only a decade later. Allowing for more than the billion-fold amplification of specific target regions, it has become instrumental in many applications including the cloning of genes, the diagnosis of infectious diseases, and the screening of prenatal infants for deleterious genetic abnormalities.

Fundamentals

The main components of PCR are a template, primers, free nucleotide bases, and the DNA polymerase enzyme. The ***DNA template*** contains the specific region that you wish to amplify, such as the DNA extracted from a piece of hair for example. ***Primers***, or oligonucleotides, are short strands of single-stranded DNA complementary to the 3' end of each target region. Both a forward and a reverse primer are required, one for each complementary strand of DNA. ***DNA polymerase*** is the enzyme that carries out DNA replication. Thermostable analogs of DNA polymerase I, such as Taq polymerase, which was originally found in a bacterium that grows in hot springs, are a common choice due to their resistance to the heating and cooling cycles necessary for PCR.

PCR takes advantage of the complementary base pairing, double-stranded nature, and melting temperature of DNA molecules. This process involves cycling through 3 sequential rounds of temperature-dependent reactions: DNA melting (denaturation), annealing, and enzyme-driven DNA replication (elongation). ***Denaturation*** begins by heating the reaction to about 95^{o}C, disrupting the hydrogen bonds that hold the two strands of template DNA together. Next, the reaction is reduced to around 50 to 65^{o}C, depending on the physicochemical variables of the primers, enabling the ***annealing*** of complementary base pairs. The primers, which are added to the solution in excess, bind to the beginning of the 3' end of each template strand and prevent re-hybridization of the template strand with itself. Lastly, enzyme-driven DNA replication, or ***elongation***, begins by setting the reaction temperature to the amount which optimizes the activity of DNA polymerase, which is around 75 to 80^{o}C. At this point, DNA polymerase, which needs double-stranded DNA to begin replication, synthesizes a new DNA strand by assembling free nucleotides in solution in the 3' to 5' direction to produce 2 full sets of complementary strands. The newly synthesized DNA is now identical to the template strand and will be used as such in the progressive PCR cycles. The steps in PCR are animated in the video below.

Figure (\PageIndex{17}\) below illustrate the step involved in PCR amplification of target DNA.

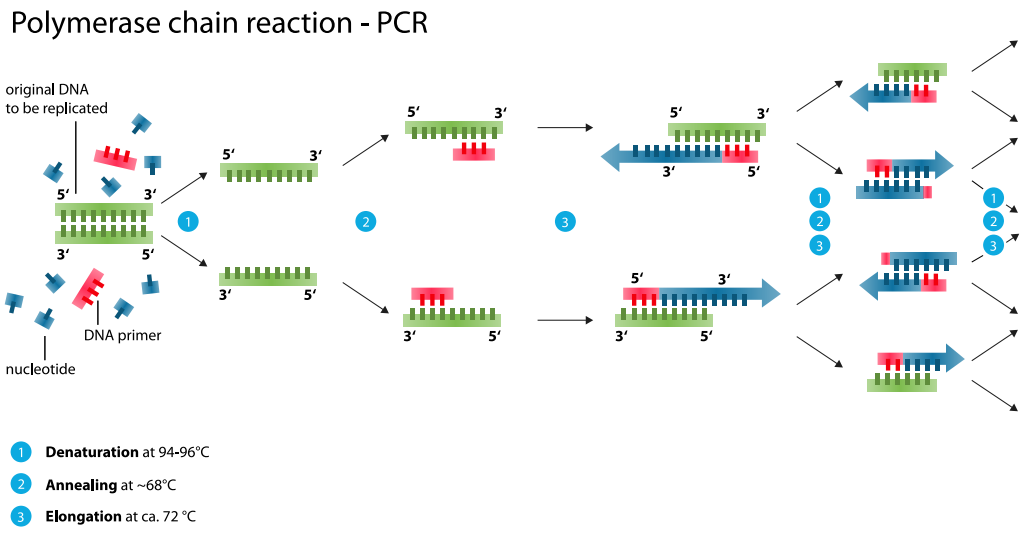


Figure (\PageIndex{17}\): Steps involved in PCR amplification of target DNA.

Figure (\PageIndex{18}\) shows a video animation of a PCR reaction.

Figure (\PageIndex{18}\): Video animation of a PCR reaction.

Given that previously synthesized DNA strands serve as templates, the amplification of DNA using PCR increases at an exponential rate, where the copies of DNA double at the end of each replication step. The exponential replication of the target DNA eventually plateaus around 30 to 40 cycles mainly due to reagent limitation, but can also be due to inhibitors of the polymerase reaction found in the sample, self-annealing of the accumulating product, and accumulation of pyrophosphate molecules.

9.1.2.2.2: Real-Time PCR

At its advent, PCR technology was limited to qualitative and or semi-quantitative analysis due to limitations on the ability to quantitate nucleic acids. At that time, to verify if the target gene was amplified successfully, the DNA product was separated by size via agarose gel electrophoresis. Ethidium bromide, a molecule that fluoresces when bound to dsDNA, could give a rough estimate of DNA amount by roughly comparing the brightness of separated bands, but was not sensitive enough for rigorous quantitative analysis.

Improvements in fluorophore development and instrumentation led to thermocyclers that no longer required measurement of only end-product DNA. This process, known as ***real-time PCR,*** or ***quantitative PCR (qPCR),*** has allowed for the detection of dsDNA during amplification. qPCR thermocyclers are equipped with the ability to excite fluorophores at specific wavelengths, detect their emission with a photodetector, and record the values. The sensitive collection of numerical values during amplification has strongly enhanced quantitative analytical power.

There are two main types of fluorophores used in qPCR: those that bind specifically to a given target sequence and those that do not. The sensitivity of fluorophores has been an important aspect of qPCR development. One of the most effective and widely used non-specific markers, SYBR Green, after binding to the minor groove of dsDNA, exhibits a 1000-fold increase in fluorescence compared to being free in solution (Video 5.1). However, if even more specificity is desired, a sequence-specific oligonucleotide, or hybridization probe, can be added, which binds to the target gene at some point in front of the primer (after the 3' end). These hybridization probes contain a reporter molecule at the 5' end and a quencher molecule at the 3' end. The quencher molecule effectively inhibits the reporter from fluorescing while the probe is intact. However, upon contact with DNA polymerase I, the hybridization probe is cleaved, allowing for the fluorescence of the dye (Video 5.1).

9.1.2.2.3: Reverse-Transcription PCR

Since its advent, PCR technology has been creatively expanded upon, and ***reverse-transcription PCR (RT-PCR)*** is one of the most important advances. Real-time PCR is frequently confused with reverse-transcription PCR, but they are separate techniques. In RT-PCR, the DNA amplified is derived from mRNA by using reverse-transcriptase enzymes, to produce a cDNA copy of the gene. Using primer sequences for genes of interest, traditional PCR methods can be used with the cDNA to study the expression of genes qualitatively. Currently, reverse-transcription PCR is commonly used with real-time PCR, which allows one to quantitatively measure the relative change in gene expression across different samples.

Figure (\PageIndex{19}\) shows a video animation video showing the use of reverse transcription Polymerase Chain Reaction (RT-PCR) in COVID-19 testing.

Figure (\PageIndex{19}\): Video animation of the Reverse Transcription Polymerase Chain Reaction (RT-PCR)

Issues of Concern

One disadvantage of PCR technology is that it is extremely sensitive. Trace amounts of RNA or DNA contamination in the sample can produce extremely misleading results. Another disadvantage is that the primers designed for PCR require sequence data, and therefore can only be used to identify the presence or absence of a known pathogen or gene. Another limitation is that sometimes the primers used for PCR can anneal non-specifically to similar sequences, but not identical, to the target gene.

Another potential issue of using PCR is the possibility of primer dimer (PD) formation. PD is a potential by-product and consists of primer molecules that have hybridized with each other due to the strings of complementary bases in the primers. The DNA polymerase amplifies the PD, leading to competition for PCR reagents that could be used to amplify the target sequences.

Clinical Significance

PCR amplification is an indispensable tool with various applications within medicine. Often, it is used to test for the presence of specific alleles, such as in the case of prospective parents screening for genetic carriers, but it can also be used to diagnose the presence of disease directly and for mutations in the developing embryo. For example, the first time PCR was used in this way was for the diagnosis of sickle cell anemia through the detection of a single gene mutation.

Additionally, PCR has greatly revolutionized the diagnostic potential for infectious diseases, as it can be used to rapidly determine the identity of microbes that were traditionally unable to be cultured, or that required weeks for growth. Pathogens routinely detected using PCR include Mycobacterium tuberculosis, human immunodeficiency virus, herpes simplex virus, syphilis, and

countless other pathogens. Moreover, qPCR is not only used for testing the qualitative presence of microbes but also to quantify bacterial, fungal, and viral loads.

The sensitivity of diagnostic tools for mutations to oncogenes and tumor suppression genes has been improved at least 10,000-fold due to PCR, allowing for earlier diagnosis of cancers like leukemia. PCR has also enabled more nuanced and individualized therapies for cancer patients. Additionally, PCR can be used for the tissue typing done that is vital to organ implantation and has even been proposed as a replacement for antibody-based tests for blood type. PCR also has clinical applications in the field of prenatal testing for various genetic diseases and/or clinical pathologies. Samples are obtained either via amniocentesis or chorionic villus sampling.

In forensic medicine, short pieces of repeating, highly polymorphic DNA, coined short tandem repeats (STRs), are amplified and used to compare specific variations within genes to differentiate individuals.[9] Primers specific for the loci of these STRs are used and amplified using PCR. Various loci contain STRs in the human genome, and the statistical power of this technique is enhanced by checking multiple sites.

9.1.2.3: Gene Synthesis

Artificial gene synthesis, sometimes known as **DNA printing** is a method in synthetic biology that is used to create artificial genes in the laboratory. Based on solid-phase DNA synthesis, it differs from molecular cloning and polymerase chain reaction (PCR) in that it does not have to begin with preexisting DNA sequences. Therefore, it is possible to make a completely synthetic double-stranded DNA molecule with no apparent limits on either nucleotide sequence or size.

The method has been used to generate functional bacterial or yeast chromosomes containing approximately one million base pairs. Creating novel nucleobase pairs in addition to the two base pairs in nature could greatly expand the genetic code.

Synthesis of the first complete gene, a yeast tRNA, was demonstrated by Har Gobind Khorana and coworkers in 1972. Synthesis of the first peptide- and protein-coding genes was performed in the laboratories of Herbert Boyer and Alexander Markham, respectively.

Commercial gene synthesis services are now available. Approaches are most often based on a combination of organic chemistry and molecular biology techniques and entire genes may be synthesized "de novo", without the need for template DNA. Gene synthesis is an important tool in many fields of recombinant DNA technology including heterologous gene expression, vaccine development, gene therapy, and molecular engineering. The synthesis of nucleic acid sequences can be more economical than classical cloning and mutagenesis procedures. It is also a powerful and flexible engineering tool for creating and designing new DNA sequences and protein functions.

Gene optimization

While the ability to make increasingly long stretches of DNA efficiently and at lower prices is a technological driver of this field, increasing attention is being focused on improving the design of genes for specific purposes. Early in the genome sequencing era, gene synthesis was used as an (expensive) source of cDNAs that were predicted by genomic or partial cDNA information but were difficult to clone. As higher-quality sources of sequence-verified cloned cDNA have become available, this practice has become less urgent.

Producing large amounts of protein from gene sequences can sometimes prove difficult. Many of the most interesting proteins are normally regulated to be expressed in very low amounts in wild-type cells. Redesigning these genes offers a means to improve gene expression in many cases. Rewriting the open reading frame is possible because of the degeneracy of the genetic code. Thus it is possible to change up to about a third of the nucleotides in an open reading frame and still produce the same protein. The available number of alternate designs possible for a given protein is astronomical. For a typical protein sequence of 300 amino acids, there are over 10^{150} codon combinations that will encode an identical protein. Codon optimization, or replacing rarely used codons with more common codons sometimes has dramatic effects. Further optimizations such as removing RNA secondary structures can also be included. At least in the case of *E. coli*, protein expression is maximized by predominantly using codons corresponding to tRNA that retain amino acid charging during starvation. Computer programs are used to optimize this task. A well-optimized gene can improve protein expression 2 to 10-fold, and in some cases, more than 100-fold improvements have been reported. Because of the large number of nucleotide changes made to the original DNA sequence, the only practical way to create the newly designed genes is to use gene synthesis.

Oligonucleotide synthesis

Oligonucleotides are chemically synthesized using building blocks called nucleoside phosphoramidites. These can be normal or modified nucleosides that have protecting groups to prevent their amines, hydroxyl groups, and phosphate groups from interacting incorrectly. One phosphoramidite is added at a time, the 5' hydroxyl group is deprotected and a new base is added, and the process is repeated. The chain grows in the 3' to 5' direction, which is backward relative to DNA biosynthesis *in vivo*. In the end, all the protecting groups are removed. The solid phase DNA synthesis reaction is shown in Figure (\PageIndex{20}\) below.

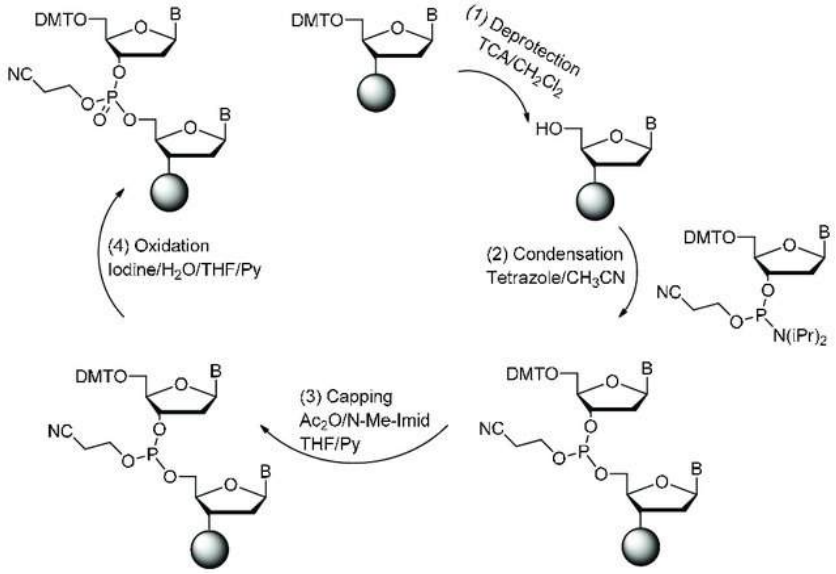


Figure (\PageIndex{20}\): **Four-step phosphoramidite oligodeoxynucleotide synthesis cycle.** The phosphoramidite method, pioneered by Marvin Caruthers in the early 1980s, and enhanced by the application of solid-phase technology and automation, is now firmly established as the method of choice. Phosphoramidite oligonucleotide synthesis proceeds in the 3 to 5 direction (opposite to the 5 to 3 direction of DNA biosynthesis in DNA replication). One nucleotide is added per synthesis cycle. The phosphoramidite DNA synthesis cycle consists of a series of steps outlined in the figure. *Image by Ni, S., et al (2017) International Journal of Molecular Sciences 18(8):1683*

Nevertheless, being a chemical process, several incorrect interactions occur leading to some defective products. The longer the oligonucleotide sequence that is being synthesized, the more defects there are, thus this process is only practical for producing short sequences of nucleotides. The current practical limit is about 200 bp (base pairs) for an oligonucleotide with sufficient quality to be used directly for a biological application. HPLC can be used to isolate products with the proper sequence. Meanwhile, a large number of oligos can be synthesized in parallel on gene chips. For optimal performance in subsequent gene synthesis procedures, they should be prepared individually and on larger scales.

9.1.2.3.1: DNA synthesis and synthetic biology

The significant drop in the cost of gene synthesis in recent years due to increasing competition of companies providing this service has led to the ability to produce entire bacterial plasmids that have never existed in nature. The field of synthetic biology utilizes the technology to produce synthetic biological circuits, which are stretches of DNA manipulated to change gene expression within cells and cause the cell to produce a desired product.

The ability to synthetically produce DNA will enable the development of environmental, medical, and commercially relevant products. For example, in 2015, Novartis, in collaboration with the Synthetic Genomics Vaccines inc. and the US Biomedical Advanced Research and Development Authority, announced that they had effectively created a synthetic DNA influenza vaccine. New synthetic DNA vaccines hold promise to provide an alternative to current egg-produced conventional vaccines that can be plagued by low efficacy.

DNA vaccines can avoid many issues associated with egg-based vaccine production by generating viral proteins within host cells. To create a DNA vaccine, an antigen-encoding gene is cloned into a non-replicative expression plasmid, which is delivered to the host by traditional vaccination routes. Host cells that take up the plasmid express the vaccine antigen which can be presented to immune cells via the major histocompatibility complex (MHC) pathways. CD4+ T helper cell activation following MHC class II presentation of secreted DNA vaccine protein is critical for the production of antigen-specific antibodies as shown in Figure (\PageIndex{21}\) below.

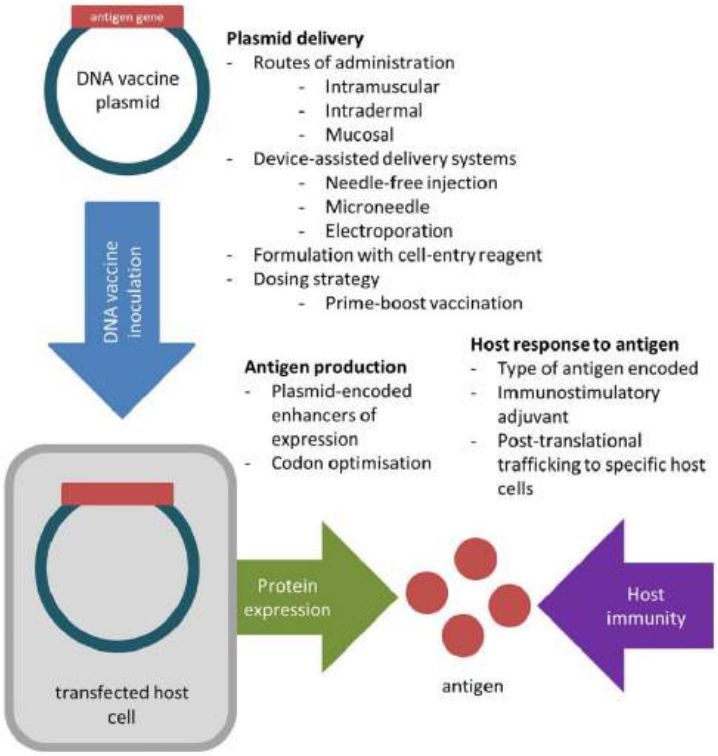


Figure by Lee, L.Y.Y., Izzard, L., and Hurt, A.C. (2018) Frontiers in Immunology, doi: 10.3389/fimmu.2018.01568

After two decades of research, DNA vaccine technology is gaining maturity—several veterinary DNA vaccines are currently licensed for West Nile virus and melanoma, and significantly, the first commercial DNA vaccine against H5N1 in chickens has recently been conditionally approved by the USDA. In addition, ongoing large animal trials of DNA vaccines against other diseases such as HIV, hepatitis, and Zika virus offer valuable insights that can be applied to influenza DNA vaccine design. Promising approaches have arisen from the numerous studies evaluating different DNA vaccine formulations and delivery systems, but a strategy that consistently elicits protection against influenza in large animal models has not yet emerged. Successful plasmid delivery and the use of appropriate adjuvants remain key challenges that need to be addressed before influenza DNA vaccines become effective for human use.

9.1.2.4: RNA Vaccines

Of course, amazing progress has been made in the creation of RNA vaccines as demonstrated by the Moderna and Pfizer RNA vaccines to the spike protein of the SARS-CoV2 virus. Models show that just in the US through 2022, Covid vaccines saved the US $1.15 trillion and 3 million lives. Just in the first year, the vaccines are estimated to have saved upwards of 20 million lives worldwide, an accomplishment worthy of every prize in the world! Many more could have been saved in the developing world if the vaccine was more widely available.

As we learned previously, RNA is much more labile to hydrolysis than DNA since RNA has a 2' OH group. Methods to stabilize RNA were required before an RNA vaccine could become a reality. More importantly, dsRNA is a danger signal that a viral infection may be present. dsRNA, a pathogen-associated molecular pattern (PAMP), binds to the Toll-like Receptor 3 (TLR3) and initiates an inflammatory response that would eliminate the RNA before it could be decoded to form a protein sequence that could elicit an antibody response, a requirement for a vaccine.

Katalin Karikó and Drew Weissman found that modifying the uracil bases to pseudouracil (Ψ) prevents the PAMP response and allows RNA to last long enough to create the protein sequence (i.e. it increases the stability of the RNA to hydrolysis). Pseudouracil (Ψ) is found in structural RNAs (transfer, ribosomal, small nuclear, and small nucleolar), and in fact is the most common modification found in RNA. It even is metabolized by a naturally occurring pathway back to uracil.

Figure (\PageIndex{22}\) below shows the structures of uracil and the N1-methyl derivative of pseudouracil attached to ribose in an RNA and their base pairing to adenine. The N1-methylpseudouracil base still bases pairs with an adenine base. Hence the RNA structure and its functional ability to be decoded into a protein sequence is **not** affected by the modification. In short, mRNAs

modified to contain pseudouracil have a much greater translational capacity as well as stability. In addition, methylation of uracil decreases the immunogenicity of RNAs that contain it.

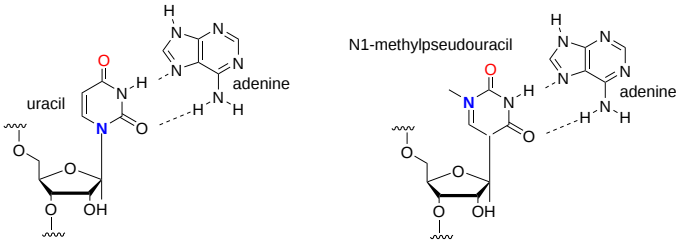


Figure (\PageIndex{22}\): Structural and base pairing features of uracil and N1-methylpseudouracil

Figure (\PageIndex{23}\) below shows a cartoon version of the modified mRNA (a) Covid-19 and its encapsulation into a lipid nanoparticle (b) for an mRNA vaccine.

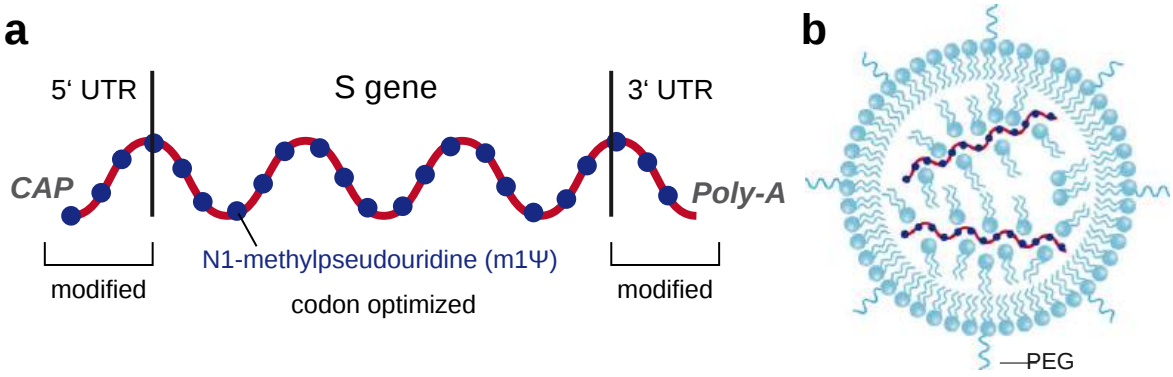


Figure (\PageIndex{23}\): Modified mRNA (a) Covid-19 and its encapsulation into a lipid nanoparticle (b) for an mRNA vaccine. Heinz, F.X., Stiasny, K. Distinguishing features of current COVID-19 vaccines: knowns and unknowns of antigen presentation and modes of action. *npj Vaccines* **6,** 104 (2021). https://doi.org/10.1038/s41541-021-00369-6. Creative Commons Attribution 4.0 International License

Panel (a) shows the generalized structure of the mRNA vaccine for the S gene which encodes the surface spike protein of the virus. Panel b shows the lipid nanoparticle that encapsulates and protects the mRNA vaccine. The lipids include proprietary mixtures of phospholipids, cholesterol, cationic (ionizable) lipids, and polyethylene glycol (PEGylated). Without the basic research of Karikó and Weissman (and many others), the world would not have had the Covid-19 vaccine in time to save so many lives. They were awarded the Nobel Prize in Medicine in 2023 for their work.

Biosynthesis of pseudouridine

The biosynthesis of pseudouridine, the most common modification of cellular RNA, is done after DNA transcription (i.e. post-transcriptionally) using the enzyme pseudouridine synthases (PUS), which is found in all kingdoms in life. The reaction involves

- cleavage of the C-N-glycosidic bond of uridine in RNA
- rotation of the cleaved uracil to align C5 of uracil and C1′ of the ribose
- formation of the C1′-C5 carbon–carbon bond.

These processes are illustrated in Figure (\PageIndex{24}\) below.

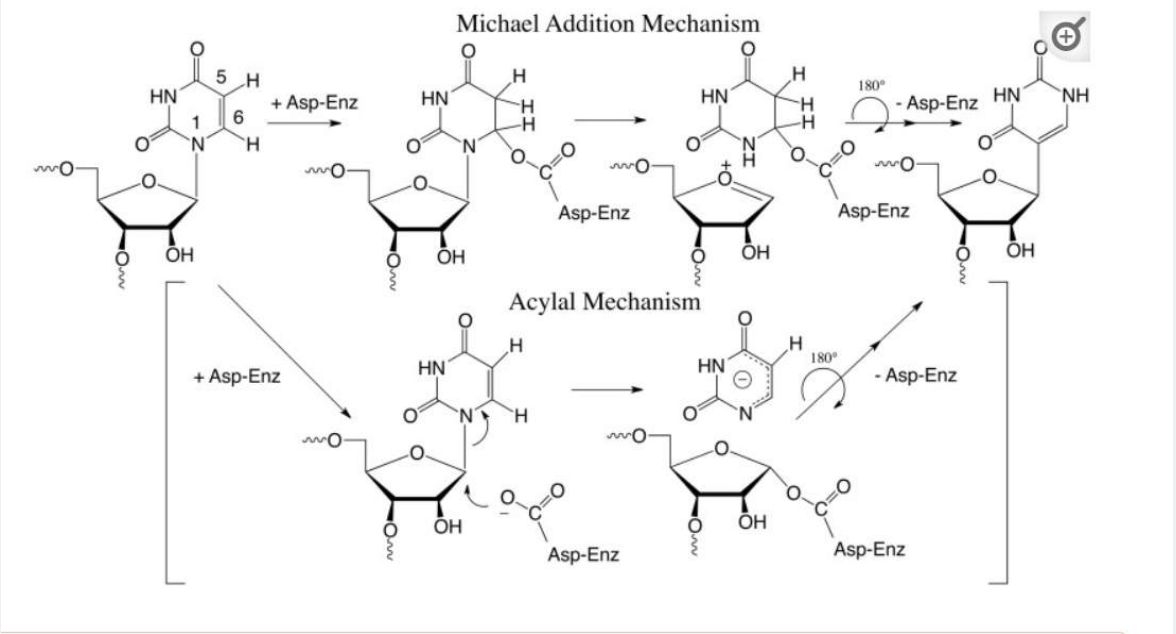


Figure (\PageIndex{24}\): Post-transcriptional modification of uridine to pseudouridine. Czudnochowski N. et al. The mechanism of pseudouridine synthases from a covalent complex with RNA, and alternate specificity for U2605 versus U2604 between close homologs. Nucleic Acids Res. 2014 Feb;42(3):2037-48. doi: 10.1093/nar/gkt1050. Epub 2013 Nov 7. PMID: 24214967; PMCID: PMC3919597. Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0/)

9.2: Bioinformatics

9.2.1: Introduction

An unprecedented revolution has been observed in science with recent technological advances, which have provided a large amount of “omic” data. The crescent generation and availability of this information available in public databases were, and still are, a challenge for professionals from different areas. However, what is the challenge? In biology, the main challenge is to make sense of the enormous amount of structural data and sequences that have been generated at multiple levels of biological systems. Still, in bioinformatics, the development of tools is necessary (statistical and computational) capable of assisting in understanding the mechanisms underlying biological questions in the study. Besides, if we consider the complexity of science, this is a highly reductionist view. The era of a “new biology” emerges accompanied by the birth/development of other sciences, such as bioinformatics and computational biology, which have an integrated interface with molecular biology. Although considered recently, bioinformatics and genomics have evolved interdependently and promoted a historical impact on the available knowledge.

Bioinformatics is a hybrid science that links biological data with techniques for information storage, distribution, and analysis to support multiple areas of scientific research including biomedicine. It uses molecular biology and genetics, computer science, mathematics, and statistics to address data-intensive, large-scale biological problems are addressed from a computational point of view. Bioinformatics is fed by high-throughput data-generating experiments, including genomic sequence determinations and measurements of gene expression patterns. Database projects curate and annotate the data and then distribute it via the World Wide Web. Mining these data leads to scientific discoveries and in the end to the identification of new clinical applications.

A bioinformatics solution usually involves the following steps:

- Collect statistics from biological data
- Build a computational model
- Solve a computational modeling problem
- Test and evaluate a computational algorithm

It also addresses the following aspects:

- Types of biological information and databases
- Sequence analysis and molecular modeling
- Genomic analysis
- Systems biology

In the field of medicine in particular, many important applications for bioinformatics have been discovered. For example, it is used to identify correlations between gene sequences and diseases, to predict protein structures from amino acid sequences, to aid in the design of novel drugs, and to tailor treatments to individual patients based on their DNA sequences (pharmacogenomics). Using bioinformatics, we can now conduct global analyses of all the available data to uncover common principles that apply across many systems and highlight novel features.

Many bioinformatic tools are available online in the fields of genomics, comparative genomics, proteomics, drug discovery, cancer research, phylogenetics, forensic sciences, biodefense, nutrigenomics, gene expressions, protein structure/function, etc. Instead of endorsing any given tool, and constantly curating web addresses, which often change, we offer a series of links that readers can explore to find the tools appropriate to their needs.

- Bioinformatics resources - NCBI/NIH
- What is bioinformatics from EMBL-EMI: A course
- Bioinformatics tools from GenScript
- Bioinformatics tools - FDA
- Bioinformatics Resources - UC Berkeley
- NCBI Bioinformatics Resources: An Introduction: NCBI/NIH Homepage - UC Berkeley

9.3: Cloning and Recombinant Expression

To clone a gene from an organism and express it in either a prokaryotic or eukaryotic cells, DNA from a target source must be isolated, purified, amplified, analyzed and sequenced as described in previous sections.

9.3.1: Cloning

In general, cloning means the creation of a perfect replica. Typically, the word is used to describe the creation of a genetically identical copy. In biology, the re-creation of a whole organism is referred to as “reproductive cloning.” Long before attempts were made to clone an entire organism, researchers learned how to copy short stretches of DNA—a process that is referred to as ***molecular cloning***.

Molecular cloning allows for the creation of multiple copies of genes, the expression of genes, and study the of specific genes. To get the DNA fragment into a bacterial cell in a form that will be copied or expressed, the fragment is first inserted into a cloning vector.

Cloning vector

A Cloning vector small piece of DNA that can be stably maintained in an organism, and into which a foreign DNA fragment can be inserted for cloning purposes. The cloning vector may be DNA taken from a virus, the cell of a higher organism, or the plasmid of a bacterium. The vector contains features that allow for the convenient insertion or removal of a DNA fragment to or from the vector, for example by treating the vector and the foreign DNA with a restriction enzyme that cuts the DNA. DNA fragments thus generated contain either blunt ends or overhangs known as sticky ends, and vector DNA and foreign DNA with compatible ends can then be joined together by molecular ligation. After a DNA fragment has been cloned into a cloning vector, it may be further subcloned into another vector designed for more specific use.

There are many types of cloning vectors, but the most commonly used ones are genetically engineered plasmids. Cloning is generally first performed using *Escherichia coli*, and cloning vectors in *E. coli* include plasmids, bacteriophages (such as phage λ), cosmids, and bacterial artificial chromosomes (BACs). Some DNA, however, cannot be stably maintained in *E. coli*, for example very large DNA fragments. For these studies, other organisms such as yeast may be used. Cloning vectors in yeast include yeast artificial chromosomes (YACs). The common bacterial cloning plasmid, pRB322, is shown in Figure 9.3.1.

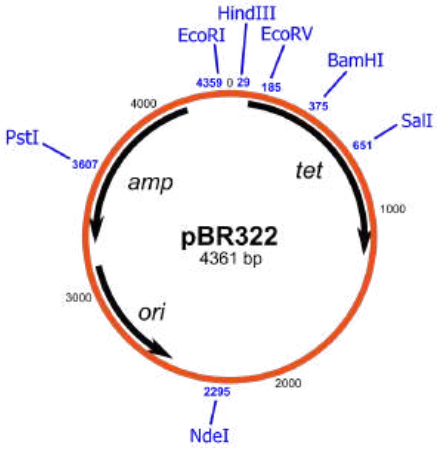


Figure 9.3.1: The cloning vector pRB322. *Image by Ayacop and Yikrazuul*

All commonly used cloning vectors in molecular biology have key features necessary for their function, such as a suitable cloning site with restriction enzymes and a selectable marker. Others may have additional features specific to their use. For reasons of ease and convenience, cloning is often performed using *E. coli*. Thus, the cloning vectors used often have elements necessary for their propagation and maintenance in *E. coli*, such as a functional ***origin of replication (ori).*** The ColE1 origin of replication is found in many plasmids. Some vectors also include elements that allow them to be maintained in another organism in addition to *E. coli*, and these vectors are called ***shuttle vectors***.

Cloning site

All cloning vectors have features that allow a gene to be conveniently inserted into the vector or removed from it. This may be a ***multiple cloning site (MCS) or polylinker***, which contains many unique restriction sites. The restriction sites in the MCS are first

cleaved by restriction enzymes, then a PCR-amplified target gene also digested with the same enzymes is ligated into the vectors using DNA ligase. The target DNA sequence can be inserted into the vector in a specific direction if so desired. The restriction sites may be further used for sub-cloning into another vector if necessary.

Other cloning vectors may use topoisomerase instead of ligase and cloning may be done more rapidly without the need for restriction digest of the vector or insert. In this TOPO cloning method, a linearized vector is activated by attaching topoisomerase I to its ends, and this "TOPO-activated" vector may then accept a PCR product by ligating both the 5' ends of the PCR product, releasing the topoisomerase and forming a circular vector in the process. Another method of cloning without the use of a DNA digest and ligase is by DNA recombination, for example as used in the Gateway cloning system. The gene, once cloned into the cloning vector (called entry clone in this method), may be conveniently introduced into a variety of expression vectors by recombination.

Restriction Enzymes

Restriction enzymes (also called restriction endonucleases) recognize specific DNA sequences and predictably cut them; they are naturally produced by bacteria as a defense mechanism against foreign DNA.

As the name implies, restriction endonucleases (or restriction enzymes) are "*restricted*" in their ability to cut or digest DNA. The restriction that is useful to biochemists is usually a ***palindromic DNA sequence***. Palindromic sequences are the same sequence forwards and backward. Some examples of palindromes: RACE CAR, CIVIC, A MAN A PLAN A CANAL PANAMA. DNA has two complementary strands. Therefore, the reverse complement of one strand is identical to the other.

Like with a palindromic word, the DNA palindromic sequence reads the same forward and backward. In most cases, the sequence reads the same forward on one strand and backward on the complementary strand. Restriction enzymes often cut DNA into a staggered pattern. When a staggered cut is made in a sequence, the overhangs are complementary as shown in Figure 9.3.2.

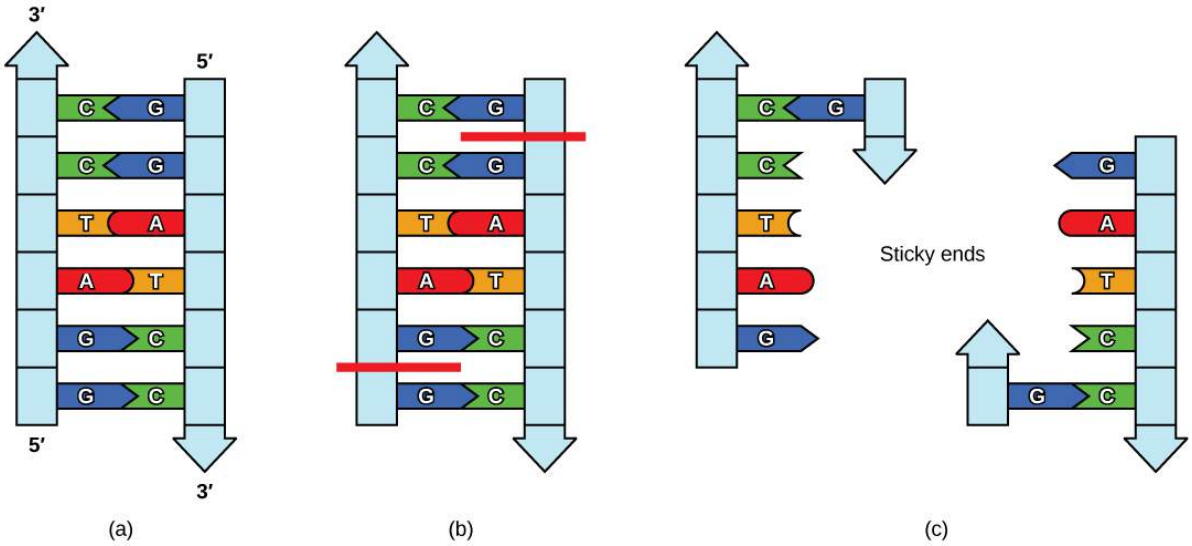


Figure_10_01_04.jpg">http://opentextbc.ca/biology/wp-cont...e_10_01_04.jpg

Figure 9.3.2: **Restriction Enzyme Recognition Sequences.** In this (a) six-nucleotide restriction enzyme recognition site, notice that the sequence of six nucleotides reads the same in the 5′ to 3′ direction on one strand as it does in the 5′ to 3′ direction on the complementary strand. This is known as a palindrome. (b) The restriction enzyme makes breaks in the DNA strands and (c) the cut in the DNA results in "sticky ends". Another piece of DNA cut on either end by the same restriction enzyme could attach to these sticky ends and be inserted into the gap made by this cut. http://opentextbc.ca/biology/wp-cont...e_10_01_04.jpg

Molecular biologists also tend to use these special molecular scissors that recognize palindromes of 6 or 8. By using 6-cutters or 8-cutters, the sequences occur rarely, but often enough, to be of utility. Figure 9.3.3 the sequence for HindII cuts.

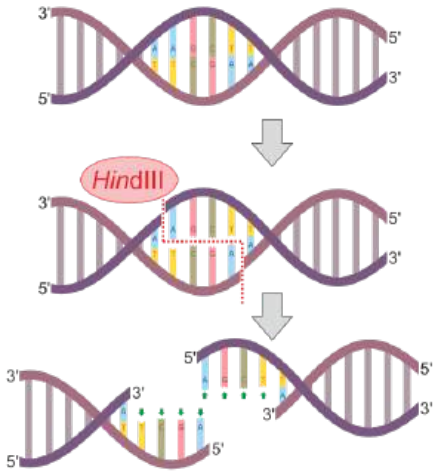
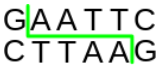


Figure 9.3.3: Sequence of HindIII stick end cuts.

Figure 9.3.4 shows restriction enzyme cuts that leave sticky or blunt end.



SmaI restriction enzyme recognition site

Figure 9.3.4: **Restriction Enzymes.** Restriction enzymes recognize palindromic sequences in DNA and hydrolyze covalent phosphodiester bonds of the DNA to leave either “sticky/cohesive” ends or “blunt” ends. This distinction in cutting is important because an *EcoRI* sticky end can be used to match up a piece of DNA cut with the same enzyme to glue or ligate them back together. While endonucleases cut DNA, **ligases** join them back together. DNA digested with *EcoRI* can be ligated back together with another piece of DNA digested with *EcoRI*, but not to a piece digested with *SmaI*. Another blunt cutter is *EcoRV* with a recognition sequence of GAT | ATC.

Selectable marker

A selectable marker is carried by the vector to allow the selection of positively transformed cells. Antibiotic resistance is often used as a marker , an example being the beta-lactamase gene, which confers resistance to the penicillin group of beta-lactam antibiotics like ampicillin. Some vectors contain two selectable markers, for example, the plasmid pACYC177 has both ampicillin and kanamycin resistance genes. Shuttle vectors that are designed to be maintained in two different organisms may also require two selectable markers, although some selectable markers such as resistance to zeocin and hygromycin B are effective in different cell types. Auxotrophic selection markers that allow an auxotrophic organism to grow in a minimal growth medium may also be used; examples of these are *LEU2* and *URA3* which are used with their corresponding auxotrophic strains of yeast.

Another kind of selectable marker allows for the positive selection of plasmid with cloned genes. This may involve the use of a gene lethal to the host cells, such as barnase, Ccda, and the parD/parE toxins. This typically works by disrupting or removing the lethal gene during the cloning process, and unsuccessful clones where the lethalstill remains intact would kill the host cells, therefore only successful clones are selected.

Reporter genes

Reporter genes are used in some cloning vectors to facilitate the screening of successful clones by using features of these genes that allow successful clone clones easily identified. Such features present in cloning vectors may be the *lacZα* fragment for α complementation in the blue-white selection, and/or marker gene or reporter genes in frame with and flanking the MCS to facilitate the production of fusion proteins. Examples of fusion partners that may be used for screening are the green fluorescent protein (GFP) and luciferase. Figure 9.3.5 shows such a construct with GFP.

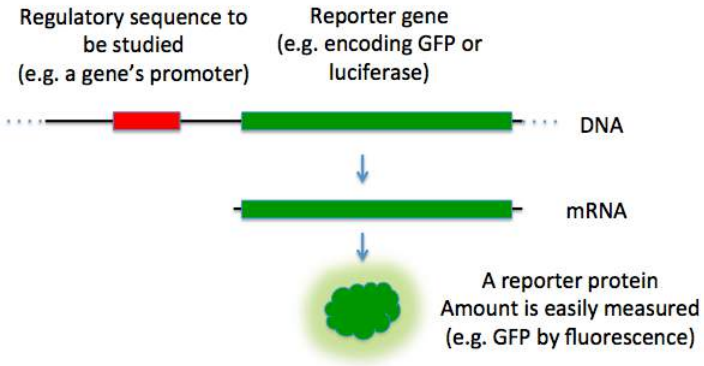


Figure 9.3.5: **Reporter Genes.** In this diagram, the green fluorescence protein is used as a reporter gene to study upstream regulatory sequences. *Image by TransControl*

Elements for expression

If the expression of the targeted gene is desired, then a cloning vector also needs to contain suitable elements for the expression of the cloned target gene, including a promoter and ribosomal binding site (RBS). The target DNA may be inserted into a site that is under the control of a particular promoter necessary for the expression of the target gene in the chosen host. Where the promoter is present, the expression of the gene is preferably tightly controlled and inducible so that proteins are only produced when required. Some commonly used promoters are the T7 and *lac* promoters. The presence of a promoter is necessary when screening techniques such as blue-white selection are used.

Cloning vectors without promoter and RBS for the cloned DNA sequence are sometimes used, for example when cloning genes whose products are toxic to *E. coli* cells. Promoter and RBS for the cloned DNA sequence are also unnecessary when first making a genomic or cDNA library of clones since the cloned genes are normally subcloned into a more appropriate expression vector if their expression is required.

9.3.2: Types of cloning vectors

A large number of cloning vectors are available, and choosing the right vector may depend on a number of factors, such as the size of the insert, copy number, and cloning method. Large DNA inserts may not be stably maintained in a general cloning vector, especially for those with a high copy number, therefore cloning large fragments may require a more specialized cloning vector.

Plasmids

Plasmids are autonomously replicating circular extra-chromosomal DNA. They are the standard cloning vectors and the ones most commonly used. Most general plasmids may be used to clone DNA inserts to 15 kb in size. Many plasmids have high copy numbers, for example, pUC19 has a copy number of 500-700 copies per cell, and a high copy number is useful as it produces a greater yield of recombinant plasmid for subsequent manipulation. However low-copy-number plasmids may be preferably used in certain circumstances, for example, when the protein from the cloned gene is toxic to the cells.

Bacteriophage

The bacteriophages most commonly used for cloning are the lambda (λ) phage and M13 phage. There is an upper limit on the amount of DNA that can be packed into a phage (a maximum of 53 kb). The average lambda phage genome is roughly 48.5 kb. Therefore to allow foreign DNA to be inserted into phage DNA, phage cloning vectors may need to have some of their non-essential genes deleted to make room for the foreign DNA. The phage sequence and cartoon structure are shown in Figure 9.3.6.

There is also a lower size limit for DNA that can be packed into a phage. This property can be used for selection - vectors without inserts may be too small, therefore only vectors with inserts may be selected for propagation.

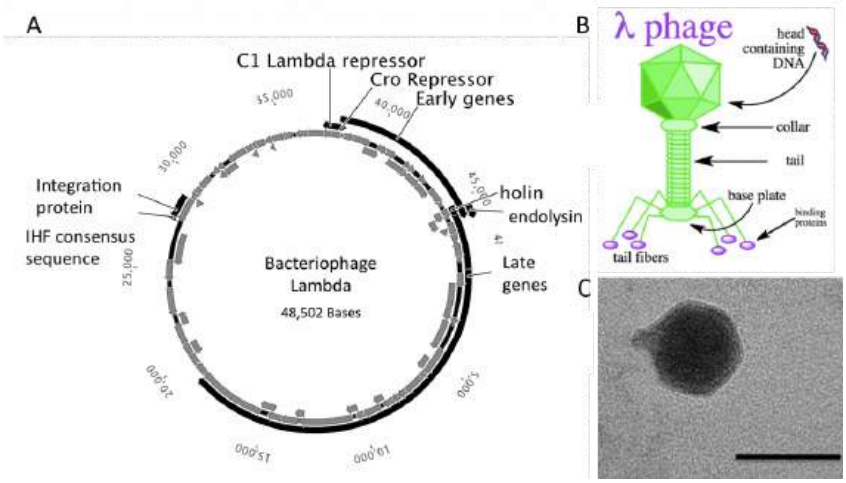


Figure 9.3.6: Lambda Phage. (A) Schematic representation of the circular genome of the lambda phage (B) Diagram of the Lambda Phage infectious particle and (C) Electron micrograph of the related bacteriophage, vibriophage VvAWI. The bar denotes 50 nm in length. *Images A and C are modified from Nigro, O, Culley, A., and Steward, G.F. (2012) Standards in Genomic Science 6(3):415-26, and image B is from Jack Potte*

Cosmids

Cosmids are plasmids that incorporate a segment of bacteriophage λ DNA that has the cohesive end sites (*cos*) which contain elements required for packaging DNA into λ particles. It is normally used to clone large DNA fragments between 28 and 45 Kb.

Bacterial artificial chromosome

Insert size of up to 350 kb can be cloned in a bacterial artificial chromosome (BAC). BACs are maintained in *E. coli* with a copy number of only 1 per cell. BACs have often been used to sequence the genome of organisms in genome projects, including the Human Genome Project. A short piece of the organism's DNA is amplified as an insert in BACs, and hen sequenced. Finally, the sequenced parts are rearranged *in silico*, resulting in the genomic sequence of the organism. BACs have largely been replaced in this capacity with faster and less laborious sequencing methods like whole genome shotgun sequencing and now more recently next-gen sequencing.

Yeast artificial chromosome

Yeast artificial chromosomes are used as vectors to clone DNA fragments of more than 1 megabase (1Mb = 1000kb = 1,000,000 bases) in size. They are useful in cloning larger DNA fragments as required in mapping genomes such as in the human genome project. It contains a telomeric sequence, an autonomously replicating sequence(features required to replicate linear chromosomes in yeast cells). These vectors also contain suitable restriction sites to clone foreign DNA as well as genes to be used as selectable markers.

Human artificial chromosome

Human artificial chromosomes may be potentially useful as gene transfer vectors for gene delivery into human cells, and a tool for expression studies and determining human chromosome function. It can carry very large DNA fragments (there is no upper limit on size for practical purposes), therefore it does not have the problem of limited cloning capacity of other vectors, and it also avoids possible insertional mutagenesis caused by integration into host chromosomes by a viral vector.

Animal and plant viral vectors that infect plant and animal cells have also been manipulated to introduce foreign genes into plant and animal cells. The natural ability of viruses to adsorb to cells, introduce their DNA and replicate has made them ideal vehicles to transfer foreign DNA into eukaryotic cells in culture. A vector based on Simian virus 40 (SV40) was used in first the cloning experiment involving mammalian Severall vectors based on viruses like Adenoviruses and Papilloma virus have been used to clone genes in mammals. At present, retroviral vectors are popular for cloning genes in mammalian cells. In the case of plant transformation, viruses including the Cauliflower Mosaic Virus, Tobcco Mosaic Virus, and Gemini Viruses have been used with limited success.

9.3.3: Summary of DNA Cloning

Figure 9.3.7 provides a summary of the basic cloning methods most widely used in biochemistry laboratories. Foreign DNA is isolated or amplified using PCR to obtain enough material for the cloning procedure. The DNA is purified and cut with restriction enzymes, and then mixed with a vector that has been cut with the same restriction enzymes. The DNA can then be stitched back together with DNA ligase. The DNA can then be transformed into a host system, often bacteria, to grow large quantities of the plasmid containing the cloned DNA.

Restriction fragment patterning and DNA sequencing can be used to validate the cloned material. For a Video Tutorial on DNA Cloning visit HHMI - BioInteractive.

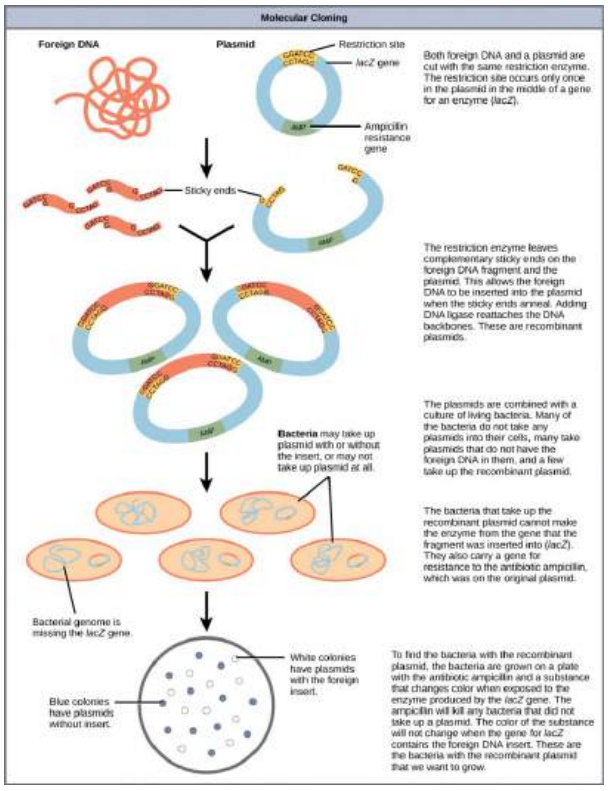


Figure 9.3.7: **Diagram Showing the Major Steps in Cloning.**

Plasmids with foreign DNA inserted into them are called recombinant DNA molecules because they contain new combinations of genetic material. Proteins that are produced from recombinant DNA molecules are called recombinant proteins. Not all recombinant plasmids are capable of expressing genes. Plasmids may also be engineered to express proteins only when stimulated by certain environmental factors, so that scientists can control the expression of the recombinant proteins.

9.3.4: Reproductive Cloning

Reproductive cloning is a method used to make a clone or an **identical copy of an entire multicellular organism**. Most multicellular organisms undergo reproduction by sexual means, which involves the contribution of DNA from two individuals (parents), making it impossible to generate an identical copy or a clone of either parent. Recent advances in biotechnology have made it possible to reproductively clone mammals in the laboratory.

Natural sexual reproduction involves the union, during fertilization, of a sperm and an egg. Each of these gametes is ***haploid,*** meaning they contain one set of chromosomes in their nuclei. The resulting cell, or zygote, is then ***diploid*** and contains two sets of chromosomes. This cell divides mitotically to produce a multicellular organism. However, the union of just any two cells cannot produce a viable zygote; there are components in the cytoplasm of the egg cell that are essential for the early development of the embryo during its first few cell divisions. Without these provisions, there would be no subsequent development. Therefore, to produce a new individual, both a diploid genetic complement and an egg cytoplasm are required. The approach to producing an artificially cloned individual is to take the egg cell of one individual and to remove the haploid nucleus. Then a diploid nucleus

from a body cell of a second individual, the donor, is put into the egg cell. The egg is then stimulated to divide so that development proceeds. This sounds simple, but it takes many attempts before each of the steps is completed successfully.

The first cloned agricultural animal was Dolly, a sheep who was born in 1996 (see **Figure 9.3.8** below). The success rate of reproductive cloning at the time was very low. Dolly lived for six years and died of a lung tumor. There was speculation that because the cell DNA that gave rise to Dolly came from an older individual, the age of the DNA may have affected her life expectancy. Since Dolly, several species of animals (such as horses, bulls, and goats) have been successfully cloned.

There have been attempts at producing cloned human embryos as sources of embryonic stem cells. In the procedure, the DNA from an adult human is introduced into a human egg cell, which is then stimulated to divide. The technology is similar to the technology that was used to produce Dolly, but the embryo is never implanted into a surrogate mother. The cells produced are called embryonic stem cells because they can develop into many different kinds of cells, such as muscle or nerve cells. The stem cells could be used to research and ultimately provide therapeutic applications, such as replacing damaged tissues. The benefit of cloning in this instance is that the cells used to regenerate new tissues would be a perfect match to the donor of the original DNA. For example, a leukemia patient would not require a sibling with a tissue match for a bone-marrow transplant.

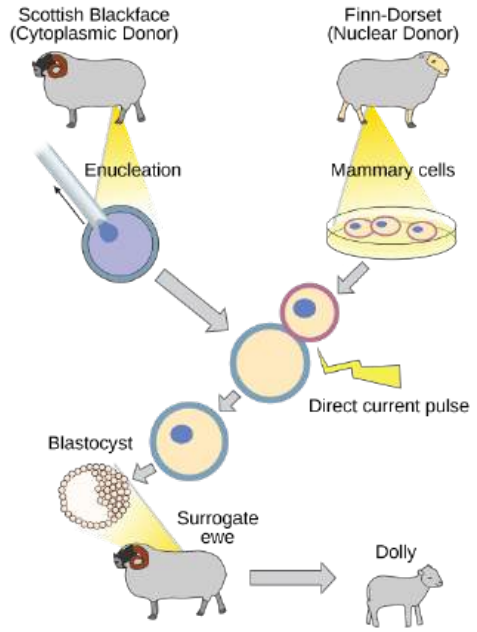


Figure 9.3.8: Cloning of Dolly, the first agricultural animal to be cloned.

To create Dolly, the nucleus was removed from a donor egg cell. The enucleated egg was placed next to the other cell, then they were shocked to fuse. They were shocked again to start division. The cells were allowed to divide for several days until an early embryonic stage was reached, before being implanted in a surrogate mother.

Why was Dolly a Finn-Dorset and not a Scottish Blackface sheep?

Because even though the original cell came from a Scottish Blackface sheep and the surrogate mother was a Scottish Blackface, the DNA came from a Finn-Dorset.

9.3.5: Genetic Engineering

Genetic engineering uses recombinant DNA technology to modify an organism's DNA to achieve desirable traits. The addition of foreign DNA in the form of recombinant DNA vectors that are generated by molecular cloning is the most common method of genetic engineering. An organism that receives the recombinant DNA is called a ***genetically modified organism (GMO)***. If the foreign DNA that is introduced comes from a different species, the host organism is called ***transgenic***. Bacteria, plants, and animals have been genetically modified since the early 1970s for academic, medical, agricultural, and industrial purposes.

Watch this short video explaining how scientists create a transgenic animal.

Although the classic methods of studying the function of genes began with a given phenotype and determined the genetic basis of that phenotype, modern techniques allow researchers to start at the DNA sequence level and ask: "What does this gene or DNA element do?" This technique, called ***reverse genetics***, has resulted in reversing the classical genetic methodology. One example of

this method is analogous to damaging a body part to determine its function. An insect that loses a wing cannot fly, which means that the wing's function is flight. The classic genetic method compares insects that cannot fly with insects that can fly, and observes that the non-flying insects have lost wings. Similarly in a reverse genetics approach, mutating or deleting genes provides researchers with clues about gene function. Alternately, reverse genetics can be used to cause a gene to overexpress itself to determine what phenotypic effects may occur.

9.3.6: CRISPR Technology

CRISPR stands for ***clustered regularly interspaced short palindromic repeats*** and represents a family of DNA sequences found within the genomes of prokaryotic organisms such as bacteria and archaea. These sequences are derived from DNA fragments of bacteriophages that have previously infected the prokaryote and are used to detect and destroy DNA from similar phages during subsequent infections. Hence these sequences play a key role in the antiviral defense system of prokaryotes. Figure 9.3.9 shows the c**rystal structure of a CRISPR RNA-guided surveillance complex, Cascade, bound to a ssDNA target,**

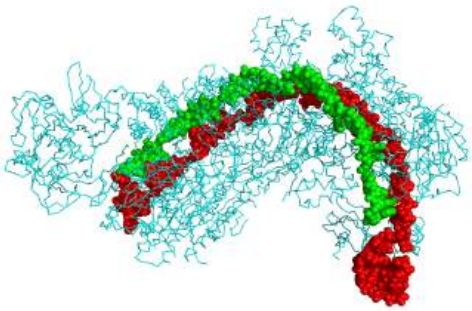
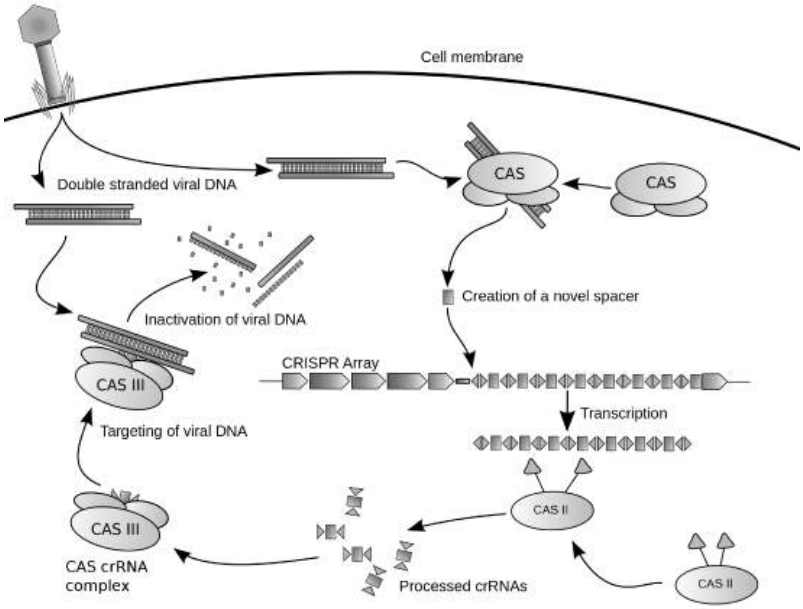


Figure 9.3.9: **Crystal structure of a CRISPR RNA-guided surveillance complex, Cascade, bound to a ssDNA target.** CRISPR system Cascade protein subunits CasA, CasB, CasC, CasD, and CasE (cyan) bound to CRISPR RNA (green) and viral DNA (red) based on PDB 4QYZ and rendered with PyMOL. *Image from Boghog*

Cas9 (or "CRISPR-associated protein 9") is an enzyme that uses CRISPR sequences as a guide to recognize and cleave specific strands of DNA that are complementary to the CRISPR sequence. Cas9 enzymes together with CRISPR sequences form the basis of a technology known as CRISPR-Cas9 that can be used to edit genes within organisms. This editing process has a wide variety of applications including basic biological research, the development of biotechnology products, and the treatment of diseases. Figure 9.3.10s shows a d**iagram of the CRISPR prokaryotic antiviral defense mechanism**



(\PageIndex{10}\): **Diagram of the CRISPR prokaryotic antiviral defense mechanism.** *Image by James Atmos*

The CRISPR-Cas system is a prokaryotic immune system that confers resistance to foreign genetic elements such as those present within plasmids and phages that provide acquired immunity. RNA harboring the spacer sequence helps Cas (CRISPR-associated) proteins recognize and cut foreign pathogenic DNA. Other RNA-guided Cas proteins cut foreign RNA. CRISPR is found in approximately 50% of sequenced bacterial genomes and nearly 90% of sequenced archaea.

9.4: DNA Microarrays

A ***DNA microarray*** (also commonly known as a DNA chip or biochip) is a collection of microscopic DNA spots attached to a solid surface. Scientists use DNA microarrays to measure the expression levels of large numbers of genes simultaneously or to genotype multiple regions of a genome. Each DNA spot contains picomoles (10^{-12} moles) of a specific DNA sequence, known as *probes* (or *reporters* or *oligos*). These can be a short section of a gene or other DNA element that is used to hybridize a cDNA or cRNA (also called anti-sense RNA) sample (called *target*) under high-stringency conditions. Probe-target hybridization is usually detected and quantified by the detection of fluorophore-, silver-, or chemiluminescence-labeled targets to determine the relative abundance of nucleic acid sequences in the target. The original nucleic acid arrays were macroarrays approximately 9 cm × 12 cm and the first computerized image-based analysis was published in 1981. It was invented by Patrick O. Brown. Figure 9.4.1 shows a schematic of DNA Microarrays.

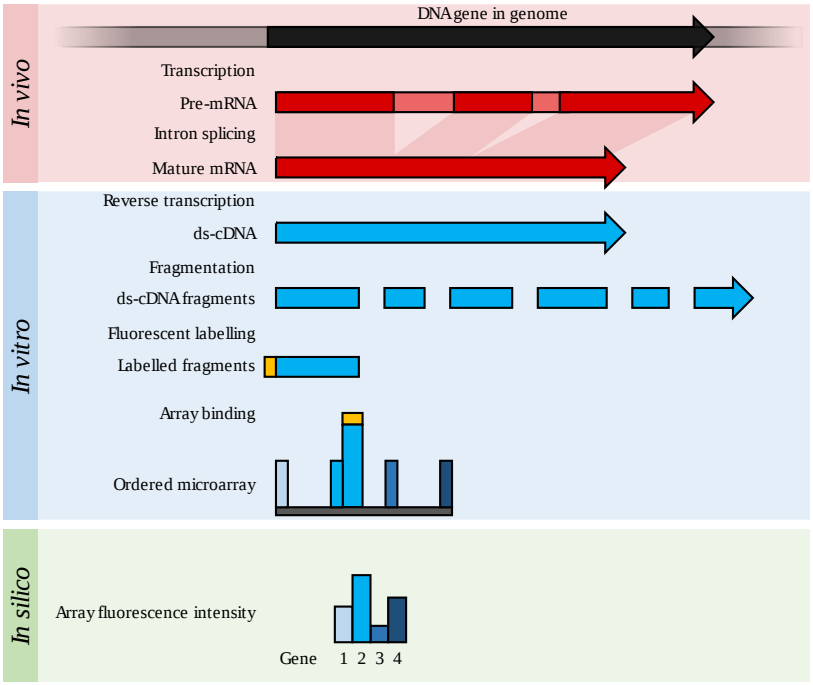


Figure 9.4.1: Schematic of DNA Microarrays. Image by Thomas Shafee

Within the organisms, genes are transcribed and spliced to produce mature mRNA transcripts (red). The mRNA is extracted from the organism and reverse transcriptase is used to copy the mRNA into stable ds-cDNA (blue). In microarrays, the ds-cDNA is fragmented and fluorescently labeled (orange). The labeled fragments bind to an ordered array of complementary oligonucleotides, and measurement of fluorescent intensity across the array indicates the abundance of a predetermined set of sequences. These sequences are typically specifically chosen to report on genes of interest within the organism's genome.

The core principle behind microarrays is a hybridization between two DNA strands, the property of complementary nucleic acid sequences to specifically pair with each other by forming hydrogen bonds between complementary nucleotide base pairs. A high number of complementary base pairs in a nucleotide sequence means tighter non-covalent bonding between the two strands. After washing off non-specific bonding sequences, only strongly paired strands will remain hybridized. Fluorescently labeled target sequences that bind to a probe sequence generate a signal that depends on the hybridization conditions (such as temperature),and washing after hybridization. The total strength of the signal, from a spot (feature), depends upon the amount of target sample binding to the probes present on that spot. Microarrays use relative quantitation in which the intensity of a feature is compared to the intensity of the same feature under a different condition, and the identity of the feature is known by its position. Figure 9.4.2s shows the hybridization of target and probe DNA on a microarray.

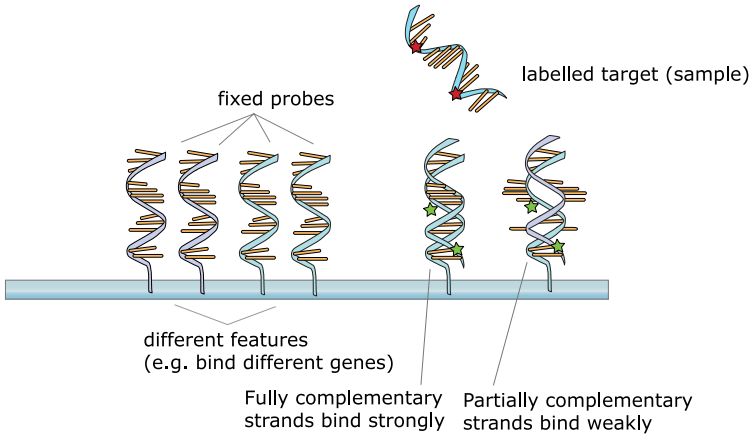


Figure 9.4.2: Hybridization of Target DNA with the Probe DNA During Microarray Analysis. Image by Squidonius

Many types of arrays exist and the broadest distinction is whether they are spatially arranged on a surface or coded beads:

- The traditional solid-phase array is a collection of orderly microscopic "spots", called features, each with thousands of identical and specific probes attached to a solid surface, such as glass, plastic, or silicon biochip (commonly known as a *genome chip*, *DNA chip* or *gene array*). Thousands of these features can be placed in known locations on a single DNA microarray.
- The alternative bead array is a collection of microscopic polystyrene beads, each with a specific probe and a ratio of two or more dyes, which do not interfere with the fluorescent dyes used on the target sequence.

DNA microarrays can be used to detect DNA (as in comparative genomic hybridization), or detect RNA (most commonly as cDNA after reverse transcription) that may or may not be translated into proteins. The process of measuring gene expression via cDNA is called expression analysis or expression profiling.

9.4.1: Fabrication

Microarrays can be manufactured in different ways, depending on the number of probes under examination, costs, customization requirements, and the type of scientific question being asked. Arrays from commercial vendors may have as few as 10 probes or as many as 5 million or more micrometer-scale probes.

9.4.2: Spotted vs. *in situ* synthesized arrays

Play media

Microarrays can be fabricated using a variety of technologies, including printing with fine-pointed pins onto glass slides, photolithography using pre-made masks, photolithography using dynamic micromirror devices, ink-jet printing, or electrochemistry on microelectrode arrays.

In *spotted microarrays*, the probes are oligonucleotides, cDNA, or small fragments of PCR products that correspond to mRNAs. The probes are synthesized before deposition on the array surface and are then "spotted" onto glass. A common approach utilizes an array of fine pins or needles controlled by a robotic arm that is dipped into wells containing DNA probes and then depositing each probe at designated locations on the array surface. The resulting "grid" of probes represents the nucleic acid profiles of the prepared probes and is ready to receive complementary cDNA or cRNA "targets" derived from experimental or clinical samples. This technique is used by research scientists around the world to produce "in-house" printed microarrays from their labs. These arrays may be easily customized for each experiment, because researchers can choose the probes and printing locations on the arrays, synthesize the probes in their labs (or collaborating facility), and spot the arrays. They can then generate their own labeled samples for hybridization, hybridize the samples to the array, and finally scan the arrays with their equipment. This provides a relatively low-cost microarray that may be customized for each study and avoids the costs of purchasing often more expensive commercial arrays that may represent vast numbers of genes that are not of interest to the investigator. Publications exist which indicate in-house spotted microarrays may not provide the same level of sensitivity compared to commercial oligonucleotide arrays, possibly owing to the small batch sizes and reduced printing efficiencies when compared to industrial manufacturers of oligo arrays.

In *oligonucleotide microarrays*, the probes are short sequences designed to match parts of the sequence of known or predicted open reading frames. Although oligonucleotide probes are often used in "spotted" microarrays, the term "oligonucleotide array" most often refers to a specific technique of manufacturing. Oligonucleotide arrays are produced by printing short oligonucleotide sequences designed to represent a single gene or family of gene splice-variants by synthesizing this sequence directly onto the array surface instead of depositing intact sequences. Sequences may be longer (60-mer probes such as the Agilent design) or shorter (25-mer probes produced by Affymetrix) depending on the desired purpose; longer probes are more specific to individual target genes, shorter probes may be spotted in higher density across the array and are cheaper to manufacture. One technique used to produce oligonucleotide arrays includes photolithographic synthesis (Affymetrix) on a silica substrate where light and light-sensitive masking agents are used to "build" a sequence one nucleotide at a time across the entire array. Each applicable probe is selectively "unmasked" before bathing the array in a solution of a single nucleotide, then a masking reaction takes place and the next set of probes are unmasked in preparation for a different nucleotide exposure. After many repetitions, the sequences of every probe become fully constructed. More recently, Maskless Array Synthesis from NimbleGen Systems has combined flexibility with large numbers of probes. Figure 9.4.3 shows a diagram of a typical dual-color microarray experiment.

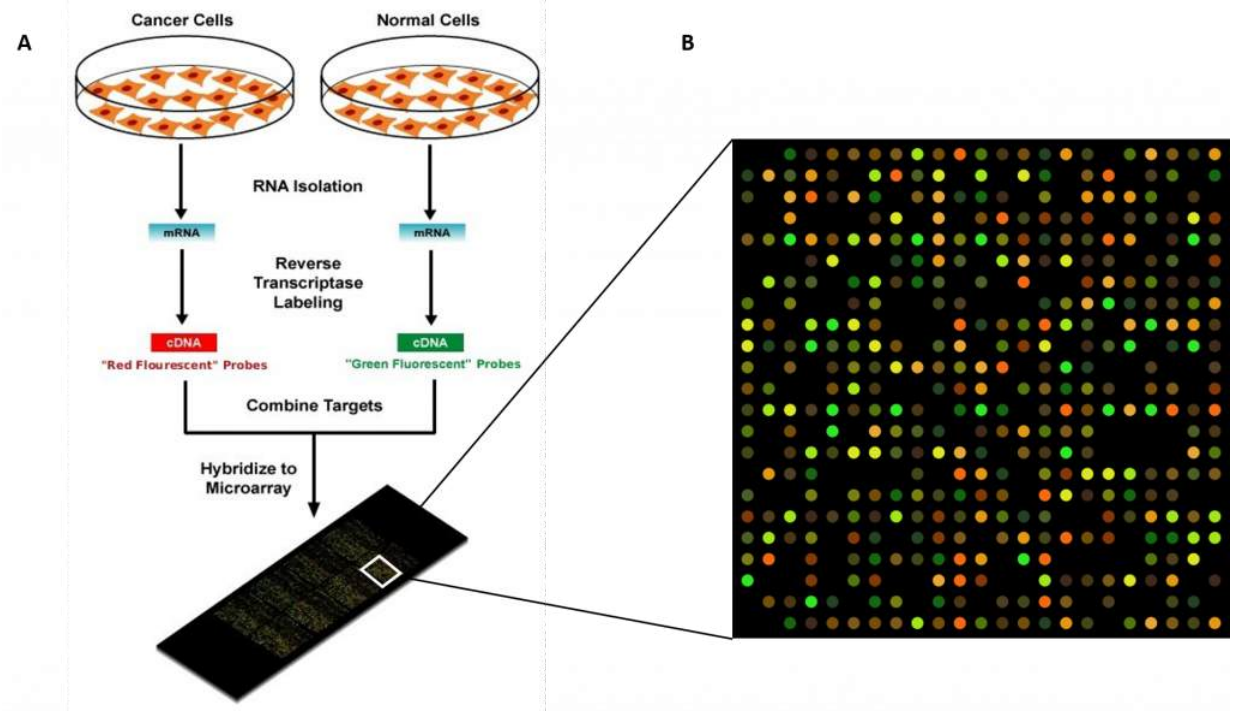


Figure 9.4.3: Diagram of a typical dual-color microarray experiment. Within a dual-color microarray, the probe DNA is typically hybridized with cDNA prepared from two different samples, each labeled with a different fluorescent probe. The analysis will yield green fluorescence for one sample that upregulates gene expression, whereas the other sample tagged with a red fluorescence marker will indicate that the other condition elicits gene expression at that location. Yellow indicates gene expression in both samples. Image A modified from Larssono and Image B from Guillaume Paumier

Two-color microarrays or ***two-channel microarrays*** are typically hybridized with cDNA prepared from two samples to be compared (e.g. diseased tissue versus healthy tissue) and that are labeled with two different fluorophores. Fluorescent dyes commonly used for cDNA labeling include Cy3, which has a fluorescence emission wavelength of 570 nm (corresponding to the green part of the light spectrum), and Cy5 with a fluorescence emission wavelength of 670 nm (corresponding to the red part of the light spectrum). The two Cy-labeled cDNA samples are mixed and hybridized to a single microarray that is then scanned in a microarray scanner to visualize the fluorescence of the two fluorophores after excitation with a laser beam of a defined wavelength. Relative intensities of each fluorophore may then be used in ratio-based analysis to identify up-regulated and down-regulated genes.

Oligonucleotide microarrays often carry control probes designed to hybridize with RNA spike-ins. The degree of hybridization between the spike-ins and the control probes is used to normalize the hybridization measurements for the target probes. Although absolute levels of gene expression may be determined in the two-color array in rare instances, the relative differences in expression among different spots within a sample and between samples are the preferred method of data analysis for the two-color system.

Examples of providers for such microarrays include Agilent with their Dual-Mode platform, Eppendorf with their DualChip platform for colorimetric Silverquant labeling, and TeleChem International with Arrayit.

In ***single-channel microarrays*** or ***one-color microarrays***, the arrays provide intensity data for each probe or probe set indicating a relative level of hybridization with the labeled target. However, they do not truly indicate abundance levels of a gene but rather a relative abundance when compared to other samples or conditions when processed in the same experiment. Each RNA molecule encounters protocol and batch-specific bias during the amplification, labeling, and hybridization phases of the experiment making comparisons between genes for the same microarray uninformative. The comparison of two conditions for the same gene requires two separate single-dye hybridizations. Several popular single-channel systems are the Affymetrix "Gene Chip", Illumina "Bead Chip", Agilent single-channel arrays, the Applied Microarrays "CodeLink" arrays, and the Eppendorf "DualChip & Silverquant". One strength of the single-dye system lies in the fact that an aberrant sample cannot affect the raw data derived from other samples, because each array chip is exposed to only one sample (as opposed to a two-color system in which a single low-quality sample may drastically impinge on overall data precision even if the other sample was of high quality). Another benefit is that data are more easily compared to arrays from different experiments as long as batch effects have been accounted for.

9.5: In Situ Hybridization

In situ hybridization (ISH) is a type of hybridization that uses a labeled complementary DNA, RNA, or modified nucleic acids strand (i.e., probe) to localize a specific DNA or RNA sequence in a portion or section of tissue (*in situ*) or if the tissue is small enough (e.g., plant seeds, *Drosophila* embryos), in the entire tissue (whole mount ISH), in cells, and in circulating tumor cells (CTCs). This is distinct from immunohistochemistry, which usually localizes proteins in tissue sections.

In situ hybridization is used to reveal the location of specific nucleic acid sequences on chromosomes or in tissues, a crucial step for understanding the organization, regulation, and function of genes. The key techniques currently in use include *in situ* hybridization to mRNA with oligonucleotide and RNA probes (both radio-labeled and hapten-labeled), analysis with light and electron microscopes, whole mount *in situ* hybridization, double detection of RNAs and RNA plus protein, and fluorescent *in situ* hybridization to detect chromosomal sequences. DNA ISH can be used to determine the structure of chromosomes. Fluorescent DNA ISH (FISH) can, for example, be used in medical diagnostics to assess chromosomal integrity. RNA ISH (RNA *in situ* hybridization) is used to measure and localize RNAs (mRNAs, lncRNAs, and miRNAs) within tissue sections, cells, whole mounts, and circulating tumor cells (CTCs). *In situ* hybridization was invented by Mary-Lou Pardue and Joseph G. Gall.

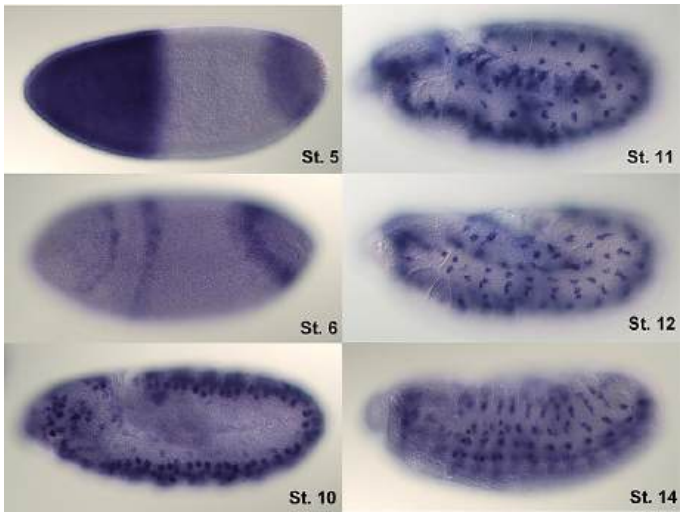


Figure 5.28 In Situ Hybridization of wild-type *Drosophila* embryos at different developmental stages for the RNA from a gene called hunchback. *Image by Nina*

For hybridization histochemistry, sample cells and tissues are usually treated to fix the target transcripts in place and to increase access to the probe. As noted above, the probe is either a labeled complementary DNA or, now most commonly, a complementary RNA (riboprobe). The probe hybridizes to the target sequence at elevated temperature, and then the excess probe is washed away (after prior hydrolysis using RNase in the case of unhybridized, excess RNA probe). Solution parameters such as temperature, salt, and/or detergent concentration can be manipulated to remove any non-identical interactions (i.e., only exact sequence matches will remain bound). Then, the probe that was labeled with either radio-, fluorescent- or antigen-labeled bases (e.g., digoxigenin) is localized and quantified in the tissue, using either autoradiography, fluorescence microscopy, or immunohistochemistry, respectively. ISH can also use two or more probes, labeled with radioactivity or other non-radioactive labels, to simultaneously detect two or more transcripts.

An alternative technology, branched DNA assay, can be used for RNA (mRNA, lncRNA, and miRNA) *in situ* hybridization assays with single molecule sensitivity without the use of radioactivity. This approach (e.g., ViewRNA assays) can be used to visualize up to four targets in one assay, and it uses patented probe design and bDNA signal amplification to generate sensitive and specific signals. Samples (cells, tissues, and CTCs) are fixed, then treated to allow RNA target accessibility (RNA un-masking). Target-specific probes hybridize to each target RNA. Subsequent signal amplification is predicated by specific hybridization of adjacent probes (individual oligonucleotides [oligos] that bind side by side on RNA targets). A typical target-specific probe will contain 40 oligonucleotides, resulting in 20 oligo pairs that bind side-by-side on the target for the detection of mRNA and lncRNA, and 2 oligos or a single pair for miRNA detection. Signal amplification is achieved via a series of sequential hybridization steps. A pre-amplifier molecule hybridizes to each oligo pair on the target-specific RNA, then multiple amplifier molecules hybridize to each pre-amplifier. Next, multiple-label probe oligonucleotides (conjugated to alkaline phosphatase or directly to fluorophores) hybridize to each amplifier molecule. A fully assembled signal amplification structure “Tree” has 400 binding sites for the label

probes. When all target-specific probes bind to the target mRNA transcript, an 8,000-fold signal amplification occurs for that one transcript. Separate but compatible signal amplification systems enable multiplex assays. The signal can be visualized using a fluorescence or brightfield microscope.

9.6: References

5.6 References

EMBL-EBI (2012) Next generation sequencing practical course. Buplished by EMBL-EBI. Available at: https://www.ebi.ac.uk/training/online/course/ebi-next-generation-sequencing-practical-course/how-take-course

Ghannam, M.G., and Varacallo, M. (2018) Biochemistry, Polymerase Chain Reaction (PCR) StatPearls Publishing. Available at: https://www.ncbi.nlm.nih.gov/books/NBK535453/

Kahn, N.T. (2018) The Emerging Role of bioinformatics in Biotechnology. J. Biotech. and Biomed. Science 1(3) ISSN: 2576-6694. Available at: https://openaccesspub.org/jbbs/article/803

Lee LYY, Izzard L and Hurt AC (2018) A Review of DNA Vaccines Against Influenza. Front. Immunol. 9:1568. doi: 10.3389/fimmu.2018.01568 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6046547/pdf/fimmu-09-01568.pdf

Molnar, C. and Gair, J. (2019) 10.1 Cloning and Genetic Engineering. Chapter in Concepts of Biology - 1st Candadian Edition. Available at: https://opentextbc.ca/biology/

Nidhi M. Bioinformatics: An Introduction. Open Acc Biostat Bioinform. 1(4). OABB.000522. 2018. DOI: 10.31031/OABB.2018.01.000522 https://pdfs.semanticscholar.org/f220/86467e2532106c8c616f03fc0a61aff9b3ea.pdf

Seto, D. (2010) Viral genomics and Bioinformatics. *Viruses* 2: 2587-2593. doi:10.3390/v2122587

Seto, J. (2019) Restriction Enzymes. Chapter in Biology OER, published by The OpenLab at City Tech. Available at: https://openlab.citytech.cuny.edu/bio-oer/analyzing-dna/restriction-enzymes/

Wikipedia contributors. (2019, July 8). Complementary DNA. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:18, September 8, 2019, from https://en.Wikipedia.org/w/index.php?title=Complementary_DNA&oldid=905281015

Wikipedia contributors. (2019, May 15). Genomic DNA. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:25, September 8, 2019, from https://en.Wikipedia.org/w/index.php?title=Genomic_DNA&oldid=897250028

Wikipedia contributors. (2019, July 10). DNA extraction. In *Wikipedia, The Free Encyclopedia*. Retrieved 05:05, September 9, 2019, from https://en.Wikipedia.org/w/index.php?title=DNA_extraction&oldid=905644602

Wikipedia contributors. (2018, May 9). Spin column-based nucleic acid purification. In *Wikipedia, The Free Encyclopedia*. Retrieved 05:26, September 9, 2019, from https://en.Wikipedia.org/w/index.php?title=Spin_column-based_nucleic_acid_purification&oldid=840437141

Wikipedia contributors. (2019, August 29). DNA sequencing. In *Wikipedia, The Free Encyclopedia*. Retrieved 05:31, September 9, 2019, from https://en.Wikipedia.org/w/index.php?title=DNA_sequencing&oldid=913023155

Wikipedia contributors. (2019, July 30). Sanger sequencing. In *Wikipedia, The Free Encyclopedia*. Retrieved 05:28, September 10, 2019, from https://en.Wikipedia.org/w/index.php?title=Sanger_sequencing&oldid=908599177

Wikipedia contributors. (2019, June 5). Artificial gene synthesis. In *Wikipedia, The Free Encyclopedia*. Retrieved 19:19, September 13, 2019, from https://en.Wikipedia.org/w/index.php?title=Artificial_gene_synthesis&oldid=900437337

Wikipedia contributors. (2019, September 16). Cloning vector. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:26, September 26, 2019, from https://en.Wikipedia.org/w/index.php?title=Cloning_vector&oldid=916010157

Wikipedia contributors. (2019, October 2). CRISPR. In *Wikipedia, The Free Encyclopedia*. Retrieved 04:12, October 8, 2019, from https://en.Wikipedia.org/w/index.php?title=CRISPR&oldid=919232229

Wikipedia contributors. (2019, September 28). DNA microarray. In *Wikipedia, The Free Encyclopedia*. Retrieved 04:53, October 8, 2019, from https://en.Wikipedia.org/w/index.php?title=DNA_microarray&oldid=918427523

Wikipedia contributors. (2019, September 22). In situ hybridization. In *Wikipedia, The Free Encyclopedia*. Retrieved 04:59, October 8, 2019, from https://en.Wikipedia.org/w/index.php?title=In_situ_hybridization&oldid=917168202

CHAPTER OVERVIEW

10: Lipids

10.1: Introduction to lipids
10.2: Lipids Aggregates in Water - Micelles and Liposomes
10.3: Membrane Bilayer and Monolayer Assemblies - Structures and Dynamics
10.4: Working with Lipids
10.5: Problems

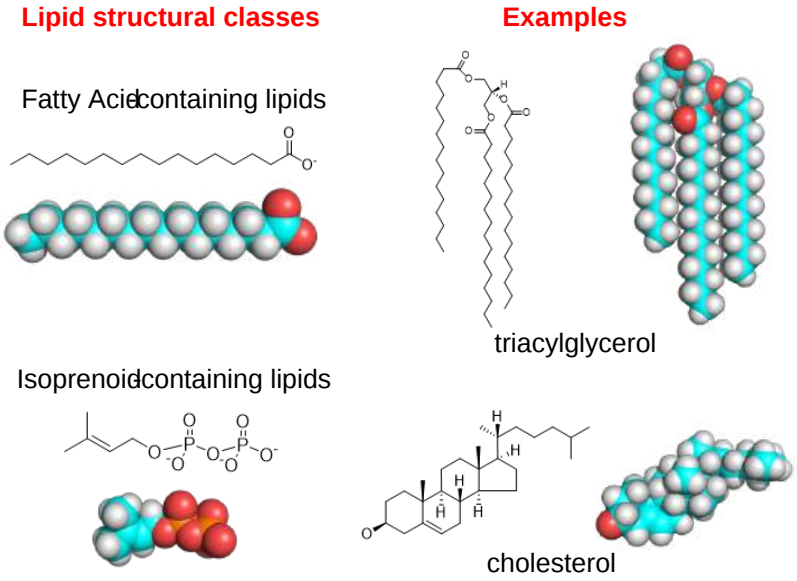
10.1: Introduction to lipids

10.1.1: Introduction to Lipids

(Thanks to Rebecca Roston for providing a cohesive organizational framework and image templates)

Lipids are organic molecule molecules that are soluble in organic solvents, such as chloroform/methanol, but sparingly soluble in aqueous solutions. These solubility properties arise since lipids are mostly hydrophobic. One type, triglycerides, is used for energy storage since they are highly reduced and get oxidized to release energy. Their hydrophobic nature allows them to pack efficiently through self-association in an aqueous environment. Triglycerides are also used for insulation since they conduct heat poorly, which is good if you live in a cold climate but bad if you wish to dissipate heat in a hot one. Triglycerides also offer padding and mechanical protection from shocks (think walruses). Another type of lipid forms membrane bilayers, which separate cellular contents from the outside environment, or separate intracellular compartments (organelles) from the cytoplasm. Some lipids are released from cells to signal other cells to change to specific stimuli in a process called cell signaling. From a more molecular perspective, lipids can act as cofactors for enzymes, pigments, antioxidants, and water repellents. As we saw with proteins, lipid structure mediates their function. So let's probe their structures.

Lipids can be split into structural classes in a variety of ways. An earlier classification divided them into those that release fatty acids on based-catalyzed hydrolysis to form soaps in a saponification reaction and those that don't. A much better and broader classification is based on if the lipids contain fatty acids or isoprenoids, as shown in Figure 10.1.1 below.



10.1.1.1:

Figure 10.1.1: Fatty acids and isoprenoid lipids

The nonpolar chains of the fatty acid are drawn in the figure above in the lowest energy zig-zag fashion as we saw when we discussed the main chain conformation of proteins (Chapter 4.1). In that chapter, we started with the exploration of a long 12 C chain carboxylic acid, dodecanoic acid. In the lowest energy conformation, the dihedral angles are all $\pm$ 180^0 to minimize torsional strain in the molecule. Rotation around one C-C bond can produce a gauche form, which introduces a kink into the chain as shown in Figure 10.1.2

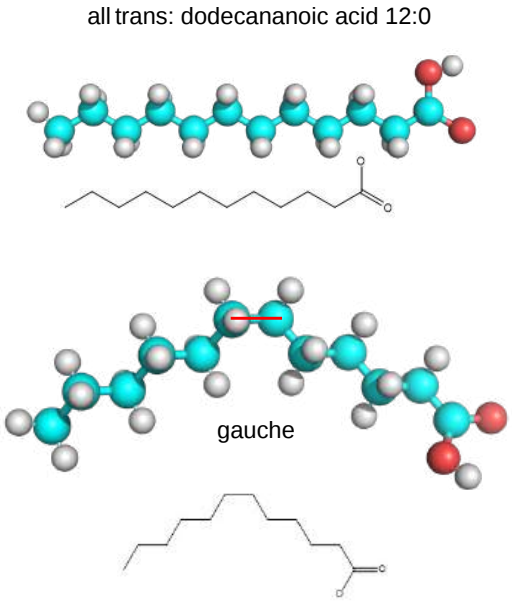


Figure 10.1.2: Trans and Gauche bonds in a fatty acids

10.1.2: Fatty Acids

Fatty acids can be free or covalently linked by ester or amide link to a base molecule like triglycerides or membrane lipids. The key principle that we learned with our study of proteins, that structure determines function, also applies to lipids. The figure below shows three different types of molecules, a free fatty acid, a wax with an esterified fatty acid, and a glycolipid with a fatty acid connected by an amide link in another type of lipid (glycosphingolipid). Each has different properties leading to different functions. Waxes for instance are very nonpolar and water-insoluble. They are amorphous solids at room temperature but, depending on their structure, can easily melt to form high-viscosity liquids. They are used as a coating on the surfaces of leaves to help prevent water loss. The glycolipids (glyco- as it contains a monosaccharide group) are constituents of membrane bilayers. Figure 10.1.3 shows the general structures of fatty acid-containing waxes and other lipids.

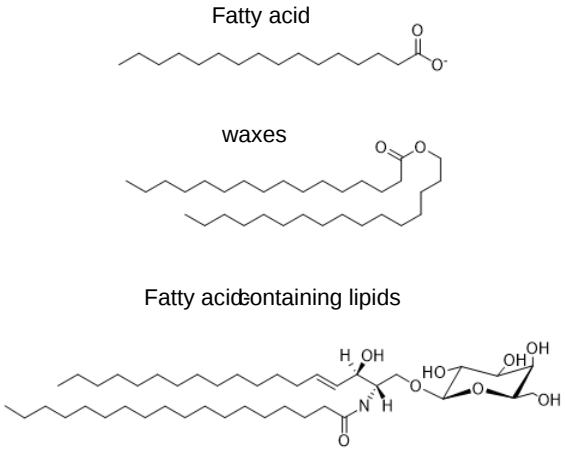


Figure 10.1.3: Fatty acids in waxes and glycolipids

Fatty acids vary in length, usually contain an even number of carbon atoms, and can be saturated (have no double bonds in the acyl chain), or unsaturated (with either one -monounsaturated - or multiple - polyunsaturated - cis double bond(s)). The double bonds are NOT conjugated as they are separated by a methylene (-CH_2-) spacer. Fatty acids can be named in many ways.

- symbolic name: given as x:y $^{\Delta\ a,b,c}$ where x is the number of carbon atoms in the chain, y is the number of double bonds, and a, b, and c are the positions of the start of the double bonds counting from C1 - the carboxyl carbon. Double bonds are usually cis (Z).
- systematic name using IUPAC nomenclature. The systematic name gives the number of carbon atoms in the chain (e.g. hexadecanoic acid for 16:0). If the fatty acid is unsaturated, the base name reflects the number of double bonds (e.g. octadecenoic acid for 18:1 $^{\Delta\ 9}$ and octadecatrienoic acid for 18:3$^{\Delta\ 9,12,15}$).
- common name: (e.g. oleic acid, 18:1$^{\Delta 9}$), which is found in high concentration in olive oil)

They can be named most easily with a symbolic name. Figure 10.1.4 shows examples of fatty acids and their symbolic names.

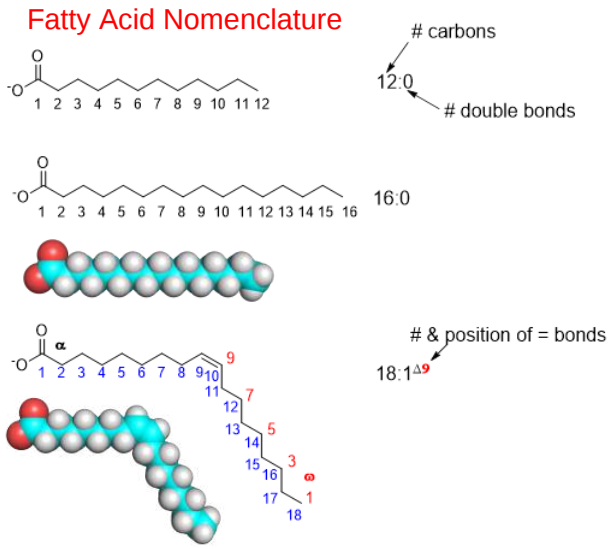


Figure 10.1.4: Some fatty acids and their symbolic names.

There is an alternative to the symbolic representation of fatty acids, in which the carbon atoms are numbered from the distal end (the n or ω end) of the acyl chain (the opposite end from the alpha carbon). Hence 18:3 $^{\Delta\ 9,12,15}$ could be written as 18:3 (ω -3) or 18:3 (n -3), where the terminal C is numbered one and the first double bond starts at C3.

The most common saturated fatty acids found in biochemistry textbooks are listed in the table below. Note how the melting point increases with the length of the hydrocarbon chain. This arises from increasing noncovalent induced dipole-induced dipole attractions between the long chains. Heat must be added to lessen these attractions to allow melting. Table 10.1.1 below shows examples of fatty acids and their symbolic names.

Symbolic	common name	systematic name	structure	mp(C)
12:0	Lauric acid	dodecanoic acid	$CH_3(CH_2)_{10}COOH$	44.2
14:0	Myristic acid	tetradecanoic acid	$CH_3(CH_2)_{12}COOH$	52
16:0	Palmitic acid	Hexadecanoic acid	$CH_3(CH_2)_{14}COOH$	63.1
18:0	Stearic acid	Octadecanoic acid	$CH_3(CH_2)_{16}COOH$	69.6
20:0	Arachidic aicd	Eicosanoic acid	$CH_3(CH_2)_{18}COOH$	75.4

Table 10.1.1: Examples of fatty acids and their symbolic names

Table 10.1.2 below shows common unsaturated fatty acids. Arachidonic acid is an (ω -6) fatty acid while docosahexaenoic acid is an (ω -3) fatty acid. Note the decreasing melting point for the 18:X series with an increasing number of double bonds.

Symbol	common name	systematic name	structure	mp(C)
16:1$^{\Delta 9}$	Palmitoleic acid	Hexadecenoic acid	$CH_3(CH_2)_5CH{=}CH{-}(CH_2)_7COOH$	-0.5
18:1$^{\Delta 9}$	Oleic acid	9-Octadecenoic acid	$CH_3(CH_2)_7CH{=}CH{-}(CH_2)_7COOH$	13.4

Symbol	common name	systematic name	structure	mp(C)
$18:2^{\Delta 9,12}$	Linoleic acid	9,12 -Octadecadienoic acid	$CH_3(CH_2)_4(CH{=}CHCH_2)_2(CH_2)_6COOH$	-9
$18:3^{\Delta 9,12,15}$	α-Linolenic acid	9,12,15 -Octadecatrienoic acid	$CH_3CH_2(CH{=}CHCH2)3(CH2)6COOH$	-17
$20:4^{\Delta 5,8,11,14}$	arachidonic acid	5,8,11,14- Eicosatetraenoic acid	$CH3(CH2)_4(CH{=}CHCH_2)_4(CH_2)_2COOH$	-49
$20:5^{\Delta 5,8,11,14,17}$	EPA	5,8,11,14,17-Eicosapentaenoic-acid	$CH_3CH_2(CH{=}CHCH_2)_5(CH_2)_2COOH$	-54
$22:6^{\Delta 4,7,10,13,16,19}$	DHA	Docosohexaenoic acid	22:6w3	

Table 10.1.2: Common unsaturated fatty acids

Let's consider how the presence of double bonds in fatty acids influences their melting points. Figure 10.1.5 shows common variants of fatty acids each with 18 carbon atoms. Compare the Lewis structure and spacefill models below. What a difference a cis double bond makes!

The double bonds in fatty acids are cis (Z), which introduces a "permanent" kink into the chain, similar to the"temporary" kink in the saturated dodecanoic acid carboxylic acid with a single gauche bond (Figure 2). The kink is permanent since there is no rotation around the double bond unless it is broken (which can happen through photoisomerization reactions). The more double bonds, the greater the kinking. The more kinks, the less chance for van der Waals contacts between the acyl chains, and reduces induced-dipole-induced dipole interactions between the chains, leading to lowered melting points.

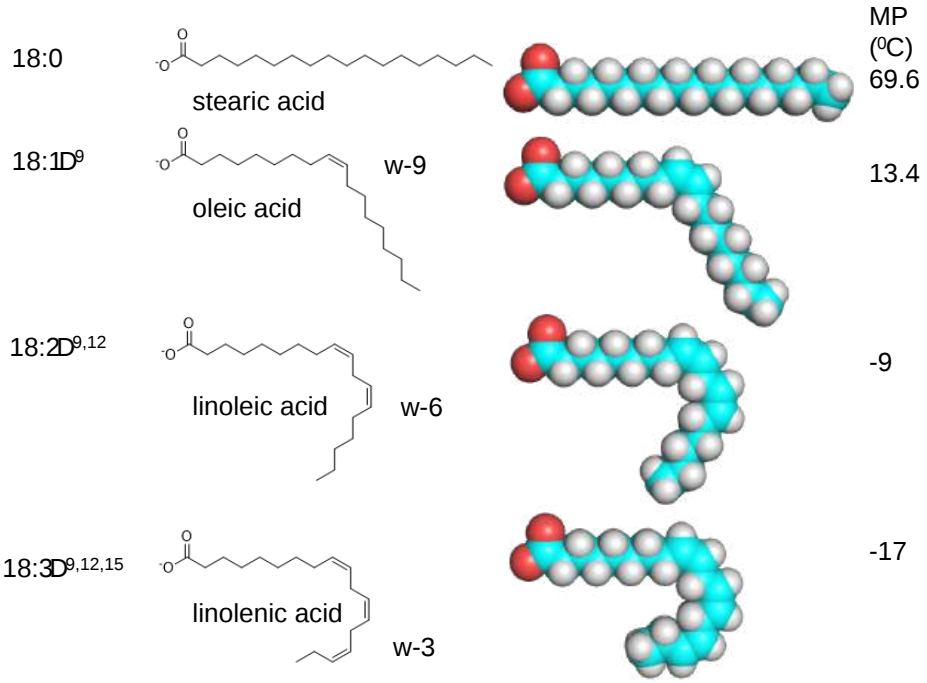


Figure 10.1.5: Fatty acids with 18 carbon atoms

Now in your mind, replace the cis double bond in oleic acid with a trans. You should now see a long "zig-zag" shaped molecule with no kinks. Trans fatty acids are rare in biology but are produced in the industrial partial hydrogenation of fats, which is done to decrease the number of double bonds and make the fats more solid-like and tastier while decreasing rancidity. These trans fatty acids would pack closer together and affect the structure and function of the lipid in a given environment (such as a membrane bilayer). Increased consumption of trans fatty acids is associated with an increased risk of cardiovascular disease.

Table 10.1.3 shows the percentages of fatty acids in different oils/fats

FAT	<16:0	16:1	18:0	18:1	18:2	18:3	20:0	22:1	22:2	.
Coconut	87	.	3	7	2	.	.	.	.	.

FAT	<16:0	16:1	18:0	18:1	18:2	18:3	20:0	22:1	22:2	.
Canola	3	.		11	13	10	.	7	50	2
Olive Oil	11	.	4	71	11	1	.	.	.	.
Butter-fat	50	4	12	26	4	1	2	.	.	.

Table 10.1.3: Percentages of fatty acids in different oils/fats

The fatty acid composition differs in different organisms:

- animals have 5-7% of fatty acids with 20-22 carbons, while fish have 25-30%
- animals have <1% of their fatty acids with 5-6 double bonds, while plants have 5-6% and fish 15-30%

As there are essential amino acids that cannot be synthesized by humans, there are also essential fatty acids that must be supplied by the diet. There are only two, one each in the n-6 and n-3 classes:

- n-6 class: α-linoleic acid (18:2 n-6, or $18:2^{\Delta 9,12}$) is a biosynthetic precursor of arachidonic acid (20:4 n-6 or $20:4^{\Delta 5,8,11,14)}$
- n-3 class: linolenic acid (18:3 n-3, or $18:3^{\Delta 9,12,15}$) is a biosynthetic precursor of eicosapentaenoic acid (EPA, 20:5 n-3 or $20:5^{\Delta 5,8,11,14,17}$) and to a much smaller extent, docosahexaenoic acid (DHA, 22:6 n-3 or $22:6^{\Delta 4,7,10,13,16,19}$).

These two fatty acids are essential since mammals cannot introduce double bonds in fatty acids beyond carbon 9. These essential precursor fatty acids are substrates for intracellular enzymes such as elongases and desaturases (to produce 20:4 n-6, 20:5 n-3, and 22:6 n-3 fatty acids), and beta-oxidation type enzymes in the endoplasmic reticulum and another organelle, the peroxisome. The peroxisome is involved in the oxidative metabolism of straight-chain and branched fatty acids, peroxide metabolism, and cholesterol/bile salt synthesis. Animals fed diets high in plant 18:2(n-6) fats accumulate 20:4(n-6) fatty acids in their tissues while those fed diets high in plant 18:3(n-3) accumulate 22:6(n-3). Animals fed diets high in fish oils accumulate 20:5 (EPA) and 22:6 (DHA) at the expense of 20:4(n-6).

Many studies support the claim that diets high in fish that contain abundant ω-3 fatty acids, in particular EPA and DHA, reduce inflammation and cardiovascular disease. ω-3 fatty acids are abundant in high-oil fish (salmon, tuna, sardines), and lower in cod, flounder, snapper, shark, and tilapia.

Some suggest that contrary to images of early hominids as hunters and scavengers of meat, human brain development might have required the consumption of fish which is highly enriched in arachidonic and docosahexaenoic acids. A large percentage of the brain consists of lipids, which are highly enriched in these two fatty acids. These fatty acids are necessary for the proper development of the human brain and in adults. Deficiencies in these might contribute to ADHD, dementia, and dyslexia. These fatty acids are essential in the diet and probably could not have been derived in high enough amounts from the eating of the brains of other animals. The mechanism for the protective effects of n-3 fatty acids in health will be explored later in the course when we discuss prostaglandins synthesis and signal transduction.

Aggregates of Fatty Acids in Aqueous Solution: Micelles

Structure determines both properties and function. It should be obvious that free, unesterified fatty acids are very (but not completely) insoluble in water. When added to water, they saturate the solution at a very low concentration and then phase separate out into aggregates called micelles. The structure of a micelle formed from dodecylsulfate, a common detergent with a sulfate instead of a carboxylate head group, is shown below. All of the nonpolar Cs and Hs of the long alkyl chains are "buried" and are not exposed to water, whereas the sulfate head groups are solvent exposed.

Figure 10.1.6 shows an interactive iCn3D model of image

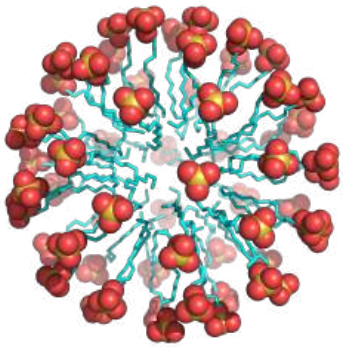


Figure 10.1.6: Sodium dodecylsulfate micelle (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...KECTiGaPPScED7

Fatty acids (carboxylates and sulfates) are **amphiphilic**, with a larger polar/charged head group that tapers down to a hydrophobic tail, forming a cone-like structure. This cone structure allows the packing of many of these single chains amphiphiles into a micelle, as shown in Figure 10.1.7.

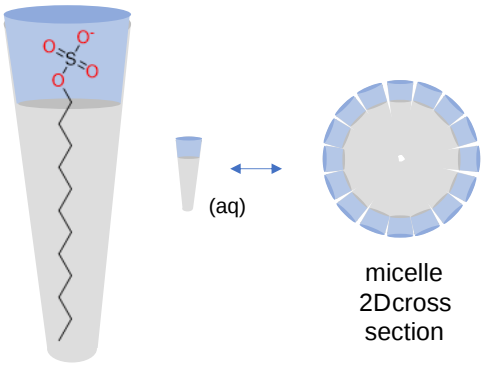


Figure 10.1.7: Cone-like representation of sodium dodecylsulfate free and in a micelle

Free fatty acids are transported in the cell and the blood not in micelles but by fatty acid-binding proteins. The most abundant protein in the blood, albumin, binds and transports fatty acids. Figure 10.1.8 shows an interactive iCn3D model of hexadecanoic acid bound to human albumin (1E7H). The fatty acids are shown in colored spacefill rendering.

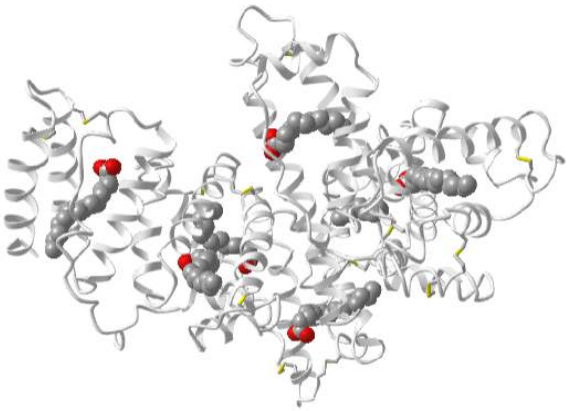


Figure 10.1.8: Hexadecanoic acid bound to human albumin (1E7H). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...CGUdMj11vNevF6

10.1.3: Waxes

You are familiar with ear wax and also the waxy surface of plants. Ear wax contains long-chain fatty acids (saturated and unsaturated) and alcohols derived from them. (They also contain isoprenoid derivatives like squalene and cholesterol). The waxy cuticle surface layers of plants contain very-long-chain fatty acids (C_{20}–C_{34}) and their derivatives including alkanes, aldehydes, primary and secondary alcohols, ketones, and esters, (They also contain isoprenoid derivatives as well). We will consider waxes as very long-chain fatty acids and their derivatives, as shown in Figure 10.1.9.

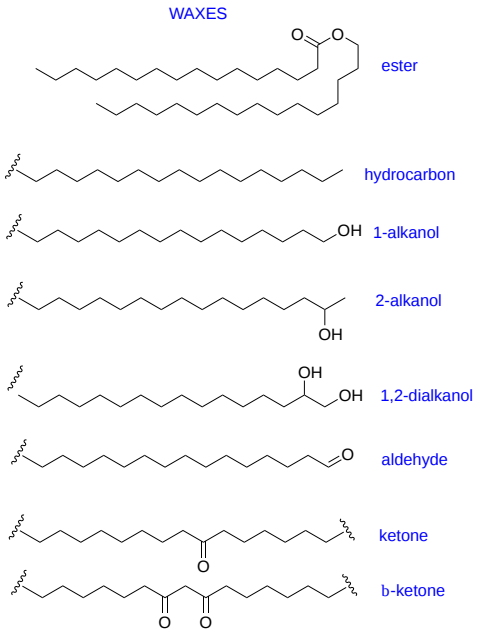


Figure 10.1.9: Waxes

These molecules are extremely nonpolar and as such makes great barriers preventing water loss through leaves and water penetration into the ear. This group of molecules clearly shows that the properties (insolubility, high melting point) and function (hydrophobic barrier/protection) arises from structure (very long chain carbon molecules, few electronegative atoms, and lack of C=C double bonds).

10.1.4: Fatty Acid-Containing Lipids

We can categorize these lipids based on function or structure (even though these are related).

Function

- storage lipids - triacylglycerols
- membrane lipids - many different lipids

Structure

- glycerolipids, which use glycerol as a backbone for fatty acid attachment
- sphingolipids, which use sphingosine as a backbone

The structures of glycerol and sphingosine are shown in Figure 10.1.10 Fatty acids are connected to these two "backbone" structures by either ester (mostly) or amide links.

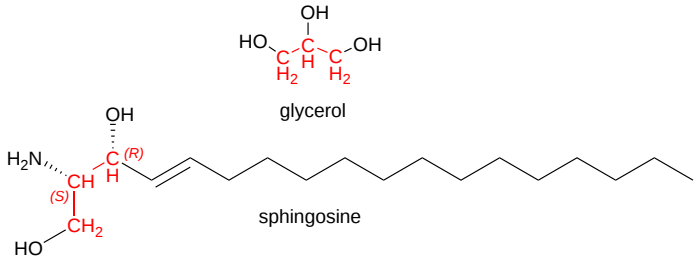


Figure 10.1.10 : Structures of glycerol and sphingosine

Let's explore the classes of fatty acid-containing lipids that use these two "backbone" structures.

10.1.5: Storage Lipids - Triacylglycerols (TAGs)

Triacylglycerols contain the majority of fatty acids in species that store fatty acids for energy. You will often see them named triglycerides or triacylglyceride, but this is a term used more in clinical chemistry and industry (and often in the media). Figure 10.1.11 shows a schematic diagram of the glycerol backbone with three fatty acids esterified to it. They are glycerolipids as they contain a glycerol base.

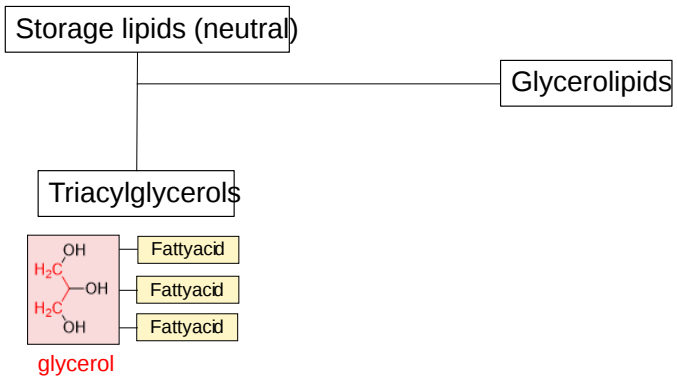


Figure 10.1.11: Structural features and nomenclature for triacylglycerols

Glycerol is not chiral but given the incredible diversity of fatty acids, glycerols likely have three different fatty acids esterified to them, making them chiral. If triacylglycerols contain predominately saturated fatty acids, they are solids at room temperature and are called **fats**. Those with multiple double bonds in the fatty acids are likely liquids at room temperature. These are called **oils**. Triacylglycerols are even more insoluble than fatty acids, which contain a polar and mostly charged carboxylate.

Figure 10.1.12 shows a triacylglycerol containing all 16:0 saturated fatty acids (left) and one containing all $16{:}2^{\Delta 9,12}$ polyunsaturated fatty acids. The figures were constructed with a specific set of dihedral angles to illustrate a point, that polyunsaturated fats in a triacylglycerol don't pack as tightly and have lower induced dipole-induced dipole attractions between the acyl chains than is possible with saturated fatty acyl chains. Hence the melting point of triacylglycerols containing polyunsaturated fatty acids is lower than for those with saturated ones.

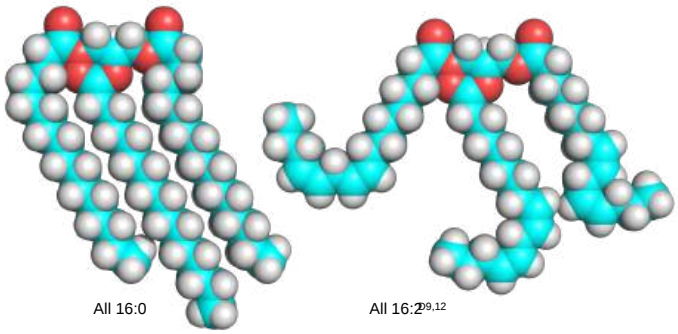


Figure 10.1.12: Triacylglycerol containing all 16:0 saturated fatty acids (left) and one containing all 16:$2^{\Delta 9,12}$ polyunsaturated fatty acids.

Let's repeat the key mantra: the structure of lipids determines their function. Consider the very insoluble triacylglycerols which are used as the predominant storage form of chemical energy in the body. In contrast to polysaccharides such as glycogen (a polymer of glucose), the carbon atoms in the acyl chains of the triacylglycerol are in a highly reduced state. The main source of energy to drive not only our bodies but also our society is obtained through oxidizing carbon-based molecules to carbon dioxide and water, in a reaction that is highly exergonic and exothermic. Sugars are already part way down the free energy "hill" since each carbon is partially oxidized. 9 kcal/mol (38 kJ/mol) can be derived from the complete oxidation of fats, in contrast to 4.5 kcal/mol (19 kJ/mol) from that of proteins or carbohydrates.

In addition, glycogen is highly hydrated. For every 1 g of glycogen, 2 grams of water is H-bonded to it. Hence it would take 3 times more weight to store the equivalent mass of carbohydrates compared to triacylglycerol, which is stored in anhydrous lipid "droplets" within cells. In addition, fats are more flexible, given the large number of conformations available to the acyl chain C-C bonds by simple rotation around the C-C bonds. Polysaccharides have monomeric cyclohexane-like chair structures and are much more rigid.

Another interesting point is that glucose and glycogen are found in cells, and they can be mobilized quickly for energy needs. Yes, fats are present in all cells as well (for example all cells have interior and exterior cell membranes). However, the major storage form of fat, triacylglycerols, is stored in special cells called adipocytes, which comprise adipose or fat tissue, and must be mobilized by signaling agents and transported in the form of fatty acids to cells for utilization. Again, triacylglycerols don't form membranes, which separate outside and inside aqueous environments. They are simply so insoluble that they phase-separate into lipid droplets. Their formation and structure are a bit more complex than that, though, and we will discuss lipid droplets more in the next section.

10.1.6: Membrane Lipids

Membranes are bilayers of amphiphilic lipids that separate the outside and inside (cytoplasm) aqueous environments in cells and the cytoplasm and interior contents of organelles within cells. In general, single-chain lipid amphiphiles form micelles. Amphiphilic membrane lipids typically have **two** nonpolar tails connected to a polar head, giving them a less conical and more cylindrical shape that disallows micelle formation while favoring bilayer formation. Figure 10.1.13 shows a short section of a bilayer membrane made from lipids with a polar (and charged) head group (phosphocholine) and two 16:0 chains. The red and blue spheres represent the O and N atoms of the head groups, which are sequestered to the exterior parts of the bilayer, where they would interact with water. The nonpolar 16:0 tails are shown in cyan, clearly illustrating the nonpolar nature of the interior of the bilayer.

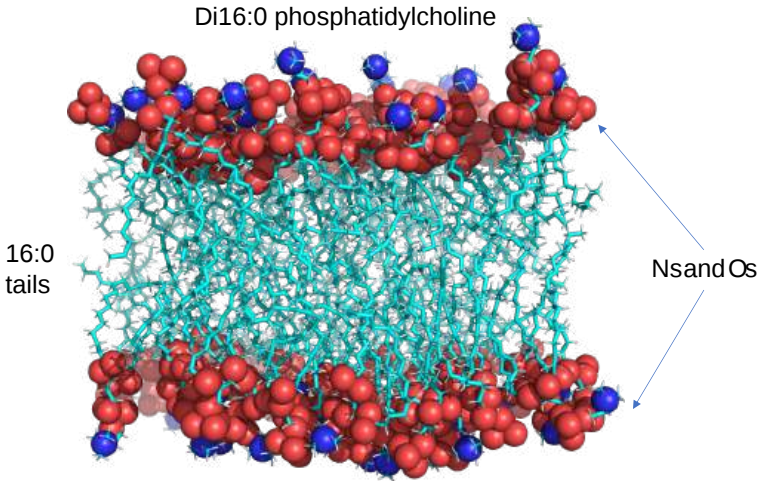
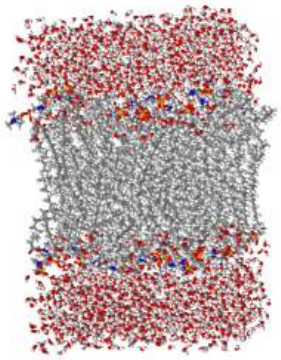


Figure 10.1.13: Bilayer membrane made from lipids with a polar (and charged) head group (phosphocholine) and two 16:0 chains.

Figure 10.1.14shows an interactive iCn3D model of a hydrated bilayer of the di16:0 phosphatidycholine bilayer.



NCBI iCn3D Figure 10.1.14 Hydrated di16:0 phosphatidycholine bilayer (Copyright; author via source).
Click the image for a popup.

We will consider two general types of fatty acid-containing membrane lipids, **glycerolipids**, with two fatty acids esterified to a glycerol base, and **sphingolipids** with one fatty acid in amide link to a different base, sphingosine. Sphingosine comes with its own built-in long alkyl chain that provides the "second" nonpolar chain. There are many different polar/charged head groups for these membrane lipids.

Now let's look in more detail at the double-chain amphiphiles comprising these membrane bilayers.

10.1.7: Glycerolipids

There are two main types of glycerolipids, glycerophospholipids, and glyceroglycolipids, which are the most common lipids in membranes.

Glycerophospholipids

Figure 10.1.15shows the structural features and nomenclature for glycerophospholipids.

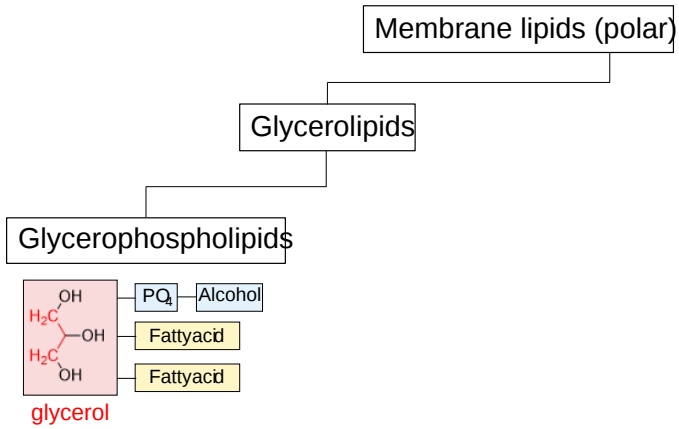


Figure 10.1.15: Structural features and nomenclature for glycerophospholipids

These lipids have enormous structural variability given the large number of different fatty acids (both saturated and unsaturated) and head groups that can be attached to a phosphate attached to the carbon 3 of glycerol. The structures of the most common glycerophospholipids are shown in Figure 10.1.16

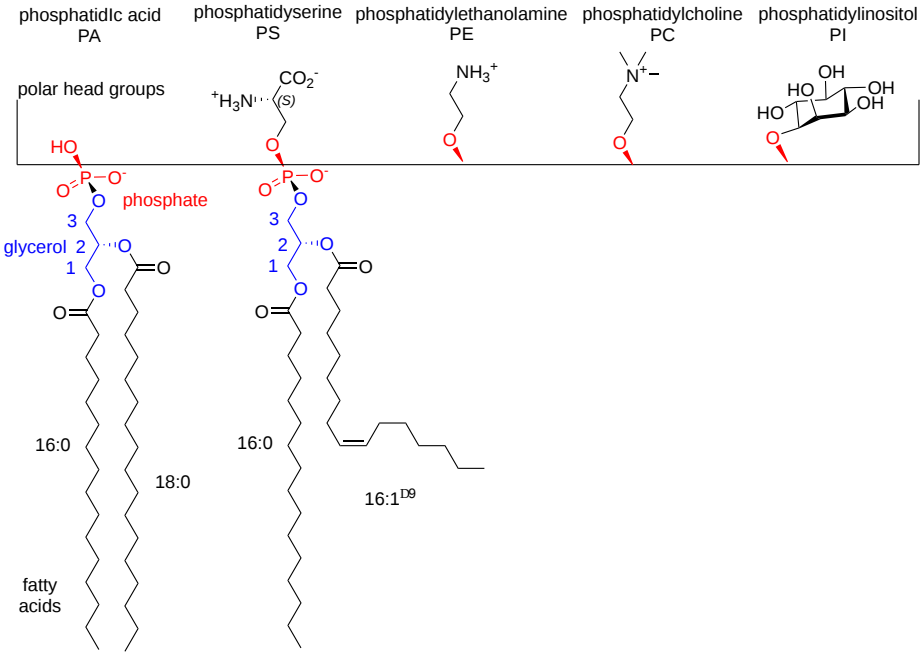


Figure 10.1.16: Common glycerophospholipids

Phosphatidylcholine (PC) has the common name lecithin while phosphatidylserine (PS) is called cephalin.

Note that the head groups all have charges since they all have a negatively charged phosphate. PS has two additional charged atoms which would effectively cancel out. PE has a charged amine but could become uncharged at pH values approaching its pKa. PC has a quaternary amine which is charged independent of pH, which would give PC a net 0 charge but with two discrete charges.

Glyceroglycolipids

These do not have a phosphate group attached to the oxygen on C3 of glycerol. Rather they have a mono- or oligosaccharide or, more loosely, a betaine group, each attached by an ether linkage to the glycerol C3 carbon. Figure 10.1.17Structural features and nomenclature for glyceroglycolipids.

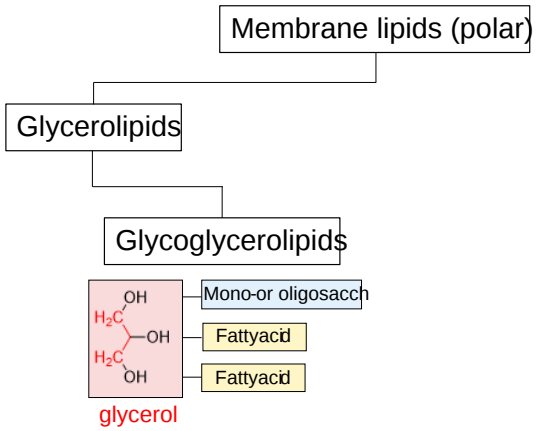


Figure 10.1.17: Structural features and nomenclature for glyceroglycolipids

Figure 10.1.18shows some examples of glyceroglycolipids.

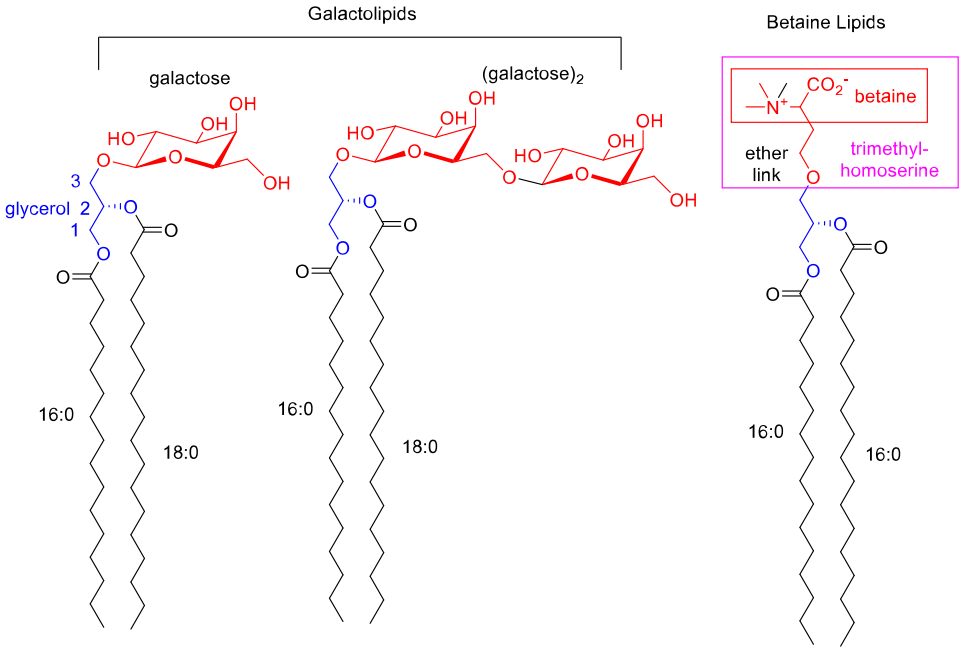


Figure 10.1.18: Examples of glycoglyerolipids.

Again there are an enormous number of different glycoglycerolipids, owing to the diversity of head groups and fatty acids esterified to glycerol at C1 and C2.

The above figure (right) shows an example that has no phosphate but also doesn't have a mono- or polysaccharide for a head group. Rather it has a betaine group. Betaine is the common name for trimethylglycine but is used for any N-trimethylated amino acids. Betaine glycerolipids are found in lower eukaryotic organisms (algae, fungi, and some protozoa), in photosynthetic bacteria, and in some spore-producing plants like ferns. Some would call these lipoamino acids.

10.1.8: Membrane Lipids - Sphingolipids

Sphingolipids contain a sphingosine backbone and a fatty acid linked through an amide bound. The simplest sphingolipids are the ceramides, which are double-chain amphiphiles but without further modification of the functional groups on the polar head of sphingosine. The structures of both sphingosine and ceramide are shown below in Figure 10.1.19

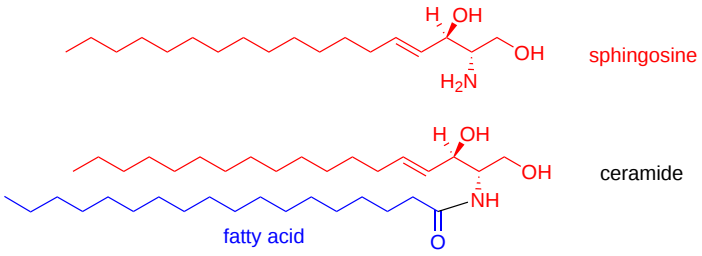


Figure 10.1.19 Structures of sphingosine and ceramide

Ceramides are abundant in the skin where they help provide a protective barrier. Skin creams also contain ceramides. These lipids make up only about 1% of body lipids and there are over 200 different ceramides in humans arising from the different fatty acids linked to sphingosine. They have a key role in cell signaling and can also affect cardiovascular health in ways that are just being appreciated. Since they are double-chain amphiphiles, they are found in membranes where they alter the properties of the bilayer (including membrane fluidity). Ceramides with long acyl tails (especially those with 16, 18, or 24 carbons) seem especially deleterious.

Phosphosphingolipids and glycosphingolipids

These membrane lipids have a ceramide base but also contain modifications at the polar functional groups of the sphingosine head. Let's consider the phosphosphingolipids and the glycosphingolipids together. These groups do not use glycerol as a base for the attachment of fatty acids and head groups. Rather they use molecule sphingosine. Figure 10.1.20 shows the structural features and nomenclature of sphingolipids.

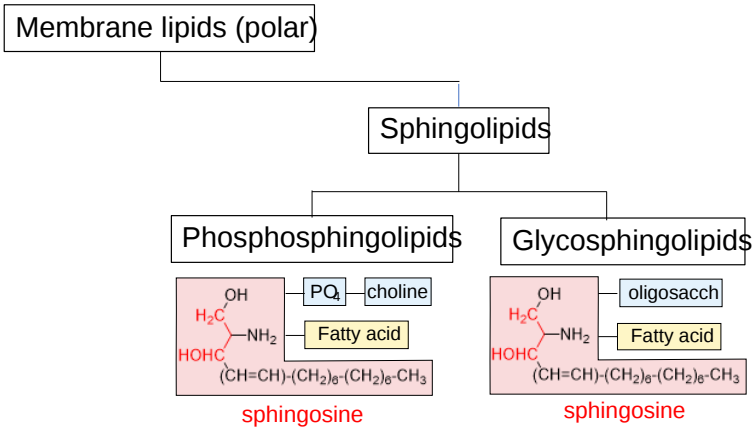


Figure 10.1.20: Structural features and nomenclature for sphingolipids

Examples of both classes are shown in Figure 10.1.21. Note the base sphingosine (in red) provides an amine to attach a fatty acid through an amide bond and an OH for attachment of the head group.

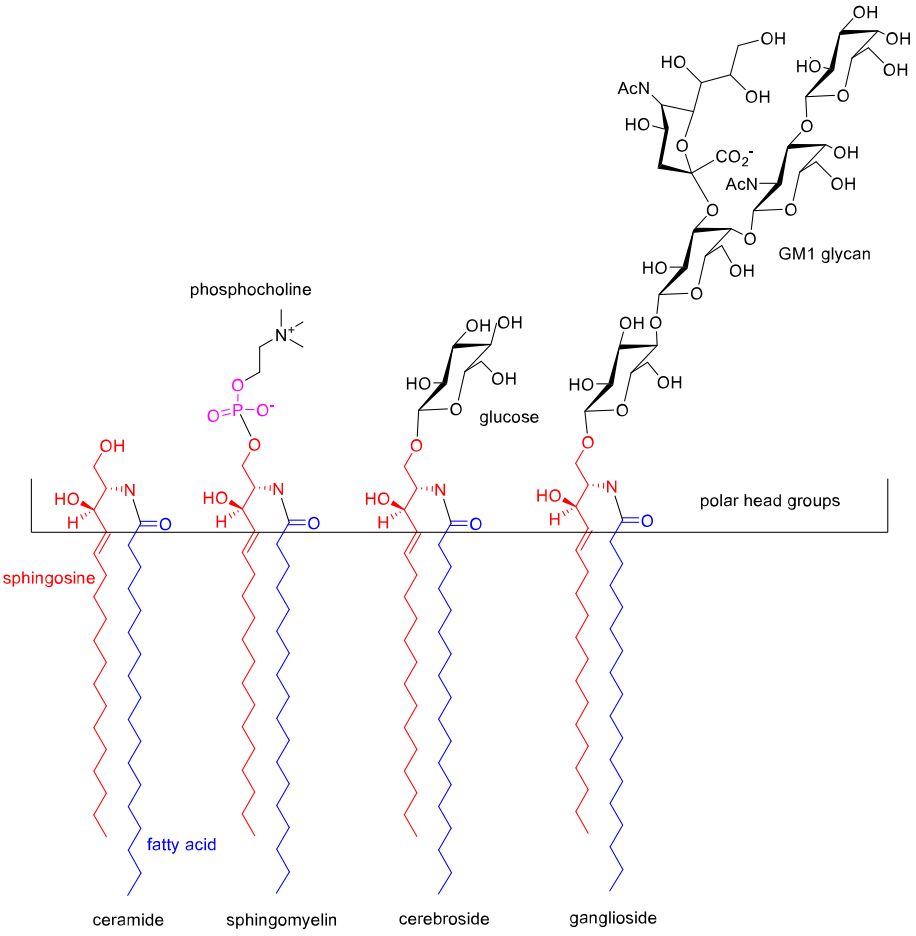


Figure 10.1.21: Examples of phosphosphingolipids and glycosphingolipids

Sugar-containing glycosphingolipids are **found largely in the outer face of plasma membranes.** The primary lipid of myelin, which coats neuronal axons and insulates them from loss of electrical signaling down the axon, is galactocerebroside

Figure 10.1.22shows a summary of all of the different types of fatty acid-containing lipids

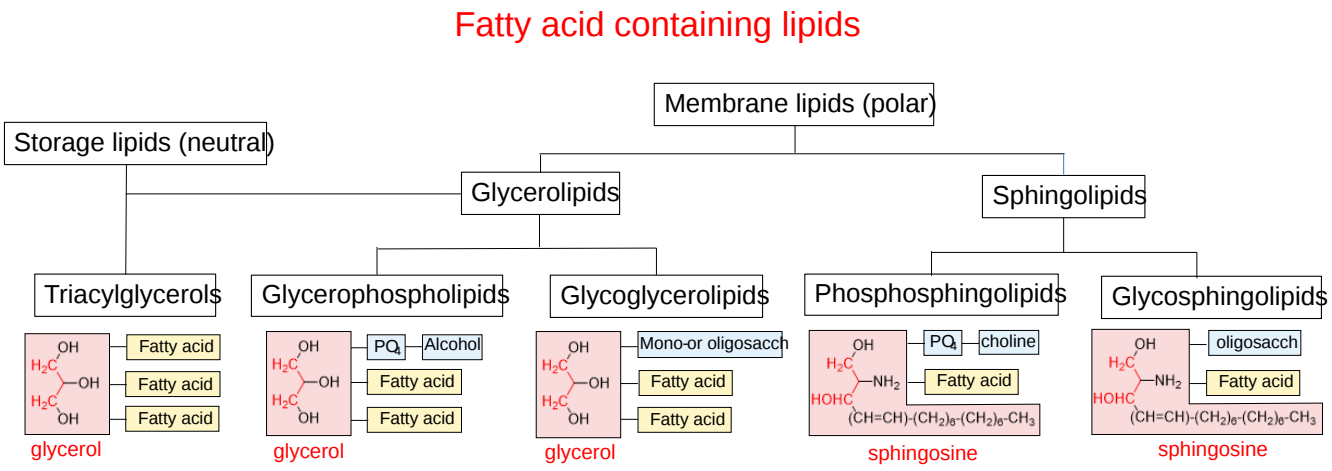


Figure 10.1.22: Summary of structural features and nomenclature for fatty acid-containing lipids

Over a 1000 different lipids are found in eukaryotic cells. This complexity has led to the development of an even more comprehensive classification system for lipids. In this system, lipids are given a very detailed as well as all-encompassing definition: "hydrophobic or amphipathic small molecules that may originate entirely or in part by carbanion-based condensations of

thioesters (fatty acyl, glycerolipids, glycerophospholipids, sphingolipids, saccharolipids, and polyketides) and/or by carbocation-based condensations of isoprene units (prenol lipids and sterol lipids)." Eight different categories of lipids are listed in the parentheses above. We will stick to the definition used throughout this chapter.

10.1.9: Shapes of membrane lipids

Let's look at the general shape of the double-chain amphiphiles that make bilayers. We saw the long-chain fatty or sulfate acids form conical structures which fit nicely together when they self-aggregate to form micelles. In contrast, membrane-forming double-chain amphiphiles have more cylindrical shapes that can't be fitted together in micelles but rather form a less curved bilayer structure, as shown in Figure 10.1.23 We will consider the variety of membrane structures in the next section.

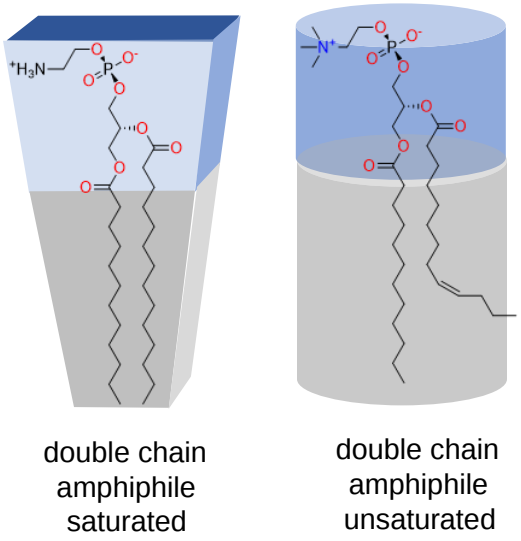


Figure 10.1.23: General shape of the double-chain amphiphiles that make bilayers

10.1.10: Triacylglyceride/Phospholipid Stereochemistry

Glycerol is an achiral molecule since C2 has two identical substituents, -CH_2OH. Glycerol in the body can be chemically converted to triglycerols and phospholipids (PL) which are chiral, and exist in one enantiomeric form. How can this be possible if the two CH_2OH groups on C2 of glycerol are identical? It turns out that even though these groups are stereochemically equivalent, we can differentiate them as described in the figure below. Let's replace the -CH_2OH in one of the end carbons with -CH_2OD. With this simple change, glycerol is now chiral. Look at the top half of Figure 10.1.24

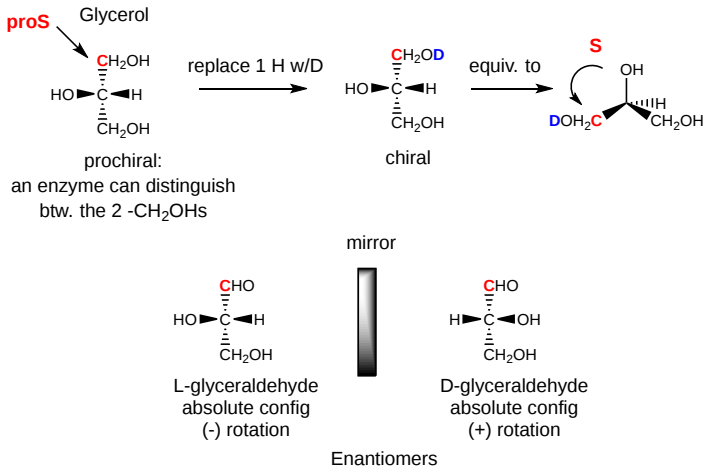


Figure 10.1.24: Potential chirality of Glycerol - ProS

Glycerol is oriented with the OH on C2 (the middle carbon) pointing to the left. The OH of the top carbon in this orientation, C1, is replaced with OD, where D is deuterium to make the molecule chiral (four different groups attached to C2). By rotating the

molecule such that the H on C2 points to the back, and assigning priorities to the other substituents on C2 (OH =1, $DOCH_2$ =2, and CH_2OH = 3), it can be seen that the resulting molecule is in the S configuration. We simply name the C1 carbon which we modified with deuterium as the **proS** carbon. Likewise, if we replaced the OH on C3 with OD, we will form the R enantiomer. Hence C3 is the proR carbon. This shows that in reality, we can differentiate between the two identical CH_2OH substituents. We say that glycerol is not chiral, but **prochiral**. (Think of this as glycerol has the potential to become chiral by modifying one of two identical substituents.)

In the bottom half of Figure 10.1.23 we can relate the configuration of glycerol above, (when OH on C2 is pointing to the left) to the absolute configuration of L-glyceraldehyde, a simple sugar (a polyhydroxyaldehyde or ketone), another 3C glycerol derivative. This molecule is chiral with the OH on C2 (the only chiral carbon) pointing to the left. It is easy to remember that any **L** sugar has the OH on the **L**ast chiral carbon pointing to the **L**eft. The enantiomer (mirror image isomer) of L-glyceraldehyde is D-glyceraldehyde, in which the OH on C2 points to the right. Biochemists use L and D for lipid, sugar, and amino acid stereochemistry, instead of the R, S nomenclature you used in organic chemistry. The stereochemical designation of all the sugars, amino acids, and glycerolipids can be determined from the absolute configuration of L- and D-glyceraldehyde.

Now let's see how an enzyme can take a prochiral molecule like glycerol and phosphorylate only one of the -CH_2OHs to make one specific isomer, glycerol-3-phosphate, a key intermediate in the biosynthesis of phosphatidic acid (PA), a glycerophospholipid, as well as chiral triacylglycerols, shown in Figure 10.1.25. The far left part of the pathway shows how the proR CH_2OH of glycerol is phosphorylated to produce one specific enantiomer, L-glycerol-3-phosphate. (The top part of the figure shows another way to make this molecule from glucose through the glycolytic pathway we will encounter in a future chapter.

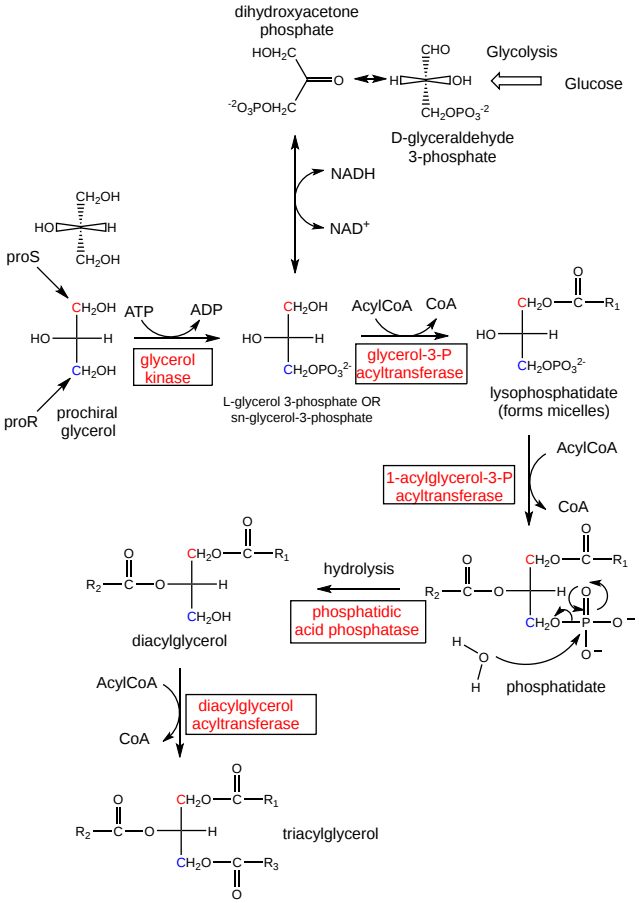
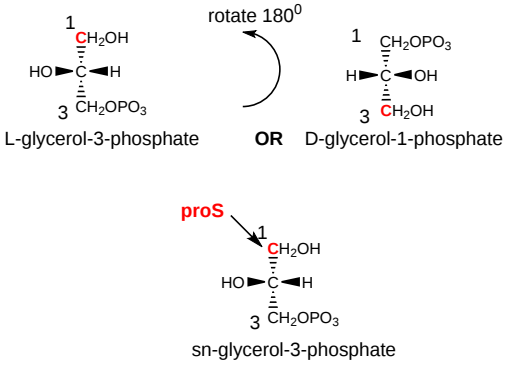


Figure 10.1.25: Synthesis of derivatives from prochiral glycerol

The first step (above figure) involves the phosphorylation of the OH on C3 by ATP (a phosphoanhydride similar in structure to acetic anhydride, an excellent acetylating agent) to produce the chiral molecule glycerol phosphate. Based on the absolute configuration of L-glyceraldehyde, and using this to draw glycerol (with the OH on C2 pointing to the left), we can see that the phosphorylated molecule can be named L-glycerol-3-phosphate. However, by rotating this molecule 180 degrees, without changing

the stereochemistry of the molecule, we don't change the molecule at all, but using the D/L nomenclature above, we would name the rotated molecule D-glycerol-1-phosphate. This is illustrated in Figure 10.1.26



Assign 1 position to group occupying the proS position of a prochiral center (glycerol). Orient molecule with C1 at top.

Figure 10.1.26: Stereospecific numbering system (sn) for glycerol

We can't give the same molecule two different names. Hence biochemists have developed the **stereospecific numbering system (sn)**, which assigns the 1-position of a prochiral molecule to the group occupying the proS position. The proS C1 is hence at the sn-1 position. With that designation, C2 is at the sn-2 position, and C3 is at the sn-3 position. Using this nomenclature, we can see that the chiral molecule described above, glycerol-phosphate, can be unambiguously named as sn-glycerol-3-phosphate. The hydroxyl substituent on the proR carbon was phosphorylated.

It is interesting to note that archaea use isoprenoid chains linked by ether bonds to *sn*-glycerol 1-phosphate in their synthetic pathways. As noted above, bacteria and eukaryotes use fatty acids attached by ester bonds to *sn*-glycerol 3-phosphate

The enzymatic phosphorylation of the proR CH_2OH of glycerol to form sn-glycerol-3-phosphate is illustrated in Figure 10.1.27. As we were able to differentiate the 2 identical CH_2OH substituents as containing either the proS or proR carbons, so can the enzyme. The enzyme can differentiate identical substituents on a prochiral molecule if the prochiral molecule interacts with the enzyme at three points. Another example of a prochiral reactants/enzyme system involves the oxidation of the prochiral molecule ethanol by the enzyme alcohol dehydrogenase, in which only the proR H of the 2 H's on C2 is removed. (We will discuss this later.)

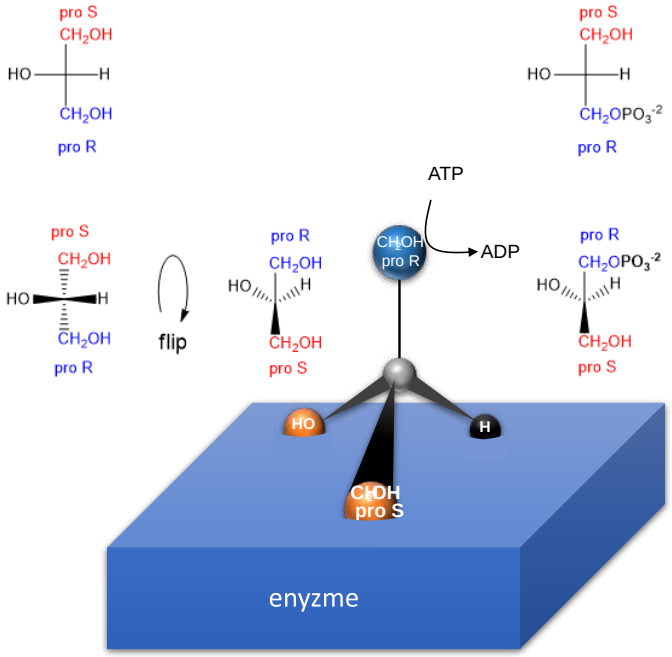


Figure 10.1.27: Enzymatic phosphorylation of the proR CH_2OH of glycerol to form sn-glycerol-3-phosphate

10.1.11: Isoprenoid-containing lipids

This is the last class of lipids we will consider. They do not contain fatty acids. Rather they contain isoprene, a small branched alkadiene, which can polymerize into larger molecules containing isoprene monomer to form isoprenoids, often called terpenes. Instead of using isoprene as the polymerization monomer, either dimethylallyl pyrophosphate (DMAPP) or isopentenylpyrophosphate (IPP) are used biologically.

Figure 10.1.28 shows how DMAPP and IPP (both containing 5Cs) are used in a polymerization reaction to form geranyl-pyrophosphate (C10), farnesyl pyrophosphate (C15) and geranyl-geranyl pyrophosphate (C20).

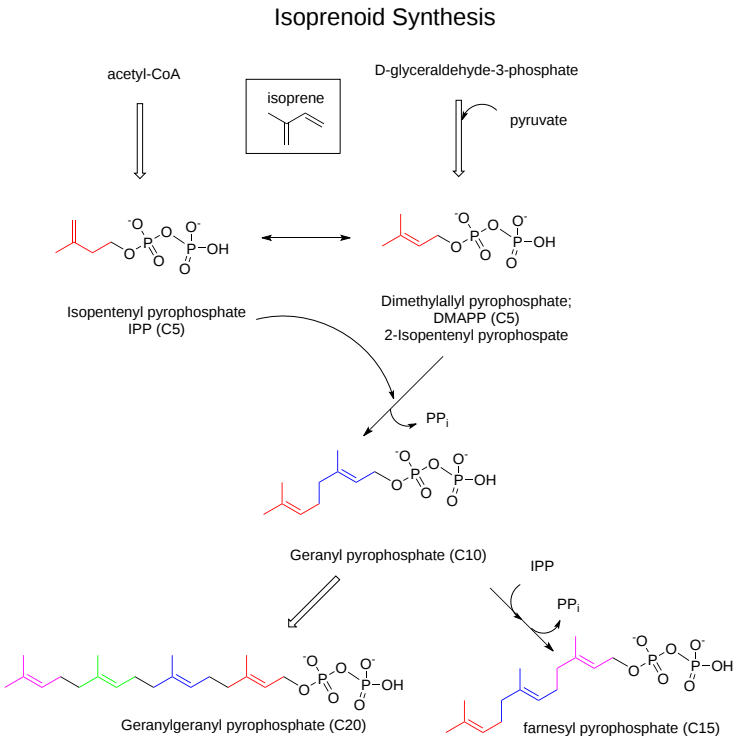


Figure 10.1.28: Synthesis of geranyl-pyrophosphate (C10), farnesyl pyrophosphate (C15) and geranyl-geranyl pyrophosphate (C20)

Many isoprenoid lipids are made from farnesyl pyrophosphate. For membrane purposes, the most important of these is cholesterol. Figure 10.1.29 shows an overview of the synthesis of cholesterol from two farnesyl pyrophosphates linking together in a "tail-to-tail" reaction to form squalene, a precursor of cholesterol. Each isoprene unit (5Cs) is shown in different colors to make it easier to see.

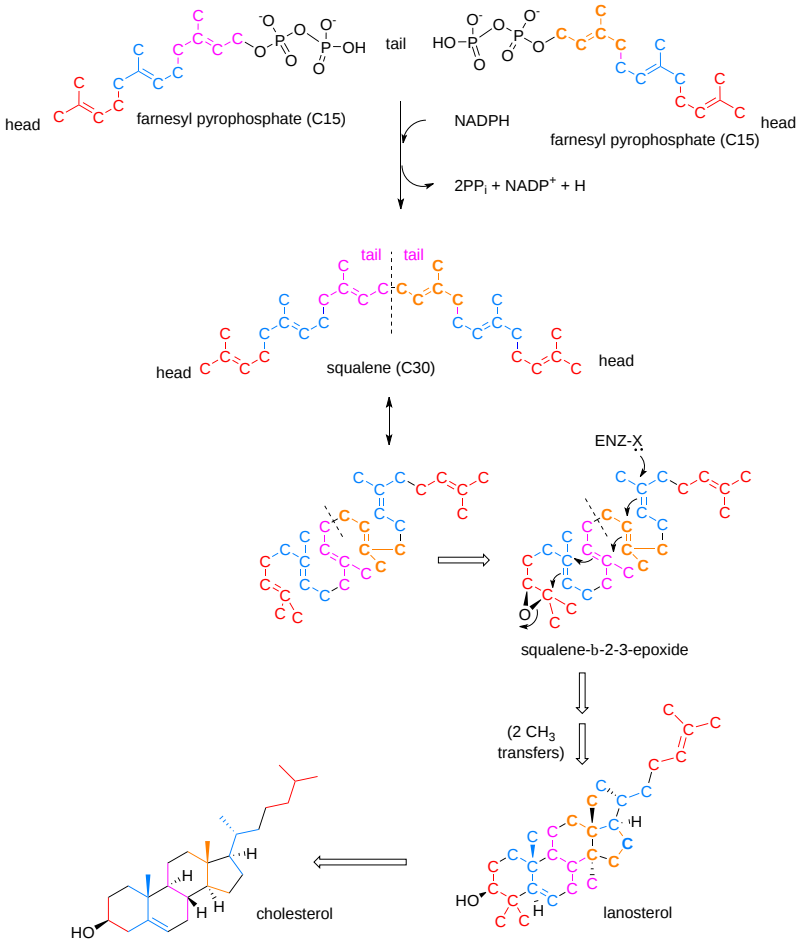


Figure 10.1.29: Synthesis of squalene from isoprene units

Other biologically important isoprenoid-containing vitamins are shown in Figure 10.1.30

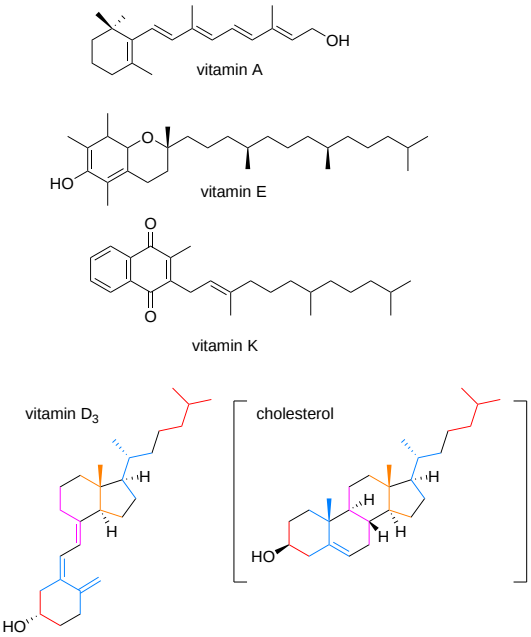
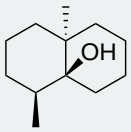


Figure 10.1.30: Isoprenoid vitamins

The smell of fresh rain

Everyone who has walked in the woods after a fresh rain knows the smell of a terpene called **geosmin**, "an Earthy-smelling substance" found in abundance in a group of substances collectively called petrichor, derived from the Greek petra (rock) and ichor (blood of gods). The structure of geosmin is shown below.



These substances are released when water falls or soils and rocks, and seeps into pores from which aerosols are released. Too hard of a rain will saturate pores in rocks and prevent the release. Some insects like the springtail are attracted to geosmin. The synthesis of geosmin by bacteria and blue-green algae is under the control of transcription factors which also affect bacterial spore formation. The attracted insects carry away the spores. Flies are repelled by geosmin, which is detected by a receptor at very low geosmin concentration. In contrast, the *Aedes aegypti* mosquito is attracted to geosmin which parallels the fact that the mosquito is not affected by bacterial toxins.

10.1.12: One last look at lipid structure and shapes

With the exclusion of waxes and triacylglycerols, the other lipids we have discussed, including the mostly planar molecule cholesterol, are amphiphilic. We have seen that single-chain fatty acids form micelles while lipids with two nonpolar chains and a polar/charged head group form bilayers. Given the relative size of the head group and the degree of unsaturation of the double bonds in fatty acids, the overall shapes of the membrane-forming lipids differ as illustrated below. They are arranged like dominos in the membrane based on their geometric volumes. Preferential clustering of identical types can cause local and extended changes in a prototypical bilayer structure.

Membranes and their components must be dynamic to enable all the functions and activities of a membrane. Ligands bind to membrane receptors (usually proteins), which can invaginate and pinch off to form an intracellular vesicle containing the receptor for processing. Likewise, vesicles can pinch off into the extracellular space. Cells must divide. Think of the membrane changes necessary for that! Also, consider that the length of cholesterol is half that of a typical double-chain amphiphile so it fits into just one of the bilayer lipid leaflets where it modulates lipid bilayer property. Figure 10.1.31 shows a series of lipids and their shape profiles.

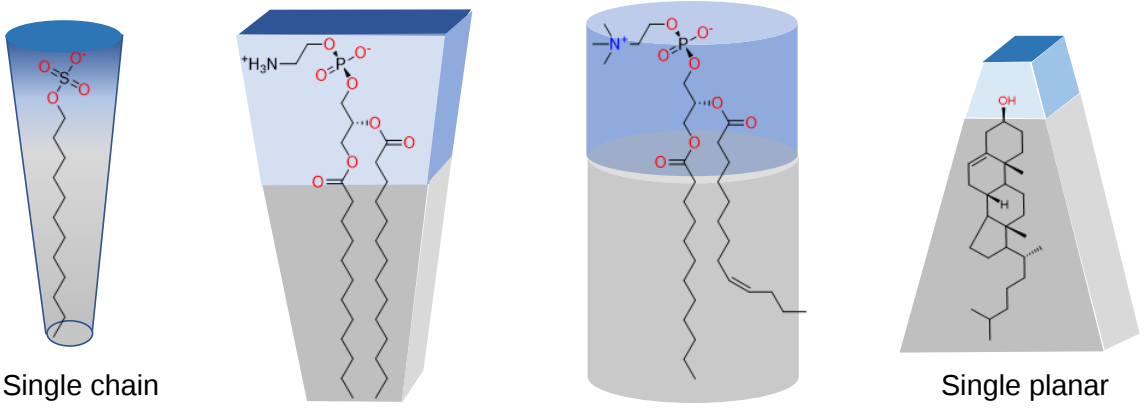


Figure 10.1.31: A series of lipids and their shape profiles

We will consider membranes in greater detail in section 10.3. Next, however, we will explore in more detail the properties of micelle and lipid droplet systems before addressing the more structurally complicated lipid bilayers.

10.1.12.1: End-of-Chapter Questions

? Exercise 10.1.1

x) A cell membrane has the ability to remodel in response to stress to promote membrane integrity. In the situations below, how could the membrane be remodeled to prevent damage? (I.e. what types of lipids could be added/removed to ensure homeostasis)

a) Increase in temperature

b)

x) What are the three classes of lipids? Explain their similarities and differences.

x) For each lipid below, name the type of lipid (membrane lipid, triacylglycerol, storage lipid, sphingolipid, wax, sterol, membrane glycerolipid, none of these), if it could be found in membrane, and if it is fatty acid or isoprene derived.

(Insert pics of lipids)

Answer

Add texts here. Do not delete this text first.

10.2: Lipids Aggregates in Water - Micelles and Liposomes

10.2.1: Single Chain Amphiphiles and Micelles

An understanding of lipids in simple solutions in the lab is incredibly helpful to understanding them *in vivo*. The same physical-chemical constraints would apply to the complex environment of the cell. What is different in the cell is that lipids are found in a cellular environment that is incredibly crowded with proteins that bind, synthesize, and breakdown lipids. Nevertheless, we can apply what we know from the test tube experiments to the cell.

To understand how molecules might react, it helps a bit to pretend you are a molecule and ask yourself what would you do! We want to know how lipid molecules, specifically single and double-chain amphiphiles, interact with each other and solvent when they are added to water. Before you read the answer, look at the image below and ask yourself the question: What would I do if I were a single chain amphiphile and jumped into water as shown in Figure 10.2.1?

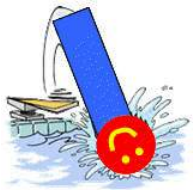


Figure 10.2.1: A cartoon single-chain amphiphile diving into a pool

Here is what they do. When added to water, some single-chain amphiphiles dissolve in water while others form a monolayer on the surface of the water. If enough enter the solution and exceed their net solubility, they self-aggregate to form micelles. These outcomes are shown in Figure 10.2.2

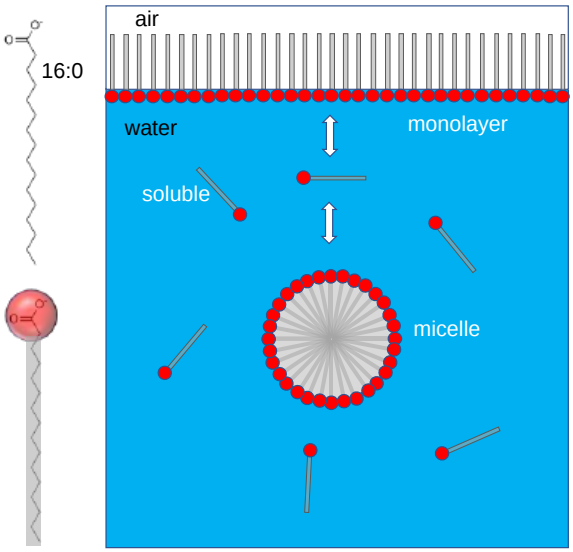


Figure 10.2.2: Distribution of single chains amphiphiles in an aqueous solution

Figure 10.2.3shows an interactive iCn3D model of an sodium dodecylsulfate (SDS) micelle

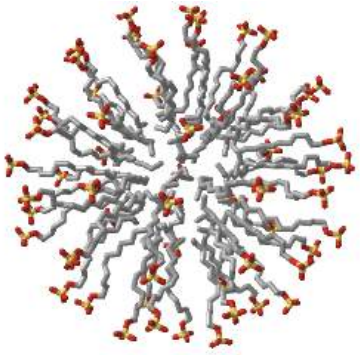


Figure 10.2.3: Sodium dodecylsulfate (SDS) micelle (Copyright; author via source).
Click the image for a popup or use this external link: not available

Double-chain amphiphiles, in contrast, form bilayers instead of micelles. (Note: single and double-chain amphiphiles can form other multimolecular aggregate structures as well, but micelles and bilayers are the most common and are the only ones we will consider.)

The micelle interior is completely nonpolar. Spherical bilayers that enclose an aqueous compartment are called vesicles or liposomes. Micelles and bilayers, formed from single and double-chain amphiphiles, respectively, represent noncovalent aggregates and hence are formed by an entirely physical process. No covalent steps are required.

Common single-chain amphiphiles that form micelles are detergents (like sodium dodecyl sulfate - SDS) as well as fatty acids, which themselves are detergents. Sodium hydroxide feels slippery on your skin since the base hydrolyses the fatty acids esterified to skin lipids. The free fatty acids then aggregate spontaneously to form micelles which act like detergents, and are also slippery.

Micelle/detergents in water are an example of an emulsion of two liquids that are generally immiscible in each other unless one is dispersed into small droplets into the other Fine oil drops can be dispersed in water and fine aqueous drops can be dispersed in a nonpolar liquid. Many vaccines are formulated as this later type of emulsion. Grease and oil in your clothes can be carried away by "diving" into the nonpolar part of the detergent micelle which is dispersed in water as an emulsion. Another example of an emulsion or more properly a colloid is a cloud, a dispersion of liquid water droplets in a solvent, the atmosphere.

The formation of these structures can be understood from the study of noncovalent interactions but also through thermodynamics. In a micelle, the buried acyl chains can interact and be stabilized by induced dipole-induced dipole forces as the nonpolar carbons and hydrogen are in van der Walls contact. They are sequestered from water. This view fits our simple axiom of "like-dissolves like". The polar head groups can be stabilized by ion-dipole bonds between charged head groups and water. Likewise, H-bonds between water and the head group stabilize the exposed head groups in water. Repulsive forces may also be involved. Head groups can repel each other through steric factors, or ion-ion repulsion from like-charged head groups. The attractive forces must be greater than the repulsive forces, which lead to these molecular aggregates.

From a thermodynamic approach, one problem arises with this simple explanation. For a micelle or bilayer to form, many monomers must aggregate to form a single micelle or vesicle, which is entropically disfavored! So let's delve into the thermodynamics of micelle formation.

ΔG, the free energy chance for a reaction, determines the spontaneity and extent of a chemical or physical reaction. The free energy of the system depends on 3 variables, temperature T, pressure P, and n, the number of moles of each substance. For the latter, think of solute X on two different sides of a permeable membrane. If the concentration of X is the same on each side, as shown in the system below, the system is in equilibrium as shown in Figure 10.2.4

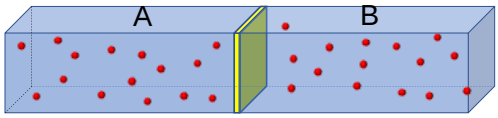


Figure 10.2.4: Equilibrium state of a species distribute in solutions separated by a semipermeable membrane.

If the system is composed of two different parts, A and B, the system is at equilibrium (ΔG=0) if $T_A = T_B$, $P_A = P_B$, and the change in the absolute free energy per mole of A is $\Delta G_A/\Delta n = \Delta G_B/\Delta n$. More precisely, using simple calculus, we would discuss incremental changes in absolute free energy/mol, dG_A/dn for A (often called the chemical potential of A, μ_A), and dG_B/dn or

μ_B)for B. At equilibrium $dG_A/dn = dG_B/dn$. (We will use the symbol G here instead of μ). G then is the absolute free energy/mol (again often called the chemical potential), where $G=G^o + RTln[A]$. The equations you used in introductory chemistry can be written.

$$\begin{aligned} \Delta G &= \Delta G^0 + RTInQr \\ \Delta G &= \Delta H - T\Delta S \\ \Delta G^0 &= \Delta H^0 - T\Delta S^0 \\ \Delta G^0 &= -RTInKeq \end{aligned} \tag{10.2.1}$$

Now let's apply this to the chemical equation for micelle formation:

n SCA <==> 1 micelle

where SCA represents a single chain amphiphile. At first glance, we might suspect that:

- $\Delta H^0 < 0$ since the induced dipole-induced dipole interactions among the buried acyl chains in the micelle would be much more favorable than the water-acyl interactions for the monomeric amphiphile in solution. This notion is supported by our aphorism, "like dissolves like". Of course, we couldn't ignore polar interactions (H bonding for example) among the head groups and water, but we might expect these to be equally favorable in both the monomeric and micellular states.
- $\Delta S^0 < 0$ since we are forming a very ordered state (a single micelle) with much less entropy from a state (single chains amphiphiles dispersed in solution) with much more entropy.

Hence it would appear that micelle formation is enthalpically favored but entropically disfavored. Let's examine this issue more closely. First, we need to obtain a greater understanding of ΔG^o which should give us a clue as to where a SCA would "want" to be in this mixture. Remember, ΔG^0 is a constant that at a given T, P, and solvent conditions and depends only on the relative stability of a molecule in a given environment and not its concentration.

Traube, in 1891, noticed that single chain amphiphiles tend to migrate to the surface of the water and decrease its surface tension (ST.) He observed that the decrease in ST is directly proportional to the amount of amphiphile, added up until a certain point, at which added amphiphile has no additional effect. In other words, the response of ST saturates at some point.

We are more interested in what happens to amphiphiles in bulk water, not at the surface. As we showed in Figure 2 above, monomeric single chain amphiphiles are in equilibrium with single chain amphiphiles in micelles. Assume you have a way to measure monomeric single chain amphiphile in solution. What happens to its concentration as you add more and more SCA to the mixture? Turns out you observe the same effect that Traube noted with changes in surface tension. This explanation goes like this: as more amphiphile is added, more goes into bulk solution as monomers. At some point, there are enough amphiphiles added to form micelles. After this point, added amphiphiles form more micelles and no further increases in monomeric single chain amphiphiles are noted. The concentration of amphiphile at which this occurs is the critical micelle concentration (CMC). Figure 10.2.5 shows a graph of monomeric single chain amphiphile in solution versus the concentration added to the solution.

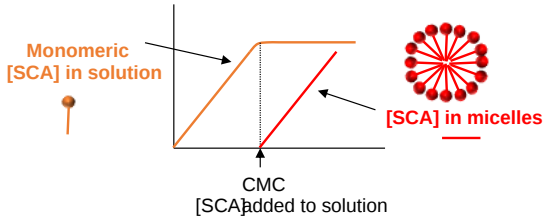


Figure 10.2.5: Monomeric single chain amphiphile in solution versus the concentration added to the solution

This saturation effect can be observed with other systems as well.

- Consider the amount of NaCl(aq) in the solution as more NaCl(s) is added to water. At some point, the water is saturated with dissolved NaCl, and no further increase in NaCl (aq) occurs.
- Consider the amount of a sparing soluble hydrocarbon (HC) in water. After saturation, phase separation occurs.

Now consider the addition of a drop of a slightly soluble hydrocarbon liquid (HC_L) into water, as pictured in the diagram below. At t=0, the system is not at equilibrium and some of the HC will transfer from the pure liquid to water, so at time t=0, $\Delta G_{TOT} < 0$.

This is illustrated in Figure 10.2.6

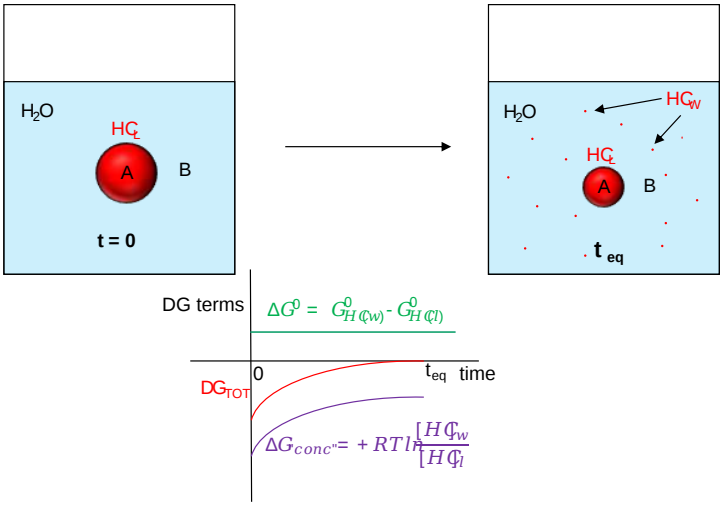


Figure 10.2.6: Addition of slightly soluble liquid hydrocarbon to water - ΔG vs time

The following equations can be derived.

$$\begin{aligned} \Delta \mathrm{G_{TOT}} = (G_{\mathrm{HC-W}}) - (G_{\mathrm{HC-L}}) = \mathrm{G^0_{HC-W}} + RT\ln[\mathrm{HC}]_{\mathrm{W}} - \left(\mathrm{G^0_{HC-L}} + RT\ln[\mathrm{HC}]_{\mathrm{L}}\right) = \\ \Delta \mathrm{G_{TOT}} = \left(\mathrm{G^0_{HC-W}} - \mathrm{G^0_{HC-L}}\right) + RT\ln([\mathrm{HC}]_{\mathrm{W}} - \ln[\mathrm{HC}]_{\mathrm{L}}) = \\ \Delta \mathrm{G_{TOT}} = \Delta \mathrm{G^0} + RT\ln\frac{[\mathrm{HC}]_{\mathrm{W}}}{[\mathrm{HC}]_{\mathrm{L}}} \end{aligned} \tag{10.2.2}$$

Now add a bit more complexity to the last example. Add a hydrocarbon x, to a biphasic system of water and octanol and shake it vigorously as shown in Figure 10.2.7. At equilibrium, x would have "partitioned" between the two mostly immiscible phases.

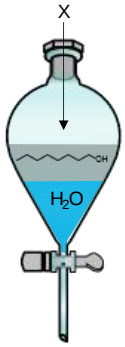


Figure 10.2.7: Addition of a hydrocarbon x to a biphasic system of water and octanol

A simple favorable reaction can be written for this system: x aq → x oct.

If X is a hydrocarbon, $\mathbf{\Delta G} < 0$. Also, $\mathbf{\Delta G^o} < 0$, since this term is independent of concentration and depends only on the intrinsic stability of x in water in comparison to that of octanol. This simple equation holds:

$$\Delta \mathrm{G_{TOT}} = \left(\mathrm{G^0_{X-oct}} - \mathrm{G^0_{X-w}}\right) + RT\ln\frac{[\mathrm{X}]_{\mathrm{oct}}}{[\mathrm{X}]_{\mathrm{w}}} = \Delta \mathrm{G^0} + RT\ln\frac{[\mathrm{X}]_{\mathrm{oct}}}{[\mathrm{X}]_{\mathrm{w}}} \tag{10.2.3}$$

At equilibrium, ΔG^0=0 and the equation can be rewritten as:

$$\Delta \mathrm{G^0} = -RT\ln\frac{[\mathrm{X}]_{\mathrm{oct}}}{[\mathrm{X}]_{\mathrm{w}}} = -\mathrm{RTlnK_{part}} \tag{10.2.4}$$

where K_{part} is the equilibrium partition coefficient for X in octanol and water. This can readily be determined in the lab. Just shake a separatory flask with a biphasic system of octanol and water after injecting a bit of X. Then separate the layers and determine the concentration of x in each phase. Plug these numbers into the last equation. You should be able to predict the sign and relative

magnitude of ΔG^o since it does not depend on concentration, but only on the intrinsic stability of the molecules in the different environments. K_{paft} values are often determined for drugs since they often must diffuse across cell membranes to move into the cytoplasm where they can act. Drugs hence must have a reasonable K_{part} to pass through the membrane but not so high that they are insoluble.

10.2.2: Double Chain Amphiphiles and Bilayers

In contrast to single chain amphiphiles, double-chain amphiphiles added to water form monolayers and vesicles called liposomes, as shown in Figure 10.2.8.

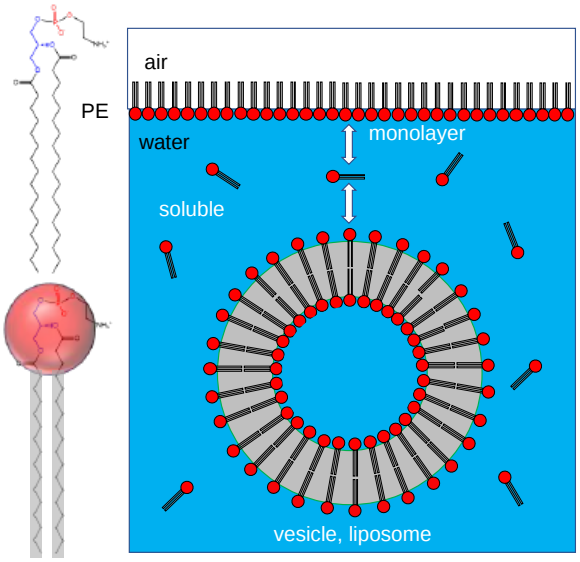


Figure 10.2.8: Distribution of double chains amphiphiles in an aqueous solution

They can be unilamellar, consisting of a single bilayer surrounding the internal aqueous compartment, or multilamellar, consisting of multiple bilayers surrounding the enclosed aqueous solution. You can image imagine that multilamellar vesicles resemble an onion with its multiple layers. Cartoons of unilamellar and multilamellar liposomes are shown in Figure 10.2.9, where each concentric circle represents a bilayer.

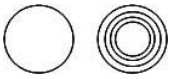


Figure 10.2.9: unilamellar and multilamellar liposomes (vesicles)

Liposomes vary in diameter. They can be generally categorized into small (S, diameter < 25 nm), intermediate (I, diameter around 100 nm), and large (L, diameter from 250-1000 nm). If these vesicles are unilamellar, they are abbreviated as SUV, IUV, and LUV

The chemical composition of liposomes made in the lab can be widely varied. Most contain neutral phospholipids like phosphatidylcholine phosphatidyl ethanolamine (PE), or sphingomyelin (SM), supplemented, if desired, with negatively charged phospholipids, like phosphatidyl serine (PS) and phosphatidyl glycerol (PG). In addition, single-chain amphiphiles like cholesterol (C) and detergents can be incorporated into the bilayer membrane, which modulates the fluidity and transition temperature (Tm) of the bilayer. If present in too great a concentration, single-chain amphiphiles like detergents, which form micelles, can disrupt the membrane so completely that the double-chain amphiphiles become incorporated into detergent micelles, now called mixed micelles, in a process that effectively destroys the membrane bilayer.

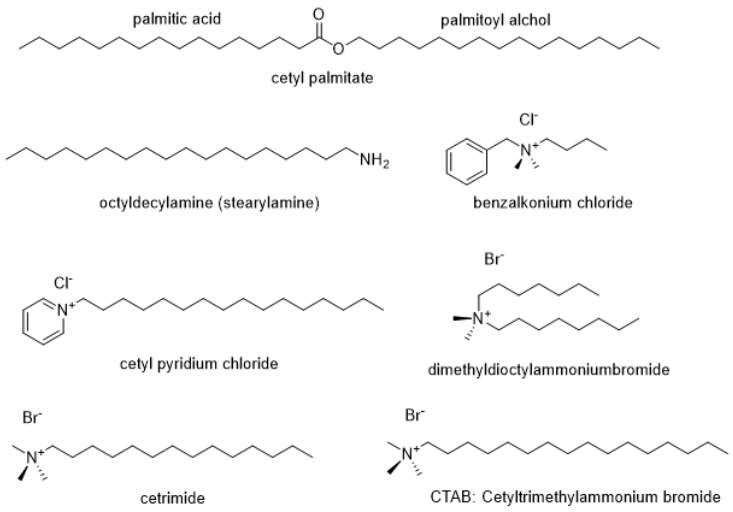
The properties of liposomes (charge density, membrane fluidity, and permeability) are determined by the lipid composition and size of the vesicle. The desired properties will be, in turn, determined by the use of the particular liposome. The vesicles offer wonderful, simple models to study the biochemistry and biophysics of natural membranes. Membrane proteins can be incorporated into the liposome bilayer using the exact method you will be using. But apart from these purposes, liposomes can be used to encapsulate water-soluble molecules such as nucleic acids, proteins, and toxic drugs. These liposomes can be targeted to specific cells if antibodies or other molecules which will bind specifically to the target cell can be incorporated into the bilayer of the

vesicle. Intraliposomal material may then be transferred into the cell either by fusion of the vesicle with the cell or by endocytosis of the vesicle.

Liposomes could also be called lipid nanoparticles as they have sizes ranging up to 1000 nm. Liposomes are vesicular - with small aqueous-filled compartments. They can be made with encapsulated drugs for delivery to target cells through blood transport. They act as an emulsion in water.

Lipid nanoparticles can also be particulate (insoluble), which slowly degrade and release their contents slowly in situ. Most recently, particulate lipid nanoparticles have helped save the world by being used to encapsulate messenger RNA (mRNA) for the spike protein from the SARS-CoV-2 virus which causes the COVID-19 pandemic. These lipid nanoparticles are used in the vaccine against the coronavirus mRNA molecules which encode the spike protein are "encapsulated" in the lipid nanoparticle. The mRNA contains specially modified nucleotides to increase their stability. The lipid nanoparticles also contain positively-charged lipids which help stabilize the negatively charged mRNA from degradation. The nanoparticles are endocytosed into cells, where the mRNA can be translated into SARS-CoV-2 spike protein, required for whole virus entry into the cell. The spike protein is then recognized by the immune system.

The lipids used in the formulation of these nanoparticles include fatty acids, mono-, di- and triglycerols, glycerophospholipids, waxes (like cetyl palmitate), and other positively charged lipids including stearyl amine, benzalkonium chloride, cetrimide, cetyl pyridinium chloride, and dimethyldioctadecylammonium bromide. These are shown in Figure 10.2.10



10.2.3:

Figure 10.2.10 Amphiphiles used to make lipid nanoparticles

10.2.4: Why Micelles and Bilayers?

Micelles and liposomes form spontaneously - i.e. $\Delta G < 0$. But why do single chain amphiphiles form micelles and double-chain amphiphiles form bilayers? Let's think about this from a thermodynamics and structural sense.

As the number of Cs in the nonpolar carbon (NC) chain increases, the ΔG for the transferring into a micelle, or by analogy, for a single chain amphiphile entering a micelle, becomes more and more negative (i.e more favored). The following equation seems to apply to the transfer of a single chain amphiphile into a micelle:

$\Delta G^o = G^o$ (mic) - G^o (aq) = + number - 709 NC

$$\Delta G^\circ = G^\circ \text{ (mic)} - G^\circ \text{(aq)} = + \text{number} - 709\text{NC} \tag{10.2.5}$$

where NC is the number of carbon atoms in the chain. The first positive term depends on the nature of the head group, while the second negative term is independent of the head group. These + and - terms bring us back to the principle of opposing forces we discussed when we looked at the noncovalent interactions involved in micelle and bilayer formation.

There are attractive interactions, including induced dipole-induced dipole interactions among the chains and dipole-ion, and H-bond interactions with water and the head groups. There are likewise repulsive interactions arising from steric hindrance with bulky

heady groups and ion-ion repulsions. Of course, there are also entropic considerations. Let us now consider these factors as we explore what might happen to a preformed micelle as we try to put more single chain amphiphiles (**sca**) into it.

As we increase the number of Cs in the SCA, the micelles would have a larger radius. For a given SCA with a fixed number of Cs, once a spherical micelle is formed, it can no longer retain its spherical shape if more SCAs are added. Imagine increasing the diameter of a spherical micelle 10x. A large part of the inside would have no atoms or be filled with water, which would not be favorable. Therefore, if the micelle is to grow, it can do so only by changing shape to something other than a sphere. By squeezing a tennis ball, one can imagine that the shape could distort to a circular cylinder with end caps. In this way, the acyl chains can still interact. The only problem is that head groups will now be closer than they were in the sphere. This is simplistically illustrated in Figure 10.2.11.

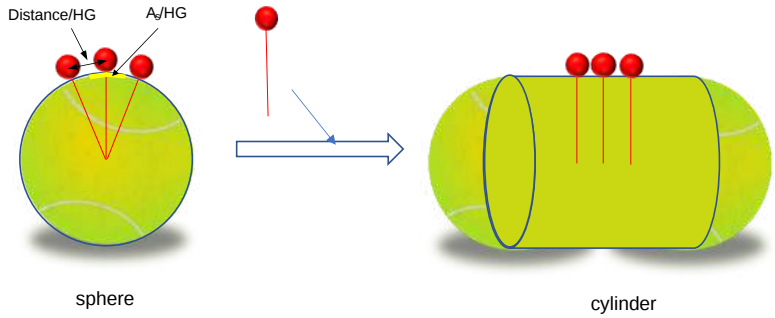


Figure 10.2.11: Hypothetical change in head group interactions on the conversion of spherical vesicle to a cylindrical vesicle

Imagine in a sphere the head groups radiating in a perpendicular direction from the sphere surface. As the sphere is distorted into a cylinder, the head groups would come closer together, and hence they will experience more steric interference. If a cylinder can be formed, however, it could continue to grow as long as needed with no further compression required. Imagine now that you further compressed the cylinder into a planar "bilayer" structure. The head groups would be even closer and experience even more repulsion. This bilayer will not form since growth can occur in the cylindrical phase without the added repulsion.

Now consider a double-chain amphiphile (DCA). In the case of a SCA, the number (N) of head groups (HG) = the number of acyl chains (CH). Hence the surface area per HG is equal to the area per HC. Or: As/N HG = As/ N CH. For a DCA, N HG = N CH/2, therefore As/N HG = 2As/NCH. There is twice the surface area available per head group compared to that of the SCA. Therefore the DCA can tolerate more compression. It can easily be compressed to a bilayer, which as we saw, has much less As/HG. The cylindrical form has too much space per head group since water can enter the structure. The extra closeness of head groups in the bilayer can be tolerated even more since the ΔG^{o} for transfer of a DCA into a micelle is 60% more negative than that of a SCA. The As/HG for closed vesicles differs only slightly from that of a truly planar bilayer since the vesicles are so large compared to a micelle.

Once again, we have discovered that structure mediates function. We can account for the fact that SCA and DCA form micelles and bilayers, respectively, by understanding the structure of the monomers!

In reality, things are more complicated.

The general rule holds that single chain amphiphiles form micelles and double-chain amphiphiles form bilayers. However, under the right condition, single-chain fatty acids can form bilayers if the pH is low enough that the head group is protonated and uncharged. Why would that make a difference? Fatty acids like oleic acid would be a prime candidates for components of the membranes of protocells in the evolution of life from abiotic conditions. Likewise, short double-chain amphiphiles can make micelles. In addition, other lipids phases can be observed. Which aggregates or phases ultimately form depends on the structure of the lipid, the solvent conditions, and the temperature. These include the following phases:

- lamellar gel (Lb) and lamellar liquid crystalline (La) phases
- hexagonal HI (cylinders packed in the shape of a hexagon with polar heads facing out into the water
- hexagonal HII (cylinders packed in the shape of a hexagon with acyl chains pointing out as in reverse micelles, and
- micellar (M).

We will discuss them in more detail in the next section.

10.3: Membrane Bilayer and Monolayer Assemblies - Structures and Dynamics

10.3.1: An overview of lipid bilayers

A membrane bilayer consists of more than just two leaflets of amphiphilic leaflets. It also contains membrane proteins (which we will discuss in the next chapter) which can also be attached to carbohydrates. Most assuredly you have seen various representations of a bilayer before. Before we proceed with a more detailed description of the lipids in the bilayer and its associated properties, we present Figure 10.3.1 to focus our discussion.

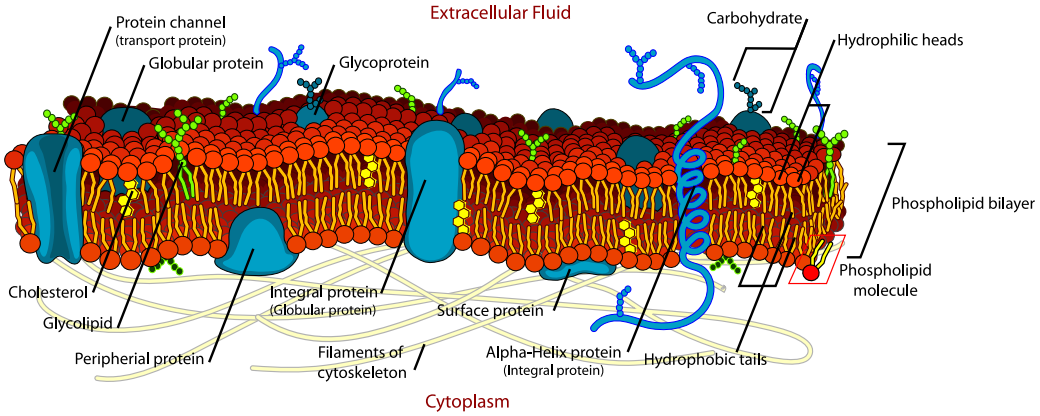


Figure 10.3.1: A representation of a membrane bilayer with peripheral and integral membrane proteins. https://cnx.org/contents/FPtK1zmh@8.108:q2X995E3@12/The-Cell-Membrane. Creative Commons Attribution 4.0 International license.

To understand the movement of lipids in an actual cell, a better understanding of lipid synthesis and trafficking in cells is important. Table 10.3.1 below shows the distribution of four classes of lipids in a macrophage, a type of immune cell (Andreyev, A.Y. et al) while the following figure shows how the lipids composition of membranes organelle membranes.

Table 10.3.1: Distribution of Lipids in Resting Macrophage

Lipid Categories	Nucleus	Mitochondria	ER	Plasma Memb	microsome	cytosol	Whole cell
Glycero-phospholipids	149	152	150	151	142	109	155
Prenol lipids	5	5	5	5	5	5	5
Sphingolipids	48	47	48	48	48	47	48
Sterol lipids	13	12	12	13	11	5	12
Total	215	216	215	217	206	166	220

Lipids in membranes are often distributed asymmetrically. The inner and outer leaflets of a biological membrane usually have different PL compositions. For example, in red blood cell membranes, the outer leaflet is composed mostly of sphingomyelin (SM) and PC, while the inner leaflet is composed mostly of PE and phosphatidyl serine (PS). This phospholipid contains the amino acid serine linked through its side chain (-CH2OH) to phosphate in position 3 of diacylglycerol. With a negative charge on the phosphate and carboxylate and a positive charge on the amine of PS, this phospholipid is acidic with a net negative charge. All the PS is located in the inner leaflet! This observation will become important later on when we discuss programmed cell death. A dying cell will expose PS in the outer leaflet. This is one of the markers of a dying cell. The membrane lipid composition in an average mammalian cell is shown in Table 10.3.2 below.

Table 10.3.2: Membrane lipid composition in an average mammalian cell.

Lipid	%
PC	45-55
PE	15-25

Lipid	%
PI	10-15
PS	5-10
PA	1-2
SM	5-10
cardiolipin (bis-PG)	2-5
cholesterol	10-20

Lipid membranes also surround the variety of intracellular organelles found in eukaryotic cells. As a refresher, Figure 10.3.2 shows the anatomy of a typical eukaryotic cell with its variety of intracellular organelles.

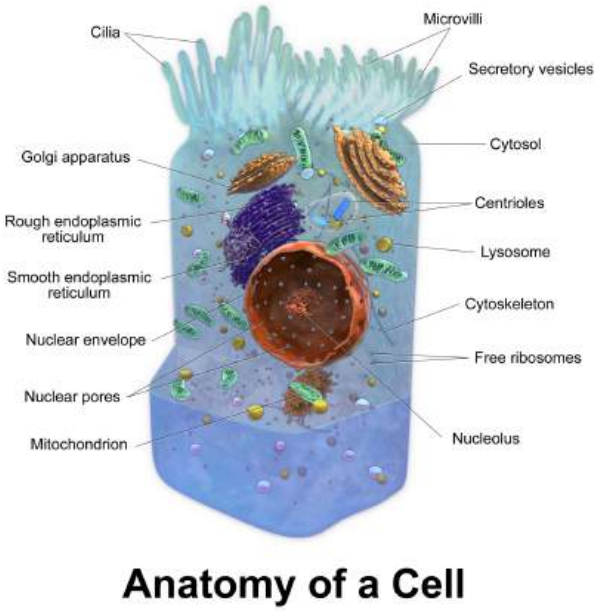


Figure 10.3.2: The Anatomy of a eukaryotic cell. https://opm.phar.umich.edu/biological_membranes https://creativecommons.org/licenses/by/3.0/

Figure 10.3.3 shows the average distribution of membrane lipids in different eukaryotic organelles.

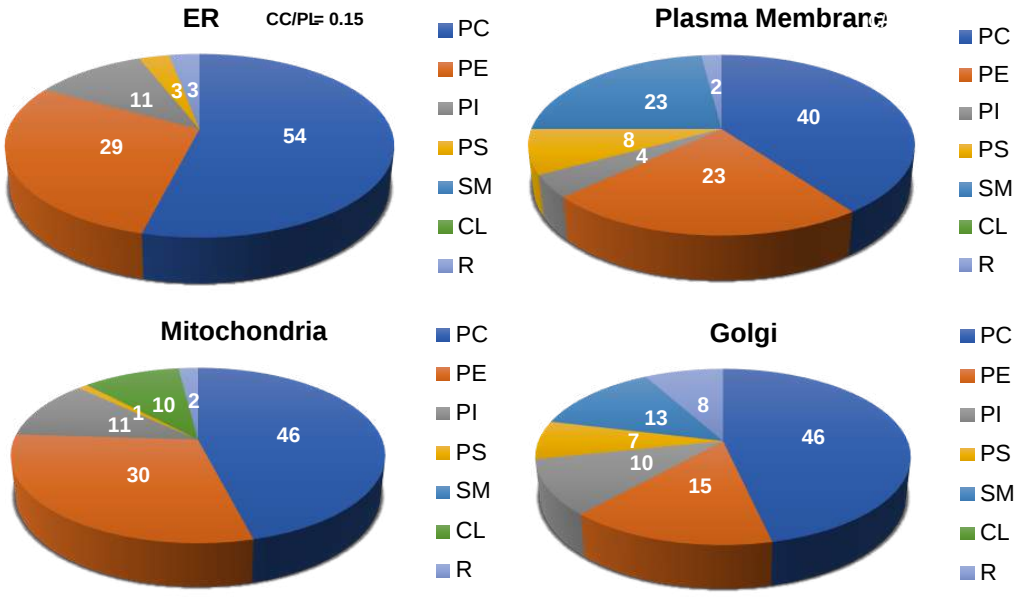


Figure 10.3.3: Average distribution of membrane lipids in different eukaryotic organelles

Figure 10.3.4, our last overview, shows the location of lipid synthesis and the resulting distribution of lipids in each leaflet. Note that most lipids are synthesized in the endoplasmic reticulum (ER).

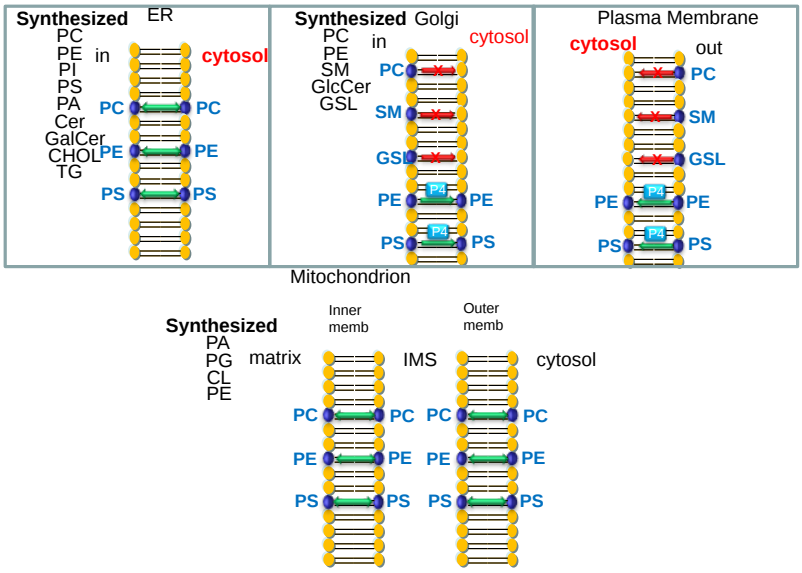


Figure 10.3.4: Location of lipid synthesis and the resulting distribution of lipids in each leaflet.

10.3.2: Dynamics of Membrane Bilayers

Molecules are not static, but rather are dynamic. This also applies to molecular aggregates. In the first part of the section, we will discuss the rigid movement of whole lipid molecules in a bilayer, within a leaflet, and between leaflets. In the second part and the following supplement, we will consider the movement of atoms within a molecule. The movements include motions like bond bending, bond stretching, and torsion angle changes as we saw in the previous chapter section on the conformations of n-butane. The position of all atoms within a molecule can be simulated as a function of time - a molecular dynamics simulation. Such motions affect the energy of the molecule, which can be calculated for given atom positions using classical molecular mechanics and electrostatics.

Liposomes and bilayers in general must be somewhat dynamic, otherwise, they would be impenetrable barriers across which nothing could pass. Cell membranes must separate the outside of a cell from the inside, but they must also allow the passage of

molecules and even ions across the membrane. What is the evidence that membranes are dynamic?

First, lipids can diffuse laterally in the membrane. This can be shown as follows. Make a liposome from phosphatidylethanolamine, PE, which has been labeled with TNBS (trinitrobenzensulfonate). The NH_2 on the head group of PE can attach the TNBS which undergoes nucleophilic aromatic substitution with the expulsion of the SO_3^{2-}. The TNB group attached to the PE head group absorbs UV light and emits light of a higher wavelength in a process called fluorescence. Next, using a fluorescent microscope, the fluorescence intensity of a region of the surface can be recorded. Then shine a laser on a small area of the surface, which can photobleach the fluorescence in the area. Over time, fluorescence can be detected from the region again. The rate at which it returns is a measure of the lateral diffusion of the labeled lipids into the region. Lipids can undergo lateral diffusion at a rate of about 2 mm/s. This implies that the lipids can transit the surface of a bacteria in 1 sec.

Transverse or flip-flop diffusion (movement of a phospholipid from one leaflet to the other, not within a given leaflet) should be more difficult. Experimentally, this is investigated as shown in the diagrams below.

Flip-Flop Diffusion in Liposomes: To test flip-flop diffusion in an artificial membrane, liposomes are made with a mixture of PC and a PC derivative with a nitroxide spin label (has a single unpaired electron) as shown in Figure 10.3.5.

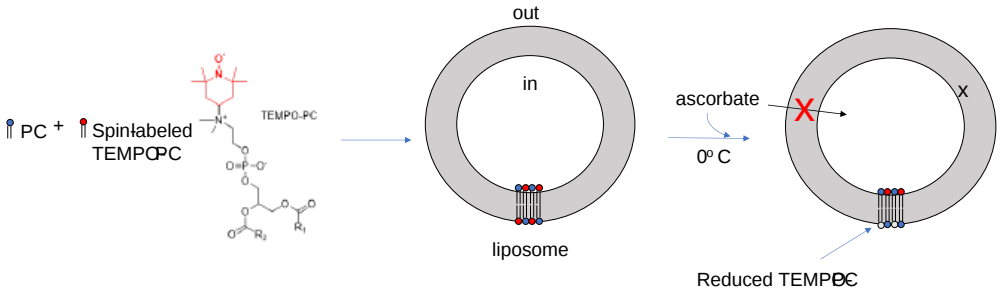


Figure 10.3.5: Flip-flop diffusion in vesicles

Both the inner and outer leaflets of the membrane have the labeled PC. Like a proton in NMR spectroscopy, a single electron has a spin that can give rise to an electron-spin resonance (ESR) signal (as a proton gives rise to a nuclear magnetic resonance signal) when irradiated with the appropriate frequency electromagnetic radiation (microwave frequency for ESR, radio frequency for NMR) in the presence of a magnetic field. The liposomes are kept at 0°C and the ESR signal is determined. Ascorbic acid, a water-soluble vitamin, and antioxidant/ reducing agent is added to the liposomes. This reduces the spin-labeled PC in the outer leaflet, but not the inner leaflet of the bilayer since ascorbic acid can not enter the liposome or otherwise interact with it. This reduces the ESR signal to a lower, constant value.

The sample is divided into two. One sample is left at 0°C, the other is raised to 30°C. The ESR signal is recorded as a function of time. The 0°C prep shows no change in ESR with time, while the 30°C prep ESR signal decreases with time. This decrease results from flip-flop diffusion of labeled PC from the inner leaflet to the outer, and subsequent reduction by ascorbic acid. These experiments in experimental bilayer systems like liposomes show that flip-flop diffusion is orders of magnitude slower than lateral diffusion.

Flip-Flop Diffusion in Bacterial Cells

An analogous experiment can be done with bacteria. Radiolabeled $^{32}PO_4^-$ is added to cells for one minute, which leads to the labeling of newly synthesized phospholipid (PL) which locates in the inner leaflet. The cells are then split into two samples. One sample is reacted immediately with TNBS, which will label only PE in the outer leaflet. The other sample is incubated for 3 minutes (to allow PL synthesis) and then reacted with TNBS. This is shown in Figure 10.3.6.

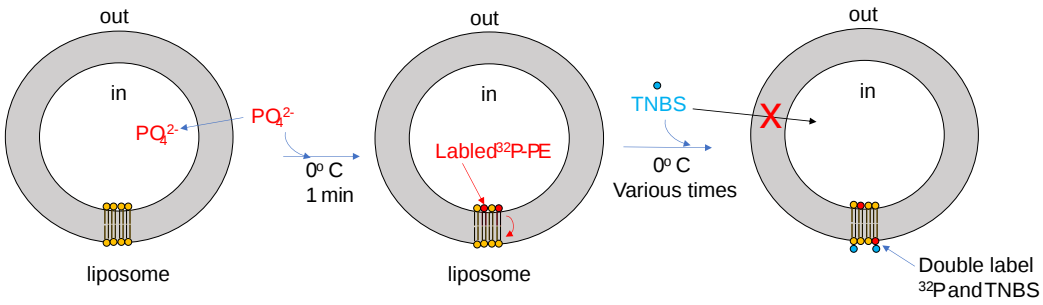


Figure 10.3.6: Flip-Flop Diffusion in Bacterial Cells

After a short labeling period, the cells are destroyed by adding organic solvents which prevent new lipids biosynthesis. The lipids are extracted into the solvent and then subjected to TLC.

The lipids can be labeled in three ways. Some will be labeled with ^{32}P alone, some with TNBS alone, and some with both ^{32}P and TNBS. TLC (or other techniques such as HPLC or GC) can easily separate PC and TNBS-labeled PC since they have different structures and hence will migrate to different places on a TLC plate. No chromatographic technique could, however, separate PC and ^{32}P-PC, since their molecular structure is the same, the only difference being in the nuclei of the P (different number of neutrons).

Those lipids with double labels (TNB and ^{32}P) must have flipped from the inner leaflet to the outer leaflet where they could be labeled with TNBS. The cells incubated for 3 minutes before the addition of TNBS have a much higher level of doubly labeled PLs. Quantitating these data as a function of differing time of incubation at elevated temperatures show that the rate of flip-flop diffusion is much higher in cells than in liposomes, which suggests that the process is catalyzed, presumably by a protein transporter (flippase or Transbilayer amphipath transporter - TAT) in cells.

The actual movement of lipids in bilayers is catalyzed by different enzymes, including flippases, floppases, and scramblases, as illustrated in Figure 10.3.7. Most required ATP hydrolysis for the physical movement of the lipid across leaflets.

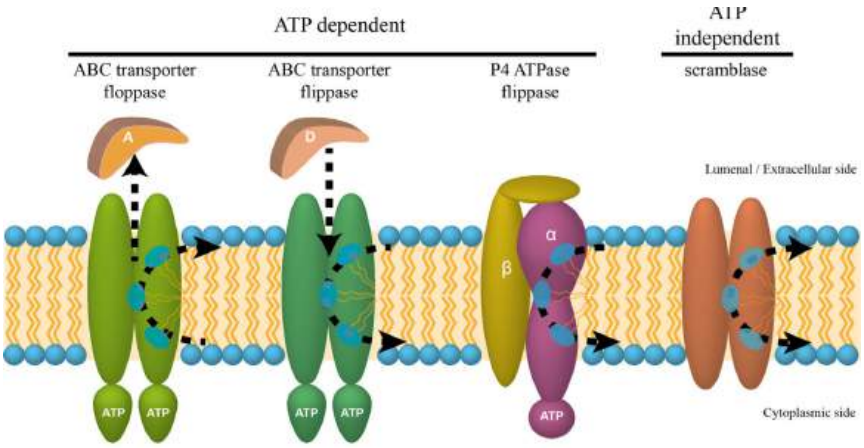


Figure 10.3.7: Enzymes that move lipids between leaflets. Juliana Rizzo et al. https://doi.org/10.1016/j.csbj.2019.09.001. Creative Commons: https://creativecommons.org/licenses/by/4.0/

Figure 10.3.8 shows an interactive iCn3D model of a human plasma membrane phospholipid flippase with bound phosphatidylserine (PS) shown in spacefill.

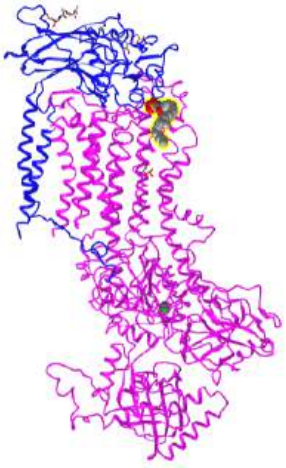


Figure 10.3.8 Human plasma membrane phospholipid flippase with bound phosphatidylserine (6lkn) (Copyright; author via source).

Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...kcJeKtZGDa53W6

The iCn3D model below shows the structure of a human plasma membrane phospholipid flippase with bound phosphatidylserine (PS) shown in spacefill. This protein moves PS from the outer to the inner plasma membrane leaflet, maintaining its asymmetric distribution. The other common aminophospholipid, PE, is also found predominantly in the inner leaflet. In contrast, PC and SM are found predominately in the outer leaflet. The movement of PS is from low to a high concentration and requires ATP. A build-up of PS in the outer leaflet is one signal that initiates programmed cell death (apoptosis) in the cell. Clotting is also initiated when cellular damage leads to exposed PS.

Flippases are proteins that move lipids from the outer to the inner leaflet, while floppases move them from the inner to the outer leaflet. Most are also ATPases. Both promote lipid asymmetry in the membrane and floppases also help move lipids out of the cell. Scramblases moved lipids in either direction and break the asymmetry of the lipid distribution. They are important in signaling. For example, they are used to expose phosphatidylserine to the outer leaflet which promotes programmed cell death.

Here are links to iCn3D models of

- a floppase MsbA (6BPP)
- scramblase (6P49 AND 7JLP)

10.3.3: Conformational Transitions in Bilayers

If a vesicle preparation is placed in a sensitive calorimeter and the temperature slowly increased, it is observed that the vesicle preparation absorbs a significant amount of heat at a temperature characteristic of the PLs which compose the vesicle. This is analogous to what would happen if solid water was heated. At the melting point of water, an increment of heat is required, the heat of fusion, to break H-bonds and cause melting. Likewise, the heat of vaporization is measured when H-bonds are broken on the liquid-gas transition. These transitions are associated with non-covalent processes, namely, breaking H-bonds. Graphs of heat absorbed (Q) as a function of temperature, or heat absorbed/T (i.e. the heat capacity) vs temperature for the melting and evaporation of water is shown in Figure 10.3.9. These transitions occur at the melting point (T_M) and the boiling point.

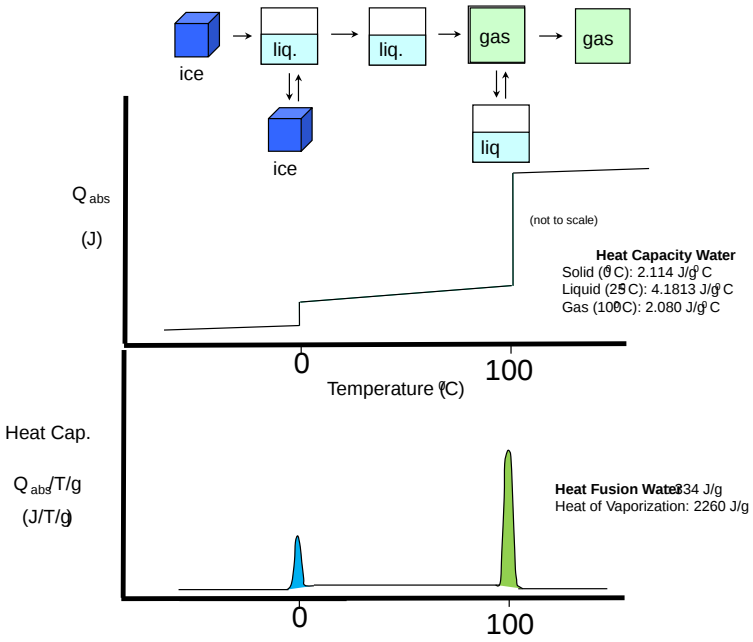


Figure 10.3.9: heat absorbed (Q) vs temperature (top) and heat absorbed/T (heat capacity) vs temperature for the melting and evaporation of water

The bottom heat capacity graph is nothing more than the derivative curve (or slope at each point) of the Q_{abs} curve!

Likewise, lipid vesicles undergo phase transitions comparable to the melting of water. One large phase transition at a "melting point" (T_M) = 42^0 C can be seen in the graph of heat capacity vs temperature for vesicles made of DPPC shown in Figure 10.3.10

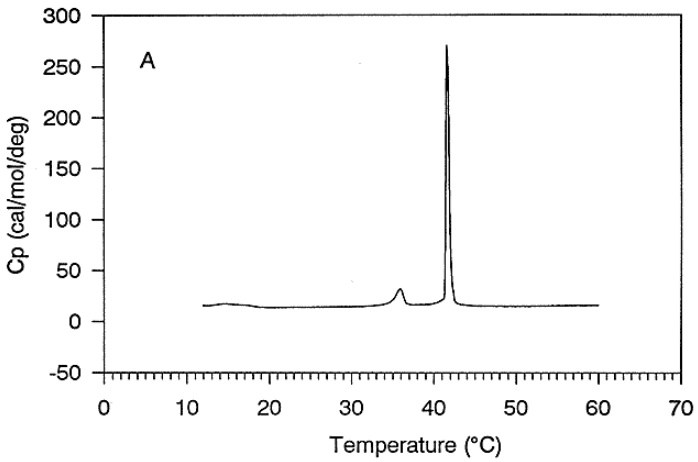


Figure 10.3.10: Heat capacity vs temperature for vesicles made of DPPC

This transition is caused by conformational chains in the packing of the acyl chains of the phospholipids as the acyl chains change from trans to gauche conformations. These changes involve not the simple translation of lipid molecules within and between bilayers but rather the movement of atoms within the molecules. These kinds of motions can be modeled using molecular dynamics simulations. Before the transition, the acyl chains are more tightly packed in the **gel phase**, and after they are less tightly packed in the **liquid crystalline phase** since many chains are in the gauche conformation. A minor transition is also noted at around 36^0 C. This is associated with changes in the orientation of head groups.

As with water going from ice to liquid, the vesicles after the phase transition are still intact. It's not like the transition of liquid to gas phase water. Vesicles in the liquid crystalline phase are more fluid, dynamic, and hence more permeable. Note that the liposomes have not been destroyed but simply have undergone a phase change, much like ice turning to liquid water.

The phases of lipid vesicles are given the names gel and liquid crystalline to reflect the rigidity of the bilayer.

- **Gel phase** (L_β): In the gel phase, which is found at temperature < T_M, the lipids are ordered with maximal packing. The acyl chains in both leaflets can be tilted so that they align in a parallel fashion (as shown in the figure below) or in a cross-tilted fashion in which they tilt toward each other. In the gel phase, the lipids diffuse slowly. This phase is sometimes called the solid phase. The gel phase is favored by low temperature and high saturation of esterified fatty acids. Saturated PC bilayers give a gel phase in the lab,
- **Liquid crystalline phase** (L_α): In this liquid crystalline phase, which is found at temperature > T_M, some saturated acyl chains have undergone all trans to gauche conformational changes. These introduce kinks into the chains which reduce packing. The notation L_α is used for bilayers of pure lipids
- **Liquid crystalline ordered** (L_0) and **Liquid crystalline disordered** (L_d): These phases typically occur with the addition of relatively high amounts of cholesterol. Cholesterol modulates the fluidity of membranes as we will see in a bit and affects bilayer properties at temperature both < and > T_M. The L_0 phase is often enriched in saturated (sphingo)lipids and cholesterol while the L_d phase is often enriched in unsaturated glycerophospholipids. The **liquid crystalline disordered** (L_d) has fast translational diffusion and lower order while the **Liquid crystalline ordered** (L_0) has fast diffusion with higher order.

Most membrane lipids *in vivo* contain unsaturated fatty acids and use specific lipids for given environments to avoid the gel phase. Figure 10.3.11shows some of these phases

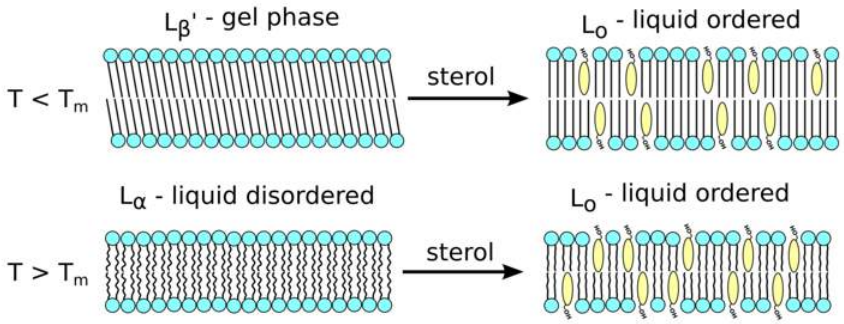


Figure 10.3.11: Common phases in membrane bilayers https://phys.libretexts.org/Courses/...se_Transitions

Figure 10.3.12shows a snapshot of a molecular dynamics simulation of a bilayer in a gel (A) and liquid crystalline (B) phase. Note that the width of the liquid crystalline phase is smaller.

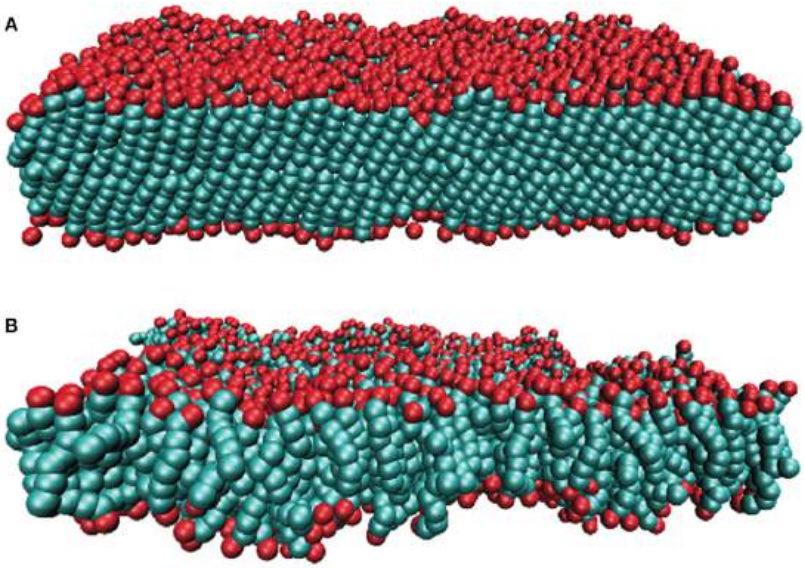


Figure 10.3.12: Snapshot of a molecular dynamics simulation of a bilayer in a gel (A) and liquid crystalline (B) phase. Tayebeh Jadidi et al. Bioeng. Biotechnol., 22 April 2014 | https://doi.org/10.3389/fbioe.2014.00008. **Creative Commons Attribution License (CC BY)**.

Vesicles made of different PL have different T_M as shown in Table 10.3.3below.

Table 10.3.3: Melting point (TM) of vesicles made with different phospholipids

Lipid	T_M	Lipid	T_M

Lipid	T_M	Lipid	T_M
12:0 PC	-1	12:0 PA	31
14:0 PC	23	14:0 PA	50
16:0 PC	41	16:0 PA	67
18:0 PC	55	18:0 PA	76
18:1 PC	-20	18:1 PA	-8
18:2 PC	-53	-	-
18:3 PC	-60	-	-

Vesicles make from phospholipids with bigger head groups have a lower T_M, since they are less "stable". For example, the Tm for vesicles of di-16:0 versions of PA, PE, and PC have T_Ms of 67, 63, and 41 degrees C, respectively, as shown in Figure 10.3.13

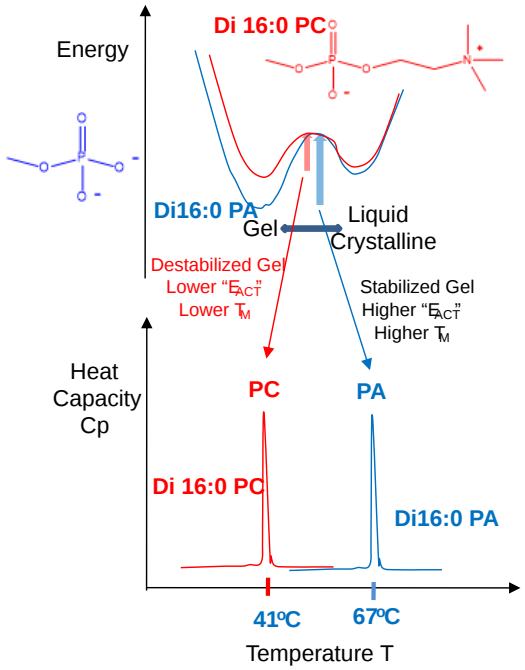


Figure 10.3.13: Melting point (T_M) of 16:0 vesicles with different head groups

10.3.4: Cholesterol and Membrane Fluidity

Cholesterol is also a ubiquitous component of animal cell membranes. Its size will allow it to fit into either leaflet with its polar OH pointed to the outside. One function of cholesterol in membranes is to keep the membrane fluid at any reasonable temperature. When a membrane is at a temperature less than the T_M, it is ordinarily in a gel, not a liquid crystalline phase. The cholesterol helps prevent the ordered packing of the acyl chains of the PLs, which increases their freedom of motion. Hence the fluidity and permeability of the membrane are increased. At temperatures greater than the T_M, the rigid ring of cholesterol reduces the freedom of the acyl chains to rotation and hence decreases the number of chains in the gauche conformation. This decreases fluidity and permeability. Cholesterol affects membrane structure at temperatures both below and above the T_M as

Figure 10.3.14shows the results of a molecular dynamics simulation depicting the relative order in a DMPC membrane with and without cholesterol

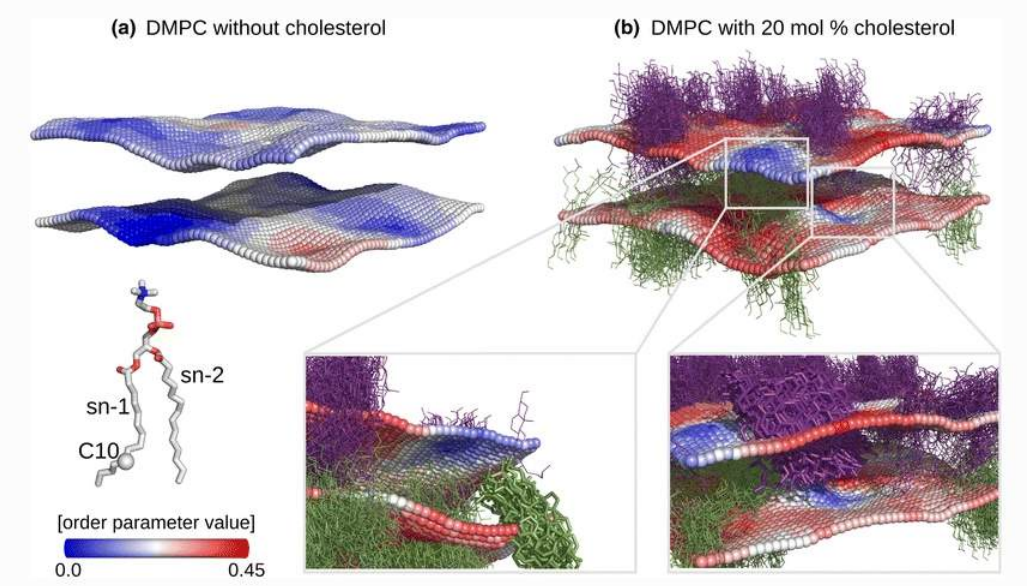


Figure 10.3.14: "S_{CD} order parameter analysis of cholesterol-enriched DMPC bilayer. Order parameter values for the C10 carbon of the acyl chain sn-1 are shown. The box area was 5.6 × 5.6 nm^2 and 50 bins were used along the *x* and *y* axes. **a** Local deuterium order parameters of a pure DMPC membrane averaged over a 100 ns MD trajectory. **b** Local S_{CD} order parameters for a 20 mol% DMPC bilayer averaged over a 10 ns MD trajectory excerpt. The cholesterol molecules represented as *lines* are overlayed for all the frames. The enlarged view figures emphasize the areas where cholesterol (visualized as sticks) decreases ordering of the lipid chains on the opposite leaflet" (from below ref)

10.3.5: Lipid Rafts and Nanodomains

Not only are lipids asymmetrically distributed between leaflets of a bilayer, but they are also distributed asymmetrically within a single leaflet. Certain lipids often cluster within a leaflet to form lipid "rafts" which can be considered to result from lateral phase separation of the lipids within one leaflet of the bilayer. Divalent cations like calcium, which can bind to negatively charged PLs like PS, can cause "rafts" of PS to form, giving rise to lateral asymmetry within a leaflet of a bilayer. Rafts also appear to be enriched in cholesterol and lipids with saturated fatty acids, especially sphingolipids, which would lead to regions of enhanced packing and reduced fluidity. Cholesterol would stabilize packing in spaces created with lipids with large head groups. You can think of these rafts as nanodomains, analogous to the domains we observed in protein structure.

Cholesterol appears to be a key player in the formation of lipid rafts. It is planar and inflexible and would pack better with saturated fatty acid chains and could also induce them to elongate to form lower energy zig-zag structures in which all the methylene groups are anti. Lipid rafts would represent a more ordered lipid phase (Lo) compared to the more disordered surrounding phase (Ld). Also compared to the structure of glycerophospholipids, the atoms in the region linking the head group and the nonpolar fatty acid chains in sphingolipids have greater potential for H bond interactions with cholesterol and other sphingolipids, as shown in Figure 10.3.15

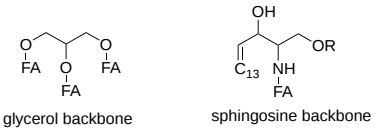


Figure 10.3.15: Comparison of glycerol and sphingosine headgroup structure. FAs are fatty acids esterified (glycerol) or amide link (sphingosine)

Rafts probably bind or exclude the binding of other biological molecules like proteins. Some proteins are chemically modified with a glycosylphosphoinositol (GPI) group at the carboxy terminus. The PI group can insert into the membrane, anchoring the protein to the bilayer. Protein also appear to induce raft formation. Lipids rafts appear to be enriched in glycosylphosphoinositol (GPI)-anchored proteins as we will see in Chapter 11.1. Recent studies have shown that the Ebola virus interacts with lipid rafts in the process of entering and exiting the infected cell. Rafts are also involved in how cells sense and respond to their environment. Signaling molecules on the outside of the cell can bind receptor proteins in the membrane. As we will see later, conformational changes in the receptor protein signal the inside of the cells that the receptor is bound with a ligand. Once bound, the receptor can move in the membrane and often cluster in outer leaflet rafts that contain cholesterol and sphingolipids. Inner leaflet rafts have also

been observed. Figure 10.3.16 shows two versions of an animated version of a lipid raft. The large shapes represent membrane proteins selectively found in the rafts. The most modern definition of a lipid raft is a nanoscale assembly of sphingolipids, cholesterol, and proteins that can be stabilized into platforms.

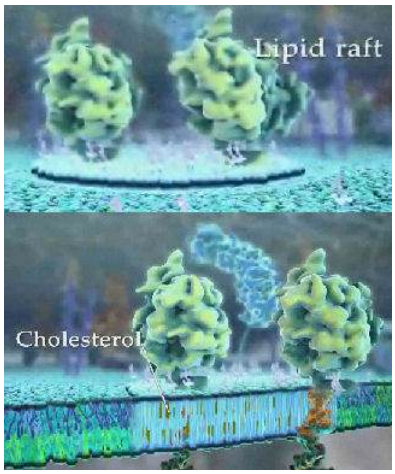


Figure 10.3.16: Lipid Rafts enriched in SM and Cholesterol. (screen capture from: http://multimedia.mcb.harvard.edu/anim_innerlife.html)

Figure 10.3.17 shows a simplified model of a lipid raft.

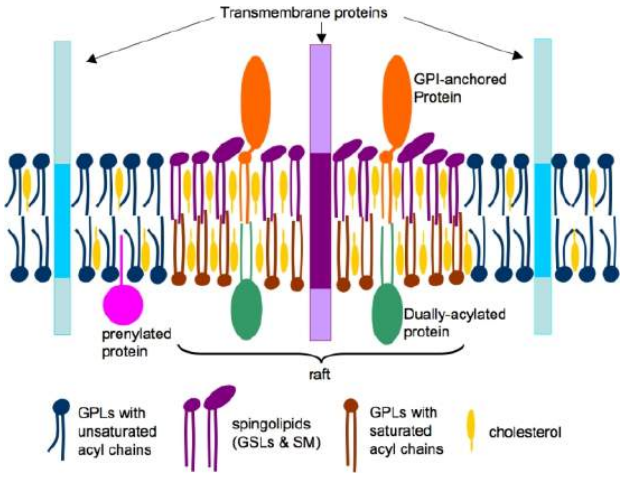


Figure 10.3.17: A simplified model of lipid rafts in cell membranes. https://www.researchgate.net/figure/..._fig4_45952334. license: https://creativecommons.org/licenses/by/3.0/

"The phospholipids (blue and brown) and cholesterol (yellow) are distributed in both the leaflets, whereas sphingolipids (violet) are enriched in the outer leaflet of the bilayer. The acyl chains of raft lipids are generally long and saturated (violet and brown), whereas those in non-raft domains are shorter and contain singly or multiply unsaturated acyl chains (blue). Raft domains contain concentrations of dually-acylated (green) and GPI-anchored (brown) proteins, whereas transmembrane (blue) and prenylated (green) proteins are usually non-raft associated."

Lipid bilayers, in contrast to single proteins, for example, are physical mixtures. In a thermodynamic sense, entropy would disfavor raft formation as random mixing is favored. Enhance enthalpic must drive the interaction between neighboring molecules to produce nanodomains and rafts.

10.3.6: Lipid Phase Diagrams

You are familiar with the phase diagrams of water from introductory chemistry classes. Phase diagrams show the different phases that are accessible under different sets of conditions such as temperature and pressure. A traditional phase diagram for water is shown in Figure 10.3.18

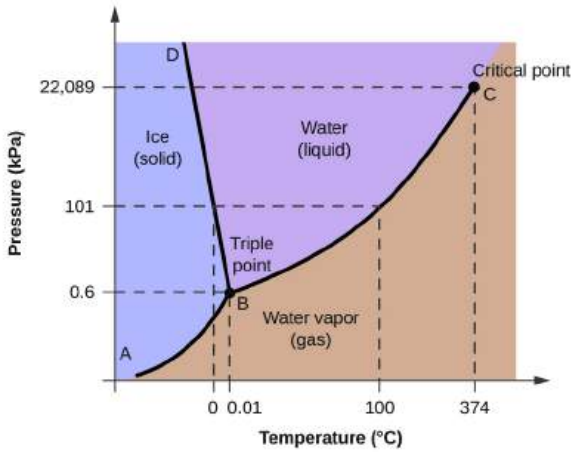


Figure 10.3.18: A traditional phase diagram for water. https://opentextbc.ca/chemistry/chap...hase-diagrams/. Chemistry by Rice University is licensed under a Creative Commons Attribution 4.0 International License.

The horizontal dotted line at 101 kPa shows the states of water as a function of temperature at 101 kPa = 1 atm pressure. The phase transition of solid to liquid water occurs at 0^0 C (freezing/melting point of water) while the liquid to gas transition occurs a 100^0 C (boiling point of water). At a reduced pressure (0.6 atm), all three phases of water can exist at 0.01 ^{0} C, the triple point of water.

In an analogous fashion, lipid bilayers have phase transition diagrams as well. Instead of showing phases as a function of temperature and pressure, they are usually shown as a function of temperature and concentration of a specific lipid component such as cholesterol, which as described above affects T_M, fluidity, and raft formation. An example of a **theoretical** phase diagram for membranes composed of saturated dipalmitoylphosphatidylcholine (16:0), the most common saturated fatty acid in animals, plants, and microorganisms vs cholesterol content is shown in Figure 10.3.19

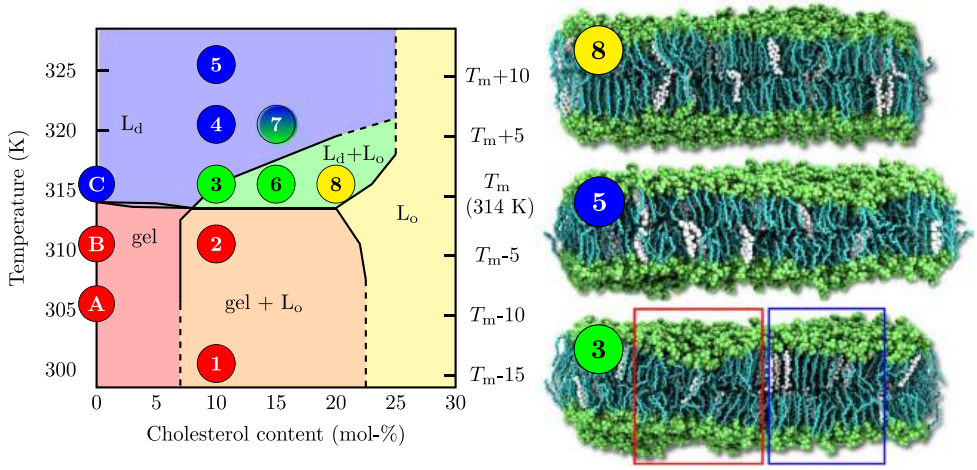


Figure 10.3.19: Javanainen, M., Martinez-Seara, H. & Vattulainen, I. Nanoscale Membrane Domain Formation Driven by Cholesterol. *Sci Rep* **7,** 1143 (2017). https://doi.org/10.1038/s41598-017-01247-9. **https://rdcu.be/chbtH.** http://creativecommons.org/licenses/by/4.0/.

DPPC vesicles have a T_M = 41 ^{0}C or 314 K. The phase diagrams show that when cholesterol is added to DPPC bilayers at temperatures above T_M, the bilayers change from the liquid-disordered (Ld) phase to the uniform liquid-ordered (Lo) phase at around 20 mol% cholesterol. At cholesterol concentrations between around 10-20 mol% and temperatures just above the T_M, both the L_0 and L_d phases coexist. The coexistence of two phases mimics a raft. At low cholesterol levels, the bilayer changes from the gel (or L_β) to the L_d phase. At really high cholesterol, only the L_0 phase exists. This makes sense as the fatty acids are all saturated and cholesterol's rigid rings reduce the freedom of the acyl chains to rotation and hence decrease the number of chains in the gauche conformation. Yet as an "impurity" (cholesterol) has been added to the system, the system is less rigid, and more fluid-like. From 0-7 mol% cholesterol, two phases exist, the gel (or L_β) and the L_d phase. Between 7-23 mol% cholesterol, a combination of phases is seen.

As mentioned above, the outer leaflet of mammalian plasma membranes is composed mostly of sphingomyelin (SM) and PC, while the inner leaflet is composed mostly of phosphatidyl ethanolamine (PE) and phosphatidyl serine (PS), along with cholesterol. These

don't appear to separate into L_d and L_0 phases and appear not to form nanodomains or rafts. In the lab, outer membrane lipids easily form vesicles but the inner leaflet polyunsaturated PEs form other phases (hexagonal or cubic). It appears that asymmetric lipid distribution is critical for biological bilayer formation.

Since nanodomains are difficult to observe and study experimentally, molecular dynamics simulations are used to provide insight into their structure and properties. The right-hand images to the right show snapshots of the membrane from molecular dynamic simulations. Cholesterol is shown in white in Figure 10.3.19 **8** shows the membrane in the Lo phase, **5** shows it in the Ld phase (note the reduced membrane width), and **3** in an ordered (blue box) and disordered (red box) phase. The simulations support the theoretical phase diagram showing the coexistence of the L_d and L_0 phases as well as finding a hexagonal-closest packed cholesterol poor with L_d domains in the region of the phase diagram showing the coexistence of both L_d and L_0 phases. In addition, cholesterol is excluded from most ordered regions. This is shown in Figure 10.3.20 showing a top-down view of the membrane. The dark green lipid headgroups are those that are hexagonally closest packed, an ideal you will remember from introductory chemistry courses.

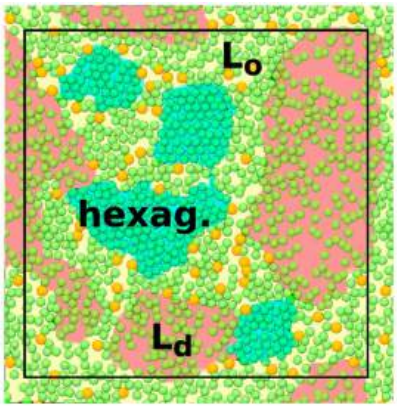


Figure 10.3.20: Top-down view showing membrane phase observed in a molecular dynamics simulation. Javanainen, M. et al. ibid.

Now imagine what a phase diagram would look like for a three (DOPC, sphingomyelin, and cholesterol) or more component system! We won't show any but from the "simple" two-component system described above, it should be evident that we have a long way to go before understanding the complexity of membrane bilayers. Ternary bilayers system often from more macroscopic (vs nanoscopic) domains (or rafts) which can be studied using fluorescence microscopy.

The actual biological membrane must be able to adopt very nonplanar shapes with positive and negative curvature. Membranes must also be able to fuse (for example the egg and sperm). Another lipid phase, which we have not yet discussed, may be involved. New lipid phases of a single membrane lipid can form based on the relative percentages of lipid and water. These include hexagonal phases. Figure 10.3.21 shows an image of a hexagonal phase of phosphatidylethanolamine with 16:0 fatty acids. Note the water inside of the middle ring of PE molecules.

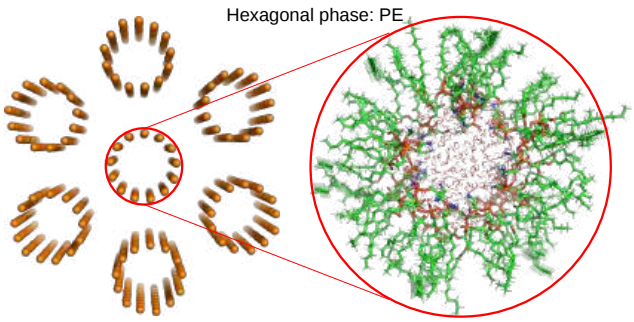
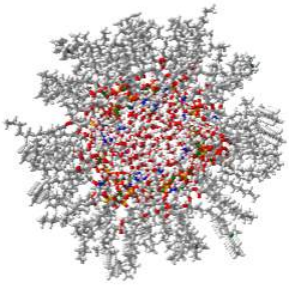


Figure 10.3.21: Hexagonal phase of phosphatidylethanolamine with 16:0 fatty acids

Figure 10.3.22 shows an interactive iCn3D model shows the inner ring outlined in the red circle. Note that water surrounds each of the closest packed lipid tubules which extend back into the figure in 3D. This phase creates an aqueous channel through the interior of each tubule.



NCBI iCn3D Figure 10.3.22 Hexagonal phase of phosphatidylethanolamine with 16:0 fatty acids (Copyright; author via source). Click the image for a popup or use this external link: not available

Figure 10.3.23shows variants of the hexagonal phase as well as a cubic phase of lipids.

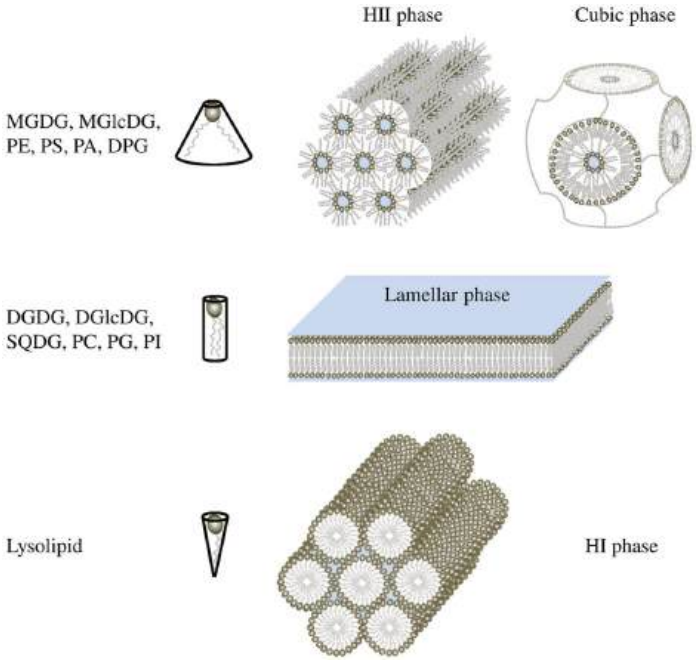


Figure 10.3.23: Hexagonal and cubic phases of lipids. Juliette Jouhet. Importance of the hexagonal lipid phase in biological membrane organization.. Frontiers in Plant Science, Frontiers, 2013, 4, pp.494. ff10.3389/fpls.2013.00494ff. ffhal-00942927f. Creative Commons.

Double-chain amphiphiles with a small polar head (like PE and PA) are more likely to form a hexagonal II phase with elongated tubules than are more cylindrical lipids like PC. Note that if a single fatty acid is removed from a double-chain amphiphile like PE, a narrow conical shape single-chain amphiphile arises, which can either form a micelle (not shown) or a hexagonal I phase. *In vitro*, equimolar amounts of PC (which forms the laminar phase) and PE (which can form an HII phase) can either form the HII phase(at low aqueous pressure) or a lamellar phase (at high aqueous pressure).

Are hexagonal lipid phases found in biological membranes? The answer is likely yes in the formation of unusual cellular structures. The plant plasmodesmata, shown in Figure 10.3.24 is one such structure. Just focus on the two bilayers connected by the membrane-lined membranes of the channel.

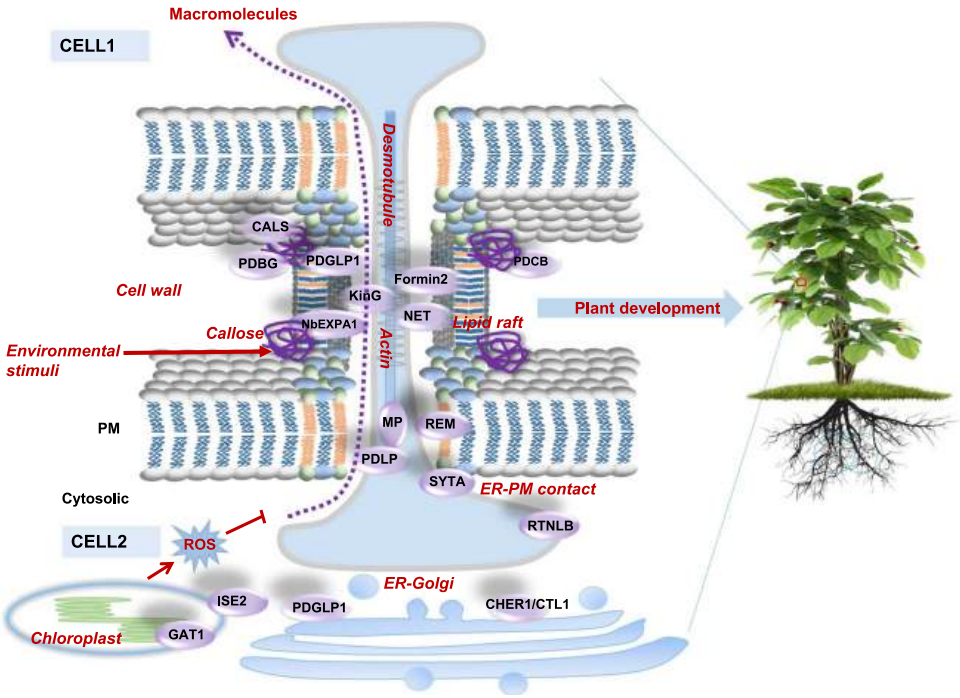


Figure 10.3.24 Plant plasmademata. Yanbiao Sun et al. https://www.nature.com/articles/s41438-019-0129-3. Creative Commons license. http://creativecommons.org/licenses/by/4.0/.

Plasmodesmata are membrane-lined channels that cut across the plant cell wall and directly connect cells, allowing the flow of water and nutrients between the cells. HII phases have been seen in the endoplasmic reticulum membrane. These membranes, along with the mitochondrial inner membrane and the inner membrane of chloroplasts are highly curved and contain higher concentrations of lipids that allow that curvature as well as the formation of H II phases. The function of some membrane enzymes as well as processes such as membrane fusion and fission are enhanced by HII-forming lipids.

10.3.7: Piecing it all together

Our emerging understanding of lipid structure has taken us from micelles and vesicles to the complexity of actual biological membranes with different phases and nanodomain structures. This complexity is needed as membranes must be dynamic in ways the proteins, for example, aren't. They must be able to pinch off either as extracellular or intracellular vesicles, for example, in the process of exocytosis and endocytosis. Both positive and negative curvatures of the membrane must be enabled. The incredible complexity of the "Lego-like" lipid monomers that assemble and rearrange into every fluctuating membrane is yet another exquisite illustration of our repeat mantra, that structure and shape mediate all function. The "Lego-like" membrane monomers in various phases and regions of membrane curvature are illustrated in Figure 10.3.25

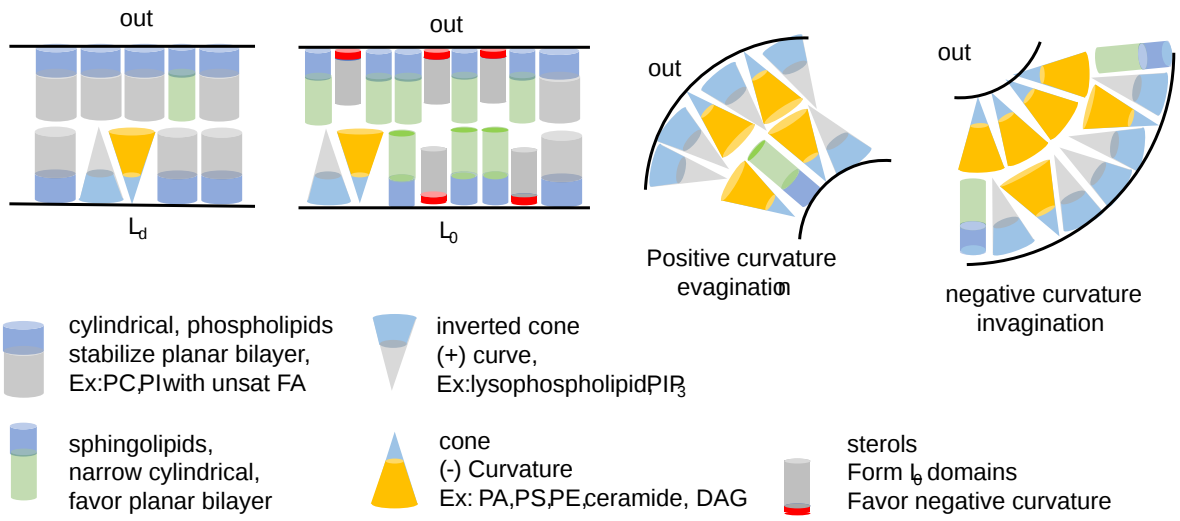


Figure 10.3.25: "Lego-like" membrane monomers in various phases and regions of membrane curvature

Given their shape, some glycerolipids don't even appear to spontaneously form bilayers by themselves. The shape and type of fatty acids found in the membrane lipids will determine the local properties of the membrane, its phases, and the presence of nanodomains. The activity of membrane enzymes, and membrane fission and fusion events, will also depend on the local properties of membranes. The figure above. doesn't even account for the presence of peripheral and integral membrane proteins, which we will discuss in the next section.

As we learned with proteins, we can be misled by looking at beautiful but static images of protein. Their dynamic motion is critical to their function. In addition to the dynamic membrane events discussed above, there is a myriad of other events that take place at and in membranes. Here are a few.

10.3.8: Other dynamic events in bilayers

Membranes are not static. They are synthesized, their contents shuffled, they fuse with other membranes, and large proteins are inserted into them. Let's explore some of these.

10.3.8.1: **Membrane Trafficking**

Movement of key "cargo" molecules into (endocytosis) and out of (exocytosis) the cells occurs mostly through membrane-encapsulated vesicles. Vesicles contain all types of biological molecules including lipids, both synthetic and absorbed. Part of the differences in lipid composition between membrane layers and between different organelles derived from this highly orchestrated and controlled movement of vesicles. Details are shown in Figure 10.3.26 below.

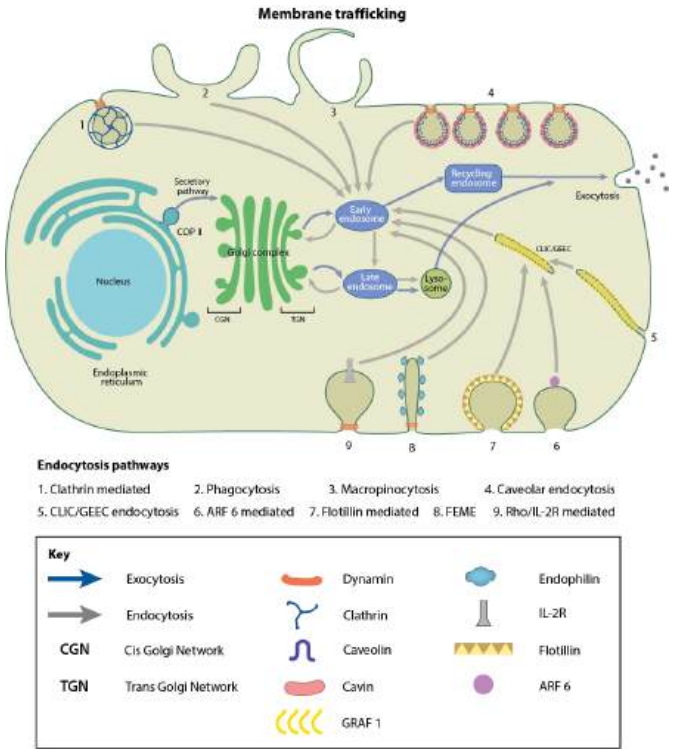


Figure 10.3.26: Players in membrane trafficking. https://www.mechanobio.info/what-is-...e-trafficking/. Creative Commons Attribution-NonCommercial 4.0 International License

In eukaryotes, the biosynthetic secretory pathways move molecules from the endoplasmic reticulum to the cis Golgi (CGN) to the trans-Golgi (TGN) and to the plasma membrane (for integral membrane proteins) or for secretion. Since most lipids are synthesized in the endoplasmic reticulum (ER), their distribution to different locations in cells is critical in maintaining the asymmetric distribution of lipids found in cells.

10.3.8.2: Fusion of membranes: Fusion Peptides

Another dynamic event in membranes is the fusion of two bilayers from two different cells vesicles or of a vesicle and cell membrane. These events are facilitated by fusion peptides. Figure (\PageIndex{27}\) below shows a molecular dynamics simulation snapshot showing how a fusion peptide in a single DMPC bilayer causes a constriction of the bilayer with the two leaflets approaching each other.

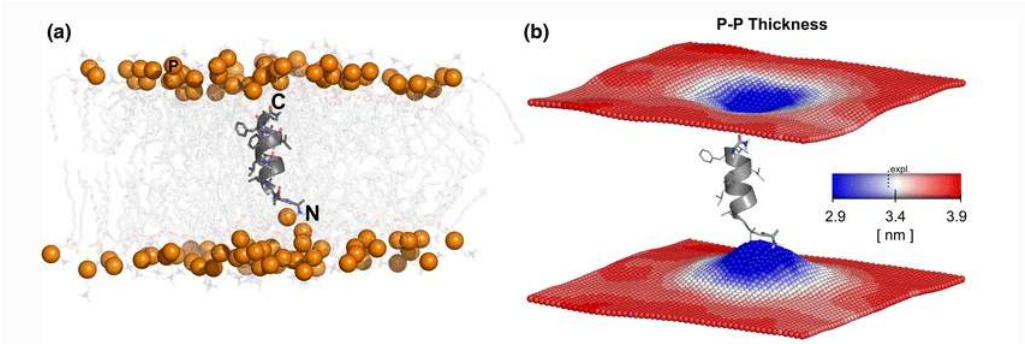


Figure (\PageIndex{27}\): A molecular dynamics simulation snapshot showing how a fusion peptide resides in a single bilayer. The first 12 amino acids of a GP41 peptide are inserted in a DMPC membrane patch.

Figure (\PageIndex{28}\) below shows the area per lipid (APL).

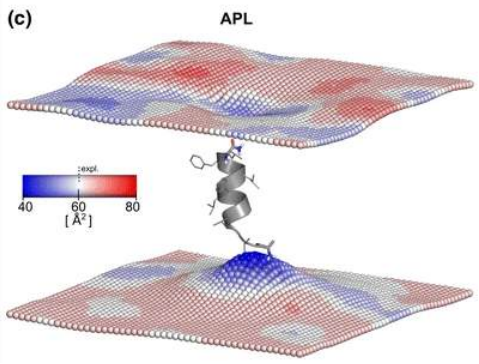


Figure (\PageIndex{28}\): Area per lipid (APL) of a DMPC membrane with an inserted fusion peptide.

10.3.8.3: Insertion of Membrane Proteins

Another dynamic event is the insertion of a membrane protein. Now let's look at changes in the bilayer on insertion of the voltage-dependent anion channel (VDAC) membrane protein, as shown in Figure (\PageIndex{29}\) below.

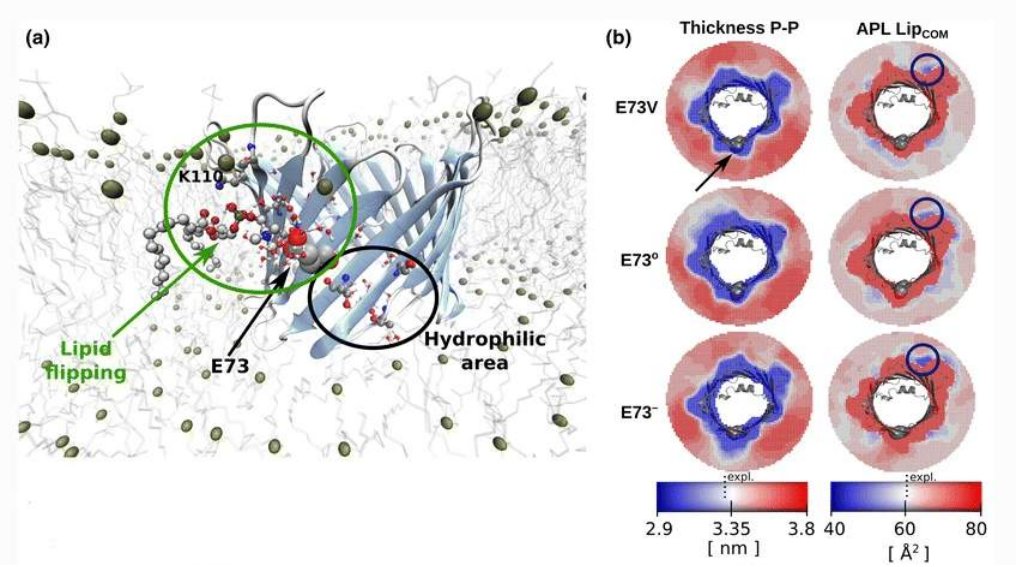


Figure (\PageIndex{29}\): Local membrane properties of a bilayer with an embedded membrane protein VDAC. CVytautas Gapsys et al, https://link.springer.com/article/10...822-013-9684-0. Creative Commons CC BY license.

Panel a shows the voltage-dependent anion channel (VDAC) is shown in a cartoon β-barrel representation. Residue E73 and the water molecules nearby are represented as spheres. Serine and threonine residues constituting hydrophilic areas close to E73 are shown in ball and stick representation. In the snapshot, a DMPC lipid is shown flipping close to the E73 and K110 residues.

Panel b shows top view perspectives of the circular representation of the local thickness, calculated considering phosphorus atoms, and area per lipid.

10.3.9: Introduction to Lipid Signaling - Chemical Cleavage of Membrane Lipids

Everything in this chapter so far describes the structure and dynamics of the components of the lipid components of a bilayer. The dynamic changes described involve the physical movement of lipids molecules in the membrane. Let's briefly introduce another type of movement that involves not only physical but chemical changes in the cell. What happens when specific lipids in membranes are chemically cleaved by **lipases**, enzymes which are analogous to protease? It turns out these changes lead to signaling within the cell. We will only briefly introduce lipid signaling in this chapter, but explore it in more detail in Chapter 12.

Lipids are not just used as a passive component of membranes, or as a source of stored energy. They are involved in the process of signal transduction at the cell membrane, a process by which the interior components of the cell respond to a signal external to the cell, allowing the cell to respond to its local environment. Usually, a chemical signal on the outside of the cell is the "primary messenger" that causes the cell to respond. Usually, the chemical transmitter of information does not get into the cell. Rather it binds to surface receptors on the cell membrane surface. Somehow, the cells sense that a ligand is bound to the outside. Enzymes,

usually in the membrane or at the intracellular surface of the lipid bilayer are activated. Many of these enzymes cleave lipids in the membrane. The cleaved fragments of the lipid molecules serve as intracellular signals or "secondary messengers", which can bind to intracellular enzymes to activate intracellular processes. Figure (\PageIndex{30}\) below shows some of the lipid mediators which are generated by the process and signal the cell to respond.

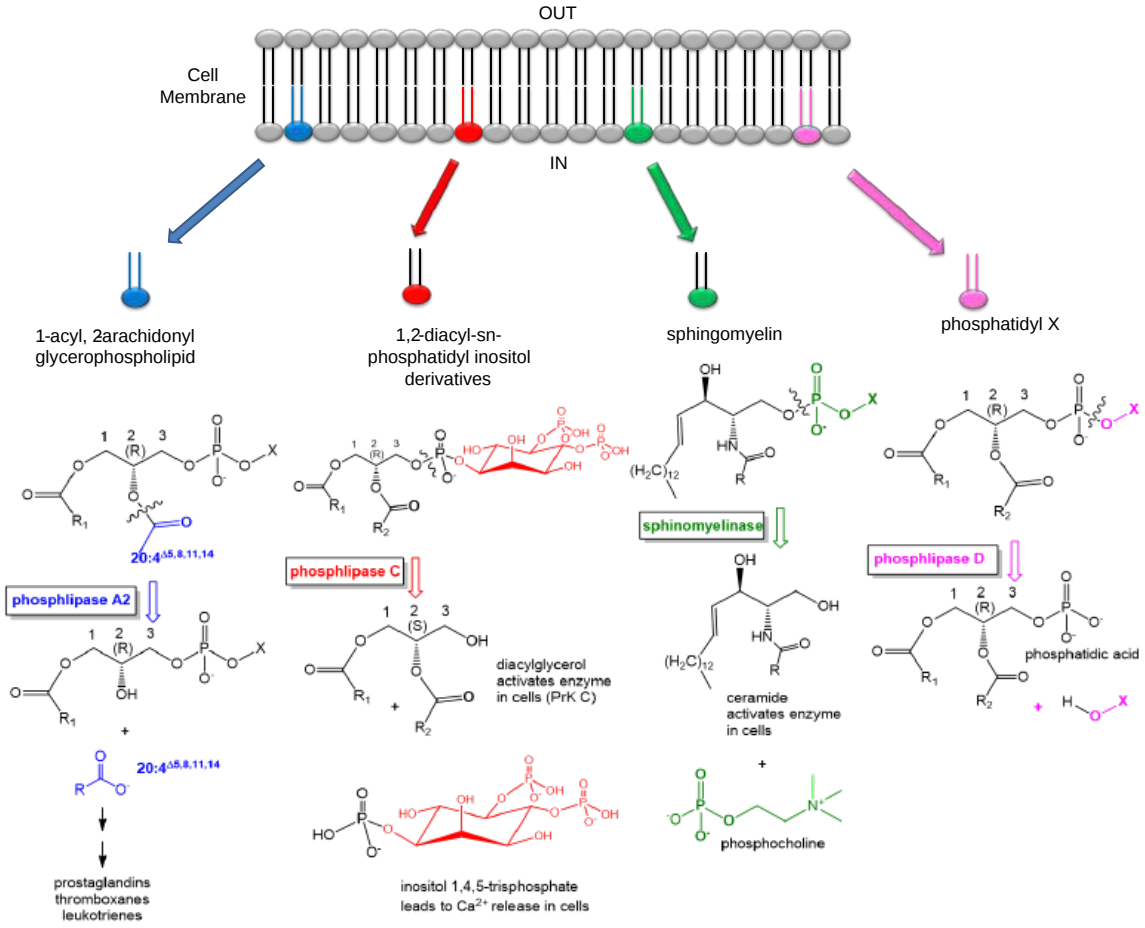


Figure (\PageIndex{30}\): Membrane lipids involved in signaling

Fatty acid amides are potent mediators of neurological processes. In one interesting experiment, sheep were sleep deprived. Reasoning that the brain might release a biochemical signal into cerebrospinal fluid to induce sleep, scientists at Scripps removed some of this fluid and isolated a substance that was not found in rested sheep. On analysis, the structure was shown to be an amide of oleic acid. Oleylethanolamide has been shown to bind to the peroxisome-proliferator-activated receptor-a (PPAR-a) which resides in the nucleus. This ligand, by affecting gene transcription, appears to regulate body weight and the feeling of fullness after eating (satiety) as it leads to reduced eating.

People have sought the natural neurotransmitter which binds to the same receptor in the brain as THC, the active ingredient of marijuana. The endogenous cannabinoid is an amide of arachidonic acid, anandamide. Figure (\PageIndex{31}\) below shows the structures of key fatty acid amides and THC.

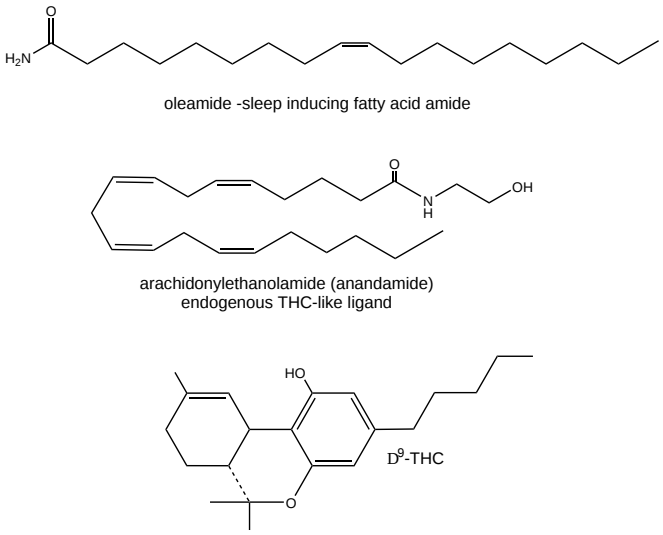


Figure (\PageIndex{31}\): Fatty Acid Amides: Neurochemical Mediators

This fatty acid amide is an example of a class of lipid derivatives called N-acylethanolamines (NAEs). These molecules, with acyl groups that vary in the number of carbons and double bonds, are found widely in organisms in nature. Naturally occurring anandamide leads to increased food intake after a short period of reduced food intake. One of the known physiological effects of THC is increased food consumption (the munchies).

Lucanic et al (2011) have shown that decreases in NAEs extend the life span of the small roundworm C. elegans, which has become a model organism to study genes in eukaryotes. Caloric restriction has been shown to increase the life span in a variety of organisms. In invertebrates, anandamide seems to inhibit food intake, even in organisms that lack a receptor similar to which cannabinoids bind. This might seem paradoxical in that anandamide (and THC) in humans seems to induce eating. However, under long periods of caloric restriction (low-level starvation) in rats, anandamide levels are suppressed, leading to a low energy-consuming state.

In general, it appears that reductions in NAEs occur during periods of caloric restriction. Mutant worms which have reduced levels of NAEs through targeted enzyme disruptions that affected either NAE synthesis or degradation have longer life spans. If normal (wild-type) worms were placed under caloric restriction but given EPA-ethanolamine (the most abundant NAE in these worms), they did not have an extended lifespan.

10.3.10: Membrane Monolayers - Lipid droplets and lipoproteins

We explored membrane bilayers that contain two leaflets above. These bilayers separate the outside and inside of cells as well as the outside and inside of internal organelles. It turns out that there are two main types of lipid-encapsulated structures in which only one phospholipid leaflet separates the interior contents from the outside. These are lipid droplets and lipoproteins. In both cases, the monolayers encapsulate TAGs and cholesterol esters - hydrophobic molecules, so there is no need for another inner leaflet to stabilize an encapsulated aqueous environment.

10.3.10.1: Lipid Droplet formation

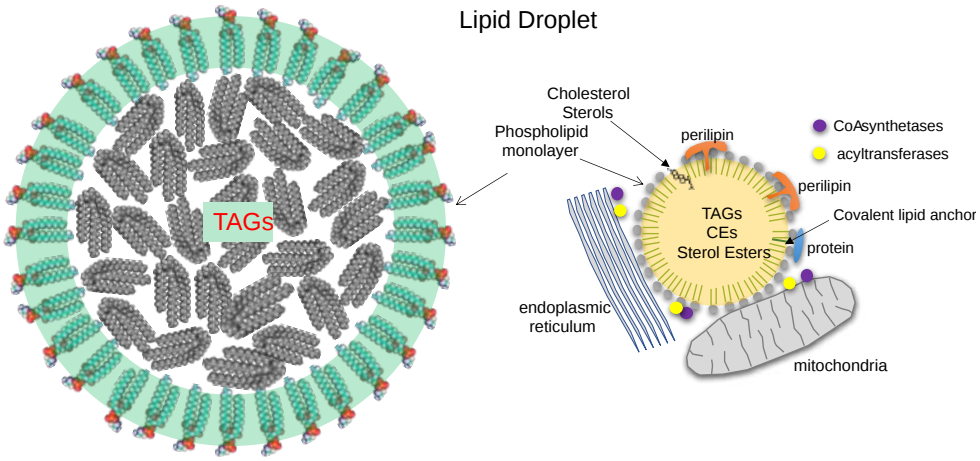
So how are triacylglycerols stored in cells? In lipid droplets! In contrast to the common single and double-chain amphiphiles, which have charged atoms in their head groups and which form micelles and bilayers, respectively, triacylglycerols (TAGs) and cholesterol esters (CEs), which are almost completely nonpolar, coalesce into lipid droplets in cells. These droplets can range from very big, which are found in adipocytes (fat cells) where they take up almost all of the available space and where they are used for energy storage, to small, which are found in all cells, where they are used mostly for membrane biogenesis and energy mobilization. When esterified into esters, fatty acids, and steroids also pose less potential toxicity to cells. Lipid droplets are often found in close approximation or attachment to mitochondria, endoplasmic reticulum (ER), and peroxisomes (where plasmalogens with ether-linked fatty acid instead of ester-linked are synthesized), all organelles intimately involved in membrane and energy biochemistry. Many of the enzymes (acyltransferases for example) required for TAG metabolism are found in the mitochondria and the ER.

Since lipids droplets are specialized to cells and don't form if TAGs are just added to water, we'll discuss their structures and their dynamic assembly in this section.

The droplets are now considered actual cellular organelles. In contrast to other organelles which are bounded by a bilayer member, lipid droplets are surrounded by a monolayer of phospholipids which prevents exposure of the nonpolar contents to the aqueous cytoplasm. PC and PE appear to be the major phospholipids in the monolayer and both are synthesized mostly in the ER. There are many different proteins involved in the formation and interaction with lipid drops, including

- perilipins: There are multiple types of perilipins. Perilipin 1 is found in adipocytes and cells that synthesize steroids (adrenals, ovaries, and testes). Perilipin 2 and 3 are found in most cells
- Acyl CoA synthetases and acyltransferase: These enzymes activate free fatty acids for metabolic processes.
- seipins: These are involved in lipid droplet shape number, and size. It appears to be involved in the formation of lipid drops and moving them from the ER to the cytoplasm.
  Facilitates initiation of LD formation, and ensures that vectorial budding of LDs from the ER is directed toward the cytoplasm

It appears that lipid droplets arise from the ER which are involved in membrane biogenesis and "trafficking" of membranes to different locations in the cell. A general structure of a **lipid droplet** is shown in Figure 10.3.32 The left figure shows just internal triacylglycerols (TAGs), but the inside would also contain cholesterol ester (fatty acid esterified to the cholesterol OH) and the monolayer would also contain unesterified cholesterol.



(\PageIndex{32}\): Cartoon view of a lipid droplet.

How might these complex lipid droplets form in a cell if not by phase separation? With the help of proteins. of course. Figure 10.3.31shows how newly synthesized TAGs (made in the ER membrane, could self-aggregate in the bilayer to form a "lens" which on further growth and addition of lipid binding proteins could bud off into the cytoplasm to form the droplets.

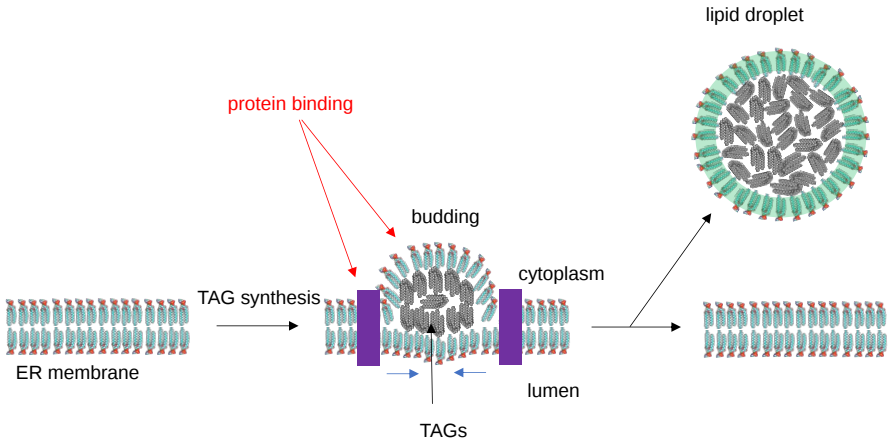


Figure 10.3.33: Formation of a lipid droplet

Figure 10.3.33 shows an incredible image of diacylglycerols (DAG) accumulating in a bilayer to form a clear lens in the membrane. This was produced by a molecular dynamics simulation.

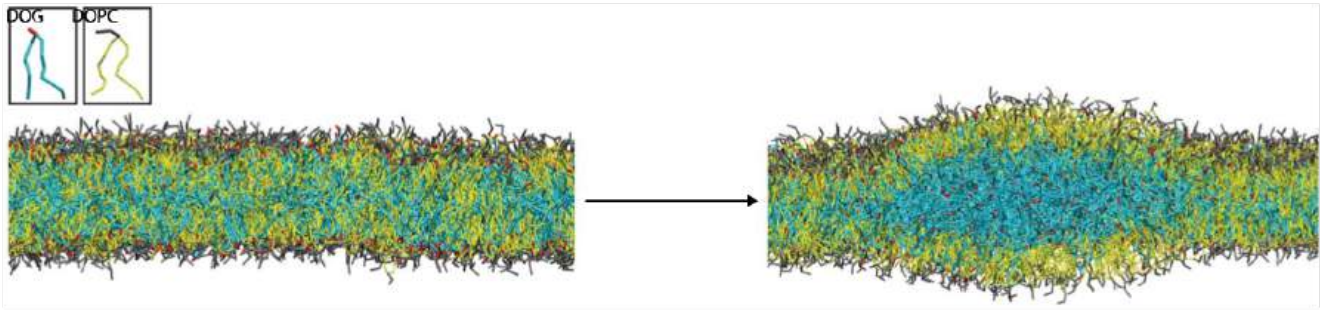


Figure 10.3.33: Molecular dynamic simulation of diacylglycerols (DAG) accumulating in a bilayer to form a clear lens in the membrane. Pablo Campomanes, Valeria Zoni & Stefano Vanni. Nature. https://www.nature.com/articles/s42004-019-0175-7.pdf. Creative Common License: http://creativecommons.org/ licenses/by/4.0/.

Figure (\PageIndex{34}\) below shows a cartoon illustrating the synthesis of a lipid droplet from the ER membrane.

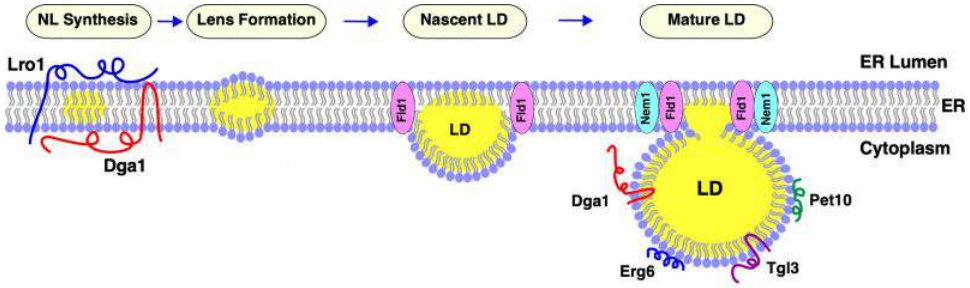


Figure (\PageIndex{34}\): **LD biogenesis in the ER.** ER-localized triacylglycerol-producing enzymes, Lro1, and Dga1 catalyze neutral lipid (NL, indicated in yellow) synthesis from opposite sides of the ER membrane. The NL then accumulates between the two leaflets of the ER membrane leading to the formation of lipid lenses. These NL lenses grow in size to become nascent LDs, which then emerge towards the cytoplasm where they further mature. Fld1 and Nem1 proteins show punctate localization at ER-LD contact sites. The acyltransferase Dga1 and the TAG lipase Tgl3 translocate onto the periphery of mature LDs. Finally, LD-marker proteins such as the perilipin ortholog Pet10 and the sterol biosynthetic enzyme Erg6 decorate the surface of the mature LD. http://microbialcell.com/researchart...icrobial-cell/ . Creative Commons Attribution 4.0 International License.

A membrane protein not displayed in the above figure, seipin, denotes the location for lipid droplet formation in the ER membrane. Mutants of the protein are associated with lipodystrophy. In yeast, seipin and another membrane protein Ldb16, associate to allow lipid droplet formation. Seipin aggregates to form a homo 10-mer in the membrane but in contrast to human cells, it alone can not concentrate triacylglycerol in the membrane. It requires the binding of another protein, Ldb16, for that to happen.

Figure (\PageIndex{35}\) below shows yeast, human and fly seipin and their properties.

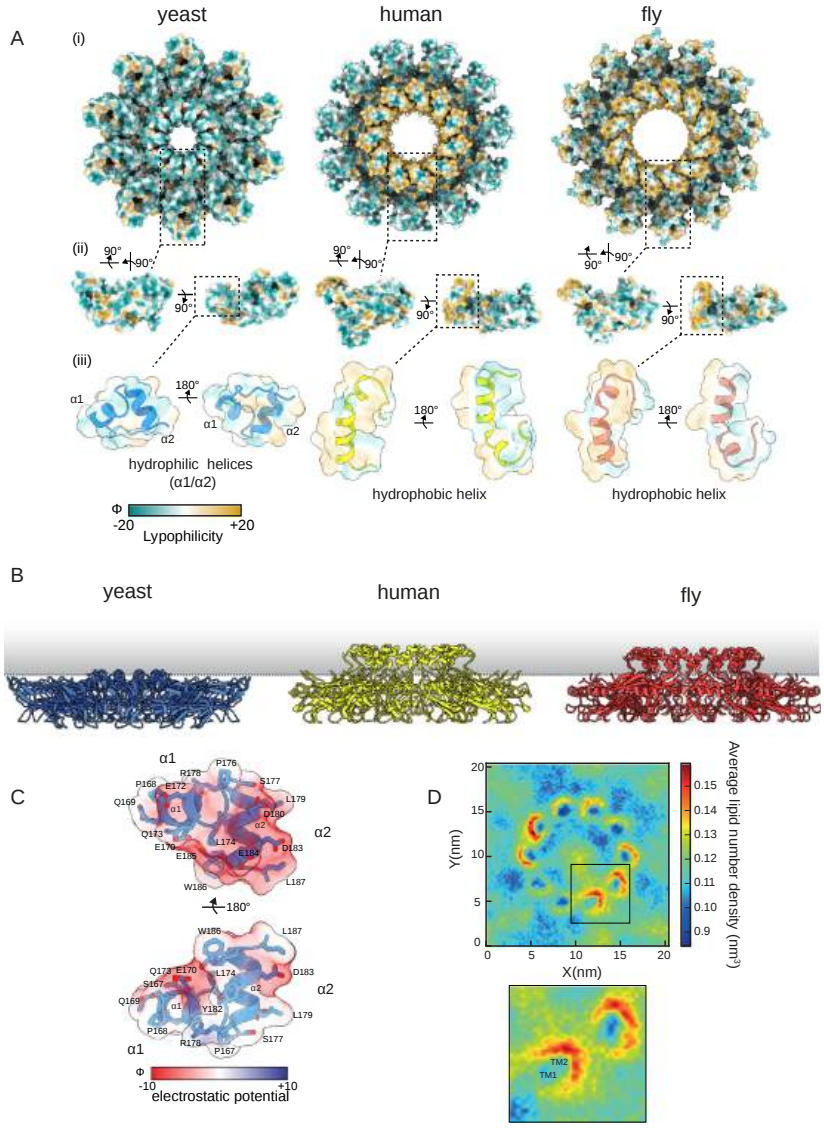


Figure (\PageIndex{35}\): Yeast, human and fly seipin and their properties. *Nat Commun* **12**, 5892 (2021). https://doi.org/10.1038/s41467-021-26162-6. Klug, Y.A., Deme, J.C., Corey, R.A. *et al.* Mechanism of lipid droplet formation by the yeast Sei1/Ldb16 Seipin complex. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Panel **A** shows a lipophilicity potential on the surface of yeast (left), human (middle), and fly (right) as viewed as (i) homodecamer assembly from the cytosol, (ii) individual protomers, or (iii) transparent overlay over zoomed in cartoon representation of the central α1-α2 helices. Surfaces are colored from hydrophilic (dark cyan) to hydrophobic (gold).

Panel B shows a side view of the luminal domains of yeast, human (PDB 6DS5) and fly (PDB 6MLU) Seipin in relation to the plane of the ER membrane (indicated by a dotted line).

Panel C shows the charge distribution of the yeast Sei1 central helices (α1, α2), depicted as a transparent Coulombic electrostatic potential surface representation (Red, negative charge; blue, positive charge; white, no charge) overlayed over a cartoon representation (light blue) to show acidic side chains.

Panel D shows a top view of molecular dynamic simulations of Sei1 in a POPC membrane with 3% trioleylglycerol. Images depict the average lipid number density of trioleylglycerol. Inset – zoom of the corresponding box showing positions of TM1 and TM2.

The lumenal domains form the ring with a floor as shown in Panel A above. In addition, the transmembrane segments for the cage top and sides. A switch area between the lumenal and transmembrane segments which can occupy two different conformations

appears important for function. The closed cage allows accumulation and hence phase-separation of triacylglycerols, while the open form allows the nascent droplet to grow and then bud.

The Ldb16 has helical regions enriched in serine and threonine, as is required for TAG loading. These -OH-containing amino acids are present in seipins in humans and flies. Site-specific mutations of the serine and threonines in the region of Ldb16 lead to problems with lipid droplet formation.

Figure (\PageIndex{36}\) below shows a cartoon of the serine- and threonine-enriched helix in Ldb16.

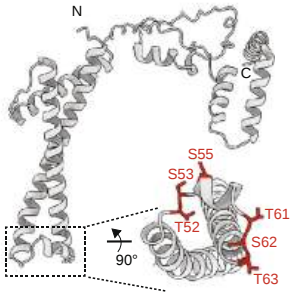


Figure (\PageIndex{36}\): Cartoon of Ldb16 predicted helix enriched in serine and threonine. Klug et al, ibid.

Another model for the assembly of lipid droplets by the Sei1-Ldb16 yeast complex is shown in Figure (\PageIndex{37}\) below.

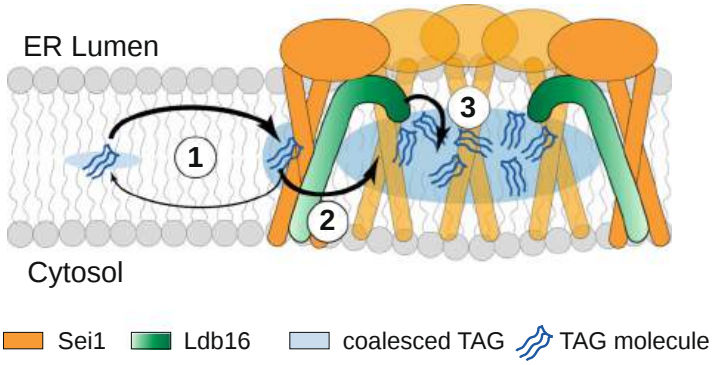


Figure (\PageIndex{37}\): Sequential TAG interactions mediate LD assembly by the Sei1-Ldb16 complex. In the ER bilayer, TAG molecules (blue) concentrate in the proximity of Seipin oligomers (orange) via weak interaction with Sei1 TMs. TAG molecules within the ring interact strongly with Ldb16 (green) hydroxyl-containing residues, facilitating TAG coalescence and lens formation. Klug et al, ibid.

Figure 10.3.38shows an interactive iCn3D model of the homo 10-mer yeast seipin membrane complex **(7OXP).**

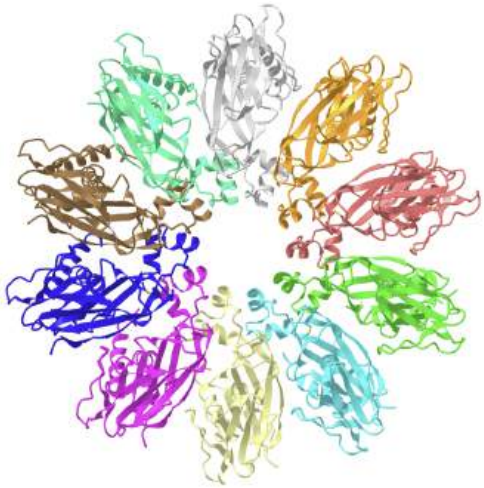


Figure 10.3.38 Yeast seipinhomo 10-mer yeast seipin membrane complex (7OXP). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...wMLHBncjFZMbFA

The two transmembrane segments (helices) from each of the monomers are evident by rotating the structure so the homo 10-mer ring is viewed from the side.

10.3.10.2: Lipoproteins

We eat, digest, and transport dietary fat. We also make fat and transport it through the blood as well. We saw that free fatty acids are carried in the blood by the most abundant protein in the blood, albumin. What about the very insoluble triacylglycerols and cholesterol esters? Turns out they are also transported in the blood by nanoparticles similar to lipid droplets. They are called lipoproteins since, like lipid droplets, they have proteins associated with them.

Lipoproteins vary in density and size. The densest is called high-density lipoproteins (HDL). As they get larger and more filled with, they form less dense lipoproteins (low density - LDL, intermediate density - IDL, and very low density - VLDL). These contain nondietary lipids made by organs like the liver. Dietary fats are processed in the intestine into very large particles called chylomicrons. Their relative size and density are shown in Figure 10.3.39

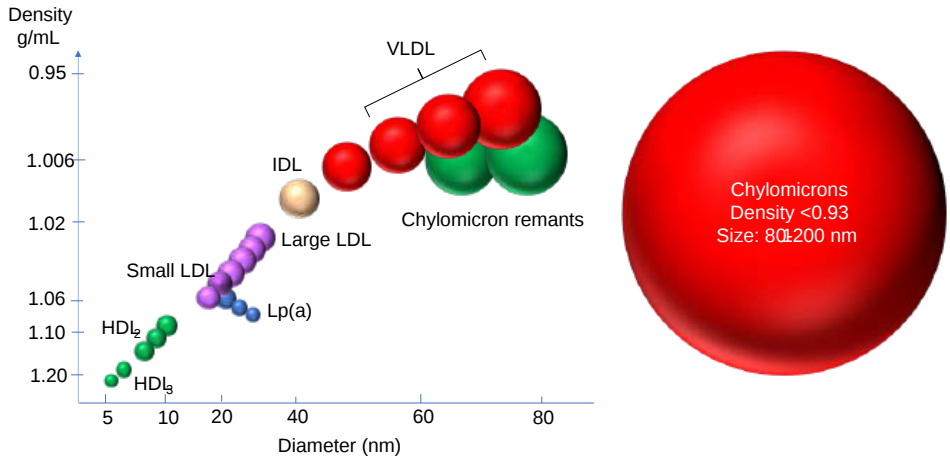


Figure 10.3.39: Size distribution of lipoproteins

Introduction to Lipids and Lipoproteins. Kenneth R. Feingold, MD. Creative Commons (CC-BY-NC-ND) license. A copy of the license can be viewed at http://creativecommons.org/licenses/by-nc-nd/2.0/. With permission.

As with lipid droplets, lipoproteins have a single outer monolayer leaflet containing double chain membrane lipids like phosphatidylcholine and free cholesterol Inside are the triacylglycerols and cholesterol esters. Proteins are bound to the outer monolayer. Figure 10.3.40 shows two renderings of discoidal HDL particles containing a single type of protein, Apo-A1. The TAGs are shown in cyan line rendering on the inside, along with cholesterol esters (in spacefill). The bottom part of the figure shows the polar Ns and Os decorating the outer part of the monolayer of phosphatidylcholine surrounding the TAGs and cholesterol esters.

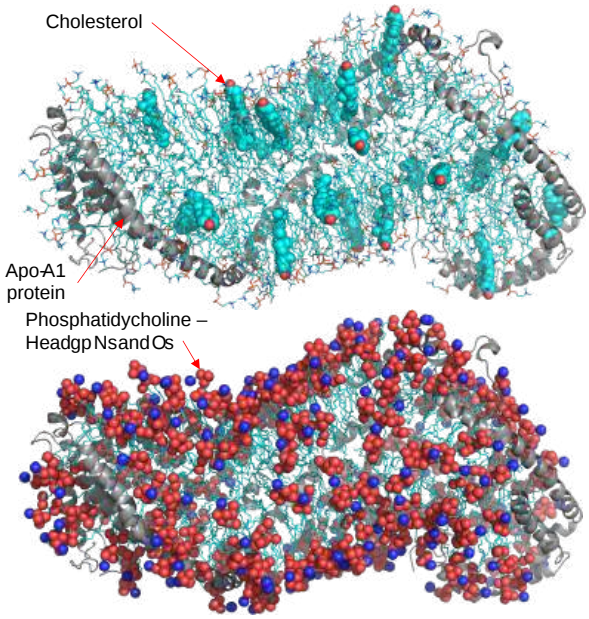
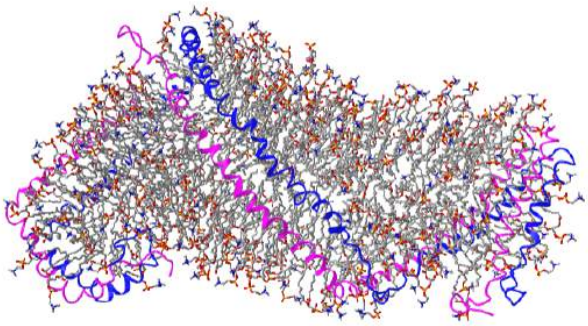


Figure 10.3.40: Two renderings of discoidal HDL particle containing a single type of protein, Apo-A1

Figure 10.3.41 shows an interactive iCn3D model of discoidal HDL (3k2s)



NCBI iCn3D Figure 10.3.41: discoidal HDL (3k2s) (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...BE5XYUopbjGALA

Table 10.3.1 below shows the proteins associated with the different types of lipoproteins.

Lipoprotein	Density (g/ml)	Size (nm)	Major Lipids	Major Apoproteins
Chylomicrons	<0.930	75-1200	Triglycerides	Apo B-48, Apo C, Apo E, Apo A-I, A-II, A-IV
Chylomicron Remnants	0.930- 1.006	30-80	Triglycerides Cholesterol	Apo B-48, Apo E
VLDL	0.930- 1.006	30-80	Triglycerides	Apo B-100, Apo E, Apo C
IDL	1.006- 1.019	25-35	Triglycerides Cholesterol	Apo B-100, Apo E, Apo C
LDL	1.019- 1.063	18- 25	Cholesterol	Apo B-100
HDL	1.063- 1.210	5- 12	Cholesterol Phospholipids	Apo A-I, Apo A-II, Apo C, Apo E
Lp (a)	1.055- 1.085	~30	Cholesterol	Apo B-100, Apo (a)

Table 10.3.1: Proteins associated with the different types of lipoproteins.Introduction to Lipids and Lipoproteins. Kenneth R. Feingold, MD. Creative Commons (CC-BY-NC-ND) license. A copy of the license can be viewed at http://creativecommons.org/licenses/by-nc-nd/2.0/.

10.4: Working with Lipids

10.4.1: Introduction

Lipids, although small compared to large biopolymers like proteins, nucleic acid, and large glycans, are very heterogenous in structure, given the large array of fatty acid and isoprenoid chain lengths, numbers of double bonds, etc that appear in different lipid classes. In addition to analyzing lipid structure, lipids are used in the laboratory to create liposomes (vesicles), which serve as models for membrane bilayers and encapsulation of chemical species (drugs, vaccines) for medical use and solubilization of membrane proteins. In this section, we will concentrate on the creation of lipid vesicles (critical for the encapsulation of RNA vaccines against the SARS CoV-2 spike protein) and the chemical analysis of biological lipids, whose composition affects health and disease states.

10.4.2: Liposomes

Liposomes produced in the lab can be unilamellar, consisting of a single bilayer surrounding the internal aqueous compartment, or multilamellar, consisting of multiple bilayers surrounding the enclosed aqueous solution. You can imagine the multilamellar vesicles resemble an onion with its multiple layers. Cartoons of unilamellar and multilamellar liposomes are shown in Figure 10.4.1, where each concentric circle represents a bilayer.

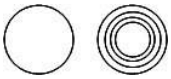


Figure 10.4.1: Cartoons of unilamellar and multilamellar liposomes (each concentric circle represents a bilayer)

Liposomes vary in diameter. They can be generally categorized into small (S, diameter < 25 nm), intermediate (I, diameter around 100 nm), and large (L, diameter from 250-1000 nm). If these vesicles are unilamellar, they are abbreviated as SUV, IUV, and LUV, respectively. Their various sizes are shown in Table 10.4.1 below, in comparison to other large biological structures.

50 nm	200 nm	2000 nm	others
SUV nuclear pore Cell membrane: 7 Ribosome: 20	LUV lysosome HIV: 100	E. Coli Mitochondria: 3000	Red blood cell: 7,000 animal cell: 10,000 – 30,000 Human egg: 120,000

Table 10.4.1: Sizes of liposomes/vesicles compared to other biological structures

The chemical composition of liposomes can be widely varied. Most contain neutral phospholipids like phosphatidylcholine (PC), phosphatidyl ethanolamine (PE), or sphingomyelin (SM), supplemented, if desired, with negatively charged phospholipids, like phosphatidyl serine (PS) and phosphatidyl glycerol (PG). In addition, single-chain amphiphiles like cholesterol (C) and detergents can be incorporated into the bilayer membrane, which modulates the fluidity and transition temperature (Tm) of the bilayer. If present in too great a concentration, single-chain amphiphiles like detergents, which form micelles, can disrupt the membrane so completely that the double-chain amphiphiles become incorporated into detergent micelles, now called mixed micelles, in a process that effectively destroys the membrane bilayer.

Given the large degree of unsaturation at C2, what do you expect the transition temperature of a liposome composed only of egg yolk PC to be? (Vesicles made using more saturated PC from mammalian sources have Tm of around 40^{o}C.) This high degree of unsaturation makes egg yolk PC very susceptible to oxidation, which could alter the properties of the liposome dramatically. Synthetic PC made with saturated fatty acids could alleviate that problem.

The properties of liposomes (charge density, membrane fluidity, and permeability) are determined by the lipid composition and size of the vesicle. The desired properties will be, in turn, determined by the use of the particular liposome. The vesicles offer wonderful, simple models to study the biochemistry and biophysics of natural membranes. Membrane proteins can be incorporated into the liposome bilayer using the exact method you will be using. But apart from these purposes, liposomes can be used to encapsulate water-soluble molecules such as nucleic acids, proteins, and toxic drugs. These liposomes can be targeted to specific cells if antibodies or other molecules which will bind specifically to the target cell can be incorporated into the bilayer of the

vesicle. Intraliposomal material may then be transferred into the cell either by fusion of the vesicle with the cell or by endocytosis of the vesicle.

Since phospholipids will spontaneously form some type of bilayer structure when placed in water, most efforts in liposome production involve producing vesicles with the desired size, lamellar structure, and physical characteristics, which as previously stated are controlled both by liposome size and chemical composition. Also, ways must be developed to entrap the desired molecule inside the vesicle in the most cost-effective manner, and with minimal leaking of contents. All methods of production involve drying of organic solvent-solubilized lipids, dispersion of the lipids in the appropriate aqueous solution, and formation of monolamellar (one bilayer) liposomes or vesicles. Finally, the vesicles are characterized (chemical composition, T_m, permeability, size, etc.)

Drying of lipids

Purified lipids of the desired composition (often egg PC:cholesterol:PS in molar ratios of 0.9:1.0:0.1) are dissolved in a purified, water-free organic solvent mixture (often chloroform/methanol, 2:1 v/v) and dried down in a round bottom flask on a rotary evaporator under reduced pressure (using a water aspirator) and slightly elevated temperature (20-40^{o}C). The rapid rotation of the flask will ensure that the lipid is dispersed over a large surface area, and will increase the rate of evaporation. To remove the last traces of solvent, the dried flask is usually placed under a high vacuum overnight. If a small volume (< 1 ml) of lipid solution is used, the solvent can be evaporated under a stream of nitrogen. To avoid entrapment of residual chloroform in the lipid film, the film is dissolved in t-butyl-methyl ether or diethylether, and dried several times. Alternatively, the residual solvent can be removed under a high vacuum.

Dispersion of the lipids

There are three main methods of dispersing lipids into an aqueous solution to form liposomes.

a. mechanical dispersion - in this method, lipid dried onto the inside glass surface of a container is hydrated with an aqueous solution, which peals off the lipid to form multilamellar - MLV - (multiple bilayers separated by water) vesicles. Only a small part of the aqueous solution is encapsulated inside the liposome, so this is not the method of choice for the encapsulation of expensive or rather insoluble solutes. Depending on the degree of agitation and the nature of the lipid used, different-sized liposomes can be prepared.

b. organic solvent dispersion - In these methods, the lipids, which are dissolved in organic solvents, are injected through a fine needle, at a slow rate, into an aqueous solution in which the organic solvent may be miscible (such as ethanol) or immiscible (such as ether). In each case, the lipids orient at the interface between the organic solvent and aqueous solution, to form bilayer structures. Injection of ethanol-dissolved lipids provides a simple way to produce SUV, but because liposome formation can not occur at an ethanol concentration greater than 7.5%, only a fraction of the total aqueous phase can be entrapped in the vesicle; hence this technique is not cost-effective for entrapment of an expensive solute. Alternatively, the lipid can be dissolved in ether and slowly injected into an aqueous solution which is warmed so that the ether evaporates at the rate at which it is injected. Since the ether is volatilized, large amounts of lipid can be introduced and the encapsulation efficiency of the aqueous solution is high.

c. detergent dispersion and solubilization - In this method, lipids are solubilized in an aqueous solution through the addition of detergents. The detergents are removed slowly from the solution, resulting in the spontaneous formation of liposomes. Detergents are single-chain amphiphiles that spontaneously form micelles in an aqueous solution when the concentration of free lipid rises to a minimum critical value, the critical micelle concentration (CMC); at this concentration, self-association of detergent results in the formation of a stable aggregate, the micelle. This is illustrated in Figure 10.4.2

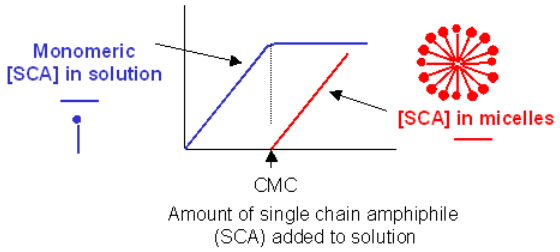


Figure 10.4.2: [Single chain amphiphile] in solution vs [Single chain amphiphile] added - the CMC

Table 10.4.2below shows the properties and CMC of various detergents.

Table 10.4.2: Properties and CMC of various detergents (data from Avanti Polar Lipids)

Name	mM	mg/ml	MW
n-hepty glucopyranoside	70	19.5	278
n-octyl glucopyranoside	23.2	6.8	292
n-nonyl glucopyranoside	6.5	2.0	306
n-decyl maltoside	2.19	1.1	499
n-dodecyl maltotrioside	0.2	0.16	825
Triton X-100 (a)	0.24	0.15	625
Nonidet P-40 (b)	0.29	0.02	603
Tween 20 (c)	0.033	0.04	1364
Brij 98 (d)	0.025	0.04	1527
sodium deoxycholate	2-6	1.7	415
sodium taurocholate	10-15	6.7	538
sodium cholate	14	6.0	431
sodium dodecyl sulfate	8.3	2.4	289

10.4.3: Making liposomes by dialysis

Lipids are deposited in a small container. An aqueous solution is then added, containing water-soluble molecules for encapsulation. Detergent is then added at a concentration in excess of the lipid concentration and greater than its CMC. The lipid molecules are then "emulsified" in the detergent micelle. The solubilized mixture is then placed in a semi-permeable dialysis bag, which is placed in a large volume of an aqueous solution. The free detergent in solution is in equilibrium with the detergent in the micelle. The bag contains microscopic holes large enough for the monomeric detergent molecule to pass through, but small enough so that the large micelle can not. The lipid, during this process, is embedded in the micelle forming a detergent-lipid mixed micelle. As dialysis continues, the monomeric detergent partitions throughout both the volume in the bag and the volume surrounding the bag, while the mixed micelle remains in the bag. If the aqueous solution surrounding the bag is changed several times with fresh solution, the equilibrium in the bag is shifted to the monomeric form. Alternatively, detergent-adsorbing beads (such as Bio Bead SM-2 by Bio-Rad) can be placed in the aqueous solution surrounding the bag to speed up the process of detergent re-equilibration. Eventually, all the detergent is in this form, and during the process, which occurs slowly, the lipid in the mixed micelle self-associates to form a liposome. A detergent of low monomer molecular weight and a high CMC is most desirable for this method of liposome production. Another method of removing the free detergent is through gel filtration chromatography. In this technique, molecules of disparate molecular weights can be separated from each other. An explanation follows this discussion. This method of formation of unilamellar liposomes is the method of choice if membrane proteins are to be inserted into the liposome bilayer to target the liposome. It is not the best method, however, for quantitative encapsulation of expensive soluble molecules.

10.4.4: Making Liposomes by Extrusion

These multilamellar liposomes can be further processed to form unilamellar liposomes by several techniques. These include probe or bath sonication of the MLVs, extrusion at high pressure of the MLVs through membrane filters of defined pore size, or pH-induced vesiculation in which a transient change in pH destabilizes the MLVs in favor of unilamellar liposomes. Another technique involves the fusion of SUVs by repeated freezing and thawing or by fusion of SUVs containing acidic phospholipids (such as PS) through Ca^{2+}-mediated aggregation.

Figure 10.4.3 below shows the structure of vesicles as they undergo multiple freeze/thaw cycles.

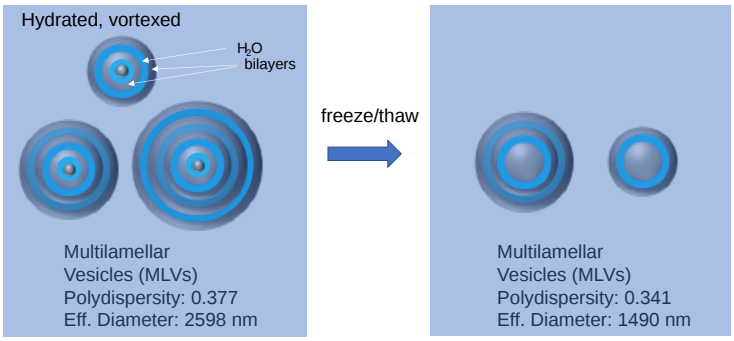


Figure 10.4.3: Structure of multilamellar vesicles as they undergo multiple freeze/thaw cycles

Figure 10.4.4shows the final step in making large unilamellar vesicles by extrusion of freeze/thaw intermediates.

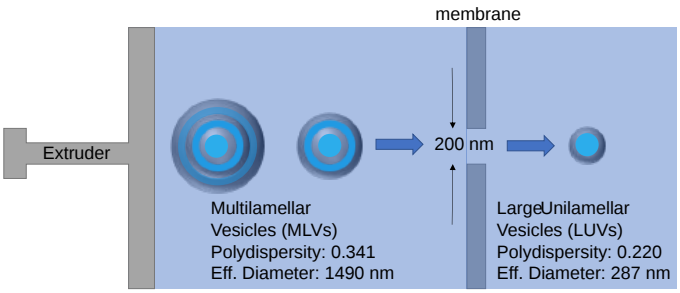


Figure 10.4.4: Conversion of multilamellar vesicles to unilamellar vesicles by extrusion

Once the liposomes are formed, they must be separated from free monomeric lipids, detergents, and unencapsulated solutes. This can be done again by dialysis, or more readily by size exclusion chromatography. Macromolecules of different sizes can be separated on a column in which the stationary phase is a polymerized agarose or acrylamide bead, which contains pores of various sizes. A small molecule (such as monomer detergent, free lipid, or small aqueous solute) in the mobile phase (aqueous buffered solution) may enter the pores in the bead, while a larger macromolecule or aggregate (such as a large protein, a micelle, or a liposome) may not, due to size restriction. The result is that a larger fraction of the overall volume of the column is available to the smaller molecules, which thus spend a longer time on the column and are eluted by the mobile solvent after the larger species. Liposomes can be characterized both chemically, to determine the average lipid and protein (if they were incorporated) composition of the bilayer, and physically, to determine the size, permeability, lamellarity, and amount of encapsulated material. Size is usually determined by electron microscopy or indirectly by light scattering from these large species.

10.4.5: Chemical Analysis of Lipids

There are so many lipids and their derivatives that are so subtlety different that the most sensitive ones are needed for separation and analysis. The field of lipidomics focuses on the analysis of the structure and function of all lipids in the cell. The most useful techniques are analysis (and separation) by gas chromatography (GC) followed by mass spectrometry (MS). NMR spectroscopy is also important. For GC analysis, the lipid must be made volatile, which limits its use in some circumstances. MS analysis is the most sensitive. MS requires ion formation, and techniques like electrospray ionization and matrix-assisted laser desorption/ionization (MALDI) are used.

10.4.5.1: Gas chromatography

(This section is adapted from https://www.intechopen.com/chapters/64008. Creative Commons Attribution 3.0 License). In GC, gas is the mobile phase that carries lipids through the stationary phase of the column. Gas chromatography (GC) is used to separate organic compounds from a mixture in the gas form. For this purpose, the GC uses interactions among the sample components and the stationary/mobile phases. After lipid extraction with chloroform and/or methanol, the samples (lipid mixture) are usually liquids and must be exposed to a high temperature at the gas chromatograph entrance (injector). Vaporized, the samples are carried by a gas, which is usually a nonheavy and inert gas (i.e., hydrogen, helium), through a long capillary column containing a high or low polarity material (stationary phase)

The gaseous compounds generated from the vaporized sample interact with the stationary phase which allows each compound to elute/separate at a different time (retention time). Because GC considers both the chemical and physical properties of the vaporized compounds, those with more chemical affinity to the stationary phase will take a longer time to be removed from the column and the temperature will influence the overall process. This explains why the column stays in an oven, which is programmed to work at different temperature ranges (i.e., temperature programming) in which the compounds are carried out by the gas according to their boiling point until they get to an electronic detector.

At the end of GC analysis, the electronic detector generates a chromatogram based on retention time by intensity. This allows a qualitative identification of the lipid compounds by comparing their retention times with certified standards using the flame ionization detector (FID) or by deduction of spectra information using a mass spectrometer as a detector. Lipid quantification can also be performed using analytical procedures of external or internal certified standards in GC analysis.

The main points to be considered when assessing FAs by GC analysis are the carrier gas flow rate, column length, and temperature because these can influence the order or retention time of the lipid compounds and then must be precisely standardized. The column length of the stationary phase influences the resolution of the analytes, once the number of theoretical plates (hypothetical zone in which two phases establish an equilibrium with each other) is respectively high in the longer column. As fat and oils have high boiling points not supported by the stationary phase, a previous derivatization reaction step is required after lipid extraction from the biological sample, in which triacylglycerol and free fatty acids are transformed into their respective free fatty esters with lower boiling points (transesterification/esterification reaction). Several methods are available for FAs derivatization.

Particularly for cholesterol analysis, the sample preparation must consider a derivatization reaction. This allows to block protic sites of steroids obtained after an unsaponifiable lipid extraction had been performed, and also, diminishes dipole-dipole interactions, increases the volatility of the compounds, and to generate products with reduced polarity. Cholesterol derivatization is usually achieved by using trimethylsilyl compounds as reagents (silylation reaction). A common method for this purpose uses *N*,*O*-bis(trimethylsilyl-trifluoroacetamide/trimethylchlorosilane).

10.4.5.2: Mass spectrometry

Mass spectrometry (MS) has become the method of choice for lipid analysis, offering label-free detection at high sensitivity and structural characterization capability. However, large-scale lipid analysis with a comprehensive capability of revealing all levels of structure information still represents a significant analytical challenge for lipidomics. General protocols, for instance, have five levels in terms of structure information, including lipid class, fatty acyl identities, fatty acyl *sn*-positions, and C=C location/geometry (viz cis/trans) in the fatty acyl. Successful attempts for determining C=C locations in fatty acyls or their *sn*-positions have already been reported for MS analysis, enabling the characterization of detailed structure moieties and identification of lipid structure isomers. An extremely useful feature offered by lipid isomer analysis is the relative quantitation achieved at high precisions without requiring the use of lipid standards, which are not readily available. Remarkably, our recent study demonstrated a close correlation between the lipid C=C location isomer compositions and Type II diabetes, which owes to tighter regulation on lipid desaturation, allowing efficient elimination of interferences due to variations among samples"

Various methods have been explored for differentiating the lipid C=C location and *sn*-position isomers. Ozone-induced dissociation (OzID) and ultraviolet photodissociation (UVPD) have been used to determine both *sn*-positions and C=C locations in GPs. By coupling the Paternò-Bŭchi (PB) photochemical reaction with tandem MS (MS/MS), qualitative and quantitative analysis of lipids with C=C specificity from complex biological samples can be accomplished. PB reaction converts the C=C to an oxetane which can be preferentially fragmented by low-energy collision-induced dissociation (CID).

An ideal analytical tool for lipidomics to survey a wide range of lipids in discovery work should not only provide detailed information at multiple lipid structure isomer levels (e.g., C=C location/geometry and *sn*-position), but also be feasible for large-scale quantitative analysis. UVPD is capable of assigning C=C locations and *sn*-positions of fatty acyls, while OzID may be the only one that has been well demonstrated for assigning C=C locations in *sn*-specific fatty acyls. One problem of OzID is the long reaction time required for the ion trap implementation, however recent work has demonstrated the ability to perform OzID on LC-compatible time scales in the high-pressure regions of the MS system. For the PB reaction method, both shotgun analysis and HPLC-PB-MS/MS workflow have been developed for identifying a large number of C=C location isomers. Figure 10.4.5 describes methods for the determination of C=C location and sn-isomers in lipids.

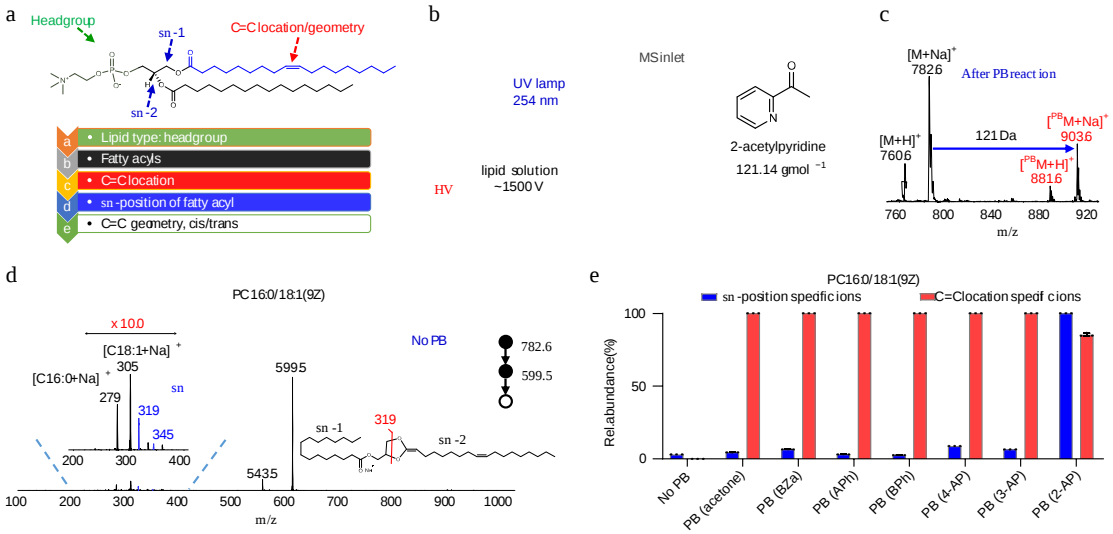


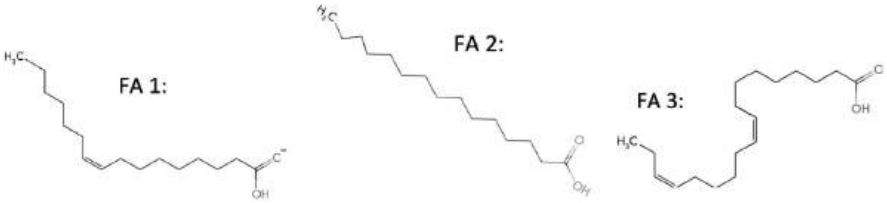
Figure 10.4.5: Methods for the determination of C=C location and sn-isomers in lipids. **a** The chemical structure of PC 18:1(9Z)/16:0. **b** Schematic of the experimental setup for online derivatization of unsaturated GPs by coupling 254 nm irradiation with nanoESI-MS. **c** MS spectrum of PC 16:0/18:1(9Z) after 30 s reaction. **d** MS^3 spectrum of PC 16:0/18:1(9Z) without PB derivatization. Sodiated lipid precursors were first fragmented to generate product ions after headgroup loss (−183 Da). Product ion at *m/z* 599.5 was further fragmented to release *sn*-1-specific diagnostic ions at *m/z* 319 (C16:0) and 345 (C18:1). **e** Comparison the relative abundance of *sn*-position and C=C location specific ions of PC 16:0/18:1(9Z) without and with PB derivatization using different PB reagents (Bza: benzaldehyde, APh: acetophenone, BPh: benzophenone, AP: acetylpyridine). The error bar represents the standard deviation, $n = 3$. Cao, W., Cheng, S., Yang, J. *et al.* Large-scale lipid analysis with C=C location and *sn*-position isomer resolving power. *Nat Commun* **11,** 375 (2020). https://doi.org/10.1038/s41467-019-14180-4. DOI: https://doi.org/10.1038/s41467-019-14180-4. Creative Commons Attribution 4.0 International License: http://creativecommons.org/licenses/by/4.0/.

10.5: Problems

10.1 Questions

1) Explain why some lipid classes are membrane-forming and some are not. Give examples that support your logic.

2) a. Using the images of fatty acids below, write the symbolic name (N:N) for each.



For the following use FA 1, 2, 3, or 4, some may have more than one answer.

b. Which fatty acid melts at the highest temperature?

c. Which fatty acid is a w-3 fatty acid and is essential for human growth and development?

d. Which fatty acid, if found in a membrane bilayer, would increase membrane fluidity?

e. Which fatty acid is the most unlikely to be found in nature, and why?

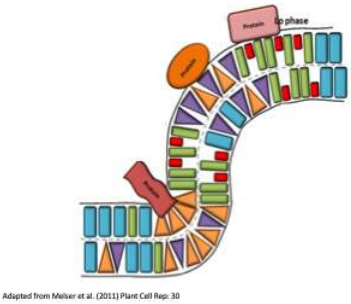
3) For the following questions, use the interactive micelle from Section 10.1:

a. When looking at the micelle in “Ball and Stick” form, do you think the core (center of the micelle) could be accessible to water molecules? Explain.

b. Examine the micelle under Style -> Surface Type -> Molecular Surface. Did your answer from part a change? Do there appear to be any solvent-accessible regions that water could use to enter the core?

c. Finally, examine the micelle under Style -> Surface Type -> Solvent Accessible. Does this view change your answer to a.? Can you think of any medical or research applications in which a micelle could be useful given what you have found about its solvent accessibility?

4)Using the figure below, hypothesize what type of lipid could be represented by the geometrical shapes in the membrane. Why are the different shapes necessary for membrane formation? Could a membrane be formed with only one shaped lipid? If so, what, and explain why.



5) A cell membrane has the ability to remodel in response to stress in order to maintain membrane integrity. In the situations below, how could the membrane be remodeled to promote an intact structure? (I.e. what types of lipids could be added or removed to maintain a functional membrane?

a. Increase in temperature

b. Decrease in temperature

6)

For each lipid below, name the type of lipid (membrane lipid, triacylglycerol, storage lipid, sphingolipid, wax, sterol, membrane glycerolipid, none of these), if it could be found in a membrane, and if it is a fatty acid or isoprenoid derived.

CHAPTER OVERVIEW

11: Biological Membranes and Transport

11.1: Membrane and Membrane Proteins
11.2: Diffusion Across a Membrane - Passive and Facilitated Diffusion
11.3: Diffusion Across a Membrane - Channels
11.4: Diffusion Across a Membrane - Pores
11.5: Active Transport
11.4: Diffusion Across a Membrane - Pores

11.1: Membrane and Membrane Proteins

11.1.1: Introduction

One easily understandable function of membrane bilayers is to separate the inside and outside of the cell or intracellular organelles. Yet as we mentioned before, such barriers can not be so rigid and impenetrable that they prevent the movement of materials across the membrane. Also, all cells must sense and respond to their environment through a process called signal transduction. We have already discussed lipid molecules involved in signaling. Now let's turn our attention to proteins that associate with the membrane and confer added functionalities to it. Figure 11.1.1 reviews some of the features of membranes we've discussed before and shows a simple bilayer (top) to the complicated membrane/cell wall of bacteria.

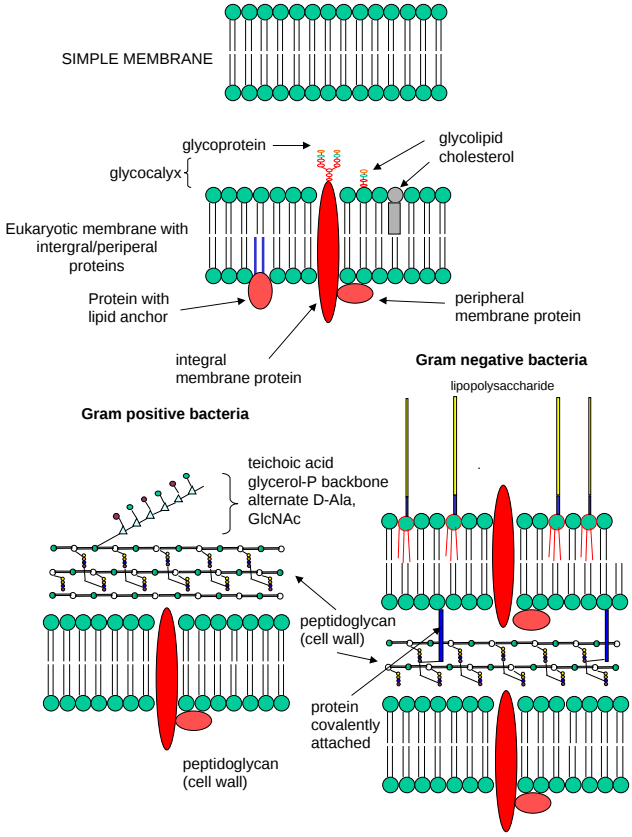


Figure 11.1.1: Membrane from simple to complex

11.1.2: Types of Membrane Proteins

Although we presented this image earlier, Figure 11.1.2 reviews the details that should now be clearer to you. In this section, we will explore membrane proteins in more detail.

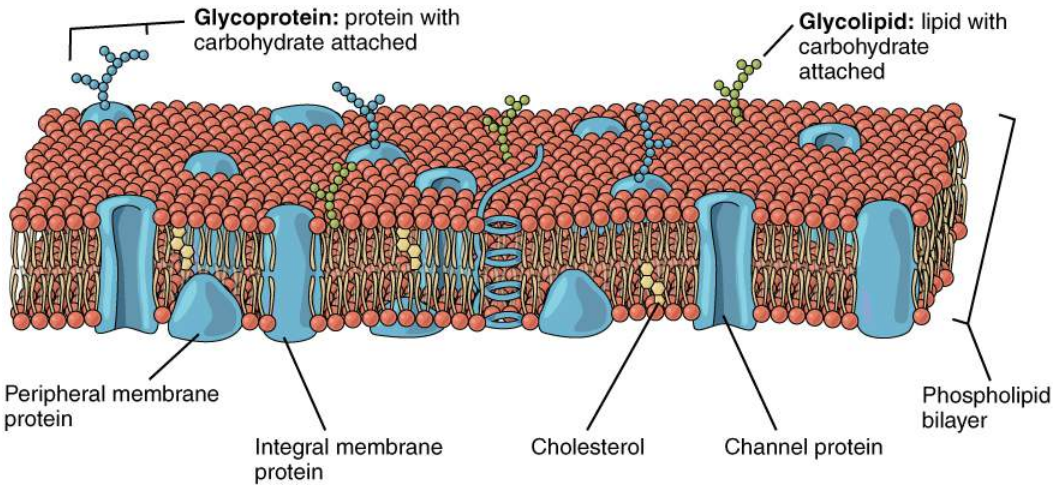


Figure 11.1.2: A cell membrane with peripheral and integral membrane proteins. https://commons.wikimedia.org/wiki/F...Components.jpg. This file is licensed under the Creative Commons Attribution 4.0 International license.

Proteins can be loosely associated with the membrane (**peripheral or extrinsic**) or can embed deeply and most typically pass through the membrane and become a transmembrane (also called i**ntegral or intrinsic)** protein. Sometimes they pass through using a single alpha helix, while other times they pass through multiple times (for example seven times in G-protein coupled receptors). They can also be classified based on the number of leaflets of the membrane they cross, as shown in Figure 11.1.3.

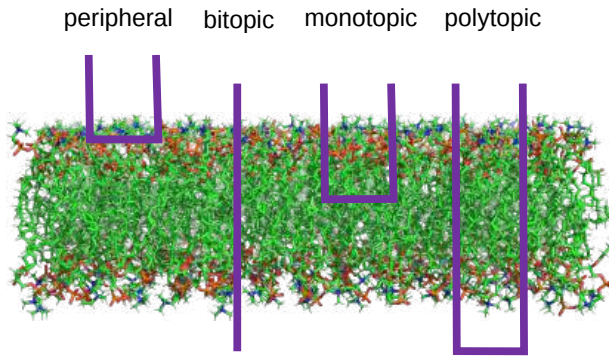


Figure 11.1.3: Peripheral and biotopic, monotopic, and polytopic integral membrane proteins

11.1.3: Peripheral Proteins

The proteins interact with a membrane through protein-lipid head group interactions, but might slightly penetrate the membrane. Those that do would be classified as monotopic peripheral. Peripheral proteins are generally easy to remove from a membrane in vitro by changing solution ions concentration as the interactions are often ion-ion nature. The first model below shows the binding of a matrix Metalloproteinase (MPP) 12 to a lipid bilayer. This protein is involved in inflammation, wound healing, arthritis, cardiovascular disease, and remodeling of neural synapses, suggesting a broad role in recovery from cell and tissue aberrations. MMP-12 is secreted by macrophages so it is considered a water-soluble (aqueous) protein. It travels to viral cells and appears to display activity not in aqueous solution but near membranes, implying activation of the enzyme through binding to the bilayer.

Studies show the catalytic domain of MMP 12 can bind bilayers through both α- and β- secondary structure regions of the protein. Figure 11.1.4 shows an interactive iCn3D model of the protein and its interaction with the membrane through the alpha-helical region. Once bound to the membranes, catalytic activity increases.

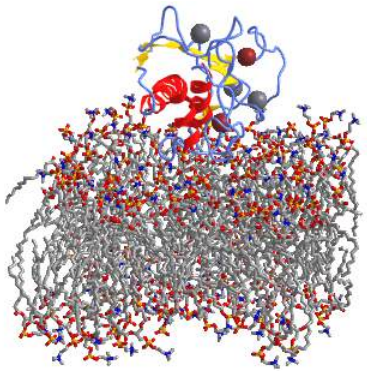


Figure 11.1.4: Catalytic domain of peripheral membrane Protein MMP 12 Interacting with a bilayer. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...wyF6WvRMx9JP6A

Other MMPs localize to membranes in other ways. MMP-7 interacts with heparan sulfate proteoglycans (CD44) and lipid rafts. Others bind transmembrane proteins like integrins.

Other examples of peripheral protein include many precursor forms of protein clotting factors. Clotting is initiated when the serine protease thrombin cleaves fibrinogen to form fibrin, which self-associates to form a fibrin clot, or when thrombin activates receptors in platelets. The soluble precursor of thrombin, prothrombin, a zymogen, is activated on membrane binding through interactions with several proteins assembled on a negatively charged phospholipid (like phosphatidylserine) bilayer in the prothrombinase complex. How does the precursor zymogen interact with the membrane? It requires calcium ions, which bind to a series of **gamma-carboxylated glutamic acid (GLA)** residues on the zymogen. The enzyme that carboxylates the zymogen depends on Vitamin K.

Figure 11.1.5 shows an interactive iCn3D model of bovine prothrombin Fragment 1 (N terminal) bound to a bilayer through its GLA domain (1NL2).

Figure 11.1.5: Bovine prothrombin Fragment 1 (N terminal) bound to a bilayer through its GLA domain (1NL2). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...SG6qByvpAsvwe6

The Gla sidechains are shown in CPK-colored sticks and interact with Ca^{2+} ions (gray spheres). Click on this link to see a zoomed view of just the calcium ions and Gla sidechains: https://structure.ncbi.nlm.nih.gov/i...xydtj8a4WHgdb7

The Gla domain in the absence of calcium ions is disordered. On binding, an ordered linear alignment of bound calcium ions is formed, stabilizing the ordered structure of the Gla domain and allowing interaction with the membrane. Three nonpolar amino acid side chains, Phe 5, Leu 6, and Val 9, are now clustered and exposed, allowing penetration of this hydrophobic patch part-way into the membrane. They are represented in cyan spacefill just underneath the surface of the red dots in the model above (the red

dots are dummy atoms that represent the outer bilayer leaflet). Given this penetration, this protein domain would then be considered monotopic.

What is not shown in the model is the role of negatively charged phosphatidyl serine. Studies have shown that the head group of serine in lysophosphatidylserine (which has only one acyl group) provides additional ion-ion interactions with the Ca^{2+} ions that also bind Gla residues 17 and 21. Arg 10 and Arg 16 also interact with the phosphatidyl serine head group. Phosphatidylcholine could also spatially fit into the active site but electrostatic interactions would prevent it. Why?

11.1.4: Lipid-Anchored Proteins

We have studied lipids, proteins, and carbohydrates. Although phospholipids can spontaneously form bilayers, the actual structure of biological membranes is made much more complicated through the addition of protein and carbohydrate substituents to the membrane. Soluble proteins can be made to insert into bilayers by the addition of nonpolar attachments. Localization to a membrane changes the functional expression of the protein. Several examples of such attachments are described below.

11.1.4.1: Fatty acid linkers

Two common covalent modifications of proteins are N-myristoylation (attached myristic acid - 14:0 - through an amide link) and S-palmitoylation (attached palmitic acid - 16:0 - through a thioester link with a Cys).

Myristoylation is usually a cotranslational modification in eukaryotic and viral proteins that occurs after cleavage of the N-terminal methionine. Figure 11.1.6 shows an image of the serine/threonine phosphatase 2C (1A6Q) with its N-terminal glycine myristoylated. It should be obvious how this post-translationally modified protein interacts with a membrane.

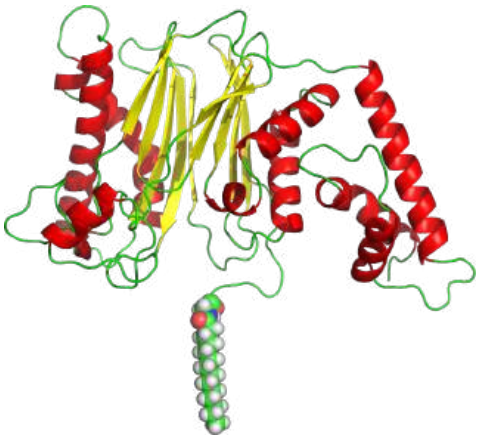


Figure 11.1.6: The serine/threonine phosphatase 2C (1A6Q) with its N-terminal glycine myrisoylated

This modification is a key part of initiating immune system signal transduction pathways. The modification is catalyzed by N-myristoyltransferase (NMT) using myristoyl-coenzyme as the fatty acid acyl donor. This activates the function of the protein in part by reducing the dimensionality of substrate diffusion to the protein to the 2D surface of the membrane instead of a 3D search in the cytoplasm. NMT acylates protein at this consensus sequence: $G^1X^2X^3X^4S/T^5X^6R^7R^8$.

Likewise, many signaling proteins are palmitoylated, leading to protein recruitment to membranes. Small G proteins like Ras, Rho, and the alpha subunit of heterotrimeric G proteins are often palmitoylated. This modification is also be found in transmembrane proteins in which localization is not an issue (see example xx below). In such circumstances, the modification might however help in targeting the proteins to rafts within the membrane. Palmitic acid is saturated and the addition of it to a protein might target it to more ordered regions of the membrane with cholesterol and sphingolipids within rafts.

11.1.4.2: Isoprenoids linkers:

The isoprenoids farnesyl (15C) or geranylgeranyl (20C)are added to a CAAX carboxy-terminal sequence in a target protein like RAS, where C is Cys, A is aliphatic, and X is any amino acid, which helps target proteins to the membrane. The enzymes used for these modifications are farnesyltransferase (FTase) and protein geranlygeranyltransferase
I (GGTase I), respectively. For this and the other modifications, it has the potential to do more than target proteins to the membrane. The modification can also modulate protein-ligand interactions and protein stability. Ras, a key signaling protein, is a target of prenylation.

Ras and other small G proteins are involved in a large percentage of human cancers. As the G protein Ras has somewhat of a billiard ball surface with obvious sites to target drugs that would affect its aberrant function in cancers, efforts have been made to target the prenyltransferases necessary to target it to the membrane.

In humans, there are 3 different genes in the Ras family, H-Ras and N-Ras, whose gene products localize to both plasma and Golgi membranes, and K-Ras, which localizes predominantly to the plasma membrane. These and other G proteins bind GTP and possess GTPase activity. The GTP-bound form is active, while the GDP form is inactive. Point mutations that attenuate or prevent GTP cleavage leave the protein continually activated which contributes to oncogenesis.

KRas has two predominant isoforms, 4A, the canonical form (also called 2A), and 4B (also called 2B) that arise from alternative splicing of the primary RNA transcript. The C-terminal protein sequences of isoform 4A and 4B differ significantly.

Isoform 4A: QYRLKKISKEEKTPGCVKIKKCIIM

Isoform 4B: KHKEKMSKDGKKKKKKSKTKCVIM

The farnesylation motif site containing the modified Cys are highlighted in yellow above. The same cysteine is also often carboxymethylated. The Cys six residues from the farnesylated Cys in isoform 4A are also often palmitoylated

Figure 11.1.7 shows isoform KRas 4B bound to a membrane bilayer through its farnesylated tail. (PDB file provided by Alemayehu (Alex) Gorfe. Viney Nair and Andrew McCammon). The tail is essential for its function at the plasma membrane where KRAS-mediated signaling events occur. Phosphodiesterase-δ (PDEδ) binds to KRAS4b and plays an important role in targeting it to cellular membranes. Note that the farnesyl attachments only penetrate part of the upper leaflet.

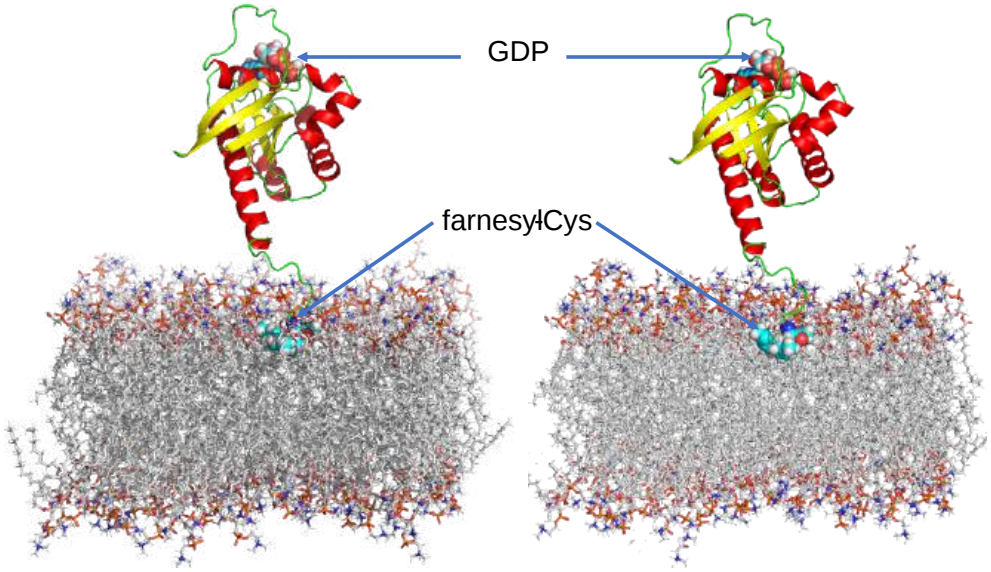


Figure 11.1.7: Isoform KRas 4B bound to a membrane bilayer through its farnesylated tail

11.1.4.3: Glycosyl-phosphatidylinositol linkers

Normally soluble cytosolic proteins can become attached to membranes through the addition of a glycosyl phophatidylinositol (GPI). The attachment usually contains a conserved tetrasaccharide core of three mannoses (Man) and one unacetylated glucosamine (GlcN) linked to the carboxy terminus of the protein. The GPI can be further modified with extra galactoses and mannoses, as well as additions to the PI group, which secures the protein in the membrane. Figure 11.1.8 shows the common backbone for GPI anchors. Note the additions of the phosphoethanolamines to the core polysaccharide.

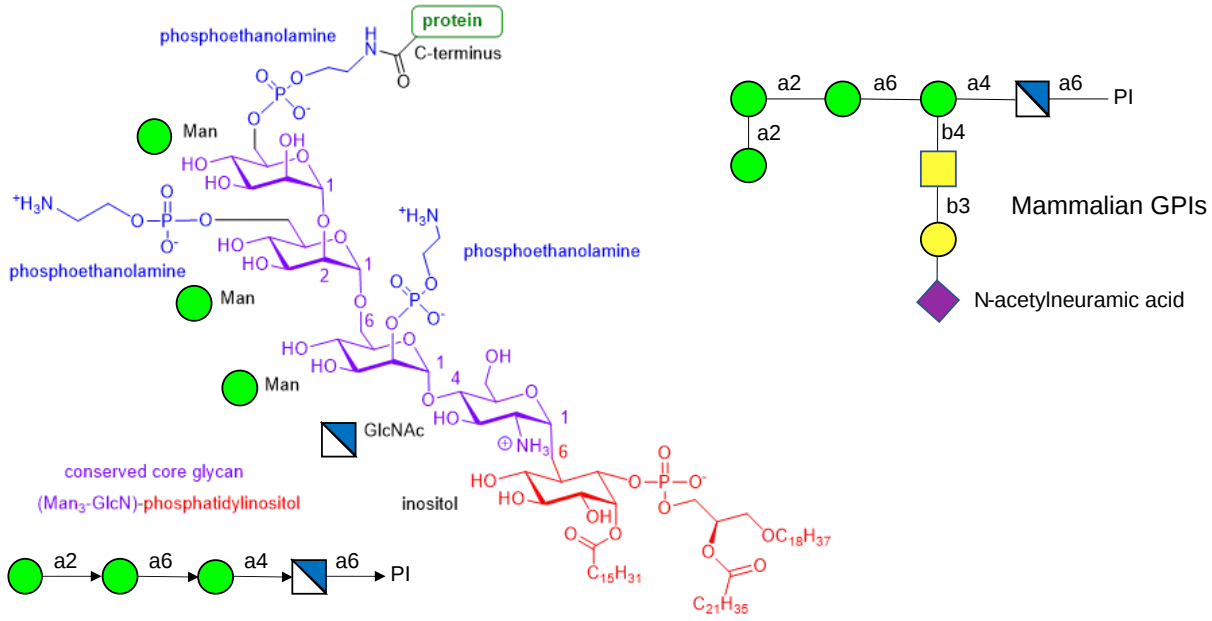


Figure 11.1.8: Common backbone of GPI anchors

GPIs are found in eukaryotic cells and link many surface antigens, adhesion molecules, and hydrolases to the membrane. GPIs from Plasmodium falciparum, the malarial parasite which kills about two million people each year, appear to act as a toxin and are the most common CHO modification of the parasite protein. Mice immunized against the GPI sequence, NH_2-CH_2-CH_2-PO_4-Man (α1-2) 6Man (α1-2) Man (α1-6) Man (α1-4) GlcNH2 (α1-6) myo-inositol-1,2-cyclic-phosphate, were substantially protected from malarial symptoms and death after they were exposed to the actual parasite.

Figure 11.1.9 shows a cross-section of a membrane (with cholesterol, PE, SM) containing the glycosylated form of the human complement regulatory protein CD59 protein (1cdr) with a GPI anchor attached at its C-terminus. Note that the middle part of the anchor (glycan) holds the actual protein well above the top of the lipid bilayer. The soluble protein is also glycosylated. The protein binds to complement proteins C8 and/or C9, which are effector immune proteins that assemble on the surface of a cell undergoing lysis.

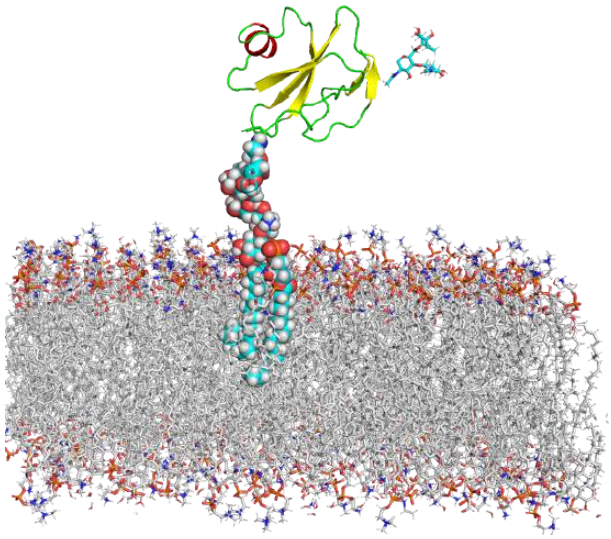


Figure 11.1.9: Cross section of a membrane containing the glycosylated form of the human complement regulatory protein CD59 protein (1cdr) with a GPI anchor attached at its C-terminus

The GPI anchor is shown in spacefill. Note that it only extends halfway into the bilayer, as you would expect from the size of the fatty acids attached to the phosphatidyl inositol. The glycan part of the GPI is shown in spacefill between the lipid and its protein attachment site. The protein is also glycosylated in the extracellular domain.

Something new!

A new (5/21) and totally unexpected type of glycosylated molecule has been found at the outer leaflet of mammalian cells - a glycosylated RNA, as shown in Figure 11.1.10 This adds RNA to the lipids and proteins as a target for glycosylation. These surface glycoRNAs interact with antibodies again ds-RNA and the Siglec lectin family. They are found in cells in vivo and cultured cells in vitro.

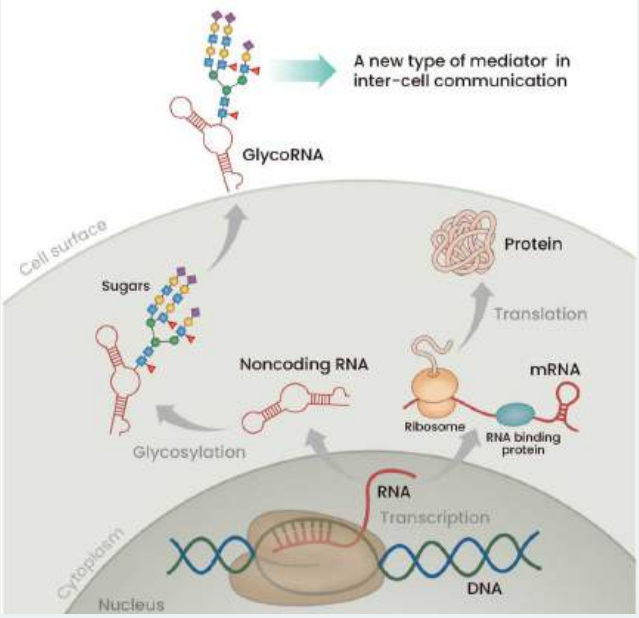


Figure 11.1.10 A glycoRNA - a small noncoding RNAs with sialylated glycans. Park. https://doi.org/10.14348/molcells.2021.0178
www.molcells.org. Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/.

Flynn et al., Small RNAs are modified with N-glycans and displayed on the surface of living cells, Cell (2021), https://doi.org/10.1016/j.cell.2021.04.023

11.1.5: Transmembrane (Integral) Proteins

These proteins pass through the membrane either one in one pass, usually with a single alpha-helix or many membrane-spanning helices. For example, G protein-coupled receptors, often called serpentine receptors, cross the membrane seven times. There are three different types based on the number of types the protein crosses the membrane and the type of secondary structure used in crossing: biotopic (single pass), alpha-helical polytopic, and beta-barrel. These proteins are found in all types of membranes and have many types of functions, from receptors, receptor ligands, structural, adhesion, transport, gene regulation, and transport.

11.1.5.1: Transmembrane Biotopic - Single Pass Proteins

The are 4 types of single-pass transmembrane proteins:

- Type I: N-terminal outside of the cell (extracellular) and the precursor signal sequence on the N-terminus which is a localization sequence is removed
- Type II: N-terminal intracellular and with the transmembrane domain close to the N-terminus
- Type III: N-terminus extracellular and no signal sequence in precursor protein
- Type IV: N-terminus intracellular and the transmembrane domain close to the C-terminus

The transmembrane domain of single-pass integral membrane proteins consists of a single alpha-helices with nonpolar side chains extending outward from the helical axis where they interact with the nonpolar lipid parts of the membrane. These nonpolar sides are more stable in nonpolar environments.

To study such proteins in a less complex environment, membranes are often "dissolved" in nonpolar, single-chain amphiphilic detergents. These single-chain amphiphiles form micelles in the absence of membrane proteins but can form **mixed micelles** in which the nonpolar part of the protein is surrounded in the detergent micelle by the nonpolar acyl chains of the detergent.

Figure 11.1.11(top) below shows just the transmembrane and juxtamembrane (next to the membrane) domains of the single pass Notch protein, which is critical in many signal transduction pathways.

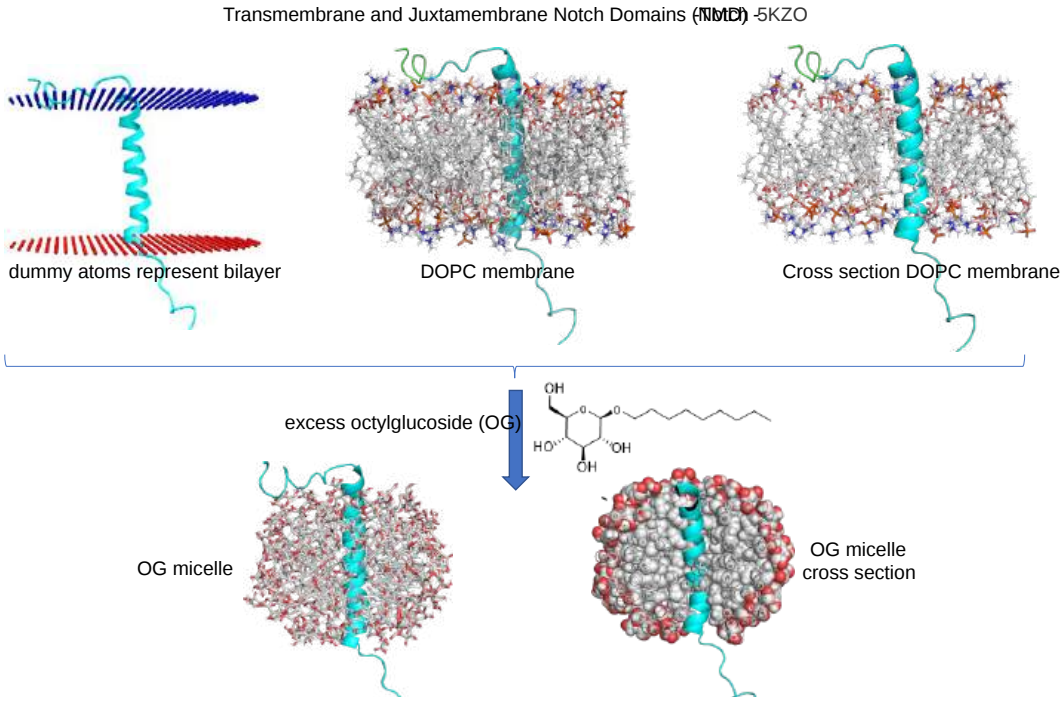


Figure 11.1.11: Transmembrane and juxtamembrane domains of the single pass Notch protein in a bilayer and an octylglucoside micelle

The top images in the figure above show different ways to represent the protein in the bilayer, with the right-hand image showing a cross-section through the membrane to better show how the protein passes through the bilayer.

The bottom images in the figure above show the protein after excess detergent, in this case, octylglucoside, is added to the protein-containing bilayer.

Here are some examples of bitopic single-pass transmembrane proteins

Cadherins

All structures need support and connections. At the macro level, the skeleton supports the mass and organization of organs and tissues in whole organisms. Within an organ, how can cells hold together? How do they adhere to each other? Certainly not through outer leaflet lipid contacts as the outer surface of the leaflet is typically charged. The extracellular matrix does provide some of the glue that holds cells together. At a more detailed level, transmembrane proteins are involved. One class of adhesive proteins is cadherins, a calcium-dependent cell adhesion molecule. There are over 100 human cadherins. They are mostly ditopic, single transmembrane pass proteins. Their cytoplasmic domains interact with proteins like catenin, which then bind to the interior cytoskeletal network composed of actin and other proteins. This provides a way for the intracellular region to regulate the extracellular interactions of the cell.

The extracellular domain is composed of five repeating "cadherin" domains, each around 110 amino acids, that can fold independently. Calcium ions bind at the domain interfaces. A cadherin can interact with other cadherin domains on other cadherins on other cells, leading to cell adhesion. Essentially, the receptor cadherin on one cell binds the ligand cadherin on the other. As metastatic tumor cells lose their adhering feature and leave the site of the primary tumor, you would expect that mutations in cadherins are often involved. They may also be involved in cell sorting during morphogenesis, "regulation of tight and gap junctions, and in the control of intercellular spacing".

Figure 11.1.12 shows a "constructed" image of cadherin-1 (1L3W) interacting with cytoplasmic β-catenin (1I7X) through a modeled transmembrane helix (amino acids QIPAILGILGGILALLILILLLLLFLRR, amino acids 706-731). No full-length structure of cadherin in a membrane is available.

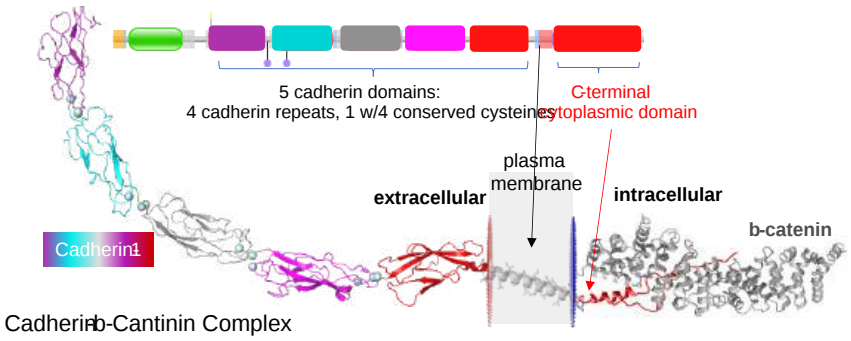


Figure 11.1.12: A constructed" image of cadherin-1 (1L3W) interacting with cytoplasmic β-catenin (1I7X) through a modeled transmembrane helix

Membrane Protein Kinases

Kinases are enzymes that phosphorylate substrates. Hexokinase is a protein enzyme that catalyzes the phosphorylation of a hexose substrate such as glucose. A **protein kinase** is a protein enzyme that phosphorylates a protein substrate. That protein could be another copy of itself or another protein. We will see in Chapter 12 that many protein kinases are involved in cell signaling. Many tyrosine protein kinases are bitopic single-pass integral membrane proteins that become active on binding a ligand. Typically, on binding an extracellular ligand, two monomeric copies of the kinase form a dimer in the membrane, activating a tyrosine kinase cytoplasmic domain, which typically phosphorylates (using ATP as a substrate) the other member of the dimer in an "autophosphorylation" reaction. Sometimes the dimers are held together by disulfide bonds.

Figure 11.1.13shows a "constructed" image of the human dimeric insulin receptor.

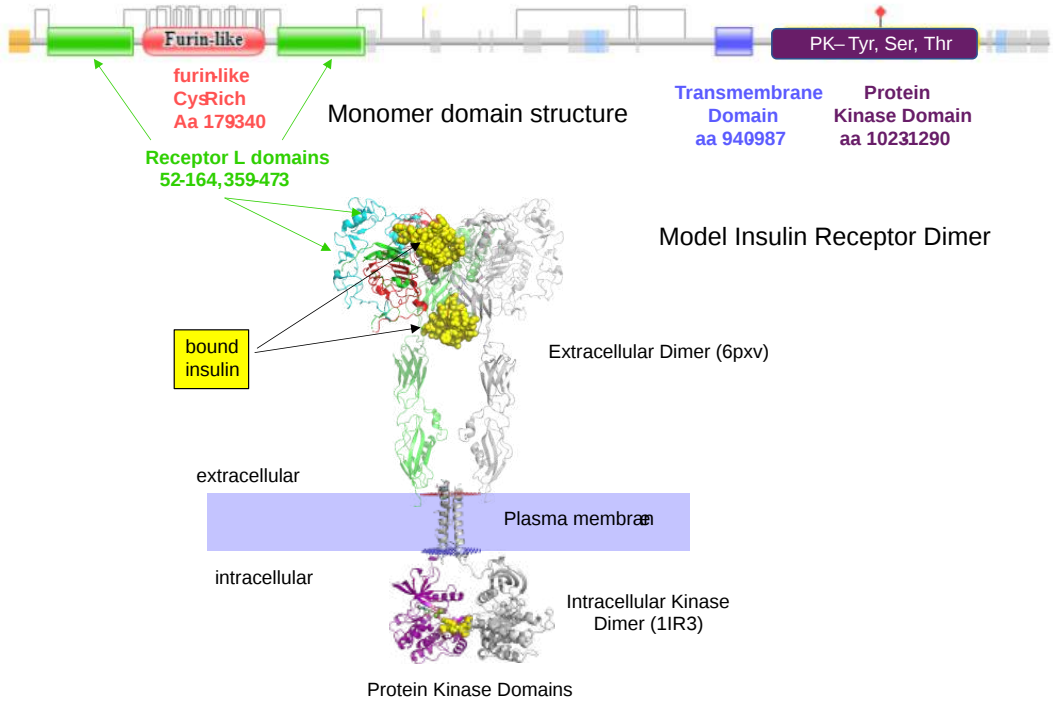


Figure 11.1.13: aA "constructed" image of the human dimeric insulin receptor

One of the monomers is shown in gray. The other monomer is shown in colors corresponding to the domain organization of the protein. Each extracellular dimer (6PXV) has two insulins bound (yellow spacefill). The intracellular domains (1IR3) are activated on insulin binding. No full-length structure of full insulin receptor in a membrane is available.

Almost half of all helical membrane proteins in humans are bitopic, compared to between 20-25% in prokaryotes. Humans have 10-20 fold more bitopic proteins than E.Coli. There appear to be about 196 bitopic proteins in E. coli (located in the inner membrane) and 70 in M. jannaschii (Archea in plasma membrane). In humans, 57% are in the plasma membrane, with the rest distributed between the Golgi, ER, nuclear, mitochondrial and chloroplast membranes. In single-celled yeast, only 8% are in the plasma membrane.

Beta-Dystroglycan

This protein is another example of a bitopic protein with a single alpha-helix membrane domain. Dystroglycan is a dimer of alpha and beta subunits. Alpha-dystroglycan is a peripheral protein that binds beta-dystroglycan, a transmembrane protein. Alpha dystroglycan also binds lassa virus and lymphocytic choriomeningitis virus glycoprotein, as it serves as viral receptors. It also binds the protein dystrophin, a protein missing in Duchenne muscular dystrophy, which affects 1 out of 5000 live male births. As an integral transmembrane protein, beta-dystroglycan connect the extracellular matrix to the cytoskeleton through dystrophin. Alpha- and beta-dystroglycan share the same gene, which codes one long protein which is proteolyzed post-translationally to form the alpha (N-terminal end) and beta subunits (C-terminal end).

Figure 11.1.14shows a schematic outline of dystrophin and the dystrophin-associated glycoprotein complex (DAGC).

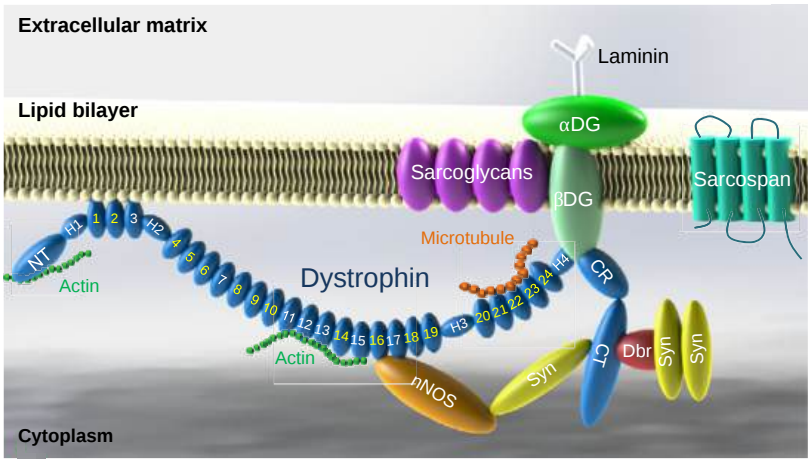


Figure 11.1.14 Schematic outline of dystrophin and the dystrophin-associated glycoprotein complex (DAGC). Dystrophin contains N-terminal (NT), middle rod,
cysteine-rich (CR), and C-terminal (CT) domains. The middle rod domain is composed of 24 spectrin-like repeats (numerical numbers in the cartoon, positively
charged repeats are marked in white color) and four hinges (H1, H2, H3, and H4). Dystrophin has two actin-binding domains located at NT and repeats 11-15,
respectively. Repeats 1-3 interact with the negatively charged lipid bilayer. Repeats 16 and 17 form the neuronal nitric oxide synthase (nNOS)-binding domain.
Dystrophin interacts with microtubules through repeats 20-23. Part of H4 and the CR domain binds to the β-subunit of dystroglycan (βDG). The CT domain of
dystrophin interacts with syntrophin (Syn) and dystrobrevin (Dbr). Dystrophin links components of the cytoskeleton (actin and microtubule) to laminin in the
extracellular matrix. Sarcoglycans and sarcospan do not interact with dystrophin directly but they strengthen the entire DAGC, which consists of dystrophin,
DG, sarcoglycans, sarcospan, Syn, Dbr, and nNOS. .Disease Models & Mechanisms (2015) doi:10.1242/dmm.018424vailable via license: CC BY 3.0

11.1.6: Transmembrane - Alpha-helical polytopic

There are so many intriguing examples of these proteins. We'll illustrate just two.

Rhodopsin-like receptors and pumps

These proteins are involved in cell signaling and are the target of most pharmaceutic drugs. **G protein-coupled receptors (GPCRs)** are incredibly important and we will discuss them extensively in Chapter 12.

GPCRs are cell receptors that span the membrane seven times in a serpentine fashion. They bind ligands (neurotransmitters, hormones, etc) in the extracellular or internal membrane domains (the latter for hydrophobic ligands), and through propagated conformations changes alter the cytoplasmic domain where they functionally interact with a heterotrimeric G protein.

Figure 11.1.15 shows an interactive iCn3D model of the human cannabinoid receptor with bound cholesterol and Δ^9-tetrahydrocannabinol (Δ^9-THC) in spacefill (5xra). The red dummy atoms represent the outer leaflet and the blue the inner.

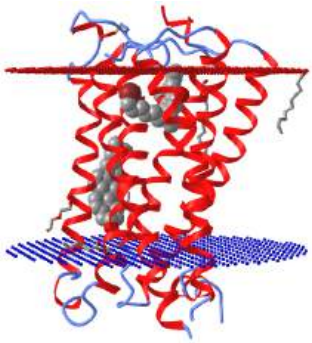


Figure 11.1.15: Human cannabinoid receptor with bound cholesterol and Δ^9-tetrahydrocannabinol (Δ^9-THC). (5xra). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...PaTFrkGFgKeP2A

Δ^9-THC is a partial agonist and tunes the response of the receptor. The active site is conformationally somewhat flexible or plastic. Other ligands bound to it act as antagonists instead of agonists and must do so by eliciting nonactive conformations.

ABC Transporter

Figure 11.1.16shows an interactive iCn3D model of the P-glycoprotein multidrug resistance transporter protein (6nf1).

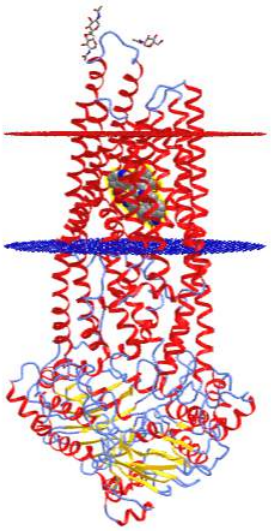


Figure 11.1.16: P-glycoprotein multidrug resistance transporter protein (6nf1). (Copyright; author via source). Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...zzXduCZjhCMaw9

The spacefill ligands represent Zosuquidar, which binds with high affinity to P-glycoprotein and inhibits its activity, making it a cancer agent as it prevents chemotherapeutic drugs that have entered the cell from being pumped out. The protein chain interacting with it on the cytoplasmic face is an antibody fragment used to stabilize the P-glycoprotein so crystals could form.

11.1.7: Transmembrane Beta-barrel transmembrane

We will focus on two of these proteins.

Outer Membrane Factor (OMF) - Gram-negative bacteria

Figure 11.1.17shows an interactive iCn3D model of a beta-barrel transmembrane protein OPRM - Outer Membrane Factor (4y1k) from Pseudomonas aeruginosa that acts as a pore. It also has a palmitoyl fatty acid in thioester linkage to Cys 1 of the protein for extra but unneeded anchorage.

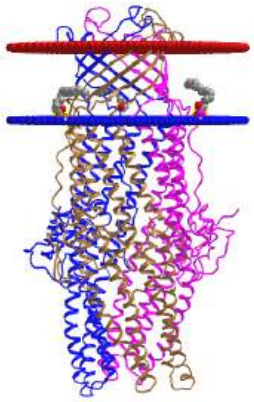


Figure 11.1.17: Beta-barrel transmembrane protein OPRM - Outer Membrane Factor (4y1k) from Pseudomonas aeruginosa (4y1k). (Copyright; author via source). Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...xrAuW1oFtY5FL7

Use this link for another view: https://structure.ncbi.nlm.nih.gov/i...FtY5FL7&t=4Y1K (OPM) in iCn3D

This protein is part of a large complex of proteins that spans both the inner and outer membranes of Gram-negative (examples E. Coli and Pseudomonas aeruginosa) bacterial cell walls. Unfortunately for humans, this protein complex pumps out toxins (to the bacteria) like antibiotics, which makes bacteria resistant to these drugs. The OPRM acts as the outer passageway or duct for the pumped molecules. The Bacterial Outer Membrane Factor (OMF) protein differs in sequence but all form the beta-barrel duct. The E. Coli version of OMF has a triacylated lipid modification of the N-terminus. The N terminal lipid modification might be necessary for the initial attachment of the protein to a membrane before the insertion of the beta-barrel. As such, the enzymes involved in the attachment of the tail could be targets for new antibiotics.

Voltage-dependent anion channel (VDAC) - mouse

This protein regulates the movement of molecules between the cytoplasm and the interior of the mitochondria across the outer mitochondrial membrane. VDAC also serves as a docking site or scaffold for the assembly of molecules into a complex that regulates mitochondrial function. The protein's conformation and hence function are regulated by changes in the transmembrane potential, which we will explore in the next sections. Hence the protein and its function are voltage-dependent. Figure 11.1.18 shows an interactive iCn3D model of mouse VDAC with a beta-barrel formed by 19 beta-strands (3emn).

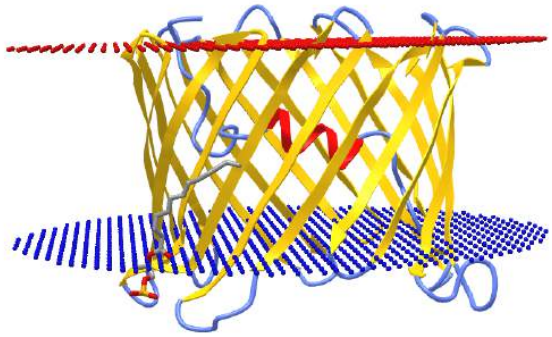


Figure 11.1.18: Mouse voltage-dependent anion channel (VDAC) (3emn). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...WxUqWaGPxgHH79

Note the N-terminal alpha-helix resident inside the channel opening. This helical section moves on changes in membrane potential, gating open and hence regulating the flow of metabolites and ions across the membrane through the pore. At a low transmembrane potential (10 mV), the conductance is high as the channel is in the open state. When the potential increases to 30 mV (either + or -) conductance drops as the protein forms the closed state.

Now that you understand the structure of membrane proteins, let's explore a key type of function of a subset of integral membrane proteins: the movement of molecules/ions across the membrane.

11.2: Diffusion Across a Membrane - Passive and Facilitated Diffusion

Learning Objectives

- Understand the structure and function of biological membranes.
- Describe the different types of diffusion and how they operate across a membrane.
- Explain the concept of concentration gradient and how it drives diffusion.
- Describe the mechanisms of passive and facilitated diffusion, including their differences and similarities.
- Understand the role of carrier proteins and channels in facilitated diffusion.
- Describe the factors that can affect the rate of diffusion across a membrane.
- Understand the importance of membrane transport in maintaining homeostasis in cells and organisms.

11.2.1: Diffusion Across a Membrane

We have studied molecular aggregates (micelles and bilayers) and macromolecular structures (mostly proteins). We also studied binding interactions which are the first steps in the expression of the biological activity of a macromolecule. For some proteins, reversible binding is their sole function (consider the binding of dioxygen to myoglobin and hemoglobin). For many others, it is not. For those, what can happen next?

You already have one possible answer. A bound reactant, which we will call a substrate, can be converted to a product in a chemical step involving the breaking and making of covalent bonds catalyzed by a protein enzyme. However, there is an even simpler process that does not involve covalent bond changes. If a small molecule is bound to a membrane protein, it could be transported in a purely physical step across a membrane. Just as reactions can proceed with and without an enzyme, a solute can move down a concentration gradient across a semi-permeable membrane, driven by diffusion alone in a thermodynamically favorable process, either by itself, in a process called **passive diffusion**, or with the assistance of a membrane protein, in a process called **facilitated diffusion**. Large **pores** made of assemblies of proteins can also be formed that allow the passage of many solutes across the membrane.

There are many occasions when it would be optimal to move a molecule across a membrane from a region of low to high concentration. This process is called **active transport**. It is not thermodynamically favored so it requires an external energy source. This is often the thermodynamically favored cleavage of ATP to ADP and Pi. The uphill transport can also be powered by the downhill diffusion of a "co-transport" molecule from high to low concentration. We will explore all of these processes in this and the remaining chapter sections. First, we should understand the simplest process, "passive diffusion", that requires no protein "help".

Let's step back and think about how difficult it is for a chemical species to cross a lipid bilayer. Chemical intuition would tell us that both size and polarity are important. The bigger the size and the greater the charge, the more difficult it would be to cross the membrane. The permeability coefficient is related to the ease with which solutes traverse the membrane. Figure 11.2.1 shows the permeability coefficients for relevant biological molecules.

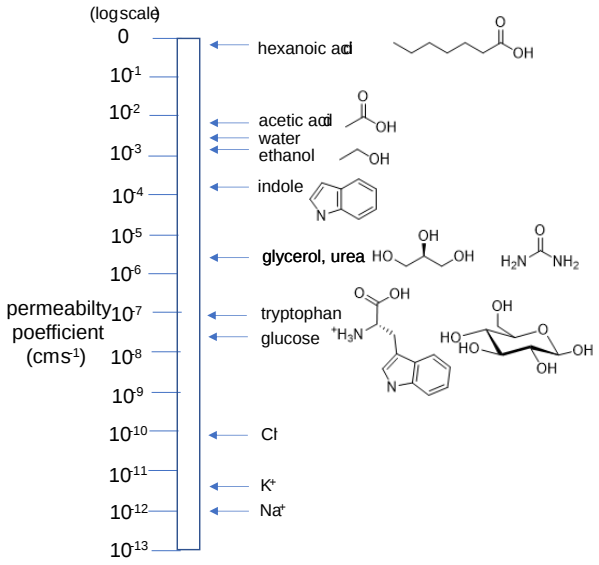


Figure 11.2.1: Permeability coefficients for relevant biological molecules

Smaller, higher charge density ions (like Na^+) have a lower permeability coefficient than do larger, lower charge density ions (like K^+) as seen in (Table 11.2.1). What about natural membranes?

Table 11.2.1: Permeability coefficient (cm/s) of natural and synthetic membranes to D-Glucose and D-mannitol at 25 °C

Membrane Preparation	D-Glucose	D-Mannitol
Synthetic Lipid Bilayer	2.4×10^{-10}	4.4×10^{-11}
Calculated Passive Diffusion	4×10^{-9}	3×10^{-9}
Intact Human Erythrocyte (red blood cell)	2.0×10^{-4}	5×10^{-9}

Looks like D-glucose gets a little help in getting across. We'll see the mechanism below.

11.2.2: Passive Diffusion

Let's start with the passive diffusion of uncharged solute A across a membrane, which can be represented by the chemical equation $A_{out} \leftrightarrow A_{in}$. Intuitively, you probably believe that the rate of **net** diffusion or the **flux** of A across the membrane is directly proportional to the concentration gradient across the membrane. If concentrations of A are identical across the membrane, the net flux J should be 0. If you double the concentration gradient, the net rate should double. We will see that the net rate is a linear function of [ΔA] across the membrane. Figure 11.2.2 shows the flux of A_{out} across a semipermeable membrane of thickness Δx (we will use dx instead when Δx is very small).

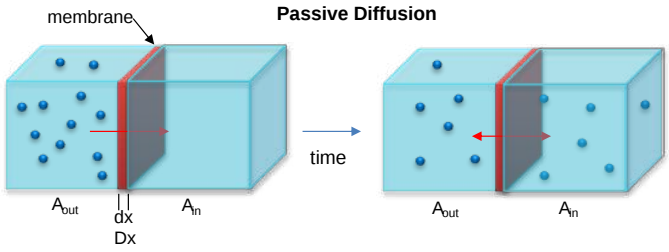


Figure 11.2.2: Flux of A_{out} across a semipermeable membrane of thickness Δx

Let's animate diffusion using a PHET simulation, as shown in Figure 11.2.3

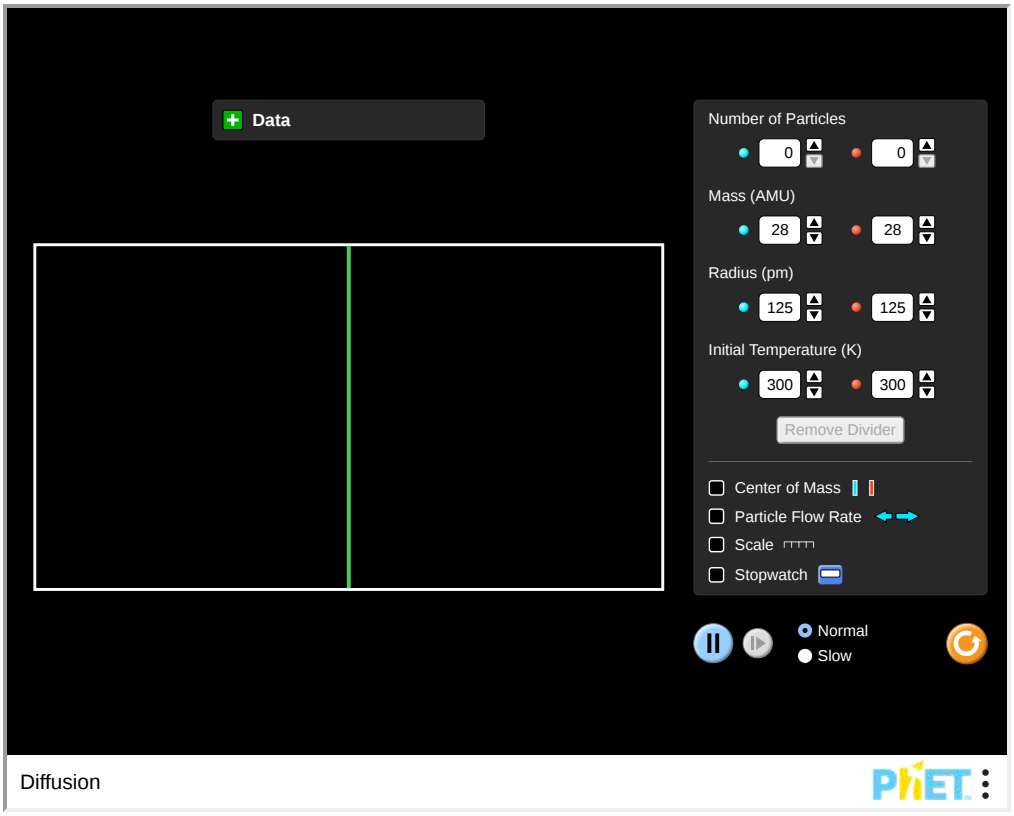


Figure 11.2.3: PHET animation of passive diffusion. PHET: https://phet.colorado.edu/

The flux of molecules A (\(J_A\)) is proportional to the concentration gradient across the membrane, $\Delta A/\Delta x$ (which we will refer to as dA/dx which is the derivative of A with respect to x). The equation below is Fick's First Law of Diffusion:

$$\mathrm{J_A} \propto \frac{d\,\mathrm{A}}{dx} = -\mathrm{D}\frac{d\,\mathrm{A}}{dx} \tag{11.2.1}$$

where D is the diffusion coefficient. The negative sign is necessary since concentration increases to the left in the figure above in the opposite direction of net flux which is to the right. For these derivations, we will assume the $\mathrm{J_A}$ is the initial flux. That is, the flux is measured for a short enough time that the relative concentrations of A on both sides of the membrane do not change significantly. It should be clear that eventually the net flux levels off to zero when the concentrations of A on both sides of the membrane are equal. Under these conditions, the free energy $G_{A\,out} = G_{A\,in}$, so $\Delta G = 0$. This thermodynamic relation can also be expressed as

$$J_A = -L\frac{dG_A}{dx} \tag{11.2.2}$$

This equation bridges the kinetic and thermodynamic aspects of diffusion.

Dimensional analysis of Fick's 1st Law (Equation 11.2.1) shows that the units of D are $\mathrm{cm^2/s}$ which gives the number (dimensionless) of molecules crossing a 1 $\mathrm{cm^2}$ surface area of membrane each second.

J = moles/area/sec = $\mathrm{mol/cm^2{\cdot}s}$ = - $(\mathrm{cm^2/s})$ $\mathrm{mol/cm^3/cm}$. Hence the units of D are $\mathrm{cm^2/s}$.

Let's rearrange Fick's 1st Law and use a bit of calculus to get Equation 4.

$$\begin{aligned} \mathrm{J_A}\int_0^{\mathrm{x}} \mathrm{dx} &= -D\int_{\mathrm{A_{out}}}^{A_{\mathrm{in}}} \mathrm{dA} \\ \mathrm{J_A x} &= -\mathrm{D}\,(\mathrm{A_{in}} - \mathrm{A_{out}}) \\ \mathrm{J_A} &= \frac{\mathrm{D}}{\mathrm{x}}(\mathrm{A_{out}} - \mathrm{A_{in}}) \end{aligned} \tag{11.2.3}$$

or

$$J_A = P\,(A_{\text{out}} - A_{\text{in}}) = P\Delta A \tag{11.2.4}$$

where P is the permeability coefficient, which has units of $cm^2/s/cm$ or cm/s. (We discussed permeability coefficients for different solutes traversing model bilayers when we discussed lipids.) That unit is less intuitive to understand but the final unit is very intuitive.

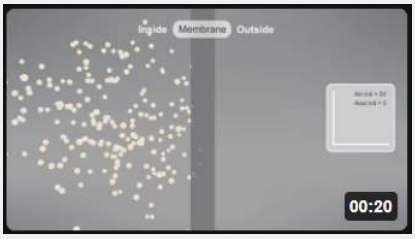
A plot of J_A vs (A_{out} - A_{in}) is linear, with a slope of P = D/x

It's important to remember that there is still diffusion of A across the membrane at equilibrium since the equilibrium is dynamic. There is no net diffusion, however. This is where animations come in handy.

The table below shows the reaction diagram (left), graphical results of the progress curve(middle), and animations for the reversible diffusion $A_{out} \leftrightarrow A_{in}$ across a membrane with the conditions shown below. The reactant A and product are called species and are shown as green spheres. The yellow square is a reaction node indicating a reaction connects A to P. The lines connect the species that participate in the reaction. The velocities (slope of the concentration vs time curve at any given time) are called **fluxes, J**, in Vcell and many other similar programs. When we get to metabolism, we will talk about fluxes of metabolites through pathways. Also, fluxes are used to describe the rate of movement of solute through membranes.

Animations

Now let's look at animations of the same passive diffusion reaction of a neutral species across a semi-permeable membrane. Animations are by Shraddha Nakak and Hui Liu.

$A_{out} \leftrightarrow A_{in}$ $[A_{in}]_{t=0} = 100$; $[A_{out}]_{t=0} = 0$; P = 2	$A_{out} \leftrightarrow A_{in}$ $[A_{in}]_{t=0} = 50$; $[A_{out}]_{t=0} = 0$; P = 2
	

Note the dynamic nature of the diffusion. An equilibrium is reached when the number of particles inside equals those outside. Diffusion occurs in both directions from compartments of equal volume so that the particles moving to the outside don't escape into a comparatively huge volume. (Animations by Shraddha Nayak and Hui Lui)

11.2.3: Passive Diffusion of ions across a membrane - Transmembrane Potentials

Before we move on to facilitated diffusion, let's alter the scenario a bit and use a charged solute. In the example of passive diffusion above, the only thermodynamic driving force for the movement of A across the membrane was the ΔG_A, the change in free energy/mol of A across the membrane (or more strictly $\Delta\mu$ = change in chemical potential). Solute A moves spontaneously across the membrane from high to low concentration. But what if A was charged?

We could add a bunch of positively or negatively charged species to Figure 11.2.2 and ask what would happen. You can't go to a chemical stockroom and find a bottle of K^+ ions but you could find a bottle of neutral KCl. Let's make our experiment system a vesicle that has 0.1 M KCl in the aqueous inner compartment with 0.1 M NaCl on the outside. We could easily prepare such vesicles by making large unilamellar vesicles (LUVs) with entrapped 0.1M KCl in a solution of 0.1 M KCl and then separate the vesicles from the 0.1 M KCl not encapsulated on liposome formation using a size exclusion gel chromatography column equilibrated and eluted in 0.1 M NaCl. These vesicles are illustrated on the left of Figure 11.2.4

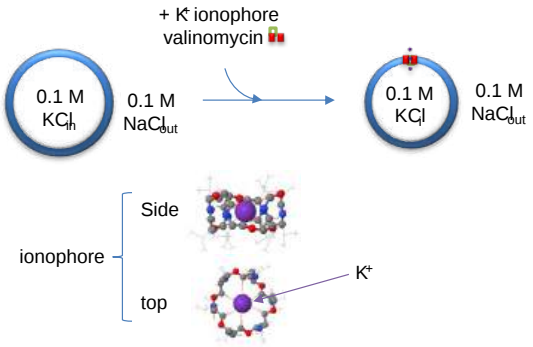


Figure 11.2.4: Making vesicles with a transmembrane potential

In these prepared vesicles is there a net thermodynamic driving for K^+ and Cl^- to move from inside to outside the vesicle? Not for Cl^- since its concentration is the same on both sides of the membrane (see Fick's 1st Law above). However, there is a clear thermodynamic driving force to take $K^+_{in} \rightarrow K^+_{out}$. If the membrane was impermeable, net outward flux would not occur even though it is favored thermodynamically. Think of this as an example of a reaction under complete kinetic control! Note that in this example there is also a net driving force to move Na^+ ion from outside to inside as well.

In our next step, let's make the membrane permeable to only K^+ ions. We can do this by adding a small antibiotic, valinomycin, which binds in the membrane and once there binds and carries K^+ ions across the membrane. It is called an **ionophore.** Figure 11.2.5 shows an interactive iCn3D model of K^+ bound to Valinomycin.

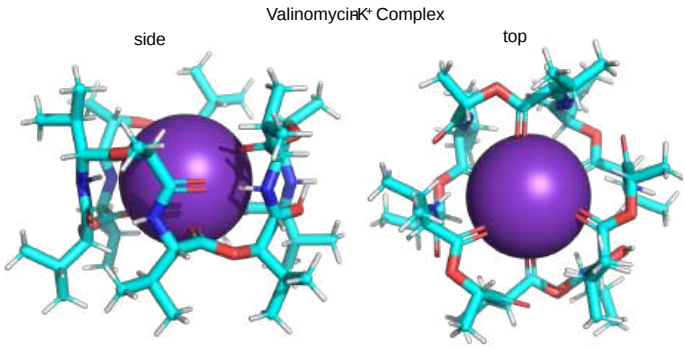


Figure 11.2.5: K^+ bound to Valinomycin (xxxx). (Copyright; author via source). Click the image for a popup or use this external link: Not available

Valinomycin, from *Streptomyces fulvissimus*, is a cyclic peptide consisting of L and D-Val along with L-lactate and D-hydroxyisovalerate, connected through both ester and amide bonds. The K^+ ion is in the center. The six valine carbonyl oxygens bind the K^+ ion. The hydrophilic groups are pointed toward the center, while the hydrophobic groups point to the outside of the structure, allowing the K^+ ion to be sequestered in a polar environment as the nonpolar exterior of the complex passes through the membrane. This ionophore is specific for K^+ and binds the smaller Na+ ion weakly. This can be accounted for by two factors. The smaller sodium ion doesn't bind as tightly to the chelating carbonyl oxygens. Also, the sodium ion has a higher charge density, so the Na^+/water interactions must be more stable and more difficult to break than those to K^+. The ion must be desolvated before it binds to the complex. Other ionophores are specific for other ions.

Once the ionophore is bound, the kinetic barriers to K^+ efflux are removed and it starts moving K^+ from inside to outside. However, as soon as it does, the charge balance across the membrane is lost, with the outside becoming net positive and the inside becoming net negative. A transmembrane electric potential develops across the membrane. This disfavors stops K^+ efflux to the outside and eventually stops it even as the concentration difference of K^+ across the membrane still favors efflux.

If you were a positive ion stuck in the middle of a membrane, as illustrated in Figure 11.2.6, which way would you move?

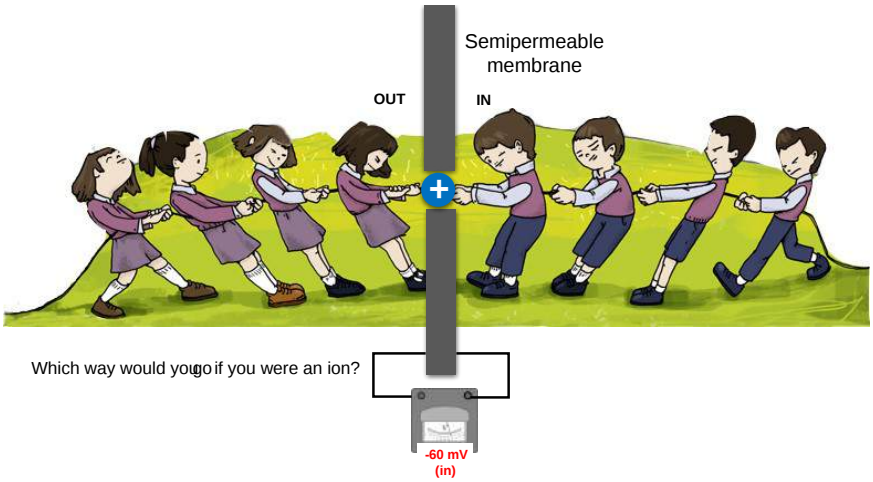


Figure 11.2.6: Which way does a positive ion move across a membrane with a transmembrane potential? (adapted from https://commons.m.wikimedia.org/wiki/File:Dib_1.jpg

There are now two thermodynamic driving forces for K^+ movement from inside to outside:

- a $\Delta G_{concentration}$ which favors K^+ efflux. At time t=0, $\Delta G_{concentration} << 0$ and it becomes a bit less negative (less favored) with efflux.
- a $\Delta G_{membrane\ pot}$ which is zero to start and slowly becomes positive, increasingly disfavoring K^+ efflux.

When these driving forces are equal and opposite, **net** K^+ movement across the membrane stops and the system is in dynamic equilbrium.

Since we use electric potential to describe electrical phenomena (electron, ion movement), we often use the word **chemical potential** in this case to describe the movement of ions across a concentration gradient. Add them together and we call it the **electrochemical potential**.

$$\begin{aligned} \Delta G_{electrochemical} &= \Delta G_{chemical} + \Delta G_{electrical} \text{ or} \\ \Delta \mu_{electrochemical} &= \Delta \mu_{chemical} + \Delta \mu_{electrical} \end{aligned} \tag{11.2.5}$$

We can use this understanding to derive an equation for flux J of a charged solute across a membrane of a given potential. The equation is called the Goldman Equation and is shown below.

$$J = \frac{P\frac{ZF}{RT}(E_1 - E_2)\, C_1 \left(1 - \frac{C_2}{C_1} e^{\frac{ZF}{RT}(E_2 - E_1)}\right)}{1 - e^{\frac{ZF}{RT}(E_2 - E_1)}} \tag{11.2.6}$$

where

- P is the permeability coefficient
- Z is the charge or valence on the ion
- F is the Faraday constant
- R is the ideal gas constant
- T is temperature
- E2-E1 and the reverse is the transmembrane potential
- C2-C1 are the concentrations of the ions across the membrane

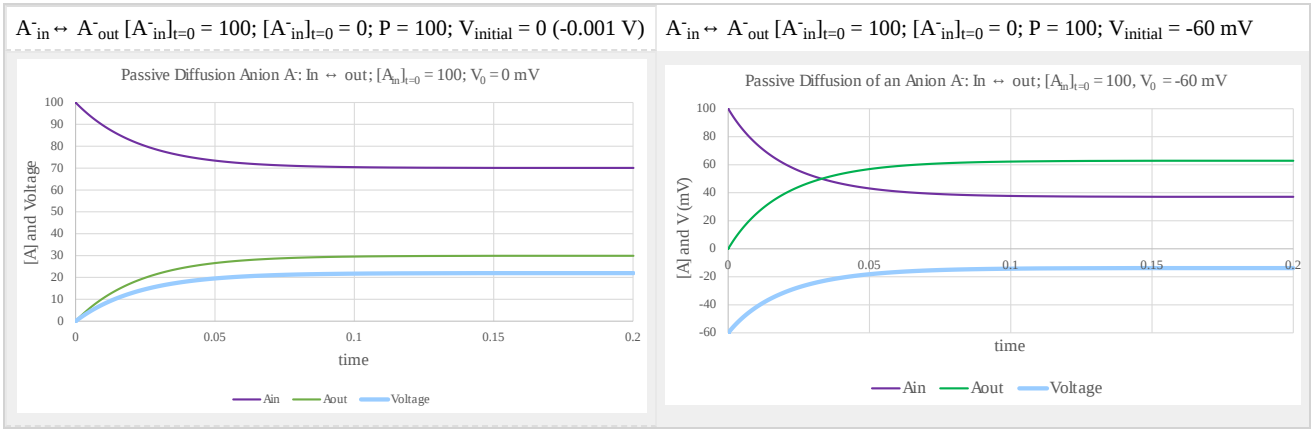
Compare this to the Nernst equation which you learned in introductory chemistry courses.

$$E = E^o - \frac{RT}{nF} \ln Q \tag{11.2.7}$$

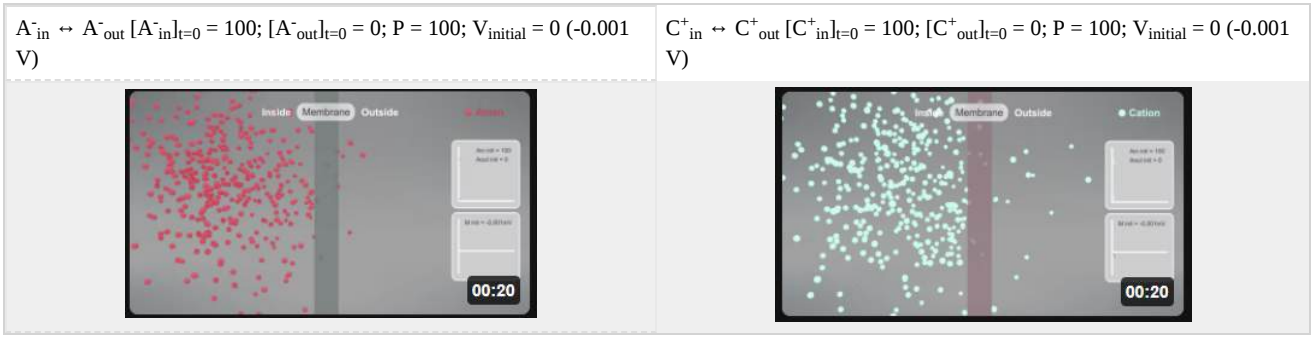
that relates the reduction potential of an electrochemical reaction to the standard electrode potential, temperature, and concentration where E is the potential difference.

Now let's run a Vcell simulation for the diffusion of an anion across a semipermeable membrane. To do so we must first set the initial transmembrane potential to solve the Goldman equation numerically. At present, this type of simulation **can not** be

embedded into this book. So instead, the concentration vs time graphs for two different simulations, one at an **initial** transmembrane potential (E_2-E_1) = - 0.001 (i.e. 0) and one at -60 mV (a typical cell resting potential), are presented in the Figure below. In each, the reaction is $A^-_{in} \leftrightarrow A^-_{out}$ $[A^-_{in}]_{t=0}$ = 100.



Here are animations that show the selective reversible movent of anions (A^-, left panel, red) and cations (C^+, right panel, cyan) as they move across a membrane from the inside to the outside. This is a very simplified simulation as it shows no counter ions on either side. Assume they exist as it would be impossible to have a "container" with just anions or cations. The initial transmembrane potential (t=0) is 0 (actually -0.001 to allow the calculations using Vcell).



As anions move to the outside (left animation), the inside becomes less negative with respect to the outside, so the membrane potential V becomes more positive. This is indicated by the membrane changing to a blue color. Conversely, as cations move to the outside (right animation), the inside would become more negative with respect to the outside, so the membrane potential V becomes more negative. This is indicated by the membrane changing to a red color. (Animations by Shraddha Nayak and Hui Lui)

11.2.4: Facilitated Diffusion

Now let's return to the diffusion of a noncharged solute down a concentration gradient (i.e. favored) after binding to a membrane receptor. The answer to that question depends on the biological function of the macromolecule. We can simplify this process by adding one additional step as reflected in the equilibrium binding expression shown below:

$$M + L \rightleftharpoons ML \rightleftharpoons M + X$$

This expression indicates that the free ligand has changed in some fashion to x. In the next two chapters, we will consider two kinds of transformations:

- L is a ligand on the outside of a biological membrane (Lout) that binds to a membrane protein receptor, R. This undergoes a conformational change (as we studied in the binding of dioxygen to hemoglobin) which leads to the expulsion of the bound ligand to the inside of the membrane (Lin). This can be modeled with the simple equation:

  $$R + L_{out} \rightleftharpoons RL \rightleftharpoons R + L_{in}.$$

  This process is called facilitated diffusion and represents a physical as opposed to chemical process since no covalent bonds are made or broken. This process proceeds down a concentration or chemical potential gradient ($\Delta\mu < 0$) and hence is spontaneous

(thermodynamically favored). If the ligand concentration is higher inside the cell, net diffusion moves it to the outside of the cell. Passive (non-facilitated) diffusion is kinetically slow in the absence of a receptor since membranes present formidable barriers to the passage of polar molecules.

- L is a ligand (or substrate S) that binds to a protein enzyme, E. The bound substrate is chemically altered to produce a new product, P, which dissociates from the enzyme. This can be expressed most simply as:
  E + S <==> ES <==> E + P .

Consider the mechanism illustrated in Figure 11.2.7.

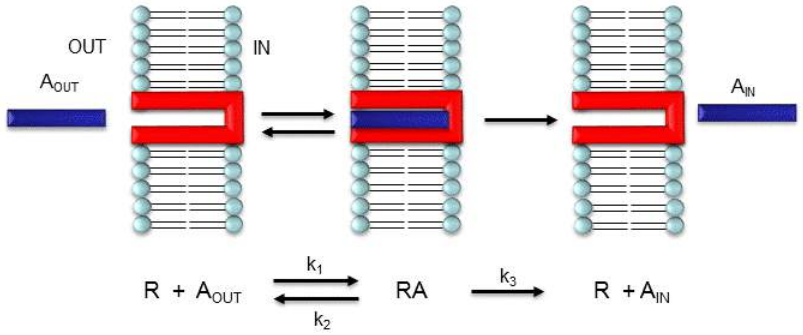


Figure 11.2.7: Mechanism for faciliatated diffusion

Let's assume that for this system the initial flux will be measured. We would like to derive equations that show J as a function of Aout (assuming that Ain is negligible over the time course of measuring the initial flux. Also, assume that the J facilitated is much greater than J passive. In contrast to passive diffusion, JA is not proportional to Aout but rather to [Abound].

Consider this example to help you understand that proportionality. Pretend that the receptor is a truck that can carry one particle across the membrane at a time (i.e. 1/1 stoichiometry). Also, assume that the particle can't get across without being carried by the truck. If there are no trucks in the membrane, no load can be delivered. If there are trucks in the membrane but no particles in them, no load will be delivered. As the number of particles available to be loaded into the truck increase, the truck will have an increased chance to be loaded (depending of course on the affinity of the particle for the truck). If the number of loaded trucks is doubled, the number of particles dumped to the other side will double. Therefore, by analogy,

JA is proportional to [RA], or

$$J_A = \text{const}[\text{RA}] = \text{k}_3[\text{RA}] \tag{11.2.8}$$

How can we calculate RA when we know A and R? Let us assume that Atotal (A0) is much greater than R0, as is the likely biological case, and Ain = 0. We can calculate RA using the following equations, and the same procedure we used for the derivation of the binding equation

$$[\text{ML}] = \frac{[\text{M}_0]\,[\text{L}]}{\text{K}_\text{D} + [\text{L}]} \tag{11.2.9}$$

The equation for the dissociation constant K_D

$$\text{K}_\text{D} = \frac{[\text{A}]_{\text{eq}}[\text{R}]_{\text{eq}}}{[\text{RA}]_{\text{eq}}} = \frac{(\text{A})(\text{R})}{\text{RA}} \tag{11.2.10}$$

The equation of mass balance of R

$$\text{R}_0 = \text{R} + \text{RA} \text{ so } \text{R} = \text{R}_0 - \text{RA} \tag{11.2.11}$$

Since we will assume that A_0 is much greater than R_0, we will not need the mass balance for A (which is Ao = A + RA).

Substitute x into x and rearrange to get:

$$\begin{aligned} \text{K}_\text{D}(RA) = (A)(\text{R}) = (A)\,((\text{R}_0) - (A)(RA) \\ \text{K}_\text{D}(RA) + (A)(RA) = (A)\,(\text{R}_0) \\ (\text{K}_\text{D} + \text{A})\,(RA) = (A)\,((\text{R}_0) \\ (RA) = \frac{(\text{R}_0)\,A}{\text{K}_\text{D} + \text{A}} \end{aligned} \tag{11.2.12}$$

Substitute x into z gives the final equation,

$$J_A = k_3[RA] = \frac{k_3 (R_0) A}{K_D + A} = \frac{J_{max} A}{K_D + A} \tag{11.2.13}$$

It should be clear to you from this equation that:

- a plot of J_A vs A is hyperbolic
- $J_A = 0$ when $A = 0$.
- $J_A = J_{max}$ when A is much greater than K_D
- $A = K_D$ when $J_A = J_{max}/2$.

These are the same conditions we detailed for our understanding of the binding equation.

This derivation is based on the assumption that the relative concentrations of A, R, and AR can be determined by the K_D for the interactions and the concentrations of each species during the early part of diffusion (i.e. under initial rate conditions). Remember under these conditions, A_{out} does not change much with time. Is this a valid assumption? Examine the mechanism shown in the above figure. A_{out} binds to R with a second order rate constant k_1. RA has two fates. It can dissociate with a first-order rate constant k_2 to A_{out} + R (to give the original species), or dissociate with a first-order rate constant of k_3 to give A_{in} + R (as A moves across the membrane). If we assume that $k_2 >> k_3$ (i.e. that the complex falls apart much more quickly than A is carried in), then the relative ratios of A, R, and RA can be described by K_D. Alternatively, you can think about it this way. If A binds to R, most of A will dissociate, and a small amount will be carried across the membrane. If this happened, then R is now free, and will quickly bind A_{out} and reequilibrate. This occurs since the most likely fate of bound A is to dissociate, not to be carried across the membrane, since $k_3 << k_2$.

11.2.5: "Receptors" in Facilitated Diffusion

Two types of proteins are involved in facilitated diffusion, **carriers and channels**. **Carrier** proteins (also called permeases or transporters) such as the glucose transporter (GLUT1) move solute molecules across a membrane while **channels/pores** facilitate the diffusion of ions down a concentration gradient by providing a pore in the membrane. We won't describe in this section the more complicated processes of phagocytosis and endocytosis. These processes are illustrated in Figure 11.2.8

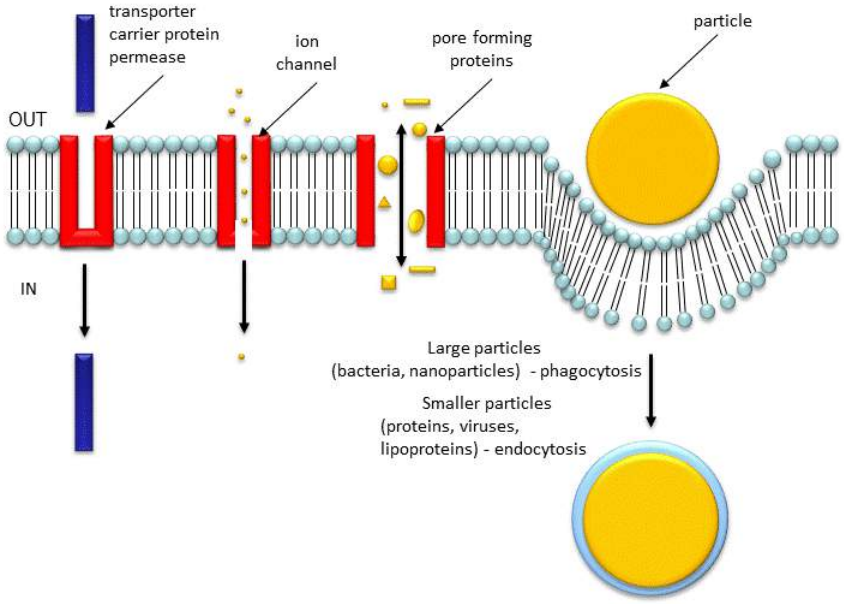


Figure 11.2.8: Movement of particles across a membrane

In the case of permeases and transport proteins, ligands bind and induce a conformational change in the receptor as illustrated in the case of the glucose transport protein shown in Figure 11.2.9

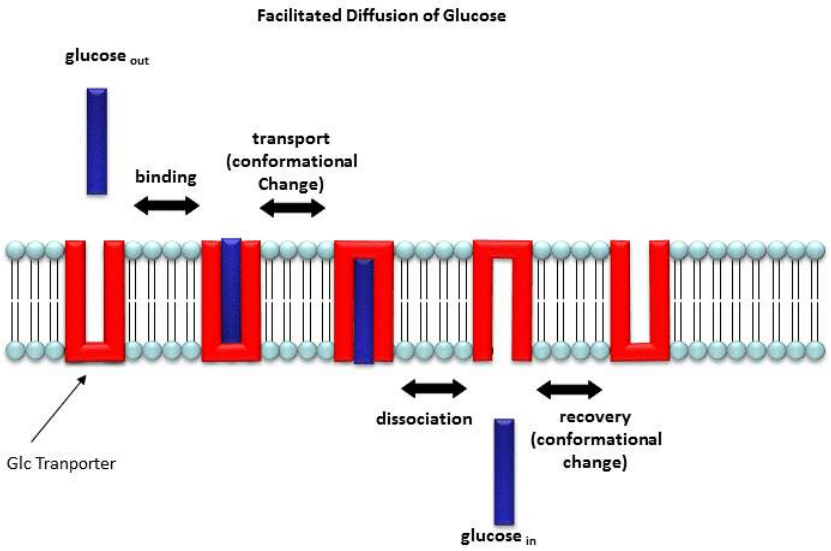


Figure 11.2.9: Model for facilitated diffusion of glucose across a membrane

In channels and pores, a ligand can bind to the receptor (channel protein), which induces a conformational change in the receptor, a "ligand-gated" channel through the membrane. This process would lead to the diffusion of many ions across the membrane (down a concentration gradient) until the channel closes (which can be induced by ligand dissociation or other events).

The mathematics we derived for the carrier proteins does not apply to the channel proteins. In addition, there are other ways to "gate" open a channel protein, which we will discuss later. Also, some transporters can move solute molecules across a membrane against a concentration gradient. These proteins require an external energy source (like ATP or coupling to the favorable collapse of a second transmembrane gradient) to drive this thermodynamically unfavored process. This is called active transport and will be discussed in the next chapter section.

- ribitol diffusion through a glycerol channel using interactive molecular dynamics

Both links above are from the Theoretical and Computational Biophysics group at the Beckman Institute, University of Illinois at Urbana-Champaign. These molecular dynamic simulations were made with VMD/NAMD/BioCoRE/JMV/other software support developed by the Group with NIH support.

11.2.6: Carrier proteins (permeases or transporters)

Now let's look at some examples of carrier proteins:

Glucose Transport Proteins

Glucose is a key metabolic fuel so its movement into cells is critical and hence highly regulated. There are multiple types of glucose transporters. **GLUT 1**, a plasma membrane protein, found in most cells, is responsible for constitutive or basal glucose uptake while **GLUT 4** is involved in insulin-regulated uptake in skeletal and heart muscles and adipose cells, of glucose after meals. Its official name is **solute carrier family 2** or, **facilitated glucose transporter member 4.** No structure is yet available for GLUT 4 but there is for GLUT1, which is highly expressed in cancer cells that have high energy demands.

Figure 11.2.10 shows an interactive iCn3D model of a glucose transporter, GLUT1 (5eqg), bound to an inhibitor, cytochalasin B (spacefill). The inhibitor binds in the inward-open state where glucose binds.

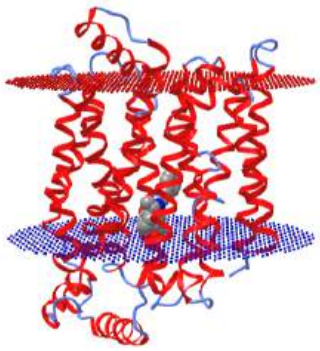


Figure 11.2.10: Glucose transporter, GLUT1 bound to the inhibitor, cytochalasin B (5eqg). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...zUTZEcTWQrQAX6

Mitochondrial ADP-ATP Carrier Protein

We will see in a few chapters that most of the ATP made in cells takes place in the mitochondrial matrix. It won't do cells much good if it stays in there since it is needed in the cytoplasm and elsewhere to drive unfavored processes. Likewise, when ATP is depleted in a cell, ADP is concomitantly high. What is needed is an inner mitochondrial membrane protein that can shuttle ATP out of the mitochondria and ADP in down concentration gradients. It would not make sense to need to power an uphill movement of ATP into the mitochondria from low to high concentration driven by ATP cleavage. Let's look at the structure of the bovine ADP-ATP carrier protein which resides in the inner membrane. Figure 11.2.11 shows an interactive iCn3D model of the bovine mitochondrial ADP-ATP carrier protein (1okc).

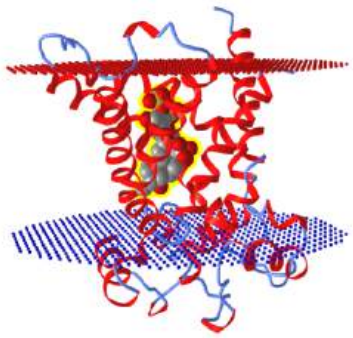


Figure 11.2.11: Bovine mitochondrial ADP-ATP carrier protein (1okc). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?Zc9iLpCncYYD7gBc6

The transmembrane domain contains six alpha-helices which form a depression leading to the inner leaflet. The cyan spacefill amino acids on the bottom of the depression are RRRMM, which is a motif found in nucleotide carrier protein. A conformational transition must transiently open the depression into a channel. The spacefill molecule in CPK colors represents carboxyatractyloside, a diterpene glycoside that inhibits the carrier protein.

A Special Case: Fatty acid carrier proteins

You might guess that free fatty acids, derived for example from lipids after the actions of lipases on triacylglycerol, would **not** need a carrier protein to move across the cell membrane since they are almost completely nonpolar. Hexanoic acid can indeed pass readily, but for solute diffusion across the membrane, size matters as well. A whole family of proteins, **Fatty Acid**

Transport Proteins (FATPs) have evolved to help long-chain fatty acids across membranes. Human fatty acid transport proteins are transmembrane proteins. Its mechanism of action is unclear. No crystal structures of these are readily available. Many proteins in this class catalyze the formation of fatty acid-CoASH derivatives, which is an endergonic reaction powered by ATP. The mechanism of fatty acid movement across the membrane probably may involve simple diffusion coupled to processes driven by ATP. However, there is still controversy on the role of passive vs facilitated diffusion for fatty acids.

First, let's consider the problems facing a cell in moving fatty acids across two aqueous environments. Figure 11.2.12 shows a mass balance depiction of the reservoirs of fatty acid in the extracellular and intracellular environment.

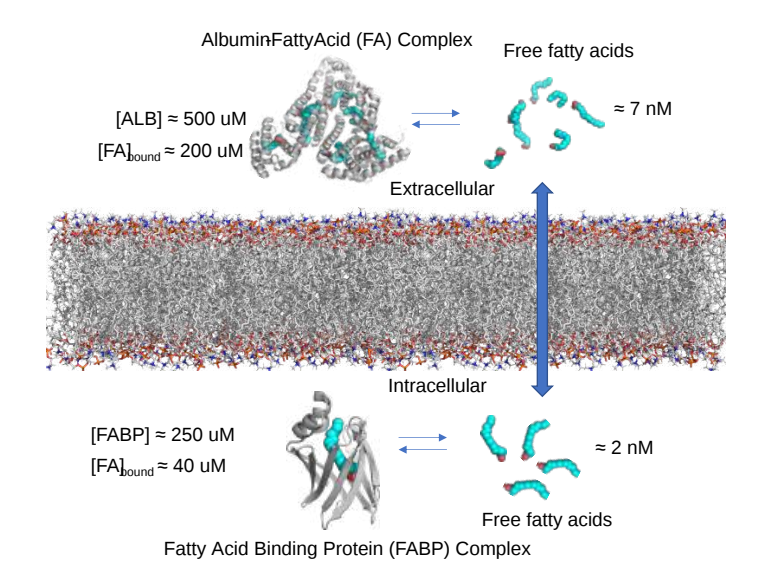


Figure 11.2.12: Mass balance depiction of the reservoirs of fatty acid in the extracellular and intracellular environment

Free fatty acids are very insoluble in aqueous solution so their concentrations on either side of the membrane are very low, in the low nanomolar range. Hence there is no great thermodynamic drive to move free fatty acids across the membrane. If you assume that fatty acids can reasonably transverse the membrane without a carrier protein, there would be no huge kinetic barriers to movement except their low concentrations.

On each side of the membrane, the free fatty acids are in an "equilibrium" with protein-bound amino acids. In the blood and interstitial fluids, albumin, which can bind multiple fatty acids, is in high concentration, so it can act as a buffer to keep free fatty acids in a useful concentration range. Likewise, in the cytoplasm, **fatty acid binding proteins (FABPs)**, which typically bind just one fatty acid, are also relatively high in concentration and buffer the free fatty acids in the cytoplasm. One other note. Free fatty acids are single-chain amphiphiles which makes them detergents, which could easily lyse cell membranes, so their free concentrations must be kept very low compared to the critical micelle concentration.

Figure 11.2.13 shows an interactive iCn3D model of the human brain fatty acid-binding protein bound to docosahexaenoic acid (1fdq).

Figure 11.2.13: Human brain fatty acid-binding protein bound to docosahexaenoic acid (1fdq). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...M2sTpEuU6hj5RA

How do long-chain fatty acids cross the membrane? We will first examine the role of proteins. Let's look at a couple of players.

Fatty acid transport proteins (FATPs): There are six members of the human FATP family, which is also known as the Solute carrier family 27. **FATP 1 (SLC27A-1)** is found in plasma and endoplasmic reticulum membranes and based on sequence analysis it is a single-pass membrane protein. Its highest expression is in muscle and adipose cells. It possesses a C-terminal AMP binding domain and acyl-CoA synthase activity. The mouse protein has an N terminal transmembrane domain and predictions from the human sequence show there is likely just one transmembrane helix at the N terminus (amino acid 13-35). Given that, and the absence of a 3D structure, it would appear that this protein would not bind and transfer the bound lipid across the membrane through a conformational change in the transmembrane domain. FATP 4, located in the ER membrane, is predicted to have 2 transmembrane helices, which still would probably be inadequate to serve as a class translocase. It is expressed in the endoplasmic reticulum cell membrane.

Platelet Glycoprotein 4 (CD36): Another candidate is platelet glycoprotein 4, which is also called CD36, Glycoprotein IIIb, fatty acid translocase, or the thrombospondin receptor. The protein has many functions and binds many types of proteins (thrombospondin, fibronectin, collagen or amyloid-beta) and lipids (oxidized low-density lipoprotein (ox-LDL), anionic phospholipids, long-chain fatty acids, and bacterial diacylated lipopeptides). It is present in plasma and Golgi membranes. Sequence analysis shows that it passes through the membrane twice (amino acids 7-29 and 441-463) and it also is palmitoylated at both N- and C-terminal ends. A PDB structure (5LGD) for most of the protein except the putative N- and C-terminal helices are known. In the structure, it is bound to a malarial protein (shown in grey) and two palmitic acids (spacefill) bound in the nonmembrane domains. Again, from this description, it doesn't appear that the bound fatty acids are translocated via a conformational change in the receptor as described above for glucose and ADP/ATP.

Fatty Acid Binding Proteins (FABPs): These proteins might also take part in the process. In addition to the cytoplasmic form, there is also a plasma membrane-associated fatty acid-binding protein ($FABP_{pm}$), also known as FABP-1. It's the same protein as mitochondrial aspartate aminotransferase (UniProtKB - P00505 (AATM_HUMAN). It has many possible functions.

Figure 11.2.14 shows a possible model for how fatty acids may transfer or be handed off from albumin to membrane-bound GP36 or FATP-1, possibly through an intermediary protein like FABPpm (for GP36).

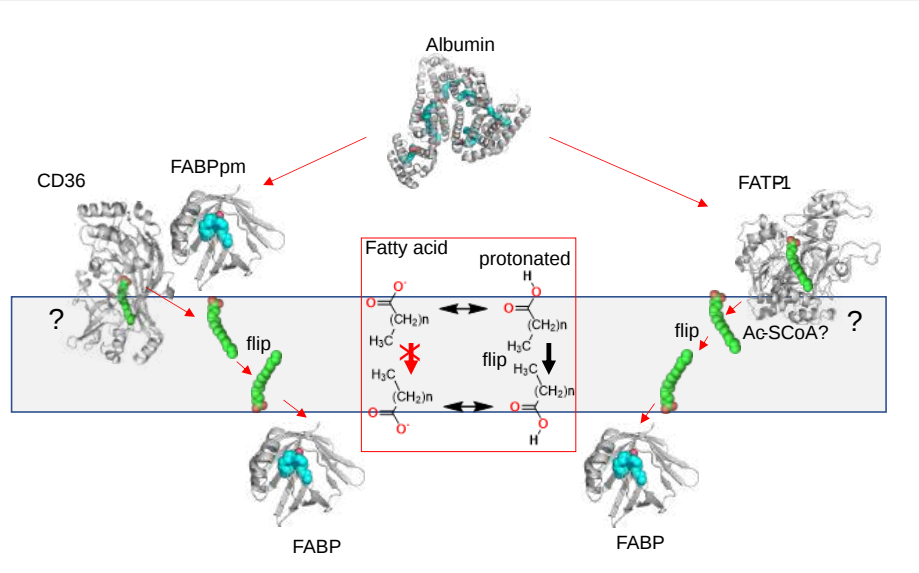


Figure 11.2.14: Possible model for how fatty acids may move from albumin to membrane-bound GP36 or FATP-1

Some proteins (albumin, FABP) deliver fatty acids to the membrane proteins (CD36, FATP-1), which deliver free fatty acids into the outer leaflet, where they flip to the inner leaflet, where they are picked up by membrane-associated cytoplasmic fatty acid binding proteins (FABP). Until structures are known for the transmembrane-bound proteins GP36 and FATP-1, whose full amino acid sequences don't suggest a classic carrier protein, this mechanism is a reasonable one. FATP-1 has acyl-CoA synthase activity so it is likely that the transferred fatty acid is converted to the acyl-CoASH before movement to the cytoplasm. Fatty acid transporters are also implicated in insulin resistance and type 2 diabetes.

The alternative model proposes that long-chain fatty acids, the preferential energy source for cardiac muscle, can cross the membrane by passive diffusion (red boxed area above) even if the activity of CD36 is inhibited. Diffusion depends on an alteration of the pKa of outer membrane adsorbed free fatty acids (around 7.5) compared to free fatty acids in solution (around

4.5). The protonated fatty acid adsorbed to the membrane would move into the outer leaflet where it would flip to the inner leaflet and be picked up by cytoplasmic and/or peripheral membrane proteins. This process would be associated with movement of H^+ across the leaflet as well. This type of diffusion has been observed in protein-free lipid vesicles and cells. Likewise, long chain fatty amines (instead of carboxylic acids) can diffuse into vesicles and cells which would support this passive diffusion if fatty acid-binding proteins don't bind the amine forms.

11.3: Diffusion Across a Membrane - Channels

11.3.1: Introduction

If you punched a hole or pore in the membrane, depending on its size, multiple types of chemical species could flow through it simultaneously. We'll talk about pores in the next section. Let's focus on channels, which have much smaller openings, which are gated open to allow ion flow through them. They are often called **ionotropic** receptors. Channels can be "gated" open by many mechanisms including ligand binding, change in membrane potential, lipid interactions, and mechanical stress. Opening a channel to ion flow allows quick passage of information (in this case an electrical signal) into the cell, leading to quick cellular responses. This is an ideal signaling mechanism for neural cells which demand quick responses.

We'll show examples of each type of gating mechanism. Before we do, it is helpful to know typical extracellular and intracellular ion concentrations in a mammalian neuron, for example (Table 11.3.1).

Table 11.3.1: Extracellular and intracellular ion concentration in a mammalian neuron

ion	extracellular = $[ion]_{out}$ (mM)	intracellular = $[ion]_{in}$ (mM)
Na^+	145 mM	5-10 mM
K^+	5 mM	140 mM
Cl^-	110 mM	10 mM
Ca^{2+} (free)	1.2 mM	100 nM

When ion channels are opened in neural cell membranes, the direction of favorable thermodynamic flow is down a concentration (chemical potential) gradient but the direction is also affected by the transmembrane potential. Typical resting potentials of neural cells are about -60 to -70 mV (negative inside). If a nonspecific **cation** channel is gated open, the kinetic barriers to diffusion are relieved and at that moment Na^+ ions would flow in due to both the chemical and electrical potential, while K^+ ions would flow out but with less driving forces since its efflux is hindered by the negative transmembrane potential. How are such large gradients of these ions formed? We'll answer that in the next section on active transport.

11.3.2: Pentameric ligand-gated ion channels (pLGICs)

These channels play a key role in neuronal signaling. They are ligand-gated channels. In neural systems, the ligands are neurotransmitters. All are comprised of five monomers, which together form the functional channel with the pore formed in the center of the pentameric structure. The subunits can be identical (homopentamer) or different (heteropentamer). All have "Cys-loop" motifs so they have been called Cys-loop receptors as well. Examples include the mammalian nicotinic acetylcholine, serotonin (5-HT), γ-aminobutyric (GABA), glycine, and glutamate receptors.

Figure 11.3.1 shows the generic structure of the pLGICs

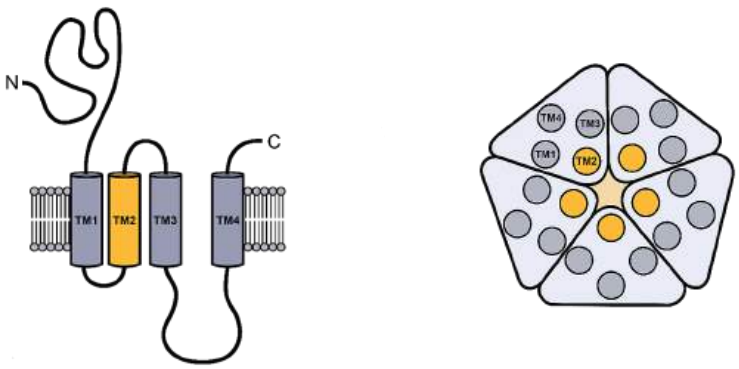


Figure 11.3.1: Generic structures of pentameric ligand-gated ion channels. Smelt, Charles L.C.; (2019) Allosteric modulation of pentameric ligand-gated ion channels. Doctoral thesis (Ph.D.), UCL (University College London). https://discovery.ucl.ac.uk/id/eprint/10075893/. Original content in this thesis is licensed under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) Licence (https://creativecommons.org/licenses/by/ 4.0/)

The monomeric structure is shown on the left. Each contains four transmembrane helices (TM1-4). A top-down view of the pentameric structure is shown to the right. The pore surface forms at the interface of the central TM2 helices. The ligand (neurotransmitter) binds to the extracellular domain with contributions from all the subunits. On ligand binding, TM2 and TM3 rearrange to allow the formation of a transient pore and passive diffusion of specific ions.

pLGICs are incredibly interesting and pharmacologically relevant. In general, they have two different types of binding sites.

Orthosteric sites bind ligands in the extracellular domains. When bound, a conformational change leads to the rearrangement of helices opening the pore. The natural ligand is also called the agonist as it promotes the function (either neuron excitation or inhibition) of the ion channel. The binding of the natural ligand/agonist opens up the channel to ion flow. This can lead to activation or excitation of the neural cell if positive cations flow into the cell, which depolarizes the cell as the transmembrane potential becomes more positive. Neurotransmitters that lead to this response are excitatory. Alternatively, the binding of inhibitor neurotransmitters in the orthosteric site can lead to inhibition of the neural cell activation if the channel is a ligand-gated anion channel. This hyperpolarizes (makes the transmembrane potential more negative), leading to inhibition of neural cell activation. Inhibitors or antagonists of channel function, whose structure typically resembles at least somewhat the structure of the endogenous ligand or agonist, also bind to the orthosteric site.

Allosteric sites are distal to the orthosteric site. Ligands that bind to allosteric sites also lead to conformational changes that either augment or diminish the effect of normal ligand/agonist binding by modulating ion flow through the pore.

pLGICs interact with analgesics and anesthetics which makes them even more interesting.

Figure 11.3.2 shows two excitatory pLGICs, the 5-hydroxytryptamine (5HT) (left) and nicotinic acetylcholine (right) receptors. The ligand binds in an orthosteric site in the extracellular domain (ECD), which is composed mostly of beta secondary structure. Allosteric sites are often found in the transmembrane domain.

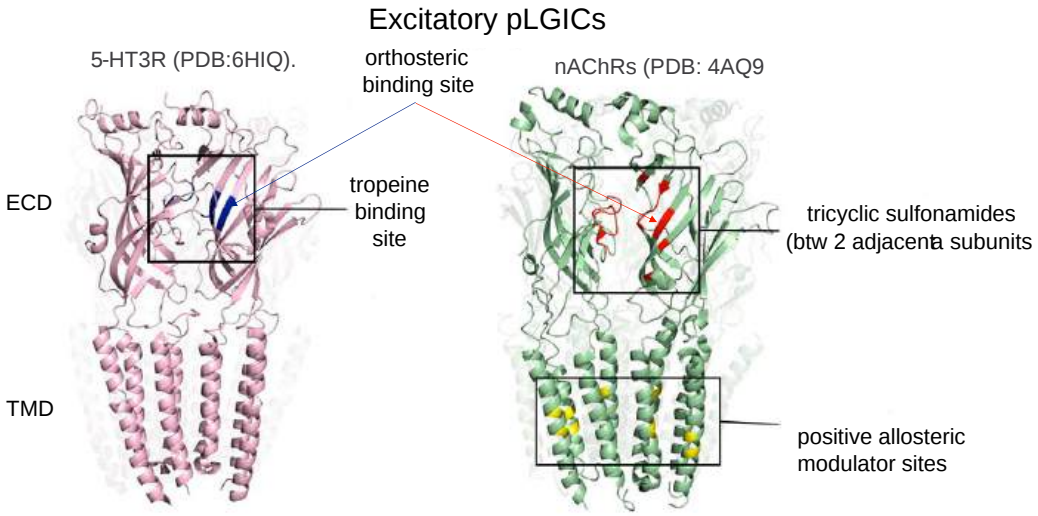


Figure 11.3.2: Two excitatory pLGICs, the 5-hydroxytryptamine (5HT) (left) and nicotinic acetylcholine (right) receptors. Adapted from Lara César O., Burgos Carlos F., Moraga-Cid Gustavo, Carrasco Mónica A., Yévenes Gonzalo E. Pentameric Ligand-Gated Ion Channels as Pharmacological Targets Against Chronic Pain, Frontiers in Pharmacology, 11 (2020), 167. https://www.frontiersin.org/article/...har.2020.00167, **Creative Commons Attribution License (CC BY)**.

Likewise, Figure 11.3.3 shows orthosteric binding sites for GABA and glycine, as well as the binding of modulators that can bind in the TMD or, in the case of benzodiazepines in the ECD as well.

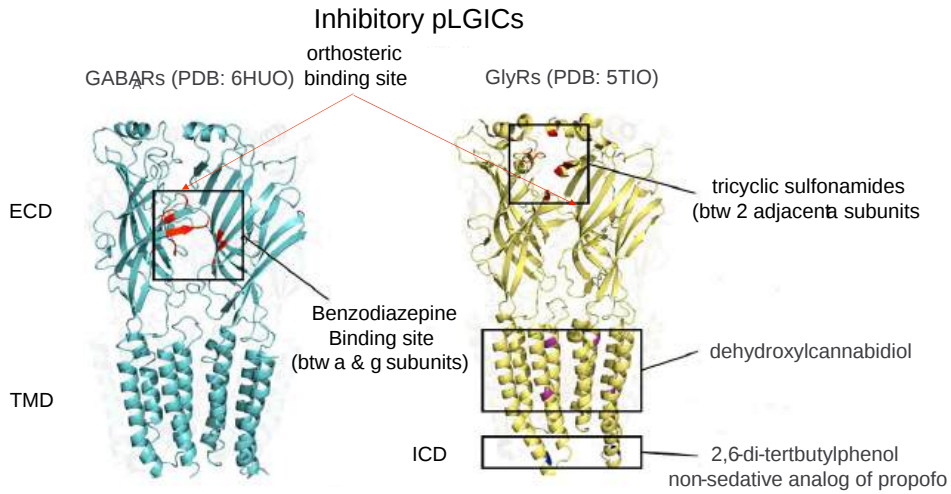


Figure 11.3.3: Orthosteric binding sites for GABA and glycine, adapted from Lara César O., Burgos Carlos F., Moraga-Cid Gustavo, Carrasco Mónica A., Yévenes Gonzalo E. Pentameric Ligand-Gated Ion Channels as Pharmacological Targets Against Chronic Pain, Frontiers in Pharmacology, 11 (2020), 167. https://www.frontiersin.org/article/...har.2020.00167, **Creative Commons Attribution License (CC BY)**.

How do inhalational aesthetics work?

The actual mechanisms of how anesthetics work are still unclear. These fascinating molecules can alter function in a variety of organisms including bacteria, yeast, worms, flies, and plants, as well as animals. Of course, their effect on consciousness appears to be found only in animals. Their selective "turning off" of a function (consciousness) of an entire organ (the brain) is stunning! One theory suggests that they exert their effects through bulk changes in the lipid bilayer, as the potency of anesthetics is generally related to their hydrophobicity. Most are nonpolar and have long been known to work on membranes, presumably altering ion flow through neural membrane ion channels. Typical inhalational and intravenous anesthetics are shown in Figure 11.3.3, along with their date of first use.

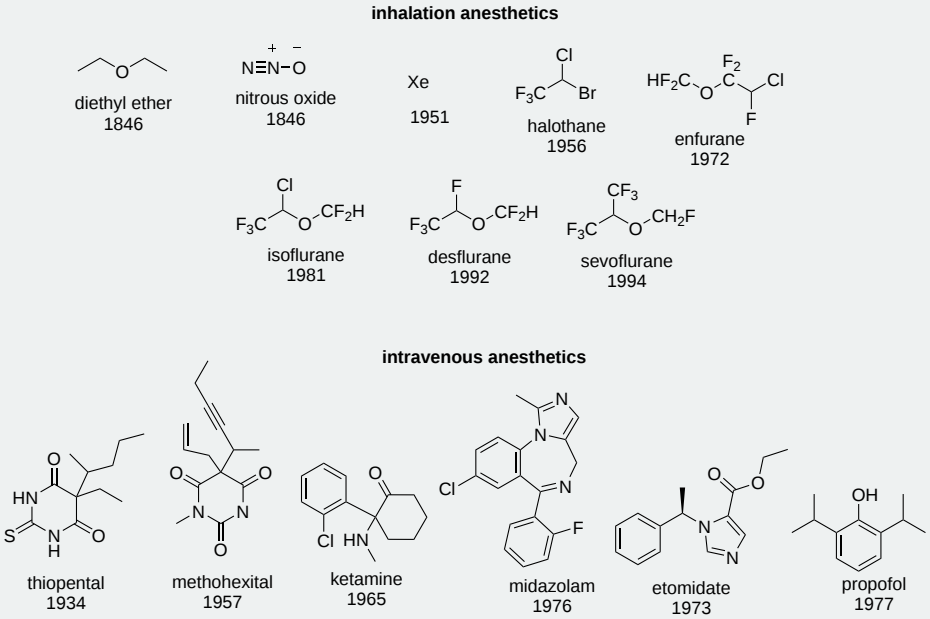


Figure 11.3.3: Structure of common inhalational and intravenous anesthetics. Adapted from Eur J Anaesthesiol. 2009 Oct; 26(10): 807–820. doi: 10.1097/EJA.0b013e32832d6b0f

A very robust correlation is found between the minimum alveolar concentration (MAC, in atmospheres) of inhalational anesthetics and their partition coefficient into olive oil (a measure of their hydrophobicity). This is illustrated in the Meyer-

Overton plot shown in Figure 11.3.3

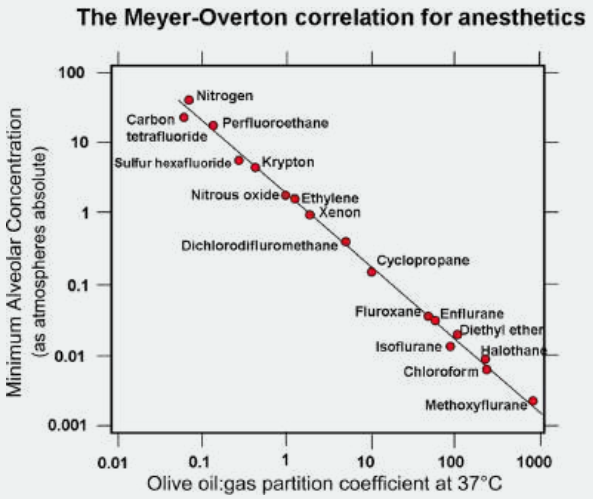


Figure 11.3.3 The Meyer-Overton plot of minimum alveolar concentrations (in atmospheres) vs partition coefficient into olive oil for inhalation anesthetics. https://commons.wikimedia.org/wiki/F...orrelation.png

The MAC is the concentration of inhaled anesthetic within the alveoli at which 50% of people do not move in response to a surgical stimulus (i.e. it is much like an IC_{50} for receptor inhibition). What's so fascinating is the range of molecular species, including N_2 and most Noble gases, that can act as anesthetics. Consider N_2, a nonpolar and nonreactive molecule, which comprises 80% of the atmosphere. From the graph, it is evident that it takes high pressure for it to exert its anesthetic properties. That can occur when scuba diving using regular air in tanks. Divers can experience nitrogen narcosis (also called depth intoxication or rapture of the deep) when using just compressed air. Hence a mixture of 21% oxygen, 35% helium, and 44% nitrogen is often used. Nitric oxide is used by dentists to alter consciousness and pain perception but does not cause general anesthesia except in some who can be hypersensitive to its use.

Additional studies suggest protein:anesthetic interactions are important. For example, the activity of the water-soluble protein luciferase is affected by them. Some molecules (like dichlorohexafluorocyclobutane), expected to have anesthetic properties based on their hydrophobicity, don't. The (S) enantiomer of isoflurane is 50% more potent than the (R) enantiomer in rats, which is hard to explain based on nonspecific partitioning into a bilayer.

Most modern theories suggest that they more directly affect specific target proteins and their proximal interacting lipids in neuromembrane bilayers. The main targets of anesthetics appear to be pLGICs. Anesthetics reduce neuron excitability and firing. Hence you could hypothesize that they inhibit excitatory pLGICs (such as the 5HT and acetylcholine receptors) and/or activate inhibitory ones such as the GABA and glycine receptors. pLGICs are pharmacological targets of many general anesthetics. However, anesthetic inhibition of certain GABA channels and potentiation of nicotinic acetylcholine channels have also been observed.

Recent elegant studies have shown that the inhaled anesthetic chloroform and isoflurane affect K^+ ion flow through the potassium channel subfamily K member 2, known also as the outward rectifying potassium channel protein TREK-1. The channel converts between a voltage-insensitive potassium leak channel and a phosphorylated voltage-dependent outward rectifying potassium channel. It doesn't affect the channel protein directly but indirectly through alteration in the local membrane which affects the location of phospholipase D2 (PLD2), a protein anchored to the membrane by covalent attachment of palmitic acid. PLD2 hydrolyzes phosphatidylcholine, with a positively charged choline head group, to choline and phosphatidic acid, with a negatively charged phosphate head group

The effect of these general anesthetics appears to be on lipid rafts in neural membranes. Lipid rafts are enriched in cholesterol and saturated lipids, especially sphingomyelins such as monosialotetrahexosylganglioside1 [GM1]. Rafts are especially important in the brain where cholesterol can reach up to 45% of plasma membrane lipids. Typical rafts are about 100 nm in diameter. In the presence of hydrophobic anesthetics, the rafts become larger and more dispersed, as the anesthetic partitions into them.

Membrane proteins also partition into rafts. One such protein is phospholipase D2 (PLD2), which is targeted to the inner leaflet rafts by post-translation palmitoylation. In the presence of general anesthetics, the PLD1 laterally translates away from the disrupted and enlarged lipid raft and binds to a disordered C-terminal region of the TREK-1 protein. This localizes the PLD2 and helps activate it to produce high local concentrations of phosphatidic acid with its negatively charged head group. That group interacts with a positive region of the TREK-1 protein, inducing a conformational change in the channel which opens it to K^+ efflux. In effect, PLD2 activates TREK-1 through the local formation of phosphatidic acids.

The opening of the channel hyperpolarizes the cell membrane (making the inside more negative, and inhibiting neural activity (a hallmark of anesthesia). These concerted actions are shown in Figure 11.3.a below.

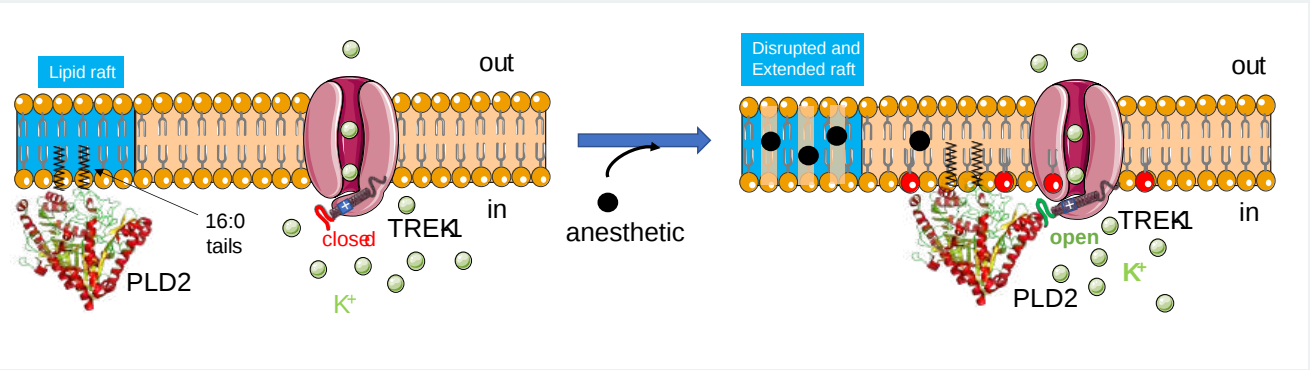


Figure 11.3.a: Effects of general anesthetics in the TREK-1 in the channel. After https://www.pnas.org/content/117/24/13757. r Creative Commons Attribution-NonCommercialNoDerivatives License 4.0 (CC BY-NC-ND)

Deletion of the TREK-1 gene decreases the effect of the anesthetics. If a catalytically inactive mutant of PLD2 (K758R) is overexpressed, all effects of chloroform were eliminated.

We will now explore two pLGIC, the eukaryotic nicotinic acetylcholine channel, and a prokaryotic analog, GLIC.

11.3.2.1: Nicotinic acetylcholine channel (6cnj)

One very interesting channel is the one involved in nicotine addiction. It binds nicotine (an exogenous alkaloid) and the normal endogenous neurotransmitter, acetylcholine. Both compete for the same orthosteric binding site. Since the binding of nicotine gates open the channel, nicotine acts as an agonist. The similarities in their structures are illustrated in Figure 11.3.4

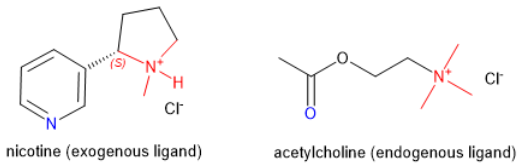


Figure 11.3.4: Structures of nicotine and acetylcholine

The membrane protein is a ligand (acetylcholine)-gated (open-close) positive ion (Na^+ or K^+) channel, involved in fast neural communication (such as at the neuromuscular junction). The quaternary structure of the pentameric receptor consists of two α4 and three β2 subunits - $(\alpha4)_2(\beta2)_3$. This isoform is the most abundant in the human brain and the one involved in nicotine addiction.

The iCn3D model (6CNJ) below has two bound nicotines (spacefill) in the extracellular domain and one Na^+ ion (spacefill) in the transmembrane domain containing the pore. The Na^+ or K^+ ions flow across the membrane down a concentration gradient in a thermodynamically favored process.

Figure 11.3.5 shows an interactive iCn3D model of the nicotinic acetylcholine channel with two bound nicotines (spacefill) in the extracellular domain and one Na^+ ion (spacefill) in the transmembrane domain containing the pore (6CNJ). (long load time)

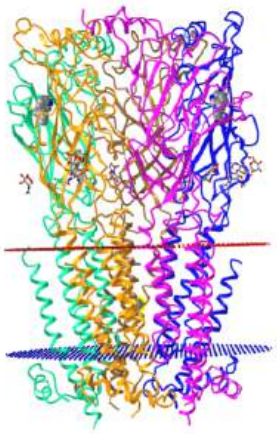


Figure 11.3.5: Nicotinic acetylcholine channel with two bound nicotines (spacefill) in the extracellular domain and one Na^+ ion (spacefill) in the transmembrane domain containing the pore (6CNJ). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...rbXmpaha6FYNB8

The Na^+ or K^+ ions flow across the membrane down a concentration gradient in a thermodynamically favored process.

The gold, blue, and brown β2 subunits are glycosylated on Asn 143 and are shown with a Man (β4) GlcNAc (β4) GlcNAc N-linked oligosaccharide. Nicotine is bound between two alpha-beta interfaces. One is shown between the green (alpha) and gold (beta) subunits and the other is between the magenta (alpha) and blue (beta) subunits.

11.3.2.2: GLIC: A prokaryotic pLGIC

This protein is a **proton**-gated cation channel with specificity for both Na^+ and K^+, which diffuse down their electrochemical gradients. In a sense, H^+s in the extracellular side act as "ligands" as the channel is opened with increasing H^+ concentration (decreasing pH) on the outside of prokaryotic cells. The protein is homologous to eukaryotic pLGICs. Structures of the protein from Gloeobacter violaceus bound to **propofol**, an anesthetic, are known. GLICs also interact with ethanol and barbiturates as well. Hence they serve as models to elucidate the binding and effects of anesthetics.

In contrast to eukaryotic pLGICs, the "ligand - H^+" does not bind in the orthosteric site in the extracellular domain occupied by traditional ligands. Rather changes in protonation states of key proton acceptors and donors in the protein lead to conformational changes analogous to those found on binding ligands to orthosteric sites on classical pLGICs. The external pH associated with half-maximal inward current, pH_{50} is approximately 5.1 ± 0.2.

Evidence suggests that when pH is lowered from 7 to 4, Glu 35 (distant from the orthosteric site), with a pKa - 5.8, becomes protonated. It connects through other H^+ acceptors and donors in the open form through a hydrogen bond network. These include two triads of amino acids found at the interface between the extracellular (ECD) and transmembrane (TMD) domains. R192-D122-D32 comprise a conserved "electrostatic triad". The second is Y197-Y119-K248. The network allows bridging of the effects starting with Glu 35 in the ECD into the transmembrane region where allosteric effectors usually interact with the protein.

Figure 11.3.6 shows an interactive iCn3D model of the open form of GLIC (3P50) with bound propofol (long load time)

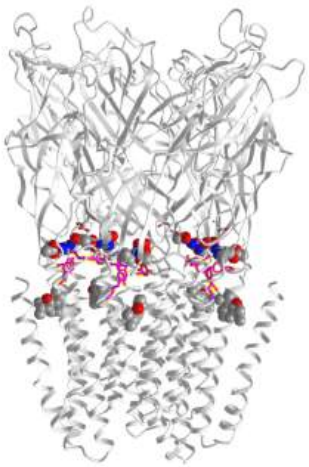


Figure 11.3.6: The open form of the prokaryotic pentameric GLIC ion channel with bound propofol (3P50). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...v92cdeunVY65R8 (long load time)

Orient the iCn3D model below with the extracellular domain (mostly beta structure) at the top and the transmembrane domain (alpha-helical) at the bottom. Key molecular players involved in the interactions described above, from the top down:

- Glu 35 (stick, color CPK)
- R192-D122-D32 electrostatic triad (sphere, CPK color)
- Y197-Y119-K248 triad (stick, color magenta)
- propofol (sphere, color CPK)

Propofol and another anesthetic, desflurane, bind at the same site localized in the upper part of the transmembrane domain of each of the five subunits.

The model below shows the mostly nonpolar (induced dipole-induced dipole) interactions between one bound propofol and side chains in the TMD. Also shown is an interaction between phosphatidylcholine and propofol.

Figure 11.3.7 shows an interactive iCn3D model showing the mostly nonpolar (induced dipole-induced dipole) interactions between one bound propofol and side chains in the TMD. Also shown is an interaction between phosphatidylcholine and propofol. (long load time)

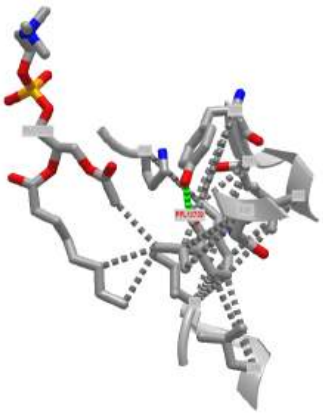


Figure 11.3.7: Propofol interactions in the transmembrane domain of the pentameric GLIC ion channel (3P50). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...s8dJnmX7SFKjLA

11.3.3: Voltage-Gated Ion Channels

In contrast to pentameric ligand-gated ion channels, which require 5 monomeric subunits to aggregate into a quaternary structure to form a pentameric pore, **voltage-gated ion channels** can form a channel from an aggregate of monomeric proteins, each containing a single 6 transmembrane helical unit, or from a longer polypeptide containing multiple repeating 6 transmembrane helical units.

Figure 11.3.8 shows a cartoon of a common voltage-gated K^+ channel. The monomer (top), denoted as the α subunit, contains a transmembrane domain containing 6 helix segments. Four of these monomers aggregate to form the actual homo- or heterotetrameric channel (bottom).

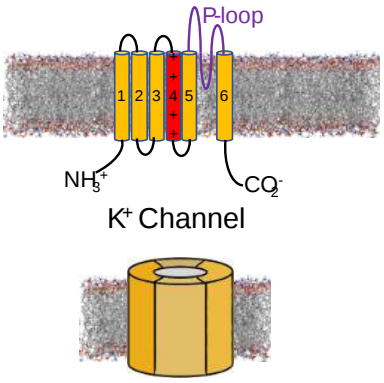


Figure 11.3.8: Cartoon of a common voltage-gated K^+ channel

The genes for the K_v channel family, which facilitate K^+ diffusion across the membrane, encode α subunits of approximately 500 amino acids and a molecular weight of about 57,000. **Four of** these α subunits come together in the membrane to form the

functional channel, a tetramer of α subunits, which together make one central pore. The α subunit can form homo- or heterotetramers since there are different α subunit encoding genes. In addition, the functional channel has smaller, regulatory β subunits as well.

A cartoon structure of a typical voltage-gated Na^+ channel is shown in Figure 11.3.9. It is a **single** polypeptide (α) chain that contains 4 sequential repeats of the 6 transmembrane helical segments (I-IV). The functional channel (bottom) has just one polypeptide chain.

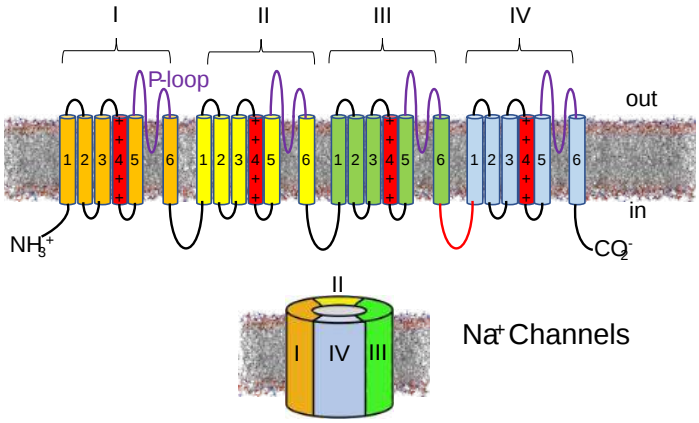


Figure 11.3.9: Cartoon structure of a typical voltage-gated Na^+ channel adapted from https://www.anaesthesiajournal.co.uk...264-4/fulltext

The Na^+ channel has a molecular weight of around 229K and about 2000 amino acids (each 4x that of the K^+ channel α subunit). It is glycosylated and subjected to multiple post-translational modifications. Usually, the protein in the central nervous system is a complex of the α subunit, and small additional regulatory β subunits, which modify the kinetics and voltage-dependency of the α subunit channel.

Segment (helix) 4 of each of the four repeat units illustrated above is the conserved "voltage" sensor. It contains multiple, charged amino acids whose disposition changes with changes in the transmembrane potential, allowing conformational changes in the protein and gating of ion flow. Each of the 4 repeating units above also contains an extracellular P-loop (colored purple in segment I in purple) connecting helix 5 and helix 6.

We will focus on K^+ channels with some additional information on the Na^+ channel below.

K^+ Permeation through Kv1.2 Channel

Voltage-dependent potassium channels (Kv) have 4 subunits and can be homo- or heterotetramers. They allow the voltage-gated flow of potassium ions through the membrane. Several obvious questions should arise. How can they be selective for K^+ ions? That is, how can they allow the larger K^+ ion to passively flow through and not the smaller Na^+ ions? Secondly, how can a change in the transmembrane potential cause the channel to open or close? That question boils down to how a change in transmembrane potential can change the conformation of proteins. We'll show several iCn3D models of this protein.

Figure 11.3.10 shows a simplified view of the rat Kv1.2 channel (3lut) from top and side views (without parts of the cytoplasmic domain), showing each of the 4 identical monomeric subunits in a different color. Each monomer has S1-S6 transmembrane segments.

Shaker Potassium Channel Kv1.2

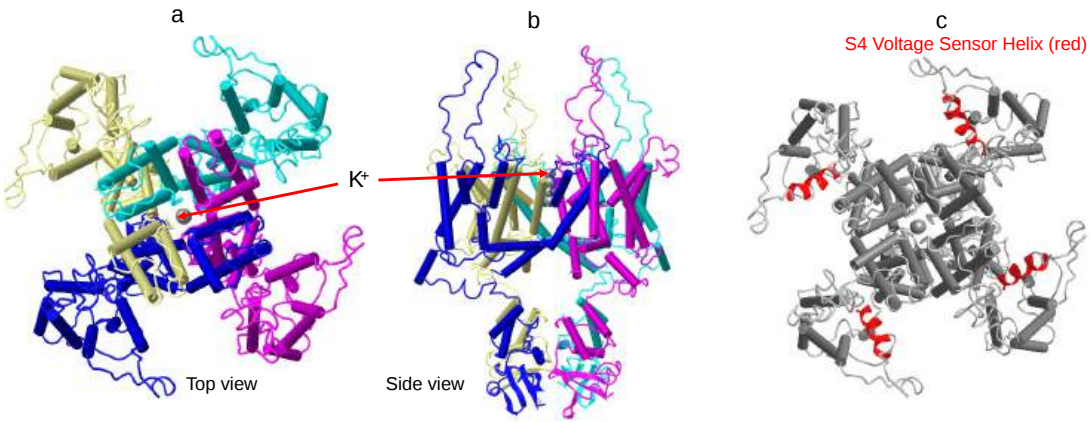


Figure 11.3.10: Rat Kv1.2 channel (3lut) from top and side views (without parts of the cytoplasmic domain), showing each of the 4 identical monomeric subunits in a different color.

The protein is found in the brain and central neuron plasma membranes and also in the cardiovascular system. K^+ ions diffuse through the center from the extracellular to intracellular side down a concentration gradient but potentially against the transmembrane potential. The four monomers in the homotetramer form one central pore.

Figure 11.3.11shows an interactive iCn3D model showing the rat Kv1.2 channel (3lut) described in Figure 11.3.10

Figure 11.3.11: Rat Kv1.2 channel (3lut) . (Copyright; author via source). Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...99VbNxVzCZyr9A

Figure 11.3.12shows an interactive iCn3D model showing the S4 voltage sensor helix in each monomer of the Kv1.2 potassium channel (3lut) .

Figure 11.3.12: S4 voltage sensor helix in each monomer of the Kv1.2 potassium channel (3lut) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...nZnNTpqSwA6Ab9

Figure 11.3.13shows a more detailed structure of the Kv1.2 potassium channel (3lut). The four monomers, which pack together to form the tetramer, are shown in light and dark grey to allow key residues highlighted in color to stand out. The details are explained below.

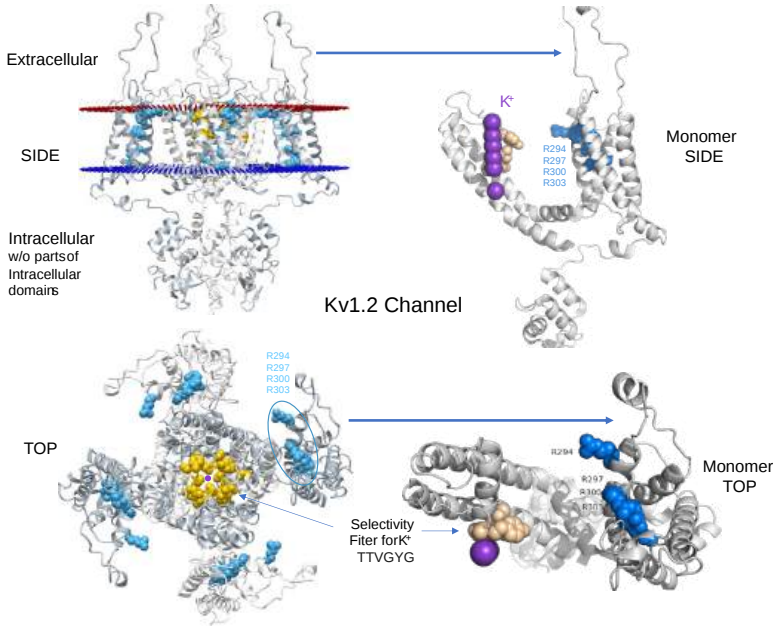


Figure 11.3.13: Detailed structure of the Kv1.2 potassium channel (3lut)

The left images show the channel from the side (top left view) and top (bottom left view). The right images show just one of the monomers with different side chains highlighted.

Figure 11.3.14shows an interactive iCn3D model detailing key residues in the workings of the Kv1.2 potassium channel (3lut)

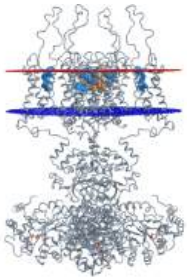


Figure 11.3.14: Key residues in the workings of the Kv1.2 potassium channel (3lut) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...wYWMWmHWp1UP6A

K^+ selectivity - All potassium ion channels, even if not voltage-gated, solve the selectivity dilemma in a similar way. All have in the narrowest part of the pore in the center of the channel this consensus sequence - Thr-Thr-Val-Gly-Try-Gly (TTVGYG) - which is found in the P-loop. These are shown in gold and brown colors in the figure above. The -OHs in the selectivity filter can interact with a dehydrated K ion but not with a dehydrated Na ion, which can not approach close enough to form significant interactions. Surrounding the filter are twelve aromatic amino acids which constrain the size of the pore opening. The interactions of the filter O's with the K ion make up for the energetically disfavored dehydration of the ion. The filter contains K^+ ions which repel each other, assisting in the vectorial discharge of the ions through the membrane. These ions must form weak interactions with the selectivity filter. The actual pore is mostly hydrophobic, which facilitates ion flow through the membrane.

Figure 11.3.15 below shows a closeup of the selectivity filter. Four Thr 374s (second Thr in the selectivity filter sequence of TTVGYG) from the four different monomers in the channel are clearly shown interacting with the top K^+ ion (gray sphere).

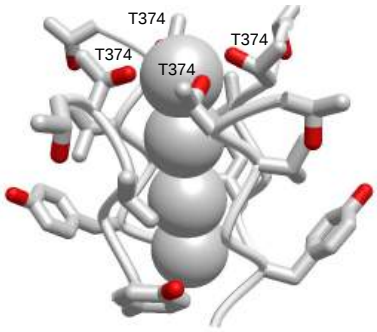


Figure 11.3.15: K^+ Selectivity filter of the Kv1.2 potassium channel (3lut)

Figure 11.3.16shows an interactive iCn3D model detailing key residues in the K^+ selectivity filter of the Kv1.2 potassium channel (3lut). Hover over the residues to identify them.

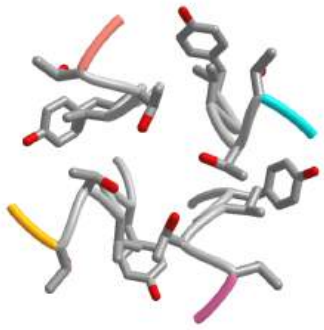


Figure 11.3.16: Key residues in residues in the K^+ selectivity filter of the Kv1.2 potassium channel (3lut) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...FFKJVVHtbpuYj6

Different voltage-gated ion channels alter ion selectivity through changes in these amino acids in the P-loop, as illustrated in Table 11.3.2below. As channels lose specificity of K^+, they gain specificity for Na^+ and Ca^{2+}. Red highlights denote conserved residues and yellow residues that are chemically similar.

sequence	specificity
TVGYG	strong K^+ channels
CIGYG	weak K^+, HCN channels
TVGDG	TRP channels
STFEG	ionotropic glutamate receptors
LCGEW	strong Ca^{2+} voltage-gated channels

Table 11.3.2 P-loop specificity side chains in voltage-gated ion channels

Voltage gating -

Helix S4 in each monomer in the transmembrane domain of the complex is the voltage sensor. The sequence of this helix is LAILRVIRLVRVFRIFKLSRH. Note the arginines and lysine highlighted in blue. They repeat every 3 amino acids. The voltage-sensor domain must be shielded from the nonpolar acyl of the bilayer. Four conserved Arg residues on S4, part of the voltage-

sensor domains, are shielded from the lipids and coupled to an amphiphilic helix running parallel to the plane of the membrane. The arginines move under the influences of forces arising from changes in the membrane's electric field initiated by ion movement through other ion channels in the membrane. Mechanical work is done by the electric field on the voltage sensor as the charged Arg residues are moved through the electric field. The movement is coupled through the amphiphilic helix to the pore which changes conformation. In turn, the S4 and coupled S5 helices of the voltage sensor do mechanical work on the pore by altering its conformation to open/close the pore, specifically at the activation gate of the pore. This seems quite similar to how iron movement into the heme plane in hemoglobin on oxygenation pulls the proximal His on the F8 helix which then transmits a conformation change to other helices in the subunit, leading to cooperative conformational changes in the tightly packed protein. About 12 charges move across the transmembrane potential field.

Channels, once open, must be inactivated. In the case of the voltage-gated potassium channel, inactivation occurs when the amino-terminal cytoplasmic domain binds to the potassium pore on the cytoplasmic side, in interaction likened to the binding of a "ball on a chain" (the cytoplasmic domain) to the pore opening. The chain acts to tether the ball domain so it may swing to bind to the pore opening. The ball domain contains both positively charged and hydrophobic regions. Where is the ball domain in the absence of inhibition? Recent studies (Oliver et al.) have shown that a positive domain can bind to proximal phosphatidyinositol 4,5-bisphosphate (PIP2) lipids on the inner leaflet of the membrane bilayer. When so bound, inactivation of the channel is prevented. As you will see in the next section, PIP2 can also be cleaved to form diacylglycerol and inositol 1,4,5-trisphosphate when cells are activated by external factors (hormones, growth factors, etc) in the process of signal transduction.

Figure 11.3.17 shows a molecular dynamics simulation showing K^+ interaction with the channel lining and the "knock-on" mechanism showing how an incoming K^+ ion can repel a K^+ ion in the pore through the channel.

Figure 11.3.17: Molecular dynamics simulation of K^+ movement through a channel

---

Recent Updates: June 2023 - Sodium Channels

Voltage-gated sodium channels (Na_V)

The eukaryotic voltage-gated sodium channels (NaV) allow inward movement of Na^+ ions which depolarizes a neuron and sets off an action potential in nerve and muscle cells. (For more information on neuron signaling, see Chapter 28.9). The NaV has an alpha subunit that forms the pore plus beta subunits that associate with it and modulate its activity. Nine eukaryotic isoforms exist. NaV has 4 domains, I-IV, each containing segments 1-6. Each of the S1-S4 segments forms a voltage-sensitive domain and each of the S5 and S6s form the pore. In contrast to eukaryotic NaVs which have a single chain, bacterial NaVs contain 4 identical subunits. Post-translation modification of the alpha subunit can regulate its activity.

As important as it is to initiate Na^+ influx to trigger neuron firing, it is equally important to turn it off to control neural signaling. The inward Na^+ current is stopped by a fast inactivation that occurs within a few milliseconds.

Many neurotoxins bind to the NaV and regulate its activity. Key examples are the α-scorpion and sea anemone toxins which both inhibit the fast inactivation of the NaV leading to prolonged or sequential action potentials. These toxins binds to the voltage sensor in domain IV which is key for the fast inactivation in the absence of the toxin.

The S4 helical segments in each domain are the key voltage sensors. Each of the S4 segments has 4-6 positively charged Arg and Lys side chains. On depolarization of the cell (when the inside becomes less negative and more positive), this helix move "up" from the cytoplasmic side (which is increasingly more positive) which opens voltage-gated channel. Specificity for the small Na^+ ion (over the K^+ and Ca^{2+} ions) is determined mainly by 4 amino acids, DEKA (Asp-400, Glu-755, Lys-1237, and Ala-1529) found in P loops of domain I-IV, respectively. This selectivity filter is conserved.

Figure 11.3.18below shows a cartoon of the NaV with two associated, regulatory beta chains.

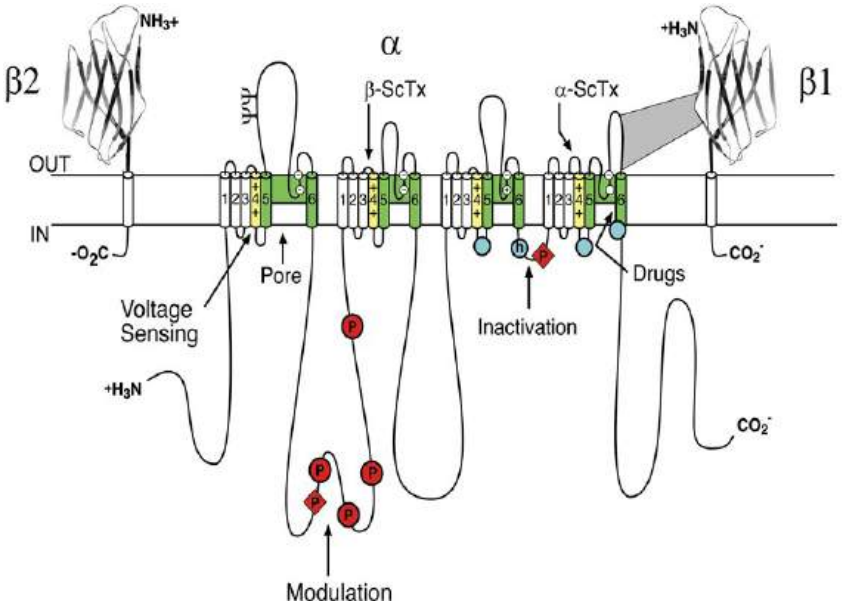


Figure 11.3.18 Cartoon of eukaryotic voltage-gated sodium channel (Nav) with two associated, regulatory beta chains. https://www.guidetopharmacology.org/...rd?familyId=82. **CC BY-SA 4.0**

Note the sites for posttranslational modification by phosphorylation and drug interactions. Cylinders represent probable α-helical segments S1-S6. Bold lines represent the polypeptide chains of the selectivity filter and tetrodotoxin binding site; The yellow S4 segments are the voltage sensors. The "h" in the blue circle is in the inactivation gate loop. Blue circles are sites implicated in forming the inactivation gate receptor. Sites of binding of α- and β-scorpion toxins (ScTX) and a site of interaction between α and β1 subunits are also shown. Tetrodotoxin is a specific blocker of the pore of sodium channels, whereas the α- and β-scorpion toxins block fast inactivation and enhance activation, respectively, and thereby generate persistent sodium current that causes depolarization block of nerve conduction.

Table 11.3.3shows the different types of neurotoxin receptors sites found in NaVs

Neurotoxin Receptor Site #	Toxin or Drug	Domains
1	Tetrodotoxin	IS2–S6, IIS2–S6
	Saxitoxin	IIIS2–S6, IVS2–S6
	μ-Conotoxin	
2	Veratridine	IS6, IVS6
	Batrachotoxin	
	Grayanotoxin	
3	α-Scorpion toxins	IS5–IS6, IVS3–S4
	Sea anemone toxins	IVS5–S6

4	β-Scorpion toxins	IIS1–S2, IIS3–S4
5	Brevetoxins	IS6, IVS5
	Ciguatoxins	
6	δ-Conotoxins	IVS3–S4
local anesthetic drug sites	Local anesthetic drugs	IS6, IIIS6, IVS6
	Antiarrhythmic drugs	
	Antiepileptic drugs	

Table 11.3.3 Neurotoxin receptor sites in NaVs. https://www.guidetopharmacology.org/...rd?familyId=82. **CC BY-SA 4.0**

A view of the effective pore in the bacterial Nav is shown below in Figure 11.3.19

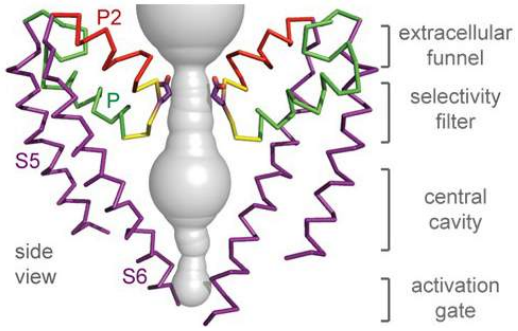


Figure 11.3.19 **Structure of the bacterial sodium channel NavAb pore B.** Architecture of the NavAb pore. Glu177 side-chains in the P loop are shown in purple. The pore volume is shown in grey. The P and P2 alpha helices that form the scaffold for the selectivity filter and outer vestibule are shown in green and red, respectively. https://www.guidetopharmacology.org/...rd?familyId=82. **CC BY-SA 4.0**

As mentioned above, the bacteria NaVs have 4 monomeric subunits. In contrast to the K^+ channel, which requires the K^+ ions to be dehydrated to make sufficient interactions with the pore and to pass through, the Na^+ ions need to be hydrated. Figure 11.3.20 shows an interactive iCn3D model of the A and B chains of the bacterial NaV voltage-gated sodium channel pore and C-terminal domain (5BZB)

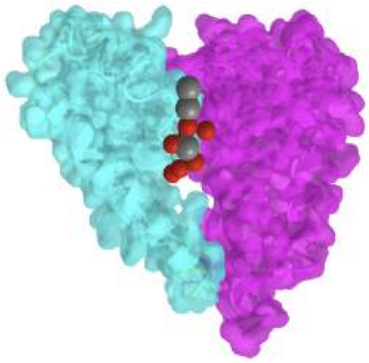


Figure 11.3.20 A and B chains of bacterial NavMs voltage-gated sodium channel pore and C-terminal domain (5BZB) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...vRmDqmDHGsYLh8

One subunit is shown in a transparent cyan surface and the second one is shown in magenta. Three Na^+ ions are shown (gray spheres). Eight water molecules are shown interacting with them in the pore. The other two subunits are not shown for clarity.

The voltage-gated sodium channel has three major conformational states:

- a basal **closed** state found at resting cell potentials in which the pore of Na^+ is occluded by an activation gate
- an **open** state found when a depolarizing potential is reached in the cell

- an **inactivated** state formed within 10 ms of opening of the channel when the inactivation gate with a a Ile-Phe-Met (IFM) sequence motif, found in the intracellular linker between domain III and IV (near the cytoplasmic face of the receptor) closes off the pore to further Na^+ entry.

The protein converts back to the closed state when the transmembrane potential is restored to its initial value (around -70 mV) and the positively charged S4 segments move back towards the cytoplasmic face.

A cartoon of the three-state model for the Na^+ channels (as other voltgage-gates ion channels in general) is shown in Figure 11.3.21below.

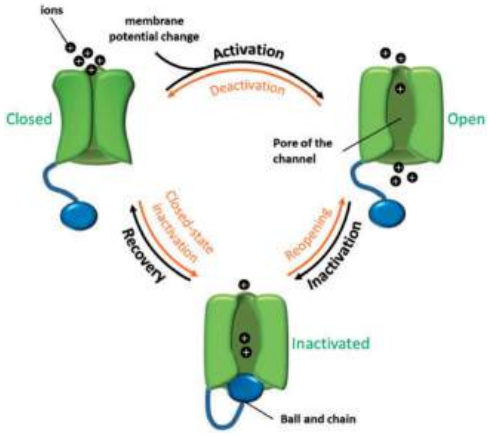


Figure 11.3.21: Three-state model for the voltage-gated ion channels. Hinard, Valerie & Britan, A & Rougier, Jean-Sébastien & Bairoch, Amos & Abriel, Hugues & Gaudet, Pascale. (2016). ICEPO: The ion channel electrophysiology ontology. Database : the journal of biological databases and curation. 2016. 10.1093/database/baw017. Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/)

The YouTube video by Pete Meighan below shows an incredibly clear description of the three conformational states of the channel and the conversion from closed to open to inactivated states.

To understand these conformations states, we need to look a the structure of the protein in greater detail. Figure 11.3.22below shows multiple representations of the voltage-gated sodium channel. The pore domain PD is formed from helix segments 4 and 5 on each of the IV domains of the protein. The voltage sensor domain is formed from the S1-S4 segments, of which segment S4, containing multiple positively-charged Arg and Lys, is key.

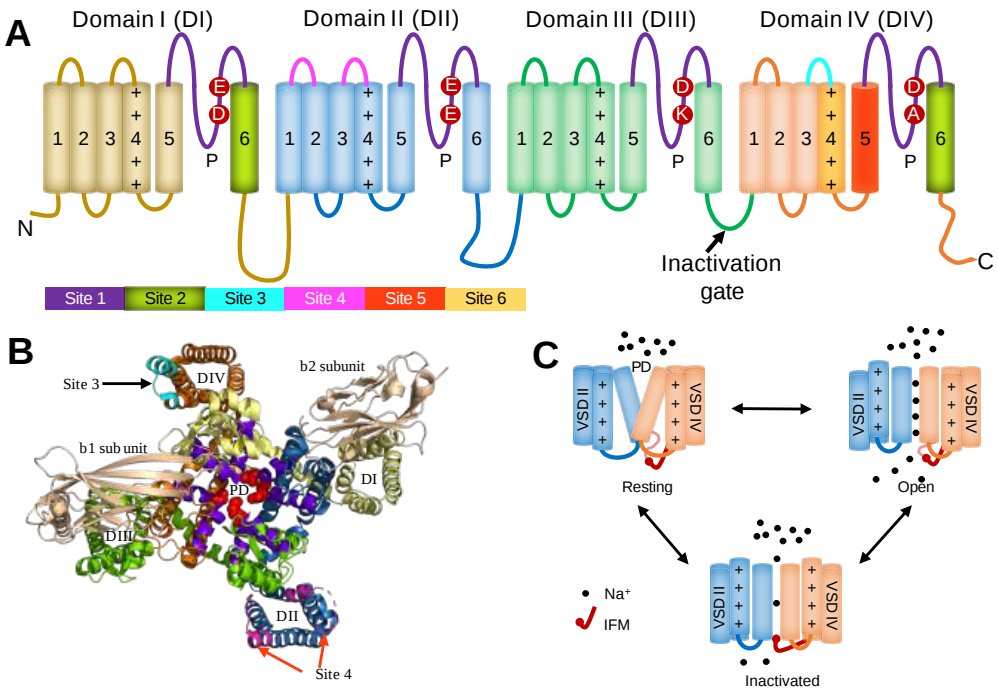


Figure 11.3.22 Structure features of the voltage-gated sodium channel. Dongol, Y.; C. Cardoso, F.; Lewis, R.J. Spider Knottin Pharmacology at Voltage-Gated Sodium Channels and Their Potential to Modulate Pain Pathways. *Toxins* **2019**, *11*, 626. https://doi.org/10.3390/toxins11110626. Creative Commons Attribution (CC BY) license (**http://creativecommons.org/licenses/by/4.0/**)

Panel (**A**) shows a schematic representation of the α-subunit of voltage-gated sodium (Na_V) channel. Four non-identical domains (DI–DIV) feature six neurotoxin receptor sites (Sites 1–6) and key residues contributing to the outer Na^+ ion selectivity filter (EEDD) and inner selectivity filter (DEKA). The connecting S5–S6 linker is called P-loop (P) which together with S5 and S6 segments from each domain contributes to the formation of the Na^+ ion selective channel pore. Sites 1-6 (colored purple, green, cyan, magenta, etc) are sites where inhibitors such as toxins bind.

Panel (**B**) shows the structure of the $Na_V1.7$ channel (PDB 6J8G). Four voltage sensing domains (VSDs), DI (yellow), DII (blue), DIII (green), and DIV (orange), are shown with their corresponding pore-forming segments (S5 and S6) arranged to form the pore domain (PD) selective to Na^+ ions. The P-loop that contributes to forming the inner selectivity filter is colored in red spheres (DEKA) and outer selectivity filter (EEDD) is colored in purple. The S6 segments of all the four domains contribute to form the intracellular region of the pore. Site 3 (cyan) and Site 4 (pink) are the major binding sites for spider knottins (neurotoxins). The β1 and β2 subunits which interact with DIII and DI, respectively, are highlighted in beige color.

Panel (**C**) shows a schematic of the three main conformational states of the protein which control gating of Na_V channels. At polarized potentials, the DI–DIV S4 segments are drawn towards the intracellular side due to the positive gating charges to render the closed conformation (down state). Upon depolarization, the forces holding the down state are relieved and DI–DIII S4 segments are rapidly released extracellularly to open the S6 channel gate in the open conformation (up state). Note the movement of the S4 helix with its positive charged toward the extracellular side of the membrane. The DIV S4 moves up slowly compared to DI–DIII S4 and drives the fast inactivation, where the channel is occluded intracellularly by the Ile, Phe, and Met (IFM) motif. After cell repolarization, the channel returns to a closed (resting) state.

Figure 11.3.23 shows an interactive iCn3D model of a ternary complex of human Nav1.2 with the beta2 regulatory subunit and conotoxin IIIA (6J8E)

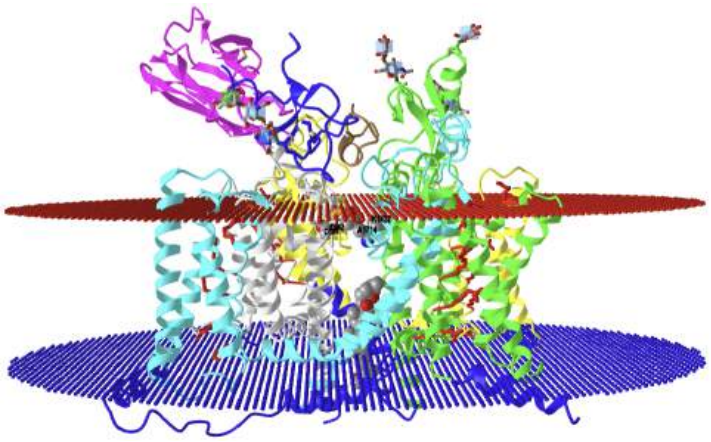


Figure 11.3.23 Human sodium channel Nav1.2-beta2-KIIIA ternary complex (6J8E) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?aYxomdUqZHqbMmE5A

Domains I-IV are shown in gray, yellow, green and cyan, respectively. The separate beta-2 regulatory subunit is shown in magenta. The positive side chains in each of the S4 segments of the four domains are shown as red sticks. The side chains of the inner 4 amino acids (DEKA) comprising the selectivity filter are shown in spacefill, CPK colors, and labeled. The brown peptide cartoon on the extracellular side (red sphere layer for outer leaflet) is the μ-conotoxin KIIIA. The spacefill molecule in the pore near the blue sphere layer representing the inner leaflet is the neurotoxin **veratridine** (VTD), which inhibits channel inactivation and lengthens the action potential with possibly fatal consequences. A single Na^+ ion is shown at the top of the pore as an orange sphere labeled Na. The μ-conotoxin blocks the pore. The Ile1488-Phe1489-Met1490 (IFM) motif, found in the intracellular linker between domain III and IV (near the cytoplasmic face of the receptor, responsible for the fast inactivation, are shown in gray spheres and labeled with single-letter codes.

The selectivity filter DEKA is different from the selectivity filter in another sodium channel (NavAb) which has 4 glutamate. Asp and Ala line the wall of the filter region and Glu and Lys can attract and release, respectively, the Na^+ ion. For fast inactivation, the IFM motif must interact with an"inactivation gate receptor" within the protein itself. Likely candidates for this are short intracelllar loops connecting all the S4 and S5 segments. This receptor appears to contain 3 amino acids, F1651, L1660 and N1662.

Figure 11.3.24 shows another interactive iCn3D model of a ternary complex of human Nav1.2 with the beta2 regulatory subunit and conotoxin IIIA (6J8E) highlighting just the DEKA selectivity filter, the IFM motif, and its receptor inactivation gate F1651, L1660 and N1662.

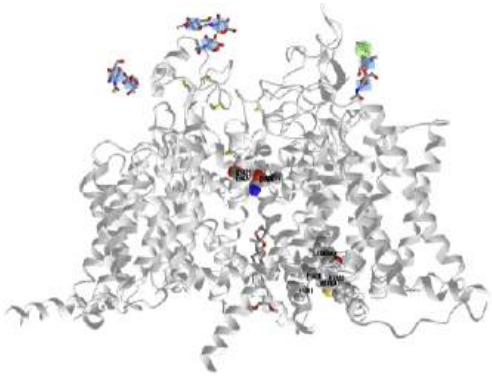


Figure 11.3.24 Human sodium channel Nav1.2-beta2-KIIIA ternary complex highlighting just the DEKA selectivity filter and the IFM and its receptor inactivation gate F1651, L1660 and N1662 (6J8E). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?tq3kXbT44GeCCUcY7

They are all shown in spacefill and labeled. Orient the molecule with the cytoplasmic domain at the top. The leaflets of the bilayer are omitted for clarity but the cytoplasmic pore inhibitor is still shown in sticks.

Figure 11.3.25below shows a closeup of the bound Na^+ ion.

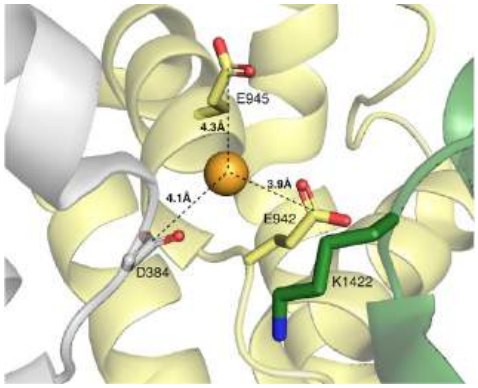


Figure 11.3.25 Sodium ion in the SF of the PDB ID 6J8E $Na_v1.2$ cryo-EM structure. The pore blocker μ-conotoxin KIIIA is not shown. Alberini et al. *J. Chem. Theory Comput.* 2023, 19, 10, 2953–2972. April 28, 2023. https://doi.org/10.1021/acs.jctc.2c00990. Creative Commons.

Structures of the closed state, the open state, and the inactivated state of NaV1.5 are now known. Key regions are shown in Figure 11.3.26below. The open state was the hardest to solve since it closes in milliseconds to form the inactive state as described above. Mutations in the Ile-Phe-Met (IFM) motif to QQQ prevented inactivation. This would ordinarily be deleterious to an organism but in the presence of a small molecule inhibitor, propafenone, it was possible to isolate this state.

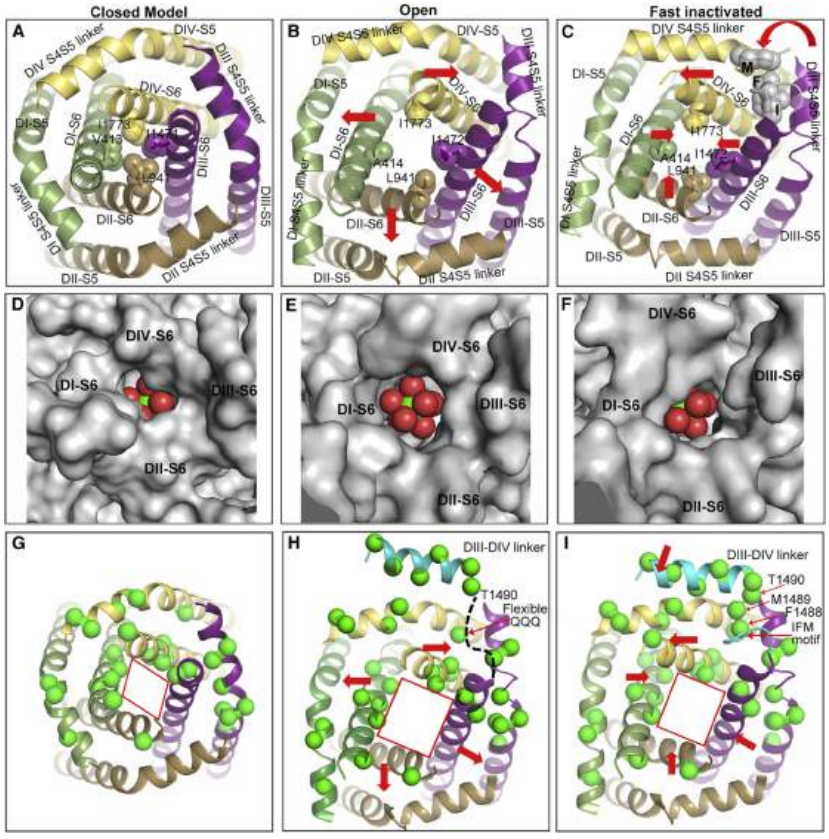


Figure 11.3.26 Closed, open, and inactivated conformations of the activation gate and the locations of arrhythmia mutations. Daohua Jiang, Richard Banh, Tamer M. Gamal El-Din, Lige Tonggu, Michael J. Lenaeus, Régis Pomès, Ning Zheng, William A. Catterall. Open-state structure and pore gating mechanism of the cardiac sodium channel. Cell, Volume 184, Issue 20, 2021, 5151-5162.e11, ISSN 0092-8674,. https://doi.org/10.1016/j.cell.2021.08.021. Reprinted with permission from Elsevier. May not be sublicensed, assigned, or transferred to any other person without publisher's written permission.

Panel (A) shows the closed activation gate of $Na_V1.5$ generated by MODELER based on the resting-state structure of Na_VAb (PDB: 6P6W), sealed by a square of hydrophobic side chains of hydrophobic residues V413, L941, I1471, and I1773 (spacefill, black) that in the closed state completely seal off the cytoplasmic opening in the pore. These ring of amino acid side chains come together on conformational changes resulting from the engagement of the IFM motif with its internal receptor.

Panel (B) shows the open activation gate of $Na_V1.5$/QQQ triple mutation. Red arrows indicate the directions of movement of S6 segments compared to the resting state.

Panel (C) shows the partially open but nonconductive activation gate of $rNa_V1.5_C$ in the inactivated state. Red arrows indicate the directions of movement of the S6 segments compared to the open state.

Panels (D–F) show structures from (A) to (C) are shown in space-filling surface representation with hydrated Na^+ placed in the central cavity behind the activation gate. Red and green spheres represent water and Na^+, respectively. van der Waals distances measured across the orifice of the activation gate are 4.3 Å (*D*I-*D*III) × 2.8 Å (*D*II-*D*IV) for the resting/closed state, 6.9 Å (*D*I-*D*III) × 5.0 Å (*D*II-*D*IV) for the inactivated state, and 7.3 Å (*D*I-*D*III) × 8.2 Å (*D*II-*D*IV) for the $Na_V1.5$/QQQ open-state structure.

Panels (G–I) show structures from (A) to (C) with the locations of arrhythmia mutations causing LQT-3 overlaid as green spheres

The spacefill models in D-E clearly show that the cytoplasmic opening is "open" in the open state E (obviously), and occluded to increasing degrees in the inactive state (F) and closed state D.

Finally, figure 11.3.27shows an interactive iCn3D model of the electrostatic surface potential of the rat $Na_v1.5$ channel (6UZ3)

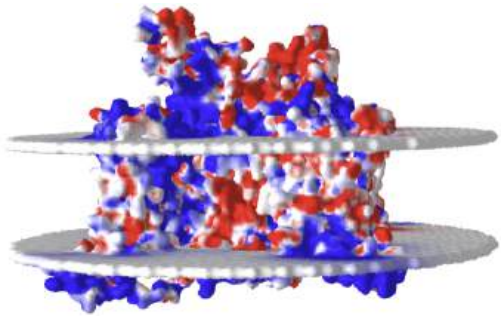


Figure 11.3.27: Electrostatic surface potential of the rat $Na_v1.5$ channel (6UZ3) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...UZdQeUUPtYGgR6

The red indicates negative potential and the blue positive potential. The white is neutral. Note that the bilayer representations are essentially neutral with respect to their effect on the overall surface potential. Tilt the model to view the top and bottom entrance to the pore and the overall negative charge density expected in the net attraction of the positive Na^+ ions.

11.3.4: Lipid-Gated Ion Channels

Membrane receptors are in the lipid bilayer, so it should not be too surprising that specific lipids might bind and trigger conformational changes in the receptor, mediating a specific biological activity. The specific lipid might form during an upstream event, then bind to the receptor, triggering function in a process of signal transduction, which we will explore in the next chapter. Let's look at receptors gated by binding of the lipid phosphatidylinositol bisphosphate (PIP_2). It is found in small concentrations in membranes and can be cleaved in by phospholipase C to form a small polar intracellular signaling molecule, IP_3 (discussed in Chapter 12). PIP_2 can also recruit proteins to the membrane and participate in signaling events that way. For this section, we will discuss how it binds to a K^+ ion channel protein and open so K^+ can flow into cells (the opposite direction of the normal efflux), which helps control transmembrane potential.

Kir2.2 -Inward rectifier potassium channel Kir2.2

The names of this channel can best be understood if you understand the meaning of the word rectify. The verb has several meanings but for us, the best definition is to **correct**. What does this channel correct? When is it functional? We'll explore the details more in the next chapter we first a better understanding of what happens to the actual Na^+ and K^+ ion concentrations outside and inside the cell during neural activation. How much do they change? We need this to understand the driving force for the Kir2.2 channels that move K^+ into the cell, which is opposite the usual direction. From that sense, it is "rectifying" the K^+ concentrations.

How much do the actual concentrations of Na^+ and K^+ ions change when a neuron is excited?

When a neural cell is activated or fires, the cell membrane potential goes from a resting potential of around -60 to -70 mV (inside negative) to a more positive potential as Na^+ ions enter the cell through voltage-gated Na^+ channels (after an early neurotransmitter-gated ion channel is open after ligand binding) with the ions flowing down a **chemical and electric potential** gradient. At a certain membrane potential (about +30 mV), K^+ channels open allowing an efflux of K^+, also down both a **chemical and electric potential** gradient, returning the cell potential close to its equilibrium value potential of around -60 to -70 mV (inside negative). But how much do the actual K^+ and Na^+ ion concentrations change in this process? The somewhat counterintuitive answer is hardly any at all!

We need to understand how the membrane acts as a capacitor first. The charge Q on a surface of a plate or side of a membrane is proportional to the voltage across the plate or membrane. Figure 11.3.28 shows how a membrane with a transmembrane potential acts as a capacitor. The dielectric medium in the capacitor will determine how quickly the charges on the plate will dissipate. When the medium is an insulator, resistant to charge flow, the plates remained charged longer. The same is true for the membrane. The hydrophobic bilayers act as an insulator and resist the discharge of the membrane potential. The bilayer offers high resistance (low conductance) to charge flow. Only when channels are open and the ions become reasonably permeable to flow does the membrane potential change over short periods. The following derivation is adapted.

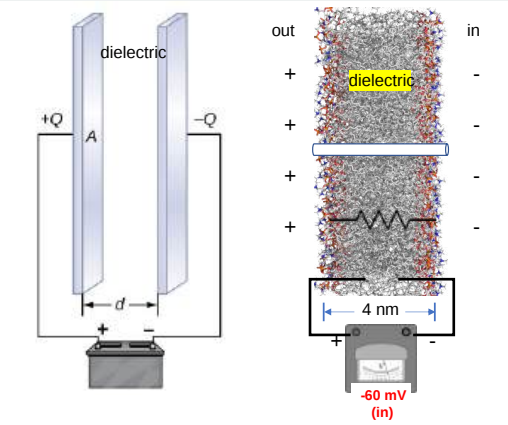


Figure 11.3.28: A membrane as a capacitor. After https://openpress.usask.ca/physics15...d-capacitance/. **License**: CC BY: Attribution

It makes sense that the stored charge (Q) on either side of the membrane is proportional to the membrane voltage. We can write the following simple equations:

$$\begin{aligned} Q &\propto V \\ Q &= \mathrm{C}V \end{aligned} \tag{11.3.1}$$

where C, the proportionality constant, is the capacitance with units of the Faraday (which you remember from introductory chemistry). Let's normalize this equation for an area of 1 cm^2. The measured capacitance of lipid bilayers is about 10^{-6}F/cm^2. Let's assume a voltage change from -70 mV to + 30 mV for a total of 0.1 V. Hence

$$Q = \frac{10^{-6}\ \mathrm{F}}{\mathrm{cm}^2}(0.1\ \mathrm{V}) = \frac{10^{-7}\mathrm{Coul}}{\mathrm{cm}^2} \tag{11.3.2}$$

Let's change that into the number of elementary charges on the surface of the membrane per μm^2 surface area.

$$\left(\frac{10^{-7}\mathrm{Coul}}{1\ \mathrm{cm}^2}\right) x \left(\frac{1\ \mathrm{ion}}{1.6\times10^{-19}\mathrm{Coul}}\right) = \frac{6.25\times10^{11}\ \mathrm{ions}}{\mathrm{cm}^2} \times \frac{1\ \mathrm{cm}^2}{10^8\,\mu\mathrm{m}^2} = \frac{6,250\ \mathrm{ions}}{\mu\mathrm{m}^2} \tag{11.3.3}$$

What does this mean for an ordinary eukaryotic cell? Let's model the cell as a sphere with a diameter of 10 µm.

Knowing the equations for the volume (V = (4/3)πr^3) and surface area ($4\pi r^2$), the volume of 10 µm cell is about 524 μm^3 and the surface area is 314 μm^2. The table below shows the resting ion concentrations in a cell and the actual number of ions in the 524 μm^3 volume of the cell (column 3).

Using the calculated value of 6250 ions moved/μm^2, the total number of K^+ and accordingly Na^+ ions that move across 314 μm^2 of total cell membrane surface area is about 2 million. The results are shown in Table 11.3.3 below.

1. ion	2. $[ion]_{intracellular}$ (mM)	3. # $ions_{intracellular}$	4. total ions moved during neuron response	5. change in $[ion]_{intracellular}$
sodium	10 mM	3.2×10^9 Na^+	~ 2,000,000 in on depolarization	~6.3 x 10^{-2} mM (~0.6% change)
potassium	150 mM	4.7×10^{10} K^+	~ 2,000,000 out on repolarization	~6.3 x 10^{-2} mM (~0.6% change)

Table 11.3.3 How many ions move across a membrane

$$\frac{2 \times 10^6 \mathrm{Na}^+ x \left(\frac{1\ \mathrm{mol Na}^+}{6.022 \times 10^{23} \mathrm{Na}^+}\right)}{524 \mu M^3 \times \left(\frac{1\ \mathrm{L}}{10^{15} \mu M^3}\right)} = 6.3 \times 10^{-6} \mathrm{M} = 6.3 \times 10^{-3} \mathrm{mM} \tag{11.3.4}$$

The actual change in intracellular Na^+ ion concentration on excitation is only about 0.6% of the initial $[Na^+]_{intracellular}$ on excitation and depolarization of the cell. We can also assume that the K^+ ion changes to the same degree in repolarization. So when the Na^+ and K^+ channels open, the "flood gates" are **not** opened. Permeability does increases, but this leads to a tiny influx of Na^+ ions, which is **not** sufficient to change the intracellular Na^+ concentration. It is, however, enough to change the transmembrane potential! It's a misconception that there are significant changes in ion concentration across the membrane on depolarization and repolarization of the cell. The Goldman-Hodgkins-Katz equation (previous section), shows that the membrane potential is a function of both concentrations and permeability coefficients.

Now we can present the inward rectifying potassium channels, which facilitate the movement of K^+ **into** the cell from the outside. This is in the opposite direction of the usual flow of the ions. Since the ions are moving towards higher levels of K^+ inside the cell, it would appear this would be an example of either passive diffusion favored only by the electrical potential or a case of active transport. None of these is true. K^+ ions diffuse from outside to inside the cell since outward diffusion is substantially blocked by molecules such as polyamines and Mg^{2+}! We don't have to invoke active transport or a violation of the basic rules of thermodynamics. The channels thus display strong inward currents and weak outwards ones.

The channels allow large conduction of K^+ ions if the membrane potential is more negative compared to the resting K^+ ion equilibrium potential but less if more positive, so the net effect is to maintain the resting K^+ potential.

There are several subfamilies of Kir channel. They are always potentially active (open) except those gated by G-proteins (see next chapter) or by ATP binding. (these are involved in metabolism). Kir activity is regulated by phospholipids and proteins. Now we will explore the Kir 2.2 channel, which is gated-open by PIP_2, a membrane lipid, not by changes in the transmembrane potential. Yet by opening this channel and allowing the inward flow of K^+ ions, PIP_2 is regulating the transmembrane potential. It is the agonist for the Kir 2.2 channel.

Instead of having monomeric units with S1-S6 transmembrane helical segments with a voltage sensor (S4) and P loop and S5-S6 selectivity filter, it has just 2 transmembrane helices. Both the N- and C-termini are in the cytoplasm, connected by an extracellular loop (H5) which helps form the prototypical K^+ selectivity filter with the consensus sequence T-X-G-Y(F)-G. Similar to the voltage-gated K^+ channel, four of these aggregate to form homo- or heterotetramers in the membrane. Figure 11.3.29 illustrates these points.

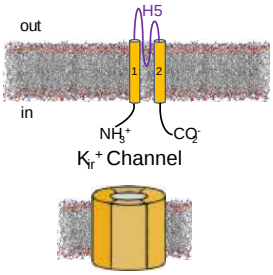


Figure 11.3.29: The Kir potassium channel

The proteins have large intracellular domains (ICD). On binding of PIP_2 to the region between the transmembrane and ICD, which produces a large conformational change, allowing K^+influx. The model below shows the R186A mutant tetrameric Kir 2.2 channel (3SPG) with four bound PIP_2 analogs containing two short fatty acids (octanol) esterified to the glycerol backbone. There appear to be two binding sites for lipids, a nonspecific site in the TMD and a specific one for PIP_2 in the ICD.

Figure 11.3.30shows an interactive iCn3D model of the inward rectifying R186A mutant tetrameric Kir 2.2 channel (3SPG) with four bound PIP_2 analogs containing two short fatty acids (octanol) esterified to the glycerol backbone. Hover over the residues to identify them.

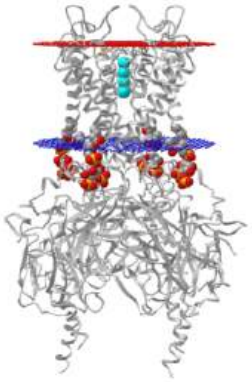


Figure 11.3.30: Inward rectifying R186A mutant tetrameric Kir 2.2 channel (3SPG). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3...S5vuBsdbSSqwV8 (long load time)

Figure 11.3.31shows an animation that shows the monomeric Kir protein morphing from the apo state (without PIP_2, 3JYC) to the PIP_2 bound state (3SPI).

Figure 11.3.31: Monomeric Kir protein morphing from the apo state (without PIP_2, 3JYC) to the PIP_2 bound state (3SPI)

Another PIP_2-gated ion channel is the transient receptor protein (TRP) channel. These channels play a role in vascular tone. Smooth muscle cells have TRPC3 and TRPC6 channels, Na^+ or Ca^{2+} channels that cause depolarization, leading to smooth muscle contraction and vasoconstriction.

11.3.5: Mechanosensitive channels

We saw in the previous chapter that some pores (not channels) are gated open not by voltage, or specific agonists such as neurotransmitters or even specific lipids like PIP_2 but by more general changes in membrane bilayers properties (membrane tension, curvature, pressure). Likewise, specific channel proteins can be gated by changes in the physical properties of the bilayer. We will consider one mechanically-active channel, TRAAK.

TRAAK - Potassium channel subfamily K member 4

The protein is a K^+ ion channel that is not sensitive to voltage changes but is open on the mechanical deformation of the bilayer. The channel can be opened by making the cytoplasm basic, raising the temperature. It is involved in pain sensation and pressure transduction. The TRAAK 4 opening is blocked by the presence of lipid from the inner leaflet which occludes the pore. If the lipid is removed by physical changes, transmembrane helix 4 rotates, which prevents the lipid from blocking the channel, which opens. Further changes in the coupled transmembrane helices 2 and 3 stabilize the opening.

Figure 11.3.32 shows an interactive iCn3D model of the TRAAK channel protein (4wff) in the **closed** state. Hover over the residues to identify them.

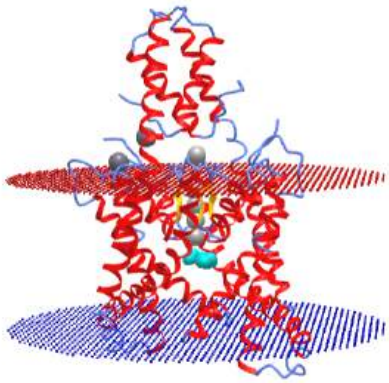


Figure 11.3.32: **Potassium channel subfamily K member 4** TRAAK channel protein in the **closed** (4wff) state. Hover over the residues to identify them. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...UzX6p5d36VicdA

The K^+ ions are aligned in the channel. Decane, a nonpolar molecule, is shown in spacefill and colored cyan. The decane is probably decanoic acid in which the carboxyl group was not defined in the structure due to high flexibility. This suggests that the "decanoic acid" is not tightly bound. It binds through the cytoplasmic side through an opening in the membrane protein. In the open state, transmembrane helix 4 rotates, blocking access to the cavity. Hence lipid binding gates the channel closed.

11.4: Diffusion Across a Membrane - Pores

11.4.1: Pores and Pore-Forming Proteins (PFPs)

If you form a pore in a cell bilayer, molecules of all sizes could move either way based on their electrochemical potential. They will move from regions of a higher to lower electrochemical potential in a thermodynamically favorable process. Hence movement through pores represents a special case of facilitated diffusion when part of the driving force is not just a concentration gradient but also an electrical potential. Several questions might come to mind.

- What proteins are involved in pore formation?
- How is the specificity of solute movement through the pore regulated?
- What is the mechanism of pore formation?

Pore formation can lead to cell death, which is the function of some pore-forming proteins (PFPs) including the toxin Hemolysin E (also known as HlyE, ClyA, SheA) secreted from E. Coli and S. Aureus. Human proteins also form a membrane attack complex (examples include the membrane attack complex-perforin/cholesterol-dependent cytolysin (MACPF/CDC) superfamily and the membrane attack complex (MAC). The MAC consists of an assembly of proteins involved in the complement system (part of the effector branch in the innate immune system), which leads to cell death of Gram-negative bacteria like E. Coli. Figure 11.4.1 illustrates the assembly process of the membrane attack complex and the complexity of interactions required to form a lethal pore in a cell.

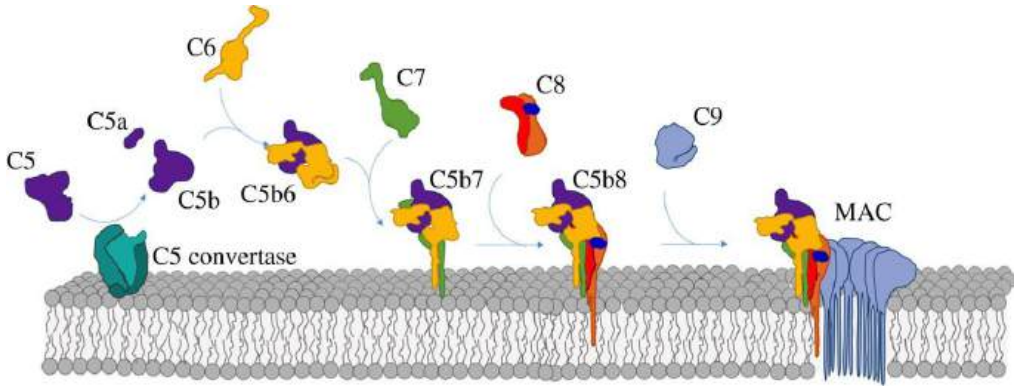


Figure 11.4.1: Human membrane attack complex. Charles Bayly-Jones et al. Philosophical Transactions of the Royal Society B: Biological Sciences, Volume 372, Issue 1726. June 2017 https://doi.org/10.1098/rstb.2016.0221. The Royal Society under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/

Complement protein C9 can adopt a soluble form or membrane form which in aggregates form the pore leading to cell death. This is a common feature of PFPs.

PFPs could create a pore by altering membrane lipid packing to form a toroid-like hole (Figure 11.4.2) and/or by inserting in a membrane and forming a pore within the protein complex itself. In either mechanism, lipid packing is altered.

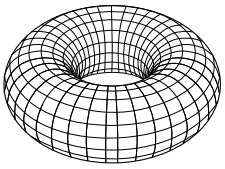


Figure 11.4.2: By GYassineMrabetTalk - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/inde...?curid=3134923

Biophysical evidence shows some support for the formation of "toroidal pores".

Lipid rearrangements in the membrane could lead to a hydrophilic or hydrophobic pore lining as shown in Figure 11.4.3

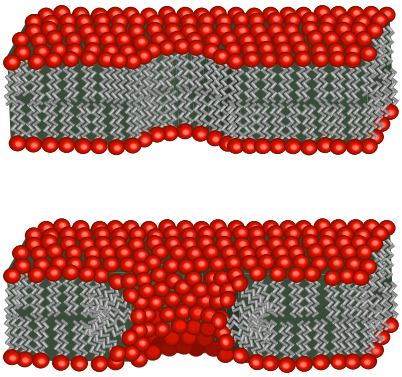


Figure 11.4.3: Diagram of possible lipid arrangements at the edge of a pore through a lipid bilayer. (above) In the absence of any rearrangement, the pore walls will be hydrophobic since the alkane tails are exposed. (below) Some researchers believe that the lipids at the edge re-orient to line the pore wall with hydrophilic head groups. Lipid bilayer. (2023, January 15). In *Wikipedia*. https://en.Wikipedia.org/wiki/Lipid_bilayer

We started our study of lipid bilayers with pure lipid systems and then added membrane proteins. Let's do the same with pore formation. A common technique to form a pore in pure lipid bilayers and also in cells is **electroporation**. This is a technique used to move a DNA with a target gene into either a prokaryotic cell (transformation) or eukaryotic cell (transfection) for exogenous gene expression. In the absence of PFPs, this requires the alteration of surface tension by applying an electrical potential. This forms depressions in the membrane, altering the nonpolar acyl chain packing. In the process, small wire-like columns of water appear, which like hydrophobic pores ultimately rearrange into hydrophilic toroidal pores. Figure 11.4.4 shows snapshots of molecular dynamics simulations as a function of the time of pore formation in electroporation.

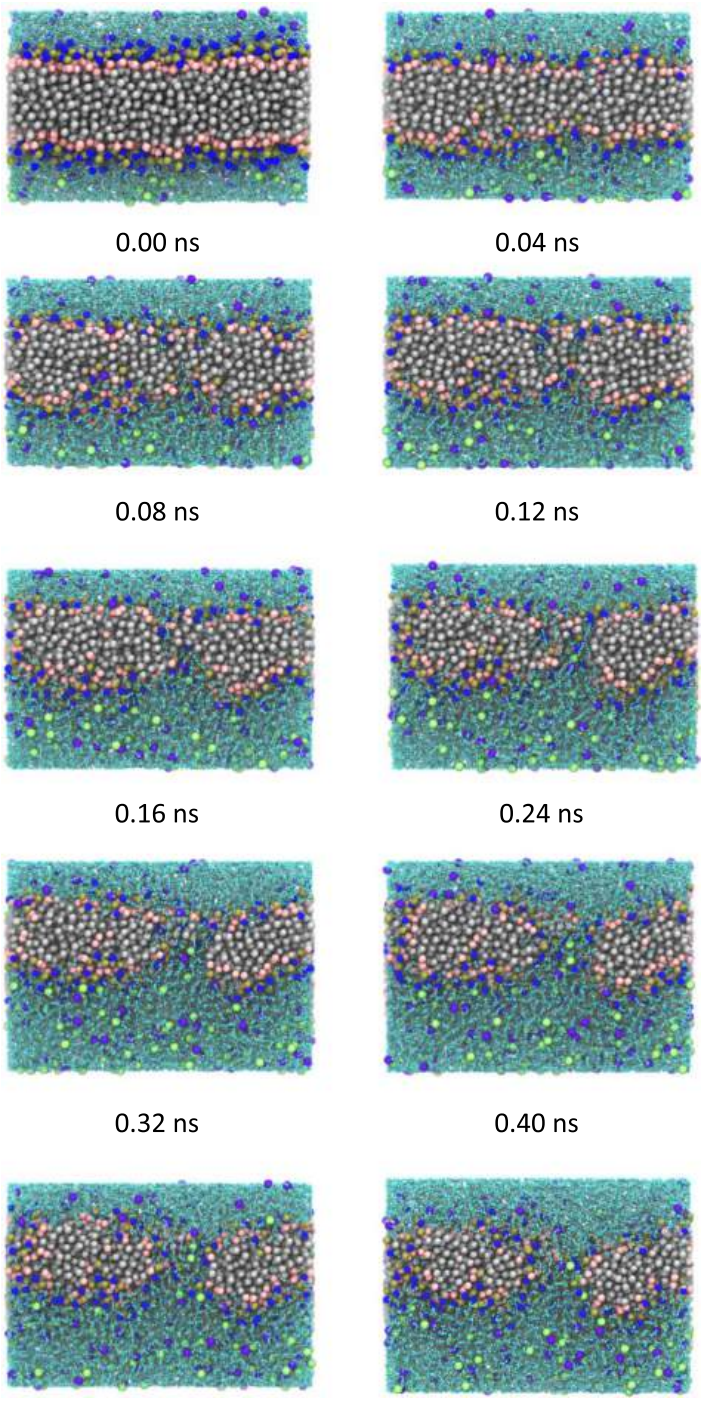


Figure 11.4.4: Zhou, C., Liu, K. Molecular dynamics simulation of reversible electroporation with Martini force field. BioMed Eng OnLine 18, 123 (2019). https://doi.org/10.1186/s12938-019-0743-1. Creative Commons Attribution 4.0 International License, http://creativecommons.org/licenses/by/4.0/

How does DNA pass through the pores in the bilayer? In pure lipid vesicles, it appears to pass through by electrophoresis. Most students are familiar with the electrophoresis of DNA fragments through pores in agarose gels. In actual living cells, small nucleic acids like small interfering RNA (siRNA) and antisense DNA molecules appear to pass through the bilayer by electrophoresis. Large DNAs like plasmids containing a gene for expression bind to the cell and form cell surface aggregates, which appear to be endocytosed into the cell. Electroosmosis, the movement of liquids under the influence of an electric field, also plays a role.

11.4.2: Pores - Outer Membrane Factor (OMF) and Voltage-Dependent Anion Channel (VDAC)

Now let's consider pores made of PFPs. We have already discussed two types of beta-barrel transmembrane proteins, the outer membrane factor (OMF) of Gram-negative bacteria and the voltage-dependent anion channel (VDAC). Both are examples of proteins called **porins** with typical beta-barrel topology.

VDAC: At low membrane potentials, VDAC (also known as mitochondrial porin), the most abundant protein in the mitochondrial outer membrane, moves metabolites and Ca^{2+} ions across the outer membrane of mitochondria. VDAC exist in an open state at 0 or very low transmembrane potentials that allows for the transfer of key metabolic anions (pyruvate, oxaloacetate, malate, succinate, ATP, ADP, and P_i, which are involved in metabolism) and Ca^{2+}, and in a closed state (above or below $\pm$ 30 mV), which is not completely closed as it allows for the transfer of ions with a preference for cations. The closed state does not allow for the transfer of ATP. The transition to the closed state is promoted by tubulin and actin (cytoskeletal proteins), negatively charged lipids such as phosphatidyl ethanolamine and cardiolipin, and also covalent phosphorylation by protein kinases. Bcl2, proteins involved in the regulation of programmed cell death (apoptosis) also interact with VDAC. In contrast to ligand-gated channels which require ligand binding to open the channel, pore complexes are unusually open. Millions of ATPs/second move across the membrane through the open pore but none through the closed pore.

Figure 11.4.5 shows an interactive iCn3D model of mouse VDAC1 (4c69) with bound ATP loosely held in the site. An alpha helix partly occludes the central pore of this β-barrel protein.

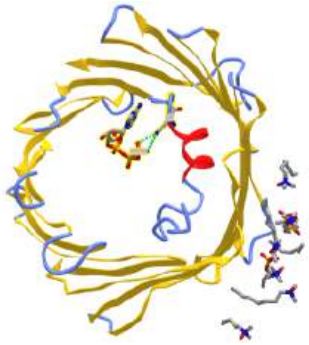


Figure 11.4.5: Mouse voltage-dependent anion channel - VDAC1 (4c69) with bound ATP loosely held in the site. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?Ueb4wg26f5oEVL4Y9

One ATP is bound in the barrel and interacts with Lys 12 and Lys 20 at each end of the cavity-bound helix. The alpha-helix narrows the pore opening and presumably changes orientation in a voltage-sensitive fashion to gate the pore open and closed, hence regulating the conductance of ions through the pore. That the orientation of the charged arginine side chains would be dependent on the transmembrane potential should be somewhat obvious.

Aquaporins

Aquaporins can move a billion water molecules per second across membranes and exclude ions including protons. Waters proceed in a single file through the pore. Instead of moving waters through, it might simply move "H^+ ions" through by the alignment of subsequent H bond donors and acceptors in a "wire" of water molecules. This is prevented by two conserved asparagine residues in the center of the channel, which disrupt water to water hydrogen bond network in the channel waters that could facilitate proton transfer. Instead, the central waters form hydrogen bonds to the central asparagines. This, along with local membrane potentials, causes opposite orientations of the water in different leaflet sides of the membrane, precluding H^+ transfer.

Here is a movie of a molecular dynamics simulation of water moving through a porin called aquaporin, GlpF, from E. Coli.

Science magazine (Tajkhorshid et al., Science Apr 2002, 296:525). Used with permission from the Theoretical and Computational Biophysics Group, the National Institutes of Health (NIH) Resource for Macromolecular Modeling and Bioinformatics, at the Beckman Institute, University of Illinois at Urbana-Champaign.

OMF: The outer membrane factor (see previous section) is one member of a class of bacterial porins, which are the most abundant proteins in the outer membrane of Gram-negative bacteria. They are classified as non-specific or specific (with respect to the solute that passes through), or monomeric, dimeric, or trimeric based on their structure. In Gram-negative bacteria, which have two lipid bilayers, the movement of solute from inside to outside includes at least three sets of proteins. Active transport (discussed in the next section) needs an energy source and is used by inner membrane transport proteins, including ATP-binding cassette (ABC)-type, resistance nodulation division (RND)-type and major facilitator superfamily (MFS)-type transporters. These are connected to membrane fusion proteins (MFP) that span the periplasm, which then interacts with at least 21 different types of porins. Molecules with molecular weights greater than 600 generally can not get through the nuclear envelope of Gram-negative bacteria, limiting the size of potential antibiotics, which must enter by passive diffusion.

11.4.3: Mechanosensitive ion channels - Mscs (which are pores!)

As the name applies, these ion channels (with openings large enough to be called pores) are gated open/closed by mechanical (physical) changes in the properties of the membrane, not extracellular/intracellular ligands or voltage changes. Certain bilayer lipids also activate the Mscs. There are two types, small (MscS) and large (MscL) mechanosensitive ion channels. Changes in local (boundary layer) and nonlocal lipids are involved in the gating of the channel (in the next section we will discuss lipid-gated ion channels). They are found in prokaryotes, archaea ,and eukaryotes. They are also called stretch-gated ion channels.

Mcss transduces a physical force (stretching and change in turgor pressure) into an electric signal - a flow of ions across the membrane. Turgor pressure is the internal pressure that "presses" the cytoplasm and cell membrane towards the cell wall in bacteria and plant cells. It arises mostly from the osmotic flow of water into the cell. If bacterial cells are placed in a high salt concentration solution (hypertonic), water flows out of the cell and the cell membrane shrinks to the inside of the volume confined by the cell wall. When placed in a hypotonic solution, water flows into the cell and the cell membrane presses out to the cell wall. The response of these channels is fast, in the millisecond range, which is about as quick as a cellular response can be. Some variants of these channels are called **piezochannels**, based on the piezoelectric effect that describes how a voltage is produced when some materials are deformed by mechanical stress which causes a redistribution of charges.

Bacteria normally have high concentrations of both K^+ and negatively charged anions, especially glutamate, which leads to a high turgor pressure from the inward osmotic flow of water. At low external osmolarity, turgor pressure in the cell could be as high as 4 atm. If placed in external solutions of high osmolarity, bacterial cells respond by the increased movement of solutes into the cell.

Mscs are particularly important when bacteria are subjected to sudden osmotic shock. If they are placed in pure water, for example, water would flow down a concentration gradient into the cell and cause the cell to swell and lyse, killing the cell. This is done in the lab to prepare almost pure hemoglobin from ruptured red blood cells. The Mscs are opened under these conditions and small species from the cytoplasm flow out, helping to keep the cell viable. Their openings must be regulated to prevent too much outward flow, which would kill the cell.

In other organisms they are also involved in touch (stretch), hearing (vibration, sound waves), and responses to gravity. Stimuli to activate them include fluid shear stress (relevant to endothelial cells that line blood vessels), membrane stretch (relevant to skeletal and cardiac muscle cells), or even indentation of a bilayer with a pipette. Changes in transmembrane turgor or other mechanical pressures cause membrane tension. Yet even in the absence of these changes, Mscs can be activated by anesthetics, phospholipids missing one fatty acid (called lysophospholipids), and certain polyunsaturated lipids. These stimuli also perturb the membrane bilayers.

Given the pore size, these proteins are less selective to ion flow compared to voltage-gate channels (see next section). Depending on the amino acids that line the pore, some Mscs would allow the preferential flow of anions while some allow cation flow.

Some somatosensory channels (i.e. not pore) proteins also respond to pressure. When open, these can be selective to specific ions like Na^+ or K^+. Examples include some variants of the Transient Receptor Potential (TRP) ion channels. Other membrane proteins can also be activated by physical force. but true Mscs have some key characteristics. If mutated or deleted, the mechanosensory response is removed. If added to a cell, a mechanosensory response results.

a. Small-conductance mechanosensitive channel

This protein is a homoheptamer with three helices from each monomer involved in the overall structure. Two (helices 1 and 2) interact more with the lipid components of the bilayer. Interactions of specific lipids with the helices seem to promote closing but changes under high pressure lead to the opening of the pore. Half of each helix 3 forms the pore, while the other half is more parallel to the membrane and interacts with a large cytoplasmic domain.

Figure 11.4.6 shows the differences between the closed form of E. Coli MscS (2oau) and the open form (2vv5) viewing down the pore axis. The heptameric protein is shown in gray. Two key valines (105 and 109) on each chain are shown in spacefill and colored cyan. These hydrophobic amino acids act like gate-keepers, helping to keep water out, acting like a "vapor seal". In the closed state, the pore is sealed by the closing of the leucine "rings" as one half of helix 3 pack more closely. Many members of the MscS family vary significantly in size and can have between 3-11 transmembrane regions. The closed pore has a diameter of about 4.8 Å, while the open pore is 13 Å across.

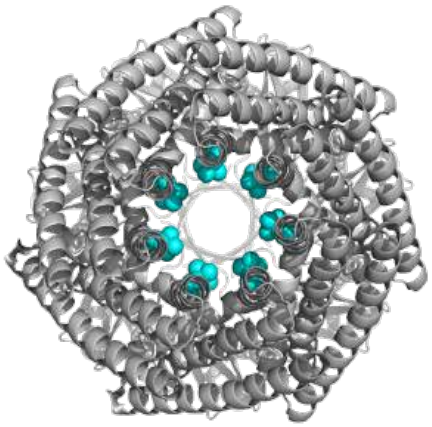


Figure 11.4.6: Differences between the closed form of E. Coli MscS (2oau) and the open form (2vv5) viewing down the pore axis.

Most of you would have studied Ohm's law, given by

$$\mathrm{V} = \mathrm{IR} \text{ or } \mathrm{I} = \frac{\mathrm{V}}{\mathrm{R}} \tag{11.4.1}$$

where I (amps) is the current, R (ohms) is the resistance and V (volts) is the voltage. A more general variant of this law used in physics is

$$\mathrm{J} = \sigma \mathrm{E} \tag{11.4.2}$$

where J is the current density, E is the electric field, and σ is the conductivity (inverse of resistance) which depends on the material. The unit of sigma σ is $ohms^{-1}$ or mhos (ohm written backward). That unit has been renamed the Siemen (S). Mscs channels have a small conductance of approximately 1 nS in 400 mM salt solution.

b. Large conductance mechanosensitive ion channel (MscL):

The channels have large conductances (3 nS) and concomitantly larger pore sizes, allowing the flow of water, ions, and even small proteins. Again they are involved in diffusion down an electrochemical gradient through pores so not active transport just gated diffusion.

Figure 11.4.7 shows an interactive iCn3D model of the pentameric MscL from Mycobacterium tuberculosis (2oar), a Gram-positive bacteria that causes tuberculosis. About 23% of the world's population is affected by this pathogen. It causes about 1.5 million deaths each year. Compare this to the total number of deaths during the COVID pandemic (2020-21) of over 3 million (as of May 2021). It has killed over 1 billion people throughout human history (but not as many as malaria). The pore diameter is about 30 Å across.

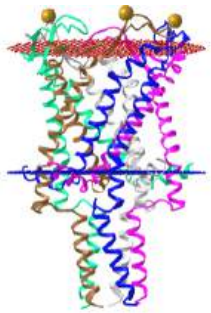


Figure 11.4.7: Closed state of the pentameric MscL from Mycobacterium tuberculosis (2oar). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...LsRZZbmFy5NRF6

11.4.4: Pore-forming alpha-helical toxins

We started this chapter section with a discussion of the major attack complex (MAC) of the innate immune system. Pathogens also employ pore formation to kill host cells. Many secrete soluble proteins which aggregate in the membrane to form either alpha-helical or beta-barrel pores. The proteins are called pore-forming toxins (PFTs). Killing occurs when either cytoplasmic components leak out or bacterial toxins, such as diphtheria and anthrax toxin, move into the cell.

Figure 11.4.8 shows an interactive iCn3D model of the pore formed by cytolysin A (ClyA, also known as HlyE), an alpha-PFT used by some E. Coi and Salmonella enterica strains. The pore is a large dodecamer that forms from soluble monomeric ClyA (2wcd). It has a pore diameter of about 40 Å,

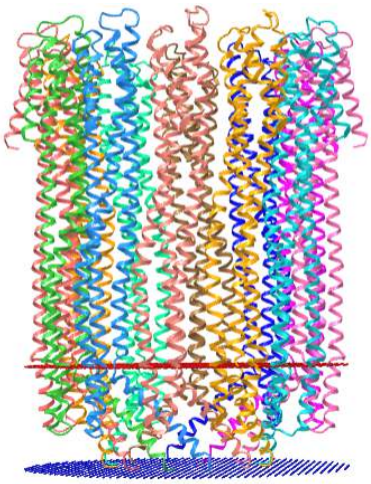


Figure 11.4.8: Pore formed by cytolysin A (ClyA, also known as HlyE). (2wcd) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...RXUsE24d9d8NT6

Figure 11.4.9 shows the soluble monomeric form of cytolysin A (ClyA, HlyE) (1QOY). Nonpolar side chains are shown in cyan. The transparent surface is mostly polar, making the monomer soluble.

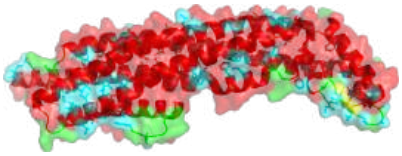


Figure 11.4.9: Soluble monomeric form of cytolysin A (ClyA, HlyE) (1QOY)

Figure 11.4.10 shows the changes in conformation between the single chain soluble form (shown in the figure above) and a single chain of the membrane oligomeric channel.

Figure 11.4.10: Changes in conformation between the single chain soluble form and the same chain in the channel

We started this chapter section by exploring electroporation and the formation of a toroidal-like hole in the lipid bilayer. In the case of ClyA, the protein aggregate itself forms the pore, not the lipids themselves, although lipid rearrangements are necessary to form the protein complex.

11.4.5: Gap Junctions

Connexins are voltage-gated channels that allow for the flow of ions, metabolites, nucleotides, and small peptides. A connexin has four transmembrane helices and two extracellular loops, Six protomers of these come together in a single cell to form a channel complex called a **connexon** or hemichannel. Beta structures in the connexin hemichannel of one cell docks with a similar channel on an adjoining cell to form a full channel passing through the membranes of both cells forming a gap junction between the cells. This is shown in Figure 11.4.11. There are 20 different connexins encoded in the human genome.

The left figure below shows the six protomers, each in a different color and a gray rectangle representing the bilayer of a single connexon or hemichannel. The right side of the figure shows a full gap junction channel between two cells, with the membranes represented by gray rectangles. The connexin 26 monomer was used to create the diagram. Mutations in this protein are associated with hearing loss.

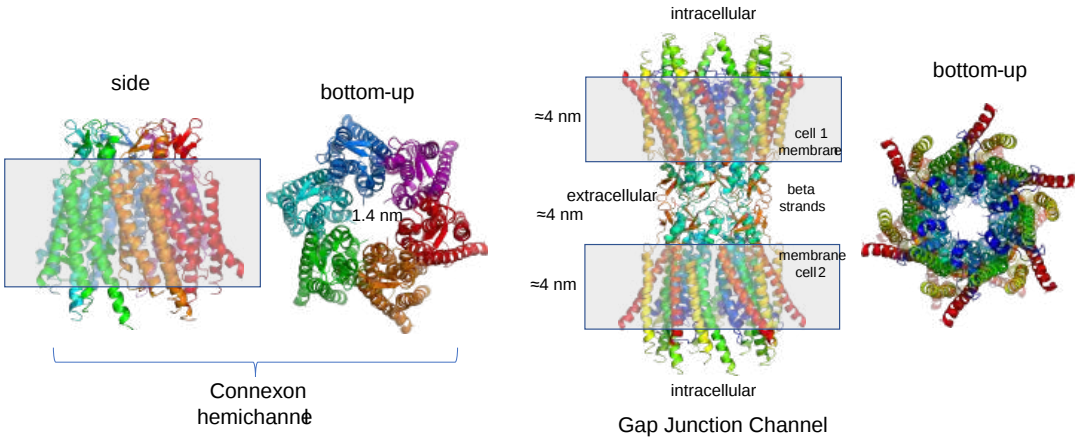


Figure 11.4.11: Connexon channel and the gap junction

The left figure above shows the six protomers, each in a different color and a gray rectangle representing the bilayer of a single connexon or hemichannel. The right side of the figure shows a full gap junction channel between two cells, with the membranes represented by gray rectangles. The connexin 26 monomer was used to create the diagram. Mutations in this protein are associated with hearing loss.

The cytoplasmic entrance of the channel is positively charged. It forms a funnel, composed of 6 amino-terminal helices, which leads into a negatively charged transmembrane lining. The entrance diameter is 14 Å.

Here is a model of a full gap junction channel connecting two membranes using the human connexin 26 monomer (2zw3). Only the bottom membrane bilayer is represented by red and blue dummy atoms.

Figure 11.4.12 shows an interactive iCn3D model of a full gap junction channel connecting two membranes using the human connexin 26 monomer (2zw3). Only the bottom membrane bilayer is represented by red and blue dummy atoms.

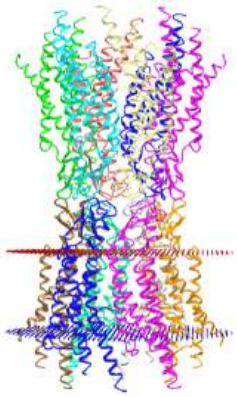


Figure 11.4.12: Gap junction channel connecting two membranes using the human connexin 26 monomer (2zw3). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?TpuQ6Agk8XWuApsWA

11.4.6: The Nuclear Pore Complex (NPC)

Channels have pores that can be gated open and allow the selective flow of ions. Pore-forming proteins have larger entrances that allow both small and large molecules to pass through the bilayer. The pore opening in even large mechanical sensitive channels (about pale in size compared to the nuclear pore complex, which has a combined molecular mass of around 125,000,000! Its outer diameter of ~1,200 Å and its inner one of about 425-Å. Figure 11.4.13 shows the relative size of nuclear pore compared to other molecular structures including the eukaryotic ribosome, nucleosome, a soluble tetrameric protein (rubisco, 270K), and MscL (shown as a circle which represents the pore diameter).

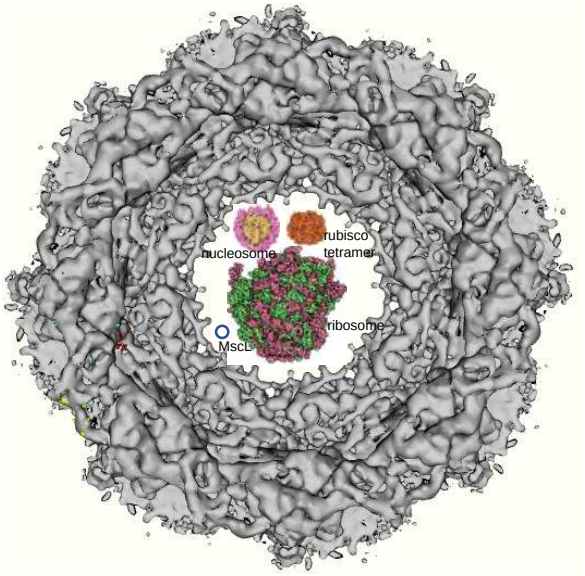


Figure 11.4.13: Relative size of nuclear pore compared to other molecular structures including the eukaryotic ribosome, nucleosome, a soluble tetrameric protein (rubisco, 70K) , and MscL (shown as a circle which rrepresentsthe pore diameter)

Its job is to shuttle small molecules by passive diffusion down a concentration gradient through the pore. In addition, it moves large molecules and molecular structures (proteins, RNA, and perhaps ribosomes) across the nuclear membrane in a process that requires energy. The proteins that comprise this complex are called **nucleoporins (nups)**, of which there appear to be around 34 in humans. Each NPC complex contains around 1000 **nucleoporins**. The complex fuses the inner and outer nuclear membranes.

We have focused so much on single bilayer membranes that comprise the plasma membrane and membranes of organelles like the Golgi complex and lysosomes, it might come as a surprise (perhaps not to biology students) that the nuclear membrane or envelope appears to consist of two bilayers. Most know that the mitochondria have two bilayers, an inner and outer membrane, similar to Gram-negative bacteria. Mitochondrial are believed to have arisen from bacteria so the double bilayer there makes sense. Figure 11.4.14shows the nuclear membrane or envelope of two bilayers (1) with an outer ring (2), spokes (3), a basket (4), and filaments (5). The NPC spans both bilayers.

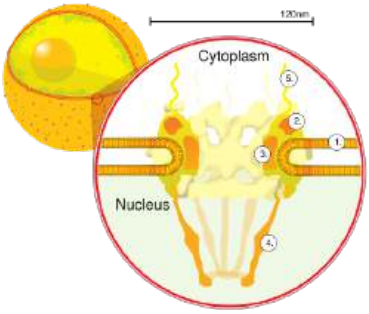


Figure \(\PageIndex{14\): Nuclear membrane or envelope of two bilayers (1) with an outer ring (2), spokes (3), a basket (4) ,and filaments (5) https://en.Wikipedia.org/wiki/File:N...e_crop.svg.png; Creative Commons Attribution-Share Alike 2.5 Generic

The outer bilayer of the nuclear envelope is continuous with the endoplasmic reticulum as shown in Figure \(\PageIndex{15\) below. The dots on the ER membrane are ribosomes, making this the rough ER (as opposed to smooth ER, which has no attached ribosomes).

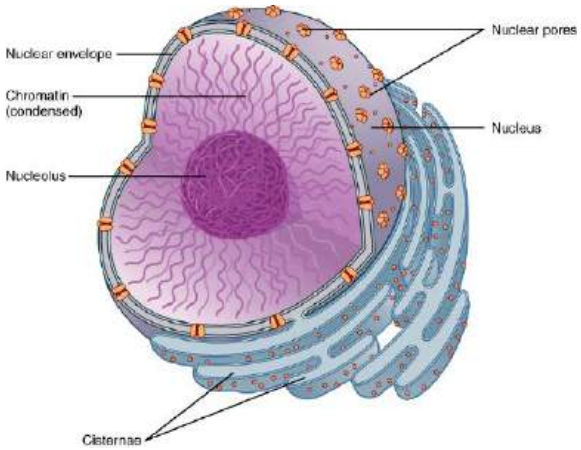


Figure \(\PageIndex{15\): The outer bilayer of the nuclear envelope is continuous with the endoplasmic reticulum https://commons.wikimedia.org/wiki/F...18_Nucleus.jpg; OpenStax, CC BY 4.0 <https://creativecommons.org/licenses/by/4.0>, via Wikimedia Commons

Figure \(\PageIndex{16\) below shows a model of the basket-like structure of the nuclear pore complex (NPC). It, as well as Figure XX, shows that instead of two separate bilayers, in actuality, there is just one bilayer with each leaflet bending around at the NPC and reversing directions! Think of the interesting lipids and protein components that enable the bend! Alternatively, you could say that 2 different membranes fuse at the NPC.

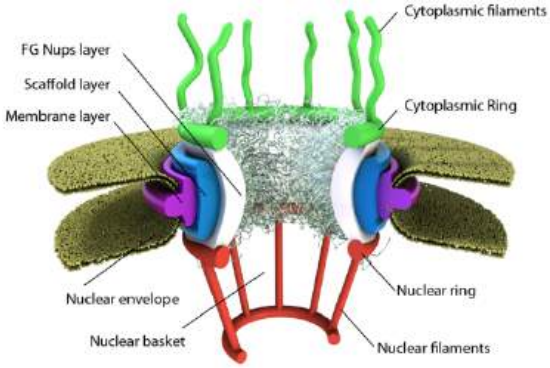


Figure \(\PageIndex{16\): A model of the basket-like structure of the nuclear pore complex **Citation:** Azimi M, Mofrad MRK (2013) Higher Nucleoporin-Importinβ Affinity at the Nuclear Basket Increases Nucleocytoplasmic Import. PLoS ONE 8(11): e81741. https://doi.org/10.1371/journal.pone.0081741. Creative Commons Attribution License,

The NPC consists of 32 copies of each specific nucleoporin (Nup) except two. One of those has 48 copies and the other 16 (even these sum to 2x32 Nups). Three rings form and surround the pore. There is a 16-membered ring of Nups facing the cytoplasm (**cytoplasmic ring**) and another 16-membered ring of Nups facing the nucleoplasm (**nuclear ring**). There is 8-fold rotational symmetry in each ring, suggesting a dimeric repeat of Nups in the rings. Eight Nups in the cytoplasmic ring have a disordered end that sticks out into the cytoplasm as filaments. In contrast, the disordered ends of eight Nups in the nuclear ring form filaments which bind together to form a ring at the bottom of the nuclear basket.

The **inner ring** (called the **FG Nups layer** in the figure above), between the cytoplasmic and nucleoplasmic rings, also consists of Nups with ordered domains and disordered parts. The disordered parts of the inner ring Nups have repetitive sequences enriched in phenylalanine and glycine (hence the name FG Nups) and stick out into the central pore. These disordered region act as a filter allowing certain molecules to pass and excluding those greater with molecular weights greater than 40K. Large molecules need transport proteins called karyopherin transport factors to move through the pore.

The core structure of the NPC, obtained through cryoelectron microscopy, is shown in Figure \(\PageIndex{17\) below.

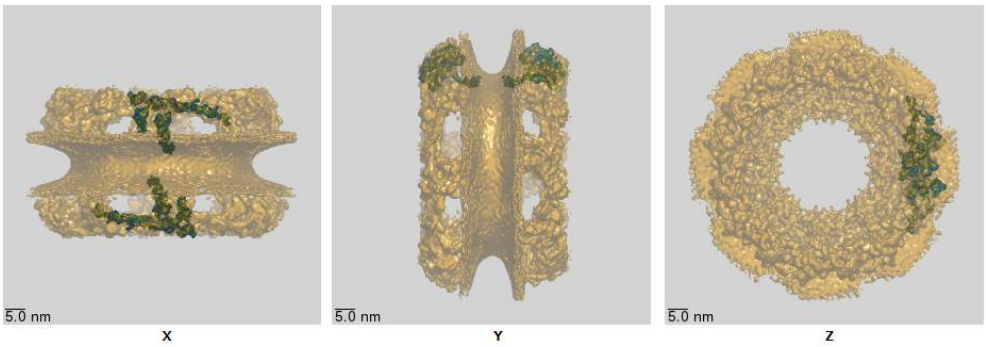


Figure \(\PageIndex{17\): The core structure of the Nuclear Pore Complex https://www.ebi.ac.uk/pdbe/entry/emd...-3103/analysis

The double bilayers are very evident. The nuclear and cytoplasmic filaments, as well as the disordered FG repetitive sequences that stick into the pore from the inner ring, are not seen since they are very flexible and don't' adopt single conformations using standard structure determination methods.

Figure 11.4.18 shows an interactive iCn3D model of the nuclear pore complex (NPC) core (5a9q), whose structure was obtained using cryoelectron microscopy. (load slowly)

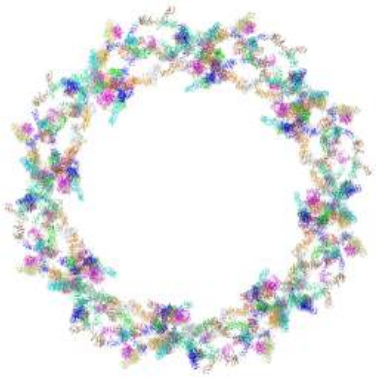


Figure 11.4.18: Nuclear pore complex (NPC) core (5a9q). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?NP7oQb235BCgPARA8

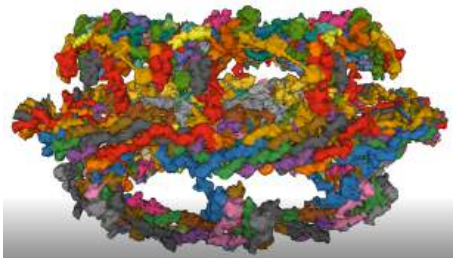
A cartoon figure showing the variety of Nups in the NPC is shown in Figure 11.4.19

Figure 11.4.19: A cartoon version of the various NUPS in the nuclear pore complex. Donnaloja F., Jacchetti E., Soncini M., Raimondi M. T.Front. Physiol., 12 July 2019 | https://doi.org/10.3389/fphys.2019.00896**Creative Commons Attribution License (CC BY)**.

In this figure, the cytoplasmic and nucleoplasmic rings are both shown in green, each mostly formed by 16 copies of the Y-complex, arranged in two eight-membered rings. The inner ring, predominantly formed by 32 copies of the Nup93 complex is shown in red. Transmembrane nucleoporins are depicted in violet, and the cytoplasmic filaments and nuclear basket structure are in orange. Attached to the inner ring are Nup62 complexes (depicted in blue), which form a cohesive meshwork within the central channel through their FG-repeat domains. Not indicated is the position of Nup98, a FG-repeat-containing nucleoporin important for the transport and exclusion function of NPCs; its position in the NPC is less defined, but it might be part of the inner ring. Similarly, Aladin (also known as AAAS), Gle1, Rae1, and Npl1 (also known as hCG1 and NUPL2) have been omitted.

Large proteins and RNA that pass through the pore must first be bound to a cargo receptor, which can move the "cargo" across the pore with concomitant GTP hydrolysis. This is a process that is closer to active transport so we will discuss that in Chapter 11.3.

The entire nuclear pore complex was solved in 2022 using cryoEM. Here are two videos showing the dilated complex (7R5J). Click on the images to download mp4 animations of the complex.



11.5: Active Transport

In the previous sections, we explored facilitated diffusion, diffusion through channels, and diffusion through larger pores. In each case, once a carrier/permease protein was available, or a channel (gated by ligand binding, change in membrane potential, lipid binding, or mechanical forces) or a pore formed, solute flows down a chemical gradient (facilitated diffusion) or electrochemical gradient in a thermodynamically favored process. But what about moving solutes from low to high concentration, against a concentration gradient, which would be necessary to capture the last bit of a vital nutrient or energy source? Active transport does just that, but it requires an energy source to do so. Many different types of chemical species are actively transported across the cell and organelle membranes, including sugars, amino acids, (deoxy)nucleotides, metabolites (like carboxylic acids),

11.5.1: Type of Active Transport

For active transport to occur, a membrane receptor is required which recognizes the ligand to be transported. Of major interest to us, however, is the energy source used to drive transport against a concentration gradient. The biological world has adapted to use almost any source of energy available.

ATP hydrolysis: One would expect that this ubiquitous carrier of free energy would be used to drive active transport. This is one of the predominant roles of ATP in the biological world. 70% of all ATP turnover in the brain is used for the creation and maintenance of a Na and K ion gradient across nerve cell membranes using the membrane protein Na^+/K^+ ATPase.

Energy released by oxidation: You may have encountered this in previous biology courses. Active transport of protons driven by oxidative processes is exergonic. In electron transport in respiring mitochondria, NADH is oxidized as it passes electrons to a series of mobile electron carriers (ubiquinone, cytochrome C, and eventually dioxygen) using protein complexes in the inner membrane of the mitochondria. The energy lost in this thermodynamically favored process is coupled to conformational changes in the complex which caused protons to be ejected from the matrix into the inner membrane space. One can imagine a series of conformation-sensitive pKa changes in various side chains in the complexes which lead in concert to the vectorial discharge of protons.

Light: Photosynthetic bacteria have a membrane protein called **bacteriorhodopsin** which contains **retinal**, a conjugated polyene derived from beta-carotene. It is analogous to the visual pigment protein rhodopsin in retinal cells. Absorption of light by retinal induces conformation changes in the retinal and bacteriorhodopsin, which leads to the vectorial discharge of protons.

The collapse of an ion gradient: The favorable collapse of an ion gradient can be used to drive the transport of a different species against a concentration gradient. We have already observed that the collapse of a proton gradient across the inner mitochondria membrane (through F_0F_1ATPase) can drive the thermodynamically unfavored synthesis of ATP. The collapse of a proton gradient provides a proton-motive force that can drive the active transport of sugars. Likewise, a sodium-motive force can drive the active transport of metal ions. Since the energy to make the initial ion gradients usually comes from ATP hydrolysis, ATP indirectly powers the transport of the other species against a gradient.

Active transport can be divided into two classes:

- primary - driven by "primary" energy sources, which for us means ATP or photons
- secondary - driven by coupled transport (also called cotransport) of solutes down a concentration gradient, which provides the thermodynamic force to move a solute "uphill".

Active transporters can also be divided into classes based on the direction of movement of the solute and any cotransported solute into uniport (no cotransported solute), symport (solute and cosolute transported in the same direction) and antiport (solute and cosolute transported in the opposite direction) as shown in Figure 11.5.1.

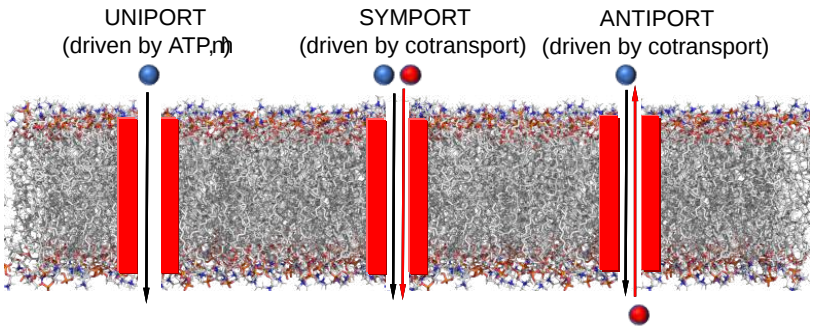


Figure 11.5.1: Uniport (no cotransported solute), symport (solute and cosolute transported in the same direction), and antiport (solute and cosolute transported in the opposite direction) transporters

Examples of these different mechanisms can be found in facilitated diffusion as well as active transport. Examples of these types include:

- uniport: An often-used example is GLUT 1, which catalyzes the FACILITATED DIFFUSION (not active transport) of glucose down a concentration gradient. A somewhat complicated example of a uniport active transporter is the Na^+/K^+ ATPase, which moves Na^+ ions out of a cell and K^+ into a cell, both against concentration gradients. The movement of one ion does not provide the chemical potential thermodynamic driving force to move the other. Rather the hydrolysis of ATP is required. So this protein can be called a uniport active transporter for two different ions.
- symport: Another glucose transporter, the glucose symporter (SGLT1), found in the small intestines, heart, and brain, cotransports one glucose or galactose for every two Na^+ ions that move into the cells down a concentration gradient.
- antiport: The sodium-calcium exchanger pumps out one Ca^{2+} ion (low to high concentration) driven by the influx of three Na^+ into the cell. This keeps Ca^{2+} low in the cytoplasm.

If the species moved is charged, two other terms are used:

- electrogenic - a net electrical imbalance is generated across the membrane by symport or antiport of charged species
- electroneutral - no net electrical imbalance is generated across the membrane by symport or antiport of charged species

In this chapter section, we will explore examples of several types of active transporters. All are polytopic with alpha-helical transmembrane domains, which through a series of conformational transitions can move chemical species that are bound to the receptor or channel protein to move through the membrane through the creation of transient openings in the protein. We will end with another look at the nuclear pore complex and see how it allows the movement of large proteins and RNA molecules through its pore through a very different type of mechanism. In that case, it is not active transport since the relative concentration of the transported species on either side of the pore is less important than the mechanism by which certain species are targeted to move into and out of the nucleus.

11.5.2: Major Facilitator Superfamily (MFS) Transporters

There are about 74 different families of active transporter in the Major Facilitator Superfamily (MFS), which is the largest of secondary transporters using symport or antiport mechanisms. There are also uniporters within this family. Different members of the MFS move a diverse array of solutes, including sugars, nucleotides, peptides, and drugs across membranes against a concentration gradient. All have two membrane domains consisting of six transmembrane helix bundles that pack to give an apparent two two-fold rotational axis of symmetry (also called pseudo C_2 axis). The ligand to be transported binds between the two domains in a central cavity potentially open to either side of the membrane. Figure 11.5.2 panel A (top) below shows a cartoon of the 12 helices and 6-helix bundles. Helix 1 and 4 in domain one and the analogous ones in domains 7 and 10 are involved in ligand movement through the membrane. Correspondingly, helices 2, 5, 8, and 11 are involved in domain-domain interactions. The bottom figure (B) shows two views of the protein with the ligand in spacefill rendering.

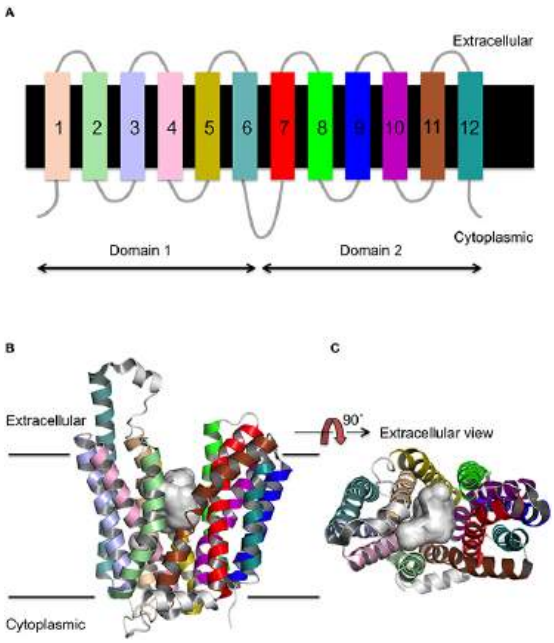


Figure 11.5.2: Structure of the Major Facilitator Superfamily (MFS). images from Lee J, Sands ZA and Biggin PC (2016) A Numbering System for MFS Transporter Proteins. *Front. Mol. Biosci*. 3:21. doi: 10.3389/fmol.2016.00021 **Creative Commons Attribution License (CC BY)**

As structure mediates function, they all display a similar "clamp and switch" mechanism, in which solute binds to the protein in an open (to binding) conformation, leading to conformation changes forming a closed form, followed by further conformational changes which cause the complex to open to the other side of the membrane. Figure 11.5.3 shows hands opening, closing, and reopening in the other direction as an analogy for this "clamp-switch" mechanism.

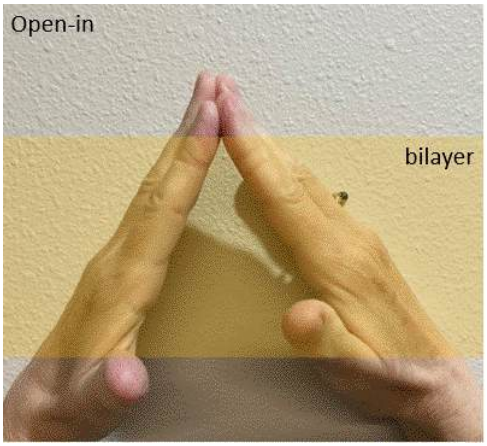


Figure 11.5.3: Hands opening, closing, and reopening in the other direction as an analogy for this "clamp-switch" mechanism of the Major Facilitator Superfamily (MFS)

We will describe two different transporters in this family.

Human monocarboxylate transporter (7BP3)

These proton-linked monocarboxylate transporters (MCTs) move many monocarboxylates (including lactate, pyruvate, and the ketone bodies acetoacetate, beta-hydroxybutyrate and acetate) out of the cell, driven by the flow of protons across the membrane in the same directions (symport). In cells engaging in glycolysis, a fundamental energy utilization pathway, it catalyzes the high-affinity transport of pyruvate, the end product of glycolysis, and lactate, a reduced derivative of pyruvate, out of the cell. In cells that are not engaged in glycolysis, their activity is effectively turned off. The flux of pyruvate/lactate has steep dependence on their concentrations, allowing the protein to act as an on-offf switch for transport. A step dependency of activity on concentration is a hallmark of cooperative interactions. Structural analyzes show conformational changes in the interface of the dimeric form of the transporter.

We will see when we study glycolysis, the conversion of the oxidized pyruvate to its reduced form lactate requires a reducing agent, in this case, NADH, which is oxidized back to NAD^+, an oxidizing agent required for glycolysis to continue. When pyruvate is high, NADH is also high, allowing for the conversion of pyruvate to lactate, which can be exported out of the cell. The interconversion is described by this reaction:

$$\text{Pyruvate} + \text{NADH} \leftrightarrow \text{Lactate} + NAD^+$$

Using MCT 2 to remove lactate from the cell pulls the above reaction to the right, regenerating NAD^+, and allowing glycolysis to continue.

Figure 11.5.4shows an interactive iCn3D model of the inward open human monocarboxylate transporter 2 (7BP3).

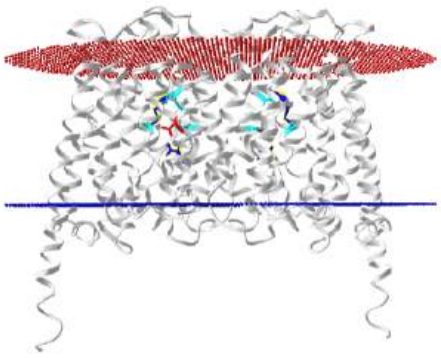


Figure \(\PageIndex{4\): Inward open human monocarboxylate transporter 2 (7BP3). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...yq5U8v4BvdDA57

Several key residues are shown. Aspartic acid 293 (side chains in red sticks) is conserved in all MCTs studied and is likely involved in H^+ transport. It is surrounded by Val 156, Met 289, Ala 290, and Phe 351 (side chains in cyan sticks), all nonpolar amino acids which would elevate the pKa of Asp 293, allowing it to stay protonated until a conformational change would alter its environment, leading to proton release. Two positive charged residues, Lys 38 and Arg 297 (side chains in blue sticks) are involved in substrate interactions. Lys 38 faces to the outside of the cell in the inward open conformation. Figure \(\PageIndex{5\) below shows the same amino acids in a zoomed view. The amino acids are labeled.

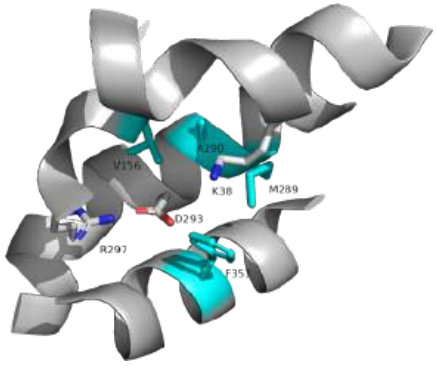


Figure \(\PageIndex{5\): Key amino acid side chains in inward open human monocarboxylate transporter 2 (7BP3)

Lactose Permease (Lactose-proton symport)

This E. Coli protein is also a symporter, which uses the movement of H^+ down a concentration gradient to drive lactose into E. Coli to capture as much lactose - an energy source - as possible. Deprotonation of a protonated Glu 269 leads to ligand binding as the protein has no binding site in the protonated site. Lactose induces the formation of its binding site. Figure \(\PageIndex{6\) below shows the change in conformation going from a more acidic pH (5.6, PDB 2CFP) to a more neutral pH (6.5, PDB 2CFQ). One lactose is transported from one H^+, which moves down a concentration gradient. Lactose moves from the outside periplasm to the cytoplasm.

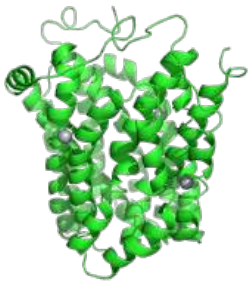


Figure \(\PageIndex{6\): Conformation changes in lactose permease from a more acidic pH (5.6, PDB 2CFP) to a more neutral pH (6.5, PDB 2CFQ)

Figure 11.5.7 shows an interactive iCn3D model of lac permease (1pv7), which highlights three residues essential for lactose binding (Glu 126, Arg 144, and Glu 269 in stick) and three involved in proton transfer (Arg 302, His 322 and Glu 325 in spacefill).

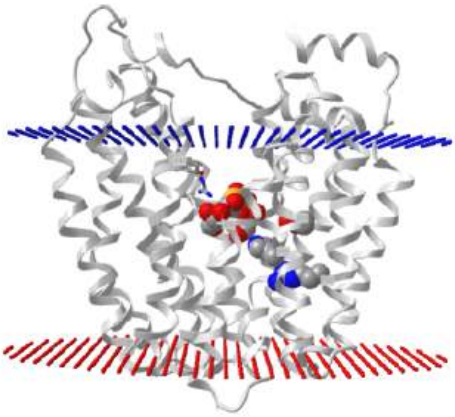


Figure 11.5.7: Lac permease (1pv7). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...9mt2juToECQPg6

It has a bound nonhydrolyzable lactose analog (beta-D-galactopyranose-(1-1)-1-thio-beta-D-galactopyranose), shown in spacefill, in the center hydrophilic cavity. The two 6-helix bundles, which give pseudo 2-fold symmetry are evident. The protein is in the open-to-inward conformation. (The blue dummy atoms represent the inner leaflet.)

Figure 11.5.8is a video of a molecular dynamics simulation showing lactose moving through lactose permease.

Figure 11.5.8 Molecular dynamics simulation showing lactose moving through lactose permease. https://www.ks.uiuc.edu/Gallery/Movi...annelProteins/

As with other members of the MFS transporters, lactose permease has 12 transmembrane helices that form two bundles. The lactose binding site is between the two bundles. When the protein is in the open-to-out conformation, it can bind protons, most likely at an exposed His. This enables lactose binding, which is followed by a conformational change to the open-to-in form. The protonated His loses a proton and lactose dissociates into the cytoplasm. Movement of both H^+ and lactose together must occur for transport.

Structural work by Singh et al on the Leu Transporter (LeuT), a member of the solute carrier 6 or sodium-coupled transporters, which is an active transporter requiring movement of Na ion into the cells to power the uptake of Leu, shows an "open-to-out" and occluded binding state for ligand (Leu). Tryptophan, a competitive nontransportable inhibitor binds to the open-to-out state but is too large for the obligate occluded state so it is not transported. Figure 11.5.9 shows inward open, closed, and outward conformations of different major facilitator superfamily transporters.

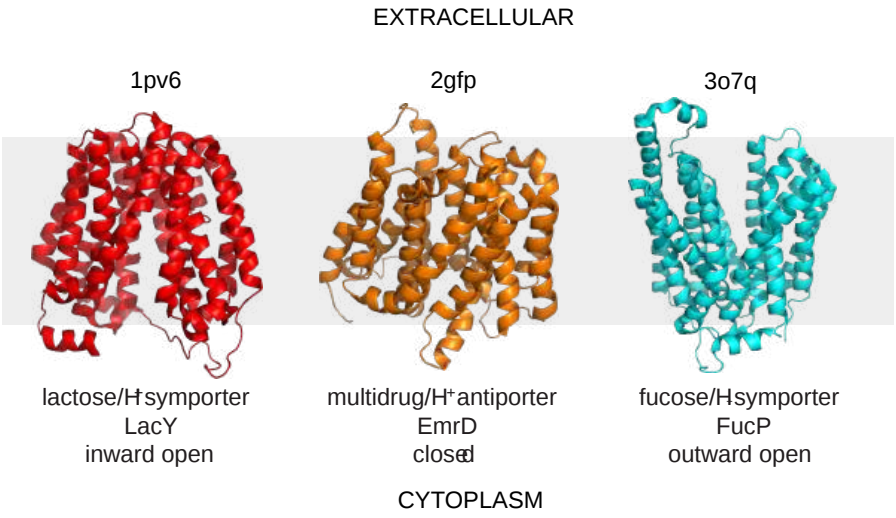


Figure 11.5.9: Inward open, closed, and outward conformations of different major facilitator superfamily transporters

Another member of the MFS Transporter family is LeuT, the leucine transport, which is similar to the neurotransmitter sodium symporter. Figure 11.5.10shows the transition between two conformational states. The first is the outward open leucine transporter (LeuT). This initial state has bound tryptophan, a competitive inhibitor, which is too big to transport, which traps the bound state in the open conformation. The second state has bound Leu and is closed. The conformational changes are subtle but sufficient to allow transient binding, occlusion, and expulsion to the other side of the membrane (not shown). Note also the two bound Na^+ ions (purple) clearly show that active transport through this transporter is driven by Na^+ ions. The ions are farther apart in the open state allowing access to the amino acid for transport.

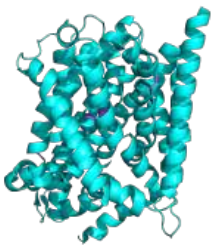


Figure 11.5.10: Transition between two conformational states of the Leu transporter

11.5.3: Membrane ATPases

Many active transporters power the uphill transport of solute through coupled exergonic cleavage of ATP. The largest family of is the **P-type ATPase**, which is used to transport a host of different solutes, including ions (Cu, Zn, Mn, Mg, Ca, Na, K, Cd, Co, Pb, Ni ,and H) and phospholipids (from one leaflet to another by translocases like flippases). They convert chemical energy (from

cleavage of phosphoanhydride bonds) to electrochemical potential energy. Over 160,000 P-type ATPases from prokaryotes and eukaryotes are known so they are highly prevalent and evolved early in time. They should be understood at both a structural and thermodynamic level.

How is energetic coupling achieved? An early model proposed by Albers and Post suggested that energetic coupling occurred by the covalent transfer of a terminal phosphate (P_i) from ATP to an Asp in the protein channel to form a covalent Asp-Pi intermediate, which is an example of a mixed anhydride. This is illustrated in Figure 11.5.11.

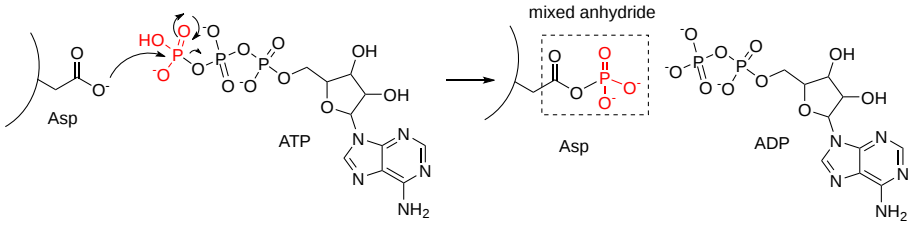


Figure 11.5.11: Formation of a high energy (compared to its hydrolysis products) mixed anhydride of Asp in membrane ATPase. (**Remember, there is no such thing as a "high energy" bond.**)

This reaction is disfavored thermodynamically since hydrolysis of ATP to form ADP proceeds with a **less negative** ΔG^0 than the same reaction with a mixed anhydride. However, ultimately hydrolysis of the mixed anhydride in the next step would make the overall coupled reactions favorable.

Another feature of the Post-Albers model is that the phosphorylated channel, P-ATPase, has two conformations:

- a high-affinity cation inward (cytoplasm) facing cation binding site called E_1
- a low-affinity cation outward (extracellular or luminal) facing cation binding site called E_2

Figure 11.5.12 is a more detailed (A) and simpler (B) set of chemical equations to represent the different states of E_1 and E_2 in the Post-Albers model.

The next three figures are from the following reference: Zhang, X.C., Zhang, H. P-type ATPases use a domain-association mechanism to couple ATP hydrolysis to conformational change. *Biophys Rep* **5,** 167–175 (2019). https://doi.org/10.1007/s41048-019-0087-1. Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/

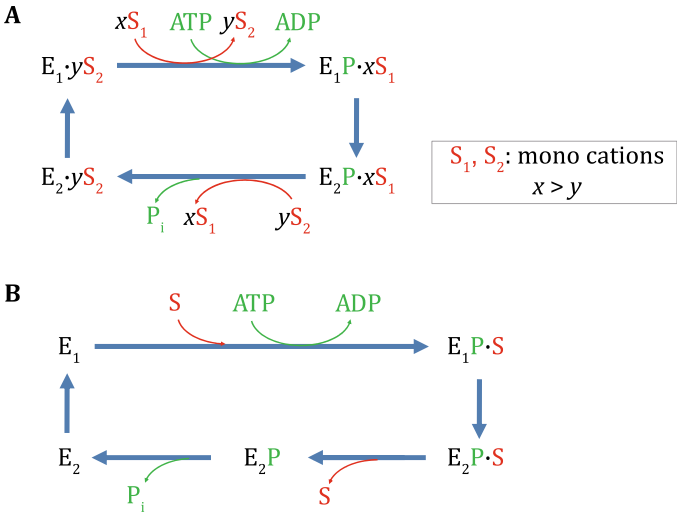


Figure 11.5.12: Chemical equations for conversion of states E_1 and E_2 for membrane ATPases in the Post-Albers mode

- The top line in both shows the reactions of E_1 (high affinity) with the substrate (ions) and ATP. Since E_1 is phosphorylated by ATP, it is considered a **kinase.**
- The bottom line in both shows the reactions of E_2 (low affinity) with the substrate (ions). Since E_2 is dephosphorylated, it is considered a **phosphatase**.

S_1 and S_2 differ for specific P-type ATPase.

Detailed Structures

Now let's look at the structures of a specific P-type ATPase, the sarcoplasmic/endoplasmic reticulum calcium ATPase 1 (or more simply a Ca^{2+} pump), which catalyzes the active uphill transport of Ca^{2+} from the cytoplasm back into the lumen of the sarcoplasmic reticulum (SR) in muscle cells and acts in the regulation of striated muscle contraction. The actual free concentration of Ca^{2+} in the SR is about 390 μM compared to 0.1-0.2 μM in the cytoplasm. On muscle excitation, Ca^{2+} floods out into the cytoplasm but must be actively transported back into the SR by the Ca^{2+} pump.

Here is the simplified chemical reaction: ATP + Ca^{2+}(cyto) + H_2O → ADP + Ca^{2+}(s. rect) + H^+ + P_i.

We'll stick with a cartoon representation of each of the various structure states of this calcium pump, as illustrated in Figure 11.5.13 (also from Zhang et al).

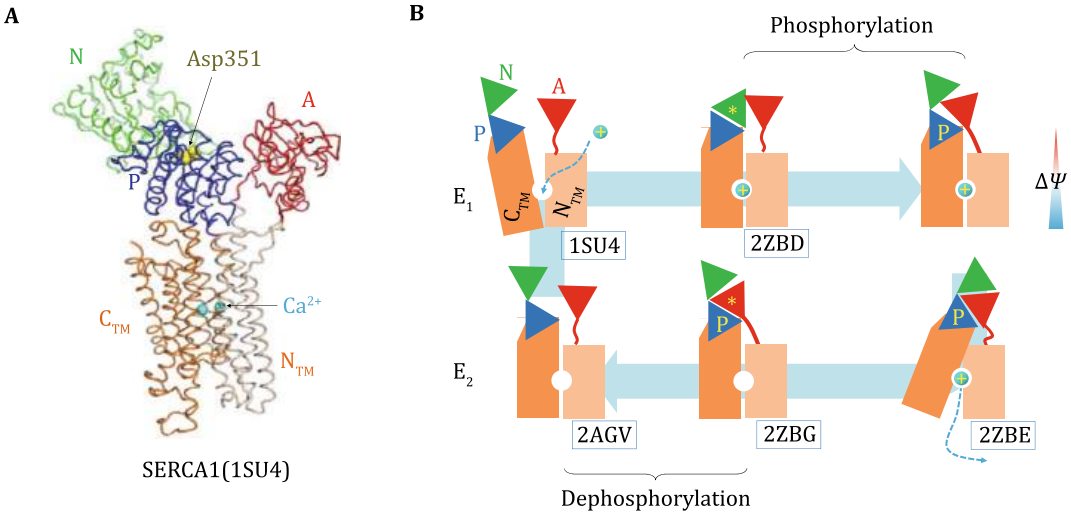


Figure 11.5.13: States of this calcium pump

Again, different states of E_1, the kinase state, are shown in the top line, while E_2, the phosphatase state, is shown in the bottom line. PDB IDs are shown for each state. Ca^{2+} is shown as a cyan sphere. Note how the cytoplasmic N (nucleotide-binding), P (phosphorylation), and A (actuator) domains and the single transmembrane domain, colored to show the C-terminal (C_{TM}) and N-terminal (N_{TM}). are represented in the cartoon version to the right. P represents the phosphorylated Asp 351 mixed anhydride. The asterisk * denotes a transition state for the phosphorylation of Asp 351 (top line) and its dephosphorylation (bottom line).

Ca^{2+} moves from the cytoplasm to the lumen of the SR. Note that the first E_1 state is open inward (to the cytoplasm) while the first E_2 state is open outward (to the lumen of the Sr, where the bulk of the Ca^{2+} ions are stored. This mimics the open-in and open-out states of the Major Facilitator Superfamily (MFS) channels discussed above.

Detailed Thermodynamics

Zhang et al provide two different views of the thermodynamics of Ca^{2+} ion pumping into the SR as shown in A (domain dissociation) and B (classic) in Figure 11.5.14 A quick view of the two "energy landscapes" show:

- Both go from higher to lower free energy as expected since ATP hydrolysis powers the process.
- They differ mostly in the energetic coupling steps

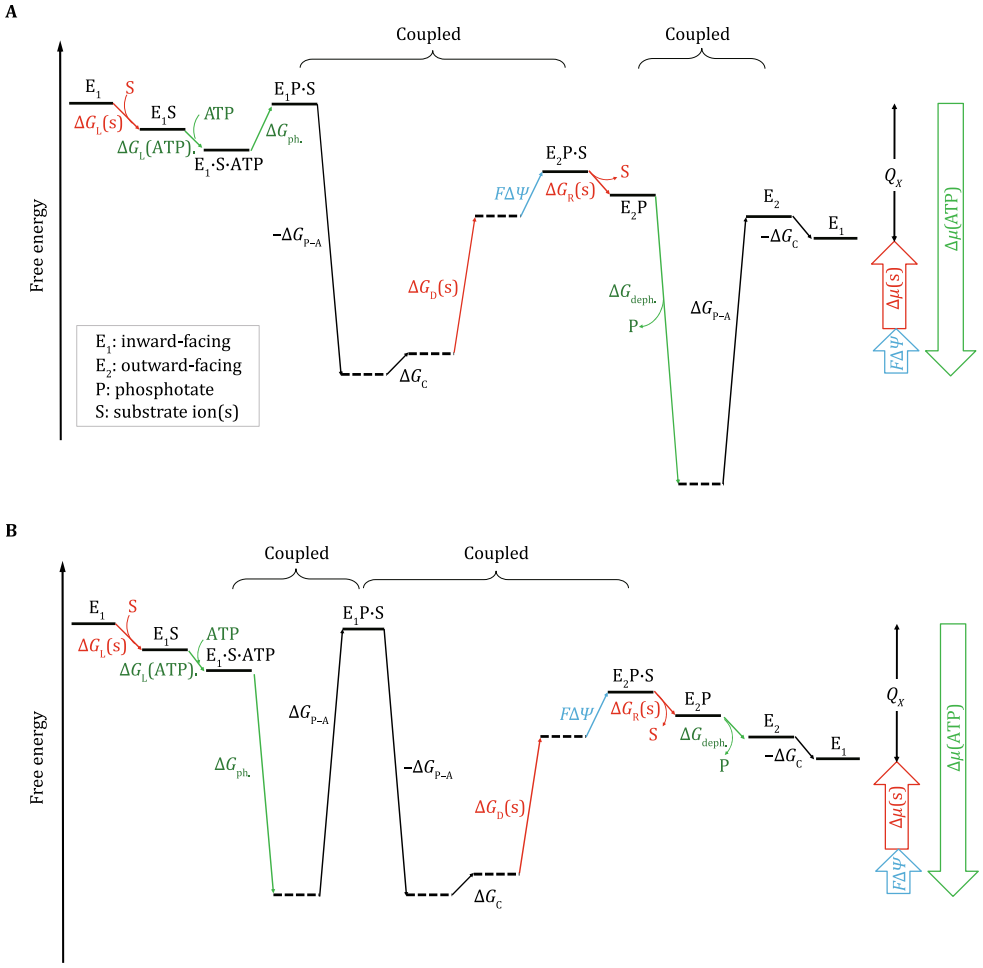


Figure 11.5.14: Thermodynamics of Ca^{2+} ion pumping into the SR as shown in A (domain dissociation) and B (classic)

In A, the actual energy released in ATP hydrolysis occurs when the Asp 351-P_i mixed anhydride is dephosphorylated (ΔG_{deph}), which is coupled to the dissociation of the P and A (** actuator) domains. This is shown in the structural cartoon near the last step in the complete cycle.

In the cartoon diagram, 6 states are represented. Any mechanism can be broken down into more and more states on a more complete description of the actual mechanism. For example, ATP cleavage to ATP can be represented by three states, ATP, ADP, and P_i, or many more if intermediates and transition states are included, the complete Ca^{2+} transport cycle in the thermodynamics diagram is broken down into 12 states, represented by 8 horizontal solid lines —— and 4 dashed ------ lines. Vertical lines with arrows show transitions between states.

- Green arrows → show changes in free energy due to the relative stability (chemical potential) of ATP.
- Red arrows → show changes in free energy due to the relative stability (chemical potential) of the substrate ions
- Cyan arrows → show changes in free energy due to the electrical potential of the substrate ions

ΔG_L are for steps involving loading of reactants, while ΔG_R are for release. In both diagrams, the conversion of state E_1·S·ATP to E_1P·S is uphill as we predicted for the formation of the

On the far right of each graph are identical sets of large unfilled arrows (since ΔG is a state function) as the beginning and ending free energies of free E_1 must be pathway (i.e. mechanism) independent. The two sets of upward unfilled arrows (red for the change in the chemical potential of the substrate ions and cyan for the change in the electric potential of the substrates are both positive as the ions move uphill from a lower to high electrochemical potential. The unfilled downward green arrow shows the change in free energy (or chemical potential) for ATP hydrolysis. Since energy cannot be created or destroyed, the narrow black arrow (Q_x) represents energy that is dissipated in the reaction cycle. The starting and ending states are identical (*i.e.*, E_1), only being differed by the energy dissipation (denoted as Q_X) of the P-ATPase transporter during one functional cycle.

The are other types of membrane ATPase whose structure and function differ from the P-type described above. Some run in reverse to synthesize ATP. They also vary in substrate ions. Here are several different types.

- F-ATPases (ATP synthases, F_1F_0-ATPases): These are used to synthesize ATP and are powered by the collapse of a H^+ gradient. They are found n mitochondria, chloroplasts, and bacterial cell membranes. We will discuss these more in chapters on mitochondrial oxidative phosphorylation and photosynthesis in chloroplasts.
- V-ATPases (V_1V_0-ATPases): These are used to pump protons into organelles to acidify them (ex. lysosome) and are also found in bacteria.
- A-ATPases (A_1A_0-ATPases): These are used to synthesize ATP in Archaea.
- E-ATPases: These are found on the cell surface and hydrolyze extracellular nucleotide triphosphate.

Let's now look at some special features of one P-type ATPase that transports both Na^+ and K^+ ions and is important in establishing their intracellular and extracellular concentration in neurons, so they are critical for neuron function.

Na^+/K^+ ATPase

This protein keeps the K^+_{in} and Na^+_{out} high compared to their respective concentrations on the other side of the neural cell membranes. ATP and **3** Na^+ ions bind to the cytoplasmic domain of the enzyme in the E_1 conformation. As described more generally above, in the presence of Na ions, the bound ATP is cleaved in a nucleophilic attack by an Asp side chain of the protein. Hence, the protein is a Na^+-activated ATPase or kinase. The phosphorylated enzyme changes conformation to the E_2 form in which Na^+ ions are now on the outside of the cell membrane, from which they dissociate. The phosphorylated protein in conformation E_2 now binds 2 $\mathbf{K}^+$ ions on the outside, which activates hydrolysis of the Asp-PO_3 mixed anhydride link. The dephosphorylated protein is more stable in the E1 conformation, to which it changes as it brings K^+ ions into the cell. Hence this protein is an electrogenic antiporter. P-Type transporters are inhibited by vanadate (VO_4^{3-}), a transition state analog of phosphate. Transport mediated by P-type membrane proteins can, in the lab, be used to drive ATP synthesis.

Detailed kinetic analysis of ATP and VO_4^{3-} interactions show there are low-affinity and high-affinity sites for each on Na/K ATPase. The high-affinity vanadate site appears to be the same as the low-affinity ATP site, which suggests that vanadate binds tightly to the E2 form of the enzyme. The low-affinity vanadate site appears to be the same (based on competition assays) as the ATP site, which is probably the E_1 form. Hence vanadate binds preferentially to the E2 form which would inhibit the transition to the E_1 form. Vanadate also inhibits phosphatases, enzymes that cleave phosphorylated Ser, Thr, and Tyr phosphoesters in proteins. This supports the notion that vanadate binds preferentially to the E2 form, which has a phosphoanhydride link (Asp-O-phosphate) that is hydrolyzed, promoting the conversion of E_2 back to E_1. Vanadate is probably a transition state analog inhibitor in that it can readily adopt a trigonal bipyramidal structure, mimicking the transition state for cleavage of the tetrahedral anhydride bonds of ATP and Asp-O-PO_3.

Figure 11.5.15is a YouTube animation of the Na^+/K^+ ATPase from xx. Permission Question?

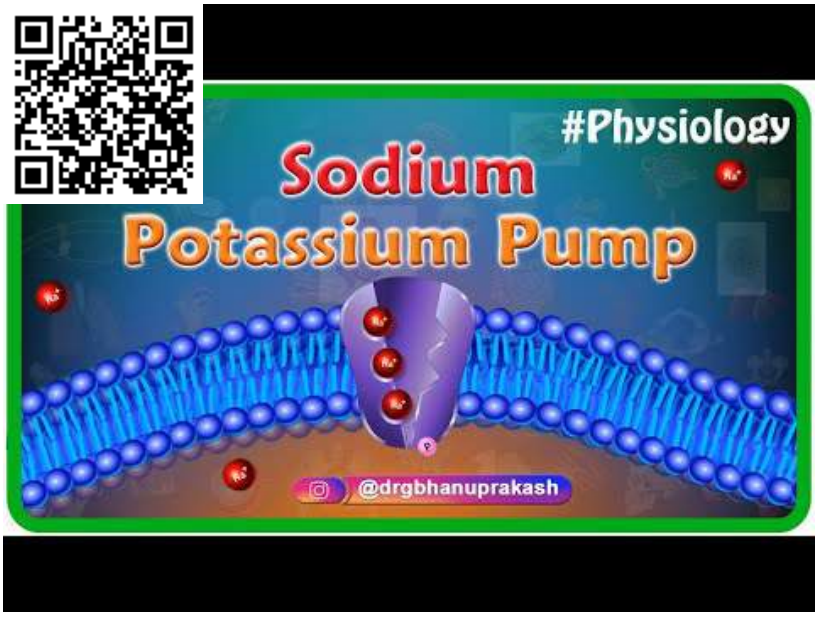


Figure 11.5.15: Animation of the Na^+/K^+ ATPase pump

These interactions are depicted in Figure 11.5.16 below (after Stryer 4th ed)

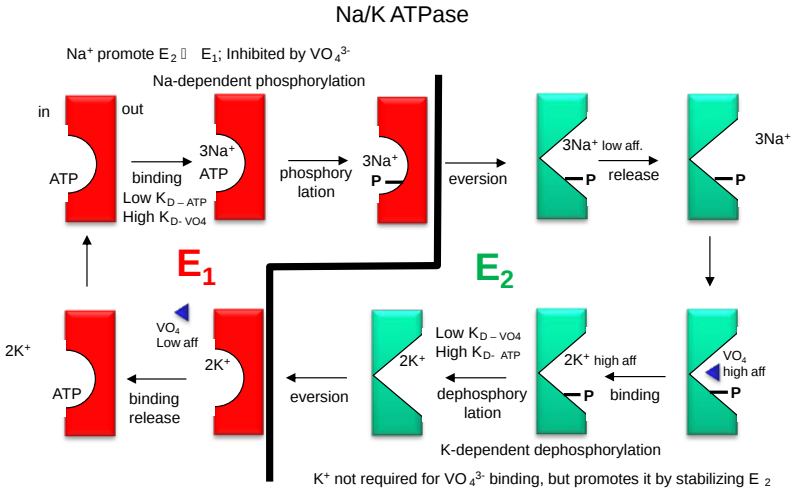


Figure 11.5.16: Steps of the Na^+/K^+ ATPase pump

ABC Transporters

The proteins comprise one of the largest families of membrane proteins with seven main families (ABCA to ABCG). Up to 3% of all bacterial proteins encode proteins associated with the ABC transporters. Different ABC transporters move a variety of required chemical species, from small (ions, sugars, amino acids, nucleosides, vitamins) to large (peptides, lipids, oligonucleotides, and polysaccharides) into the cell. They also remove toxic (to the cell) molecules such as xenobiotics (molecules foreign to the cell like drugs, toxins, etc) and potentially toxic metabolites. All of these require ATP hydrolysis. Moving toxic molecules out of the cell is beneficial to the health of the cell, but in the case of a tumor cell, not to the benefit of the organism. TAs in other active transporters using ATP, its hydrolysis leads to two different conformations, open-outward and open-inward. All ABC transporters have a LSGGQ amino acid sequence in the NBD. They also have a phosphate-binding loop (P-loop or Walker A motif). Most eukaryotic ABC transporters move solutes from the inside to the outside of the cell.

The various structures within the ABC transporter superfamily are shown in Figure 11.5.17. The transmembrane domain (TMD, in green) and the nucleotide-binding domain (NBD, blue) are present in most versions of the gene. PK represents prokaryotic and EK eukaryotic organisms (shown at the very right for each).

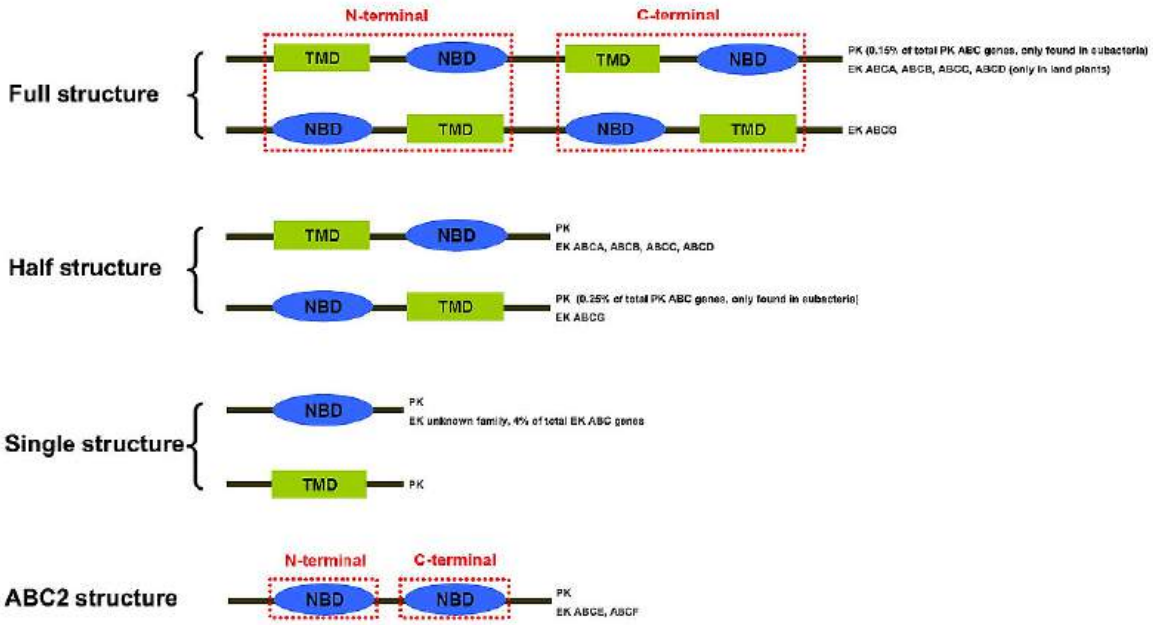


Figure 11.5.17: Variations in the structures within the ABC transporter superfamily ng, J., Feng, J., Yuan, D., *et al.* Tracing the structural evolution of eukaryotic ATP binding cassette transporter superfamily. *Sci Rep* **5,** 16724 (2015). https://doi.org/10.1038/srep16724. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/

The ABC transporter genes denoted as full structures have 2 TMDs and 2 NMDs while half structures have one of each. Some of the genes encode either a single NBD or a single TMD (prevalent in prokaryotes, along with half structures). The ABC2 structure has only two NBDs. The single structure represents the ABC transporter gene with a single TMD or NBD; ABC2 structure represents the ABC transporter gene with only two NBDs. The families possessing certain structures are listed on the right. TMDs typically have 6-10 transmembrane α-helices. Those that are involved in the export of chemical species have six.

Structural cartoons representing domain and ligand binding for a variety of ABC transporters are shown in Figure 11.5.18below.

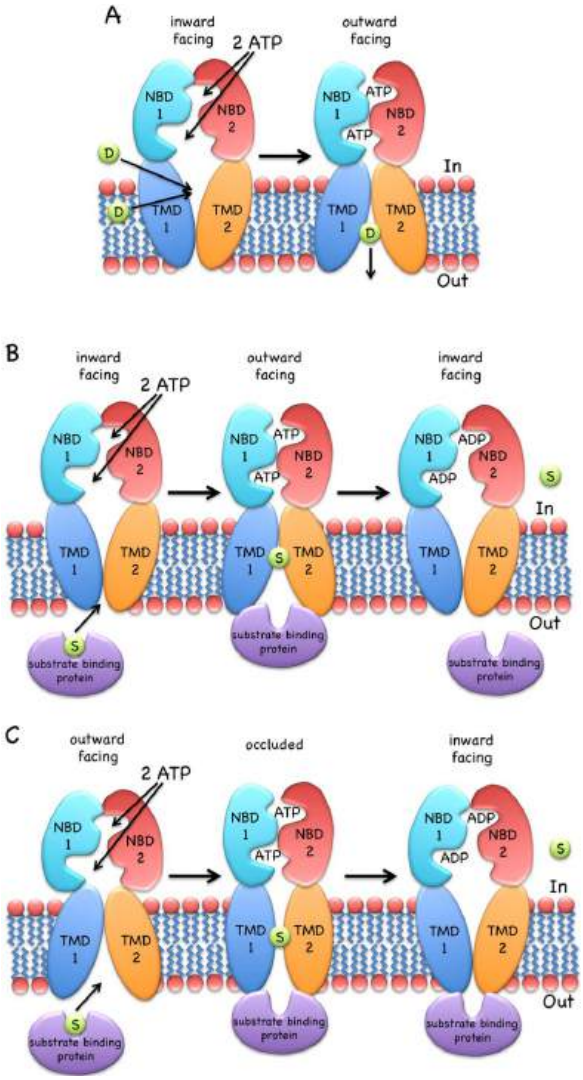


Figure 11.5.18: Domain and ligand binding for a variety of ABC transporters Structure and mechanism of ABC transporters. Stephan Wilkens. F1000Prime Reports 2015, 7:14 (doi:10.12703/P7-14). Creative Commons Attribution-Non Commercial License (http://creativecommons.org/licenses/.../3.0/legalcode).

(A) shows an inward-open drug (D) exporter. Binding of 2 ATPs cause dimer formation between the two NBD domains, resulting in a conformation change to outward facing which allows dissociation of the drug. Hydrolysis of ATP enables the release of ADP/P_i and dissociation of the 2 NBD domain, which results in conformational change to the inward-open form.

(B) shows the delivery of a substrate (S) from its binding protein (ex. part of the ABC transporter complex MalEFGK involved in maltose/maltodextrin import) in the periplasm of E. Coli for delivery into the cell.

(C) shows an outward-open variant of (B) (ex. part of the Vitamin B12 import system permease protein BtuC)

The mammalian protein multi-drug resistance (MDR), also known as P-glycoprotein, Phospholipid transporter ABCB1 or ATP-dependent translocase ABCB1**,** is an example of an ABC Transporter. It moves both drugs across the membrane as well as phospholipids, including phosphatidylcholine, phosphatidylethanolamine, ceramides, and sphingomyelins, from the inner to the outer leaflet of the membrane. The gene for this protein is often mutated in tumor cells which moves chemotherapeutic drugs from the cell.

Figure 11.5.19 shows an interactive iCn3D model of the structure of the mouse P-glycoprotein (4M2T) bound to a cyclic peptide inhibitor in the inward open conformation is shown below. Most proteins bind substrates specifically but P-glycoprotein binds them quite indiscriminately which makes this protein so useful in pumping drugs out of the cell.

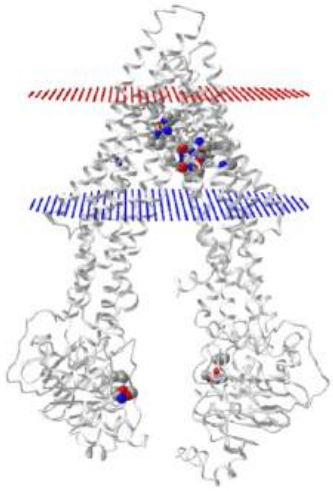


Figure 11.5.19: Mouse P-glycoprotein bound to a cyclic peptide inhibitor in the inward open conformation (4M2T). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...1bHUEmryErK5PA

The iCn3D model below shows the conserved (mouse and human) transmembrane domain aromatic amino acids (H60, F71, T114, F299, Y303, Y306, F332, F339, F724, F728, F766, F938, Y949, F953, F974, F979) involved in the transport pathway in colored sticks. The inhibitor (cyclic-tris-(S)-valineselenazole; QZ59-SSS) is shown in spaceiflll. The consensus LSGGQ (527-531 and 1172-1176) sequences in the nucleotide-binding domain are shown in colored spheres.

Here is a view showing just the transmembrane domain with the conserved aromatic amino acids again.

Figure 11.5.20 shows an interactive iCn3D model of just the transmembrane domain with the conserved aromatic amino acids of the mouse P-glycoprotein bound to a cyclic peptide inhibitor in the inward open conformation (4M2T).

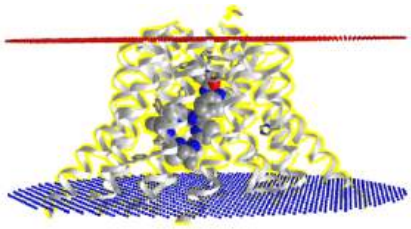


Figure 11.5.20: Transmembrane domain of the mouse P-glycoprotein bound to a cyclic peptide inhibitor in the inward open conformation (4M2T). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...PdHtggbLFWuCo7

11.5.4: Another look at the Nuclear Pore Complex

We examined the incredibly complicated structure of the nuclear pore in the previous section. Both small molecules and large proteins (synthesized in the cytoplasm) and RNAs (synthesized in the nucleus) move through its large pore. The movement of solutes through its gaping pore does not require either ATP cleavage in a primary active transport or the collapse of a chemical gradient in a secondary active transport process. Then why study it in this chapter section? It turns out that the regulation of the process requires **GTP** cleavage and a gradient of a particular protein called **Ran**.

What's fascinating about the nuclear pore is its selectivity for protein transfer across the pore. What proteins are allowed in and out? What is the origin of the specificity? The specificity is determined in part by a protein family called the karyopherin-βs (22) of receptors, which bind and transport nuclear proteins. There are two types of these nuclear **transport receptors or transport factors:**

- **Importins** facilitate the movement of proteins into the nucleus (ex. nuclear proteins like histones, DNA and RNA polymerases, etc).
- **Exportins** facilitate the movement of RNA (except mRNA) out of the nucleus.

Large proteins destined to be moved through the nuclear pore are called **cargo** proteins. They have a molecular signal that differentiates them from molecules that should stay in the cytoplasm or move into organelles or be secreted from the cell. The signals are called:

- **NLS** or nuclear localization sequences
- **NES** or nuclear export sequences

There is no obvious NLS consequence and different motifs are used by different cargos. A classical NLS is enriched in Lys and Arg residue. Others appear enriched in Pro-Tyr or Ile-Lys. Large domain structures might also participate in the NLSs. Prediction programs are used to identify clusters of lysines and arginines with gaps between the clusters. Classical NESs appear to be rich in leucines in an 8-15 amino acid sequence with regularly spaced conserved hydrophobic amino acids.

A third player, **RAN** (Ras-related nuclear protein), is involved that determines which way the transport receptor:cargo complex moves. Ran is a small protein that binds and can hydrolyze GTP to GDP. It is a member of the small G proteins that are GTPase.

Ran itself can run (a pun) or move across the nuclear membrane and is found in both the cytoplasm and nucleus. What determines which way a Cargo:receptor complex goes? It depends on whether GTP or GDP is bound to RAN! There are higher concentrations of RanGTP in the nucleus and lower concentrations in the cytoplasm.

- Importins bind cargo proteins with a NLS in the cytoplasm where RanGTP is low, move into the nucleus, and release the cargo protein when the abundant RanGTP displaces the bound cargo protein
- Exportins bind both a cargo protein with a NES and Ran:GTP to form a ternary Exportin:Cargo:RanGTP complex in the nucleus. This moves into the cytoplasm, where the Ran bound GTP is hydrolyzed to GDP, causing the complex to dissociate, freeing the exported cargo protein from the complex.

Importins are dimeric structures consisting of an alpha and beta subunit. The alpha subunit binds to the cargo protein through the NLS sequence on the cargo protein. The beta subunit interacts with the nucleoporin proteins (Nups) in the

These interactions are shown in Figure 11.5.21

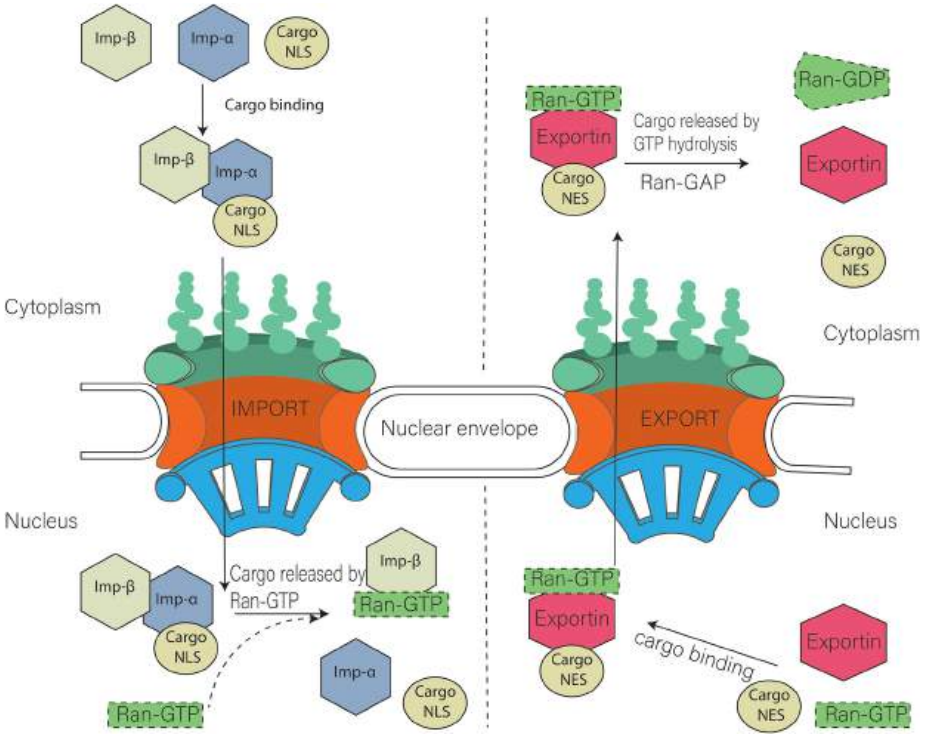


Figure 11.5.21: Model of nuclear import and export. Cargo containing NLS (Nuclear localization signal) is imported with the help of Importin α and importin β heterodimer. Nuclear export of cargo having NES (nuclear export signal) is carried out with the help of exportins. Ran GTP is also required during that process. Khan Asmat Ullah, Qu Rongmei, Ouyang Jun, Dai Jingxing. Front. Physiol., 03 April 2020 | https://doi.org/10.3389/fphys.2020.00239. **Creative Commons Attribution License (CC BY)**.

Figure 11.5.22 shows an interactive iCn3D model of the Delpi electrostatic surface potential map of the putative NLS from the carboxy-terminal of a cargo protein, the W protein of the Hendra virus (4M2T)which is imported into the nucleus.

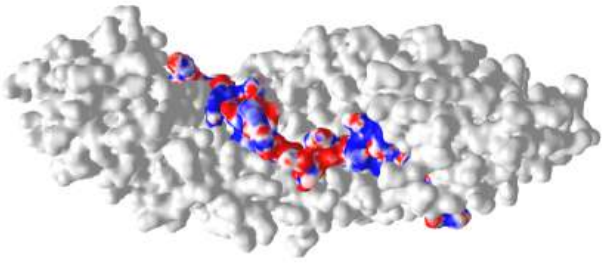


Figure 11.5.22: Delphi electrostatic surface potential map of the putative NLS from the carboxy-terminal of a cargo protein, the W protein of the Hendra virus (6BW9) which is imported into the nucleus (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...mUkDXG1KoiieTA

The virus derives from bats and has recently emerged. The peptide sequence containing the NLS has the following sequence: 419 CLGRRVVQPGMFADYPPTKKARVLLR 444. The red surface indicates the negative potential and the blue positive potential, which is associated with the Lys and Arg side chains in the consensus sequence. It is shown bound to the human importin-α3 subunit (molecular surface shown in white), which binds the NLS sequence with high affinity.

An astute reader might ask why GTP stays bound and is not cleaved into GDP in the nucleus by the intrinsic GTPase activity of Ran. It turns out there is yet another protein found only in the cytoplasm (i.e. it doesn't have a nuclear import signal), which binds to Ran:GTP and promotes GTP → GDP exchange. So oddly, hydrolysis of a nucleotide triphosphate (GTP, not ATP as in the case of most transporters discussed above) is required for nuclear import. A different but very fascinating mechanism!

The cargo complexes described above must still pass through the nuclear pore. In section 11.3 we discussed the structure of the nuclear pore which consists of many nucleoporin proteins (NUPs). The **inner ring** is called the **FG Nups layer.** The disordered parts of the inner ring Nups have repetitive sequences enriched in phenylalanine and glycine (hence the name FG Nups) and stick out into the central pore. The FG repeats can take multiple forms including Phe-Gly (FG), Gly-Leu-Phe-Gly (GLFG), or Phe-any-Phe-Gly (FXFG). These are intrinsically disordered proteins and the disordered regions in the pore act as a filter allowing certain molecules to pass and excluding those greater with molecular weights greater than 40K. Large molecules complexed to importins/exportins move through the pore. This disordered mesh prevents passive diffusion of molecules through it but allows protein complexes with cargo:exportin/importin complex through. It's a bit like electrophoresis of small protein complexes through the pores of an acrylamide or agarose gel polymerized matrix, only without the "push" of an electric field. Presumably, the FG-Nups make transient hydrophobic (induced-dipole:induced dipole) interactions between the FGs on the Nups and the nuclear transport receptors.

Figure 11.5.23 shows an **export** complex (5XOJ) from yeast of Ran (cyan with bound GTP in spacefill) bound to an **exportin** (Xpo1p, white surface) and 3 Nup42p peptides containing SxFG/PxFG repeats (spacefill, side chains, which are mostly nonpolar). To bind 3 Nup peptides, the exportin Xpo1 also contains repeating binding sites on domains called HEAT repeats 14–20 of Xpo1p, The exportin Xpo1p is shaped like a toroid with 21 HEAT repeat domains that have two antiparallel sheets with connecting loops of different sizes.

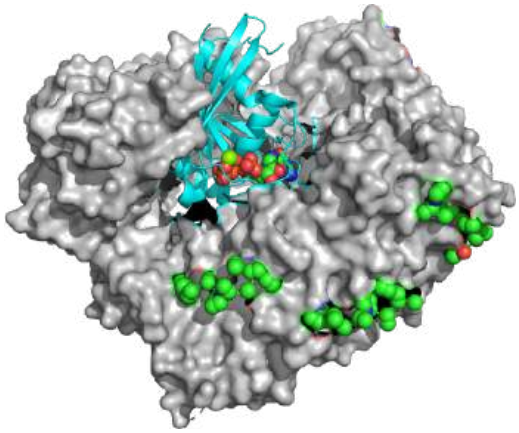


Figure 11.5.23: Yeast export complex (5XOJ) with Ran (cyan with bound GTP in spacefill) bound to an **exportin** (Xpo1p, gray surface) and three Nup42p peptides containing SxFG/PxFG repeats (spacefill, side chains))

Some believe the unfolded mesh of FGs in the core might condense on the binding of motifs on importins/exportins or through FG:FG domain associations. But what is the basis for importins/exportins interactions with the FG structures? To study it in more detail Fragasso et al designed and made an artificial FG-Nup that binds to an importin transport receptor **Kap95** that interacts with a cargo protein with a nuclear localization sequence (NLS) in cargo proteins. They attached them to the inside of solid-state nanopores and demonstrated the fast movement of Kap95 through the derivatized nanopore while a control "non cargo" without a NLS, BSA) was blocked. Underivatized nanometer-sized pores are made in a silicon nitride synthetic membrane (SiN) by using an ion or electron beam to tune the size of the hole.

Figure 11.5.24(top image) shows colored code structures of the inner ring (top of figure) snapshots of different yeast GLFG NUPs. These are named after their Gly-Leu-Phe-Gly repeating motifs and are especially cohesive. The GLFG-Nups, shown in red, are mostly found in the inner ring compared to FxFG/FG-Nups shown in green.

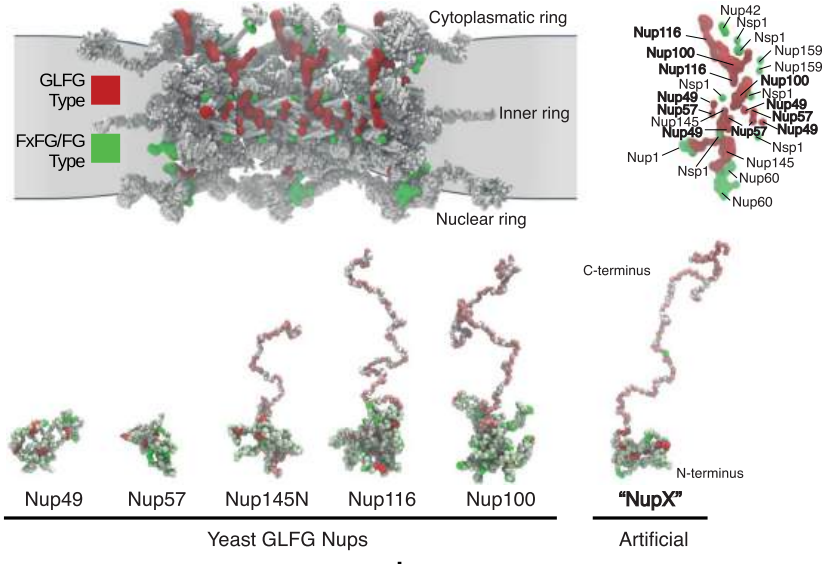


Figure 11.5.24: Inner ring (top of figure) snapshots of different yeast GLFG nups. Fragasso, A., de Vries, H.W., Andersson, J., *et al.* A designer FG-Nup that reconstitutes the selective transport barrier of the nuclear pore complex. *Nat Commun* **12,** 2010 (2021). https://doi.org/10.1038/s41467-021-22293-y. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

The bottom part of Figure 11.5.24shows snapshots of molecular dynamics simulations of various yeast GLFG-Nups compared to the NupX synthetic one. Three of the yeast GLFG Nups and the synthetic NupX show a compact and extended domain Each has a collapsed cohesive domain at one end (characterized by a low charge to hydrophobic amino acid residue ratio (C/H), lots of alternating FG and GLFG repeats, with light green showing non-FG/GLFG/charged residues) and an extended domain at the other end (high ratio of charged/hydrophobic amino acids, no FG repeats, pink-red showing non-FG/GLFG/charged residues). Bright green shows the FG repeats, red the GLFG repeats, and white the charged groups.

Figure 11.5.25shows the pore in the silicon nitride synthetic membrane (SiN) membrane before (left) and after derivitization with NupX.

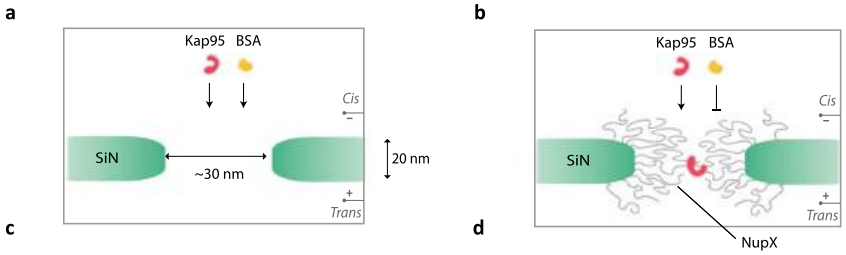


Figure 11.5.25: Pore in the SiN silicon nitride synthetic membrane before (left) and after derivatization with NupX

Figure 11.5.26shows derivatized pores in the SiN membrane of increasing size. In the smallest particle, the NupXs don't fit and are expelled from the pore. In the 30 nm particle, the NupX cohesive domains plug the hole. In the 45 nm pore, a hole appears.

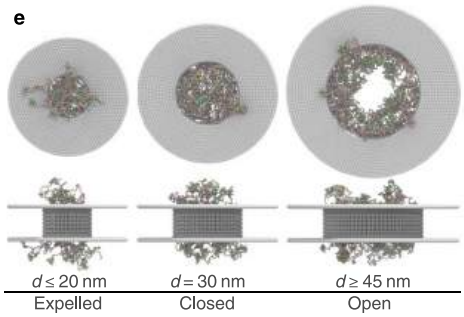


Figure 11.5.26 Derivatized pores in the silicon nitride (SiN) membrane of increasing size

Lastly, Figure 11.5.27 shows snapshots of 30 nm pores lines with NupXs. In panel c, an importin, Kap95 (spheres with orange binding spots) is shown forming transient interaction with and translocating through the NupX-lined hole. BSA (sphere without binding sites) may interact weakly but does not translocate through the pore.

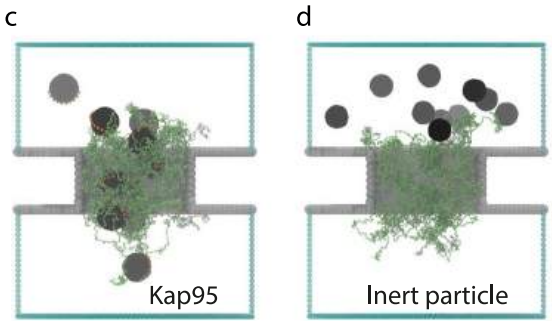


Figure 11.5.27: Snapshots of 30 nm pores lines with NupXs. and interacting with an importin, Kap95 (spheres with orange binding spots, panel C) and BSA (sphere, panel D)

11.4: Diffusion Across a Membrane - Pores

11.4.1: Pores and Pore-Forming Proteins (PFPs)

If you form a pore in a cell bilayer, molecules of all sizes could move either way based on their electrochemical potential. They will move from regions of a higher to lower electrochemical potential in a thermodynamically favorable process. Hence movement through pores represents a special case of facilitated diffusion when part of the driving force is not just a concentration gradient but also an electrical potential. Several questions might come to mind.

- What proteins are involved in pore formation?
- How is the specificity of solute movement through the pore regulated?
- What is the mechanism of pore formation?

Pore formation can lead to cell death, which is the function of some pore-forming proteins (PFPs) including the toxin Hemolysin E (also known as HlyE, ClyA, SheA) secreted from E. Coli and S. Aureus. Human proteins also form a membrane attack complex (examples include the membrane attack complex-perforin/cholesterol-dependent cytolysin (MACPF/CDC) superfamily and the membrane attack complex (MAC). The MAC consists of an assembly of proteins involved in the complement system (part of the effector branch in the innate immune system), which leads to cell death of Gram-negative bacteria like E. Coli. Figure 11.4.1 illustrates the assembly process of the membrane attack complex and the complexity of interactions required to form a lethal pore in a cell.

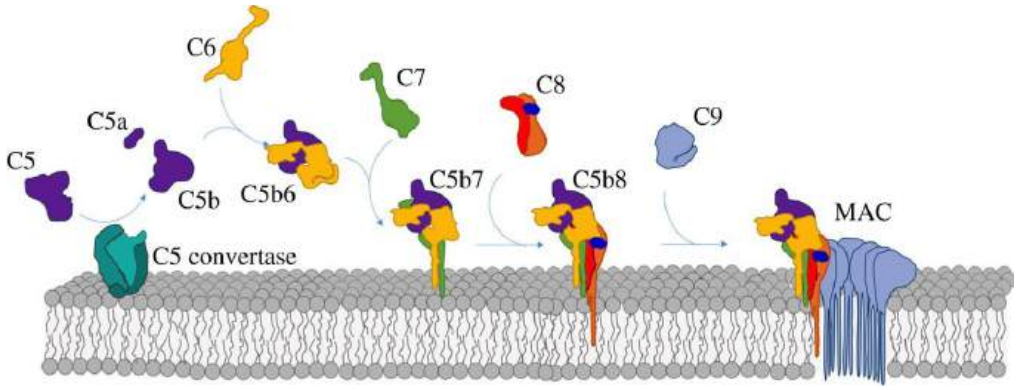


Figure 11.4.1: Human membrane attack complex. Charles Bayly-Jones et al. Philosophical Transactions of the Royal Society B: Biological Sciences, Volume 372, Issue 1726. June 2017 https://doi.org/10.1098/rstb.2016.0221. The Royal Society under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/

Complement protein C9 can adopt a soluble form or membrane form which in aggregates form the pore leading to cell death. This is a common feature of PFPs.

PFPs could create a pore by altering membrane lipid packing to form a toroid-like hole (Figure 11.4.2) and/or by inserting in a membrane and forming a pore within the protein complex itself. In either mechanism, lipid packing is altered.

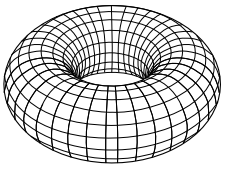


Figure 11.4.2: By GYassineMrabetTalk - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/inde...?curid=3134923

Biophysical evidence shows some support for the formation of "toroidal pores".

Lipid rearrangements in the membrane could lead to a hydrophilic or hydrophobic pore lining as shown in Figure 11.4.3

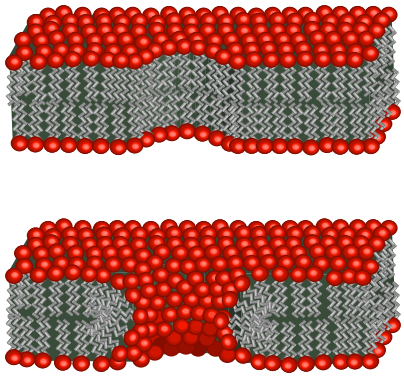


Figure 11.4.3: Diagram of possible lipid arrangements at the edge of a pore through a lipid bilayer. (above) In the absence of any rearrangement, the pore walls will be hydrophobic since the alkane tails are exposed. (below) Some researchers believe that the lipids at the edge re-orient to line the pore wall with hydrophilic head groups. Lipid bilayer. (2023, January 15). In *Wikipedia*. https://en.Wikipedia.org/wiki/Lipid_bilayer

We started our study of lipid bilayers with pure lipid systems and then added membrane proteins. Let's do the same with pore formation. A common technique to form a pore in pure lipid bilayers and also in cells is **electroporation**. This is a technique used to move a DNA with a target gene into either a prokaryotic cell (transformation) or eukaryotic cell (transfection) for exogenous gene expression. In the absence of PFPs, this requires the alteration of surface tension by applying an electrical potential. This forms depressions in the membrane, altering the nonpolar acyl chain packing. In the process, small wire-like columns of water appear, which like hydrophobic pores ultimately rearrange into hydrophilic toroidal pores. Figure 11.4.4 shows snapshots of molecular dynamics simulations as a function of the time of pore formation in electroporation.

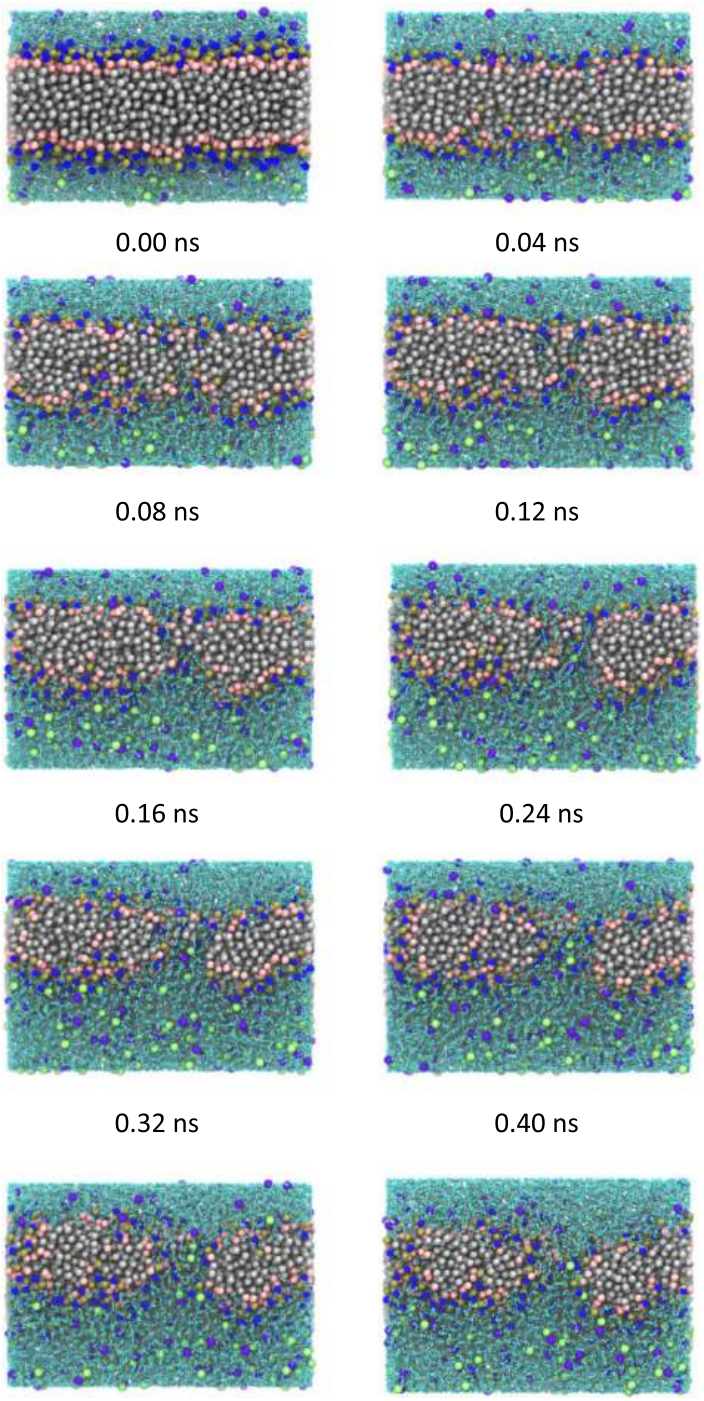


Figure 11.4.4: Zhou, C., Liu, K. Molecular dynamics simulation of reversible electroporation with Martini force field. BioMed Eng OnLine 18, 123 (2019). https://doi.org/10.1186/s12938-019-0743-1. Creative Commons Attribution 4.0 International License, http://creativecommons.org/licenses/by/4.0/

How does DNA pass through the pores in the bilayer? In pure lipid vesicles, it appears to pass through by electrophoresis. Most students are familiar with the electrophoresis of DNA fragments through pores in agarose gels. In actual living cells, small nucleic acids like small interfering RNA (siRNA) and antisense DNA molecules appear to pass through the bilayer by electrophoresis. Large DNAs like plasmids containing a gene for expression bind to the cell and form cell surface aggregates, which appear to be endocytosed into the cell. Electroosmosis, the movement of liquids under the influence of an electric field, also plays a role.

11.4.2: Pores - Outer Membrane Factor (OMF) and Voltage-Dependent Anion Channel (VDAC)

Now let's consider pores made of PFPs. We have already discussed two types of beta-barrel transmembrane proteins, the outer membrane factor (OMF) of Gram-negative bacteria and the voltage-dependent anion channel (VDAC). Both are examples of proteins called **porins** with typical beta-barrel topology.

VDAC: At low membrane potentials, VDAC (also known as mitochondrial porin), the most abundant protein in the mitochondrial outer membrane, moves metabolites and Ca^{2+} ions across the outer membrane of mitochondria. VDAC exist in an open state at 0 or very low transmembrane potentials that allows for the transfer of key metabolic anions (pyruvate, oxaloacetate, malate, succinate, ATP, ADP, and P_i, which are involved in metabolism) and Ca^{2+}, and in a closed state (above or below $\pm$ 30 mV), which is not completely closed as it allows for the transfer of ions with a preference for cations. The closed state does not allow for the transfer of ATP. The transition to the closed state is promoted by tubulin and actin (cytoskeletal proteins), negatively charged lipids such as phosphatidyl ethanolamine and cardiolipin, and also covalent phosphorylation by protein kinases. Bcl2, proteins involved in the regulation of programmed cell death (apoptosis) also interact with VDAC. In contrast to ligand-gated channels which require ligand binding to open the channel, pore complexes are unusually open. Millions of ATPs/second move across the membrane through the open pore but none through the closed pore.

Figure 11.4.5 shows an interactive iCn3D model of mouse VDAC1 (4c69) with bound ATP loosely held in the site. An alpha helix partly occludes the central pore of this β-barrel protein.

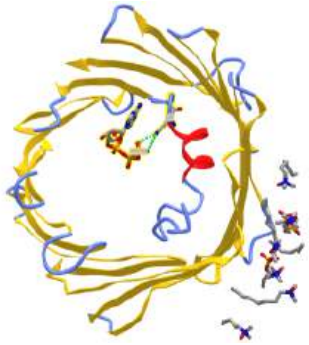


Figure 11.4.5: Mouse voltage-dependent anion channel - VDAC1 (4c69) with bound ATP loosely held in the site. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?Ueb4wg26f5oEVL4Y9

One ATP is bound in the barrel and interacts with Lys 12 and Lys 20 at each end of the cavity-bound helix. The alpha-helix narrows the pore opening and presumably changes orientation in a voltage-sensitive fashion to gate the pore open and closed, hence regulating the conductance of ions through the pore. That the orientation of the charged arginine side chains would be dependent on the transmembrane potential should be somewhat obvious.

Aquaporins

Aquaporins can move a billion water molecules per second across membranes and exclude ions including protons. Waters proceed in a single file through the pore. Instead of moving waters through, it might simply move "H^+ ions" through by the alignment of subsequent H bond donors and acceptors in a "wire" of water molecules. This is prevented by two conserved asparagine residues in the center of the channel, which disrupt water to water hydrogen bond network in the channel waters that could facilitate proton transfer. Instead, the central waters form hydrogen bonds to the central asparagines. This, along with local membrane potentials, causes opposite orientations of the water in different leaflet sides of the membrane, precluding H^+ transfer.

Here is a movie of a molecular dynamics simulation of water moving through a porin called aquaporin, GlpF, from E. Coli.

Science magazine (Tajkhorshid et al., Science Apr 2002, 296:525). Used with permission from the Theoretical and Computational Biophysics Group, the National Institutes of Health (NIH) Resource for Macromolecular Modeling and Bioinformatics, at the Beckman Institute, University of Illinois at Urbana-Champaign.

OMF: The outer membrane factor (see previous section) is one member of a class of bacterial porins, which are the most abundant proteins in the outer membrane of Gram-negative bacteria. They are classified as non-specific or specific (with respect to the solute that passes through), or monomeric, dimeric, or trimeric based on their structure. In Gram-negative bacteria, which have two lipid bilayers, the movement of solute from inside to outside includes at least three sets of proteins. Active transport (discussed in the next section) needs an energy source and is used by inner membrane transport proteins, including ATP-binding cassette (ABC)-type, resistance nodulation division (RND)-type and major facilitator superfamily (MFS)-type transporters. These are connected to membrane fusion proteins (MFP) that span the periplasm, which then interacts with at least 21 different types of porins. Molecules with molecular weights greater than 600 generally can not get through the nuclear envelope of Gram-negative bacteria, limiting the size of potential antibiotics, which must enter by passive diffusion.

11.4.3: Mechanosensitive ion channels - Mscs (which are pores!)

As the name applies, these ion channels (with openings large enough to be called pores) are gated open/closed by mechanical (physical) changes in the properties of the membrane, not extracellular/intracellular ligands or voltage changes. Certain bilayer lipids also activate the Mscs. There are two types, small (MscS) and large (MscL) mechanosensitive ion channels. Changes in local (boundary layer) and nonlocal lipids are involved in the gating of the channel (in the next section we will discuss lipid-gated ion channels). They are found in prokaryotes, archaea ,and eukaryotes. They are also called stretch-gated ion channels.

Mcss transduces a physical force (stretching and change in turgor pressure) into an electric signal - a flow of ions across the membrane. Turgor pressure is the internal pressure that "presses" the cytoplasm and cell membrane towards the cell wall in bacteria and plant cells. It arises mostly from the osmotic flow of water into the cell. If bacterial cells are placed in a high salt concentration solution (hypertonic), water flows out of the cell and the cell membrane shrinks to the inside of the volume confined by the cell wall. When placed in a hypotonic solution, water flows into the cell and the cell membrane presses out to the cell wall. The response of these channels is fast, in the millisecond range, which is about as quick as a cellular response can be. Some variants of these channels are called **piezochannels**, based on the piezoelectric effect that describes how a voltage is produced when some materials are deformed by mechanical stress which causes a redistribution of charges.

Bacteria normally have high concentrations of both K^+ and negatively charged anions, especially glutamate, which leads to a high turgor pressure from the inward osmotic flow of water. At low external osmolarity, turgor pressure in the cell could be as high as 4 atm. If placed in external solutions of high osmolarity, bacterial cells respond by the increased movement of solutes into the cell.

Mscs are particularly important when bacteria are subjected to sudden osmotic shock. If they are placed in pure water, for example, water would flow down a concentration gradient into the cell and cause the cell to swell and lyse, killing the cell. This is done in the lab to prepare almost pure hemoglobin from ruptured red blood cells. The Mscs are opened under these conditions and small species from the cytoplasm flow out, helping to keep the cell viable. Their openings must be regulated to prevent too much outward flow, which would kill the cell.

In other organisms they are also involved in touch (stretch), hearing (vibration, sound waves), and responses to gravity. Stimuli to activate them include fluid shear stress (relevant to endothelial cells that line blood vessels), membrane stretch (relevant to skeletal and cardiac muscle cells), or even indentation of a bilayer with a pipette. Changes in transmembrane turgor or other mechanical pressures cause membrane tension. Yet even in the absence of these changes, Mscs can be activated by anesthetics, phospholipids missing one fatty acid (called lysophospholipids), and certain polyunsaturated lipids. These stimuli also perturb the membrane bilayers.

Given the pore size, these proteins are less selective to ion flow compared to voltage-gate channels (see next section). Depending on the amino acids that line the pore, some Mscs would allow the preferential flow of anions while some allow cation flow.

Some somatosensory channels (i.e. not pore) proteins also respond to pressure. When open, these can be selective to specific ions like Na^+ or K^+. Examples include some variants of the Transient Receptor Potential (TRP) ion channels. Other membrane proteins can also be activated by physical force. but true Mscs have some key characteristics. If mutated or deleted, the mechanosensory response is removed. If added to a cell, a mechanosensory response results.

a. Small-conductance mechanosensitive channel

This protein is a homoheptamer with three helices from each monomer involved in the overall structure. Two (helices 1 and 2) interact more with the lipid components of the bilayer. Interactions of specific lipids with the helices seem to promote closing but changes under high pressure lead to the opening of the pore. Half of each helix 3 forms the pore, while the other half is more parallel to the membrane and interacts with a large cytoplasmic domain.

Figure 11.4.6 shows the differences between the closed form of E. Coli MscS (2oau) and the open form (2vv5) viewing down the pore axis. The heptameric protein is shown in gray. Two key valines (105 and 109) on each chain are shown in spacefill and colored cyan. These hydrophobic amino acids act like gate-keepers, helping to keep water out, acting like a "vapor seal". In the closed state, the pore is sealed by the closing of the leucine "rings" as one half of helix 3 pack more closely. Many members of the MscS family vary significantly in size and can have between 3-11 transmembrane regions. The closed pore has a diameter of about 4.8 Å, while the open pore is 13 Å across.

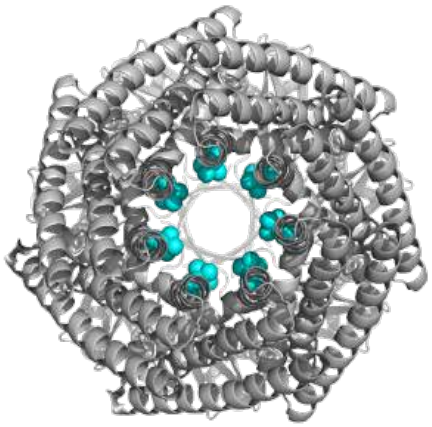


Figure 11.4.6: Differences between the closed form of E. Coli MscS (2oau) and the open form (2vv5) viewing down the pore axis.

Most of you would have studied Ohm's law, given by

$$\mathrm{V} = \mathrm{IR} \text{ or } \mathrm{I} = \frac{\mathrm{V}}{\mathrm{R}} \tag{11.4.1}$$

where I (amps) is the current, R (ohms) is the resistance and V (volts) is the voltage. A more general variant of this law used in physics is

$$\mathrm{J} = \sigma \mathrm{E} \tag{11.4.2}$$

where J is the current density, E is the electric field, and σ is the conductivity (inverse of resistance) which depends on the material. The unit of sigma σ is $ohms^{-1}$ or mhos (ohm written backward). That unit has been renamed the Siemen (S). Mscs channels have a small conductance of approximately 1 nS in 400 mM salt solution.

b. Large conductance mechanosensitive ion channel (MscL):

The channels have large conductances (3 nS) and concomitantly larger pore sizes, allowing the flow of water, ions, and even small proteins. Again they are involved in diffusion down an electrochemical gradient through pores so not active transport just gated diffusion.

Figure 11.4.7 shows an interactive iCn3D model of the pentameric MscL from Mycobacterium tuberculosis (2oar), a Gram-positive bacteria that causes tuberculosis. About 23% of the world's population is affected by this pathogen. It causes about 1.5 million deaths each year. Compare this to the total number of deaths during the COVID pandemic (2020-21) of over 3 million (as of May 2021). It has killed over 1 billion people throughout human history (but not as many as malaria). The pore diameter is about 30 Å across.

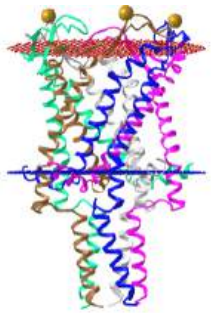


Figure 11.4.7: Closed state of the pentameric MscL from Mycobacterium tuberculosis (2oar). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...LsRZZbmFy5NRF6

11.4.4: Pore-forming alpha-helical toxins

We started this chapter section with a discussion of the major attack complex (MAC) of the innate immune system. Pathogens also employ pore formation to kill host cells. Many secrete soluble proteins which aggregate in the membrane to form either alpha-helical or beta-barrel pores. The proteins are called pore-forming toxins (PFTs). Killing occurs when either cytoplasmic components leak out or bacterial toxins, such as diphtheria and anthrax toxin, move into the cell.

Figure 11.4.8 shows an interactive iCn3D model of the pore formed by cytolysin A (ClyA, also known as HlyE), an alpha-PFT used by some E. Coi and Salmonella enterica strains. The pore is a large dodecamer that forms from soluble monomeric ClyA (2wcd). It has a pore diameter of about 40 Å,

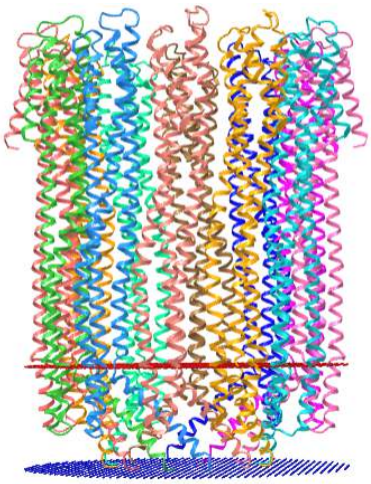


Figure 11.4.8: Pore formed by cytolysin A (ClyA, also known as HlyE). (2wcd) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...RXUsE24d9d8NT6

Figure 11.4.9 shows the soluble monomeric form of cytolysin A (ClyA, HlyE) (1QOY). Nonpolar side chains are shown in cyan. The transparent surface is mostly polar, making the monomer soluble.

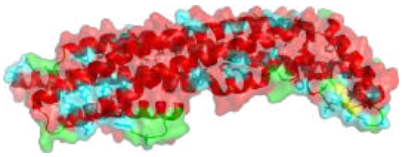


Figure 11.4.9: Soluble monomeric form of cytolysin A (ClyA, HlyE) (1QOY)

Figure 11.4.10 shows the changes in conformation between the single chain soluble form (shown in the figure above) and a single chain of the membrane oligomeric channel.

Figure 11.4.10: Changes in conformation between the single chain soluble form and the same chain in the channel

We started this chapter section by exploring electroporation and the formation of a toroidal-like hole in the lipid bilayer. In the case of ClyA, the protein aggregate itself forms the pore, not the lipids themselves, although lipid rearrangements are necessary to form the protein complex.

11.4.5: Gap Junctions

Connexins are voltage-gated channels that allow for the flow of ions, metabolites, nucleotides, and small peptides. A connexin has four transmembrane helices and two extracellular loops, Six protomers of these come together in a single cell to form a channel complex called a **connexon** or hemichannel. Beta structures in the connexin hemichannel of one cell docks with a similar channel on an adjoining cell to form a full channel passing through the membranes of both cells forming a gap junction between the cells. This is shown in Figure 11.4.11. There are 20 different connexins encoded in the human genome.

The left figure below shows the six protomers, each in a different color and a gray rectangle representing the bilayer of a single connexon or hemichannel. The right side of the figure shows a full gap junction channel between two cells, with the membranes represented by gray rectangles. The connexin 26 monomer was used to create the diagram. Mutations in this protein are associated with hearing loss.

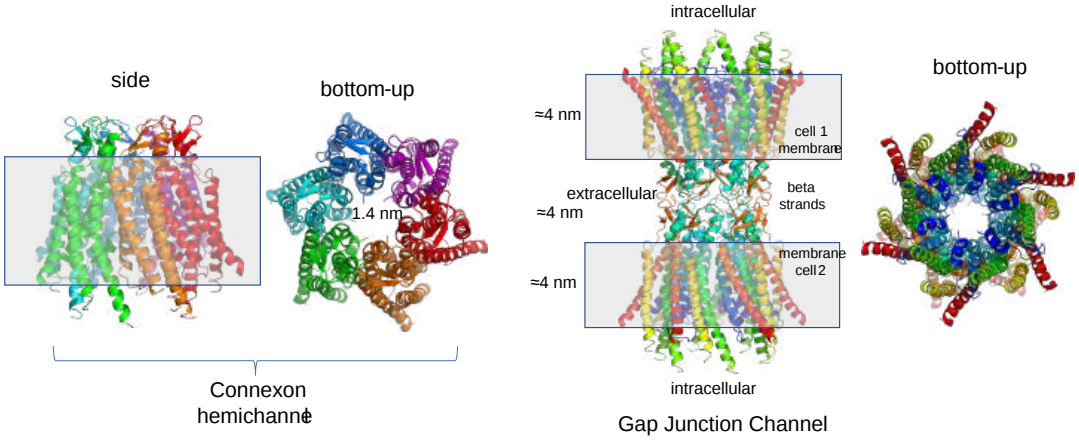


Figure 11.4.11: Connexon channel and the gap junction

The left figure above shows the six protomers, each in a different color and a gray rectangle representing the bilayer of a single connexon or hemichannel. The right side of the figure shows a full gap junction channel between two cells, with the membranes represented by gray rectangles. The connexin 26 monomer was used to create the diagram. Mutations in this protein are associated with hearing loss.

The cytoplasmic entrance of the channel is positively charged. It forms a funnel, composed of 6 amino-terminal helices, which leads into a negatively charged transmembrane lining. The entrance diameter is 14 Å.

Here is a model of a full gap junction channel connecting two membranes using the human connexin 26 monomer (2zw3). Only the bottom membrane bilayer is represented by red and blue dummy atoms.

Figure 11.4.12 shows an interactive iCn3D model of a full gap junction channel connecting two membranes using the human connexin 26 monomer (2zw3). Only the bottom membrane bilayer is represented by red and blue dummy atoms.

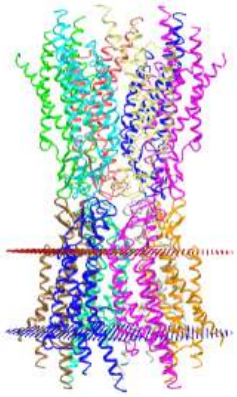


Figure 11.4.12: Gap junction channel connecting two membranes using the human connexin 26 monomer (2zw3). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?TpuQ6Agk8XWuApsWA

11.4.6: The Nuclear Pore Complex (NPC)

Channels have pores that can be gated open and allow the selective flow of ions. Pore-forming proteins have larger entrances that allow both small and large molecules to pass through the bilayer. The pore opening in even large mechanical sensitive channels (about pale in size compared to the nuclear pore complex, which has a combined molecular mass of around 125,000,000! Its outer diameter of ~1,200 Å and its inner one of about 425-Å. Figure 11.4.13 shows the relative size of nuclear pore compared to other molecular structures including the eukaryotic ribosome, nucleosome, a soluble tetrameric protein (rubisco, 270K), and MscL (shown as a circle which represents the pore diameter).

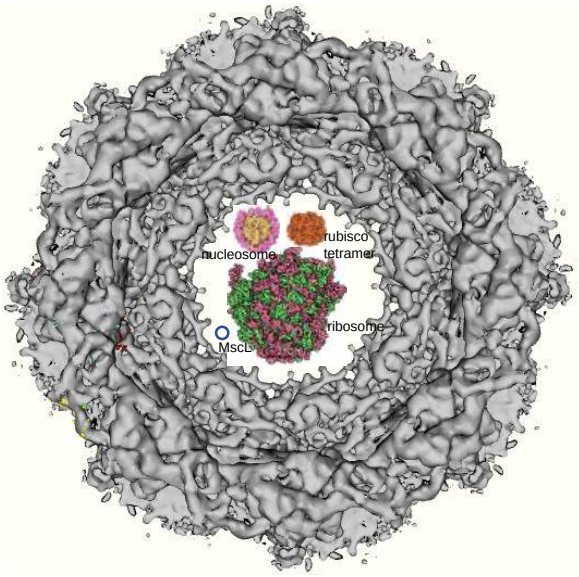


Figure 11.4.13: Relative size of nuclear pore compared to other molecular structures including the eukaryotic ribosome, nucleosome, a soluble tetrameric protein (rubisco, 70K) , and MscL (shown as a circle which rrepresentsthe pore diameter)

Its job is to shuttle small molecules by passive diffusion down a concentration gradient through the pore. In addition, it moves large molecules and molecular structures (proteins, RNA, and perhaps ribosomes) across the nuclear membrane in a process that requires energy. The proteins that comprise this complex are called **nucleoporins (nups)**, of which there appear to be around 34 in humans. Each NPC complex contains around 1000 **nucleoporins**. The complex fuses the inner and outer nuclear membranes.

We have focused so much on single bilayer membranes that comprise the plasma membrane and membranes of organelles like the Golgi complex and lysosomes, it might come as a surprise (perhaps not to biology students) that the nuclear membrane or envelope appears to consist of two bilayers. Most know that the mitochondria have two bilayers, an inner and outer membrane, similar to Gram-negative bacteria. Mitochondrial are believed to have arisen from bacteria so the double bilayer there makes sense. Figure

11.4.14shows the nuclear membrane or envelope of two bilayers (1) with an outer ring (2), spokes (3), a basket (4), and filaments (5). The NPC spans both bilayers.

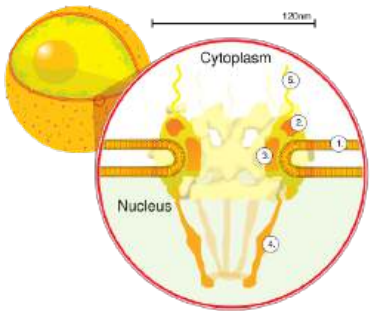


Figure \(\PageIndex{14\): Nuclear membrane or envelope of two bilayers (1) with an outer ring (2), spokes (3), a basket (4) ,and filaments (5) https://en.Wikipedia.org/wiki/File:N...e_crop.svg.png; Creative Commons Attribution-Share Alike 2.5 Generic

The outer bilayer of the nuclear envelope is continuous with the endoplasmic reticulum as shown in Figure \(\PageIndex{15\) below. The dots on the ER membrane are ribosomes, making this the rough ER (as opposed to smooth ER, which has no attached ribosomes).

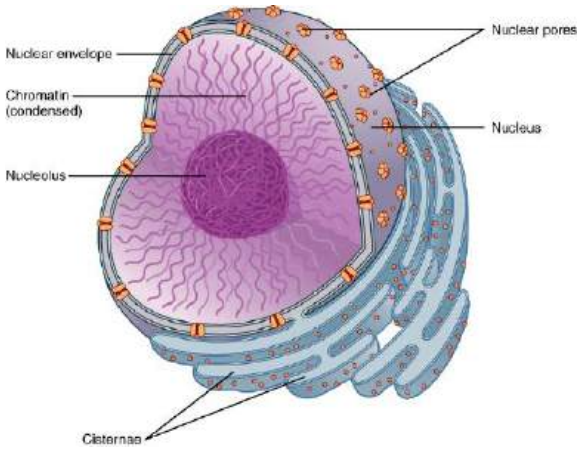


Figure \(\PageIndex{15\): The outer bilayer of the nuclear envelope is continuous with the endoplasmic reticulum https://commons.wikimedia.org/wiki/F...18_Nucleus.jpg; OpenStax, CC BY 4.0 <https://creativecommons.org/licenses/by/4.0>, via Wikimedia Commons

Figure \(\PageIndex{16\) below shows a model of the basket-like structure of the nuclear pore complex (NPC). It, as well as Figure XX, shows that instead of two separate bilayers, in actuality, there is just one bilayer with each leaflet bending around at the NPC and reversing directions! Think of the interesting lipids and protein components that enable the bend! Alternatively, you could say that 2 different membranes fuse at the NPC.

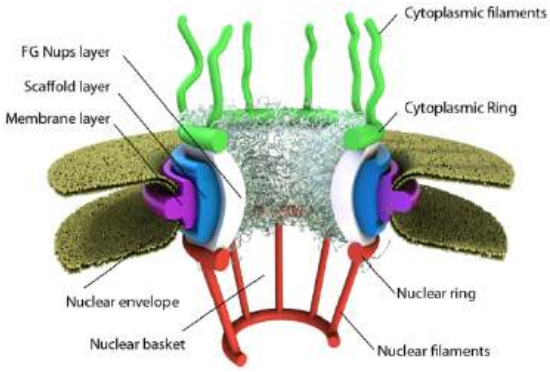


Figure \(\PageIndex{16\): A model of the basket-like structure of the nuclear pore complex **Citation:** Azimi M, Mofrad MRK (2013) Higher Nucleoporin-Importinβ Affinity at the Nuclear Basket Increases Nucleocytoplasmic Import. PLoS ONE 8(11): e81741. https://doi.org/10.1371/journal.pone.0081741. Creative Commons Attribution License,

The NPC consists of 32 copies of each specific nucleoporin (Nup) except two. One of those has 48 copies and the other 16 (even these sum to 2x32 Nups). Three rings form and surround the pore. There is a 16-membered ring of Nups facing the cytoplasm

(**cytoplasmic ring**) and another 16-membered ring of Nups facing the nucleoplasm (**nuclear ring**). There is 8-fold rotational symmetry in each ring, suggesting a dimeric repeat of Nups in the rings. Eight Nups in the cytoplasmic ring have a disordered end that sticks out into the cytoplasm as filaments. In contrast, the disordered ends of eight Nups in the nuclear ring form filaments which bind together to form a ring at the bottom of the nuclear basket.

The **inner ring** (called the **FG Nups layer** in the figure above), between the cytoplasmic and nucleoplasmic rings, also consists of Nups with ordered domains and disordered parts. The disordered parts of the inner ring Nups have repetitive sequences enriched in phenylalanine and glycine (hence the name FG Nups) and stick out into the central pore. These disordered region act as a filter allowing certain molecules to pass and excluding those greater with molecular weights greater than 40K. Large molecules need transport proteins called karyopherin transport factors to move through the pore.

The core structure of the NPC, obtained through cryoelectron microscopy, is shown in Figure \(\PageIndex{17}\) below.

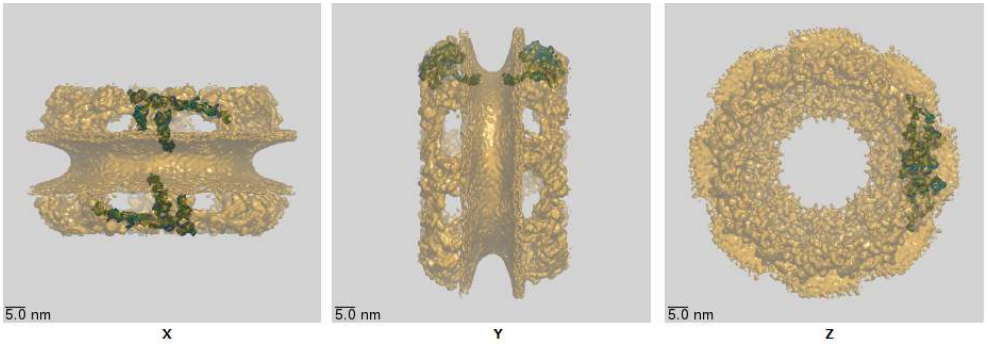


Figure \(\PageIndex{17}\): The core structure of the Nuclear Pore Complex https://www.ebi.ac.uk/pdbe/entry/emd...-3103/analysis

The double bilayers are very evident. The nuclear and cytoplasmic filaments, as well as the disordered FG repetitive sequences that stick into the pore from the inner ring, are not seen since they are very flexible and don't' adopt single conformations using standard structure determination methods.

Figure 11.4.18 shows an interactive iCn3D model of the nuclear pore complex (NPC) core (5a9q), whose structure was obtained using cryoelectron microscopy. (load slowly)

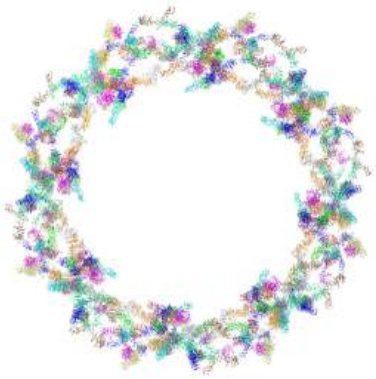


Figure 11.4.18: Nuclear pore complex (NPC) core (5a9q). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?NP7oQb235BCgPARA8

A cartoon figure showing the variety of Nups in the NPC is shown in Figure 11.4.19

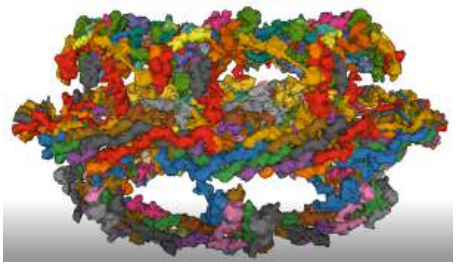
Figure 11.4.19: A cartoon version of the various NUPS in the nuclear pore complex. Donnaloja F., Jacchetti E., Soncini M., Raimondi M. T.Front. Physiol., 12 July 2019 | https://doi.org/10.3389/fphys.2019.00896**Creative Commons Attribution License (CC BY)**.

In this figure, the cytoplasmic and nucleoplasmic rings are both shown in green, each mostly formed by 16 copies of the Y-complex, arranged in two eight-membered rings. The inner ring, predominantly formed by 32 copies of the Nup93 complex is shown in red. Transmembrane nucleoporins are depicted in violet, and the cytoplasmic filaments and nuclear basket structure are in orange. Attached to the inner ring are Nup62 complexes (depicted in blue), which form a cohesive meshwork within the central channel through their FG-repeat domains. Not indicated is the position of Nup98, a FG-repeat-containing nucleoporin important for

the transport and exclusion function of NPCs; its position in the NPC is less defined, but it might be part of the inner ring. Similarly, Aladin (also known as AAAS), Gle1, Rae1, and Npl1 (also known as hCG1 and NUPL2) have been omitted.

Large proteins and RNA that pass through the pore must first be bound to a cargo receptor, which can move the "cargo" across the pore with concomitant GTP hydrolysis. This is a process that is closer to active transport so we will discuss that in Chapter 11.3.

The entire nuclear pore complex was solved in 2022 using cryoEM. Here are two videos showing the dilated complex (7R5J). Click on the images to download mp4 animations of the complex.



CHAPTER OVERVIEW

12: BIOENERGETICS AND BIOCHEMICAL REACTION TYPES

12.1: Biochemical Reactions and Energy Changes
12.2: Phosphoryl Group Transfers and ATP
12.3: The Chemistry and Biochemistry of Dioxygen
12.4: Biological Oxidation-Reduction Reactions

12.1: BIOCHEMICAL REACTIONS AND ENERGY CHANGES

12.1.1: INTRODUCTION

We have already discussed How Enzymes Work and Enzymatic Reaction Mechanisms in great detail in Chapter 6. Here we will focus on a lighter, less granular review of some key reaction mechanisms and the changes in Gibb's free energy associated with them, both in an uncatalyzed and enzyme-catalyzed reaction. Consider it a simple review of very basic organic reactions and thermodynamics in preparation for the comprehensive focus on reaction mechanisms in Unit 2: Bioenergetics and Metabolism.

12.1.2: BREAKING C-X BONDS

Many of the organic reactions involved in metabolism involve making and breaking bonds to carbon. There are three ways to break a bond to a C-X bond, producing either a carbocation, carbanion, or free radical intermediate, all of which are unstable and reactive, as illustrated in Figure 12.1.1. Both the carbocation and free radical are electron deficient, and the carbanion, although not electron deficient, has a negative on C, an atom that has a relatively low electronegativity.

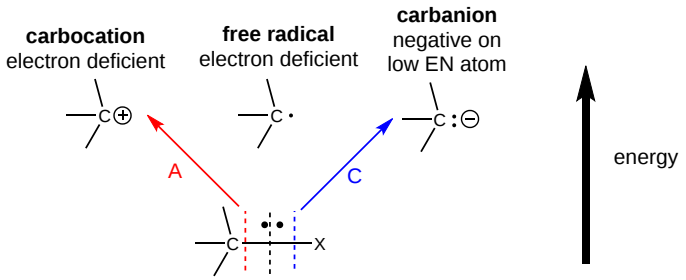


Figure 12.1.1: Ways to break a C-X bond

These unstable intermediates are higher in energy than the reactants, and hence the transition state, which is even higher in energy than the intermediates, must have a structure that resembles the intermediates more than the reactants, as shown in Figure 12.1.2. For the charged carbanion and carbocation intermediates, there is a developing charge in the transition state.

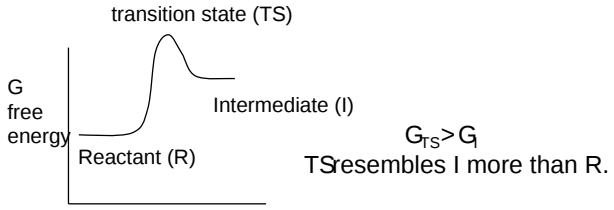


Figure 12.1.2: Gibb's free energy of reactants, transition state, and intermediate in breaking a bond.

The thermodynamics of the reactions is determined by the change in free energy between the intermediates and the reactants, while the kinetics of the reaction is determined by the difference in free energy between the transition states and the reactants, as shown in Figure 12.1.3. A catalyst lowers the energy of the transition state without affecting the energies of the reactants or intermediates (assuming that these are free and not bound to the catalyst.

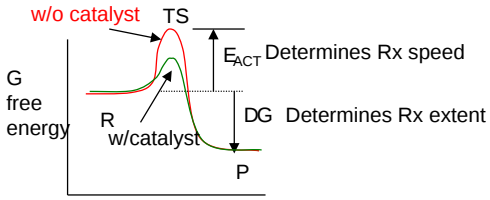


Figure 12.1.3: Activation energy and ΔG for a simple uncatalyzed and catalyzed reaction

The free energy diagram shown in Figure 3 is very simplistic. We need a diagram that better fits an enzyme-catalyzed reaction, using the simple reaction equation below.

$$E + S \leftrightarrow ES \rightarrow EP \leftrightarrow E + P$$

A free energy diagram taking into account the binding of E and S followed by the conversion of bound S to bound and then the free product is shown in Figure 12.1.4.

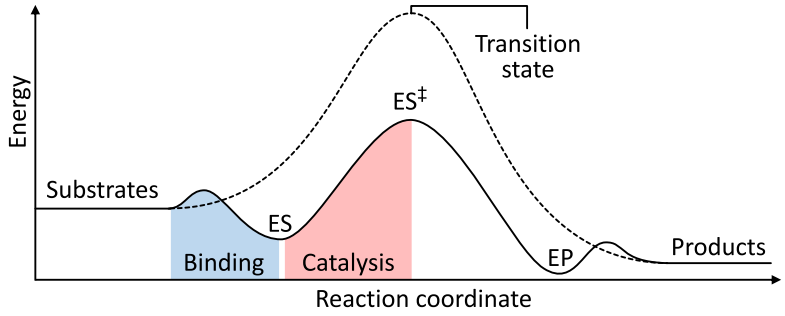


Figure 12.1.4: Simple free energy curve for the conversion for the enzyme-catalyzed conversion of substrate to product. https://commons.wikimedia.org/wiki/F...y_levels_2.svg

Even this diagram is overly simplified since it suggest that the bound substrate in the ES complex is converted to the bound product in one step with no intermediates.

The free energy diagram should include intermediates along the reaction pathway. An example of this is shown in Figure 12.1.5 for the reversible conversion of a 3-carbon sugar, dihydroxyacetone phosphate (DHAP) to another 3-carbon sugar, glyceraldehyde-3-phosphate, in a reaction catalyzed by the enzyme triose phosphate isomerase (we will study enzyme in the next chapter).

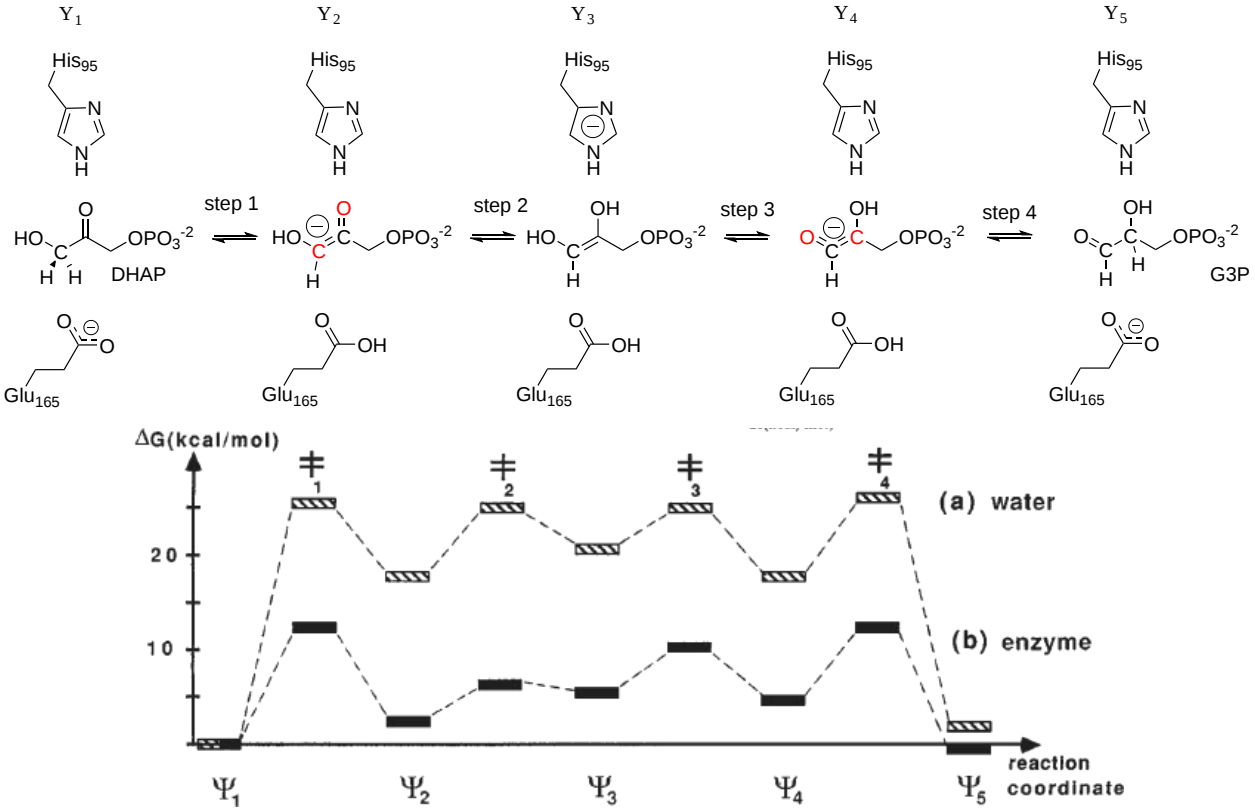


Figure 12.1.5: Complete free energy profile for all the elementary steps of the triosephosphate isomerase catalyzed reaction. Aqvist J, Fothergill M. Computer simulation of the triosephosphate isomerase catalyzed reaction. J Biol Chem. 1996 Apr 26;271(17):10010-6. doi: 10.1074/jbc.271.17.10010. PMID: 8626554. Creative Commons Attribution (CC BY 4.0)

Note that the enzyme lowers the activation energy of each of the steps in the overall reaction. Enzymes can also catalyze reactions by altering the reaction pathway although in this case, all the intermediates in the conversion are the same in both the uncatalyzed and catalyzed pathways.

A comparison of the thermodynamic reactivity of molecules of similar structures can be made by determining the relative stability of the reactants and products from structural considerations. Consider two reactants, R_1 and R_2, which produce products P_1 and P_2, respectively. Any structural features that preferentially stabilize R_2 compared to R_1, or P_2 compared to P_1, but don't stabilize R_2 and P_2 to the same extent, will lead to a greater driving force for $R_2 \rightarrow P_2$ compared to R_1. This is shown graphically in Figure 12.1.6:

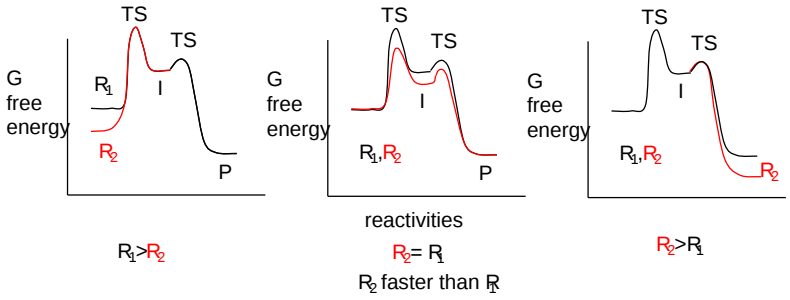


Figure 12.1.6: Free energy reaction digrams for two similar molecules

Mechanisms that lead to the stabilization of a reactant, intermediate, or product include resonance and inductive effects (electron release or withdrawal).

An example of how the comparative acidity of two similar molecules can be determined through a comparison of their structures is shown below for acetic acid and ethanol is shown in Figure 12.1.7.

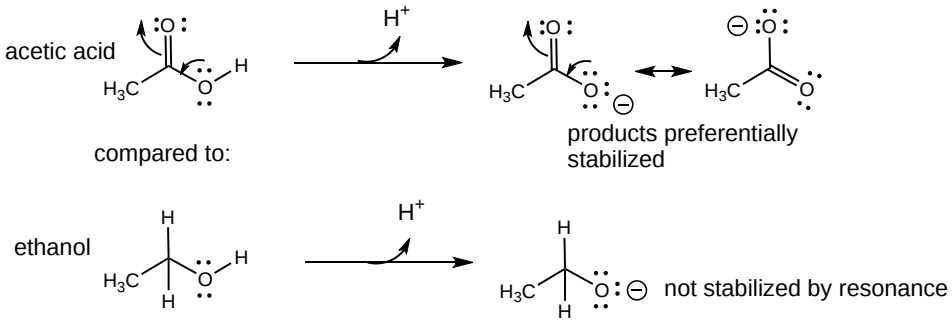


Figure 12.1.7: Comparative acidity of two similar molecules

The stronger acid, acetic acid, has the more stable (and hence less basic) charged product (the conjugate base)

An example of how an intermediate can be stabilized through resonance is shown in Figure 12.1.8 for the keto-enol tautomerization reaction, which is favored in the direction of the keto form, a weaker acid than the enol.

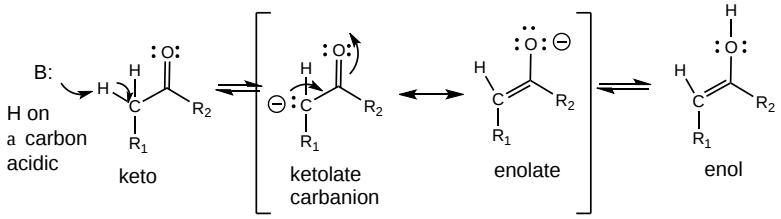


Figure 12.1.8: Comparative stability of keto and enol intermediates

An example of how the inductive effect (electron release and withdrawal) stabilizes/destabilizes carbocations and cations is shown in Figure 12.1.8.

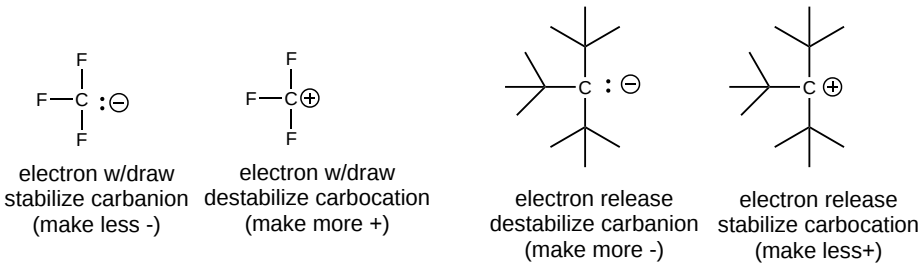


Figure 12.1.8: Factors contributing to stabilization of carbocations and carbanions

Also, remember that electron-withdrawing by the F's in the negatively charged conjugate base of trifluoracetic acid helps explains its lower pKa compared to acetic acid.

Lastly consider how the stabilization of a tertiary carbocation below helps explain the preferential formation of the tertiary alcohol over the secondary alcohol, as shown in Figure 12.1.9.

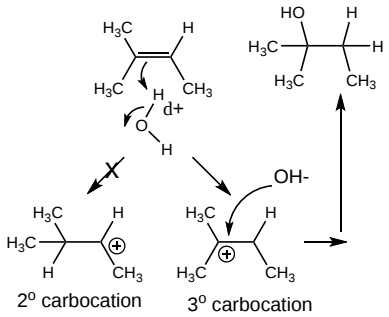


Figure 12.1.9: Preferential formation of tertiary alcohol

12.1.3: OXIDATION OF ORGANIC MOLECULES

Organic molecules are usually oxidized in two-electron steps. Two methods can be used to determine if a C atom in an organic molecule has been oxidized.

- If the number of bonds from C to oxygen increase, or the number of bonds to H decrease, the C is oxidized, More generally, if the number of bonds from C to a more electronegative atom increases, or the number of binds from C to a less electronegative atom decrease, the carbon is oxidized
- A more powerful method involves determining the oxidation number of the carbon atoms in the reactant and product. If the oxidation number becomes more positive, the C is oxidized.

The general rules for determining the oxidation numbers of the atoms in a molecule are:

1. O is generally 2-
2. H is usually 1+
3. in molecules consisting of one type of atom, (like O2) - i.e. a polyatomic element, the atoms have an oxidation number of 0.
4. the sum of the oxidation numbers of the atoms in a molecule equals the net charge on the molecule or ion.

In general, the oxidation number can be calculated as follows:

1. assign all nonbonded electrons of an atom to that atom
2. assign all bonded electrons to the more electronegative atom of the two atoms bonded
3. assign one electron of a bond to each atom if the two atoms are identical.
4. sum up the assigned electrons from 1-3. Subtract this number from the total number of electrons usually present in the outer shell of the atom (the group number). The result is the oxidation number.

An illustration of the sequential two-step oxidation of ethane to acetic acid and assigned oxidation numbers are shown in Figure 12.1.10.

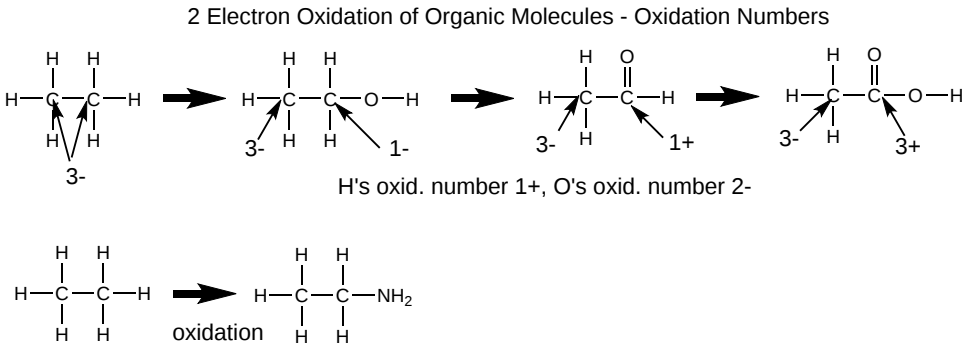


Figure 12.1.10: Change in oxidation number on the stepwise conversion of ethane to acetic acid

Reactions of Carbonyls: Aldehydes and Ketones

When water reacts with an aldehyde in a nucleophilic addition reaction, a 1,1 diol, or a geminal diol results. This reaction can be catalyzed by a base, which acts as the nucleophile (it's a stronger nucleophile than water) and adds to the carbonyl C. OH- is regenerated when the alkoxide produced abstracts a proton from water, regenerating OH-.

When an alcohol adds to an aldehyde or ketone, a hemiacetal or hemiketal, respectively, is formed. In the presence of an acid catalyst, the acid protonates the carbonyl oxygen, making the carbonyl more electrophilic. After the alcohol adds and forms the hemiacetal or hemiketal, the acid can protonate the OH group, leading to its expulsion as water in an acid-catalyzed elimination. The carbocation or resonant-form oxonium ion can react with another ROH to form an acetal or ketal. These steps are summarized in Figure 12.1.11.

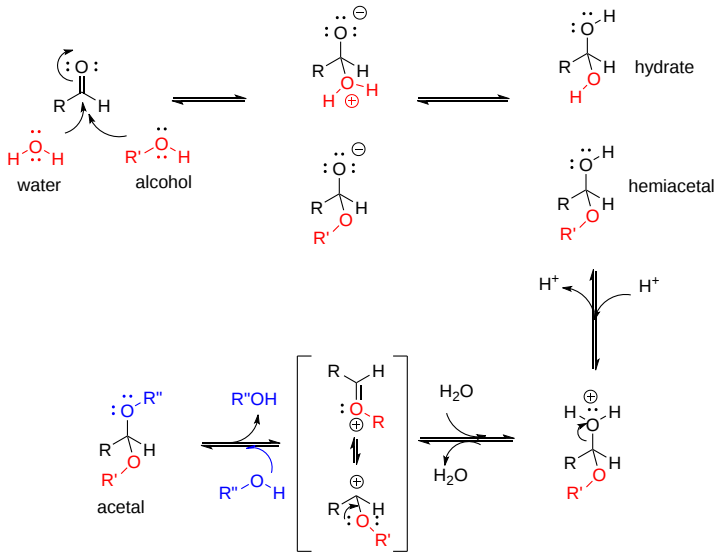


Figure 12.1.11: Nucleophilic addition to an aldehyde

If the nucleophile is an amine, an addition can occur, followed by an elimination to form an imine or Schiff base, as shown in Figure 12.1.12.

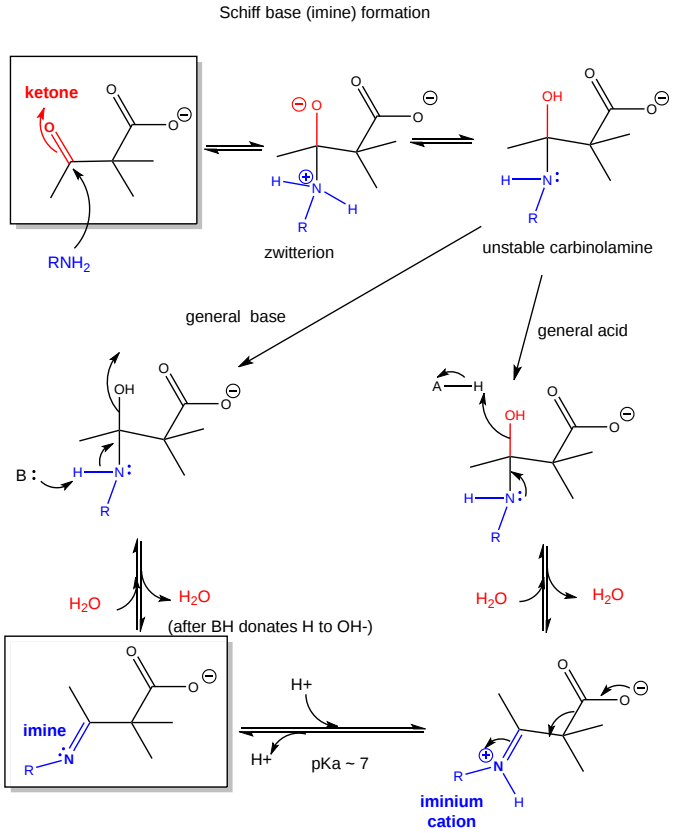


Figure 12.1.12: Schiff base formation

An acetal or ketal are geminal ethers (as water addition to aldehydes or ketones produced geminal diols). As with other ethers, these geminal ethers are stable to base and are hence often used as protecting groups to keep aldehydes and ketones from undesired reactions in

basic solution. Acetal formation is favored by excess anhydrous alcohol in acetic conditions, while acetal breakdown is accelerated by high concentrations of water and the presence of an acid catalyst.

Why are ethers and hence acetals/ketals resistant to bases? They are resistant to nucleophilic attack, such as by base, since the expelled group (alkoxide) is unstable. (Epoxides, in contrast, will react with OH- nucleophiles since the epoxide ring is strained and of high energy.). Ethers can react with acids, however, which protonate the ether O to form an oxonium ion. Nucleophilic attack (such as by Br-) on an adjacent C can occur (S_N2), with electrons flowing to the protonated oxonium ion (a great electron sink) as it departs.

12.1.4: REACTIONS OF CARBOXYLIC ACID DERIVATIVES

Carboxylic acids undergo nucleophilic substitution reactions, assisted by the fact that compared to aldehydes and ketones, they have good leaving groups. With the substitution reaction, the stability of the double bond in the carbonyl is retained. Two things control the reactivity of these derivatives: the stability of the reactants and the stability of the products. The relative reactivity of carboxylic acid derivatives is shown in Figure 12.1.13.

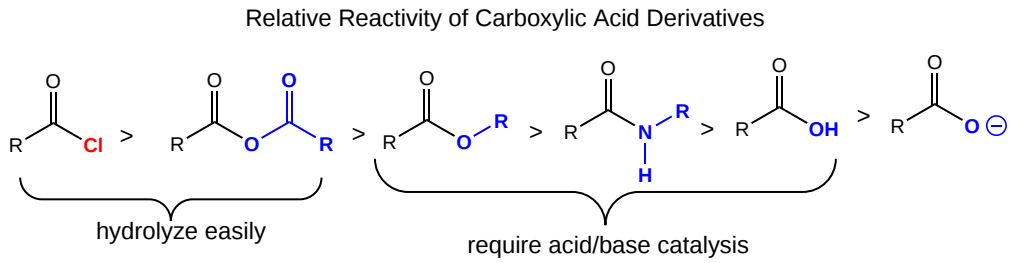


Figure 12.1.13: Relative reactivity of carboxylic acid derivatives

A reactant is less reactive if stabilized by resonance. Hence the relative reactivity is as follows: amide < ester < anhydride < acid chloride

The nonbonded electron pair on N of the amide, a less electronegative atom than O, can delocalize and form a resonant structure with a double bond with the carbonyl C more readily than the O in the ester. An electron pair on the bridging O in the anhydride, since its ability to form a double bond in a resonance structure, is split between the two carbonyls C is less effective in stabilizing either side than in the. (This is called competing resonances.) The reactant less stabilized by resonance is the acid chloride since a nonbonded pair of electrons on the larger chlorine molecule can't delocalize as readily given the C-Cl bond distance.

Notice this order of decreased stability based on resonance stabilization is also the order of increased electrophilicity of the carbonyl C (which is most electrophilic in the absence of electron delocalization from the adjacent N, O, or Cl.

The stability of the products also is important. If the deprotonated leaving group is considered as one of the products (which differentiates the different reactions), then the order of decreased stability of products is Cl^- > $RCOO^-$ > RO^- > RHN^-. This is shown in Figure 12.1.14. (Note: the pKa of ROH = 16 and R_2NH = 40)

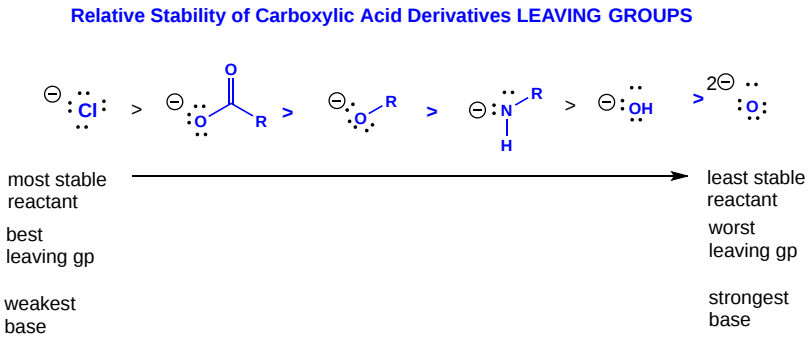


Figure 12.1.14: Relative stability of carboxylic acid derivative leaving groups

What determines the stability of products compared to reactants is the strength of bonds made and broken during the reaction.

In nucleophilic addition to aldehydes and ketones, the strength of the bond to the nucleophile must be greater than the strength of the pi bond broken in the carbonyl. A C-Cl bond strength is 81 kcal/mol (340 kJ/mol) compared to a pi C-O bond strength of 93 kcal/mol (389 kJ/mol). Hence a Cl- is not likely to add to a carbonyl C. Consider the hydration of formaldehyde (carbonyl with 2 H's), acetaldehyde (with 1 H and 1 methyl group, and acetone (with 2 methyl groups). The ΔGo for hydration of these is -19, -1, and +15 kcal/mol (63 kJ/mol), respectively, showing that increased electron release toward the carbonyl C, which makes it less electrophilic and more stable, decreases the reactivity of the carbonyl.

In nucleophilic substitution, the leaving group (anion) must be more stable than the nucleophile.

12.1.5:
KINETICS OF REACTIVITY OF CARBONYLS

The relative kinetic reactivity of various carbonyls toward nucleophiles follows the order of electrophilicity of the C. (i.e the extent of the positive charge on the carbonyl C.) The slow step in a nucleophilic attack is breaking the pi-carbonyl bond. If the reactant is stabilized by resonance in ways that reduce the electrophilicity of the carbonyl C, the reaction is slowed.

Nucleophilicity is a measure of the "affinity" of an atom or ion on an electrophilic C for the nucleophilic lone pair. This is similar to basicity which is a measure of the "affinity" of an atom or ion for a proton. Halides are not good nucleophiles for reactions with acid derivatives since the halide (like Cl-) is a better leaving group than the actual leaving group.

Making C-C Bonds

Metabolism can be divided into catabolic (breaking down) and anabolic (synthetic) reactions. To obtain energy, sugars, and fatty acids are converted to carbon dioxide. Hence C-C bonds must be broken. In contrast, C-C bonds must be synthesized in photosynthesis. In all reactions, electrons from bond broken flow to atoms where bonds will be made. Flow is from a source (a pair of electrons possibly with a negative charge) to a sink (a slightly or fully positive atom). Figure 12.1.15 shows a couple of ways to make a C-C bond using either the reaction of two carbon-centered radicals (free radical mechanisms are uncommon biologically) or a carbocation with a carbanion.

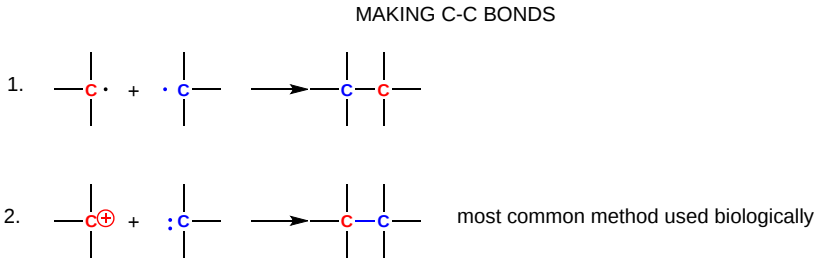


Figure 12.1.15: Making C-C bonds

A carbocation is unstable unless incorporated into a molecule in which it is stable, so instead of using them, the carbonyl C is used as the electrophilic carbon. (Instability of carbocations is reflected in their propensity to rearrangement.) Taking into account the resonance form of the C=O carbonyl bond with a positive on C and a negative on O, the net charge on the carbonyl is about +0.5. A carbanion, often stabilized as an enolate resonant form, is used as the negatively charged carbon. These features are illustrated in Figure 12.1.16.

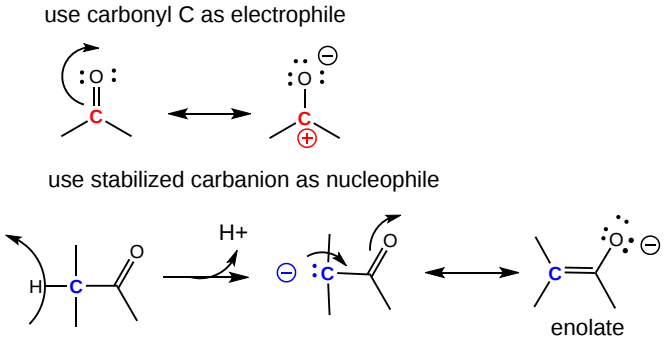


Figure 12.1.16: Carbonyl carbons as electrophiles and carbanions as nucleophiles

One method of making a C-C bond is an aldol condensation, in which a carbanion formed by the deprotonation of a C-H alpha to a carbonyl (which is stabilized by the enolate resonance form) acts as a nucleophile which adds to a carbonyl C in an aldehyde or ketone. The reaction is illustrated in Figure 12.1.17.

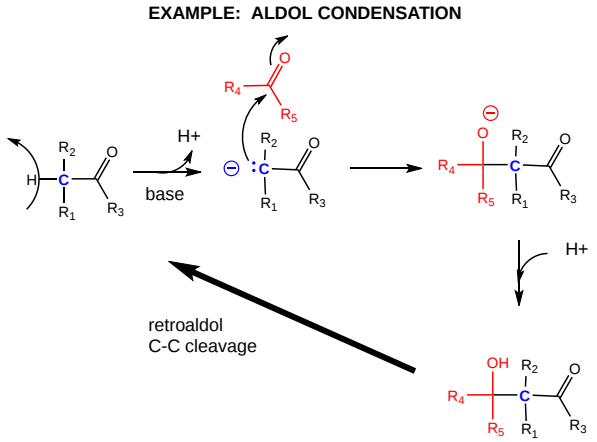


Figure 12.1.17: Aldol condensation

In another C-C bond synthesis reaction, a Claisen Condensation, a carbanion formed by the deprotonation of a C-H alpha to a carbonyl (which is stabilized by the enolate resonance form) acts as a nucleophile that substitutes at a carbonyl C in a carboxylic ester or thioester. This is illustrated in Figure 12.1.18.

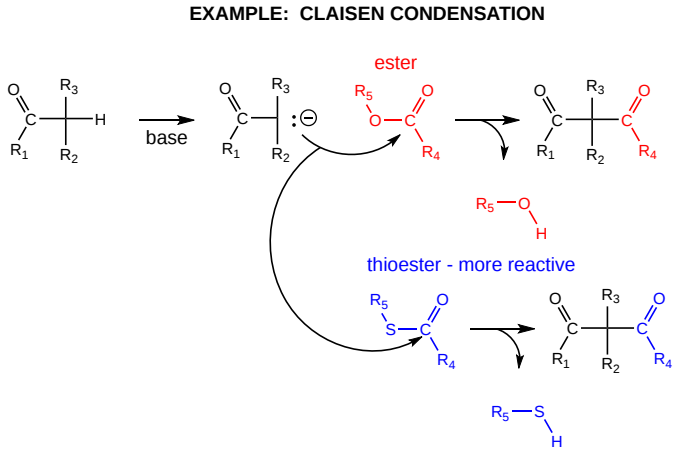


Figure 12.1.18: Claisen condensation

12.1.6: BREAKING C-C BONDS

In addition to a retroaldol condensation, a common method to break a C-C is through a decarboxylation reaction at a beta-keto acid. Notice in Figure 12.1.19 that the analogous reaction at an alpha-keto acid is unlikely since the electrons from the C-C bond that is cleaved have no "sink" to which to flow.

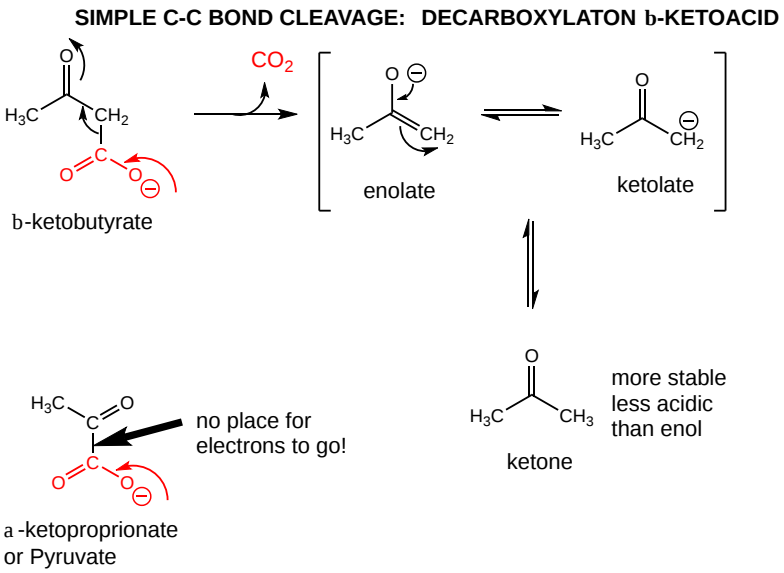


Figure 12.1.19 C-C bond cleavage by decarboxylation of beta-keto acids.

Alpha-keto acids can be decarboxylate using thiamine cofactors as discussed in Chapter 6.

12.2: PHOSPHORYL GROUP TRANSFERS AND ATP

12.2.1: ATP AND PHOSPHORYL TRANSFER REACTIONS

Biological oxidation reactions serve two functions, as described in the previous chapter. Oxidation of organic molecules can produce new molecules with different properties. For example, increases in solubility are observed on the hydroxylation of aromatic substrates by cytochrome P450 (which we will explore in another chapter section). Likewise, amino acids can be oxidized to produce neurotransmitters. Many biological oxidation reactions occur, however, to produce energy to drive thermodynamically unfavored biological processes such as protein and nucleic acid synthesis, or motility. Chemical potential energy is not just released in biological oxidation reactions. Rather, it is transduced into a more useful form of chemical energy in the molecule ATP (adenosine triphosphate). We will discuss the properties that make ATP so useful biologically, and how exergonic biological oxidation reactions are coupled to the synthesis of ATP.

ATP is used universally as a "carrier" of free energy, which effectively means that it has a high free energy compared to its hydrolysis reaction products. Using alternative language, ATP has a high-energy transfer potential. The structure of ATP and a simplified reaction mechanism for the cleavage of the terminal (γ) phosphoanhydride bond by water (hydrolysis) or an alcohol (alcoholysis) is shown in Figure 12.2.1 below.

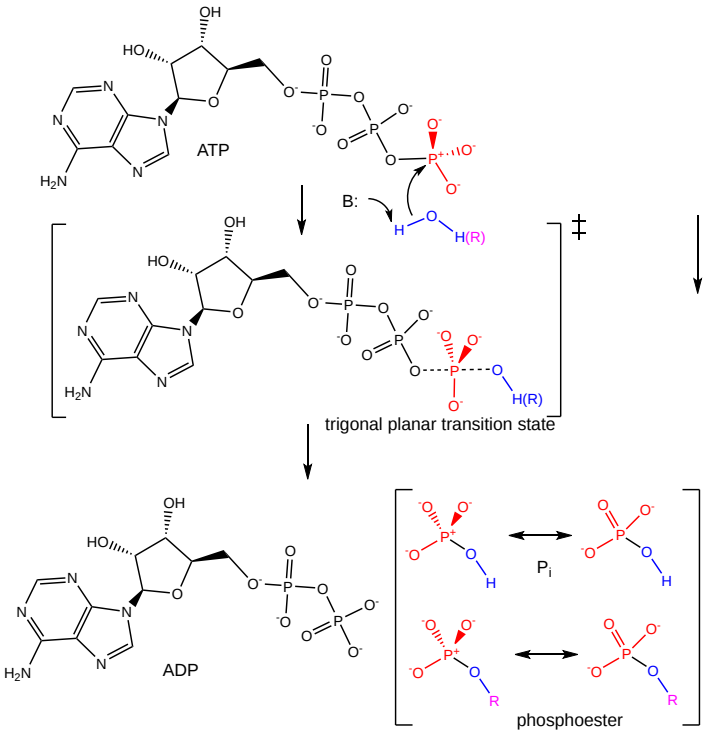


Figure 12.2.1: Mechanism for the hydrolysis and alcoholysis of the terminal phosphoanhydride bond in ATP

The reaction is also a nucleophilic substitution. Changing perspectives, the water or alcohol is phosphorylated by ATP. Hence the reaction can also be called a **phosphoryl transfer**. In Chapter 6.5, we can see that two different Enzyme Commission numbers might apply to reactions involving ATP. These are

- EC2 - transferases: transfer/exchange of group from one molecule to another. More specifically they are in category EC2.7 - transferring phosphorus-containing groups
- EC3- hydrolases: hydrolysis reactions. More specifically they are in category EC3.6 - acting on acid anhydrides

Most of the examples in this book would best be classified as **phosphoryl transfer** reactions.

The terminal phosphate is shown in an alternative resonance form with the P atom sp^3 hybridized with tetrahedral geometry and a +1 formal charge. The attack on the electrophilic P atom by the incoming nucleophile leads to the formation of a trigonal planar transition state with the dashed lines representing bond formation and breaking in an SN2-like reaction. The final products are ADP and either inorganic phosphate (Pi) or a phosphoester.

The hydrolysis reaction can be represented by a chemical equation or by a more typically written "biochemical" equation. Here is the chemical equation as recommended by the IUBMB/IUPAC:

$$\mathrm{ATP^{4-}(aq) + 2\,H_2O(l) \rightleftharpoons ATP^{3-}(aq) + HPO_4^{2-}(aq) + H_3O^+(aq)}$$

This equation is written to have both charge and mass balance. For example, the sum of charges on the left-hand side (-4) is the same as on the right-hand side (-4).

Of course, this is even a simplified equation since the reaction depends on the pH and the presence of divalent cations such as Mg^{2+}. Other species that could be included just for ATP^{4-} include $HATP^{3-}$, H_2ATP^{2-}, $MgHATP^{-}$, and Mg_2ATP, for example. The actual equilibrium constant would depend on the pH, the concentration of Mg^{2+}, and the total ionic strength of the solution. If these were all fixed, the reaction could be written as a simplified biochemical equation as shown below:

$$ATP + H_2O \rightleftharpoons ADP + Pi$$

We will most often use simplified biochemical equations when discussing metabolism.

Just as there are standard state conditions for chemical reactions (1 bar pressure for a gas, 1 M for a solute in solution), there are biochemical standard states for biochemical reactions. They are pressure = 1 bar, pH = 7 (i.e. $H_3O^+=10^{-7}$ M), Mg^{2+} = 1 mM, and ionic strength of either 0 or 0.25 M.

ATP contains two phosphoanhydride bonds (connecting the 3 phosphates) and one phosphoester bond (connecting a phosphate to the ribose ring). The pK_as for the reactions $HATP^{3-} \rightarrow ATP^{4-} + H^+$ and $HADP^{2-} \rightarrow ADP^{3-} + H^+$ are about 7.0, so the overall charges of ATP and ADP at physiological pH are -3.5 and -2.5, respectively. Each of the phosphorous atoms is highly electrophilic and can react with nucleophiles like the OH of water or an alcohol.

As we discussed earlier, anhydrides are thermodynamically more reactive than esters which are more reactive than amides. The large negative ΔG^o (-7.5 kcal/mol, -31kJ/mol) for the hydrolysis of one of the phosphoanhydride bonds can be attributed to relative destabilization of the reactants (ATP and water) and relative stabilization of the products (ADP = Pi). Specifically

- The reactants can not be stabilized to the same extent as products by resonance due to competing resonance of the bridging anhydride O's.
- The charge density on the reactants is greater than that of the products
- Theoretical studies show that the products are more hydrated than the reactants.

The ΔG^o for hydrolysis of ATP is dependent on the divalent ion concentration and pH, which affect the stabilization and the magnitude of the charge states of the reactants and products.

Carboxylic acid anhydrides are even more unstable to hydrolysis than ATP (-20 kcal/mol, -84 kJ/mol), followed by mixed anhydrides (-12 kcal/mol, -50 kJ/mol), and phosphoric acid anhydrides (-7.5 kcal/mol. -31 kK/mol). The terminal anhydride bond is often called a "high energy bond". **This is absolutely wrong and has created deep misconceptions about molecules like anhydrides**. What is true is that the anhydride reactants are high energy but *only compared to the energy of their cleavage products*, such that the reaction proceeds with a large negative ΔG^o.

There is no such thing as a high-energy bond. All covalent bonds lower the energy of a system of two separated atoms. Figure 12.2.2 shows molecules that are high energy compared to their hydrolysis products. Think of a graph of free energy. Carboxylic acid anhydrides and water (the reactants) have high energy compared to their collective products, two acetic acids, for example.

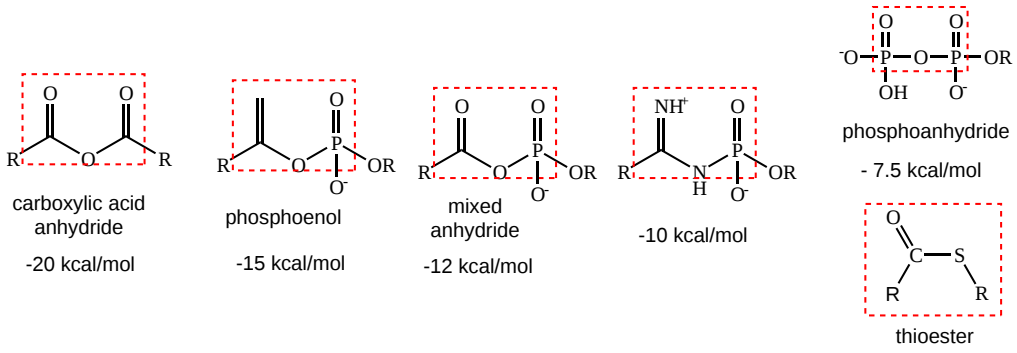


Figure 12.2.2: Molecules that are high energy *compared to their hydrolysis products*.

Some older books state that the terminal anhydride bond is "high energy". There is no such thing as a "high energy" bond. When a bond forms between two molecules, the energy is lowered. Energy input is required to break any bond.

Each of the molecules above except the thioester has a similar motif outlined with the red dotted rectangle. The thioester also is considered high energy compared to its hydrolysis product since the reactant is effectively destabilized compared to a carboxylic acid ester. This arises because the sulfur atom is larger in the thioester than the oxygen atom in the carboxylic acid ester. Hence the bond length of C-S (1.82 Å)is larger than for C-O (1.43 Å), so the C-S bond is weaker. In addition, the lone pairs on the S, which is more polarizable than O, are less likely to be shared with the C as part of the resonance stabilization of the ester. Both effects raise the energy of the thioester compared to the

carboxylic acid ester. The hydrolysis products of both esters are of similar energy. Hence the ΔG^o for the hydrolysis of the thioester is more negative and about the same for the hydrolysis of ATP.

How can ATP be used to drive thermodynamically unfavored reactions? First consider how the hydrolysis of a carboxylic acid anhydride, which has a ΔG^o = -12.5 kcal/mol (-52 kJ/mol) can drive the synthesis of a carboxylic acid amide, with a ΔG^o = + 2-3 kcal/mol (+ 4-12 kJ/mol). The reaction is:

anhydride + amine --> amide + carboxylic acid

This can be broken into two reactions, the hydrolysis of the anhydride, and the synthesis of the amide as shown in Figure 12.2.3.

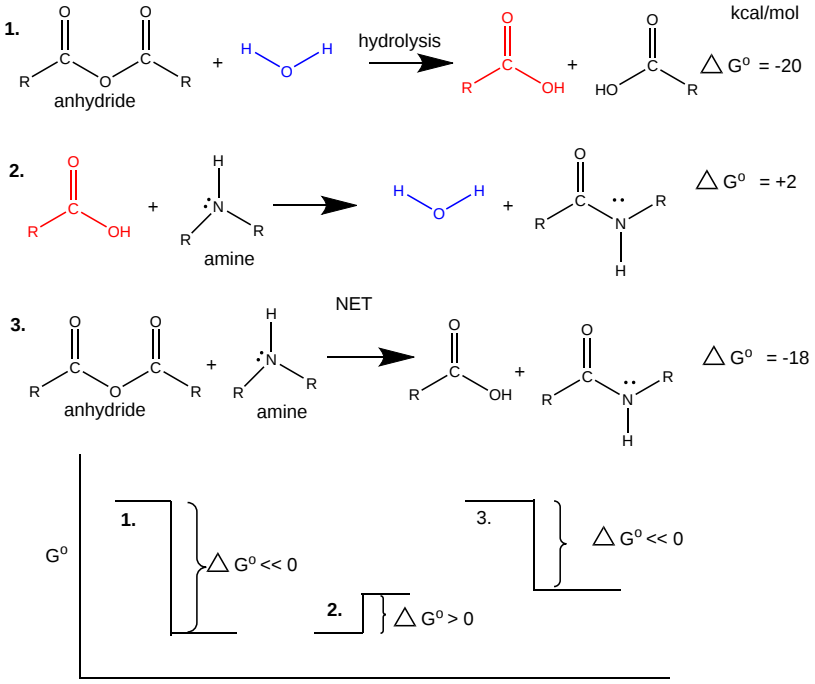


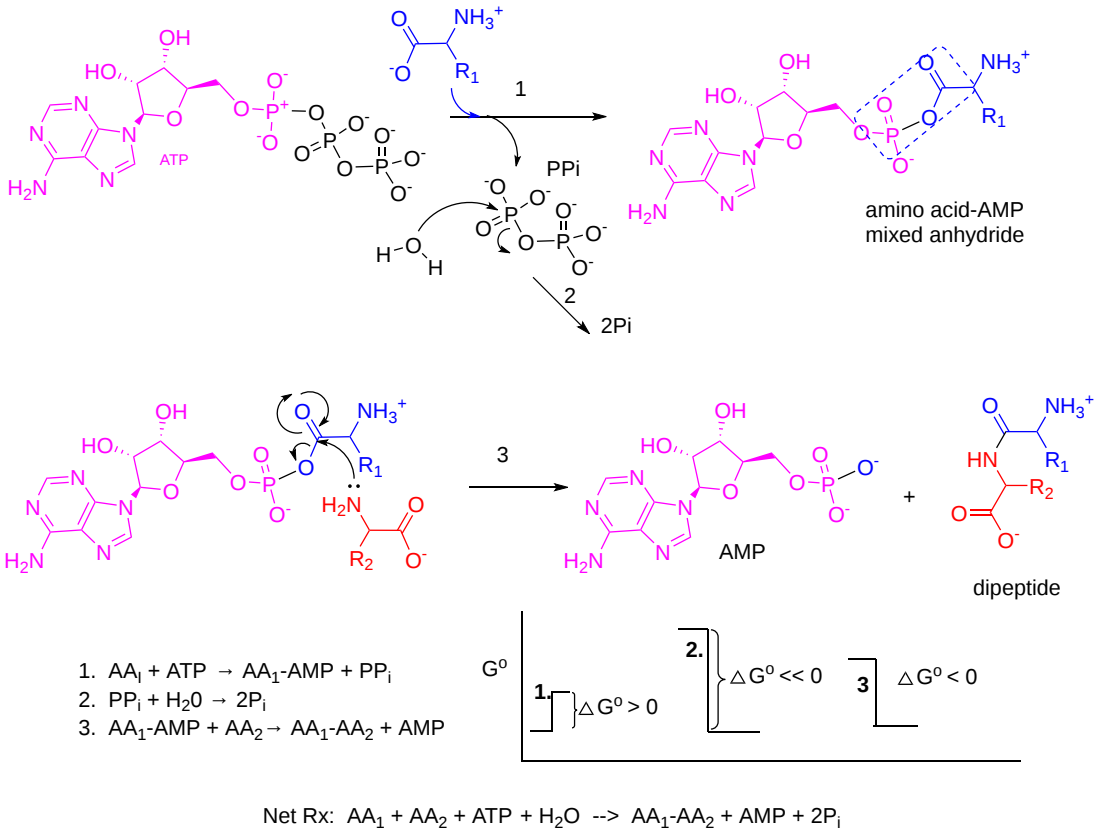
Figure 12.2.3: Individual and net reactions for the conversion for the formation of an amide from an anhydride and an amine

Now consider the reaction of glucose + Pi to form glucose-6-P. In this reaction, a phosphoester is formed, so the reaction would proceed with a positive ΔG^o = 3.3. Now if ATP was used to transfer the terminal (gamma) phosphate to glucose to form Glc-6-P, the reaction proceeds with a ΔG^o = -4.2 kcal/mol (-17.6 kJ/mol). This can be calculated since ΔG and ΔG^o are state functions and path independent. Adding the reactions and the ΔG^os give:

- glucose + Pi → glucose-6-P, ΔG^0 = 3.3
- ATP + H_2O → ADP + Pi , ΔG^0 = -7.5
- NET: glucose + ATP -→ Glucose-6-P + ADP, ΔG^o = -4.2

In most biological reactions using ATP, the terminal phosphate of ATP is transferred to a substrate using an enzyme called a kinase. Hence, hexokinase transfers the gamma phosphate from ATP to a hexose sugar. Protein kinase is an enzyme that transfers the gamma phosphate to a protein substrate.

ATP is also used to drive peptide bond (amide) synthesis during protein synthesis. From an energetic point of view, anhydride cleavage can provide the energy for amide bond formation. Peptide bond synthesis in cells is accompanied by cleavage of both phosphoanhydride bonds in ATP in a complicated set of reactions that are catalyzed by ribosomes in the cells. (This topic is considered in depth in Unit III). Figure 12.2.4 is a grossly simplified mechanism of how peptide bond formation can be coupled to ATP cleavage.



Net Rx: $AA_1 + AA_2 + ATP + H_2O$ --> AA_1-$AA_2 + AMP + 2P_i$

Figure 12.2.4: Individual and net reactions for the formation of a dipeptide from separate amino acids coupled with the cleavage of ATP. f

In the first reaction, the carboxylic acid end of amino acid 1 is activated to form a mixed carboxylic ester. The leaving group, Pi, is hydrolyzed in reaction 2 to help drive the reaction. An amide bond is formed in reaction 3, with the expulsion of an excellent leaving group, AMP. Phosphorylation reactions using ATP are nucleophilic substitution reactions that proceed through a pentavalent transition state. These reactions are also called phosphoryl transfer reactions.

One last note. ATP exists in cells as just one member of a pool of adenine nucleotides which consists of not only ATP but also ADP and AMP (along with Pi). These constituents are readily interconvertible. We break down an amount of ATP each day equal to our body weight. Likewise, we make about the same amount from the turnover products. When energy is needed, carbohydrates and lipids are oxidized and ATP is produced, which can then be immediately used for motility, biosynthesis, etc. It is very important to realize that although ATP is converted to ADP in a thermodynamically spontaneous process, the process is kinetically slow without an enzyme. Hence ATP is stable in solution. However, its biological half-life is not long since it is used very quickly as described above. This recapitulates a theme we have seen before. Many reactions (like oxidation with dioxygen, denaturation of proteins in nonpolar solvent, and now ATP hydrolysis) are thermodynamically favored but kinetically slow. This kinetic slowness is a necessary but of course insufficient condition, for life.

12.2.2: INTRODUCTION TO ACTIVE TRANSPORT

We have previously discussed how chemical potential energy in the form of reduced organic molecules can be transduced into the chemical potential energy of ATP. This ATP can be used to drive reductive biosynthesis and movement (from individual cells to whole organisms). ATP has two other significant uses in the cell.

Active Transport: Molecules must often move across membranes against a concentration gradient - from low to high chemical potential - in a process characterized by a positive ΔG. As protons could be "pumped" across the inner mitochondrial membrane against a concentration gradient, powered by the ΔG associated with electron transport (passing electrons from NADH to dioxygen), other species can cross membranes against a concentration gradient - a process called active transport - if coupled to ATP hydrolysis or the collapse of another gradient. Remember that active transport is different from facilitated diffusion we studied earlier, which occurs down a concentration gradient across the membrane. Many such species must be transported into the cell or intracellular organelles against a concentration gradient as illustrated in Figure 12.2.5:

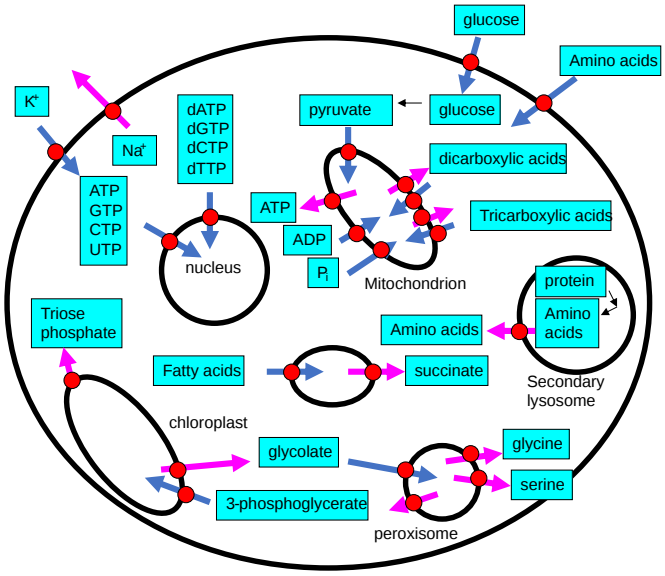


Figure 12.2.5: Examples of active transport reactions (source unfortunately lost)

Signal Transduction: All cells must know how to respond to their environment. They must be able to divide, grow, secrete, synthesize, degrade, differentiate, cease growth, and even die when the appropriate signal is given. This signal invariably is a molecule that binds to a receptor, typically on the cell surface. (Exceptions include light transduction in retinal cells when the signal is a photon and lipophilic hormones which pass through the membrane.) Binding is followed by shape changes in transmembrane protein receptors which effectively transmits the signal into the cytoplasm. We will discuss two main types of signal transduction pathways:

- nerve conduction, in which a presynaptic neuron releases a neurotransmitter causing a postsynaptic neuron to "fire";
- signaling at the cell surface which leads to the activation of kinases within the cytoplasm;

We will discuss signal transduction in the final two chapters.

For active transport to occur, a membrane receptor is required which recognizes the ligand to be transported. Of major interest to us, however, is the energy source used to drive transport against a concentration gradient. The biological world has adapted to use almost any source of energy available.

Energy released by oxidation: We have already encountered the active transport of protons driven by oxidative processes. In electron transport in respiring mitochondria, NADH is oxidized as it passes electrons to a series of mobile electron carriers (ubiquinone, cytochrome C, and eventually dioxygen) using Complex 1, 3, and 4 in the inner membrane of the mitochondria. Somehow the energy lost in this thermodynamically favored process was coupled to conformational changes in the complex which caused protons to be ejected from the matrix into the inner membrane space. One can imagine a series of conformation-sensitive pKa changes in various side chains in the complexes which lead in concert to the vectorial discharge of protons.

ATP hydrolysis: One would expect that this ubiquitous carrier of free energy would be used to drive active transport. This is one of the predominant roles of ATP in the biological world. 70% of all ATP turnover in the brain is used for the creation and maintenance of a Na and K ion gradient across nerve cell membranes using the membrane protein Na+/K+ ATPase.

Light: Photosynthetic bacteria have a membrane protein called bacteriorhodopsin which contains retinal, a conjugated polyene derived from beta-carotene. It is analogous to the visual pigment protein rhodopsin in retinal cells. Absorption of light by the retinal induces a conformation change in the retinal and protein, which leads to the vectorial discharge of protons ;

Collapse of an ion gradient: The favorable collapse of an ion gradient can be used to drive the transport of a different species against a concentration gradient. We have already observed that the collapse of a proton gradient across the inner mitochondria membrane (through FoF1ATP synthase) can drive the thermodynamically unfavored synthesis of ATP. The collapse of a proton gradient provides a proton-motive force that can drive the active transport of sugars. Likewise, a sodium-motive force can drive the active transport of metal ions. Since the energy to make the initial ion gradients usually comes from ATP hydrolysis, ATP indirectly powers the transport of the other species against a gradient.

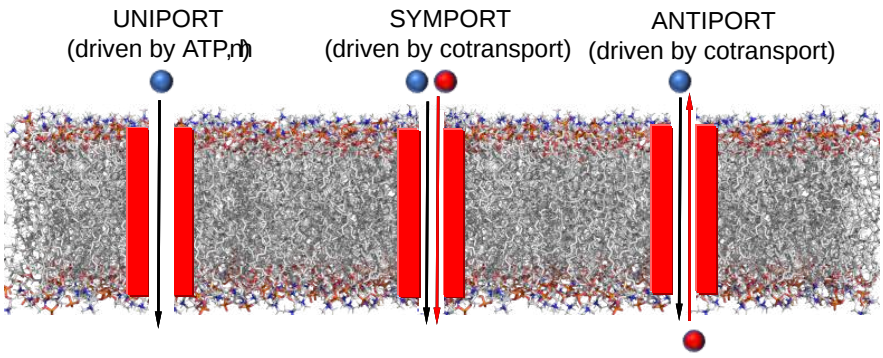


Figure 12.2.6: Types of coupled active transport

Often, the transport of one species is coupled to transport of another. If the species are charged, a net change in charge across the membrane may occur. Several terms are used to describe various types of transport, as we saw previously in Chapter 12, and which are illustrated in Figure 12.2.6.

- symport - two species are cotransported in the same direction by the same transport protein
- antiport - two species are cotransported in opposite directions by the same transport protein
- electrogenic - a net electrical imbalance is generated across the membrane by symport or antiport of charged species
- electroneutral - no net electrical imbalance is generated across the membrane by symport or antiport of charged species

---

12.3: THE CHEMISTRY AND BIOCHEMISTRY OF DIOXYGEN

12.3.1: THE HISTORY OF OXYGEN

Oxygen may be considered one of the most important elements in chemistry. Not counting hydrocarbons, there is a greater diversity of molecules with oxygen than with carbon. Given its role in the molecular world, very little time is spent on the chemistry of oxygen in undergraduate chemistry classes. Why is oxygen so special?

Oxygen reacts with atoms of all elements except the Noble gases to form molecules. One of the most important molecules of course, from a biological sense, is water. It :

- provides a perfect solvent for biomolecules
- moderates the earth's climate
- is the source of almost all the dioxygen in the air

From a chemical point of view, water is a(n):

- nucleophile and electrophile
- acid and base
- oxidizing agent and reducing agent
- a protic solvent that can form H-bonds

The formation of the earth and the development of life:

The gaseous and dusty environment from which the earth was formed contained metals and water, which as you remember from introductory chemistry, can react to form hydrogen gas. H_2 reacts with nonmetals (under various conditions of temperature and pressure) to form H_2S, HCl, CH_4, and NH_3 which contributed to the reducing nature of the early atmosphere. This kept the transition metals in their lowest oxidation states. Many metals, including the coinage metals (Cu, Ag, and Au) and the platinum group (Ru, Rh, Pd, Pt) were stable in elemental form.

Then, around 2.7-2.8 billion years ago, photosynthetic organisms (blue/green algae- also called cyanobacteria) developed which could oxidize water to form dioxygen. Oxygen was generally unavailable for redox chemistry before then as photosynthesis, the process that would evolve to oxidize water to produce dioxygen, was unavailable. Remember that to oxidize water to dioxygen, itself a strong oxidizing agent, requires a stronger oxidizing agent than dioxygen and lots of energy. Fossilized remains of cyanobacteria are found in stromatolites. Using knowledge of how atmospheric oxygen can alter the chemistry of different sulfur isotopes of SO_2, it has been shown that O_2 did not exist in the atmosphere as a whole above 1 ppm earlier than 2.4 billion years ago, although there might have been isolated pockets with higher concentrations. After that, it rose, presumably as a result of cyanobacteria. Before this time, bacteria oxidized a similar molecule, H_2S to form elemental sulfur. It could do this through the photosynthetic reduction of CO_2 by H_2S. Volcanic gases like H_2 might have kept oxygen levels from rising between 2.7 billion years ago and 2.4 billion years ago when its build-up started. Hydrogen in the form of H_2 and methane probably decreased around 2.4 billion years ago as methane with its hydrogen atoms escaped to the upper atmosphere and space. Methane levels would also be decreased by its easy reaction with dioxygen in the presence of UV light to form CO2. This would paradoxically lead to a cooling of the earth and pronounced glaciation as a more potent greenhouse gas, methane, was replaced with a less potent one, carbon dioxide.

Over the next billion years, dioxygen rose to perhaps 0.2 - 2% (compared to the present levels of 20%) Why? Because the early atmosphere was reducing, the added oxygen combined with a large "sink" of reduced metals (like elemental Cu and Fe) or nonmetals (like C and ammonia), preventing a large buildup. Only after these reduced substances were "titrated" did dioxygen build up to present levels. In addition, the oxygen might have increased weathering (by oxidation) of sulfur deposits which can lead to sulfides entering the ocean, where they could precipitate ocean iron ions that are necessary for cyanobacterial chemistry. This would place constraints on cyanobacterial growth until dioxygen levels in the atmosphere increased enough so sulfides were converted to sulfates. This first increase in atmospheric oxygen is often called the Great Oxidation Event as it correlated and presumably caused one of the greatest mass extinctions (of anaerobic organisms) of all time.

Around 2.3 billion years ago, as trace dioxygen had accumulated in the atmosphere, redox chemistry changed, although isotope evidence suggests that little dioxygen was found in water. Around 1.8 - 1.5 billion years ago, the earth's atmosphere became somewhat oxygenated, which was also coincident with the development of eukaryotic organisms. Until then, life was restricted to the oceans since there was no ozone to absorb dangerous UV radiation. The buildup of dioxygen in the air must have led to another extinction of anaerobic organisms since as we shall see, products of oxygen metabolism are very toxic. Some evolved to use dioxygen. Ozone developed, and life could then migrate from the sea to the land. It wasn't until around 600 million years ago that animals arose, however. Was this event associated with the development of a fully oxygenated (20%) atmosphere? Recent evidence, which shows that substantial oxygen wasn't available in the deep sea until about 600 million years, seems to suggest that. Based on an analysis of iron compounds in waters in Newfoundland, it appears that oxygen was very low in the sea 580 million years ago, during the Gaskier's glaciation period. Immediately after that it rose to levels

consistent with atmospheric dioxygen levels of 15%, levels necessary for large animals. Similar trends in carbon and sulfur isotopes in marine rocks in Oman also suggest large increases in oxygen at the end of the Gaskiers glaciation period. What caused this second great oxygenation event? One possibility is that organic matter was sequestered from a reaction with atmospheric dioxygen, as clays bound organic molecules in the ocean and lichens and zooplankton facilitated weather and production of insoluble organic material in the oceans.

Dioxygen is critically important for higher organisms, so an understanding of its chemistry becomes important. This chapter will show that dioxygen is a ground state diradical that has low solubility in an aqueous solution, reacts in a kinetically sluggish fashion in the oxidation reaction, and forms toxic byproducts as it gets reduced. Life forms hence evolved ways to deal with these problems, including ways to increase its solubility (with dioxygen binding and transport proteins), and enzymes (that could activate it kinetically and also detoxify oxygen by-products). Dioxygen is toxic to many cells. Obligate aerobes die in an oxygen environment as many of their cellular components get oxidized by this excellent oxidizing agent. Several strains of bacteria swim away from high levels of dioxygen. A graph showing the log of survival vs log pO2 is linear with a negative slope for a variety of organisms, including mice, fish, rats, rabbits, and insects. Pure oxygen can induce chest soreness, coughs, and sore throats in people. Premature infants put in pure dioxygen environments often developed blindness due to retrolental fibroplasia (a build-up of fibrous tissue behind the lens). The trade-off for this toxicity is clear. Energy is derived from organic molecules through oxidation. Before dioxygen became available to power aerobic catabolism of reduced molecules like fatty acids and less reduced sugars, such molecules were only partially oxidized. The glycolytic pathway, found in most organisms, oxidizes glucose (6 Cs) to two molecules of pyruvate (3 Cs). It was only with the availability of dioxygen did pathways evolve (Kreb Cycle, mitochondrial electron transport/oxidative phosphorylation) that allowed pyruvate to be fully oxidized to carbon dioxide, with the release of much more energy.

12.3.2: THE PROPERTIES OF DIOXYGEN

It is important to understand the properties of dioxygen since oxidation reactions using its power not only our bodies but our entire civilization. We will concentrate on biological reactions, but even these show the same characteristics as non-biological ones.

- oxidation of organic molecules by oxygen is thermodynamically favored but kinetically slow.
- pure oxygen environments are toxic to cells and organisms.

First, we will try to understand these properties of oxygen, and then we will see how organisms overcome these problems to use dioxygen.

We can understand both of these properties by looking at the molecular orbitals of oxygen and its reduction products as shown in the diagrams below. Ground state oxygen is a diradical, which explains the paramagnetic behavior of oxygen. The two unpaired oxygen each have a spin state of 1/2 for a total resultant spin S of 1, making ground state oxygen a triplet (2S+1) = 3. Organic molecules typically undergo 2 electron oxidation steps. Consider the stepwise oxidation of methane below. The oxidation number of C in methane is -4, -2 in methanol, 0 in formaldehyde, +2 in formic acid, and finally +4 in carbon dioxide, indicating two electron losses in each step. These states are shown in Figure 12.3.1.

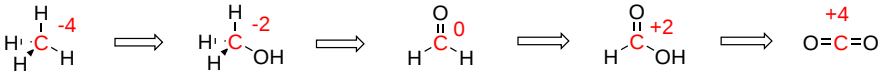


Figure 12.3.1:

The two electrons lost by the organic substrate are added to oxygen, but since the two lost electrons are spin paired, a spin flip must occur to allow the electrons to enter the unfilled oxygen orbitals. Alternatively, energy can be put into ground state dioxygen to produce excited state singlet oxygen (S=0, 2S+1 = 1). The source of the large activation energy required (about 25 kcal or 105 kJ/mol) to flip the electron spin accounts for the kinetic sluggishness of reactions of dioxygen with organic reactants.

A traditional Lewis structure for ground state dioxygen can not be easily written since the electrons are added in pairs, and dioxygen is a diradical. There are 6 electrons in the sigma molecular orbitals from second shell electrons (two each in σ2s, σ2s*, and σ2p,) and 6 electrons in the pi molecular orbitals from second shell electrons (two each in two different π2p orbitals, and one electron each in two different π2p*), so the net number of electrons in bonding orbitals is 4, giving a bond order (or number of 2). In contrast, it is easy to write the Lewis structure of singlet, excited state oxygen, since all electrons can be viewed as paired, with two net bonds (1 sigma, 1 pi) connecting the atoms of oxygen. Figure 12.3.2 shows the molecular orbital diagram for ground and excited state dioxygen.

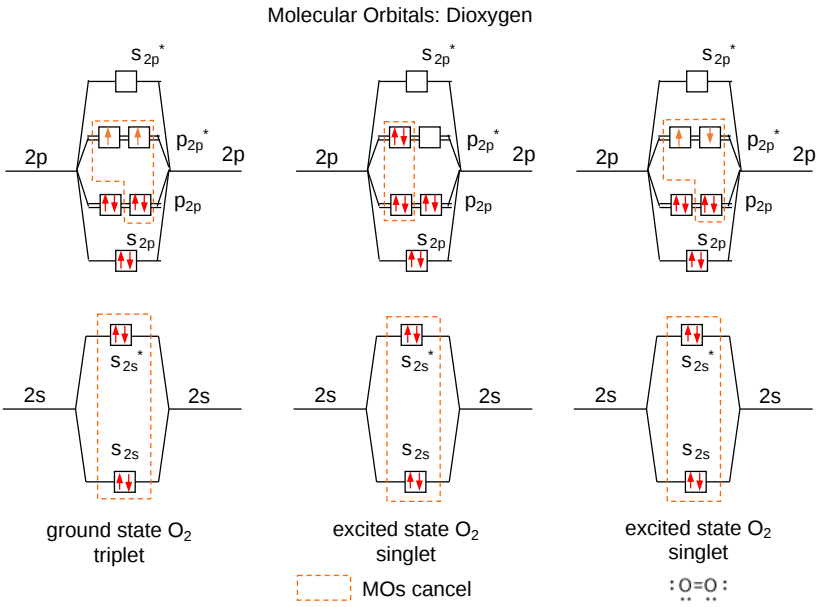


Figure 12.3.2: Molecular orbital diagram for ground and excited state dioxygen

This Lewis structure will be used to represent singlet, excited oxygen, which should react more quickly with organic molecules. The excited state singlet on the right is unstable and decays to the middle singlet state. The middle state is approximately 94.3 kJ/mol higher in energy than the ground state triplet (on the left). In quantum mechanical parlance, the transition from the ground state triplet to the singlet state is forbidden for several reasons, making it unlikely that absorption of a photon will induce the transition.

12.3.3: THE REDUCTION OF DIOXYGEN

When oxygen oxidizes organic molecules, it is reduced. By adding electrons one at a time to the molecular orbitals of ground-state dioxygen we produce the step-wise reduction products of oxygen. On the addition of one electron, superoxide is formed. A second electron produces peroxide. Two more produce 2 separated oxides since no bonds connect the atoms (the number of electrons in antibonding and bonding orbitals is identical). Each of these species can react with protons to produce species such as HO_2, H_2O_2 (hydrogen peroxide), and H_2O. It is the first two reactive reduction products of dioxygen that make it potentially toxic. Figure 12.3.3 shows the MO diagrams for the reduction products of dioxygen.

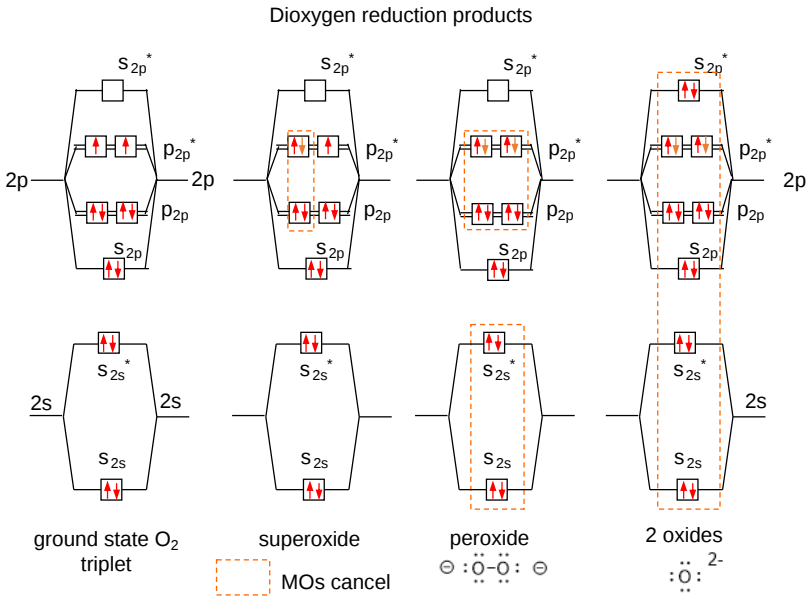


Figure 12.3.3: MO diagrams for the reduction products of dioxygen

How are the potential problems in oxygen chemistry dealt with biologically?

Kinetic sluggishness: Enzymes that utilize dioxygen must activate it in some way, which decreases the activation energy. Enzymes that use dioxygen typically are metalloenzymes, and often heme-containing proteins. Since metals such as Fe^{2+} and Cu^{2+} are themselves free radicals (i.e. they have unpaired electrons), they react readily with ground-state oxygen which itself is a radical. The molecular orbitals of the metal and oxygen combine to produce new orbitals which for oxygen are more singlet-like. Likewise, dioxygen reacts more readily with organic molecules which can form reasonably stable free radicals, such as flavin adenine dinucleotide (FAD), as we shall see later.

Dioxygen toxicity: Since toxicity arises from the reduction products of oxygen, enzymes that use oxygen have evolved to bind oxygen and its reduction products tightly (through metal-oxygen bonds) so they are not released into the cells where they can cause damage. In addition, enzymes that detoxify free dioxygen reduction products are widely found in nature. For example:

- superoxide dismutase catalyzes the dismutation (self-redox) of 2 superoxides into dioxygen and hydrogen peroxide;
- catalase converts hydrogen peroxide into water and oxygen;
- peroxidase catalyzes the reaction of hydrogen peroxide with an alcohol to form water and an aldehyde
- peroxiredoxins react with peroxides and thioredoxin (a small electron donor) to form water and oxidized thioredoxin.

Finally, free radical scavengers such as vitamins A, C, E, and selenium can react with reactive free radicals to produce more stable free radical derivatives of the vitamins and Se. More on this later.

12.3.4: THE REACTIONS OF DIOXYGEN AND ITS REDUCTION PRODUCTS

12.3.4.1: TRIPLET O_2 - GROUND STATE

Here are some reactions for the ground state (triplet O_2):

a. Metals ions - Metal ions are radicals themselves, so can easily react with dioxygen (think about rust). Here is one example

$$Fe^{2+} + O_2 \rightleftharpoons [Fe^{2+} - O_2 \rightleftharpoons Fe^{3+} - O_2^{-\cdot}] \rightleftharpoons Fe^{3+} + \underbrace{O_2^{-\cdot}}_{\text{superoxide}}$$

b. Autoxidation of organic molecules to produce peroxides - These are multistep reactions that have initiation, propagation, and termination steps.

$$RH \rightarrow R^{\cdot} \text{ (Initiation)}$$
$$R^{\cdot} + O_2 \rightarrow ROO^{\cdot} \text{ (Propagation)}$$
$$ROO^{\cdot} + RH \rightarrow R^{\cdot} + ROOH \text{ (Propagation)}$$
$$R^{\cdot} + R^{\cdot} \rightarrow R\text{-}R \text{ (Termination)}$$
$$ROO^{\cdot} + ROO^{\cdot} \rightarrow ROOR + O_2 \text{ (Termination)}$$
$$ROO^{\cdot} + R^{\cdot} \rightarrow ROOR \text{ (Termination}$$

Figure 12.3.4 summarizes the reactions of triplet ground state dioxygen.

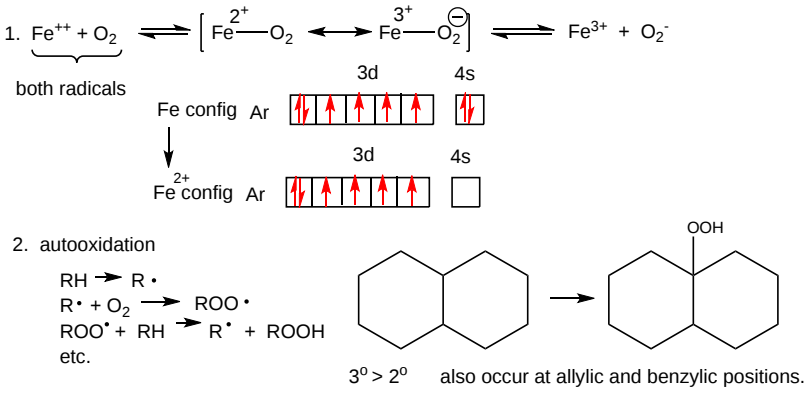


Figure 12.3.4: Reactions of triplet ground state dioxygen

The initiation step above occurs mostly at C atoms which can produce the most stable free radicals (allylic, benzylic position, and $3^o > 2^o >> 1^0$ carbons).

12.3.4.2: SINGLE O_2 - EXCITED STATE

Figure 12.3.5 shows some reactions for singlet dioxygen in which dioxygen is shown as with a double bond and two lone pairs on each oxygen.

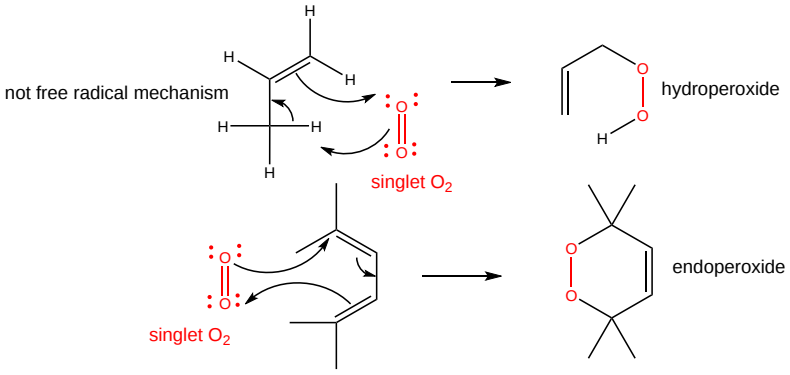


Figure 12.3.5: Some reactions for singlet dioxygen

Alkenes react with oxygen to form hydroperoxides, potentially through an epoxide intermediate. Dienes reacts with oxygen in a Diels-Alder pericyclic reaction to form endoperoxides. A molecular orbital perspective (that you may remember from chemistry classes) on this cycloaddition reaction is shown in Figure 12.3.6.

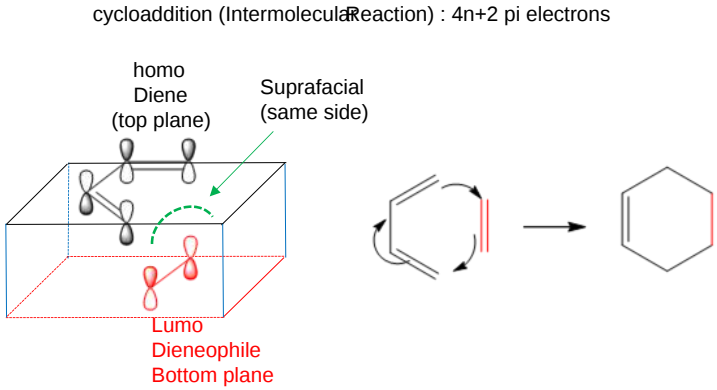


Figure 12.3.6: Pericyclic cycloaddition reactions using molecular orbitals

Singlet oxygen can be made from triplet oxygen by photoexcitation. Alternatively, it can be made from triplet oxygen through collision with an excited molecule which relaxes to the ground state after a radiationless transfer of energy to triplet oxygen to form reactive singlet oxygen. This later process accounts for the photobleaching of colored clothes when the conjugated dye molecules absorb UV and Vis light and relax to the ground state by transferring energy to triplet oxygen to form singlet oxygen. That can more readily react with the conjugated double bonds in the dye. These processes are summarized in Figure 12.3.7.

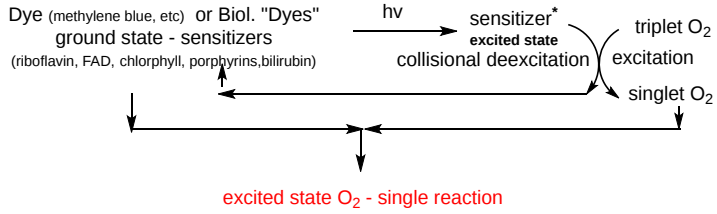


Figure 12.3.7: Photosensitization and the reaction of dyes with singlet state dioxygen

12.3.4.3: SUPEROXIDE

Common reactions of superoxide are shown below.

a. **Dismutation:** This reaction involves a specified reactant undergoing an oxidation reaction, followed by another molecule of the same reactant undergoing a reduction.

$$O_2^{-\cdot} + O_2^{-\cdot} + 2\,H^+ \rightleftharpoons H_2O_2 + O_2 \qquad \text{(slow)}$$

$$HO_2^{\cdot} + O_2^{-\cdot} + H^+ \rightleftharpoons H_2O_2 + O_2 \qquad \text{(fast)}$$

b. Acid/Base:

$$HO_2^{\cdot} \rightleftharpoons O_2^{-\cdot} + H^+ \qquad (pKa = 4.8)$$

c. With metal ions:

$$\underbrace{Fe^{3+}}_{\text{as in heme}} + O_2^{-\cdot} \rightarrow O_2 + Fe^{2+}$$

The enzyme superoxide dismutase catalyzes the dismutation reaction. The common eukaryotic cytosolic form contains Cu^{2+} and Zn^{2+} ions, which are coordinated by histidine side chains. The reaction proceeds in two steps or half-reactions. The first is the removal (oxidation) of an electron from superoxide ($O_2^{-\cdot}$) through its reduction by Cu^{2+}. This reaction forms Cu^{1+} and nontoxic O_2.

$$Cu^{2+} - SOD + O_2^{-\cdot} \rightarrow Cu^{+} - SOD + O_2$$

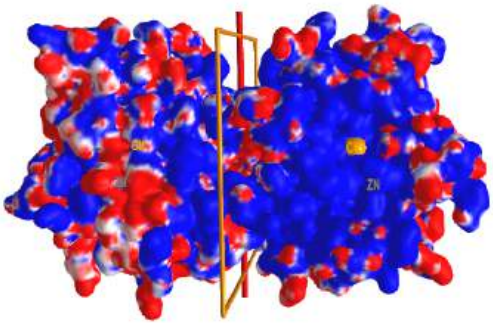
(reduction of Cu; oxidation of superoxide)

The second is the addition (reduction) of an electron from a second superoxide to Cu^{1+} to reform the catalytic Cu^{2+} and in the process form the reactive peroxide O_2^{2-} (unfortunately), which when protonated forms H_2O_2.

$$Cu^{+} - SOD + O_2^{-\cdot} + 2\,H^{+} \rightarrow Cu^{2+} - SOD + H_2O_2$$

(oxidation of copper; reduction of superoxide)

Figure 12.3.8 shows an interactive iCn3D model of the electrostatic potential surface of the superoxidase dismutase dimer showing Cu and Zn ions (2SOD).



NCBI iCn3D Figure 12.3.8: Electrostatic potential surface of the superoxidase dismutase dimer showing Cu and Zn ions (2SOD) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...zozKPVKr4MzES9

Two dimers are shown with a rotation C2 axis separating them. Red indicates the negative surface potential and blue the positive. Catalytic Cu^{2+} and Zn^{2+} ions are in each subunit and are shown in orange (Cu) and gray (Zn) spheres and labeled (disregard the label not centered on the orange and gray spheres). Note that the Cu^{2+} and Zn^{2+} ions are in the center of a large blue surface (positive potential), which helps "sweep up" any negatively charged superoxide O_2^{-} nearby. Rotate the complex around the C2 axis and you will see another large positive blue patch on the backside. These two positive electrostatic surfaces facilitate the electrostatic attractions and binding of the dangerous superoxide anion from the larger 3D region around the enzyme.

The Zn^{2+} in SOD is not redox active. What is its role? Both metals appear to increase the thermostability of the individual monomer and the dimer. In the presence of metal ions (a holo form of the enzyme), the dimer dissociates into monomers at a higher urea concentration than does the apo-dimer. Ion binding reduces the flexibilities of groups found in the dimer interface. The protein has a disulfide bond between Cys 57 and Cys 146. This bond stabilizes a metal ion binding loop that contributes to the binding interface in the dimer. The enzyme operates at diffusion-controlled rates (k_{cat}/K_M is between 10^8 and 10^9 $M^{-1}s^{-1}$), in part due to the attraction of any negatively charged superoxide by the electrostatic field around the dimer.

There are two other types of superoxidedismutates, one that uses either Mn^{2+} or Fe^{2+} (bacterial, mitochondrial, chloroplasts, protists) and another that uses Ni^{2+}(some prokaryotes).

12.3.4.4: PEROXIDE

In contrast to dioxygen which contains multiple bonds between the O atoms, peroxide has only one bond. It is quite weak and requires only 38 kcal/mol (160 kJ/mol) to break it. Remember, bonds can be broken in a heterolytic way (both electrons in a bond go to one of the atoms, or in a homolytic fashion, in which one electron goes to each atom.

Figure 12.3.9 shows typical reactions of peroxides.

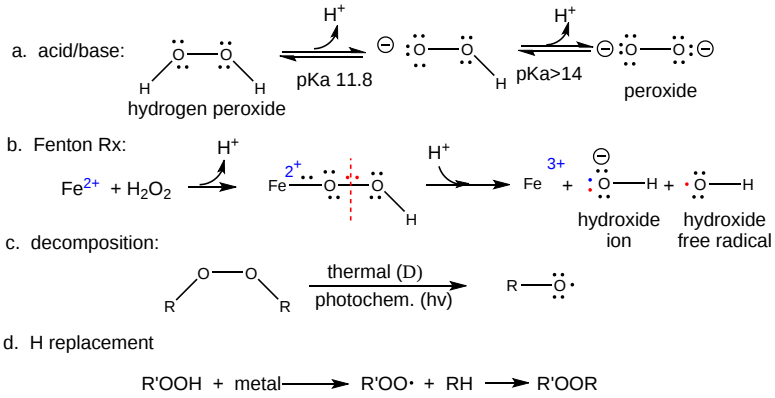


Figure 12.3.9: Reactions of peroxides

The reaction with Fe^{2+}, the **Fenton Reaction**, is similar to the reaction of triplet O_2 with Fe^{2+}. In this reaction, homolytic cleavage of the O-O bond occurs generating OH^- and the hydroxy free radical, $OH\cdot$, which will react with any molecule it encounters. Thermal or photochemical homolytic cleavage of peroxide also forms free radicals which react like the hydroxy free radical.

The enzyme **catalase** facilitates the decomposition of the reactive H_2O_2 to water and dioxygen. As such offers protection similar to that against superoxide offered by superoxide dismutase. This is the next reaction catalyzed by the enzyme catalase:

The human enzyme is a homotetramer with each monomer having a heme at the active site. The tetramer also binds $NADP^+$ (2/tetramer) but its function is unclear. A potential mechanism for catalase is shown in Figure 12.3.10.

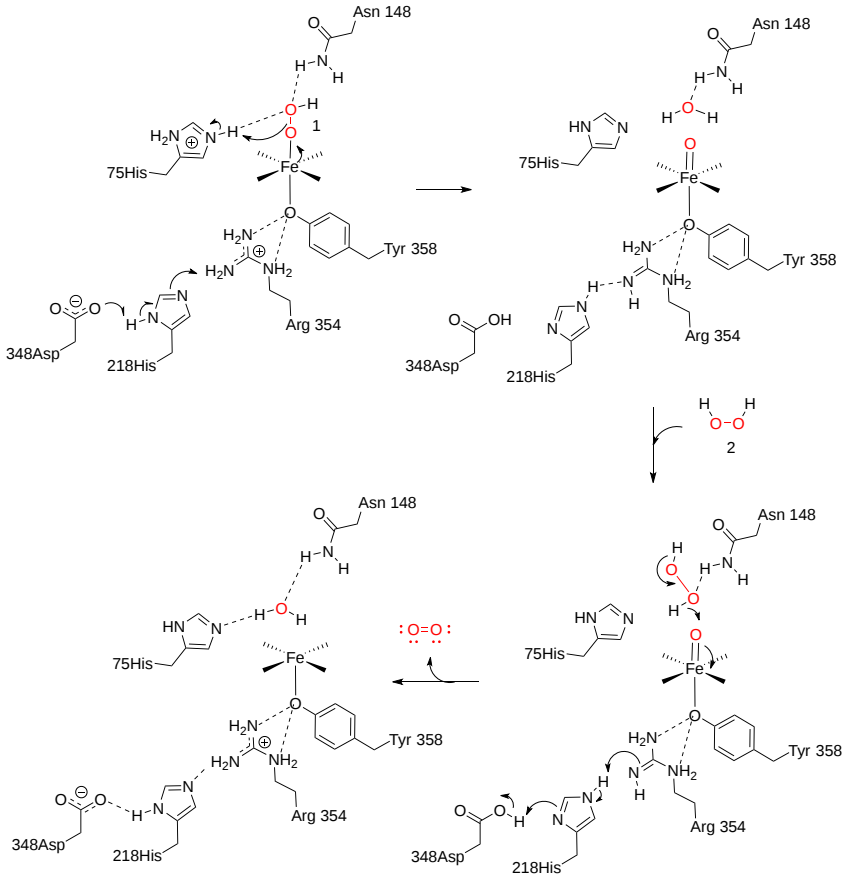


Figure 12.3.10: Human catalase reaction mechanism (adapter from Putnam et al. J. Mol. Biol. 296, 2000,https://doi.org/10.1006/jmbi.1999.3458)

Figure 12.3.11 shows an interactive iCn3D model of human erythrocyte catalase with bound ligand (CN^-) and NADPH (1DGG)

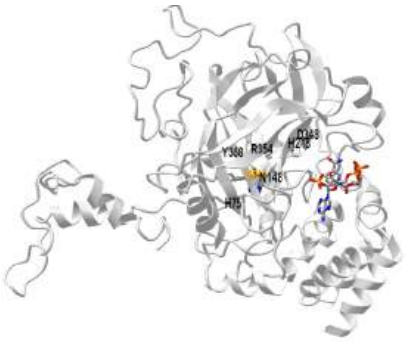


Figure 12.3.11: human catalase with bound ligand (CN^-) and NADPH (1DGG). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...TD7yJsfrNcwSJ7

The cyanide ligands in this structure bind to the heme in place of peroxide. The reaction in mammalian catalases appears to involve a tyrosine radical

H_2O_2 is very similar in structure to H_2O. How does the enzyme differentiate between them? Both approach the heme through a long 25 Å water channel with a hydrophobic constriction part way into the channel leading to the active site heme. Four waters in the crystal structure are positioned in the constricted opening and in a widened open just at the heme. The amino acid side chains forming the constriction are Val 74, Val 116, Pro 129, Phe 153, Phe154 and Trp 186. These allow only small molecules to enter. Two of the four waters form hydrogen bonds to side chains (one to His 75 and Asn 148 and another to Gln 168 and Asp12. The others don't form hydrogen bonds, are more dynamic, and more likely to leave the channel. H_2O_2 is bigger and can form bridging hydrogen bonds to side chains that the mobile waters can't. They probably leave the opening. H_2O_2 is more polar and has a higher dipole moment (2.26 Debye) compared to water (1.86 Debye) which implies that it would be differentially stabilized by hydrogen bonds in this hydrophobic site compared to the less polar water.

Figure 12.3.12 shows an interactive iCn3D model of human erythrocyte catalase (monomer) showing the hydrophobic constricture leading to the active site (1DGF**)**

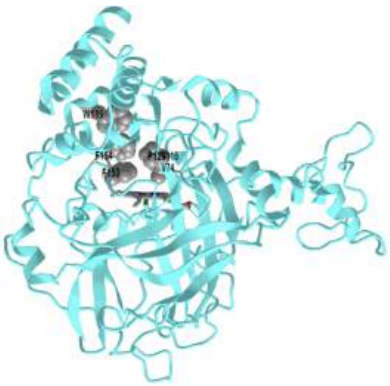


Figure 12.3.12: Human erythrocyte catalase (monomer) showing hydrophobic constriction leading to the active site (1DGF**)**. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...e48pAa9X4zhcM7

12.3.5: HYDROXYL FREE RADICAL

We won't specifically discuss the reaction of the hydroxyl free radical ($\cdot OH^-$) since it will react with anything nearby to produce another free radical. General reactions of the radical are shown in Figure 12.3.13.

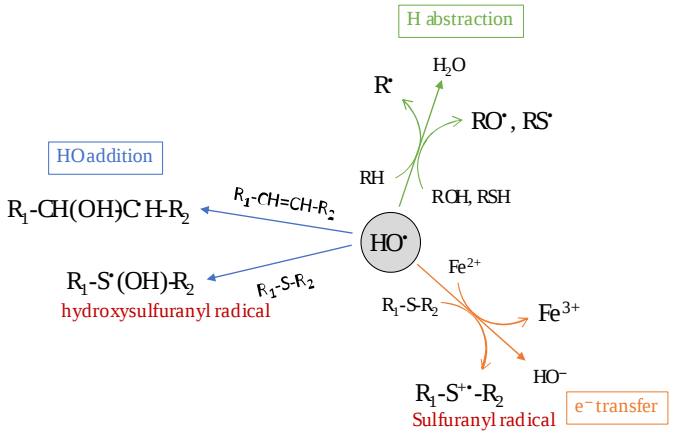


Figure 12.3.13: Chemical reactions of the hydroxyl radical. Fabrice Collin. Int. J. Mol. Sci. 2019, 20, 2407; doi:10.3390/ijms20102407. Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

In summary, we reap the benefits of using dioxygen as an oxidizing agent as it allows the aerobic and hence complete oxidation of carbohydrates and lipids to CO_2 and H_2O. Yet we pay the price for using O_2 as an oxidizing agent as its partial reduction products, the superoxide radical, peroxides, and hydroxy free radical, collectively known as **reactive oxygen species (ROS)** can react with and damage proteins, nucleic acids, and lipids, as we will see below. Figure 12.3.13 shows a summary of their reactions.

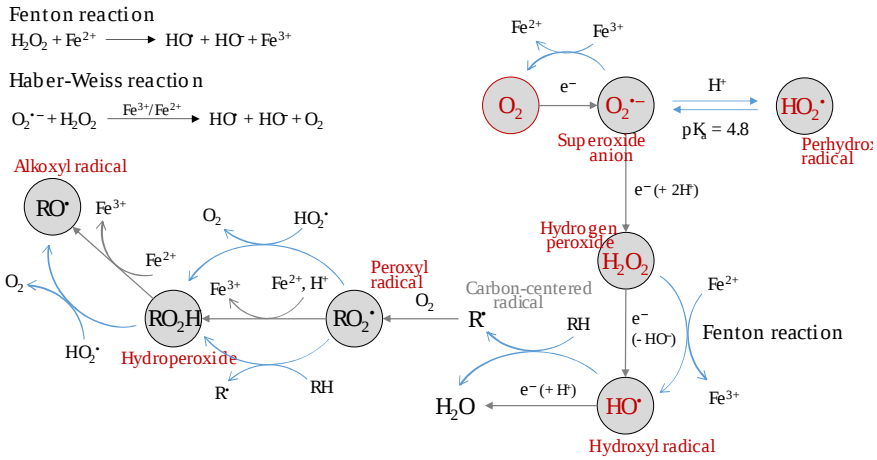


Figure (\PageIndex{14}\): Basis of Reactive Oxygen Species (ROS) generation. Fabrice Collin, ibid.

We have presented the role of superoxide dismutase and catalase in the removal of ROS in cells. In addition to these elegantly designed proteins, there is another simpler biomolecule, the tripeptide glutathione (γGlu-Cys-Gly), that can reduce ROS in cells and prevent oxidative damage. Glutathione serves as a major antioxidant in cells. It exists in reduced (GSH) and oxidized (GSSG) states, the ratio of which depends on the cellular oxidative stress. It can react with and detoxify peroxides and alkyl free radicals by the following net reactions:

$$2\ GSH + R_2O_2 \rightarrow GSSG + 2\ ROH$$

$$2GSH + 2R^{\cdot} \rightarrow GSSG + 2RH$$

12.3.6: PROTECTION FROM FE^{2+} - FERRITIN AND TRANSFERRIN

The Fenton reaction shows the potential problem with having free Fe^{2+} ions in a chemical state that would easily allow this reaction and the generation of ROS. Hence much of the Fe^{2+} in the body is sequestered in Fe^{2+} binding cofactors like heme and FeS clusters. It is transported in the blood by the protein **transferrin** and stored in cells like the erythrocyte in **ferritin**. The iron ions in both transferrin and ferritin are in the +3 oxidation state (Fe^{3+}). This strongly positive cation is quite insoluble in the presence of anions like hydroxide, phosphate, and carbonate. The K_{sp} values for the hydroxide salt of Fe^{3+} is $3.8x10^{-38}$. Let's detour for a bit and look at the structures of ferritin and transferrin and how they work, starting with ferritin.

12.3.6.1: FERRITIN

The biologically functional form of ferritin is a 24-mer. The structure encapsulates a large volume that can hold many Fe ions (up to 4500) in the central cavity. The ions are stored in the more insoluble form, Fe^{3+}, in complexes of oxide and hydroxide. Mammalian ferritin contains both heavy (H) and light (L) chains so they are hetero 24-mers.

Figure 12.3.15 shows an interactive iCn3D model of ferritin, the intracellular Fe storage protein (1fha).

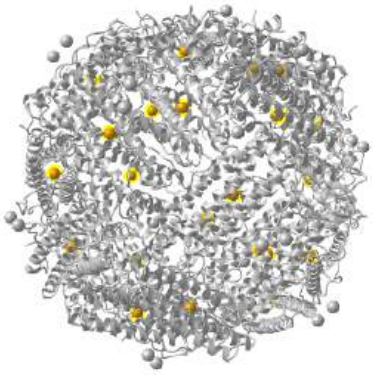


Figure 12.3.15: Ferritin, the intracellular Fe storage protein (1fha). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...MQXAbPAwNnVya8

The ferritin chains are shown in purple. This structure shows only 24 Fe ions (in orange spheres with yellow halo).

Two questions arise. How does Fe^{2+} get into the internal volume of ferritin and how does it get converted to Fe^{3+}? There must be a channel through which Fe^{2+} can diffuse. Its conversion to Fe^{3+} requires catalysis by an Fe cluster in the heavy chain (H) subunits. Figure 12.3.16s shows a generic diagram outlining the processes of uptake and conversion to Fe^{3+} salts inside the central cavity of ferritin.

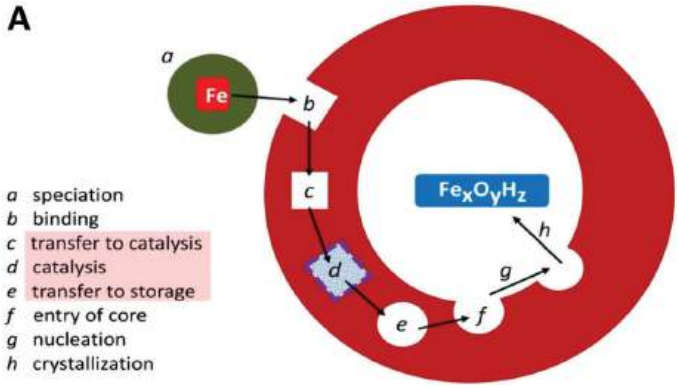


Figure 12.3.16: (A) Schematic outline of phases involved in the uptake, oxidation, and storage of Fe(II). Hagen et al. *Metallomics*, 9,(2017). https://doi.org/10.1039/c7mt00124j. Creative Commons Attribution Non-Commercial License (https://creativecommons.org/licenses/by/3.0/),

Heart and brain ferritin are enriched in the heavy (H) chain. These two organs clearly require safety from toxic Fe^{2+} ions. Ferritins in organs like the liver and spleen, which store lots of iron, are enriched in the light (L) chain. The two chains are about 50% homologous, but the H chain has a dinuclear **ferroxidase** iron site which catalyzes the Fe^{2+} to Fe^{3+} conversion. Once inside the L chain surface provides a nucleation site for the deposition of Fe^{3+} into a ferrihydrite "precipitate" ($(Fe^{3+})_2O_3 \bullet 0.5H_2O$). The general reaction is:

$$2Fe^{2+} + O_2 \rightarrow [Fe^{3+}\text{-O-O-}Fe^{3+}] \rightarrow [Fe^{3+}\text{-O(H)-}Fe^{3+}]$$

Figure 12.3.17 shows a possible generic mechanism for the oxidation of the Fe cluster from Fe^{2+} to Fe^{3+}.

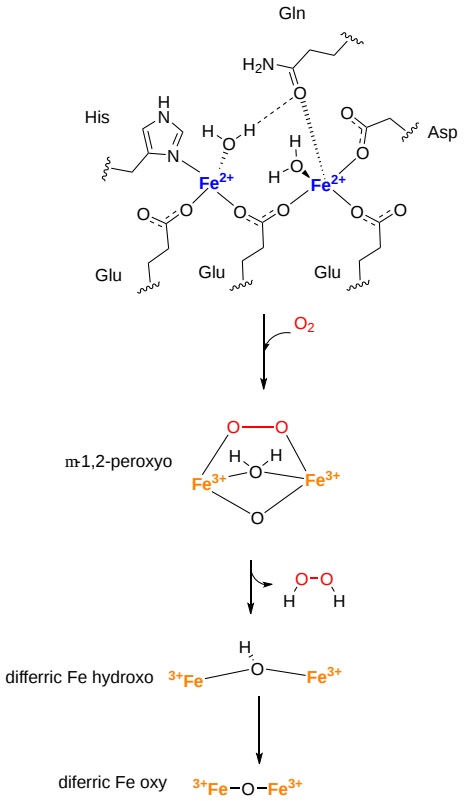


Figure 12.3.17: Possible mechanism for ferroxidase activity of ferritin heavy chain. (adapted from Hagen et al. *Metallomics*, 9 (2017). https://doi.org/10.1039/c7mt00124j)

It acts as a ferroxidase that suggests that dioxygen is involved as a ligand in the oxidation of the two Fe^{2+} ions in the cluster to Fe^{3+}.

Figure 12.3.18s shows a closeup of the interactions of the di-Fe cluster and two other Fe ions bound in the human H chain with water ligands.

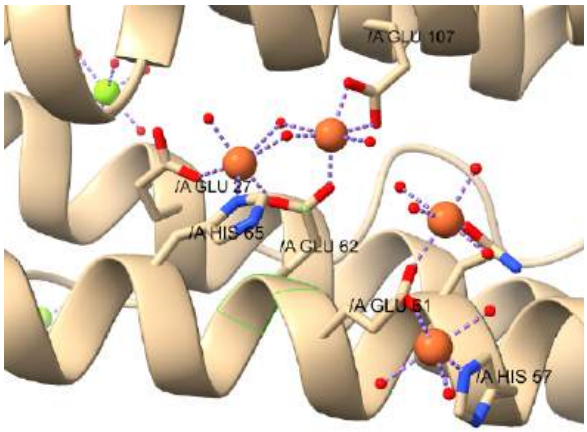
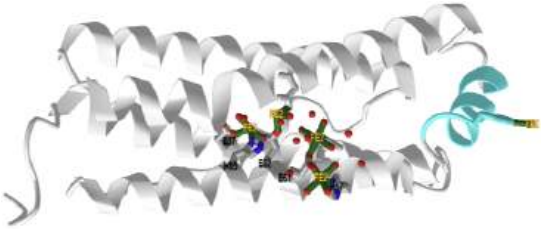


Figure 12.3.18: Interactions of the diFe cluster and two other Fe ions bound in the site with water ligands (image made with ChimeraX)

The dinuclear Fe core is shown in the central area of the figure.

Figure 12.3.19 shows an interactive iCn3D model of human heavy-chain ferritin monomer with bound Fe (4zjk)



NCBI iCn3D Figure 12.3.19: Human heavy-chain ferritin monomer with bound Fe (4zjk). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?Ux6ED11dT5dyBb2XA

The crystal structure of the full ferritin structure shows multiple binding sites and a channel to the oxidase site.

Now let's look at how the light chain L might nucleate the formation of the iron precipitates. Crystal structures show how mineralization probably occurs at a specific site on the light chains that present themselves on the inside surface of ferritin. Figure 12.3.20 shows a closeup of the interactions of a di-Fe cluster and two other Fe ions bound in the human L chain with water and peroxide ligands. This site probably represents the nucleation and mineralization site.

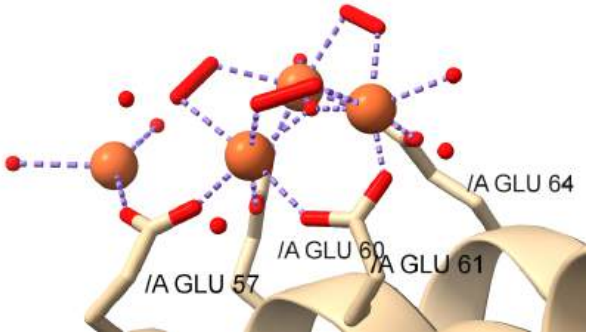
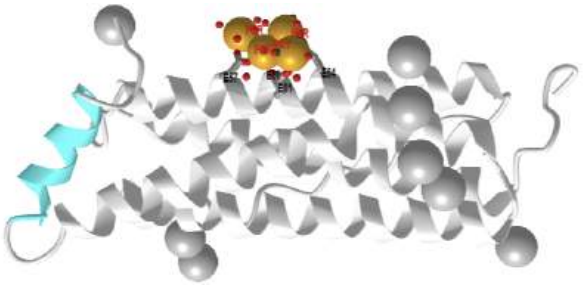


Figure 12.3.20: Closeup of the interactions of a di-Fe cluster and two other Fe ions bound in the human L chain with water and peroxide ligands at the probable mineralization site

Figure 12.3.21 shows an interactive iCn3D model of human light chain ferritin with a possible nuclear site for mineralization (5LG8). The structure was made after 60 minutes of mineralization.



NCBI iCn3D Figure 12.3.21: Human human light chain ferritin with a possible nuclear site for mineralization (5LG8). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...LKo7WJtoHeLBq9

The structure suggests the presence of a μ(3)-oxo)Tris[(μ (2)-peroxo)] triiron(III) cluster assembled at subsite on the L chains containing the carboxylate ligands Glu60, Glu61, and Glu64 side chains. A Glu57, which is along the incoming path of Fe ions, is involved in Fe delivery and coordination. Figure 12.3.22 shows the electrostatic surface around the nucleation site.

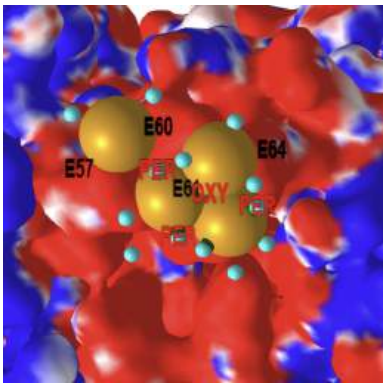


Figure 12.3.22: Electrostatic surface around the nucleation site in light chain ferritin. Here is a link to an interactive iCn3D model: https://structure.ncbi.nlm.nih.gov/i...9LKs4NFkurnhX8

Note that the Fe ions are embedded in a site of negative electrostatic potential arising, in part, from the localization of the glutamic acid side chains in the site. this

Why doesn't the heavy chain of ferritin perform the same nucleation function in preparation for the crystallization of ferrihydrite? A comparison of Figures 18 and 20 shows that the H chain (which has the ferroxidase activity) has a His 65 ligand instead of a glutamic acid

(position 60) as one of the coordinating ligands. This gives Glu 61 more flexibility which must inhibit the nucleation and mineralization process.

Figure 12.3.23 shows two top views (left and center image) and one side view of three contiguous subunits of human heavy-chain ferritin monomer with bound Fe (4zjk). The gray/black sphere is actually a Ca^{2+} ion (which is larger than Fe^{2+}) from the crystal structure. This three-monomer cluster would be replicated 8 times to form the full ferritin shell. There is a three-fold C3 axis going through the central calcium in the figure where all the monomers meet. Fe^{2+} must move through these central ions into the internal cavity.

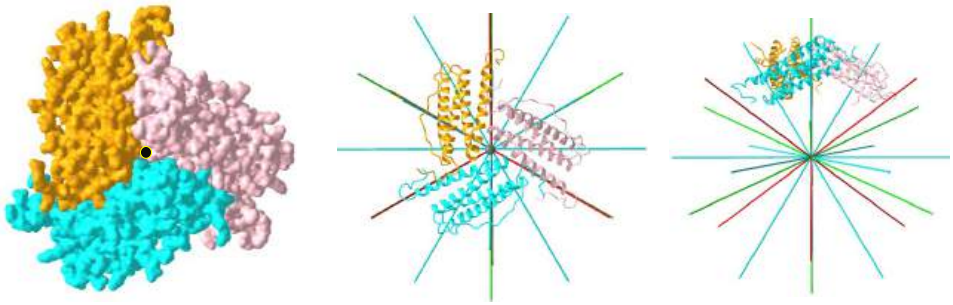
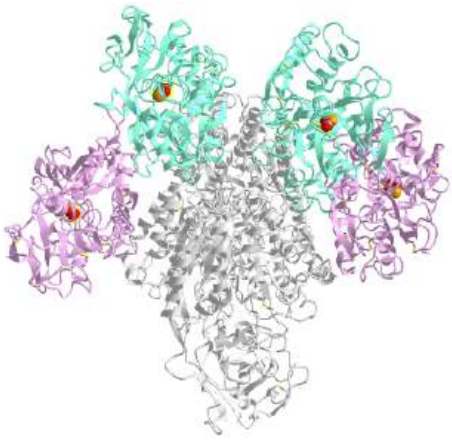


Figure 12.3.23: Opening in ferritin formed by 3 contiguous subunits of human heavy-chain ferritin monomer (4zjk)

12.3.6.2: TRANSFERRIN

Iron ions are moved in the circulation bound to the iron-binding protein **transferrin**. It binds to a transferrin receptor and can be endocytosed into the cell, where the Fe ions are transferred and stored in ferritin. The transferrin receptor can also bind and internalize circulating ferritin (see pdb 6GSR).

Figure 12.3.24 shows an interactive iCn3D model of transferrin (2N and 2C-lobes) binding to the ectodomain of the transferrin receptor (1SUV).



NCBI iCn3D Figure 12.3.24: Transferrin (2N and 2C-lobes) binding to the ectodomain of the transferrin receptor (1SUV). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...cBhTyxoSDQgH47

The transferrin receptor is shown in gray (2 monomers) bound to the N-lobes (cyan) and C-lobes (magenta) of transferrin. Each transfer lobe has a bound CO_3^{2-} (spacefill with CPK colors) and a Fe^{3+} ion (orange).

12.3.7: OXIDATIVE MODIFICATION OF PROTEINS

Many amino acid side chains can be oxidized in cells, as shown in Table 12.3.1.

Table 12.3.1: Amino acid side chain oxidation products. Berlett and Stadtman. JBC, 272, 20313-20316 (1997).
DOI:https://doi.org/10.1074/jbc.272.33.20313.

Amino acids	Oxidation products
Cysteine	Disulfides, cysteic acid
Methionine	Methionine sulfoxide, methionine sulfone
Tryptophan	2-, 4-, 5-, 6-, and 7-Hydroxytryptophan, nitrotryptophan, kynurenine, 3-hydroxykynurinine, formylkynurinine
Phenylalanine	2,3-Dihydroxyphenylalanine, 2-, 3-, and 4-hydroxyphenylalanine
Tyrosine	3,4-Dihydroxyphenylalanine, tyrosine-tyrosine cross-linkages, Tyr-*O*-Tyr, cross-linked nitrotyrosine
Histidine	2-Oxohistidine, asparagine, aspartic acid
Arginine	Glutamic semialdehyde
Lysine	α-Aminoadipic semialdehyde
Proline	2-Pyrrolidone, 4- and 5-hydroxyproline pyroglutamic acid, glutamic semialdehyde
Threonine	2-Amino-3-ketobutyric acid
Glutamyl	Oxalic acid, pyruvic acid

We will focus just on one here, the oxidation of the ε-amino group of lysine by H_2O_2 to form α-aminoadipic semialdehyde, where the amine is replaced with an aldehyde. This reaction is called the Fenton reaction as is actually a **carbonylation** reaction. A hypothetical mechanism for the oxidation of lysine side chains is shown in Figure 12.3.25.

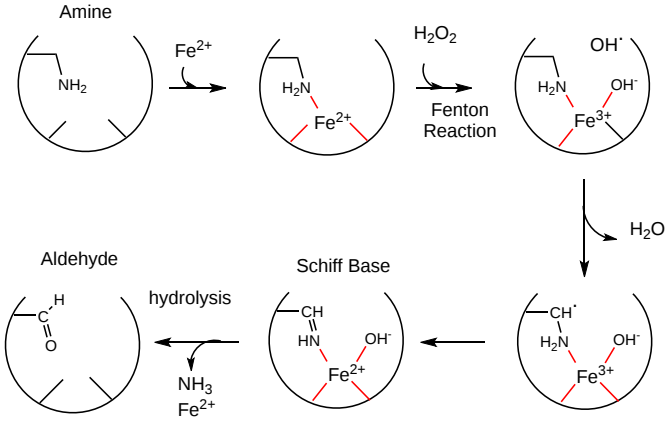


Figure 12.3.25

The next result is the oxidation of the lysine ε-amino group to an aldehyde. Oxidized levels of proteins (as evidenced by increased levels of carbonylation) increase dramatically with age (especially after age 40). The reactions seem to be catalyzed by metals and may proceed by the generation of hydroxy free radicals. Diseases associated with premature aging (Werner's Syndrome, another link to Werner's Syndrome, Progeria) show very high levels of oxidized proteins at an early age. Fibroblasts from 10 yr. old children with progeria have levels of oxidized proteins usually not seen until the age of 70. Beta-amyloid protein deposits (found in Alzheimer's and Down's Syndrome) cause neurotoxicity and death, partly by increasing superoxide production by endothelial cells, causing vasoconstriction/dilation, and ultimately disease progression. Beta-amyloid aggregates appear to increase H_2O_2 levels, in a process facilitated by Fe^{2+} and Cu^{+}. Free radical scavengers (antioxidants) may help to prevent this damage.

Carbonylation of proteins appears to be irreversible and nonrepairable. Increased carbonylation leads to misfolding and protein aggregation in ways in which protein chaperones can not reverse. The graphs in Figure 12.3.26 (the summation of many experiments) show the correlation (negative) of increasing carbonylation of proteins (red line), a measurement of oxidative damage, and the resulting decrease in protein function (green).

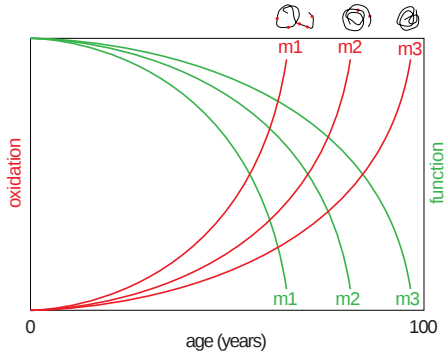


Figure 12.3.26: Susceptibility of proteoforms to oxidative modifications determines the onset of their oxidation over time and leads to their functional decline. Krisko and Radman, (2019) *Open Biol.***9.** http://doi.org/10.1098/rsob.180249. Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/.

Red dots in the top representations of protein show carbonylation. Variants of the same protein (proteoforms) that have more intrinsically disordered regions (m1) are more susceptible to carbonylation compared to more ordered variants (m3). Cancer starts to increase around 40 years of age, and the levels of carbonylation correlate with increased cancer rates and may, in part, cause it.

Lou Gehrigs Disease (Amyotrophic Lateral Sclerosis) is a disease of progressive motor neuron degeneration, which affects 1/100,000 people, and is 10-15% familial. Of the familial cases, about 25% have a mutation in superoxide dismutase I, a copper-zinc enzyme. About 2-3% of ALS patients carry 1 of 60 different dominant mutations in this enzyme. Mutations often decrease the stability of the protein which decreases Zn^{2+} affinity 5-50 fold. The A4V mutation (valine at amino acid 4 substituted for Ala) has the weakest Zn affinity and causes rapid disease progression. In the absence of Zn^{2+}, the apoprotein somehow seems to induce cell death in neurons. This superoxide dismutase also expresses a second activity. It also acts as a peroxidase which takes ROH + H_2O_2 to form an RHO (an aldehyde) plus water. In some cases, the enzyme retains normal activity against superoxide but altered peroxidase activity.

Figure 12.3.27s shows the extent of carbonylation of wild type (WT) and two mutant forms α-synuclein after exposure (in vitro) to increasing doses of γ-radiation. α-synuclein forms aggregates (Lewy bodies) in Parkinson's Disease.

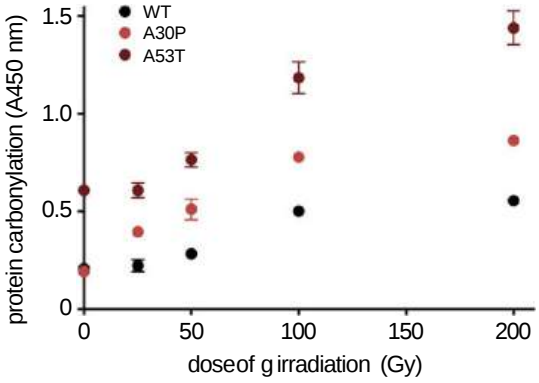


Figure 12.3.27: Wild-type version of α-synuclein is characterized by the highest oxidation resistance. The plot displays the differential resistance to γ radiation of human α-synuclein protein isomorph. Krisko and Radman, ibid.

The mutants, especially the A53T one, show significantly high extents of oxidative damage. This particular mutation is associated with the early onset (around 30) of Parkinson's Disease.

Is your hair going white?: Wood et al have shown that millimolar concentrations of hydrogen peroxide builds up in hairs that have grayed and whitened. This was associated with a decrease in catalase and in increases in Met oxidation (to Met-sulfoxide) in proteins, also associated with a decrease in the repair enzyme Met-sulfoxide reductase, Met 374 in the active site of tyrosinase, an enzyme required for the production of melanin in hair follicles, is also damaged, leading to lack of melanin, a pigment necessary for hair coloration and "senile hair graying".

ROS and Protein Folding

As discussed earlier, the cytoplasm has sufficient concentrations of "β-mercaptoethanol"-like molecules (used to reduce disulfide bonds in proteins in vitro) such as glutathione (γ-Glu-Cys-Gly) and reduced thioredoxin (with an active site Cys) to prevent disulfide bond formation in cytoplasmic proteins. Disulfide bonds in proteins are typically found in extracellular proteins, where they serve to keep multisubunit proteins together as they become diluted in the extracellular milieu. These proteins destined for secretion are cotranslationally inserted into the endoplasmic reticulum (see below) which presents an oxidizing environment to the folding protein and where sugars are covalently attached to the folding protein and disulfide bonds are formed (see Chapter 3D: Glycoproteins - Biosynthesis and Function). Protein

enzymes involved in disulfide bond formation contain free Cys which form mixed disulfides with their target substrate proteins. The enzymes (thiol-disulfide oxidoreductases, protein disulfide isomerases) have a Cys-XY-Cys motif and can promote disulfide bond formation or their reduction to free sulfhydryls. They are especially redox-sensitive since their Cys side chains must cycle between and free disulfide forms.

Reactive oxygen species (ROS) can significantly affect redox chemistry, and if present in excess can place the cell in a condition of "oxidative" stress. ROS can indiscriminately oxidize lipids, nucleic acids, and proteins, but more specifically, they may also oxidize proteins involved in creating and maintaining the normal disulfide bond formation in proteins. As the concentration of ROS increases, the concentration of cytoplasmic proteins with incorrect disulfides should increase. Using a two dimension PAGE system (first dimension run under nonreducing and the second reducing conditions) of neural cell proteins derived from cells exposed to normal and differing oxidative conditions (hydrogen peroxide or decreased intracellular glutathione levels, Cumming et al showed that oxidizing stress increased the levels of disulfide bonds in redox sensitive enzymes and, unexpectedly, among other cytoplasmic proteins involved in many aspects of life, affecting the activity of many cellular processes, suggesting that disulfide bond formation may have not only a structural but regulatory role.

12.3.8: OXIDATIVE MODIFICATION OF LIPIDS:

Figure 4 shows the most likely position in organic molecules that can form stable free radicals (allylic, benzylic position, and $3^o > 2^o >> 1^0$ carbons) are likely targets for reaction with ROS. Hence unsaturated fatty acids are extra reactive at the methylene C that separates the double bonds as shown in Figure 12.3.28.

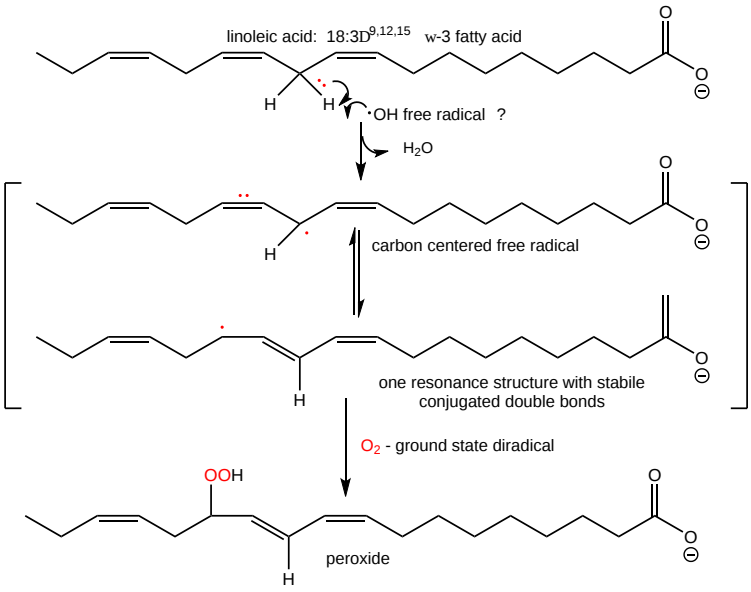


Figure 12.3.28: Reactions of fatty acids with ROS

12.3.9: LIPID AND PROTEIN OXIDATION - CARDIOVASCULAR DISEASE

The initial stages of cardiovascular disease appear to involve the development of fatty acid streaks under the artery walls. Macrophages are immune cells that have receptors that recognize oxidized lipoproteins in the blood, which they takeup. The cells then further differentiate into fat-containing **foam cells** which form the streaks. Oxidation of fatty acids in lipoproteins could produce lipid peroxides and along with the Fenton reaction lead to the oxidation of apoproteins in LDL. Cortical neurons from fetal Down's Syndrome patients show 3-4 times levels of intracellular reactive oxygen species and increased levels of lipid peroxidation compared to control neurons. This damage is prevented by treatment of the neurons in culture with free radical scavengers or catalase. Key events in atherosclerotic plaque initiation are shown in Figure 12.3.29.

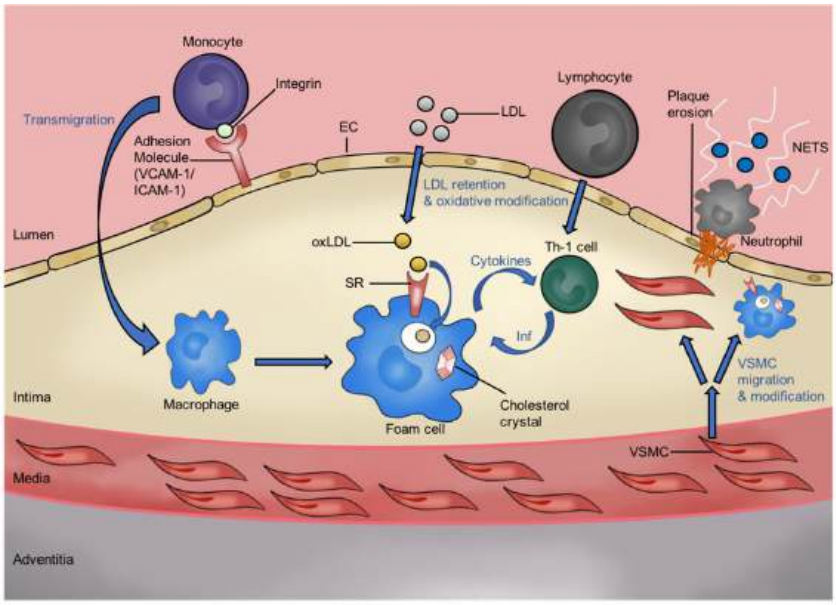


Figure 12.3.29: Key events in atherosclerotic plaque initiation. Blood LDL can accumulate within the arterial wall, in the sub-endothelial intima. This accumulated LDL can be chemically modified or oxidized: this new lipid particle species (e.g., oxLDL) promotes chronic inflammation, which promotes the trans-endothelial migration of immune cell types and foam cell development. Abbreviations: endothelial cell, EC; intercellular adhesion molecule-1, ICAM-1; interferon-γ, IFNγ; low-density lipoprotein, LDL; Neutrophil extracellular traps, NETS; oxidized low-density lipoprotein, oxLDL; scavenger receptor, SR; T- helper 1 cell, Th1; vascular cell adhesion molecule-1, VCAM-1; vascular smooth muscle cell, VSMC. Cuthbert et al. *Cells* **2020**, *9*(11), 2453; https://doi.org/10.3390/cells9112453. Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

How do fatty streaks appear under the endothelial cells? LDL oxidized in the lipid monolayer and through carbonylation of lysine side chains of the apoproteins in LDL binds to "scavenger" receptors in macrophages which have moved into the intima below the endothelial cell barrier. Scavenger receptors were first recognized to bind acetylated LDL (conversion of ε-amino groups of lysines, for example, to acetylated and uncharged derivatives). This mimics to some degree the carbonylation of the ε-amino groups to aldehydes, as shown in Figure 12.3.30.

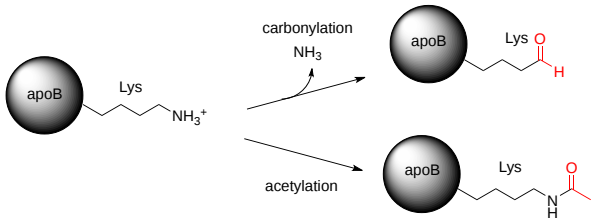


Figure 12.3.30: Carbonylation and acetylation of lysine side chains

Either modification would make the apoprotein more acidic with a lower isoelectric point since positive lysine side chains are replaced with neutral derivatives. Assuming an asymmetric distribution of the negatively charged side chains (Asp and Glu) on the apoprotein, any pre-modification negative electrostatic potential surfaces on the apoprotein would become more negative, enhancing binding to positive clusters displaying positive electrostatic potentials on scavenger receptors. Scavenger receptors often bind polyanions.

There are many 12 different classes (A-L) of scavenger receptor classes that have been identified. One, class C, is only found in drosophila. They bind a variety of polyanionic ligands and display broad binding specificity. Many in a single class have multiple names that makes their designation even more confusing. They bind a diverse set of ligands including those from bacteria and yeast (in a way similar to pathogen-associated molecular patterns - PAMPs - in the innate immune system) as well as self and modified self ligands (such as oxidized LDL - oxLDL). Once bound the ligand and scavenger receptors are taken into the cell by endocytosis for removal and degradation of the bound ligand. They can also act in signaling pathways.

The members of the scavenger receptor family are designated as illustrated in this example, **S**R-F1.1, where **S** is Scavenger, **R** is Receptor, **F** is Class, **1** is Order in class and **1** is alternatively spliced forms. Figure 12.3.31 shows domain structures of the different classes of scavenger receptors.

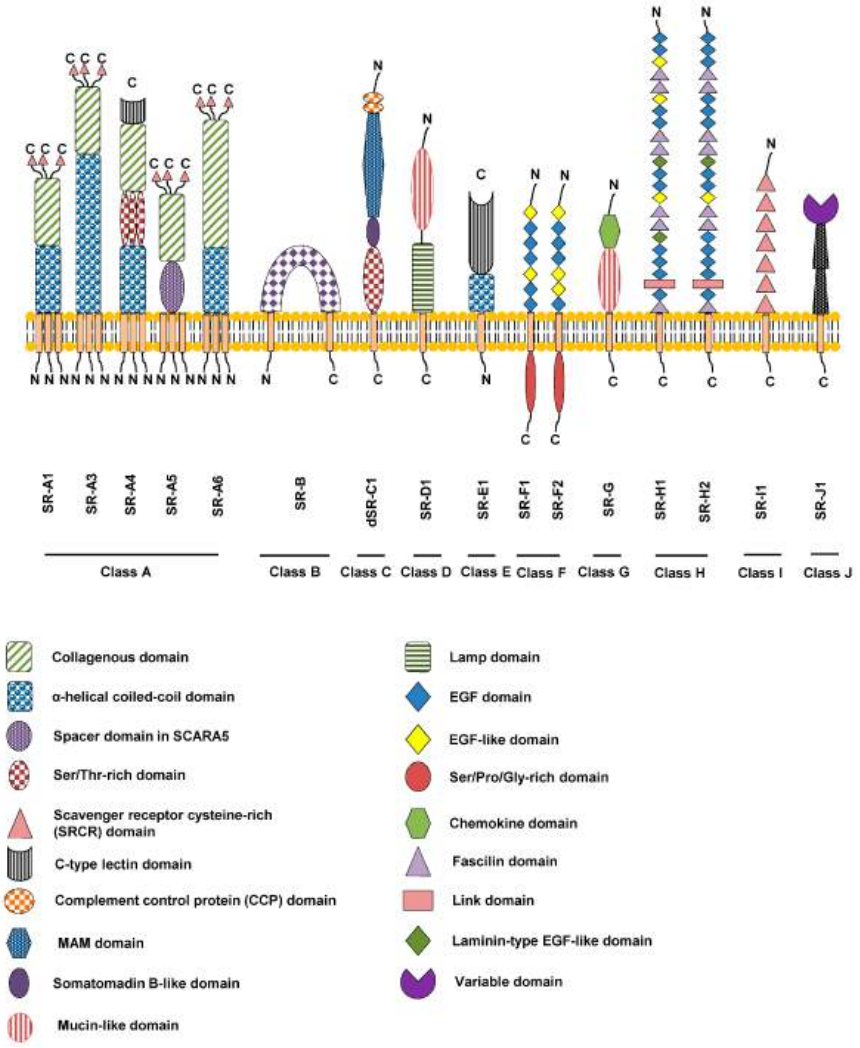


Figure 12.3.31: Domain structures of the different classes of scavenger receptors. Zani et al. *Cells* **2015**, *4*, 178-201. https://doi.org/10.3390/cells4020178. Creative Commons Attribution 4.0 International

A more detailed cartoon showing examples from a few different scavenger receptor classes involved in cardiovascular disease is shown in Figure 12.3.32.

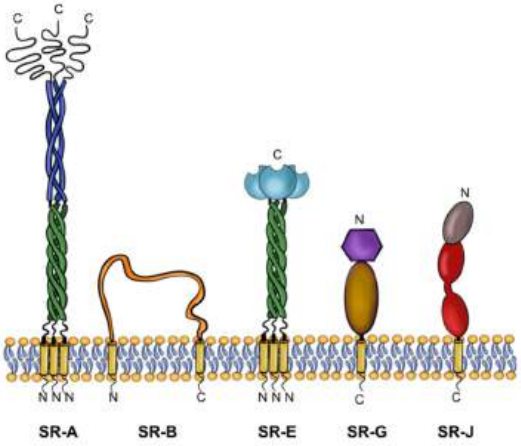


Figure 12.3.32 Five different SR classes (A, B, E, G, J) are most closely linked to the initiation and progression of atherosclerosis. These molecules are potential biomarkers and therapeutic targets. Their structures are represented in the sketches above. Abbreviations: N-terminus, N; C-terminus, C. Cuthbert et al. ibid.

Modified LDL binds to several different scavenger receptors, including SR-A1 (also called SCARA1 or CD204), SR-A2 (also called MARCO), and SR-E1 (also called Lectin-like oxidized LDL receptor 1 or LOX-1) binds oxidized and acetylated LDL.

Let's look in greater detail at two scavenger receptors that recognized oxLDL

SR-A2 (MARCO)

This scavenger receptor is a trimer that binds oxLDL, polyanions, and pathogens. It has an extracellular domain (ectodomain) formed from three monomers that are **c**ysteine **r**ich, so it's abbreviated SRCR. A five-stranded antiparallel β-sheet, and an α-helix with a large loop covering it, while the dimer has a larger 8-stranded eight-stranded β-sheet. The polyanion ligands bind presumably to the surface of the receptors with a positive (blue in figures) electrostatic potential associated with an arginine cluster. Crystal structures show that the protein also has a region of negative (red in figures) electrostatic potential which most likely is involved in metal ion binding and in the self-association of monomers to form trimeric receptors.

Given the size of the oxidized LDL (250 Å in diameter), it would not be unexpected that oxLDL binding would promote the formation of clusters of the normally trimer scavenger receptor. This is illustrated in Figure 12.3.33.

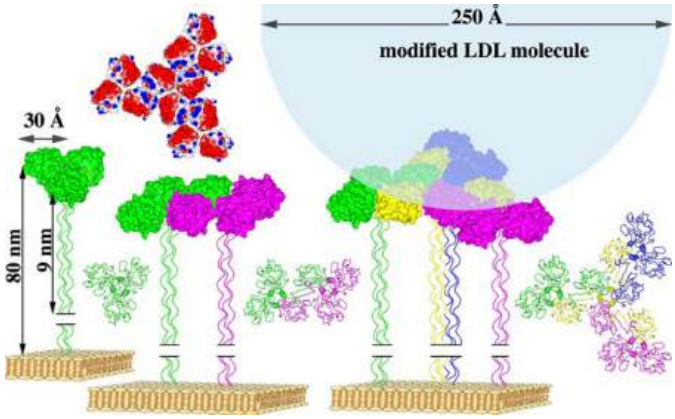
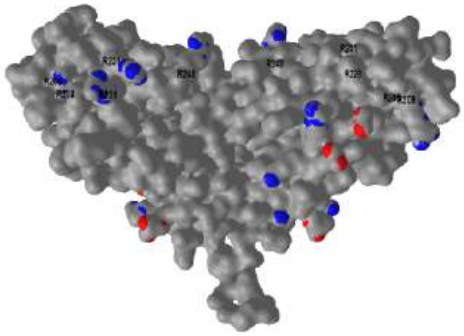


Figure 12.3.33: Model for dimerization and oligomerization of trimeric MARCO molecules. Ojalaet al. JBC **282,16654**-16666 (2007) DOI:https://doi.org/10.1074/jbc.M701750200. Creative Commons Attribution (CC BY 4.0)

Monomeric, dimeric, and oligomeric forms of **SR-A2 (**MARCO) are shown in the bottom part of the figure. The trimeric receptor molecules can form dimers and multimers by swapping domains. Multiple interactions would promote tighter binding of large ligands such as LDL and even bacteria (0.2-2 μm diameter). The assembly could proceed to the formation of oligomers (the *yellow molecule* has swapped domains with three other molecules), thus resulting in the creation of a large surface capable of interacting with large ligands, such as modified LDL (250 Å in diameter) or bacteria (0.2-2 μm). The red and blue surfaces shown above the trimer represent the negative (red) and positive (blue) electrostatic surface electrostatic potential of the oligomeric from top down.

SR-E1 (Lox1 or Lectin-like oxidized LDLR or Oxidized low-density lipoprotein receptor 1.

LOX-1 is expressed on macrophages, dendritic cells, endothelial cells, platelets, smooth muscle cells, and adipocytes. It binds oxLDL, some bacteria (through their negatively charged cell walls), and even apoptotic cells. Figure 12.3.34 shows an interactive iCn3D model of the extracellular C-type lectin-like domain of dimeric human Lox-1 (1YPQ)

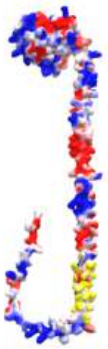


NCBI iCn3D Figure 12.3.34: Extracellular C-type lectin-like domain of dimeric human Lox-1 (1YPQ). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...qGY4HBzPjjs7P7

The dimer in this structure is connected by a disulfide bridge. The two monomers are shown in gray spheres. They come together to form a heart-like structure A series of arginines (with blue spheres for the surface Ns and labeled) are shown in the cradle of the heart shape. These most likely interact with the oxidized apoprotein B of the oxLDL. Other blue (N) and red (O) spheres near to each other can form salt

bridges and may interact with zwitterion heads of LDL surface lipids such as phosphatidylcholine, sphingomyelin, or phosphatidylethanolamine.

Figure 12.3.35 shows an interactive iCn3D model of an AlphaFold predicted model of human oxidized LDL receptor - LOX (P78380)



NCBI iCn3D Figure 12.3.35: . (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...rLoi7ZK4K3SsH8

The yellow spacefill indicates the transmembrane segment. Rotate the model along the long axis and you will see one face of the top domain has a red (negative electrostatic potential) face while the opposite side has blue (positive potential).

12.3.10: OXIDATIVE MODIFICATION OF DNA

Significant evidence suggests oxygen free radicals are linked to aging and diseases. Mutations caused by hydroxylation reactions (presumably from the generation of hydroxyl free radicals as shown above) can potentially lead to cancer. A particularly nasty reaction is the insertion of the hydroxy radical into bases in DNA. Figure 12.3.36 shows the hydroxylation at position 8 of guanine to produce 8-oxy-G and at positions 5 and 6 in thymine.

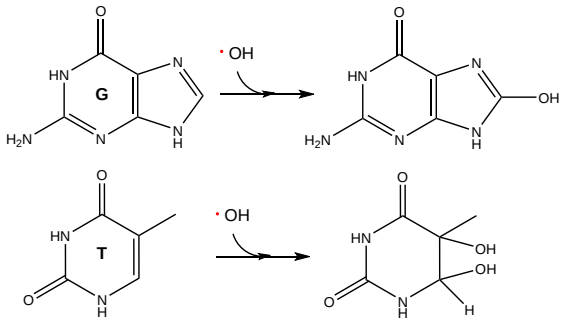
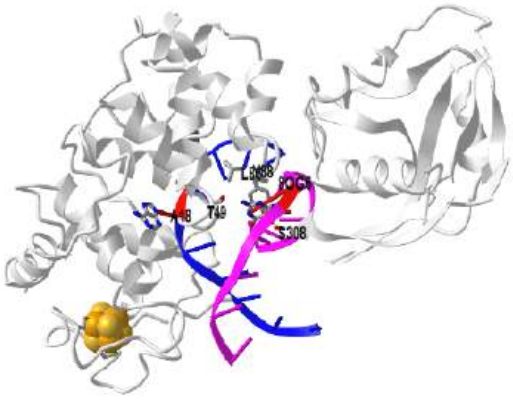


Figure 12.3.36 Hydroxylation of G and T bases by the hydroxyl free radical

Mitochondrial DNA is more susceptible to oxidation than is nuclear DNA. The human mitochondrium has a small genome (16.5 Kb compared to the nuclear genome of 3 Gb) which code 13 protein subunits involved in respiration, 22 tRNAs and two ribosomal RNAs. (The mitochondria presumably are vestiges of a bacteria which invaded an early cell and established a symbiotic relationship with the cell). There is an inverse correlation of oxidized mitochondrial DNA [8-oxoG] with the maximal life span of an organism, but this correlation is not seen with nuclear DNA. Presumably, the nuclear DNA is somewhat protected from oxidative damage since it is bound to histone proteins (which form nucleosome core particles with DNA) and by DNA repair enzymes. DNA repair enzymes that are encoded in the nucleus are found in the mitochondria and mitochondrial DNA is packaged with mitochondrial transcription factor A (TFAM). Examination of human bladder, head and neck, and lung primary tumors reveals a high frequency of mitochondrial DNA mutations. In addition, most dioxygen use by the cell occurs in the mitochondria. Hence this organelle probably faces the highest concentration of toxic oxygen reduction products. Recently, the crystal structure of an enzyme, adenine DNA glycosylase (MutY), that repairs 8-oxyG modified DNA has been determined in complex with the oxidatively damaged DNA. If not repaired, the 8-oxyG base pairs with adenine instead of cytosine, causing a GC to AT mutation on DNA replication.

Figure 12.3.37 shows an interactive iCn3D model of adenine mispaired with 8-oxoguanine by MutY adenine DNA glycosylase (1RRQ)



NCBI iCn3D Figure 12.3.37: Adenine mispaired with 8-oxoguanine by MutY adenine DNA glycosylase (1RRQ) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?4RiEc6TuqmexiW1r7

The protein MutY, which catalyzes the base excision and repair, is shown in gray. The DNA strands are shown in magenta and blue. 8-OxyG on the magenta strand is labeled 8OG7 and is shown in CPK-colored sticks. Its mismatched adenine base pair partner, labeled A18 on the blue strand, is shown in CPK-colored sticks. Notice its orientation is kinked away from the orientation in a canonical base pair. Key amino acid side chains (Thr49, Leu86, Tyr88, and Ser308) interacting with the 8-oxyG are shown in CPK-colored sticks and labeled.

Although oxidative damage in mitochondria clearly can promote premature aging, other independent mechanisms may also. Kujoth et al. developed a mouse model that expressed a mutant form of mitochondrial DNA polymerase that was defective in the proofreading activity of the enzyme. These mice displayed premature aging but showed no increased levels of oxidized mitochondrial lipids or hydroxylated G residues in mitochondrial DNA. They did show significant activation of a cytosolic enzyme called caspase-3, which when active lead to the programmed death of cells (a process called apoptosis). This calcium-activated aspartic acid protease (with an active site Asp) is activated by binding mitochondrial cytochrome C that has "leaked" into the cytoplasm from its normal location in the intermembrane space in mitochondria. The process is usually associated with DNA damage (mutations, fragmentation) that would arise if the proofreading function of DNA polymerase was defective. This was indeed found in these mice.

Oxidative damage to biomolecules might not initiate aging and disease processes, but rather might be a secondary effect of other initiating events. .Reversing or preventing oxidative damage might slow the progression of aging and disease. Aging is a complex feature of organisms and would be expected to have complex causes and biological effects. At the organismal level, aging has been studied in the roundworm C. elegans which lives for only a few weeks. Genetic analyses can be easily used to find gene alterations associated with premature aging. One hormonal system that has recently been associated with aging in eukaryotes (and in C. elegans) involves the signaling pathways for insulin and insulin growth factor I (IGF-1), which regulate carbohydrate, lipid, and reproductive pathways in C. elegans. Mutations that decrease signaling from this pathway increase C. elegans life span. These mutations lead to increased activity of the DAF16 transcription factor, which upregulates the expression of many genes. In contrast, wild-type organisms, when exposed to insulin or IGF-1, decrease the activity of DAF16. Using DNA microarrays, investigators determined which DAF16-controlled genes were upregulated in mutant worms in the mid-life point of the organism. These genes included, among others, peroxisomal and cytosolic catalase, Mn-superoxide dismutase, cytochrome P450s, metallothionein-related Cd-binding protein, and heat shock proteins. We will investigate the function of several of these gene products in the next section, but needless to say, they are all involved in cellular responses to stress, often involving dioxygen metabolites. The over-expression of mitochondrial catalase in mice increased their lifespan by 20%. It has also been showed that decreased levels of insulin-like growth factor also promote longevity in mice, indicating again that mechanisms in addition to oxidative damage by ROS are involved in aging.

12.3.11: BENEFICIAL OXIDATION OF PROTEINS: OXIDATIVE BURST IN MACROPHAGES

There are cases in which oxidative damage to protein and lipids is desirable. One example involves the role of macrophages in the immune system in eliminating foreign microorganisms. When macrophages recognize and engulf microbes, one mechanism deployed in killing the microorganism is through oxidative damage. The stimulated macrophages undergo an oxidative burst which leads to increased oxygen

utilization. One outcome of this is the generation of ROS. The activation of the ROS-generating system can also kill the macrophage (which is OK).

In addition, immune function decreases with age. This probably also occurs through damage from ROS. Telomeres at the end of chromosomes are also shortened by oxidative stress and irradiation. They are enriched in guanine bases the have may repeat (thousands) of TTAGGG sequences and hence are susceptible to oxidation. However, as you might expect, macrophages have developed ways to limited self-damage by ROS.

In the presence of ROS, the macrophage or mitochondrial kinase Mst1/2 are recruited to their respective membranes from the cytosol. The enzyme acts as a ROS sensor and "attenuator" though the phosphorylation and stabilization of a protein Keap1 that binds to a transcription factor Nrf2 (also called NFE2L2), a transcription factor. When bound to unphosphorylated Keap1, the transcription factor Nrf2 is target for proteolysis. Keap1 phosphorylation prevents its binding to Nrf2. Free Nrf2 then can translocate to the nucleus where it promotes transcription of antioxidant proteins such as glutamate-cysteine ligase catalytic subunit (Gclc), which catalyzes the first step and rate-limiting step of glutathione (g-glutamyl-cysteinyl-glycine) synthesis. These processes are illustrated in Figure 12.3.38.

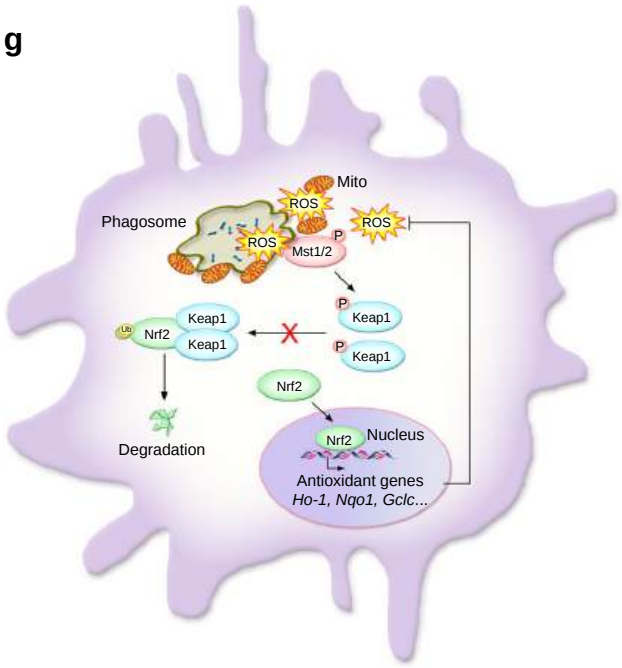


Figure 12.3.38 A proposed working model for kinases Mst1/2 sense ROS and maintain cellular redox balance by modulating the stability of Nrf2. Wang, P., Geng, J., Gao, J. *et al.* Macrophage achieves self-protection against oxidative stress-induced aging through the Mst-Nrf2 axis. *Nat Commun* **10,** 755 (2019). https://doi.org/10.1038/s41467-019-08680-6. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Phagosomal or mitochondrial ROS release attracts Mst1/2 to the membrane of phagosome or mitochondrion from the cytosol and activates Mst1/2; Mst1/2
phosphorylate Keap1 to stabilize Nrf2 and regulate the expression of antioxidant enzymes to protect the cell against oxidative damage.

12.4: BIOLOGICAL OXIDATION-REDUCTION REACTIONS

12.4.1: GENERAL OXIDIZING AGENTS

Before we consider common biological oxidizing agents, lets look back at ones you saw in other chemistry classes. Oxidizing agents are required to oxidize organic molecules. In organic lab, you never used dioxygen as an oxidizing agent. It is difficult to limit the extent of oxidation using dioxygen. In addition, side reactions are likely given the nature of the reactive oxygen reduction products. The mechanisms of combustion reactions of organic molecules with dioxygen to produce carbon dioxide and water are very complicated. Table 12.4.1 belows some key steps in the combustion of methane.

Initiation

$$\mathrm{CH_4 \rightarrow CH_3^{\cdot} + H^{\cdot}}$$

$$\mathrm{O_2 \rightarrow 2\,O^{\cdot}}$$

Propagation

$$\mathrm{CH_4 + H^{\cdot} \rightarrow CH_3^{\cdot} + H_2}$$

$$\mathrm{CH_4 + HO^{\cdot} \rightarrow CH_3^{\cdot} + H_2O}$$

$$\mathrm{CH_3^{\cdot} + O^{\cdot} \rightarrow CH_2O + H^{\cdot}}$$

$$\mathrm{CH_2O + HO^{\cdot} \rightarrow CHO^{\cdot} + H_2O}$$

$$\mathrm{CH_2O + H^{\cdot} \rightarrow CHO^{\cdot} + H_2}$$

$$\mathrm{CHO^{\cdot} \rightarrow CO + H^{\cdot}}$$

$$\mathrm{CO + HO^{\cdot} \rightarrow CO_2 + H^{\cdot}}$$

Branching

$$\mathrm{H^{\cdot} + O_2 \rightarrow HO^{\cdot} + O^{\cdot}}$$

Termination

$$\mathrm{H\cdot + R^{\cdot} + M \rightarrow RH + M^{\cdot}}$$

In chemistry, other oxidizing agents are often used, including permanganate and chromate. Their mechanism of action is illustrated in Figure 12.4.1.

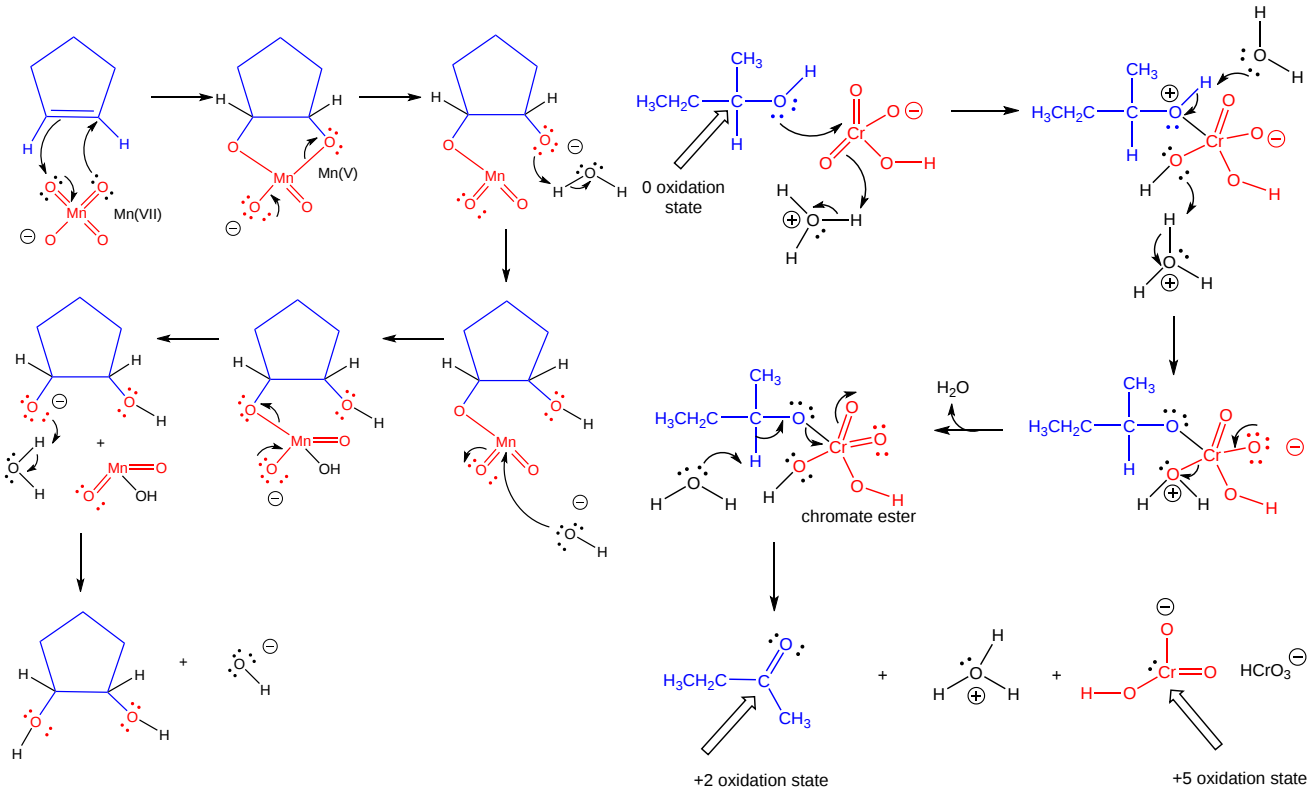


Figure 12.4.1: Mechanism of oxidation of organic molecules by permanganate and chromate

Oxygen can often be inserted into a molecule in a nonoxidative process by hydration of an alkene to an alcohol (a readily reversible reaction), which could then be oxidized to either an aldehyde/ketone or carboxylic acid using an appropriate oxidizing agent.

Most biological oxidation reactions (such as those found in glycolysis, Kreb Cycle, and fatty acid oxidation) do not use dioxygen as the immediate oxidizing agent. Rather they use nicotinamide adenine dinucleotide (NAD+) or flavin adenine dinucleotide (FAD) as oxidizing agents, which get reduced. Enzymes that uses these oxidizing agents are usually called **dehydrogenases**. Dioxygen can also be used to introduce oxygen atoms into biological molecules in oxidative reactions. Enzymes that introduce one oxygen atom of dioxygen into a molecule (and the other oxygen into water) are called **monooxygenases**. (Note: some monooxygenase that hydroxylate biomolecules are called **hydroxylases**.) Those that introduce both atoms of dioxygen into a substrate are called **dioxygenases**. These oxygenases are not usually used to oxidize organic molecules for energy production. Rather they introduce O atoms for other reasons, including increasing the solubility of nonpolar aromatics to facilitate secretion, and to produce new molecular species which have different biological activities. Finally, biological molecules can be oxidized by dioxygen in which no atoms of oxygen are added to the substrate. Rather, electrons lost from the oxidized substrate are passed via intermediate electron carriers to dioxygen , which get reduced to superoxide (if one electron is added), hydrogen peroxide (if two electrons are added) or water (if 4 electrons are added). These enzymes are called oxidases. (Note: The letters oxi- or oxygen- are used in all the enzymes that use dioxygen as the oxidizing agent.)

In this chapter section, we will discuss biological oxidation reactions. Most introductory biochemistry texts don't approach oxidation reactions in one cohesive chapter. Probably because of that, when I was learning biochemistry, I found the presentation of these different enzymes involved in redox reactions to be very confusing. Hopefully this section will alleviate that problem. First the chemistry of NAD^+ and FAD will be discussed. Then the enzymes using dioxygen in oxidative reactions (monooxygenases, dioxygenases, and oxidases) will be explored.

12.4.2: THE CHEMISTRY OF NAD^+ AND FAD

NAD^+ is a derivative of nicotinic acid or nicotinamide, as illustrated in Figure 12.4.2.

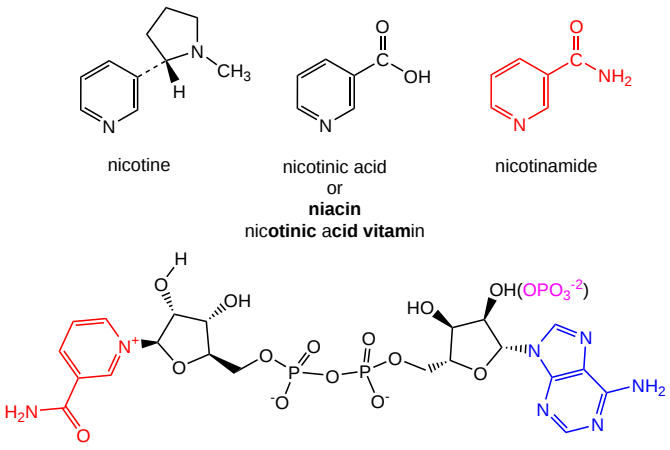


Figure 12.4.2: Structure of niacin derivatives

It and its reduction product, NADH, exists in the cells as interconvertible members of a pool whose total concentration does not vary significantly with time. Hence, if carbohydrates and lipids are being oxidized by NAD+ to produce energy in the form of ATP, levels of NAD+ would begin to fall as NADH rises. A mechanism must be be present to regenerate NAD+ from NADH if oxidation is to continue. As we will see later, this happens in the muscle under anaerobic conditions (if dioxygen is lacking as when you are running a 100 or 200 m race, or if you are being chased by a saber-toothed tiger) when pyruvate + NADH react to form lactate + NAD^+. The reaction is shown in Figure 12.4.3.

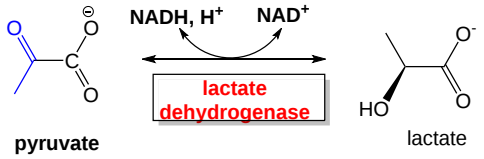


Figure 12.4.3: Conversion of pyruvate to lactate

Under aerobic conditions (sufficient dioxygen available), NADH is reoxidized in the mitochondria by electron transport through a variety of mobile electron carriers, which pass electrons to dioxygen (using the enzyme complex cytochrome C oxidase) to form water.

NAD^+/NADH can undergo two electron redox steps, in which a hydride is transferred from an organic molecule to the NAD^+, with the electrons flowing to the positively charged nitrogen of NAD^+ which serves as an electron sink. NADH does not react well with dioxygen, since single electron transfers to/from NAD^+/NADH produce free radical species which can not be stabilized effectively. All NAD^+/NADH reactions in the body appear to involve 2 electron hydride transfers. Figure 12.4.5 shows both 1 and 2 electrons to NAD^+.

Figure 12.4.4:

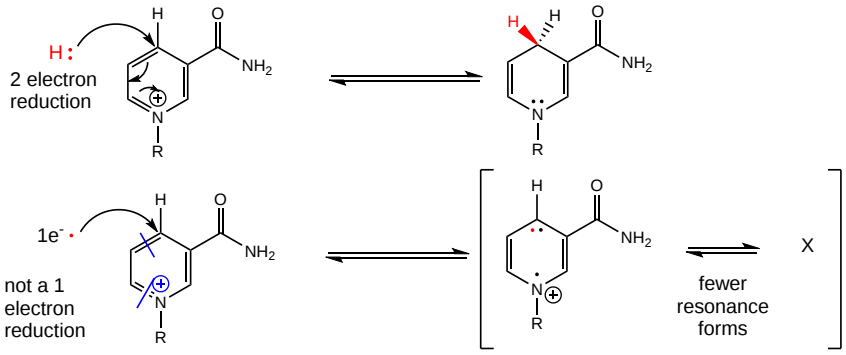


Figure 12.4.4: 1 and 2 electrons to NAD^+

FAD (or flavin mononucleotide-FMN) and its reduction product, $FADH_2$, are derivatives of riboflavin, as shown in Figure 12.4.5.

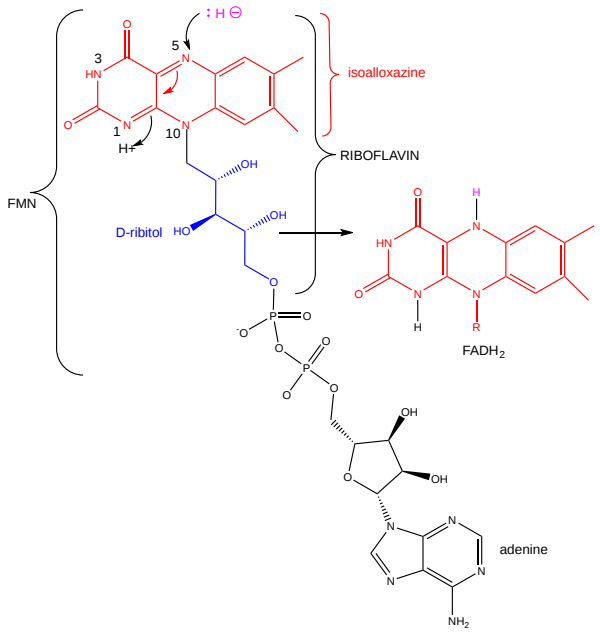


Figure 12.4.5: Structures of riboflavin, FMN and FAD

$FAD/FADH_2$ differ from $NAD^+/NADH$ since they are bound tightly (K_D approx 10^{-7} - 10^{-11} M) to enzymes which use them. This is because $FADH_2$ is susceptible to reactions with dioxygen, since $FAD/FADH_2$ can form stable free radicals arising from single electron transfers. $FAD/FADH_2$ can undergo 1 OR 2 electrons transfers. This is illustrated in Figure 12.4.6.

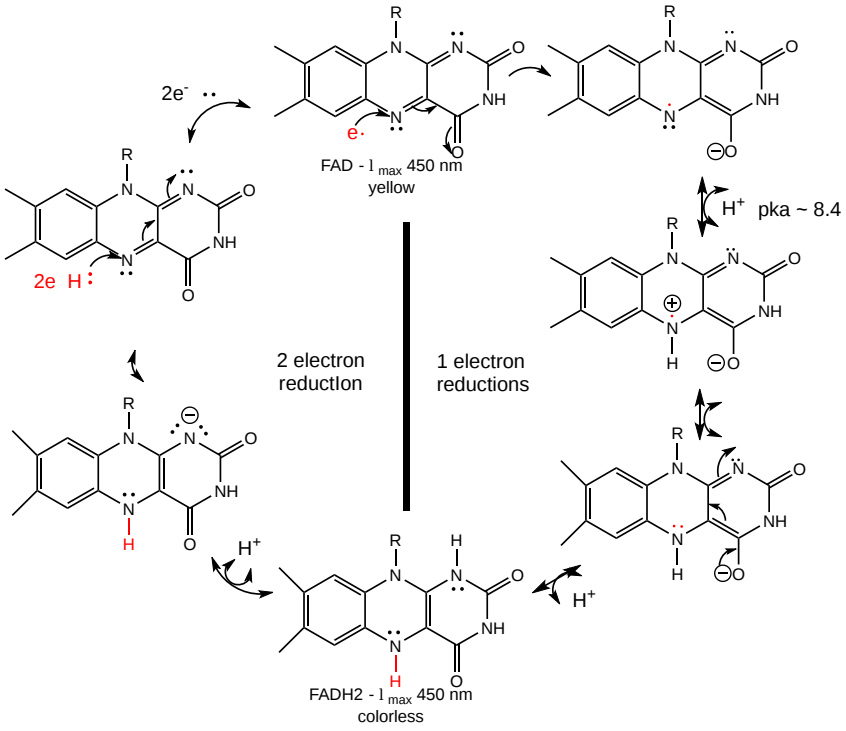


Figure 12.4.6: 1 and 2 electrons reduction of FAD

$FAD/FADH_2$ are tightly bound to enzymes so as to control the nature of the oxidizing/reducing agents that interact with them. (i.e. so dioxygen in the cell won't react with them in the cytoplasm.) If bound FAD is used to oxidize a substrate, the enzyme would be inactive in any further catalytic steps unless the bound $FADH_2$ is reoxidized by another oxidizing agent.

12.4.3: DEHYDROGENASES

These enzymes use NAD^+/NADH or FAD/$FADH_2$ and are named for the substrate that is oxidized. For instance in the reaction:

$$\text{pyruvate} + \text{NADH} \rightleftharpoons \text{lactate} + \text{NAD}^+$$

which is used to regenerate NAD^+ under anerobic conditions, the enzyme is named lactate dehydrogenase. As in acid/base reactions, when the preferred direction for the reaction (from a ΔG^o perspective) is from stronger acid to weaker (conjugate) acid, the preferred direction for a redox reaction is in the direction from strong to weak oxidizing/reducing agents. This can easily be determined from charts of standard reduction potentials, and using the equation: $\Delta G^o = -nFE^o$,

- where F is the Faraday constant (96,494 Coulombs/mol e^- = 96, 494 J/(V.mol) = 23.06 kcal/(V·mol) or 96 kJ/(V·mol). One Faraday is the charge per one mol of electrons).
- and E^o, the standard EMF or standard cell potential (total voltage at standard state conditions), which can be determined by adding the standard reduction potentials (**E^o**) for the two appropriate half-reactions, after reversing the equation for the half-reaction that represents the oxidation. Hence $E^o = E^o_{reduction} - E^o_{oxidation}$.

When n=2 (number of electrons) which is common for oxidations of organic molecules,

$$\Delta G^o \text{ (kcal/mol)} = -46.12E^o \text{ or approximately } -50E^0 \text{ (or } -193E^o \text{ kJ/mol)}$$

Notice when $E^o > 0$, $\Delta G^o < 0$, the reaction as written is favored under standard conditions. Note in the table below that many of the half reactions involve protons. For biological reactions involving free protons, the standard state concentration for the protons are not 1 M as for other solutes in solution, but defined to be the hydronium ion concentration at pH 7.0. The E^o and ΔG^o values for the reactions involving hydrogen ions at a standard state of pH 7.0 are usually written as $E^{o'}$ and $\Delta G^{o'}$.

Common standard reduction potentials are shown in Table 12.4.2 below.

Table 12.4.2: Standard reduction potentials for common biomolecules

oxidant	reductant	n (electrons)	$E^{o'}$ (volts), 25°C
Acetate + carbon dioxide	pyruvate	2	-0.70
succinate + CO_2 + $2H^+$	α−ketoglutarate + H_2O	2	-0.67
acetate	acetaldehyde	2	-0.60
glycerate-3-P	glyceraldehyde-3-P + H_2O	2	-0.55
O_2	O_2^-	1	-0.45
ferredoxin (ox)	ferredoxin (red)	1	-0.43
Carbon dioxide	formate	2	-0.42
$2H^+$	H_2	2	-0.42
α-ketoglutarate + CO_2 + $2H^+$	isocitrate	2	-0.38
acetoacetate	β-hydroxybutyrate	2	-0.35
Cystine	cysteine	2	-0.34
Pyruvate + CO_2	malate	2	-0.33
NAD^+ + $2H^+$	NADH + H^+	2	-0.32
$NADP^+$ + $2H^+$	NADPH + H^+	2	-0.32
FMN (enzyme bound)	$FMNH_2$	2	-0.30
Lipoic acid, ox	Lipoic acid, red	2	-0.29
1,3 bisphosphoglycerate + $2H^+$	glyceraldehyde-3-P + Pi	2	-0.29
Glutathione, ox	Glutathione, red	2	-0.23
FAD (free) + $2H^+$	$FADH_2$	2	-0.22
Acetaldehyde + $2H^+$	ethanol	2	-0.20
Pyruvate + $2H^+$	lactate	2	-0.19
Oxalacetate + $2H^+$	malate	2	-0.17
α-ketoglutarate + NH_4^+	glutamate	2	-0.14
FAD + $2H^+$ (bound)	$FADH_2$ (bound)	2	0.003-0.09
Methylene blue, ox	Methylene blue, red	2	0.01
Fumarate + $2H^+$	succinate	2	0.03
CoQ (Ubiquinone - UQ) + H^+	UQH	1	0.031
UQ + $2H^+$	UQH_2	2	0.06
Dehydroascorbic acid	ascorbic acid	2	0.06
Ubiquinone; ox	red	2	0.10
Cytochrome b2; Fe^{3+}	Cytochrome b2; Fe2+	1	0.12
Cytochrome c1; Fe3+	Cytochrome c1; Fe^{2+}	1	0.22
Cytochrome c; Fe^{3+}	Cytochrome c; Fe^{2+}	1	0.25
Cytochrome a; Fe^{3+}	Cytochrome a; Fe^{2+}	1	0.29
1/2 O_2 + H_2O	H_2O_2	2	0.30
Cytochrome a3; Fe^{3+}	Cytochrome a3; Fe^{2+}	1	0.35
Ferricyanide	ferrocyanide	2	0.36
Cytochrome f; Fe^{3+}	Cytochrome f; Fe^{2+}	1	0.37
Nitrate	nitrite	1	0.42
Photosystem P700	.	.	0.43
Fe^{3+}	Fe^{2+}	1	0.77
1/2 O_2 + $2H^+$	H_2O	2	0.816

The mechanism for the oxidation of a substrate by NAD^+ involves concerted hydride transfer to one face of NAD^+. Hydride transfer is possible since water is excluded from the active site. It is facilitated by removal of a proton from an oxygen on an alcoholic substituent, for example, adjacent to the departing hydride. The negative charge on the oxide acts as a "source" of electrons, which can then flow through the hydride transfer to the positively charged ring nitrogen of NAD^+, which acts as an electron "sink". This is crudely illustrated in the cartoon shown in Figure 12.4.7, which shows the oxidation of ethanol by the enzyme alcohol dehydrogenase.

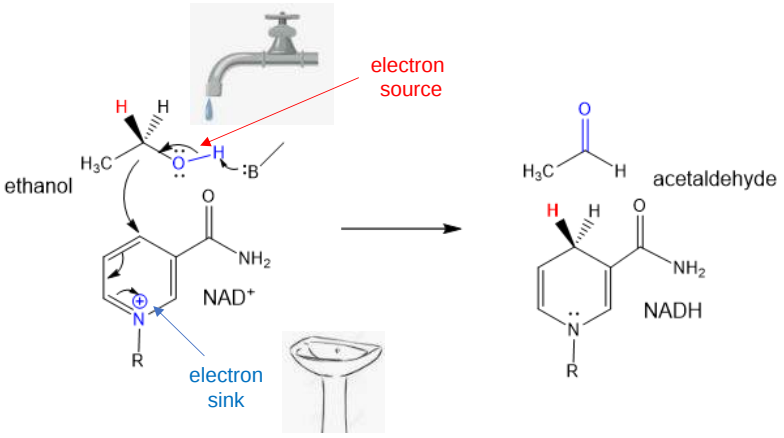


Figure 12.4.7: Source and sink analogy for hydride transfer in alcohol dehydrogenase

For substrates like ethanol that lose a hydride from a methylene carbon atom that has two hydrogens, only one of them is lost (either the proR or proS) from the prochiral center.

The site on NAD^+ that receives the hydride, as well as the entire ring, is planar with sp^2 hybridization. When bound to the enzyme, the hydride is transfer to the Re face of he ring. The same occurs in the reverse reaction when the hydride from NADH is transferred to the re face of acetaldehyde. Re and Si faces can be determined using by prioritizing the substituents attached to the sp^2 carbon using the Cahn-Ingold rules. The reversible transfer of the proR hydrogens to the Re faces of the reactants in the reversible conversion of ethanol to acetaldehyde are shown in Figure 12.4.8.

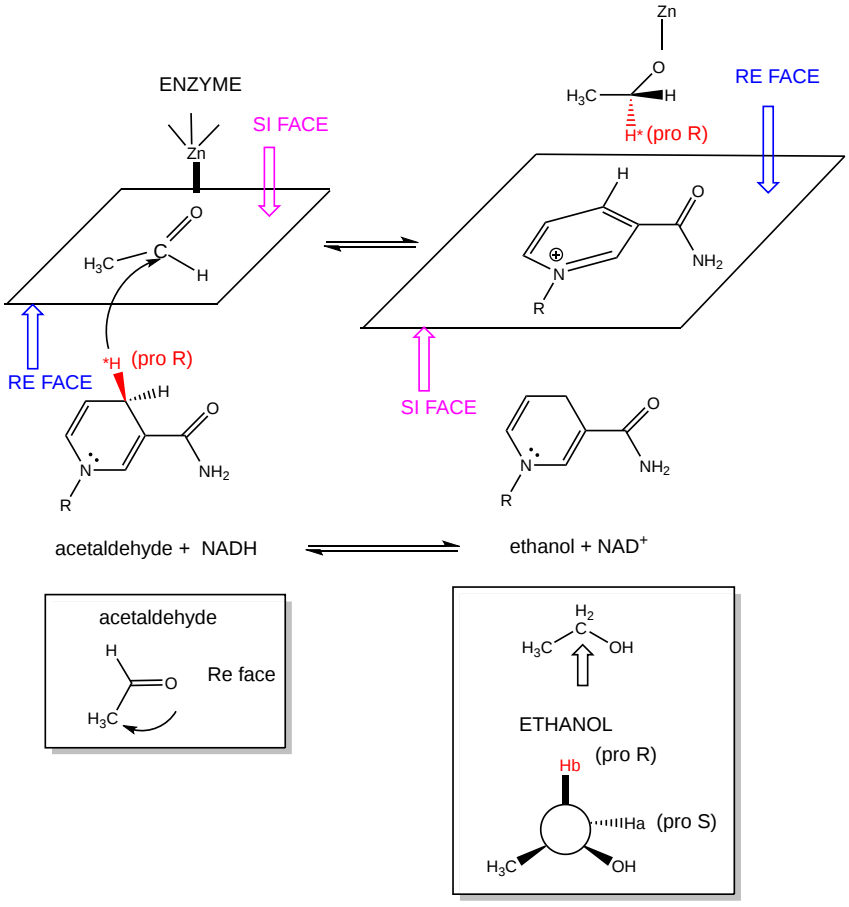


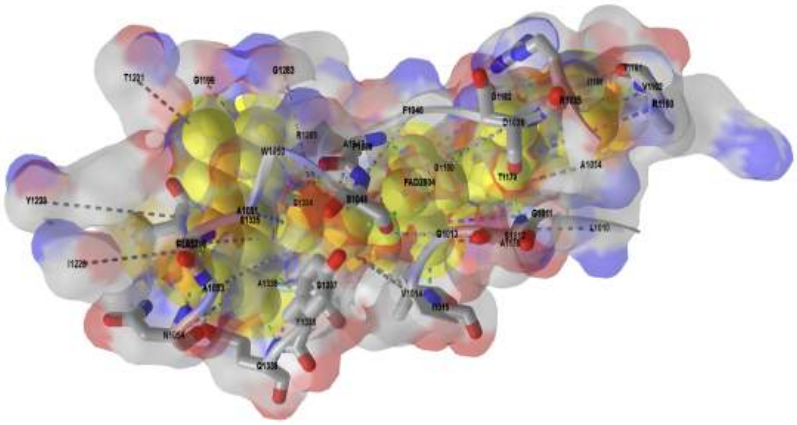
Figure 12.4.8: Reversible transfer of the proR hydrogens to the Re faces of the reactants in the alcohol dehydrogenase reaction

FAD has a more positive reduction potential than NAD^+ so it is used for more "demanding" oxidation reactions, such as dehydrogenation of a C-C bond to form an alkene. You will notice on standard reduction potential tables that the potential of FAD is often listed several times and depends on the enzyme. This is because the FAD is tightly bound to the enzyme so its tendency to acquire electrons depends on its

environment, in much the same fashion as the pKa of an amino acid side chain (which reflects is tendency to release protons) is affected by the environment of the amino acid side chain in the protein. The standard reduction potential for flavin enzymes varies from -465 mV to + 149 mV. Compare this to the reduction potential of free FAD/$FADH_2$, which in aqueous solution is -208 mV. The standard reduction potential of the flavin in D-amino acid oxidase, a flavoprotein, is about 0.0 V. Remember, the more positive the standard reduction potential, the more likely the reactant will be reduced and hence act as an oxidizing agent. Hence the FAD in D-amino acid oxidase is a better oxidizing agent than free FAD. The K_D for binding of FAD to the enzyme is 10^{-7} M compared to the K_D for binding of $FADH_2$, which is 10^{-14} M. By gaining electrons, the flavin binds more tightly, which preferentially stabilizes the bound $FADH_2$ compared to the bound FAD. This shifts the equilibrium of FAD ↔ $FADH_2$ to the right, making the bound FAD a stronger oxidizing agent.

A mechanism for the 2-electron hydride reduction of FAD is shown in Figure 6 above.

Figure 12.4.9 below shows an interactive iCn3D model of D-amino acid oxidase bound to FAD and a trifluoroalanine (1C0L) .



NCBI iCn3D **Figure 12.4.9 :** D-amino acid oxidase bound to FAD and a trifluoroalanine (1C0L). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...wYfRAK9TBSt2c8

FAD is shown in spacefill and colored yellow. Only the amino acids interacting with FAD are shown as a surface with underlying CPK colored sticks. The noncovalent interactions are shown as sticks. Note that you rotate the molecule, you will see the the FAD is almost completely buried which, along withe the extensive interactions with protein contributes to its low K_D.

Can the standard reduction potential of a redox active center in a protein be tuned by changing the environment of that center, much as the pKa of an acid side chain can by changing the polarity of the environment? The answer is yes. The active site of azurin, a cupredoxin, has a redox active copper ion coordinated by a Cys and two His residues in a trigonal planar fashion. Met 121 serves as a weak axial ligand. Marshall et al. have reported a feasible method to manipulate the redox potential (E^o) of this active site. The wild type azurin was mutated to alter the hydrophobicity and hydrogen bonding capabilities, while maintaining the overall architecture of the metal binding site. Ser 46 was selected for mutation since it occupied a position similar to Asn in another cupredoxin that was involved in an important H bond binding two ligand binding loops. An N47S-mutation, which strengthened the hydrogen bond between the two ligand-containing loops increased E^o by ~130 mV while preserving metal binding site architecture as determined by UV-Vis spectroscopy. They also compared a M121Q mutant with wild-type M121 and with a M121L mutant. A plot of E^o vs log partition coefficient for transfer of the side chain from water to octanol was essentially linear with a positive slope, showing that the standard reduction potential depended on the hydrophobicity of the weakly coordinating ligand in the metal binding region. This behavior extended to double mutants (where one set of mutants involved M121). The investigators were able to tune the E^o over a 700 mv range!

12.4.4: MONOOXYGENASES

An examples of monooxygenases are the hydroxylases which hydroxylate amino acids like tryptophane and tyrosine to form 5-hydroxytryptophan and 3-4-dihydroxyphenylalanine or dopa, respectively. These latter two substances can be decarboxylated using PLP-dependent enzymes to form the neurotransmitters 5 hydroxytryptamine (5HT or serotonin) and dopamine. The latter can be hydroxylated again to form norepinephrine, and subsequently methylated to form epinephrine. LSD and amphetamine are analogs of serotonin and dopamine, respectively. The derivative of tryptophan and tyrosine are shown in Figure 12.4.10.

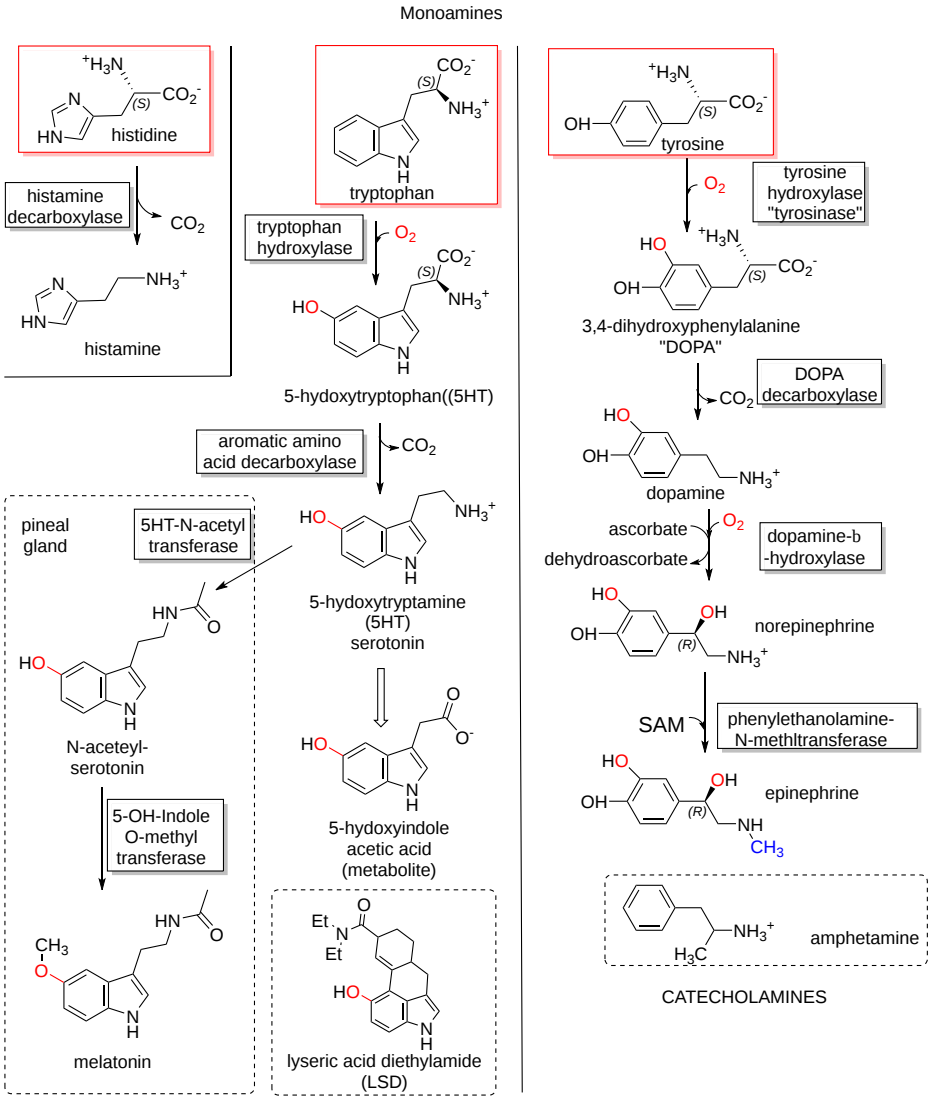


Figure 12.4.10: Derivative of tryptophan and tyrosine derived from monooxygenases

Since these monooxygenases use dioxygen, you might expect that the enzymes would use the motifs described in the previous section to facilitate its reaction with dioxygen. In fact, the enzyme contains a metal ion (Fe2+) bound to a heme in the protein. In addition, the reduction products of dioxygen that are eventually used to hydroxylate the substrate stay bound to the enzyme.

Tyrosine 3-monooxygenase (Tyrosine hydroxylase)

Tyrosine hydroxylase is a homotetramer which uses an Fe ion and biopterin cofactor for hydroxylation. In the central nervous system is used in the biosynthesis of dihydroxyphenylalanine (DOPA), which is the rate limiting step in the catecholamine synthesis.

A possible mechanism for the rat enzyme is shown in Figure 12.4.11.

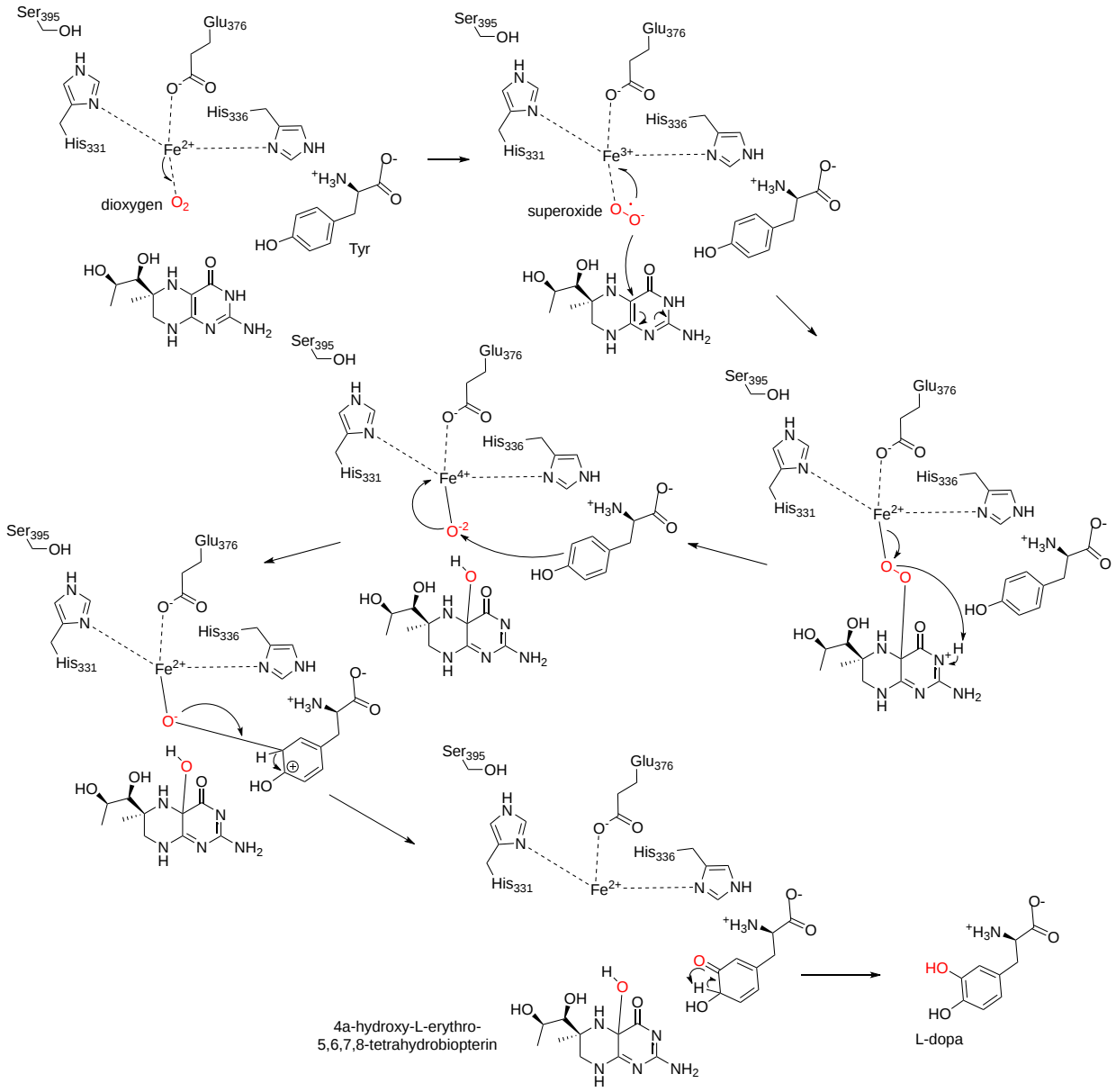


Figure 12.4.11: A possible mechanism for rat tyrosine hydroxylase, a monooxygenase. https://www.ebi.ac.uk/thornton-srv/m-csa/entry/134/

The Fe^{2+} ion binds O_2. Some mechanisms show a single electron transfer to the bound O_2 from the pterin ring to form Fe^{2+}-$O2^-$(bound superoxide) and a radical cation pterin. This is unstable and forms Fe^{2+}-μ-peroxypterin, followed by heterolytic cleaves of the peroxo O-O bond. Ultimately, hdroxpternin and an $Fe^{4+}O$ oxospecies form which then hydroxylates tyrosine.

Figure 12.4.12 below shows an interactive iCn3D model of rat tyrosine hydroxylase with bound cofactor analogue and iron (2TOH).

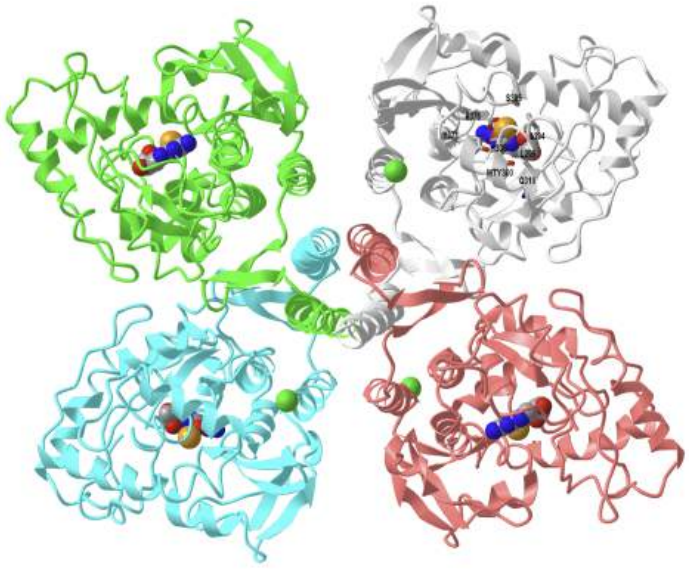


Figure 12.4.12 : Rat tyrosine hydroxylase with bound dihydrobiopterin analogue and iron (2TOH). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...KeKiLDfmeNFfr7

Four monomers in the homotetramer are shown. The cofactor analog (HBI), 7,8-dihydrobiopterin, is shown in spacefill. Phenylalanine 300 has been self-hydroxylated by the enzyme to produce meta-tyrosine (MTY300), which along with the other key amino acids in the active site are shown labeled in CPK-colored sticks in the gray monomer. The pterin ring interacts through pi stacking with Phe 300, which facilitates self-hydroxylation. The Fe ion is far (5.6 Å) from the carbon on pterin which gets phosphorylated. This suggests that O_2 might bridge the reacting pterin C- and the Fe ion in a Fe^{2+}-μ-peroxypterin complex.

Tryptophan hydroxylase

This enzyme also uses a tetrahyropterin cofactor and Fe ion in the formation of he hydroxylating intermediates As in the case of tyrosine hydroxlase, both the amino acid substrate and the cofactor must be present for oxygen to be activated. Other, it only oxidizes Fe^{2+} to Fe^{3+}. This serves as a protective mechanism so that O_2 is not activated unnecessarily, which avoids the formation of soluble ROS.

Figure 12.4.13 shows a possible mechanism without catalytic residues for tryptophan hydroxylase

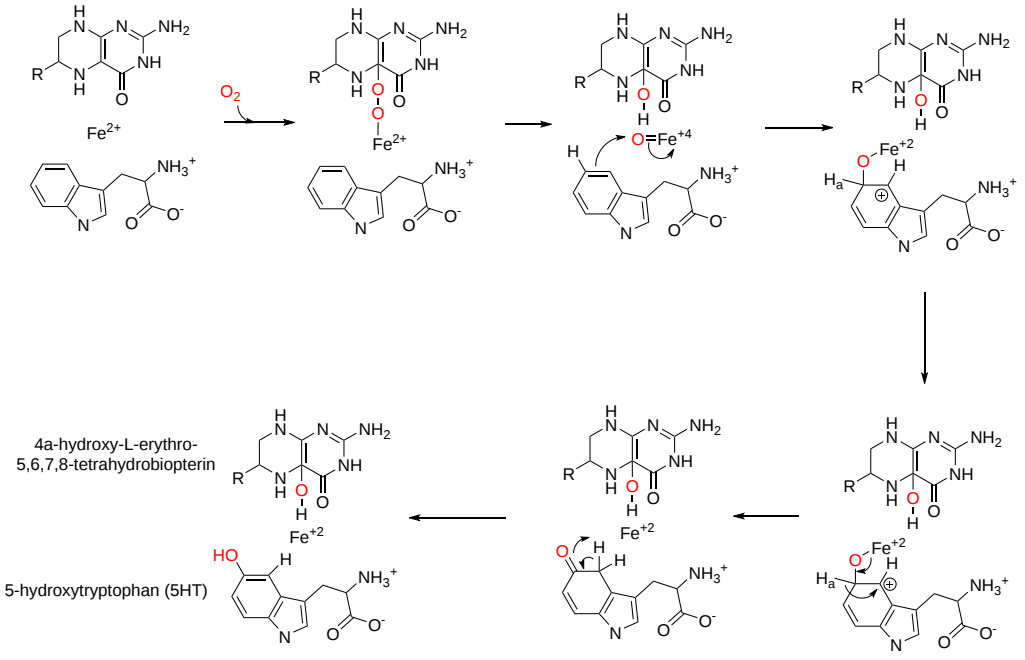


Figure 12.4.13: possible mechanism without catalytic residues for tryptophan hydroxylase adapted from Kenneth M. Roberts, Paul F. Fitzpatrick, https://doi.org/10.1002/iub.1144

Cytochrome P450s

The **cytochrome P450 (CYP)**consists of a large group of monooxygenases that contain a heme that absorbs maximally at 450 nm. They catalyze the following reaction:

$$RH + NAD(P)H + H^+ + O_2 \rightarrow ROH + NAD(P)^+ + H_2O$$

An example includes cytochrome P450cam that hydroxylates camphor, a large aromatic completely nonpolar molecule. The enzyme has been called the"biological equivalent of a blowtorch" as it can, at room temperature, stereospecific hydroxylate nonactivated hydrocarbons at physiological temperature. This reaction proceeds without stereospecificity only at high temperatures in the absence of a catalyst. Remember from the previous chapter section that dioxygen is a ground state triplet radical which can't react well with carbons atoms (which are singlets) unless converted to a singlet state, a process which requires significant energy. Cytochrome P450s get around this problem by binding to an Fe ion in the heme to form common intermediates such as oxos, oxides, and peroxides. The molecular orbitals of the bound dioxygen are "singlet-like".

Remember from introductory chemistry that transition metal ligands are names with a specific nomenclature. Some are shown in Figure 12.4.14.

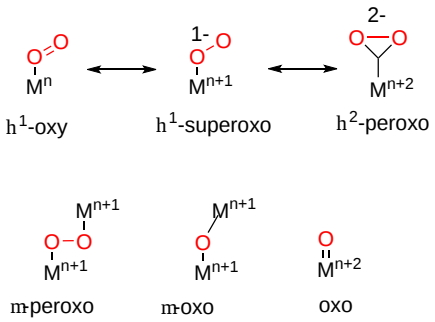


Figure 12.4.14: Common oxygen transition metal ligands

In naming η is the hapticity (the number of atoms of a ligand attached to a metal) and μ is the number of metal atoms bridged by a ligand.

For Fe ions and oxygen ligands, some examples include $Fe^{III}–O_2^-$ (superoxo) and Fe^{IV}=O (ferryl-oxo)

If the aromatic substrate get oxidized, something must get reduced. That something is of course O_2. Electrons for the reduction of O_2 come from the oxidized substrate but also by injection of electrons through NAD(P)H. Cytochrome P450 are microsomal protein and most require another protein, **NADPH-cytochrome P450 reductase (CPR)**. Before we look at the structure and mechanism of cytochrome P450, lets looks at this microsomal protein first. CPR catalyzes this reaction:

$$NADPH + \text{oxidized CytoP450 } (Fe^{3+}\text{-heme b}) + H^+ \leftrightarrow NADP^+ + \text{Oxidized CytoP450 } (Fe^{2+}\text{-heme b})$$

The protein contains multiple domains and resulted from a fusion of gene from flavodoxin, which binds FMN) and a FAD reductases. In addition is contains an NADP binding domain. It ultimately accepts a hydride ion (2 electrons) from NADPH and then transfers electrons to FMN (in one electron steps). These are used to reduce the heme which then activates dioxygen (oxidation number of 0) for hydroxylation reactions by cytochrome P450. In the hydroxylated organic product and in water, oxygen has an oxidation number of -2.

Figure 12.4.15 below shows an interactive iCn3D model of rat NADPH-cytochrome P450 reductase (1AMO).

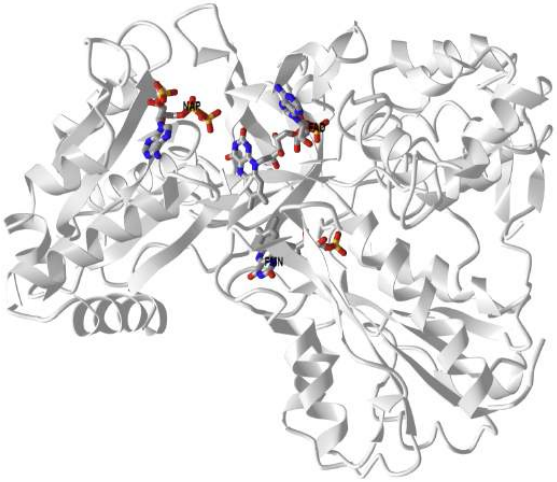


Figure 12.4.15 : Rat NADPH-cytochrome P450 reductase (1AMO). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...nMh9ByRbsWk4B8

$NADP^+$ is labeled as NAP. The enzyme is quite unique in that it has binding sites for both FAD and FMN (similar to nitric-oxide synthase). The flavins are aligned for electron transfer. The cleft in the enzyme presumably binds cytochrome P450. The protein has short (20 amino acid) lumenal sections, followed by 20 amino acid membrane-spanning helix (not shown in the above structure) followed a large cytosolic domain, where it can interact with cytochrome P450, which is found in multiple places including the ER, microsomal, and mitochondrial membranes as well as the cytoplasm..

The cytochromes P450s containing a heme, which instead of reversibly carrying dioxygen (as in myoglobin and hemoglobin), activates dioxygen for hydroxylation reactions involving aromatic, nonpolar substrates. Hydroxylation of these substrates increases their solubility which facilitates their elimination from the body. Epoxide intermediates or products are often produced, which can open up through nucleophilic attack using an alcohol (sugar derivative), as illustrated in Figure 12.4.16.

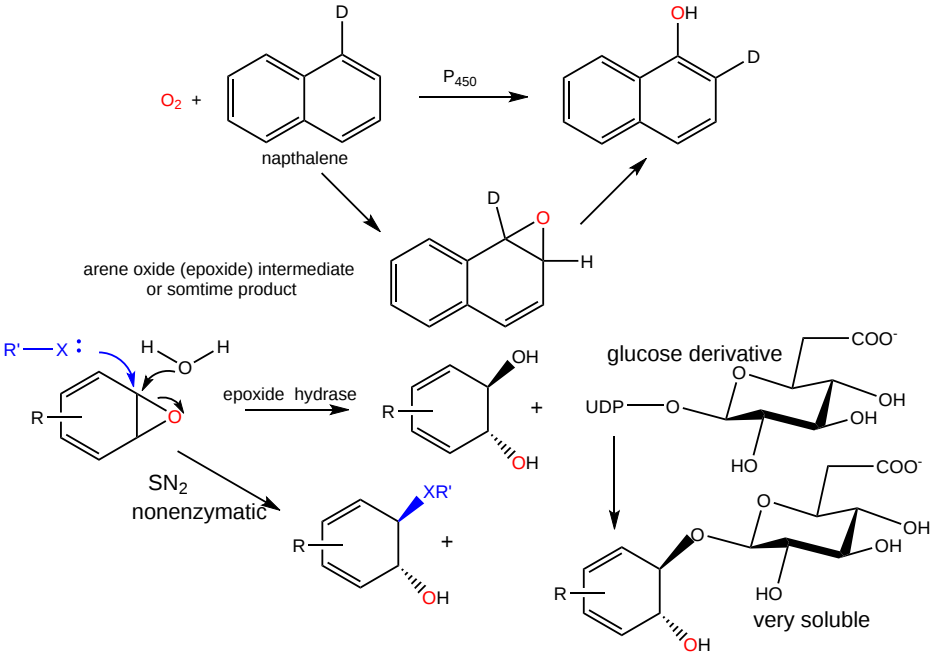


Figure 12.4.16: Cytochrome P450 hydroxylation reaction to increase target solubility

This type of reaction also converts steroids to different biologically active one.

The hydroxylation reaction and can also lead to reactions with amines, including those on nucleotide bases in DNA, resulting in the formation of large adducts, as illustrated in Figure 12.4.17.

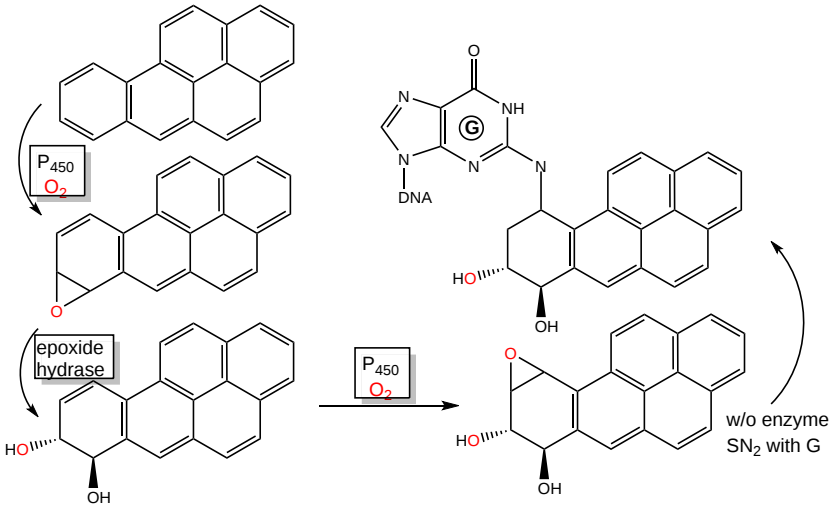


Figure 12.4.17: Cytochrome P450-mediated formation of carcinogens.

Hence cytochrome P450 can actually activate aromatic substrates to become carcinogens.

The cytochrome P450s family of genes/proteins are inducible on exposure to nonpolar aromatic molecules such as dioxin. These nonpolar molecules can enter the cytoplasm where they bind to the arylhydrocarbon receptor (AhR) which is bound to a heat shock protein, Hsp90. Upon binding of dioxin, TCDD, for example, the AhR.TCCD complex dissociates from Hsp90, and migrates to the nucleus where it binds a protein called Amt. The AhR-Amt complex serves as an enhancer/transcription factor, facilitating the transcription of the cytochrome P450 genes.

Figure 12.4.18 illustrates the activation of cytochrome P450 gene expression on exposure to nonpolar aromatic molecules such as dioxin.

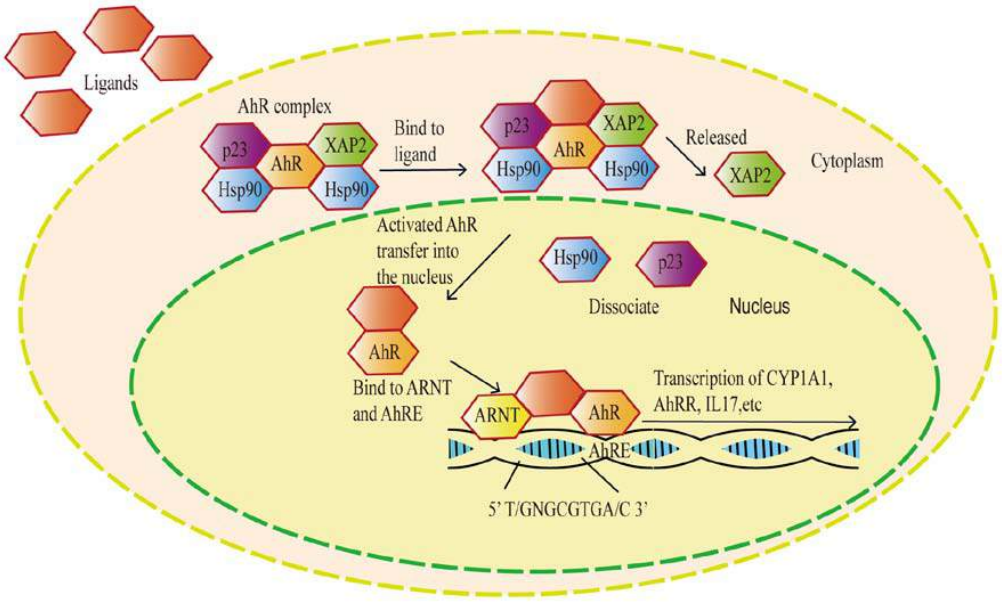


Figure 12.4.18: Activation of cytochrome P450 and other gene expression on exposure to nonpolar aromatic molecules such as dioxin

On binding to a ligand, the AhR is activated and enters the nucleus, where it binds to ARNT on the aryl hydrocarbon response element (AhRE) and promotes transcription of downstream genes including cytochrome P450 family 1 subfamily A member 1 (CYP1A1) and interleukin-1 (IL-1). CYP1A1 is in the CYP1A family that promotes activation of procarcinogens a well as hydroxylation of steroid hormones like estrogens. It also participates in the metabolism of steroidal hormones including estrogens. ARNT is the aryl hydrocarbon receptor nuclear translocator, AhRE the aryl hydrocarbon response element; CYP1A1, XAP2 is the aryl hydrocarbon receptor interacting protein, AHRR is the aryl-hydrocarbon receptor repressor, IL-11 is interleukin 17, Hsp90 is heat shock protein 90 and p23 is prostaglandin E synthase 3

Dioxin has been shown to affect estrogen-mediated activities. Estrogens, small hydrophobic hormones derived from cholesterol, enter cell and bind to cytoplasmic estrogen receptors, which then dimerize and bind to the estrogen response element (ERE), initiating transcription. Tamoxifen, a drug derived from the yew plant, blocks the biological effects of the estrogen receptor. Although it binds to the estrogen receptor, it doesn't elicit the same conformational changes in the protein, which prevents the bound receptor from binding to the estrogen response element and recruiting other proteins needed for estrogen-dependent gene transcription. It is used in chemotherapy and prevention of estrogen-dependent breast cancer cells.

How does dioxin interfere with estrogen signaling? Ahr and ARnt contain a basic helix-loop-helix motif which mediate their interaction with DNA. Upon binding of the complex, detoxification genes are activated. The dioxin-Ahr-Arnt complex can also bind to the estrogen receptor, which can then lead to activation of genes containing an estrogen response element (ERE) in the absence of estrogen. However, if estrogen is present, inhibition of gene expression from ERE is observed. Dioxins can be potent dysregulators of estrogen-induced gene expression. Such changes in estrogen activity could help to explain the pro- and inhibitory effects of dioxin on estrogen-mediated cellular responses and possible effects of dioxin on the immune system and on cancer development.

Given the importance of the cytochrome P450s, we'll offer two variant portrayals of their mechanism. Figure 12.4.19 shows the overall catalytic cycle of the enzyme with associated redox changes in the generic substrate, RH, and the Fe heme ion.

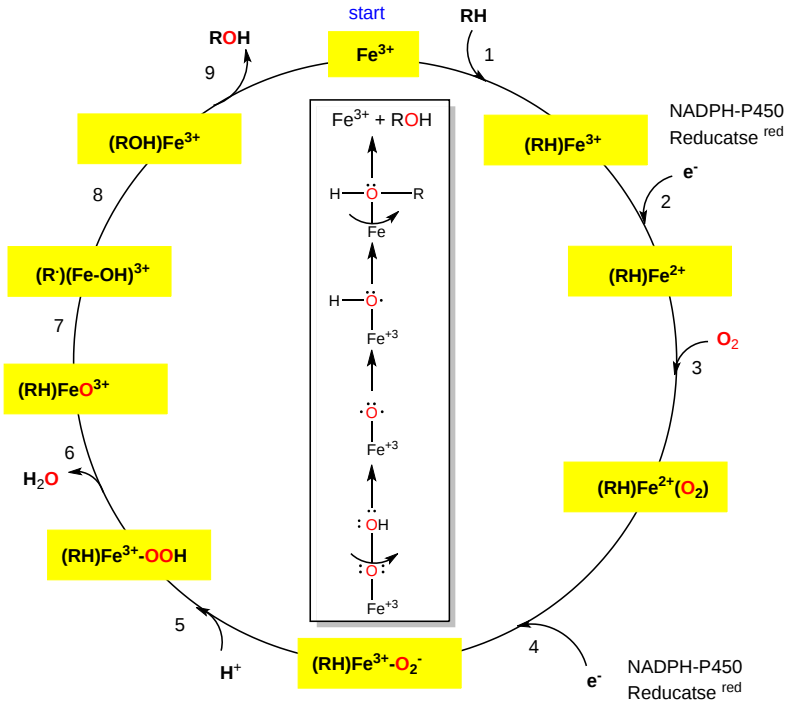


Figure 12.4.19: Catalytic cycle of cytochrome P450 with the generic substrate RH and the Fe heme ion

The key hydroxylating agent appears to be formal FeO^{3+} shown in step 7.

A possible mechanism for the hydroxylation of estrone by cytochrome P450 from Bacillus megaterium is shown below in Figure 12.4.20. It follows the general catalytic cycle shown above.

$$\text{Estrone} + \text{NADH} + H^+ + O_2 \rightarrow \text{2-hydroxyestrone} + NAD^+ + H_2O$$

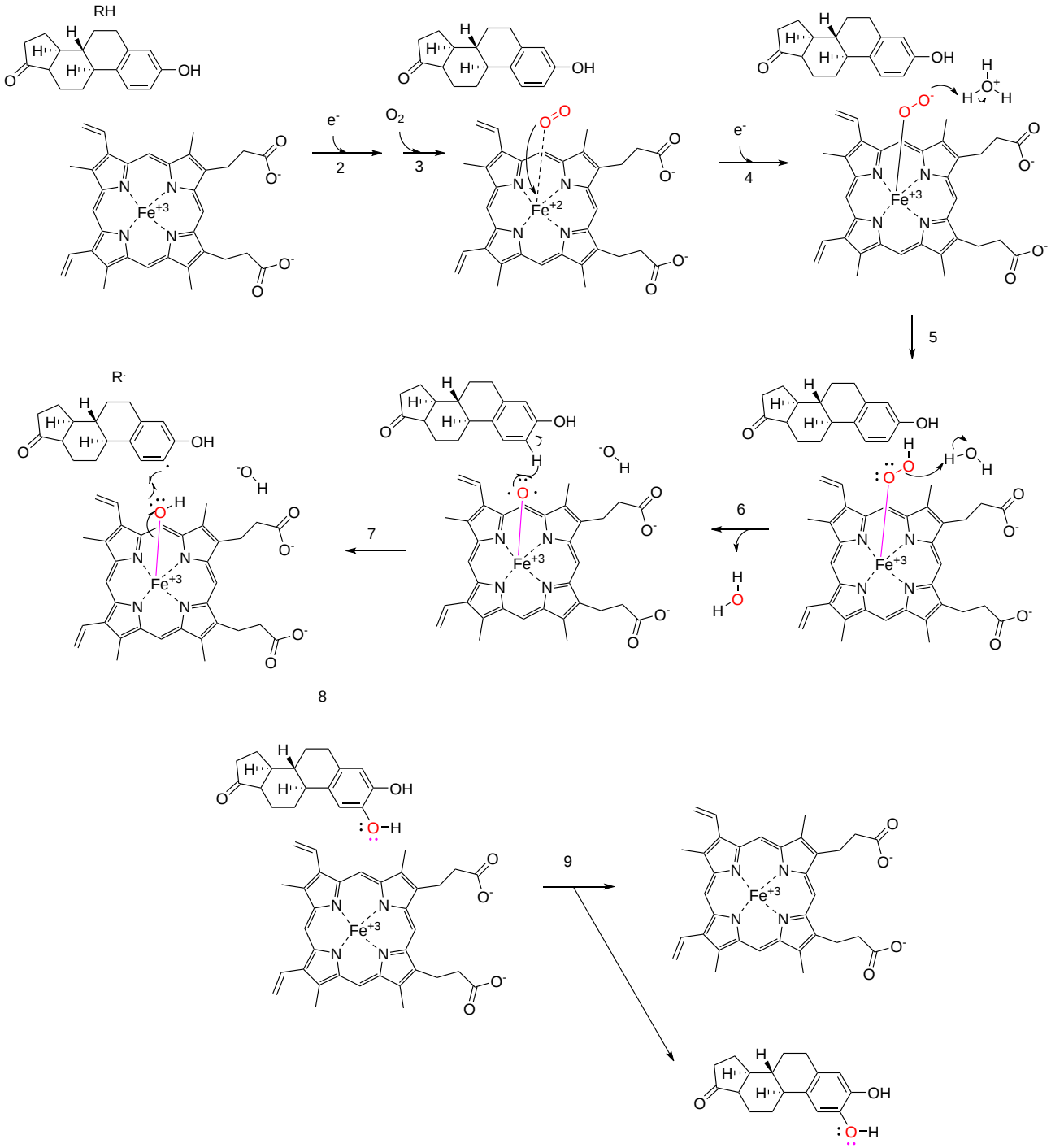


Figure 12.4.20: Possible mechanism for the hydroxylation of estrone by cytochrome P450 from Bacillus megaterium after https://www.ebi.ac.uk/thornton-srv/m...csa/entry/699/

As with the previous generic mechanism, the two added electrons derive from NADH which passed a hydride to FAD which passes single electrons on to FMN which them passes them onto the heme. The exact species for several parts of the mechanism are not completely clear. For instance, a Fe^{4+}-oxo complex has been proposed.

Let's shown the structure of a human cytochrome P450 1A1 that is active in drug metabolism and also in the activation of benzo[a]pyrene, a component of cigarette smoke, into a carcinogen. **Figure 12.4.21** below shows an interactive iCn3D model of Human Cytochrome P450 1A1 in complex with alpha-naphthoflavone (4I8V).

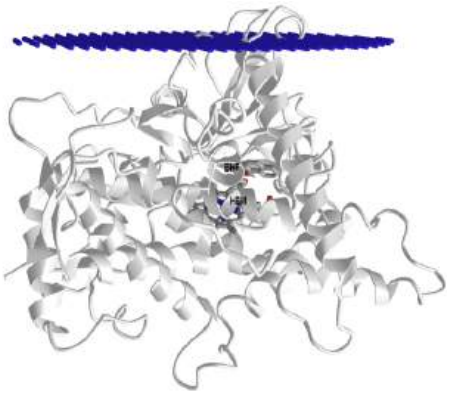


Figure 12.4.21 : Human Cytochrome P450 1A1 in complex with alpha-naphthoflavone (4I8V). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...JGtJeHG1qswN59

The shows in sticks and labeled the active site heme and an inhibitor, α-naphthoflavone (labeled BHF) bound and enclosed in the active site. It's 2-phenyl group points toward the heme Fe ion.

12.4.5: DIOXYGENASES

An example of a dioxygenase is the cyclooxygenase activity of prostaglandin synthase. This enzyme, often just called cyclooxygenase or COX, is an integral membrane protein found in the ER membrane, and is a homodimer (with two hemes). It catalyzes two different reactions. One is the addition of two dioxygens to arachidonic acid - 20:4$^{\Delta}$5, 8, 11, 15 (which is liberated from the C2 position of phospholipid membranes by phospholipase A2 upon appropriate signaling) to form prostaglandin PGG2. This molecule, with 5 chiral centers, arises from arachidonic acid, which has one. The cyclooxygenase activity is buried in the membrane, from where the arachidonic acid can readily access its active site. The active site is at the end of a hydrophobic channel (arachidonic acid binding site) and stretches from the membrane-binding region to a buried heme. PGG2 can be further metabolized to PGH2 by the addition of two electrons to PGG2 by the hydroperoxidase activity of the enzyme, located at the other end of the enzyme. This activity forms an alcohol from the peroxide functional group in PGG2. There is one heme per monomer, which acts in both the cyclooxygenase and peroxidase activities. Each monomer of the dimer has both enzymatic activities. A possible abbreviated reaction mechanism is shown in Figure 12.4.22.

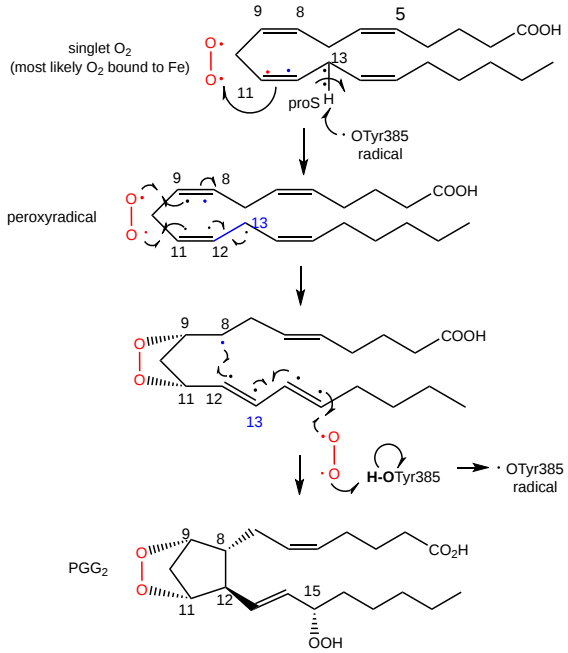


Figure 12.4.22: A possible abbreviated reaction mechanism for dioxygenase cyclooxygenase.

In summary:

- The carboxylate of arachidonic acid is coordinated to Arg 120 and Tyr 355.
- The C13 pro(S) H atom of arachidonic acid is close to Tyr 385 which allows its abstraction.
- This results in a radical centered on C11 that reacts with dioxygen to form a peroxyl radical

- Attack by dioxygen at C11 occurs from the side of the substrate opposite to that of hydrogen abstraction
- The oxygen radical at C11 cyclizes by attacking C9.

The C13 proS hydrogen atom (not proton) is remove from bound arachidonic acid by a free radical form of Tyr 385, which acts as an oxidizing agent. A site-specific mutants in which Tyr 385 is replaced by Phe is inactive. How is the Tyr free radical formed? Based on single electron standard reduction potential (0.9 V for Tyr and -0.2 to + 0.2 V for Fe^{3+} in the bound heme), it appears that the heme iron is not a potent enough oxidizing agent to accomplish this task. However, oxygen bound to the heme iron could be converted to a peroxide and form an Fe^{4+}-oxo complex (which has also been proposed for cytochrome P450). The Fe^{4+} ion is a more potent oxidizing agent (standard reduction potential of approximately 1 V, sufficient for oxidation of Tyr 385. Another possibility is that the peroxide activator (in the formation of the ferryl-oxo ligand) is NO (nitric oxide, a free radical). NO is formed by immune cells (like macrophages) on immune activation. The NO might react with superoxide (also a radical, possibly formed during an oxidative burst in macrophages during immune stimulation) to form peroxynitrite (NO^{3-}). This can donate an oxo group to the Fe^{3+} to form the Fe^{4+}-oxo complex, which could then oxidize Tyr to the free radical form. There might be other mechanisms as well to generate the Tyr free radical, since just adding organic peroxides to the enzyme will generate it.. After abstraction of the proS H atoms, a carbon-centered free radical at C11 results which reacts with oxygen as shown below. The exact form of oxygen that reacts is unclear, but presumably is either a peroxy or activated singlet form.

12.4.6: OXIDASES

This class of enzymes does not incorporate dioxygen into an organic substrate. Rather it accepts electrons released from an organic substrate, through intermediate electron carriers (such as ubiquinone and cytochrome C) to form superoxide (as in NADPH-oxidase), hydrogen peroxide (as in xanthine oxidase) or water (as in cytochrome C oxidase). The mechanism of cytochrome C oxidase again supports our expectations about enzymes that use dioxygen. Dioxygen binds metals in the enzyme. One oxygen atom binds a heme Fe^{2+} of cytochrome a3 which is bound to the enzyme, while the other binds a Cu^{1+} of Cu B. All oxygen reduction intermediates remain bound to the enzyme. Four electrons are added from four different cytochrome C molecules, which serve as mobile carriers of electrons.

We will explore cytochrome C oxidase in great detail in Chapter 19.1: Electron-Transfer Reactions in Mitochondria. Figure 12.4.23 shows cartoon versions of several oxidases.

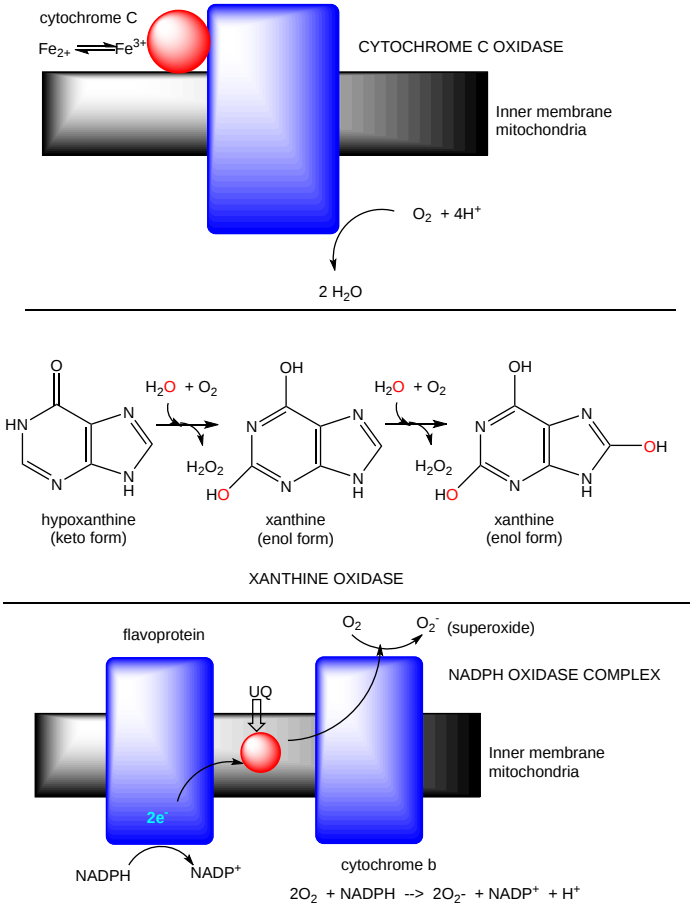


Figure 12.4.23: Examples of oxidases

Another example of an oxidase is monoamine oxidase. Mitochondrial monoamine oxidase catalyzes the oxidative deamination of certain neurotransmitters after they have been taken up by post-synaptic neurons, in a process of inactivation. The reaction is shown in Figure 12.4.24.

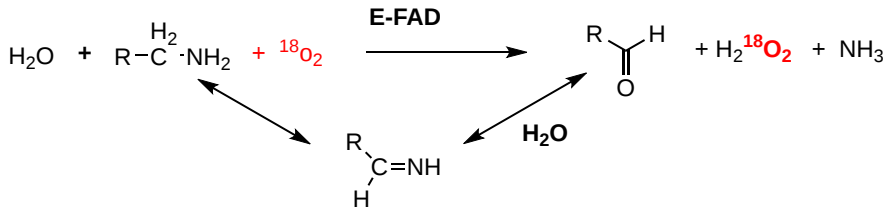


Figure 12.4.24: Monoamine oxidase reaction

A Schiff base is formed which is then hydrolyzed, incorporating unlabeled oxygen into the oxidized molecule.

12.4.7: BIOLOGICAL OXIDATIONS: METHANE TO CO_2

In the previous chapter section, we discussed the progressive oxidation of methane by 2 electron loses to form methanol, formaldehyde, formic acid and CO_2, with a progressive increase in oxidation number for the C by +2 (from -4 in methane to +4 in CO_2), as reviewed in Figure 12.4.25.

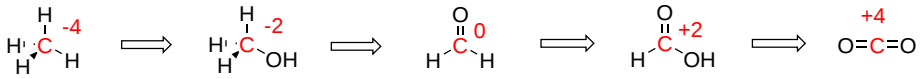


Figure 12.4.25:: Progressive stages in the oxidation of methane

Methanotrophs are aerobic bacteria that use methane as a source of energy, converting it in a series of two electron oxidations as shown above, to carbon dioxide. The enzymes involved in this sequential process are methane monooxygenase, methanol dehydrogenase, formaldehyde dehydrogenase, and formate dehydrogenase. Methane monooxygenase exists in a soluble and membrane form, both of which are part of a larger complex. Both have a hydroxylase (which uses dioxygen to add O to methane) and the membrane form has recently been shown to be associated with have methanol dehydrogenase in a larger complex consisting of trimers of each enzyme (the hydroxylase and the dehydrogenase).

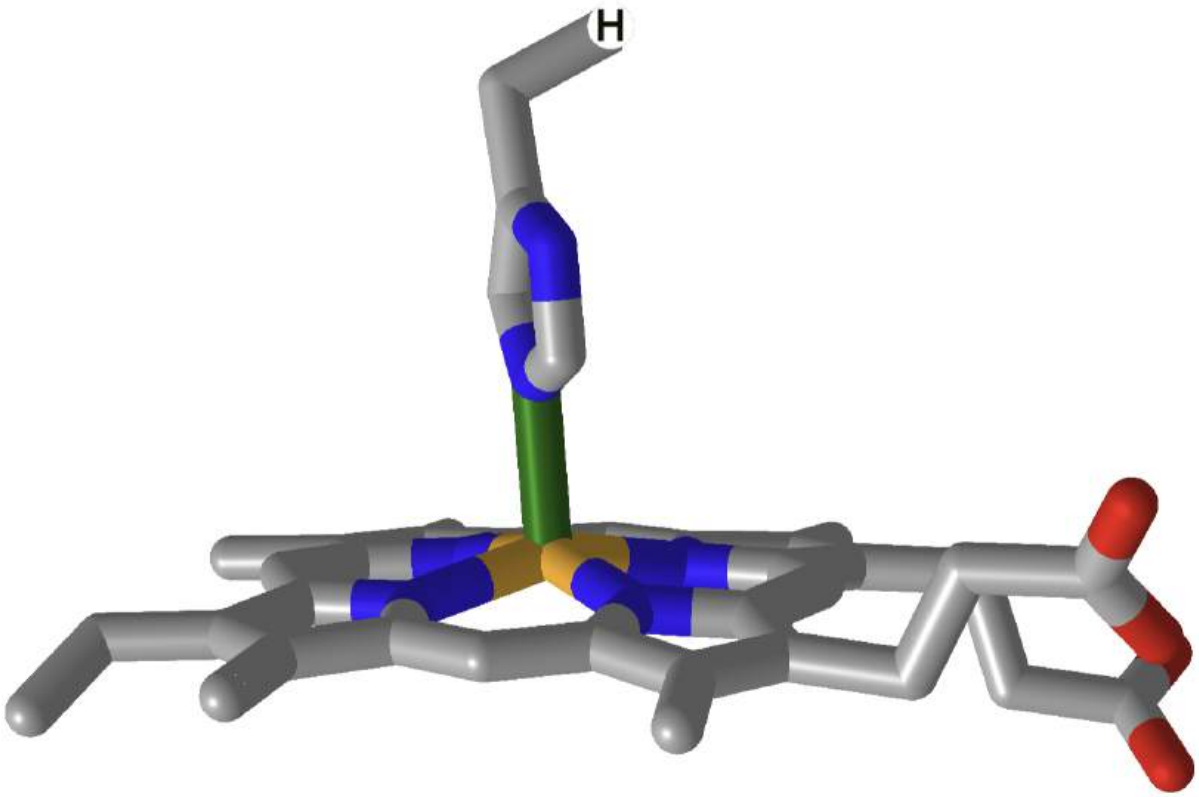
12.4.8: HEME PROTEINS

So far in this course, we have examined three different kinds of heme proteins..

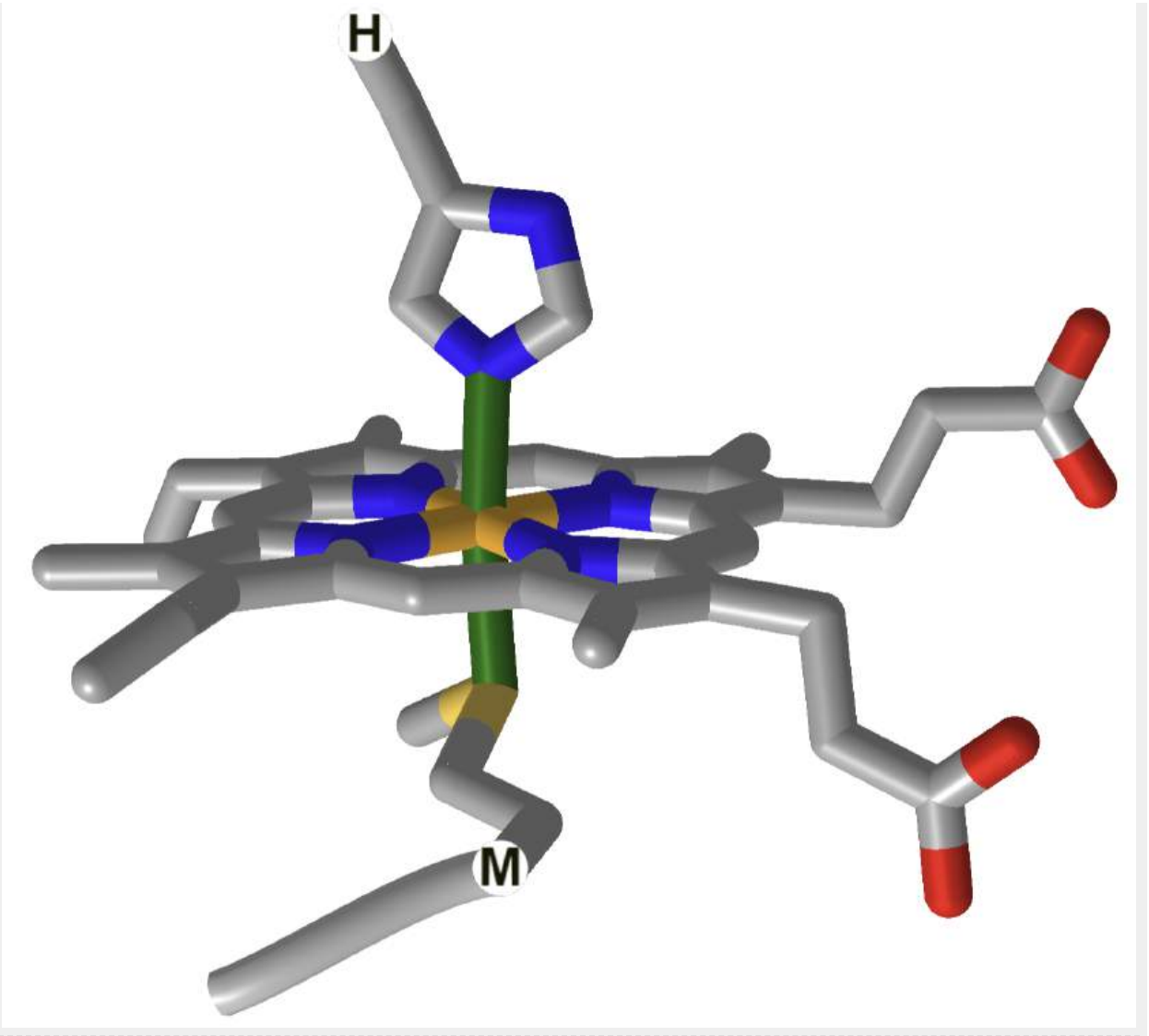
- The first, hemoglobin (and myoglobin) serve as carriers of **dioxygen**. Even though they bind one of the best oxidizing agents around (dioxygen), the heme Fe^{2+} does not get oxidized to Fe^{3+}. If it does, as in the case of met-Hb, the protein looses it ability to carry oxygen.
- Cytochrome C, on the other hand, does not bind dioxygen but rather serves as a carrier of **electrons** which get passed to dioxygen in Cytochrome C oxidase. Its Fe ion readily cycles between the 2+ and 3+ states as it serves as an electron carrier.
- Finally, the Fe^{2+} in the heme of the cytochrome P450s (so named since they have an absorbance maximum at 450 nm when they bind CO) does **both**. It binds dioxygen and cycles between the 2+ and 3+ states as it activates dioxygen for hydroxylation reactions.

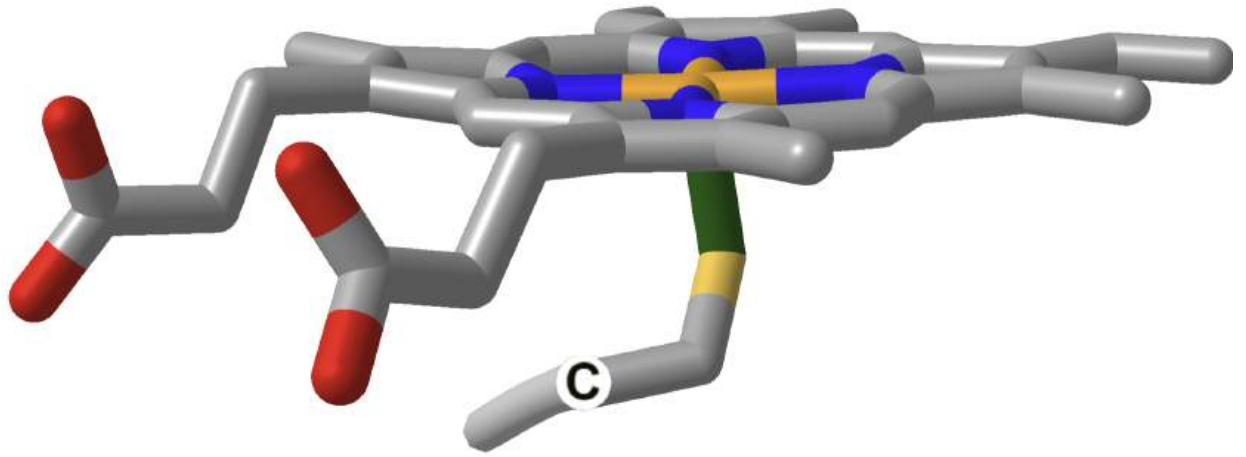
The structure of the heme and amino acid ligands, along with their absorbance spectra, are shown in Table 12.4.3 below.

hemoglobin



cytochrome C





Cytochrome P450

P450: Fujishiro et al. J Biol Chem. 2011 Aug 26; 286(34): 29941–29950. Published online 2011 Jun 30. doi: 10.1074/jbc.M111.245225. CC BY license.

Cyto C: Hulko et al. December 2011. Sensors 11(6):5968-80. DOI:10.3390/s110605968. Creative Commons Attribution 3.0 Unported

Hemoglobin: Nitzan et al. July 2014, Medical Devices: Evidence and Research 7(1):231-9. DOI:10.2147/MDER.S47319. Creative Commons Attribution-NonCommercial 3.0 Unported

How could heme serve such diverse functions? We can explain this by referring to one of the main themes of the course - structure mediates function. The environment of each heme must be different. Clearly the protein ligands coordinating the Fe ions are different. The 5th ligand is the proximal His in hemoglobin while dioxygen binds to the 6th site. In cytochrome C, the 5th and 6th ligands are His and Met, respectively. In cytochrome P450, the 5th site is occupied by Cys, and the 6th by dioxygen. Presumably the environments surrounding the hemes are different as well. Once again, we have seen analogous example in which chemical properties are influenced by the microenvironment. The pKa of a given amino acid side chain can vary considerably depending on the polarity of the local environment. Likewise, the standard reduction potential of tightly bound FAD/FADH2 depends on the microenvironment.

As we have seen (from the study of heme proteins and the oxidative enzymes of cells), transition metals such as Fe, Zn, and Cu have vital biological roles as binding sites and cofactors in many reactions. Yet they also pose problems since they can lead to oxidative damage in cells. As we saw with cytoplasmic metallothioneins, which bind to heavy metals and protect the cell from such damage, many proteins are involved in binding and regulation of transition metals in the cell. Integral membrane proteins are required to bind and transport these cations into the cytoplasm. Other proteins act as sensors of transition ion concentration (such as latent transcription factors which bind heavy metals and become active transcription factors for metallothioneins. Others act as chaperone proteins which bind metal ions and transfers them to apometalloproteins. Recent work has suggested transporters and chaperones involved in metal ion biology bind these ions with unusual coordination geometry, which presumably facilitates transfer of the ion to the apo-target protein.

The transition metals Zn and Fe are often found in E. Coli at a concentration of 0.1 mM, compared to Cu and Mn which are present at concentrations from 10 to 100 μM. Also, about one third of all proteins demonstrate specific binding of metal ions and can be classified as metalloproteins. Mass balance suggests that metal ions would be distributed in proteins with low, intermediate, and high metal binding affinity as well as in free pools, which which potentially be toxic to cells. Metalloproteins, depending on their Kd for metal ion binding, would hence be in various state of ligation. The free concentration of some ions (Cu and Zn) is so low that newly synthesized apoproteins which bind these ions would not obtain the ion from the free pool. In such cases, metal chaperones would be required.

CHAPTER OVERVIEW

13: GLYCOLYSIS, GLUCONEOGENESIS, AND THE PENTOSE PHOSPHATE PATHWAY

13.1: Glycolysis

13.2: Fates of Pyruvate under Anaerobic Conditions- Fermentation

13.3: Gluconeogenesis

13.4: Pentose Phosphate Pathway of Glucose Oxidation

Thumbnail: Due to its role in the pentose phosphate pathway and the Calvin cycle, D-Ribose-5-phosphate isomerase is highly conserved in most organisms, such as bacteria, plants, and animals. It plays an essential role in the metabolism of plants and animals, as it is involved in the Calvin cycle which takes place in plants, and the pentose phosphate pathway which takes place in plants as well as animals. (CC BY-SA 4.0; Ishikawa, K., Matsui, I., Payan, F., Cambillau, C., Ishida, H. et al. via Wikipedia)

13.1: GLYCOLYSIS

13.1.1: INTRODUCTION

In this chapter, we will provide you with a historical overview of glycolysis and introduce you to the 10 enzymatic reactions in the pathway. Our main goal is to understand how the oxidation of our major food molecules, sugars in the case of glycolysis, can lead to ATP synthesis.

Before we begin our journey into the glycolytic pathway, it is useful to review the concept of free energy within reactions. For a reaction to be spontaneous, the change in the free energy within the system must be negative. From the equation in **Figure 13.1.1** you can see that the change in free energy of a reaction is dependent on the concentration of the reactants and the products, as well as the temperature within the system.

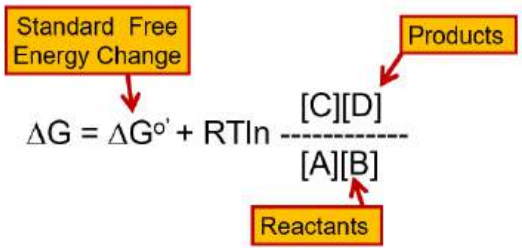


Figure 13.1.1 : The Change in Free Energy of a Reaction. The change in the free energy of a reaction is equal to the standard free energy change for the reaction plus the gas constant (R) multiplied by the Temperature in Kelvin (T) and the natural log (ln) of the concentrations of the products of the reaction (noted as C and D in the example) over the concentration of the reactants (noted as A and B in the example).

You will note that some reactions of glycolysis and other metabolic pathways that we will investigate are not favored. Thus, it is necessary to drive the reaction to become spontaneous by either coupling the reaction with a spontaneous reaction that can generate enough free energy to drive the nonspontaneous reaction forward, or by using Le Chatelier's principle and removing the products from the enzyme's area as soon as they are made. This will help drive the reaction in the forward direction as it will reestablish equilibrium. In this way, a small amount of product can be formed spontaneously. Essentially, the continued removal of the product or the addition of excess reactant will drive the reaction forward. It is of note that the temperature of the reaction can also influence the change in free energy. However, in biological systems, temperature changes that will significantly affect the change in free energy usually are not compatible with maintaining most life forms. Thus, it will not be a large consideration in the context of our metabolic discussions, here.

When reactions are coupled together to obtain a spontaneous reaction, the overall free energy change for a chemically coupled series of reactions is equal to the sum of the free energy changes of the individual steps, as noted in **Figure 13.1.2 .**

Figure 13.1.2 : The Coupling of Two Reactions. The top reaction is nonspontaneous with a positive change in free energy of 21 kJ/mol. This reaction can be driven in the forward direction, by coupling it with a strongly spontaneous reaction that releases -34 kJ/mol. The overall net reaction has a negative change in free energy and is spontaneous

The metabolic reactions of carbohydrates and other food molecules play an important role in generating energy within living systems in the form of ATP. For carbohydrates, this begins with the metabolic process known as glycolysis (or the breakdown, "lysis", of sugars, "glyco") At the end of the 1910s Otto Meyerhof mapped some of these metabolic conversions by measuring heat trends and oxygen consumption in frog muscles. When the muscle is working, lactic acid is formed from carbohydrates, and Otto Meyerhof showed that during recovery, this is followed partly by the burning of lactic acid and partly by reprocessing of lactic acid to carbohydrates.

Concurrently, Archibald Hill was also outlining these processes in the muscles of frogs. In opposition to the prevailing view that mechanical movement and chemical processes were parallel sequences, Hill was able to show through measurements of heat generated by the mechanical processes that these were delayed compared to the movements. The chemical sequence consists of a work phase, which is not dependent on oxygen supply, and a recovery phase when oxygen is required. Together, their work has opened the door to understanding aerobic and anaerobic metabolism beginning with the process of glycolysis. They both shared the Nobel Prize in Physiology and Medicine in 1922 for their work in these processes **Figure 13.1.3** .

Figure 13.1.3: The Nobel Prize for Physiology and Medicine in 1922. Archibald V. Hill and Otto Fritz Meyerhof received this award for their work on sugar metabolism and breakdown. Photos from Nobel Prize.Org

The glycolytic pathway consists of 10 enzymatic steps that convert glucose to pyruvate. This conversion generates a small amount of energy. The pyruvate can then be converted to lactic acid (lactate) in vertebrates or to ethanol in yeast in an anaerobic (or oxygen-independent) pathway or it can be fully oxidized to carbon dioxide in an aerobic (or oxygen-requiring) pathway that takes place within the mitochondria (which is shown in green in **Figure 13.1.4:**). Aerobic oxidation yields about 18 times as much energy as the anaerobic pathways causing them to be favored over anaerobic pathways. Most animal tissues can only survive short anaerobic bursts that occur in isolation and don't involve the entire organism. The aerobic pathway is required to sustain life. Yeast also prefers to grow using the aerobic, mitochondrial pathway. However, if oxygen is unavailable, yeast and other fungi can switch to anaerobic growth and produce ethanol as a byproduct. The production of alcoholic beverages through this fermentation process is quite popular.

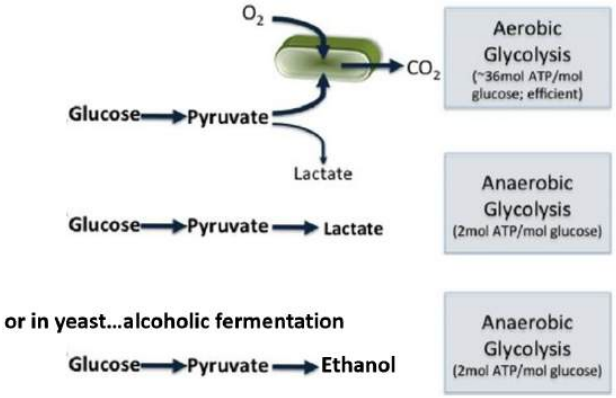


Figure 13.1.4: Overview of Aerobic and Anaerobic Oxidation of Glucose. Figure modified from Kim, Y. et al (2011) PLoS ONE 6(12):e28293

Figure 13.1.5 provides a summary of the glycolytic pathway coupled to the oxidative phosphorylation pathway that occurs within the mitochondria.

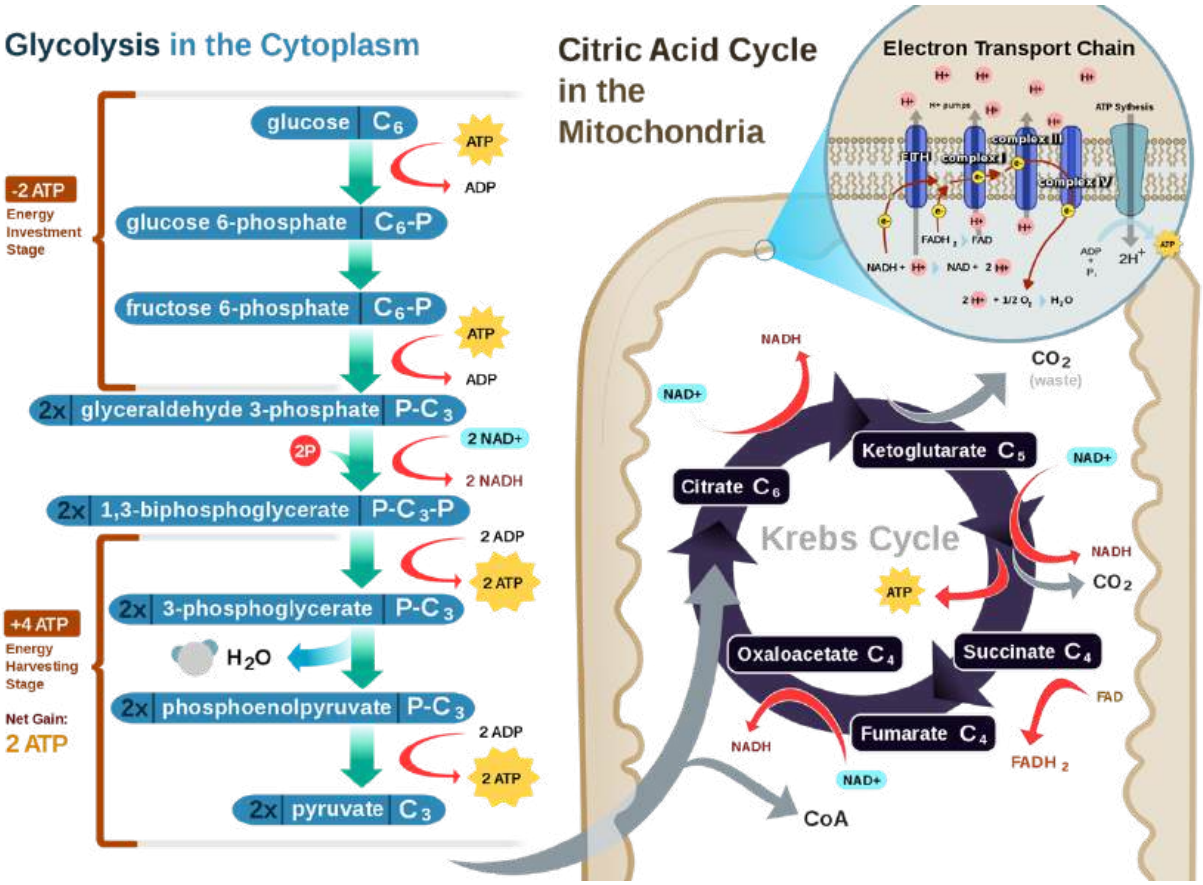


Figure 13.1.5 : A Summary of the Glycolytic and Oxidative Phosphorylation Pathways. The glycolytic pathway is shown on the left-hand side in blue. During aerobic metabolism, pyruvate would be converted to acetyl-CoA which then enters the Kreb cycle within the matrix of the mitochondria. Future chapters will focus on the reactions inside the mitochondria. Our focus in this chapter will be on the cytosolic reactions of the glycolytic pathway.

The major reactions of glycolysis are shown in **Figure 13.1.6** . The pathway can be broken down into two major sections: (1) the energy-consuming reactions and the (2) the energy-generating reactions. An adage states that 'It takes money to make money'. The same can be thought about the glycolytic pathway. The first section requires an investment of energy to generate energy in the second half of the reaction pathway. In this pathway, glucose, a 6-carbon hexose, is converted to two, 3C molecules - pyruvate. Note that **Figure 13.1.6** shows the entire pathway using Lewis wedge/dash representations plus the anaerobic conversion of pyruvate to lactate.

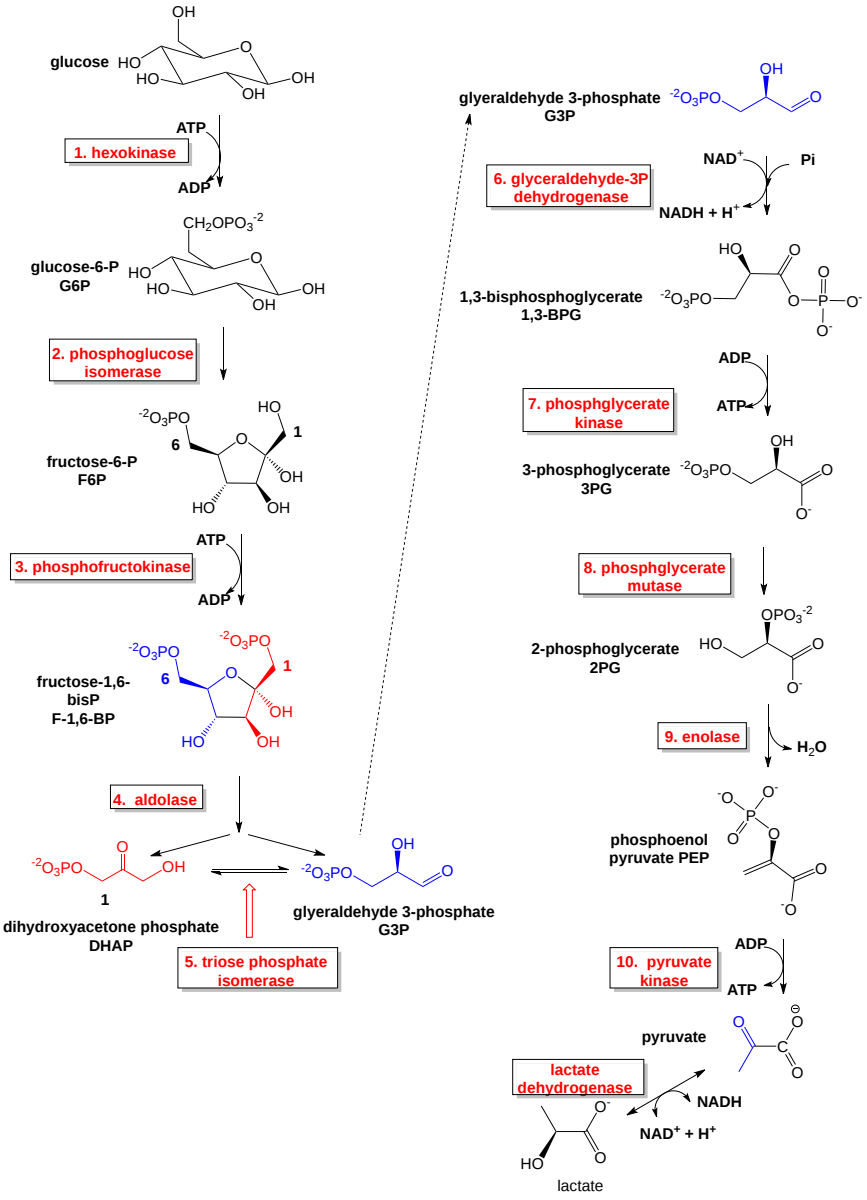


Figure 13.1.6 : Glycolytic pathway. The chemical steps of the glycolytic pathway are shown using Lewis wedge/dash representations. Enzymes required at each step are labeled in red. The energy-consuming steps encompass reactions 1 - 3, whereas the energy-producing steps occur in the second half of the pathway from reactions 4 - 10. The conversion of pyruvate to lactate by lactate dehydrogenase represents anaerobic respiration as it occurs within mammalian species.

Glycolysis is the key anaerobic pathway for energy products in all organisms, except in lithotropes that use the oxidation of inorganic molecules for energy production. In aerobic systems, glycolysis provides the release of fast energy within the body as glycogen metabolism can quickly release free glucose for utilization. Given the centrality of glycolysis to all of life, we will explore each of the reactions in detail below.

13.1.2: **REACTION 1: GLUCOSE → GLUCOSE-6-PHOSPHATE**. ΔG^{O}=-4.0 KCAL/MOL (-16.7 KJ/MOL).

The first step in glycolysis is catalyzed by enzymes known as hexokinases. The hexokinase family enzymes typically have broad specificity for various hexoses and catalyze the phosphorylation of carbon 6. Hexokinase phosphorylates glucose using ATP as the source of the phosphate, producing glucose-6-phosphate, a more reactive form of glucose. Notably, this reaction prevents the phosphorylated glucose molecule from continuing to interact with the GLUT transport proteins that can shuttle glucose into and out of the cell. Thus, once glucose is phosphorylated it can no longer leave the cell. Do note that both of these sugar forms (the free sugar and the phosphorylated version) can shift back and forth between the ring-closed and ring-opened conformation. Hexokinases require glucose to be in the closed conformation

for the phosphorylation reaction. **Figure 13.1.7** shows the reaction which is catalyzed by **hexokinase** (a kinase that transfers the γ-phosphate from ATP to a hexose in the closed-ring form).

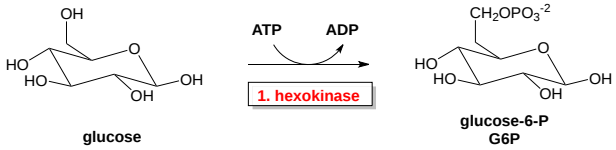


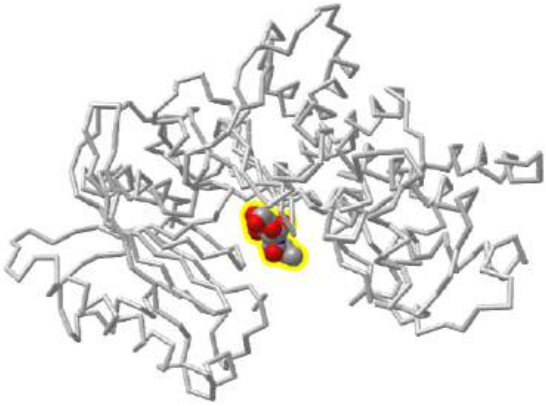
Figure 13.1.7 : Summary reaction - hexokinase

This reaction is a nucleophilic substitution reaction on the gamma phosphate of ATP. A phosphoanhydride bond is broken in ATP as a phosphoester bond is made producing glucose 6-phosphate. Hence the reaction proceeds with a negative ΔG^o.

Vertebrates have 4 different types of hexokinases, I-IV (also called A-D). The regulation of these enzymes is presented in more detail in Chapter 15.5. Hexokinases I-III bind glucose more tightly as reflected by low K_M values. Type IV or glucokinase binds it less tightly, and is found in high concentrations in vertebrate livers. Yeast has 3 isozymes, P1, PII or hexokinase B, and glucokinase. Most prokaryotic hexokinases are in the glucokinase (Type IV) class.

In addition to hexokinases, there are also many other different types of sugar kinase enzymes. Notably, all sugar kinases must prevent the spurious hydrolysis of ATP by water, which can also be considered a phosphotransfer to water, instead of the preferred transfer of the γ phosphate of ATP to a sugar ROH (alcohol), an "alcoholysis" reaction. Hexokinase does this through an "induced fit" mechanism, in which glucose binding triggers a large conformation change to close off the active site to water.

This conformational change is illustrated in **Figure 13.1.8** below which shows an interactive iCn3D model of the yeast hexokinase PI in the absence (2YHX) and presence (3B8A) of glucose.



NCBI iCn3D **Figure 13.1.8 : Yeast hexokinase PI in the absence (2YHX) and presence (3B8A) of glucose.** (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...84YGEd1ob3th7A

Within **Figure 13.1.8 ,** the gray structure is hexokinase without glucose (2YHX). However, it has a glucose analog bound, o-tolyoylglucosamine (spacefill, yellow highlights), which binds at the glucose binding site but does not cause a subsequent conformation change. The cyan structure (in the actual iCn3D model) show the enzyme with glucose shown as colored sticks. Note the large conformational change on the actual binding of glucose, a classical example of an "induced fit" mechanism.

Figure 13.1.9 shows the surface of both enzymes using the same color coding as in the above figure These images better show how the active site of hexokinase is occluded in the glucose-bound form. This prevents water access and hydrolysis of bound ATP instead of "alcoholysis" of ATP (transfer of phosphate to glucose). The gray structure on left has the glucose analog bound which doesn't alter the global conformation of the protein.

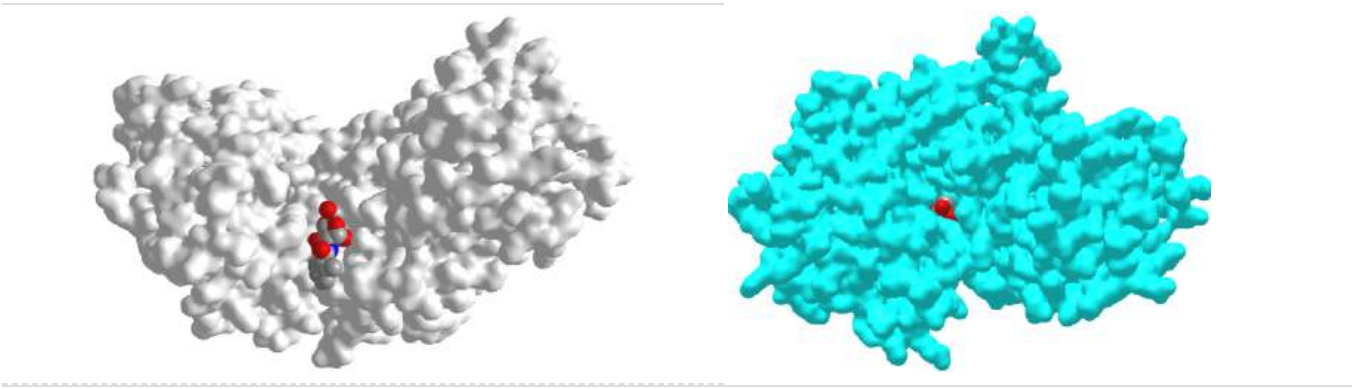


Figure 13.1.9 : Surface of human hexokinase I with bound o-tolyoylglucosamine (spacefill, left, 2YHX) and with bound glucose (spacefill, right, 3B8A)

Hexokinase I is the key form in the brain. It forms a complex with porin in the mitochondrial outer membrane and ATP/ADP translocase or carrier protein in the mitochondrial inner membrane which facilitates the hexokinase reaction. The mechanism for the reaction of human hexokinase I is shown in **Figure 13.1.10** below.

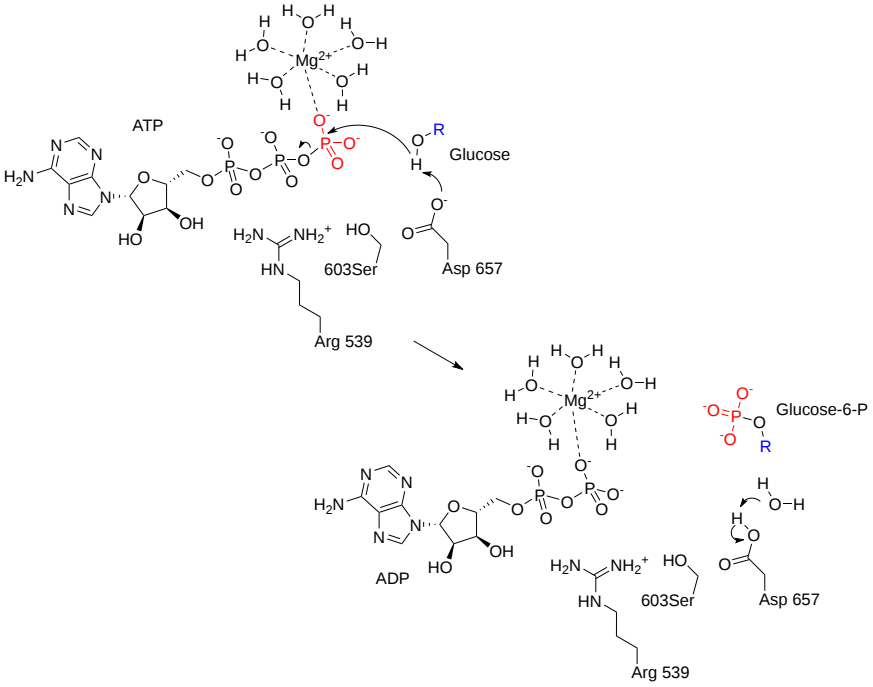


Figure 13.1.10 : Mechanism of human hexokinase I.

Ribeiro AJM *et al.* (2017), *Nucleic Acids Res*, **46**, D618-D623. Mechanism and Catalytic Site Atlas (M-CSA): a database of enzyme reaction mechanisms and active sites. DOI:10.1093/nar/gkx1012. PMID:29106569. https://www.ebi.ac.uk/thornton-srv/m-csa/entry/696/. Creative Commons Attribution 4.0 International (CC BY 4.0) License.

Figure 13.1.11 below shows an interactive iCn3D model of human hexokinase I with glucose and ADP in the active site (1dgk).

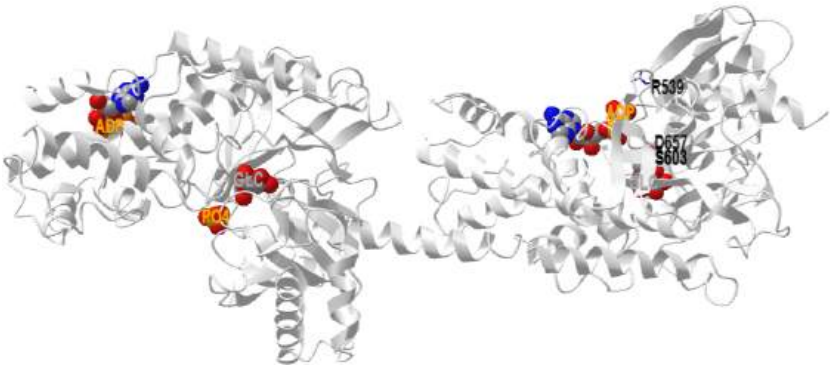


Figure 13.1.11 : Human hexokinase I with glucose and ADP in the active site (1dgk). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...aQy4gf6A2bWtp8

The key catalytic residues are shown in colored sticks and labeled. Glucose, ADP, and PO_4^{2-} are shown as colored spheres.

Note that hexokinase I, as well as forms II and III, have spatially distinct halves similar to those in the yeast enzymes. Forms I-III have a molecule weight of about 100K. Form IV (glucokinase) has a molecular weight of 50K and is similar to the distinct halves of I-III. suggesting that I-III arose through gene duplication.

Each half also binds glucose and ADP but the N-terminals of I and III are catalytically inactive. Hexokinase II, with both halves active, is the main enzyme involved in glycolysis in most mammalian tissues. Other differences are that glucose-6-phosphate, a product, inhibits I-III. PO_4^{2-} relieves G6P product inhibition in form I but not the others. ADP binding at multiple sites in hexokinase I likely causes a conformational change that affects structure and activity.

13.1.3: SUGAR KINASES IN GENERAL

We will see many kinases that use ATP to phosphorylate sugars, so it is useful to explore both their differences and similarities at the beginning of our studies on carbohydrate metabolism.

Figure 13.1.12 below shows the comparative structures of the five carbohydrate kinase classes.

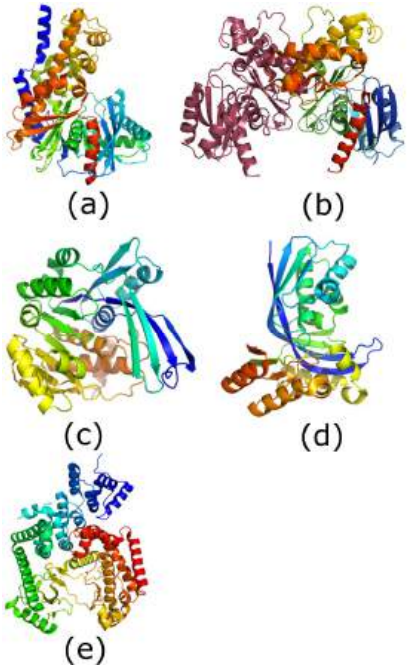


Figure 13.1.12 : Structures of the five carbohydrate kinase classes. All images are shown in rainbow format (blue: *N*-terminus, red: *C*-terminus). Roy, S.; Vivoli Vega, M.; Harmer, N.J. Carbohydrate Kinases: A Conserved Mechanism Across Differing Folds. *Catalysts* **2019**, *9*, 29. https://doi.org/10.3390/catal9010029. **Creative Commons Attribution License**

Panel (**a**) shows the structure of human glucokinase (hexokinase class; PDB (protein data bank) ID: 4IWV).

Panel (**b**) shows the structure of *Bacillus subtilis* fructokinase dimer: the second molecule shown in raspberry (ROK kinase class; PDB ID: 1XC3 [17]).

Panel (**c**) shows the structure of *Escherichia coli* ribokinase (ribokinase class; PDB ID: 1RKD; [18]).

Panel (**d**) shows the structure of *Aquifex aeolicus* IspE (GHMP kinase class; PDB ID: 2V2Z; [19]).

Panel (**e**) shows the structure of human PIK3C3 (phosphatidylinositol phosphate kinase class; PDB ID: 3IHY).

Table 13.1.1 : Overview of the Five Carbohydrate Kinase Classes.

Carbohydrate Kinase Family	Common Substrates	Native Phosphate Donors (Minor Donors in Parentheses)	Pfam ID
Hexokinase	Glucose, mannose, fructose	ATP (ITP)	PF00349, PF03727, PF02685
ROK Kinase	Glucose, allose, fructose, *N*-acetylglucosamine, *N*-acetylmannosamine	ATP (polyphosphate)	PF00480
Ribokinase	Ribose, 2-deoxy-d-ribose, adenosine	ATP, ADP (GTP, ribonucleotide)	PF00294
GHMP Kinase	Galactose, *N*-acetylgalactosamine,	ATP (GTP, ITP)	PF00288
Phosphatidylinositol kinase	Phosphatidylinositol, phosphatidylinositol phosphates	ATP (GTP)	PF00454, PF01504

ROK is a bacterial Repressor, Open reading frame, Kinases are predominantly bacterial enzymes. Ribokinases include adenosine kinases, fructokinases, and phosphofructokinases. GHMP Kinases include Galactokinase, Homoserine kinase, Mevalonate kinase, and Phosphomevalonate kinase. Roy et al. Ibid

Many of these kinases are regulated by the binding of allosteric effectors.

Figure 13.1.13 shows a common mechanism for all, using glucose as an example substrate.

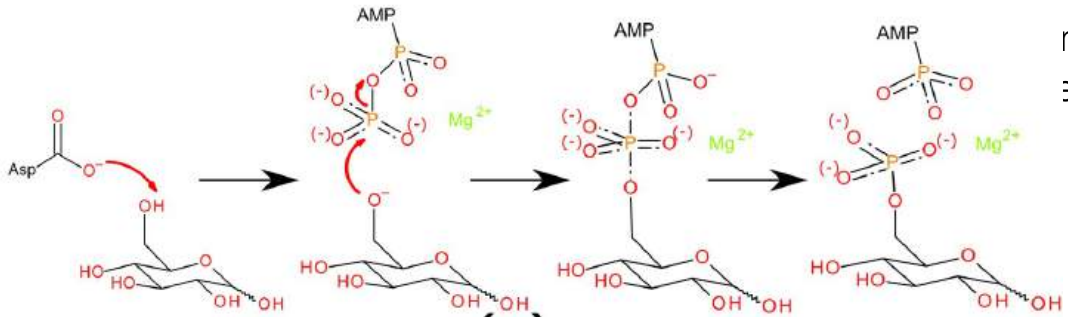


Figure 13.1.13 : Generic reaction mechanism for carbohydrate kinases (hexokinase shown). Roy et al, Ibid. The generic reaction is initiated by a catalytic base abstracting a proton from the reactive hydroxyl (left). The oxygen atom then attacks the -phosphate of ATP (second left), forming a pentacoordinate transition state (second right). This is stabilized by a divalent cation, and by the protein (not shown). This transition state resolves, leaving ADP and the phosphorylated carbohydrate (right).

13.1.4: **REACTION 2**: **GLUCOSE-6-PHOSPHATE ↔ FRUCTOSE-6-PHOSPHATE**. ΔG^{O}=+0.4 KCAL/MOL (+1.7 KJ/MOL)

In the second step of glycolysis, the **phosphoglucose isomerase (PGI)** converts glucose-6-phosphate into one of its isomers, fructose-6-phosphate. Recall that an isomerase is an enzyme that catalyzes the conversion of a molecule into one of its isomers. In this reaction, the aldose, glucose-6-phosphate, is converted to the ketose, fructose-6-phosphate. This conversion is essential to allow the eventual split of the sugar into two three-carbon molecules. We'll present this reaction, catalyzed by **phosphoglucose isomerase (PGI)**, in several different types of representations (chair, wedge/dash, and Fischer projection), as shown in **Figure \(\PageIndex{14}\)** below.

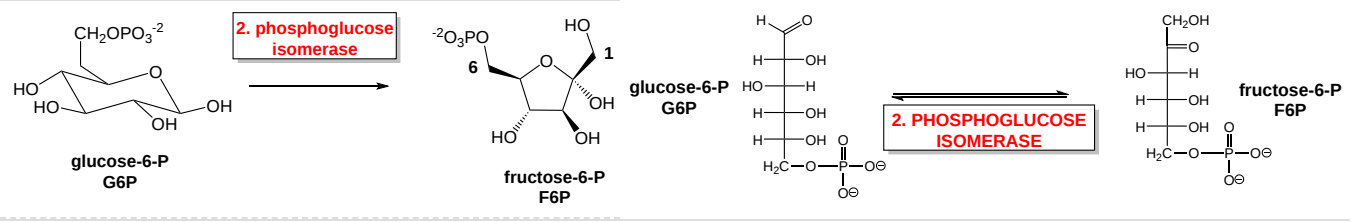


Figure 13.1.14 : Summary reaction, phosphoglucoisomerase

From a functional perspective, seen more clearly from the linear Fischer structure, it is evident that the C=O has been moved to the C2 position to create the ketose structure. The carbonyl O is now positioned to be an electron sink facilitating electron flow for reaction 4. This isomerization reaction would be expected to have a ΔG^o of about 0.

It has been proposed that the mechanism of PGI requires the enzyme to open the Glucose-6-Phosphate ring before the actual isomerization step, which proceeds through the formation of a cis-enediol intermediate and keto-enol tautomers before conversion back to the cyclic form.

Figure 13.1.15 below shows a proposed mechanism for rabbit phosphoglucose isomerase

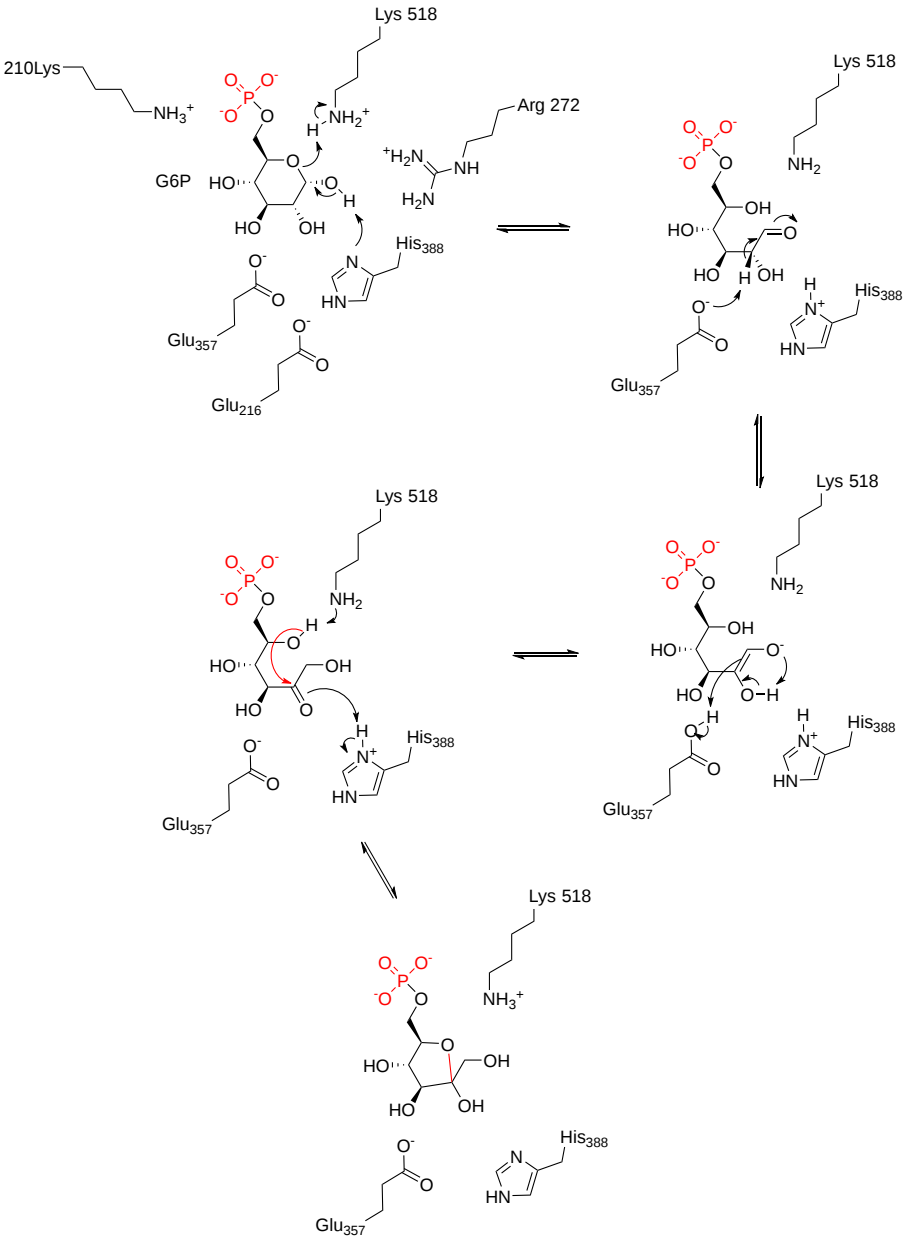


Figure 13.1.15 : Mechanism for phosphoglucose isomerase. Ribeiro AJM *et al. Ibid*

Figure 13.1.16 below shows an interactive iCn3D model of rabbit phosphoglucose isomerase with bound 6-phosphogluconic acid (1DQR)

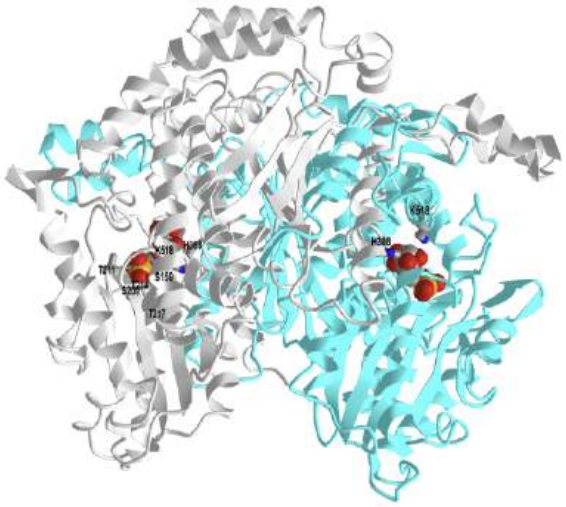


Figure 13.1.16 : Rabbit phosphoglucose isomerase with bound 6-phosphogluconic acid (1DQR). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...xo1SgF1ZpAySm9

One monomer of the dimer is shown in gray and the other in cyan. The catalytic residues Lys518 and His388 are shown in both subunits in colored sticks and labeled. Additional residues that contribute to specificity (Ser209, Ser159, Thr214, Thr217, and Thr211) are shown in the gray subunit with color sticks and labeled. 6-phosphogluconic acid, a competitive inhibitor, is shown in color spacefill.

The enzyme also has other functions in addition to its role in glycolysis, so it is a member of a group called "moonlighting" (a term that refers to working at a secondary job). proteins. Outside of the cell, phosphoglucoisomerase acts as a nerve growth factor and cytokine. It is also called autocrine motility factor (PGI/AMF) and its cytokine activity is associated with aggressive cancers.

13.1.5: **REACTION 3: FRUCTOSE-6-PHOSPHATE → FRUCTOSE-1,6-BISPHOSPHATE.** ΔG^{O}=-3.4 KCAL/MOL (-14 KJ/MOL)

The third step is the phosphorylation of fructose-6-phosphate, catalyzed by the enzyme **phosphofructokinase-1 (PFK1)**. A second ATP molecule donates a high-energy phosphate to fructose-6-phosphate, producing fructose-1,6-bisphosphate. The term bisphosphate is used when two phosphate groups are joined to a molecule at different positions on the molecule. In this case, one phosphate is at the 1-carbon position and the other is at the 6-carbon position. This differs from the term diphosphate, which is used when the phosphate groups are joined in a sequence, as in the molecule ADP. In ADP, both phosphate groups are joined in tandem from the 5-carbon position on the ribose ring structure.

In the glycolytic pathway, PFK1 is a rate-limiting enzyme. The mechanism of PFK1 regulation is discussed in greater detail in Chapter 15.5. However, we will introduce the process here. Essentially, the enzyme is sensitive to the energy load within the cell. Recall that ATP is a recycled molecule and exists in a pool of interconverting ATP, ADP, and AMP. A chief outcome of glycolysis is to shift the pool towards increased levels of ATP to drive endergonic processes like muscle contraction. PFK1 activity is sensitive to the ATP:ADP ratio within the cell. PFK1 is more active when the concentration of ADP is high and the concentration of ATP is low, and it is conversely less active when ADP levels are low and the concentration of ATP is high. This is a type of end-product inhibition since ATP is the end product of glucose catabolism. Note, however, that ATP is also a substrate for PFK1. ATP serves as the phosphate donor in the reaction and is required in the process. This is the second energy-intensive step in the glycolytic pathway. At the end of the PFK1 step, a total of 2 ATP molecules have been broken down in the glycolytic pathway.

The PFK1 enzymatic step is also important within the glycolytic pathway as it is the committed step in the pathway. Glucose-6-phosphate may be used for other purposes within the cell, and the isomerase step that converts glucose-6-phosphate to fructose-6-phosphate is readily reversible. The PFK1 enzyme only works in the forward direction to create fructose 1,6-bisphosphate. It cannot in the reverse direction and recover the reactant. Thus, it is thought of as the committed step in the glycolytic pathway as the fructose-1,6-bisphosphate will predominantly be converted into pyruvate through the remaining enzymatic steps.

The reaction catalyzed by PFK1 is shown in **Figure 13.1.17** below.

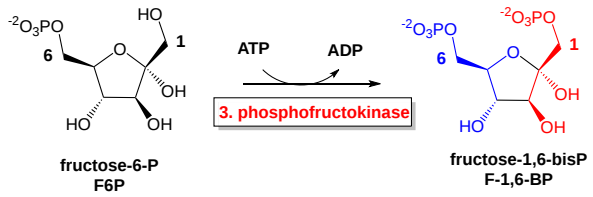


Figure 13.1.17 : Summary reaction - phosphofructokinase

By phosphorylating this intermediate, both products of the cleavage of this 6C molecule will be phosphorylated, keeping both more readily inside the cell. This reaction is a nucleophilic substitution reaction on the gamma phosphate of ATP. A phosphoanhydride bond is broken in ATP as a phosphoester bond is made producing fructose-1,6-bisphosphate. As in reaction 1, this reaction proceeds with a negative ΔG^o

The PFK1 enzyme exists as a tetramer and as with another "famous" tetramer, hemoglobin, it can exist in T and R symmetry states. Hence it is an allosteric enzyme and has many allosteric regulators. One important allosteric activator of eukaryotic (but not prokaryotic) PFK is fructose-2,6-bisphosphate. This is formed by a separate **phosphofructokinase enzyme named PFK2**. This regulatory pathway will be described in greater detail in Chapter 15.5. Hence the number 1 is added to the name of the glycolytic enzyme, PFK1, that forms fructose 1,6-bisphosphate, to indicate the position of phosphorylation. Mammals have three isoforms of PFK1, muscle (PFKM), liver (PFKL), and platelet (PFKP).

The mechanism for E. Coli phosphofructokinase is shown in **Figure 13.1.18** below. A metal cofactor (Mg2+) coordinates the positioning of the ATP and stabilizes the gamma phosphate for nucleophilic attack by the fructose alcohol group at position-1. Activation of the fructose alcohol group is mediated by proton abstraction with a coordinated Asp_{127} of PFK1.

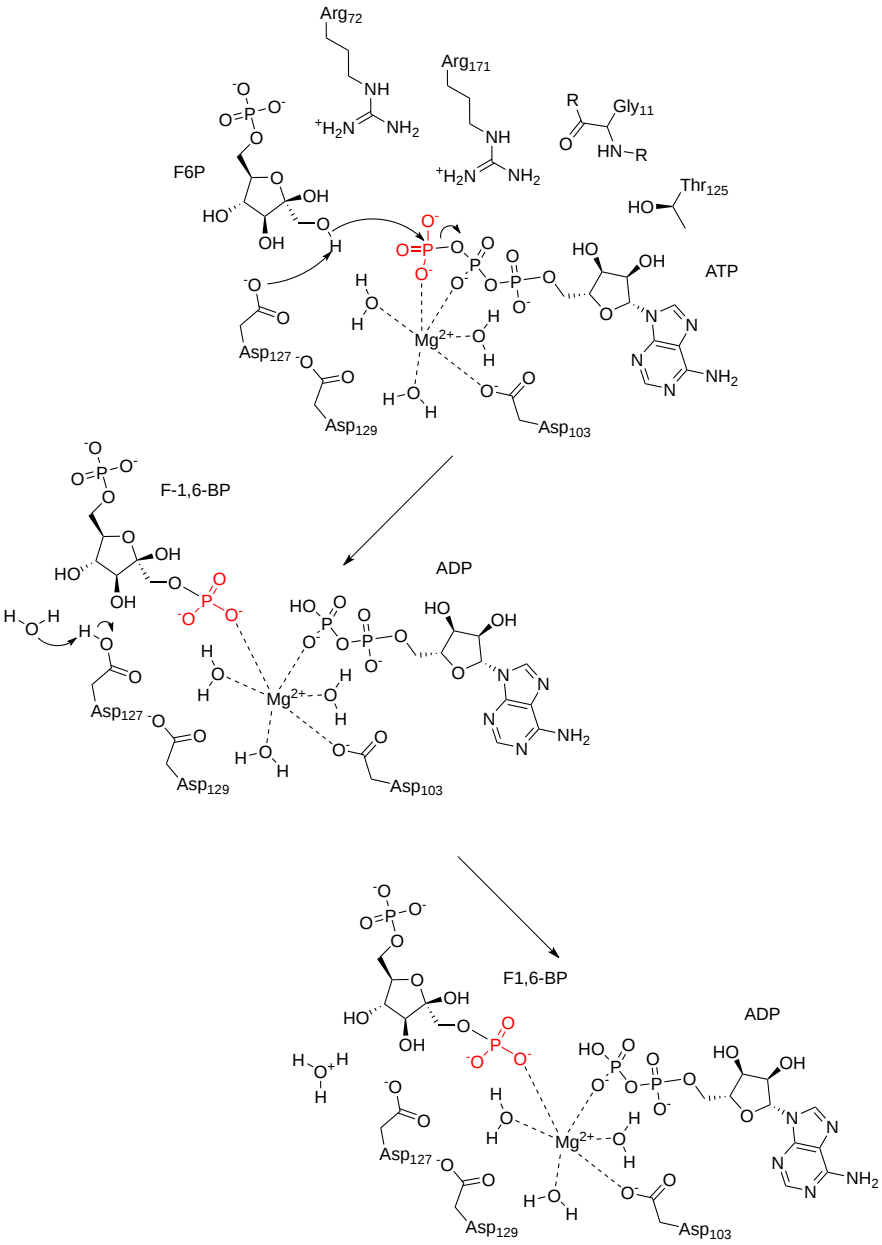


Figure 13.1.18 : Proposed mechanism for E. Coli phosphofructokinase. Ribeiro AJM *et al. Ibid*

Figure 13.1.19 below shows an interactive iCn3D model of the *E. Coli* phosphofructokinase with bound F1,6-bisphosphate and ADP products (1PFK). (long load)

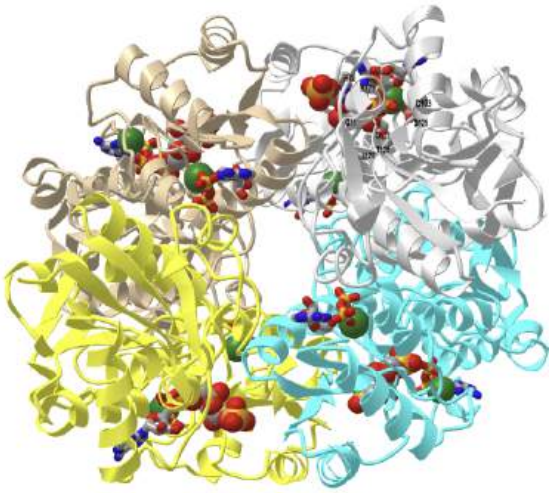


Figure 13.1.19 : E. Coli phosphofructokinase with bound F1,6-bisphosphate and ADP products (1PFK). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...t4KKgLAGzHrzo8

Each monomer of the tetramer is shown in a different color. The gray monomers show key binding and catalytic residues described in the mechanism above. F1,6-bisphosphate is shown as spacefill and ADP is shown as sticks.

The structure of the human platelet PFK1 tetramer has been determined in the presence of ATP and ADP. **Figure 13.1.20** below shows an interactive iCn3D model of the human phosphofructokinase-1 dimer (for clarity) in complex with ATP and Mg (4XYJ). (long load)

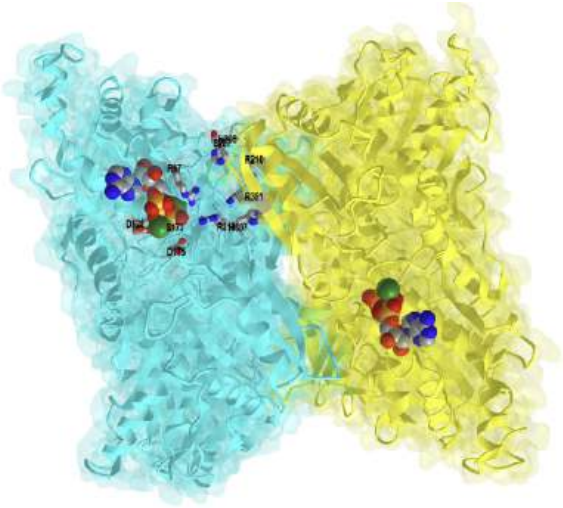
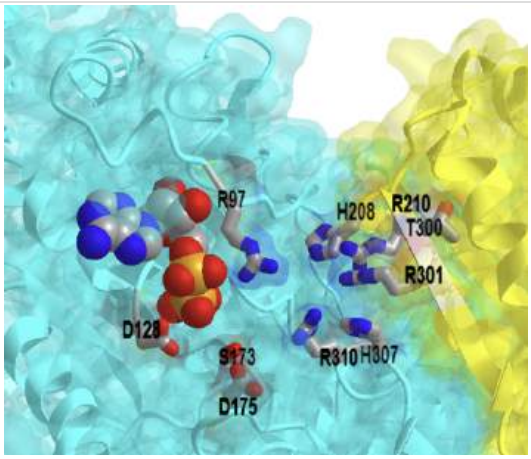


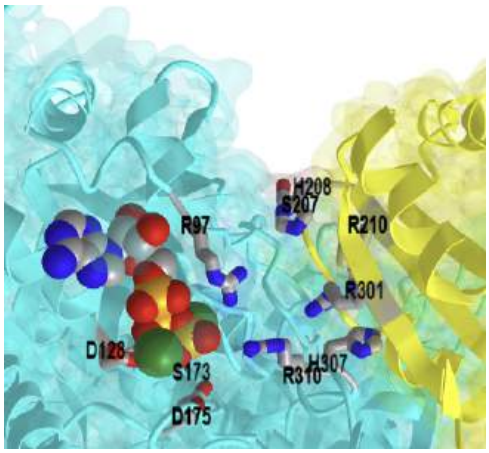
Figure 13.1.20 : Human phosphofructokinase-1 dimer in complex with ATP and Mg (4XYJ). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...AKfonyGsNq6Cs9 (long load). Here is a link to the ADP bound structure: https://structure.ncbi.nlm.nih.gov/i...N81MBS4PZfSQp9

The monomers are shown as colored surfaces with secondary structures underneath. Note that side chains from each monomer in the dimer contribute to binding and catalysis. The actual PDB structure is tetrameric. The structures have an E173S mutation.

Figure 13.1.21 shows the differences in orientation of the key side chains in the ADP structure (left) and the ATP structure (right) that are key in the conformation changes in the enzyme on binding substrate.



ADP bound form of human PFK (4XKJ)



ADP bound form of human PFK (4XYJ)

Figure 13.1.21 : Conformation changes in the active site of PFK on ATP hydrolysis

13.1.6: REACTION 4: **FRUCTOSE-1,6-BISPHOSPHATE → DIHYDROXYACETONE PHOSPHATE (DHAP) + GLYCERALDEHYDE 3-PHOSPHATE (G3P).** ΔG^{O} = + 5.7 KCAL/MOL (+24 KJ/MOL)

The newly added high-energy phosphates further destabilize fructose-1,6-bisphosphate. The fourth step in glycolysis employs an enzyme, aldolase, to cleave fructose-1,6-bisphosphate into two three-carbon isomers: dihydroxyacetone phosphate and glyceraldehyde-3-phosphate. This reaction is shown in **Figure 13.1.22** below in both wedge dash and Fischer projections.

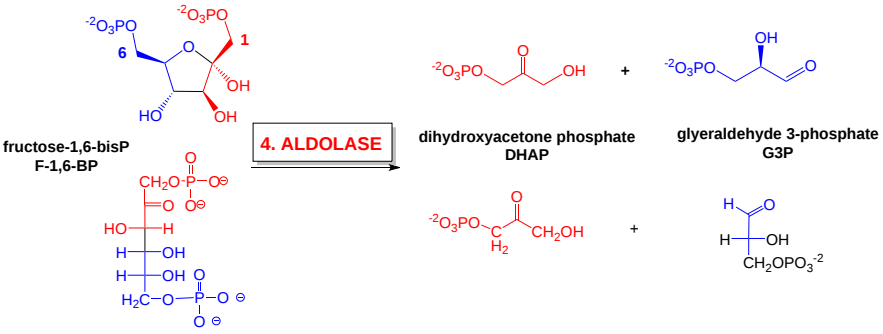


Figure 13.1.22 : Summary reaction - aldolase

This is the first C-C bond cleavage within glucose on the path to complete cleavage during aerobic respiration and release of 6 carbon dioxide molecules per glucose metabolized. This reaction is the reverse of an aldol condensation when an enol or enolate reacts with a carbonyl C to form an adduct. Note that both products are phosphorylated. The reaction is not thermodynamically favored but is pulled in the forward direction by the utilization of the product in subsequent reactions in the pathway.

Table 13.1.2 Characteristics of the Three Classes of Aldolases (I, IA, and II) and the Organisms in Which They are Found.

Aldolase class	Organisms	Characteristics
I	Higher eukaryotes (plants, animals), some protozoans (e.g., Trypanosoma brucei, Leishmania donovani, Plasmodium falciparum), some bacteria, some algae	• Mechanism of action: Schiff base forms substrate intermediate • Structure: Typically forms homo-tetramers, contains TPI β/α-barrel fold
IA	Archaea	• Mechanism of action: Schiff base forms substrate intermediate • Structure: Typically forms homo-octamers/ decamers/higher oligomers, contains TPI β/α-barrel fold
II	"Lower" eukaryotes (some protozoans (e.g., Giardia lamblia), fungi, yeasts, algae), most bacteria, some archaea	• Mechanism of action: Divalent ion forms substrate intermediate, metal-dependent • Structure: Typically forms homo-dimers, contains TPI β/α-barrel fold

Class I Aldolase

These enzymes proceed through a Schiff base intermediate between a reactive lysine and the reactant/product. The enzyme is favored in the reverse direction and it's perhaps easier to see the mechanism presented in that fashion. In rabbits, the muscle Class I aldolase (RAMA) uses Lys-229 to form a Schiff base with DHAP as shown in a mechanism presented in **Figure 13.1.23** below.

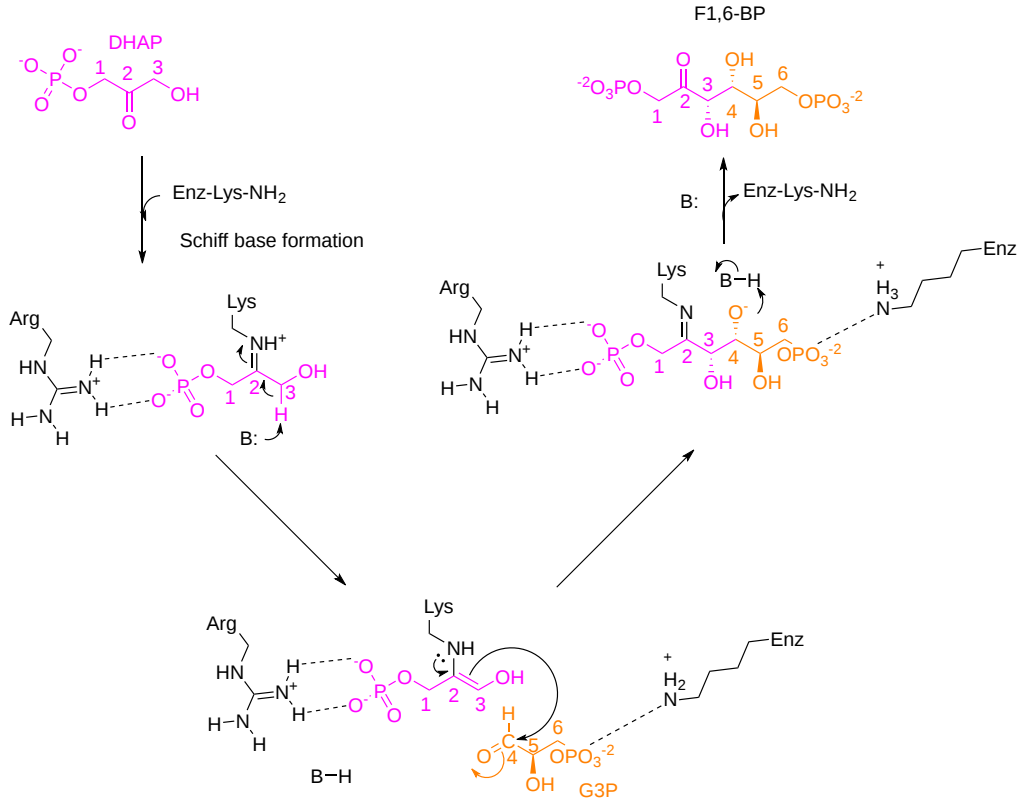


Figure 13.1.23 : Class I aldolases - general mechanism (after Bolt et al., Arch Biochem Biophys. 2008 June 15; 474(2): 318–330)

Only the pro(S) proton of the dihydroxyacetone phosphate C3 carbon is removed and effectively exchanged with the glyceraldehyde-3-phosphate substrate. **Figure 13.1.24** below reviews how the pro(R) and pro(S) hydrogens can be visually differentiated by replacing one

with a deuterium and determining the stereochemistry of the now chiral C3.

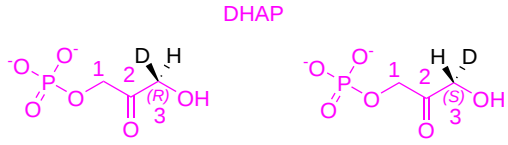


Figure 13.1.24 : Visualization of the Pro(R) and Pro(S) Hydrogens on DHAP

Figure 13.1.25 shows key active site residues in the active site of a Class I aldolase from rabbit muscle. The first step in Schiff base formation with Lys229 is shown.

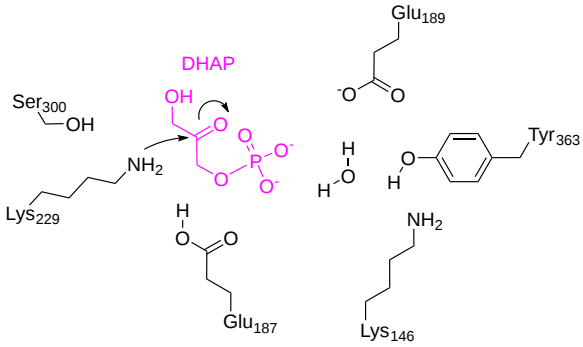
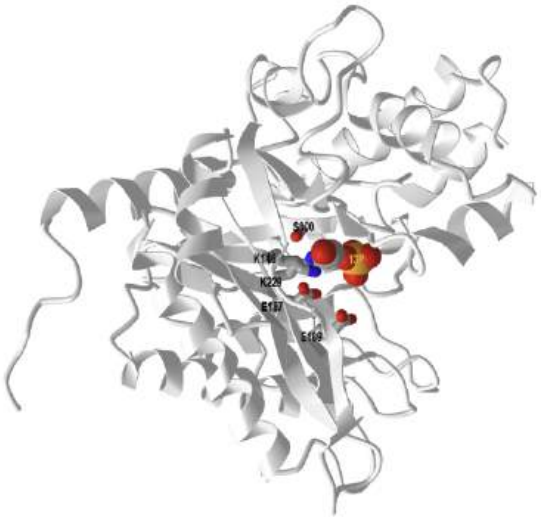


Figure 13.1.25: Catalytic residues in rabbit muscle Class I enolase.

Figure 13.1.26 shows an interactive iCn3D model of the dihydroxyacetone phosphate enamine intermediate in fructose-1,6-bisphosphate aldolase from rabbit muscle (2QUT)



NCBI iCn3D Figure 13.1.26: Dihydroxyacetone phosphate enamine intermediate in fructose-1,6-bisphosphate aldolase from rabbit muscle (2QUT). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...uR5KZBsfN8QQh9

Only one monomer of the four in the homotetramer is shown. The ligand, 1,3-dihydroxyacetonephosphate, is shown in spacefill with CPK colors. It is Schiff base linkage with Lys229. The key catalytic residues are shown and labeled.

The reaction proceeds between eneamine and iminium covalent intermediates. A K146M mutation (in the active site) decreases enzyme activity and allows the trapping of the K229 eneamine intermediate. A key tyrosine (Y363) in its deprotonated state formed in the presence of the iminium phosphate and a local water appears to abstract the C3 pro(S) proton to form the enamine.

Class II Aldolase

Figure 13.1.27:

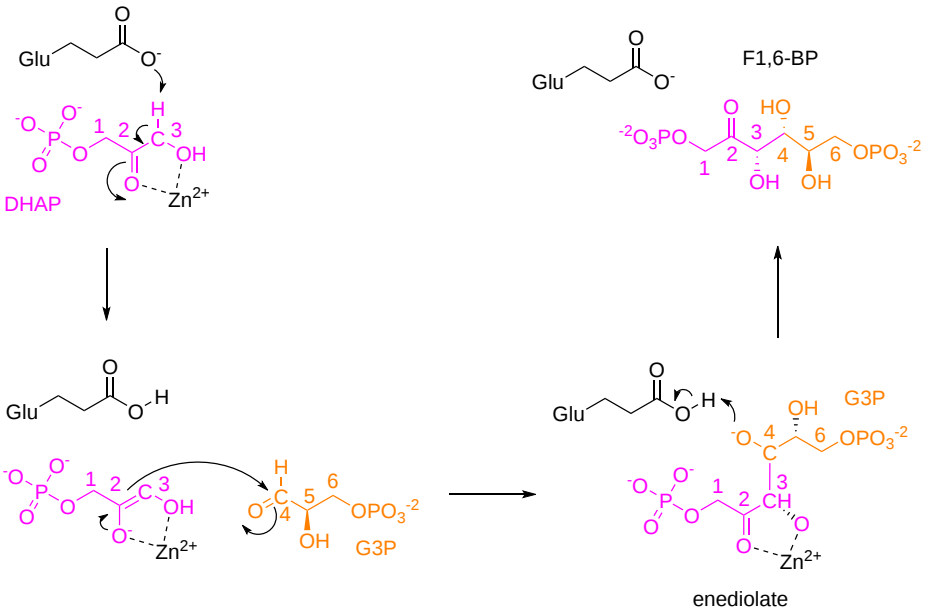


Figure 13.1.27: Class II aldolases - general mechanism (after Bolt et al, ibid)

13.1.7: **REACTION 5: DHAP ↔ G3P**. ΔG^{O}= +1.8 KCAL/MOL (+7.5 KJ/MOL)

This reaction, catalyzed by triose phosphate isomerase (TPI), is shown in Figure 13.1.28.

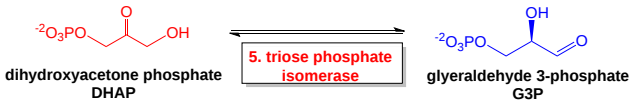


Figure \(\PageIndex{28}\: Summary reaction - triose phosphate isomerase.

This is another simple isomerization reaction. Only one product, glyceraldehyde-3P continues on in glycolysis, so only one enzyme is needed to metabolize the cleavage products of this reaction further. As in other isomerization reactions, the ΔG^{o} is close to 0.

A proposed mechanism for this reaction is shown in Figure \(\PageIndex{29} below.

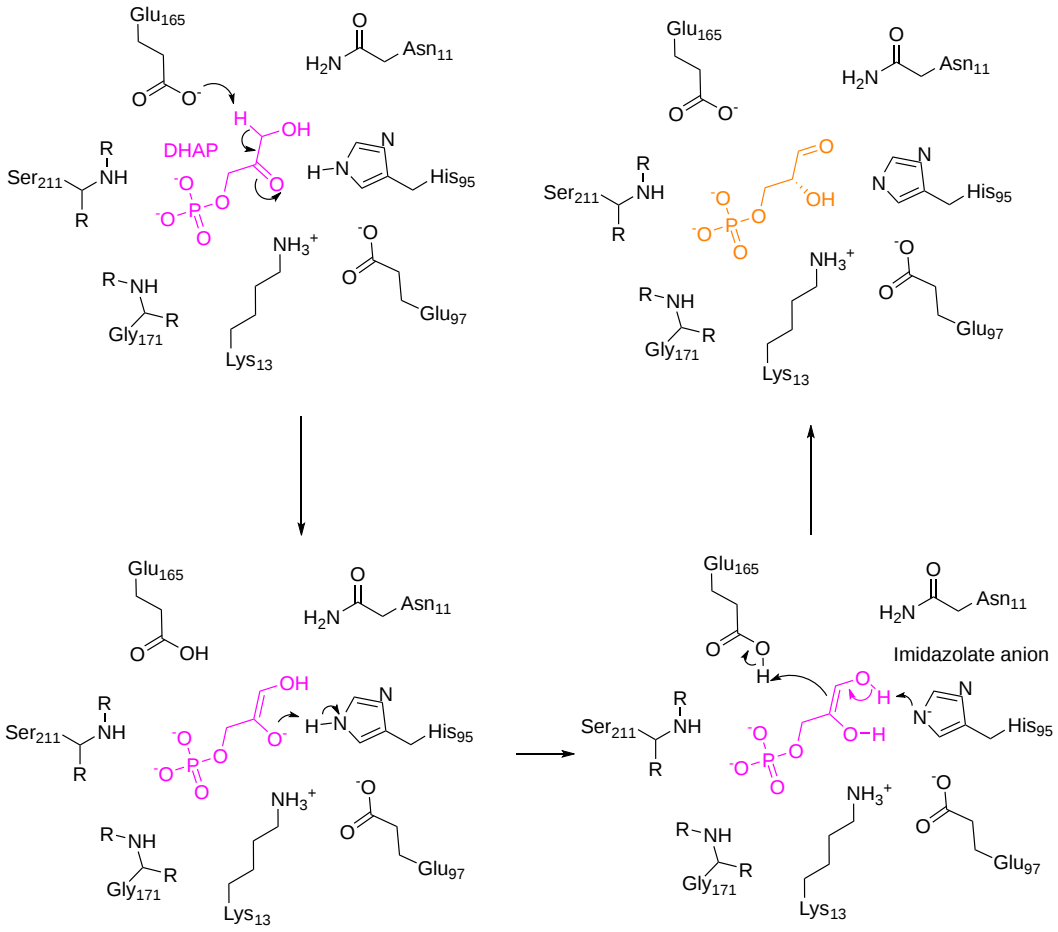
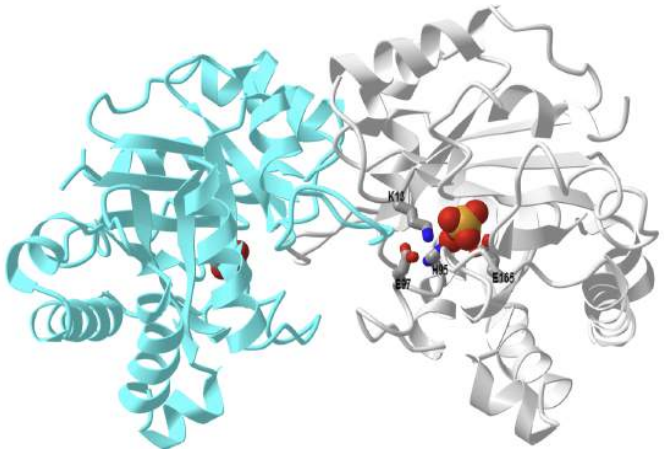


Figure 13.1.29: Proposed mechanism for triose phosphate isomerase (after Bolt et al, ibid)

Note that the reaction employees the deprotonation of a neutral imidazole side chain to form an imidazolate anion. This would not likely be favorable, given its pKa value.

The are four possible conserved proton donors with more reasonable pKas in the active sites of TPIs, including K12, H95, E97, and E165. Mutations of E97 to E97Q and E97D lead to a 4000-fold reduction in k_{cat} (E97Q) but only a 100-fold reduction for E97D suggesting tha30}\) shows an interactive iCn3D model of the chicken triosephosphate isomerase-phosphoglycolohydroxamate complex (1TPH)



NCBI iCn3D Figure 13.1.30: Chicken triosephosphate isomerase-phosphoglycolohydroxamate complex (1TPH). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...AFXpt9kg5aki5A

The four key conserved residues are shown in the gray monomer of the homodimer.

13.1.8: **REACTION 6: G3P ↔1,3 BPG.** ΔG^{O}= +1.5 KCAL/MOL (+6.1 KJ/MOL)

This reaction, catalyzed by **glyceraldehyde-3-phosphate dehydrogenase**, is shown in Figure 13.1.31

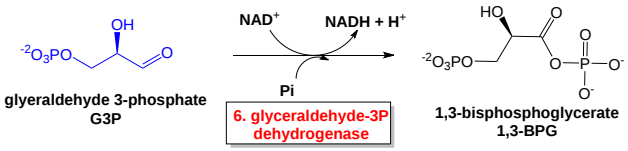


Figure \(\PageIndex{31}\: Summary reaction - glyceraldehyde-3-phosphate dehydrogenase

This is a big reaction! Note that the ΔG^{o} is close to 0 but look at what happened. The carbonyl O in G3P has been oxidized to the form of a mixed anhydride which can donate a phosphate to ADP (in the next step) to form ATP. That this is an oxidation reaction should be obvious from the fact that the carbonyl C in G3P has two bonds to O but 3 bonds in 1,3 BPG.

To carry out an oxidation reaction, you need an oxidizing agent. In comes NAD^{+}, a modest but very prevalent oxidizing agent in biology. When glucose is oxidized completely by O_2 to CO_2 during combustion, much energy is released so we can surmise that oxidation reactions, if carried out by powerful oxidizing agents like O_2, proceed with a large negative ΔG^{o}. For every oxidation reaction, the oxidizing agent is reduced. All reactions are potentially reversible so the products formed are new potential oxidizing and reducing agents. As in acid/base reactions, which proceed from stronger acid to weaker conjugate acid, redox reactions proceed from a stronger to a weaker oxidizing agent. For reaction 6, we must use tables of redox potentials to calculate the actual ΔG^{o}. It turns out to be close to 0, which is great since in the same reaction, a substrate-level phosphorylation reaction (using inorganic phosphate - P_i instead of ATP) occurs. In summary, this reaction catalyzes the first and only oxidation of glucose in glycolysis which has paid (thermodynamically) for the generation of a mixed anhydride whose phosphorylating potential is higher than that of ATP.

Figure 13.1.32 shows a proposed mechanism for glyceraldehyde-3-phosphate dehydrogenase from Trypanosoma cruzi

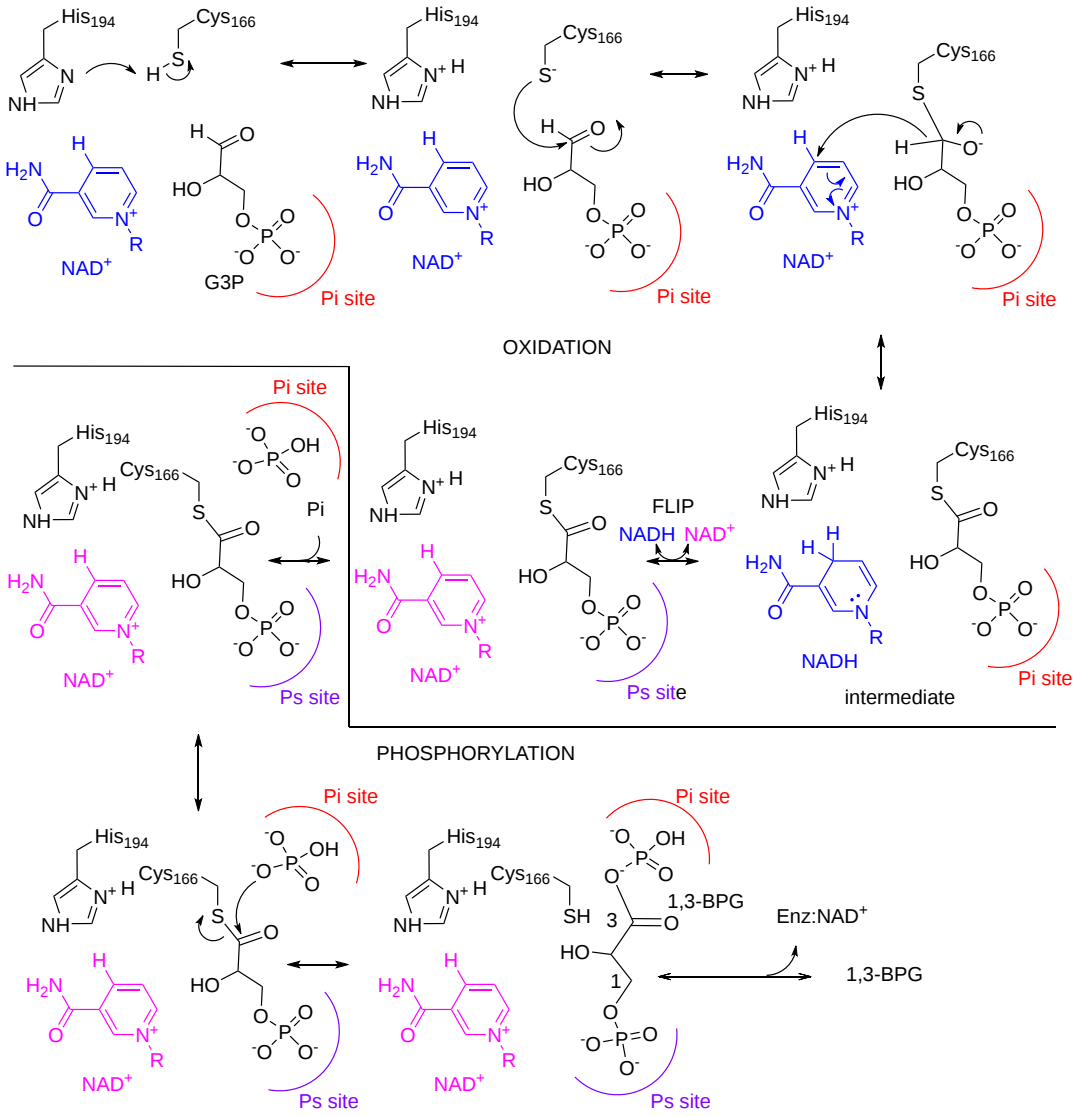
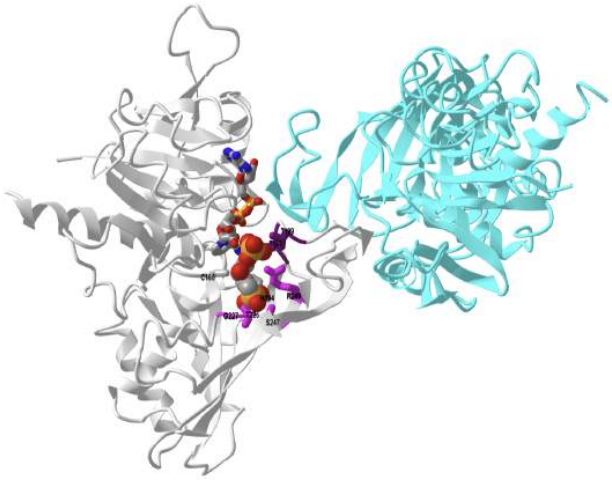


Figure 13.1.32: Proposed mechanism for glyceraldehyde-3-phosphate dehydrogenase from Trypanosoma cruzi (after Reis et al. Phys. Chem. Chem. Phys., 2013, 15, 3772. https://pubmed.ncbi.nlm.nih.gov/23389436/)

The reaction proceeds in two parts. The first (top section) is the oxidation of glyceraldehyde-3-phosphate (G3P) by NAD^+ to the state of a thioester attached to Cys166. The now reduced NADH dissociates and is replaced by a new NAD^+ for another cycle of catalysis. In the meantime, inorganic phosphate, Pi, binds and reacts with the thioester to form 1,3-bisphosphoglycerate (1,3-BPG).

The mechanism is not entirely clear. There are two phosphate binding sites, Pi (red) and Ps (purple) which interact with phosphate groups on the substrates. The oxidation part of the reactions appears to take place at the Pi site. Pi is composed of the side chains of Thr197, Thr199, and the 2-hydroxyl group of the ribose in NAD^+. The Pi site in the T. cruzi enzyme consist of Thr226 and Arg249, Gly227, and Ser 247. The mechanism appears to involve a flip of the orientation of substrates and intermediates after the dissociation of NADH from the enzyme.

Figure 13.1.33 shows an interactive iCn3D model of the Trypanosoma cruzi glyceraldehyde-3-phosphate dehydrogenase with bound NAD and a 1,3-bisphosphoglycerate analogue (1QXS)



NCBI iCn3D Figure 13.1.33: Trypanosoma cruzi glyceraldehyde-3-phosphate dehydrogenase with bound NAD and a 1,3-bisphosphoglycerate analogue (1QXS) . (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...WnsNSgmA3AJyw6

Only two monomers of the homotetramer are shown for clarity. The Pi site is shown in magenta and the Ps site is in purple. The phosphonic acid analog is shown in spacefill and NAD in sticks. The active site His194 and Cys166 are also shown in sticks and labeled.

13.1.9: **REACTION 7: 1,3 BPG + ADP + H^+ → 3PG + ATP** ΔG^O= -4.5 KCAL/MOL (-19 KJ/MOL)

This reaction, catalyzed by **phosphoglycerate kinase**, is shown in Figure 13.1.34

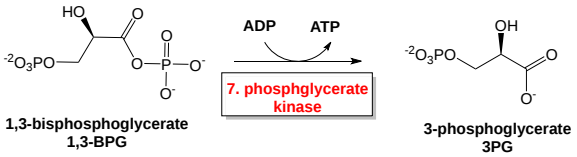


Figure 13.1.34: Summary reaction - phosphoglycerate kinase

It's finally happened! An ATP has been made for each of the two 1,3-BPG molecules derived from glucose. We've made back the ATP used in steps 1 and 3. A mixed phosphoanhydride bond is broken in 1,3 BPG as a phosphoanhydride bond is made in ATP. As the mixed phosphoanhydride has higher energy than its hydrolysis product compared to the phosphoanhydride in ATP, the reaction proceeds with a negative ΔG^o

A mechanism for the reaction is shown in Figure 13.1.35.

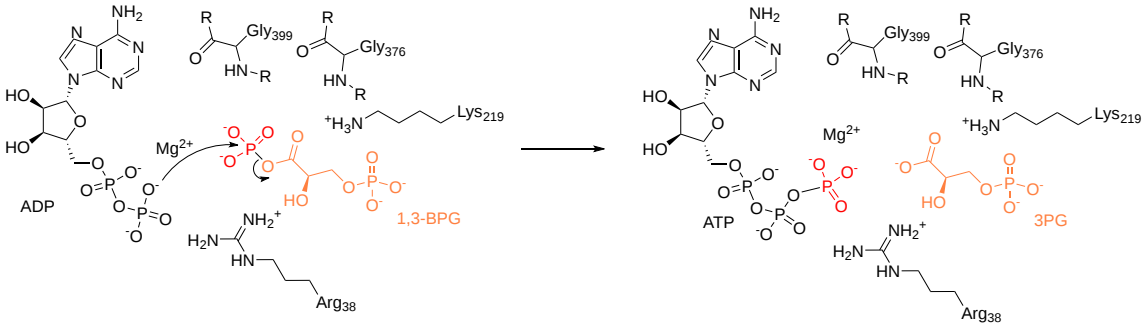
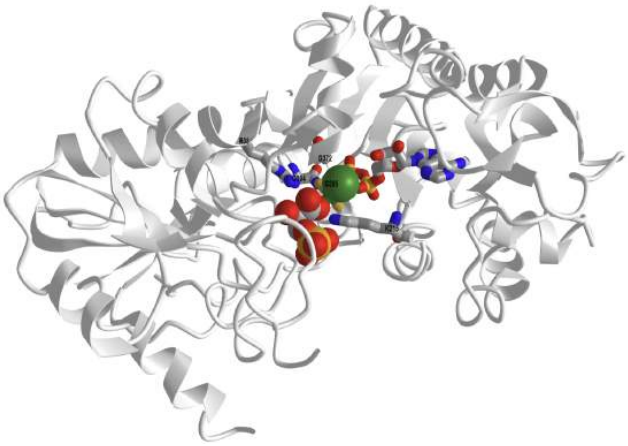


Figure 13.1.35: Reaction mechanism for phosphoglycerate kinase

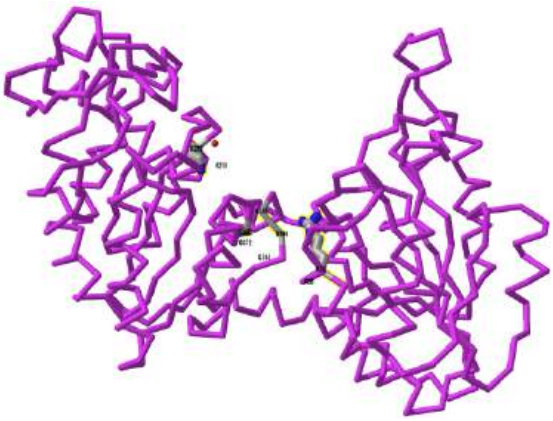
Figure 13.1.36 shows an interactive iCn3D model of human phosphoglycerate kinase in complex with ADP, 3PG, and magnesium trifluoride (2WZB).



NCBI iCn3D Figure 13.1.36: Human phosphoglycerate kinase in complex with ADP, 3PG, and magnesium trifluoride (2WZB). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...w5yvu2VuGhmQXA

3-phosphoglycerate is shown in spacefill with CPK colors. ADP plus the adjacent MgF_3 is a mimetic for the transition state for ATP synthesis. Hence this structure shows the products/transition state in a closed active site, which as we have seen before prevents spurious hydrolysis of 1,3-BPG or ATP.

Figure 13.1.37 shows an interactive iCn3D model of the alignment of the open form of human phosphoglycerate kinase (2XE7) with bound 3PG and ADP with the closed form with bound 3PG, ADP, **and MgF_3** (2WZB).



NCBI iCn3D Figure 13.1.37: Alignment of the open form of human phosphoglycerate kinase (2XE7) with bound 3PG and ADP with the closed form with bound 3PG, ADP, **and MgF_3** (2WZB). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...n5FfsJeYitiWr9

Use the "a" key to toggle back between the open form (magenta) and closed forms (cyan). In the closed state most closely representing the bond transition state of ATP, the two lobes of the enzyme clamp together.

13.1.10: **REACTION 8: 3PG ↔ 2PG** ΔG^{O}= +1.1 KCAL/MOL (4.6 KJ/MOL)

This reaction, catalyzed by **phosphoglycerate mutase (PGM)**, is shown in Figure 13.1.38

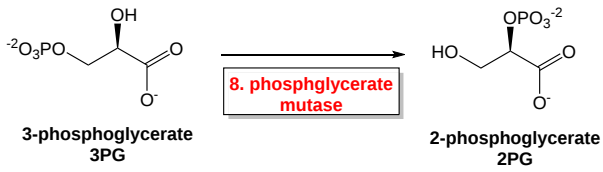


Figure 13.1.38: Summary reaction - phosphoglycerate mutase

This isomerization reaction proceeds with little thermodynamic barrier. Its function is to locate the phosphate on C2 which on the next reaction (dehydration) will form a molecule whose phosphoryl transfer potential is greater than ATP. It seems so simple but the enzymes that catalyze this reaction are diverse and quite complicated from a mechanistic perspective.

There are two types of PGMs, bisphosphoglycerate and monophosphoglycerate mutases that carry out three different reactions involving shuffling of phosphates from one position to another in 3C sugars or cleavage of a phosphate from a sugar

- 3-phosphoglycerate ↔ 2-phosphoglycerate (reaction 8 of glycolysis) catalyzed by bisphosphoglycerate and monophosphoglycerate (the glycolytic enzyme) mutases
- 1,3-bisphosphoglycerate ↔ 2,3-bisphosphoglycerate by bisphosphoglycerate mutases
- 2,3-phosphoglycerate ↔ 3-phosphoglycerate + Pi by bisphosphoglycerate mutases

Within the monophosphoglcyerate mutases, and more specifically the phosphoglycerate mutase (PGM) of glycolysis, there are two types

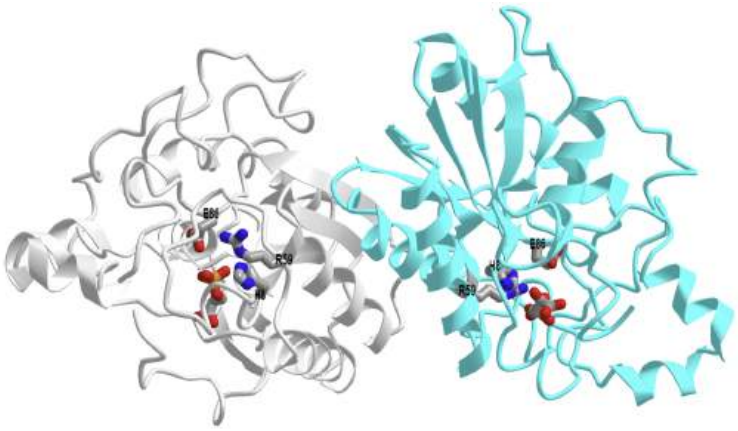
- one that depends on the cofactor 2,3-phosphoglycerate. These are called cofactor-dependent phosphoglycerate mutase (dPGM) and are found in mammals, yeast, and some bacteria. They do not require metal ions.
- one that does not depend on the cofactor 2,3-phosphoglycerate. These are called cofactor-independent phosphoglycerate mutase (iPGM) and are found in plants and some bacteria. These can only interconvert 3PGA and 2PGA. One family of enzymes in the class requires Mn^{2+} while the other requires Mg^{2+} or Zn^{2+}. These enzymes are often structurally similar to alkaline phosphatases

The cofactor-independent and cofactor-dependent monophosphoglycerates (such as the phosphoglycerate mutase of glycolysis) are very different structurally and mechanistically so we will look at both types of mechanisms. Within each type, the enzyme sequences are very conserved.

Mechanism of cofactor (2,3-BPG) dependent phosphoglycerate mutase (dPGM)

The reaction is much less complicated than the cofactor-independent PGM. In E. Coli, the reaction involves the transfer of the phosphate on the C3-OH to the nucleophilic nitrogen on histidine 8 (His 10 in other enzymes) in the active site to form a covalent pHis8 intermediate. The phosphate on pHis could then be transferred to the O on carbon C2 of the substrate. An active His 181in E. Coli may also act as a general acid/base and is adjacent to the Glu 88 in the active site.

Figure 13.1.39 shows an interactive iCn3D model of yeast phosphoglycerate mutase (cofactor dependent) bound to 3-phosphoglycerate (1QHF)



NCBI iCn3D Figure 13.1.39: Yeast phosphoglycerate mutase (cofactor dependent) bound to 3-phosphoglycerate (1QHF). (Copyright; author via source). A dimer of the active tetramer is shown. Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...9F4oc6mmr5DFs9

Mechanism of cofactor-independent phosphoglycerate mutase (iPGM)

Let's consider the mechanism for the Mn^{2+}-requiring iPGM from Geobacillus stearothermophilus. Two Mn^{2+} ions are in the active site. The mechanism for the first half of the cofactor-independent phosphoglycerate mutase (iPGM) reaction is shown in Figure 13.1.40.

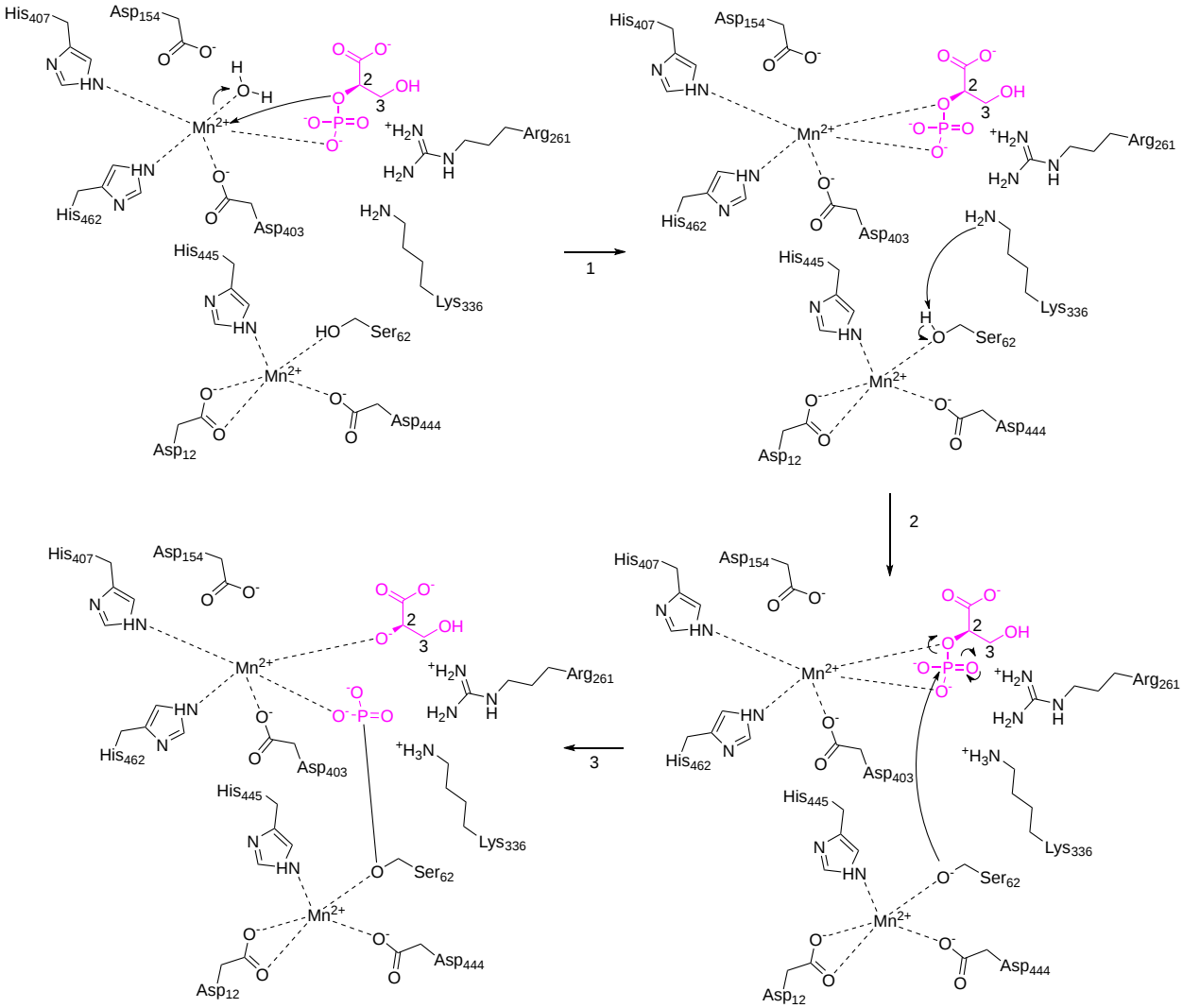


Figure 13.1.40: Part A - Mechanism for cofactor independent phosphoglycerate mutase (iPGM) from Geobacillus stearothermophilus (after Bolt et al, ibid)

Arg 261 interacts with the substrate, stabilizing the negative charge in it and its transition state. It also makes the target phosphorous more electrophilic. Ser62 is activated by a Mn^{2+} ion to become more nucleophilic on the abstraction of a proton by Lys336. Reaction 3 probably proceeds through an S_N2 mechanism.

The rest of the mechanism is shown in Figure 13.1.41.

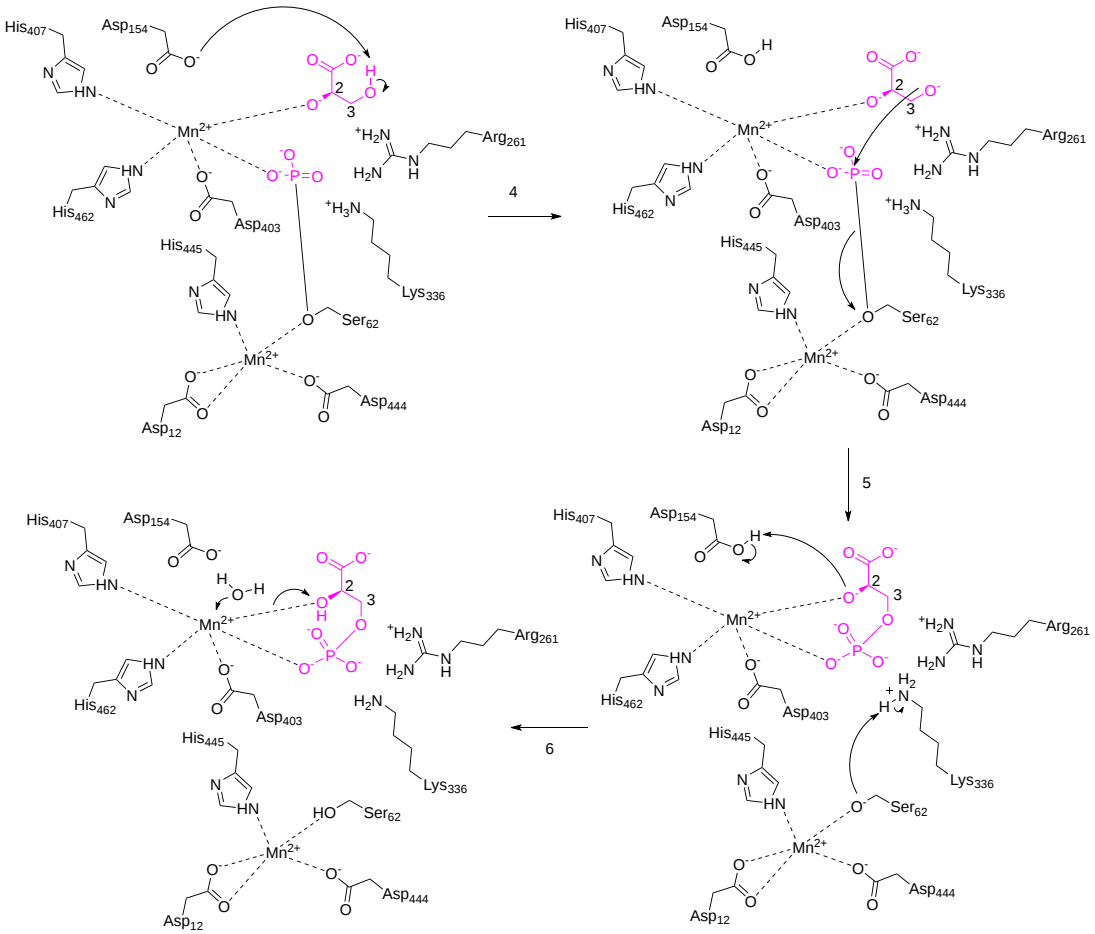
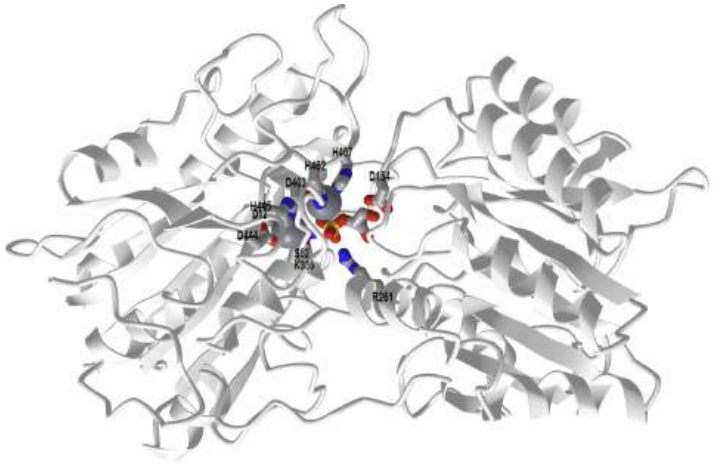


Figure 13.1.41: Part B - Mechanism for cofactor independent phosphoglycerate mutase (iPGM) from Geobacillus stearothermophilus (after Bolt et al, ibid)

The iPGMs are monomers with two distinct domains (lobes) containing a substrate binding site separated from an active site. They are connected by flexible sequences that bend to produce either an open or closed form of the enzymes (as we have seen before). In some enzymes, the two sites are merged at the interface between the domain. The active site Ser62 becomes phosphorylated and then transfers its phosphate to the new site in the substrate which is oriented differently in the enzyme.

Figure 13.1.42 shows an interactive iCn3D model of the Geobacillus stearothermophilus cofactor-Independent Phosphoglycerate Mutase (iPGM) with bound 2-phosphoglycerate (product) (1O98)



NCBI iCn3D Figure 13.1.42: Cofactor-Independent Phosphoglycerate Mutase with bound 2-phosphoglycerate (product) (1O98). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...71iYHXAEKMGsEA

13.1.11: **REACTION 9: 2PG ↔ PEP + H_2O** ΔG^O= +0.4 KCAL/MOL (+1.7 KJ/MOL)

This reaction, catalyzed by **enolase**, is shown in Figure 13.1.43

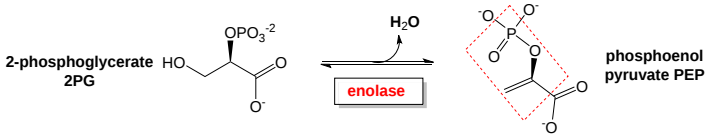


Figure \(\PageIndex{43}\: Summary reaction - enolase

Now you can see the rationale for reaction 8. With a simple dehydration, a molecule with high phosphoryl transfer potential has been produced which in the next and final step of glycolysis produces ATP.

This enzyme has an active site Mg^{2+} that is required for catalysis. Mammals have three forms of the enzyme, α-enolase (ENO-1) found in most tissues, β-enolase (ENO-3) found mostly in muscle, and γ-enolase (ENO-2) found in neurons. A possible mechanism for yeast enolase is shown in Figure 13.1.44.

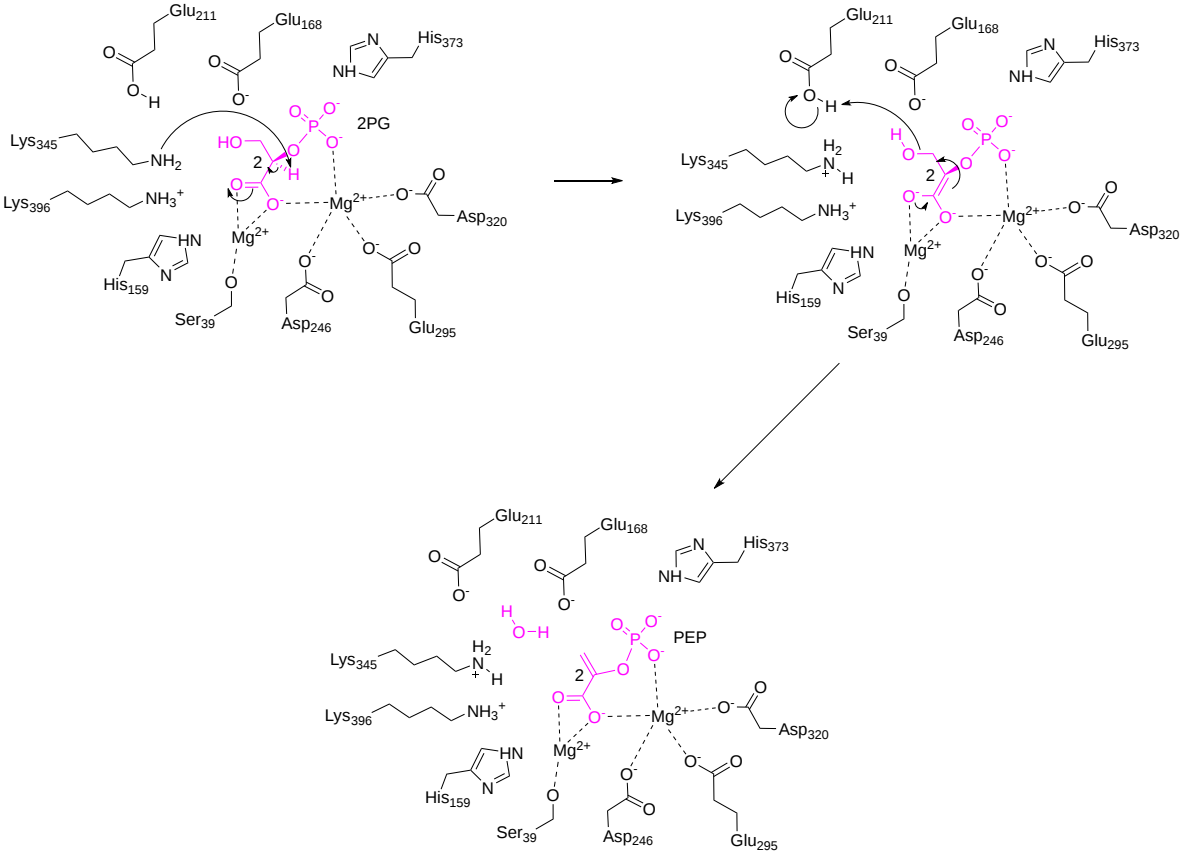


Figure 13.1.44: A possible mechanism for yeast enolase

Figure 13.1.45 shows an interactive iCn3D model of Yeast enolase with bound 2-phosphoglycerate (7ENL)

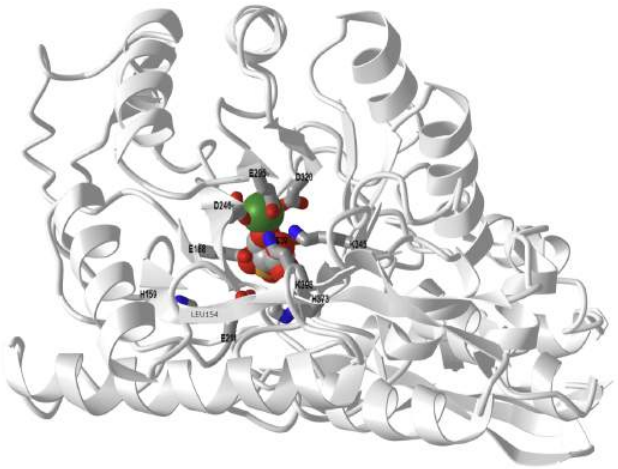


Figure 13.1.45: Yeast enolase with bound 2-phosphoglycerate (7ENL). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...TgwYV9vF8cYtPA

13.1.12: **REACTION 10: PEP + ADP → PYR + ATP** ΔG^{O}= -7.5 KCAL/MOL (-31 KJ/MOL)

This reaction, catalyzed by **pyruvate kinase**, is shown in Figure 13.1.46

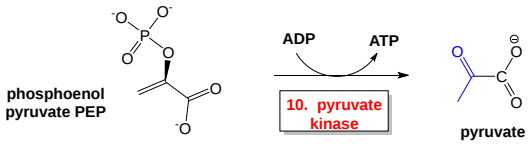


Figure 13.1.46: Summary reaction - pyruvate kinase

In this step, 1 more ATP is made for each PEP consumed (hence 2 ATPs for both 3C PEPs). The phosphoryl transfer potential for PEP is higher than for ATP, which allows this reaction to proceed with a large negative ΔG^{o} (-7.5 kcal/mol, -31 kJ/mol).

The mechanism for rabbit pyruvate kinase is shown in Figure 13.1.47.

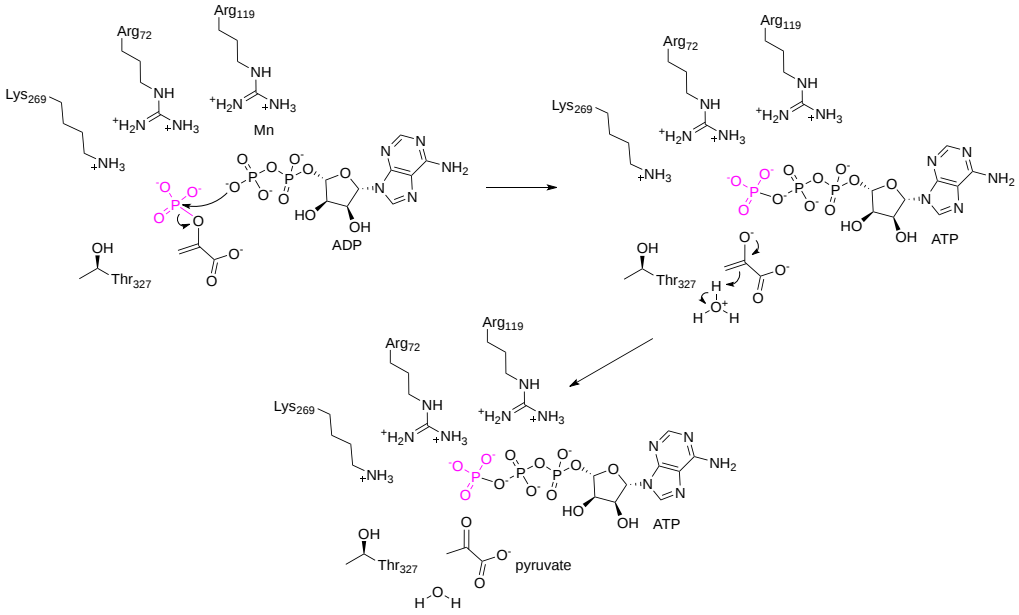


Figure 13.1.47: Mechanism for rabbit pyruvate kinase

Figure 13.1.48 shows an interactive iCn3D model of rabbit muscle pyruvate kinase complexed with Mn^{2+}, K^{+}, and pyruvate (1PKN)

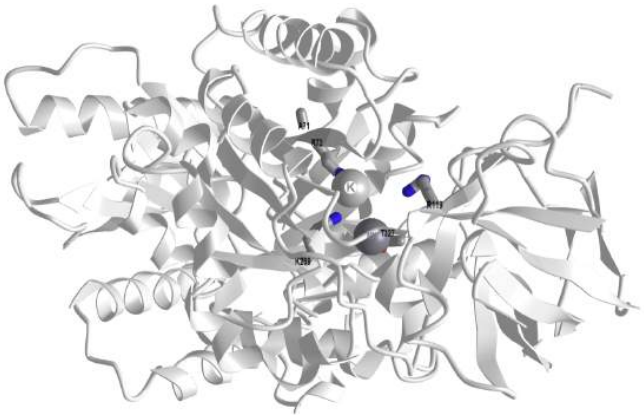


Figure 13.1.48: **Rabbit muscle pyruvate kinase complexed with Mn^{2+}, K^{+}, and pyruvate (1PKN)**. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...cZkm8DdEw2N3C6

We are done! Given that glycolysis is the central anaerobic energy-extracting pathway in all life, it is important that we examined each enzyme in detail.

This is the net reaction of the glycolytic pathway:

$$\mathrm{Glc} + \mathrm{Pi} + 2\,\mathrm{ADP} + 2\,\mathrm{NAD}^{+} \longrightarrow 2\,\mathrm{Pyr} + 2\,\mathrm{ATP} + 2\,\mathrm{NADH} + 2\,\mathrm{H}^{+} + 2\,\mathrm{H_2O}$$

13.1.13: UNCOUPLING GLYCOLYTIC OXIDATION AND PHOSPHORYLATION (ATP FORMATION)

Since we are most interested in energy transduction at this point, let's consider just two important steps in glycolysis that directly lead to ATP synthesis. Only one oxidative step is found in this pathway, namely the oxidative phosphorylation of the 3C glycolytic intermediate glyceraldehyde-3-phosphate, to 1,3-bisphosphoglycerate, a mixed anhydride (see link below for mechanism). The oxidizing agent is NAD^{+} and the phosphorylating agent is NOT ATP but rather Pi. The enzyme is named glyceraldehyde-3-phosphate dehydrogenase. It contains an active site Cys, which helps explain how the enzyme can be inactivated with a stoichiometric amount of iodoacetamide. A general base in the enzyme abstracts an H+ from Cys, which attacks the carbonyl C of the glyceraldehyde, forming a tetrahedral intermediate. Instead of the expected reaction (which would be the protonation of the alkoxide in an overall nucleophilic addition reaction at the aldehyde), a hydride leaves from the former carbonyl C to NAD+ in an oxidation step. Notice, this is a two-electron oxidation reaction similar to that seen in alcohol dehydrogenase. An acyl-thioester intermediate has formed, much like the acyl intermediate that formed in Ser proteases. Next inorganic phosphorous, Pi, attacks the carbonyl C of the intermediate in a nucleophilic substitution reaction to form the mixed anhydride product, 1,3-bisphosphoglycerate. Although we have formed a mixed anhydride, we cleaved a sulfur ester, which is destabilized with respect to its hydrolysis products (since the reactant, the thioester, is not stabilized by resonance to the extent of regular esters owing to the poor donation of electrons from the larger S to the carbonyl-like C.) In the next step, catalyzed by the enzyme phosphoglycerate kinase, ADP acts as a nucleophile that attacks the mixed anhydride of the 1,3-bisphosphoglycerate to form ATP. Note that the enzyme is named for the reverse reaction. We have coupled the oxidation of an organic molecule (glyceraldehyde-3-phosphate) to phosphorylation of ADP through the formation of a "high" energy mixed anhydride, 1,3-bisphosphoglycerate.

The linkage between the oxidation of glyceraldehyde-3-phosphate and the phosphorylation of ADP by 1,3-bisphosphoglycerate can be artificially uncoupled by adding arsenate, which has a similar structure as phosphate. The arsenate can form a mixed anhydride at C1 of glyceraldehyde-3-phosphate, but since the bringing O-As bond is longer and not as strong as in the mixed anhydride, it is easily hydrolyzed. This prevents the subsequent transfer of phosphate to ADP to form ATP.

Figure 13.1.49 shows a summary of oxidation and substrate-level phosphorylation in glyceraldehyde-3-phosphate dehydrogenase.

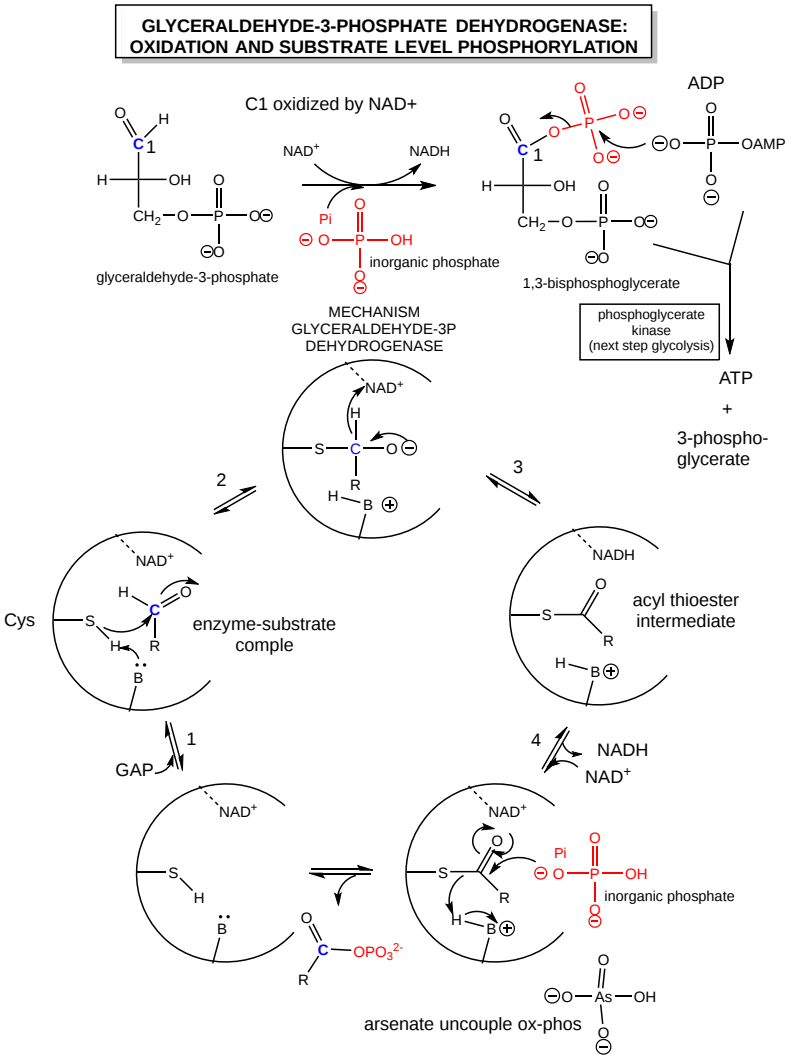


Figure 13.1.49: Oxidation and substrate-level phosphorylation in glyceraldehyde-3-phosphate dehydrogenase (after Voet and Voet)

Summary: Under anaerobic conditions, glucose (6Cs) is metabolized through glycolysis which converts it to two molecules of pyruvate (3Cs). Only one oxidation step has been performed when glyceraldehyde 3-phosphate is oxidized to 1,3-bisphosphoglycerate. To regenerate NAD^+ so glycolysis can continue, pyruvate is reduced to lactate, catalyzed by **lactate dehydrogenase**. These reactions take place in the cytoplasm of cells actively engaged in the anaerobic oxidation of glucose (muscle cells for example during sprints). Note that t50}\).

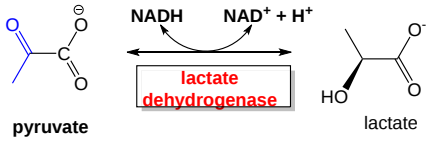


Figure 13.1.50: Conversion of pyruvate to lactate

We will explore the fate of pyruvate under anaerobic conditions more in the next chapter section.

13.2: FATES OF PYRUVATE UNDER ANAEROBIC CONDITIONS- FERMENTATION

13.2.1: INTRODUCTION TO FERMENTATION

Fermentation is a process in which a fuel molecule is broken down anaerobically (or without oxygen) to meet ATP demands. One of the most notable pathways that can utilize fermentation is glycolysis, which we have just described. During glycolysis, glucose is converted to pyruvate, yielding a total of 2 ATP energy-rich molecules. This occurs in anoxic conditions (or without the requirement of oxygen). However, the process of glycolysis cannot be sustained if the end product is pyruvate. This is because the production of pyruvate also yields 2 molecules of reduced NADH in addition to the production of the two ATP molecules. The pool of available NAD^+/NADH is limited within the body, and thus, this electron acceptor/donor molecule must continually be recycled for metabolic pathways to remain functional. For example, if NADH is not oxidized back into NAD^+ in a timely manner, the glycolytic pathway can slow. Therefore, to be able to maintain the process of energy generation through glycolysis, the pool of NADH needs to be converted from the reduced form, back into the oxidized form (NAD^+) form where it can then accept electrons from intermediates in the glycolytic pathway. This process of recycling NADH to NAD^+ can occur aerobically through the Citric Acid Cycle, which is described in more detail in Chapter 16, or it can be processed via anaerobic fermentation (Figure 13.2.1).

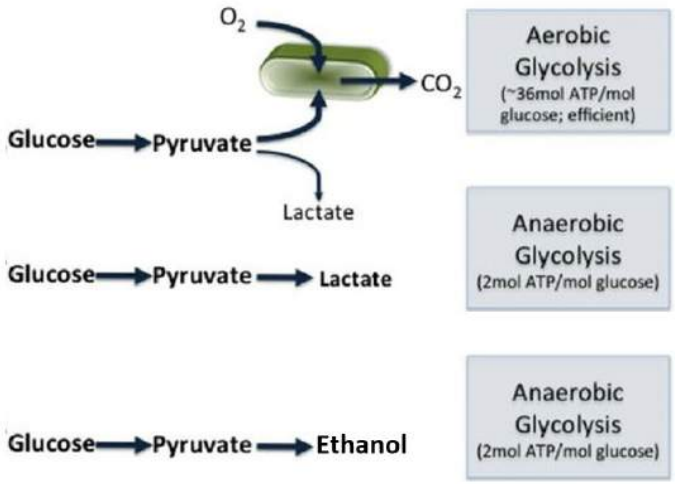


Figure 13.2.1 : Fates of Pyruvate via Aerobic or Anaerobic Pathways.

In the 1860s, Louis Pasteur first described fermentation very narrowly, as the process that yeast use to convert glucose into ethanol when they are grown in the absence of air. Subsequently, it was discovered that other microorganisms could convert pyruvate to lactate (or lactic acid), instead of ethanol, during the process of anaerobic respiration. These microorganisms became known as lactic acid bacteria and are currently utilized heavily within the food industry to produce a wide array of fermented food products including cheese, yogurt, and sauerkraut, to name a few. As the enzymes within the glycolytic pathway were discovered, it became apparent that muscle tissue could also engage in anaerobic respiration-producing lactate. Thus while fermentative processes are most often described in microbial organisms, the definition of fermentation has since been broadened to include any enzymatic, energy-yielding pathways that occur in the absence of oxygen, including the production of lactate in muscle tissue of animals during the glycolytic process.

Thus, NADH recycling to NAD^+ typically occurs in anaerobic systems by two different routes: ethanolic fermentation or lactate fermentation (Figure 13.2.1). We will focus on the details of these two systems here.

13.2.2: LACTATE FERMENTATION

In lactate fermentation, pyruvate is converted to lactate by the enzyme lactate dehydrogenase. In the process, NAD^+ is regenerated. The reaction catalyzed by lactate dehydrogenase (LDH) is shown in Figure 13.2.2.

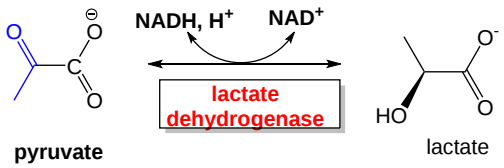


Figure 13.2.2: Reaction catalyzed by lactate dehydrogenase

It is named for the reverse reaction which, as with other dehydrogenases, uses NAD^+ as an oxidizing agent. The reaction is reversible with the ΔG^0 for conversion of pyruvate to lactate of -3.76 kcal/mol (-15.7 kJ/mol). Under anaerobic conditions when glycolysis is the major source of ATP, pyruvate levels increase, further driving the reaction towards lactate formation and NAD^+ regeneration so glycolysis can continue.

The enzyme is found in the cytoplasm but a mitochondrial form also exists. It is most abundant in muscle, liver, kidney, and also in erythrocytes. Interestingly, mature red blood cells are enucleated and do not contain any mitochondria. Their lifespan is limited to approximately two weeks. During this time, their primary energy resources are generated through the process of anaerobic fermentation via the glycolysis-lactate pathway.

The active enzyme is a tetramer of various compositions of two different subunits, the heart (H) and muscle (M) forms. The quaternary structures consist of 5 different isozyme forms containing the H and M subunits. LDH-1, found most abundantly in the heart, is a tetramer of 4H subunits (H_4). The other forms are as follows: LDH-2 (H_3M, prevalent in RBCs), LDL-3 (H_2M_2, prevalent in lungs), LDH-4 (HM_3, prevalent in the kidney), and LDH-5 (M_4, prevalent in muscle).

During anaerobic metabolism, lactate is produced by muscle tissue and released into the bloodstream where it can travel back to the liver. Once in the liver, lactate is converted back into pyruvate and can be utilized to produce glucose through a pathway called gluconeogenesis. The liver can then export the glucose into the blood from where it can be taken up by the muscle for ATP production. This cycle is called the **Cori cycle** and is illustrated in Figure 13.2.3.

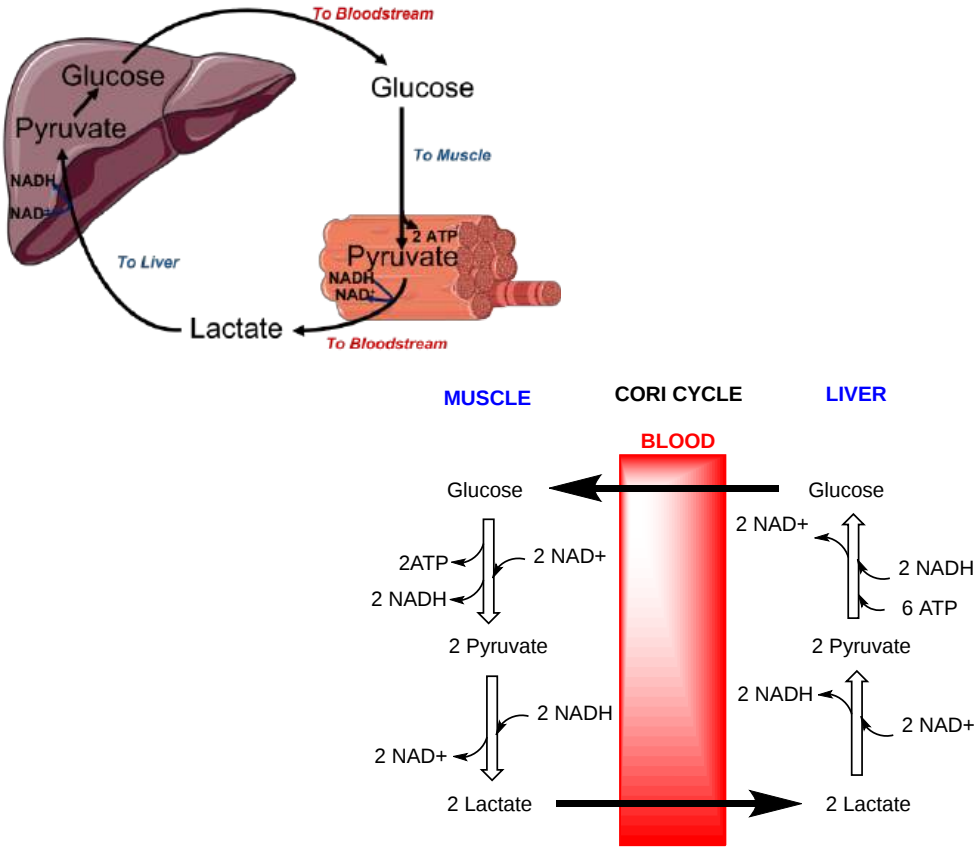


Figure 13.2.3 : Cori cycle *Upper diagram* is a cartoon image of the Cori cycle and the *lower diagram* demonstrates the recycling of NAD^+/NADH in different locations within the body. Figure modified from Servier Medical Art

Notably, anaerobic metabolism can only be sustained for short periods in animals due to its high energy demand. Aerobic respiration is required to maintain an adequate ATP supply. However, the production of lactate by certain tissues, such as white, or fast-twitch muscle cells occurs regularly and releases lactate into the bloodstream, where it can be taken up and used as an energy source by other neighboring tissues such as red, slow-twitch muscle. Brain tissue can also effectively use lactate as an energy source as well. Within these 'consumer' tissues, lactate is converted back into pyruvate using the mitochondrial LDH enzyme where it can then be processed by aerobic respiration producing high levels of ATP.

Thus, the production of lactate can be thought of as a strain response that occurs during times of metabolic stress, such as intense cardiovascular exercise. It has also been noted that after an injury or head trauma, the activation of epinephrine will cause an increase in lactate production and blood levels of lactate will increase. Thus, it has been hypothesized that lactate may play a role in the repair of

damaged tissue. Clinical trials and experiments are currently underway to determine if lactate can help in the healing and recovery process for conditions such as traumatic brain injury, myocardial infarction, and sepsis. Early studies have shown that lactate can increase the production of Brain-Derived Neurotropic Factor (BDNF) which supports neuronal growth, providing further support for the role of lactate in recovery and repair.

Spectroscopy analyses show the geometry and electrostatics of the active sites of LDHs from all organisms are essentially the same. The mechanism for the reverse oxidation of lactate by NAD^+ by spiny dogfish lactate dehydrogenase (LDH) is shown in Figure 13.2.4

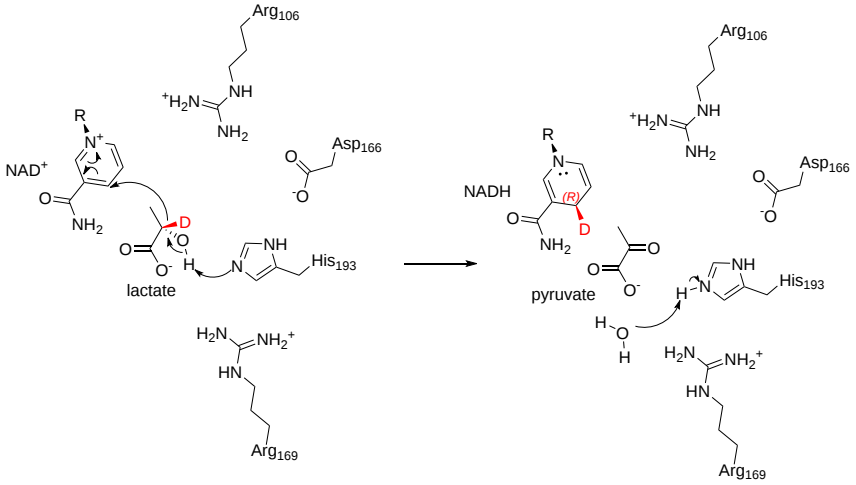
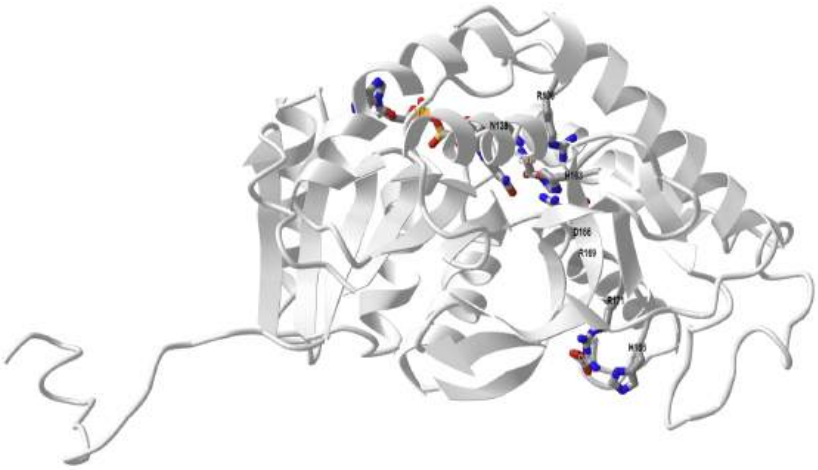


Figure 13.2.4 : Mechanism for the reduction of lactate by NAD^+ by spiny dogfish lactate dehydrogenase (LDH)

Lactate is shown with a deuterium (D), which moves as a deuteride to NAD^+ to form NADH, simply to illustrate the stereochemistry of the reaction. In the reverse reaction, the reduction of pyruvate to lactate (to regenerate more NAD^+ so anaerobic glycolysis can continue), the deuterium (or the proR H of non-deuterated NADH) is removed.

Figure 13.2.5 shows an interactive iCn3D model of dogfish M4 apo-lactate dehydrogenase (1LDM).



NCBI iCn3D **Figure 13.2.5 : Dogfish M4 apo-lactate dehydrogenase (1LDM).**(Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...FBzuqJ2gEPUHM8

Only the monomer of the active tetramer is shown. NAD and oxamic acid, structurally similar to lactate/pyruvate and a competitive inhibitor of the enzyme, are labeled. His193, the active site general base/acid, is near Arg 169 and to Asp 166 and Asn 138, to which it is hydrogen bonded.

LDH exists in two major conformational states, T (inactive) and R (active) state, as we have seen with other allosteric proteins. Figure 13.2.6 shows an interactive iCn3D model of the T state (2ZQY) and R state (2ZQZ) of Lacticaseibacillus casei L-lactate dehydrogenase.

Figure 13.2.6 : T state (2ZQY) and R state (2ZQZ) of Lacticaseibacillus casei L-lactate dehydrogenase. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...zjrtVYjLXdhDp8

Toggle between the magenta T state and cyan R state using the "a" key. Orient the proteins to see the best view of the allosteric changes caused by binding regulators. The spheres show Arg 171 whose disposition changes significantly in the T and R states. Pyruvate (a reactant/product) activates the enzyme (conversion of T to R state) and is considered a **homotropic regulator** since it is a substrate. This is similar to the allosteric "activation" of hemoglobin by its ligand O_2, which preferentially binds to the R state and promotes the T-to-R state transition. Fructose 1,6-bisphosphate, a glycolytic intermediate, also activates the Lacticaseibacillus casei L-lactate dehydrogenase and since it is not a reactant/product of the enzyme, it is considered a heterotropic allosteric regulator. The enzyme hence is effectively regulated by the concentration of substrate (pyruvate) and by NADH levels. High concentrations of ethanol consumption lead to high levels of NADH through the activity of alcohol dehydrogenase. High levels of NADH would lead to increased lactate production as well.

13.2.3: ETHANOL FERMENTATION IN YEAST

In this process, pyruvate is decarboxylated first to acetaldehyde by the thiamine pyrophosphate (TPP)-requiring enzyme **pyruvate decarboxylase**. The resulting product, acetaldehyde, is then reduced by NADH to form ethanol by the enzyme **ethanol dehydrogenase** in a process that reforms NAD^+. Yeast are facultative (not obligate) anaerobes in that they can produce energy by glycolysis and ethanol fermentation in the absence of oxygen Figure 13.2.7:. Of course, in the presence of oxygen, the pyruvate produced from glycolysis in yeast is preferentially converted to acetyl-CoA which enters the citric acid cycle and oxidative phosphorylation pathways to maximize ATP production.

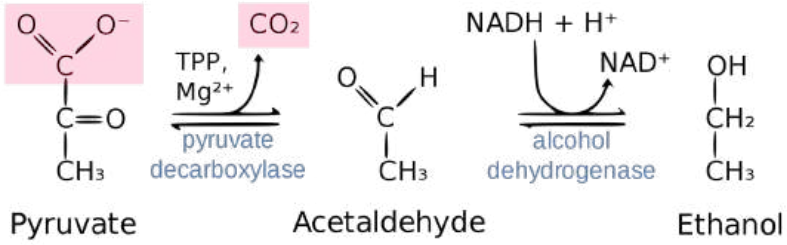


Figure 13.2.7 : Summary of Ethanol Fermentation in Yeast

Figure from Thomas Baldwin

Pyruvate decarboxylase

Pyruvate decarboxylase catalyzes the first step in the alcoholic fermentation pathway converting pyruvate into acetaldehyde and carbon dioxide. Interestingly, in addition to yeast, some species of fish, such as goldfish and carp, have a homologous enzyme that allows for the production of ethanol when oxygen is scarce. Pyruvate decarboxylase is tetrameric occurring as a dimer of dimers with two active sites shared between the monomer subunits of each dimer. The decarboxylation reaction requires two cofactors, thiamine pyrophosphate (TPP) and magnesium (Figure 13.2.8). Within the active site, the acidic Glu-477 and Glu-51 residues and the Mg^{2+} cofactor interact with and stabilize the TPP cofactor. The aminopyrimidine ring on TPP acts as a base and enables the formation of the TPP nucleophile with the removal of the C2 proton. This reaction is stabilized by the protonation of Glu-51. The nucleophilic attack of pyruvate causes the release of carbon dioxide. The acetaldehyde intermediate is still covalently attached to the TPP cofactor following the release of carbon dioxide.

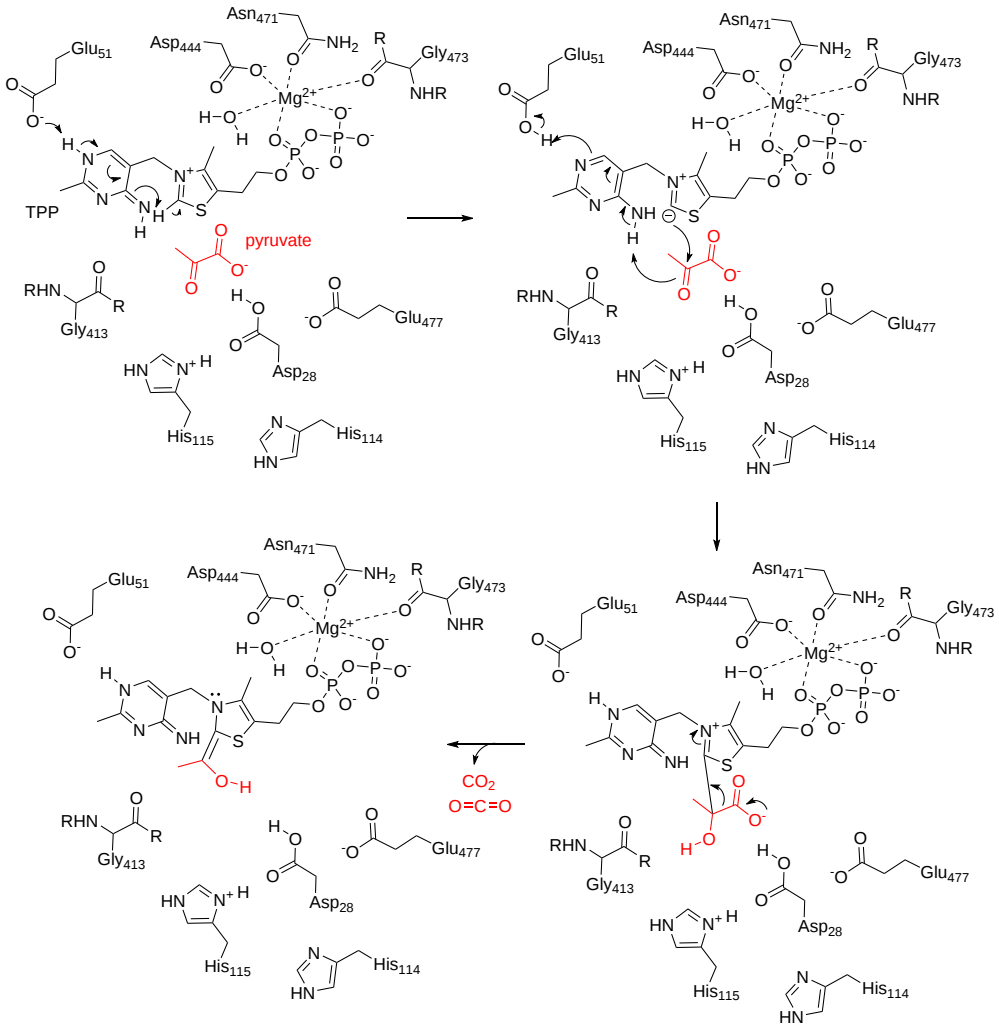


Figure 13.2.8 : Part 1 Pyruvate Decarboxylase mechanism - decarboxylation

Once carbon dioxide diffuses away from the active site, the double bond of the enol-intermediate abstracts a proton from Asp-28, which is stabilized by a neighboring His-115 residue. TPP can then serve as a good leaving group during the formation of the carbonyl functional group and causes the release of acetaldehyde (Figure 13.2.9:).

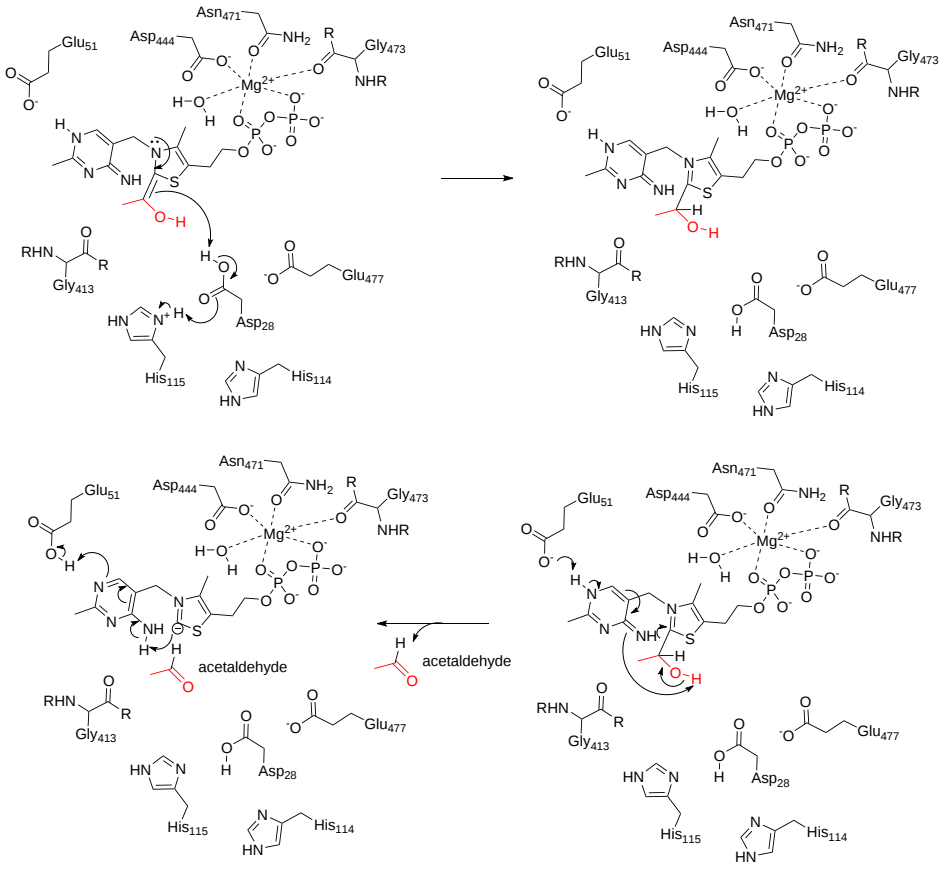


Figure 13.2.9 : Part 2 Pyruvate Decarboxylase mechanism - acetaldehyde generation

Figure 13.2.10 shows an interactive iCn3D model of the thiamin diphosphate-dependent enzyme pyruvate decarboxylase from the yeast Saccharomyces cerevisiae (1PVD).

NCBI iCn3D **Figure 13.2.10 : Thiamin diphosphate-dependent enzyme pyruvate decarboxylase from the yeast Saccharomyces cerevisiae (1PVD).**(Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...DMM4gjodX4q1G9 TTP is shown in spacefill. The enzyme is a homotetramer.

Alcohol dehydrogenase (ADH)

Alcohol dehydrogenase enzymes catalyze the interconversion of aldehydes or ketones with alcohol functional groups. In humans and other animals, they are utilized to break down alcohols that are ingested, as well as participate in the biosynthesis of numerous metabolites. There are at least 5 classes of alcohol dehydrogenases that use NAD^+ as an oxidizing agent to convert alcohols to aldehydes or ketones. We are

mainly concerned with the yeast ADH, a member of Family I, that is involved in ethanol fermentation. Figure 13.2.11 shows the reaction catalyzed by alcohol dehydrogenase.

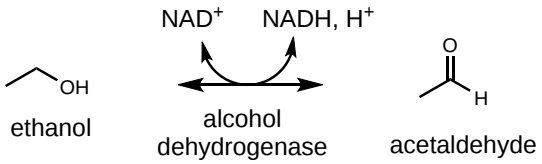


Figure 13.2.11 : Reaction catalyzed by alcohol dehydrogenase

Figure 13.2.12A below shows a mechanism for yeast alcohol dehydrogenase (ADH1). ADH1 requires zinc as a metal cofactor and utilizes NADH as the reducing agent in the conversion of acetaldehyde into ethanol.

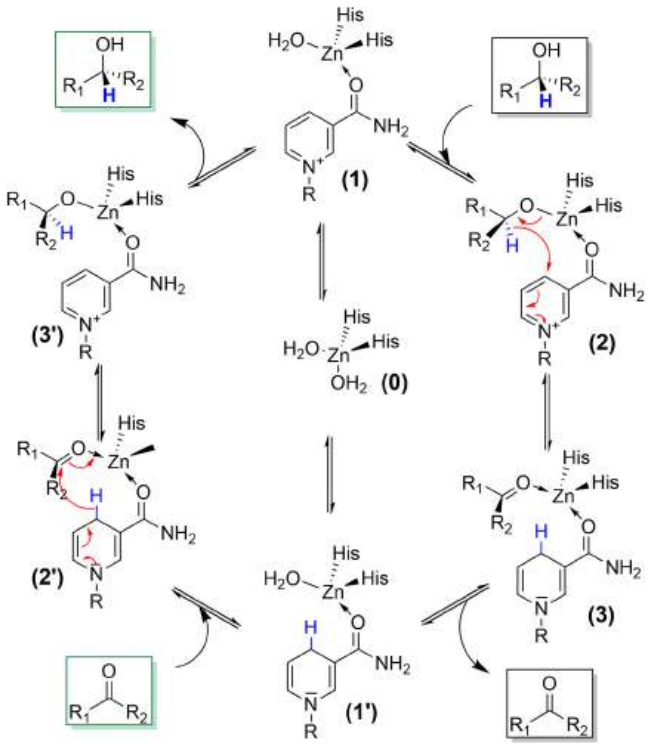


Figure 13.2.12A : Simplified catalytic mechanism of ADH reactions. Upon binding of both substrates (e.g., NAD(P)$^+$ and alcohol; 0->1->2) a hydride transfer occurs from the alcohol-carbon atom to the oxidized nicotinamide moiety yielding the Zn-coordinated carbonyl product and NAD(P)H (3). Both can dissociate from the active site yielding apo-ADH (0). Alternatively, only NAD(P)H stays bound and the reduced ADH can undergo a reductive conversion (1'->2'-> 3'-> 1). Amanda Silva de Miranda, et al. Front. Catal., 10 May 2022. Sec. Biocatalysis. https://doi.org/10.3389/fctls.2022.900554. Creative Commons Attribution License (CC BY).

Figure 13.2.12B shows key amino acids in the active site of yeast ADH1. Note that in this figure, two cysteines are the coordinating ligands.

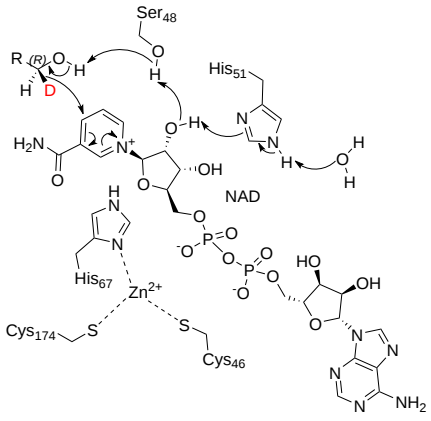


Figure 13.2.12B : Abbreviated mechanism for yeast alcohol dehydrogenase (ADH1)

If you look at the reverse reduction reaction, the acetaldehyde (or a ketone for 2^0 alcohols) carbonyl is sp^2 hybridized and planar, so the hydride could be added to either of the two faces, re or si, of the plane (using the same rules used to define R or S enantiomers). Likewise, the methylene carbon atom with two hydrogens, one of which is transferred as a hydride, is prochiral. Hence four different stereochemical hydride transfer pathways are possible, as shown in **Figure 13.2.13A** below. Hydride transfers from the *si*-face of the prochiral aldehyde/ketone result in (*R*)-configured alcohols whereas hydride attacks from the *re*-face yield (*S*)-alcohols. In both cases, the hydride transferred can stem from either the *re*- or *si*-face of the nicotinamide ring.

Figure 13.2.13A : Possible stereochemical courses of the hydride transfer from NAD(P)H to the ketone. Attacks from the *si*-face of the ketone result in (*R*)-alcohols (E1 and E2) whereas hydride attacks from the *re*-face (E3 and E4) result in (*S*)-alcohols.

ADHs catalyzing hydride addition from the *re*-face of the ketone (or abstraction of a hydride from (*S*)-alcohols) are termed Prelog-selective ADHs whereas those ADHs attacking from the *si*-face (or abstracting a hydride from the (*R*)-alcohol) are termed *anti*-Prelog ADHs. Different ADHs are available for the different reaction stereochemistries.

The stereochemistry for the reaction of ADH1 with ethanol is shown in **Figure 13.2.13B** below**.** Deuterium (**D**) is shown in the figure to better illustrate the stereochemistry of the reaction. With the **deuterium** label, the carbon is now chiral and the enantiomer shown is the R isomer. The corresponding **H** in the undeuterated and prochiral form of ethanol is removed by the enzyme, as shown in Figure 13.2.13B below.

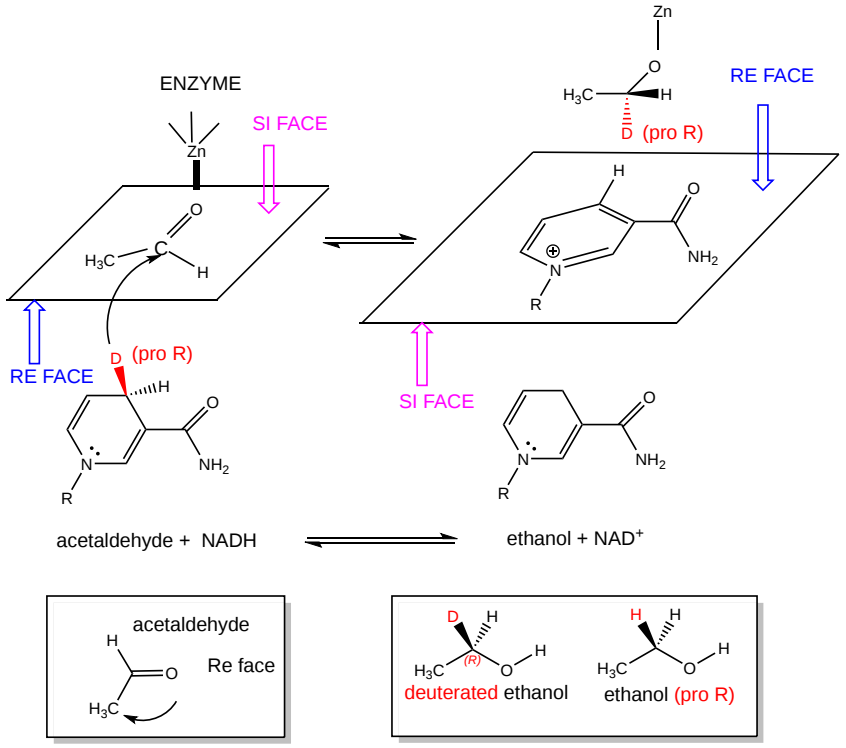


Figure 13.2.13*B* : Stereochemistry of human alcohol dehydrogenase reaction.

Figure 13.2.14 shows an interactive iCn3D model of the Yeast alcohol dehydrogenase (ADH I) with bound substrate analogs- ADH1 (4W6Z).

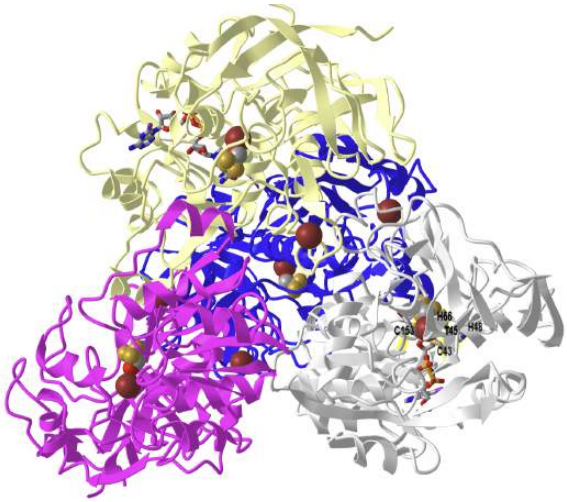


Figure 13.2.14 : Yeast alcohol dehydrogenase structure with bound substrate analogs (4W6Z). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...feE7toF6wU6MB9

All four monomers of the homotetramer are shown. The light gray subunit has the key catalytic and binding residues shown in sticks and labeled. The numbering for the key residues is a bit different than in the mechanism shown above. The cofactor analog, nicotinamide-8-iodo-adenine-dinucleotide, is shown in sticks and the substrate analog, trifluoroethanol is shown in spacefill.

Only two subunits (a homodimer) have both bound cofactor and substrate. The substrate trifluoroethanol in these subunits is ligated to the catalytic Zn^{2+} through its oxygen, with the other ligands provided by the side chains of Cys 43, Cys 153, and His 66. The other two monomers (the other homodimer) have a different conformation and use Glu 67 to coordinate the Zn^{2+} instead of His 66 and no substrate is bound. This may be an intermediate in the process that displaces water bound to Zn^{2+} with the substrate.

13.2.4: ALCOHOL METABOLISM IN HUMANS:

When ethanol is consumed, it is oxidatively metabolized primarily in the liver. Cytosolic Alcohol Dehydrogenase (ADH) and mitochondrial Acetaldehyde Dehydrogenase 2 (ALDH2) are the main enzymes involved in this metabolic pathway, first converting ethanol to acetaldehyde and then acetaldehyde to acetate (Figure 13.2.15:). Liver mitochondria have a limited capacity to use the acetate in the Kreb cycle because the enzyme needed to convert acetate to acetyl-CoA (acetyl-CoA synthase 2) is almost absent in the liver, but is abundant in the heart and skeletal muscles. Thus, most of the acetate resulting from ethanol metabolism escapes the liver into the blood circulation and is eventually metabolized to CO_2 by way of the Kreb cycle in cells with mitochondria that contain enzymes to convert acetate to acetyl CoA, such as heart, skeletal muscle, and brain.

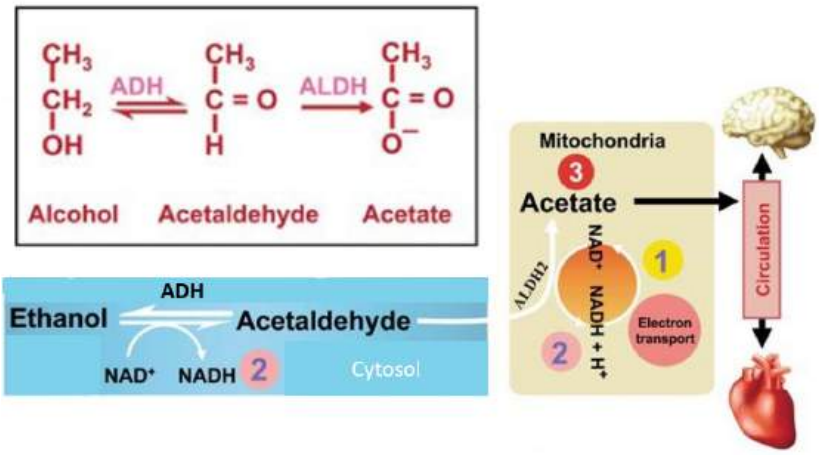


Figure 13.2.15 : Summary of Ethanol Metabolism in Humans

Figure modified from Zakhari, S. (2013) Alcohol Research: Current Reviews 35(1):6-16

Like many substances, the consumption of alcohol can be both beneficial and detrimental, depending on the quantity and frequency of consumption. This is known as a hormetic response (Figure 13.2.16:). At low doses (up to 2 daily drinks for men and 1 daily drink for women) the consumption of alcohol can be cardioprotective. At higher consumption rates this protective effect is lost and the detrimental effects of alcohol consumption become apparent and include cardiotoxicity, liver damage, and increased cancer risk. Not to mention the debilitation that can accompany addiction.

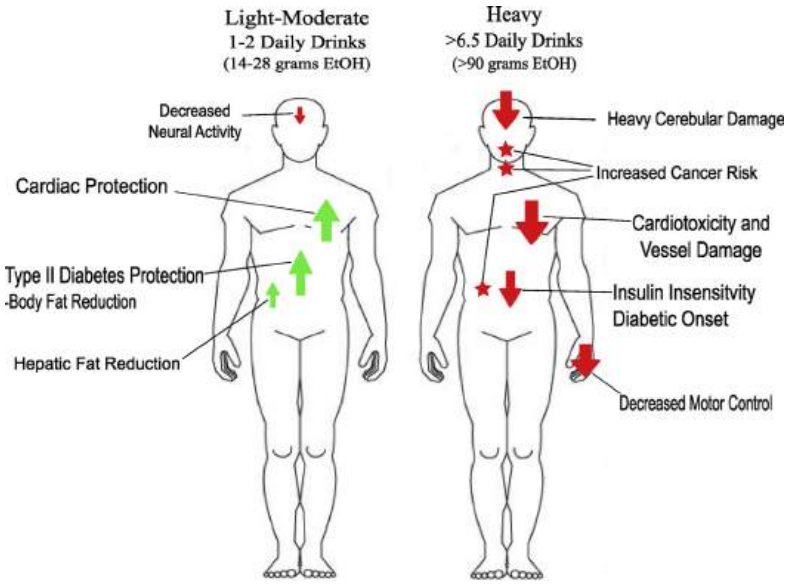


Figure 13.2.16 : Alcohol Consumption and Hormesis

Figure from Adamson, S.S., et al (2017) Translational Medicine of Aging 1:18-23.

The graph presented in Figure 13.2.17 exemplifies the hormetic nature of alcohol consumption. Maintaining ethanol consumption between one (for women) and two (for men) daily drinks will significantly reduce the overall risk of mortality. However, mortality risk skyrockets at high ethanol consumption.

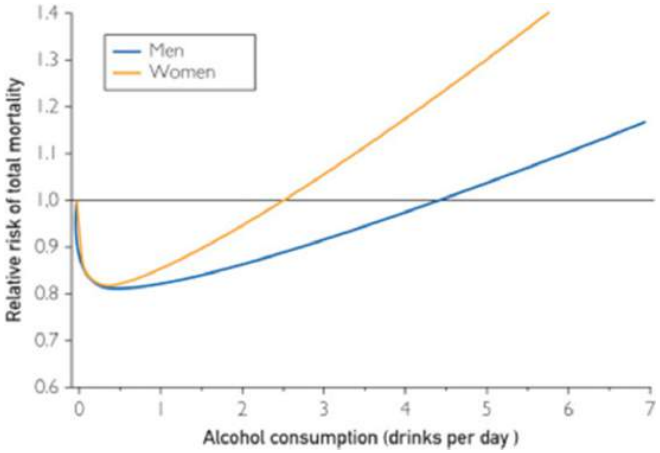


Figure 13.2.17 : Risk of Mortality and Alcohol Consumption

Figure from Adamson, S.S., et al (2017) Translational Medicine of Aging 1:18-23.

Moderate alcohol use of no more than 2 drinks per day for men, and no more than 1 drink per day for women can reduce the risk of coronary artery disease by 29%, and reduce the risk of dying from any cardiovascular disease by 25%. These statistics are significant. Do note that at these levels that there is an increased risk of dying from hemorrhagic stroke. Increased alcohol consumption, however, can have dire effects on health and can lead to unwanted addiction. The Centers for Disease Control estimates that alcohol abuse leads to approximately 80,000 deaths annually in the United States and that up to 40% of deaths related to liver disease are caused by alcohol abuse.

Harvard Medical School has evaluated thousands of studies and has tried to come up with a method for alcohol risk assessment. They categorize low-risk drinking behavior for men as no more than four drinks in a single day or 14 drinks in a week and no more than 3 drinks in a single day or a total of 7 drinks in a single week for women. Women have lower drinking tolerance due to their smaller sizes, but also due to metabolic differences. Women have less ADH and typically higher fat levels which disperse and retain ethanol longer. Addiction and damage to the liver tend to occur more quickly in women as well. Harvard ranks people at increased risk if their drinking is above either the single day or the weekly limit (which is estimated at 29% of drinkers). High-risk drinkers break both the daily and weekly limits (estimated at 14% of drinkers).

During ethanol metabolism, when circulating ethanol is in the millimolar range, acetaldehyde is in the micromolar range, and acetate is in the millimolar range (Figure 13.2.18:). When heavy drinking or chronic drinking occurs the Acetaldehyde Dehydrogenase enzyme cannot keep up with the Alcohol Dehydrogenase enzyme and the pool of Acetaldehyde increases. This aldehyde has many toxic effects within biological systems (Figure 13.2.19:).

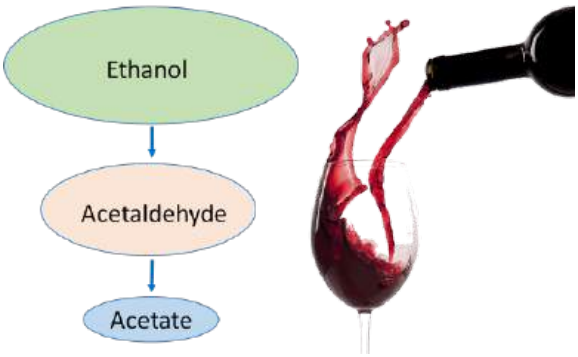


Figure 13.2.18 : The Effects of Binge or Chronic Drinking on Ethanol Metabolites.

Figure modified from Zakhari, S. (2013) Alcohol Research: Current Reviews 35(1):6-16

Acetaldehyde can form adducts with proteins. For example, acetaldehyde adducts on cytoskeletal components such as microtubulin, lead to the swelling of hepatocytes (liver cells). If secreted from the cell, these protein adducts can also be recognized as foreign by the immune system and cause an autoimmune response causing further inflammation and damage to the liver. Acetaldehyde also causes oxidative damage to lipids and DNA and can alter mitochondrial function. Overall, the liver is stressed and unhappy when too much ethanol is consumed.

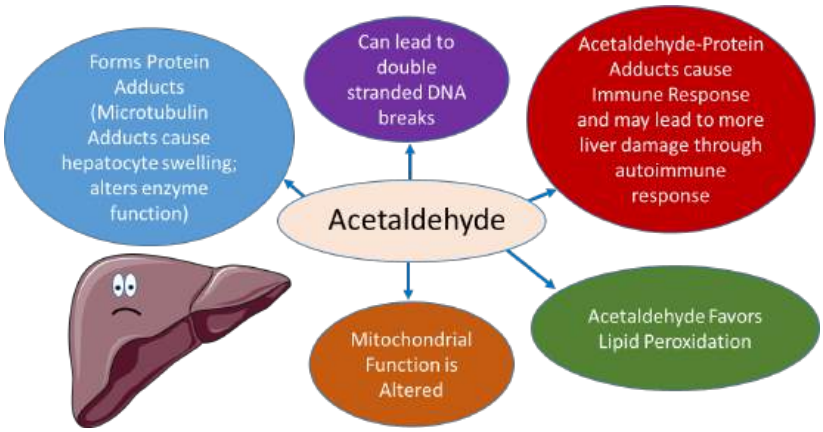


Figure 13.2.19 : Acetaldehyde Toxicity

Figure modified from Servier Medical Art

Heavy chronic drinking can also lead to epigenetic modifications that alter protein expression patterns within the cell. Due to the oxidation of alcohol to acetate, the metabolism of alcohol also leads to an increase in the NADH: NAD+ ratio altering carbohydrate metabolism. Heavy ethanol consumption promotes the formation of reactive oxygen species and can also promote apoptosis. Overall, heavy drinking is extremely hard on liver function. It can lead progressively from the formation of a fatty liver to liver cirrhosis and increased risk of liver cancer. The risk for several other types of cancer is also heightened with heavy alcohol use.

Chronic heavy alcohol use can also alter gene expression, especially in the liver. Figure 13.2.20 is a schematic representation of DNA methylation, which converts cytosine to 5′methyl-cytosine via the actions of DNA methyltransferase (DNMT). DNA methylation typically occurs at cytosines that are followed by a guanine (i.e., CpG motifs). Within the liver, chronic heavy drinking reduces pools of S-adenosylmethionine (SAM) while increasing homocysteine and S-adenosylhomocysteine (SAH). SAH further inhibits DNA methyltransferases (DNMTs) by negative feedback inhibition, ultimately resulting in global hypomethylation of DNA. This hypomethylation leads to the inappropriate expression of many genes, especially within the liver tissue.

Figure 13.2.20 : Effects of Chronic Alcohol Use on DNA Methylation

Figure modified from Zakhari, S. (2013) Alcohol Research: Current Reviews 35(1):6-16

Two genes that are upregulated due to this hypomethylation express the Catalase enzyme and the p450 oxidoreductase enzyme, CYP2E1. Both of these enzymes are involved in oxidative pathways of alcohol metabolism that produce the toxic acetaldehyde intermediate (Figure 13.2.21). Expression of both of these proteins become more prevalent in chronic alcohol consumption or when blood alcohol levels are high, as in cases of binge or heavy drinking. The activity of these enzymes can also lead to the formation of reactive oxygen species that contribute to global cellular damage (lipid peroxidation, DNA damage, protein damage, etc).

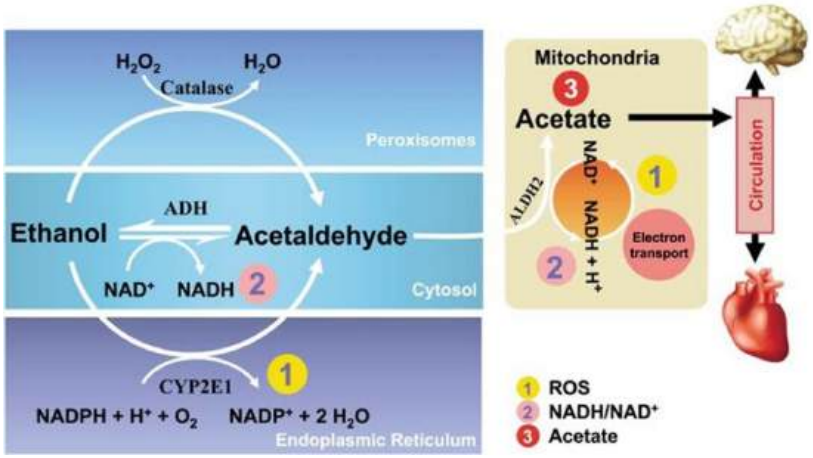


Figure 13.2.21 : Oxidative Pathways of Alcohol Metabolism

Figure from Zakhari, S. (2013) Alcohol Research: Current Reviews 35(1):6-16

Gene regulation in other areas of the body is also affected in response to chronic heavy alcohol consumption. This is due to the production of acetate during the metabolic pathway of alcohol that is released from the liver into the bloodstream. In other areas of the body, acetate is converted to acetyl-CoA by the enzyme Acetyl-coenzyme A (acetyl-CoA) synthetase (AceCS). AceCS is activated by Sirtuin 1, also known as NAD-dependent protein deacetylase (SIRT1). Acetyl-CoA is used by histone acetyltransferase (HAT) to acetylate the lysine residues in histone proteins. Histone acetylation causes these proteins to release the bound DNA, allowing regions to be opened up for transcription (Figure 13.2.22). Thus, higher levels of acetate promote histone acetylation and increased gene expression.

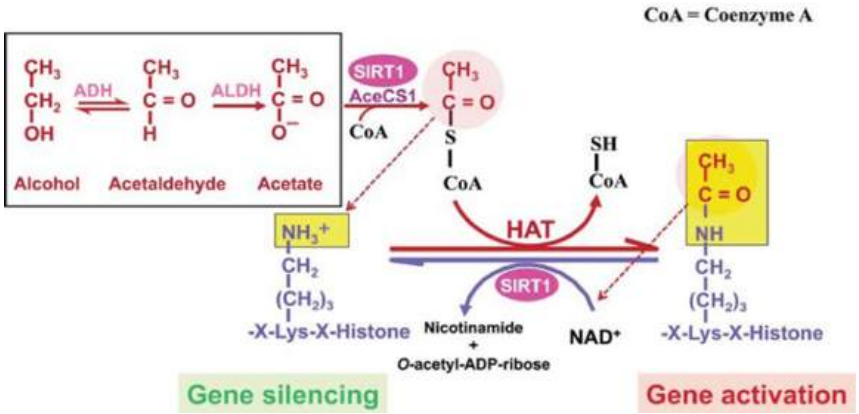


Figure 13.2.22 : The Effects of Chronic Alcohol Consumption on Histone Acetylation. Note in the diagram shown, that SIRT1 also deacetylates histones, resulting in gene silencing. Thus, SIRT1 is a sensor that balances gene activation and silencing in the cell based on the cell's energy status. Alcohol metabolism results in acetate formation, which is used in ***extrahepatic tissues*** to produce acetyl-CoA, upregulating histone acetylation within those tissues. NOTES: AceCS1 = Acetyl-CoA synthase 1; ADH = alcohol dehydrogenase; ALDH = Aldehyde dehydrogenase. Figure from Zakhari, S. (2013) Alcohol Research: Current Reviews 35(1):6-16

In addition to these effects, both ADH and ALDH utilize the cofactor nicotinamide adenine dinucleotide (NAD+), which is reduced to NADH; as a consequence, during ethanol oxidation the ratio NADH/NAD+ is significantly increased, altering the cellular redox state and triggering several adverse effects, related to alcohol consumption. Glycolysis and the Kreb Cycle are downregulated due to low NAD+ levels. This results in lower pyruvate levels, and lower conversion of pyruvate to acetyl-CoA and also causes a decrease in gluconeogenesis (ie there is not enough pyruvate to drive glucose production). Thus, the pyruvate that does form favors anaerobic conversion to lactate and can result in lactic acidosis or lowering of the blood pH levels. Low rates of gluconeogenesis can also contribute to hypoglycemia which can be seen during binge drinking.

Hyperuricemia or an increase in blood levels of uric acid can also occur due, in part, to increased production of ketone bodies and lactic acid. Both the ketone bodies and lactate can compete with uric acid for excretion into the urine within the kidney. The uric acid gets retained and heightens blood levels.

Ketone bodies typically form during periods of starvation when carbohydrate stores have been depleted. When the liver can no longer efficiently maintain blood glucose levels, it will break down fatty acids into ketone bodies and secrete these into the bloodstream (Figure

13.2.23). Ketone bodies, such as acetoacetate, acetone, and D-beta-hydroxybutyrate (which isn't a real ketone, but is still referred to as a ketone body) are released into the bloodstream to compensate for the reduced glucose levels. Brain, heart, and skeletal muscle tissue can utilize ketone bodies as an energy source and this is a good short term solution to starvation. However, the formation of lactate and ketone bodies can severely reduce blood pH levels and induce a life-threatening state known as ketoacidosis. In addition to starvation, heavy alcohol consumption can induce ketogenesis inappropriately.

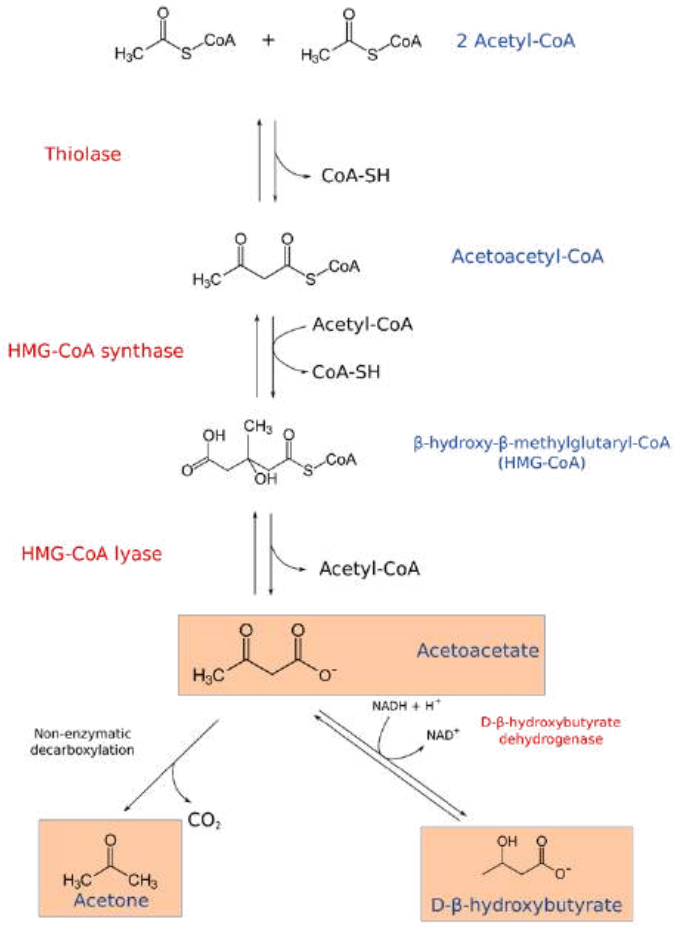


Figure 13.2.23 : Formation of Ketone Bodies. Figure from Sav vas

As noted before, high levels of lactate and ketone bodies within the bloodstream can result in dehydration and reduced excretion of uric acid, leading to hyperuricemia. Over time, high uric acid levels can cause uric acid to precipitate, especially in joints where it causes painful gout flare-ups (Figure 13.2.24).

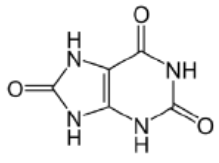


Figure 13.2.24 : Hyperuricemia and Gout. Image from www.scientificanimations.com

According to the American Cancer Society, excessive alcohol use can also increase the risk for several different forms of cancer.

- **Cancers of the mouth, throat, voice box, and esophagus:** Alcohol use raises the risk of these cancers. Drinking and smoking together raise the risk of these cancers even more than drinking or smoking alone. This might be because alcohol can help harmful chemicals in tobacco get inside the cells that line the mouth, throat, and esophagus. Alcohol may also limit how these cells can repair damage to their DNA caused by the chemicals in tobacco.
- **Liver cancer:** Long-term alcohol use has been linked to an increased risk of liver cancer. Regular, heavy alcohol use can damage the liver, leading to inflammation and scarring. This might raise the risk of liver cancer.
- **Colon and rectal cancer:** Alcohol use has been linked with a higher risk of cancers of the colon and rectum. The evidence for this is generally stronger in men than in women, but studies have found the link in both sexes.
- **Breast cancer:** Even a few drinks a week are linked with an increased risk of breast cancer in women. This risk may be especially high in women who do not get enough folate (Vitamin B12) in their diet or through supplements. Alcohol can also raise estrogen levels in the body, which may explain some of the increased risks. Cutting back on alcohol may be an important way for many women to lower their risk of breast cancer.

So how does alcohol consumption contribute to increased risks of cancer? Alcohol may help other harmful chemicals, such as those in tobacco smoke, enter the cells lining the upper digestive tract more easily. This might explain why the combination of smoking and drinking is much more likely to cause cancers in the mouth or throat than smoking or drinking alone. In other cases, alcohol may slow the body's ability to break down and get rid of some harmful chemicals. Alcohol might affect the body's ability to absorb some nutrients, such as folate. Folate is a vitamin needed as a cofactor for enzymes involved in amino acid biosynthesis. Absorption of nutrients can be even worse in heavy drinkers, who often have low levels of folate. These low levels may play a role in the risk of some cancers, such as breast and colorectal cancer. Alcohol can raise the levels of estrogen, a hormone important in the growth and development of breast tissue. This could affect a woman's risk of breast cancer. Too much alcohol can also add extra calories to the diet, which can contribute to weight gain in some people. Being overweight or obese is known to increase the risks of many types of cancer.

Overall, heavy ethanol consumption produces a wide spectrum of hepatic lesions (Figure 13.2.25). Fatty liver (i.e., steatosis) is the earliest, most common response that develops in more than 90 percent of problem drinkers who consume 4 to 5 standard drinks per day. With continued drinking, alcoholic liver disease can proceed to liver inflammation (i.e., steatohepatitis), fibrosis, cirrhosis, and even liver cancer (i.e., hepatocellular carcinoma). Heavy drinking can also damage other organs, such as the pancreas and the brain, and can raise blood pressure. It also increases the risk of heart disease and stroke. In pregnant women, alcohol use, especially heavy drinking, may lead to birth defects or other problems with the fetus.

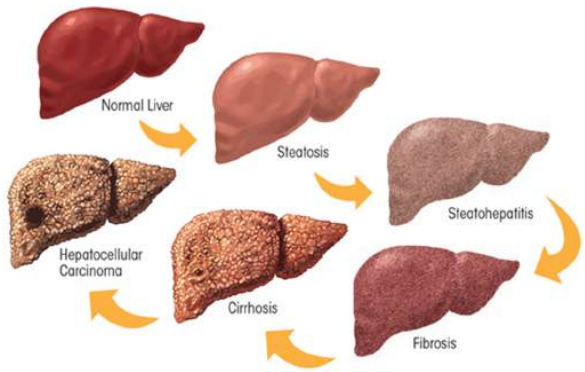


Figure 13.2.25 : The Effects of Heavy Chronic Drinking on the Liver

Figure from Osna, N.A. et al (2017) Alcohol Research: Current Reviews 38(2)

In 1951, the FDA approved disulfiram for the treatment of alcoholism in the US (Figure 13.2.26). Initially, the drug was prescribed in very high doses, often as high as 3,000 mg per day. The high doses led to reports of severe reactions to alcohol, some of which were fatal. It is an inhibitor of the ALDH-2, which will lead to an even higher increase in acetaldehyde concentration if alcohol is consumed. Thus, alcoholics that are taking this drug to quit drinking need to maintain sobriety. If they drink while taking this drug, they will become very ill due to the accumulation of acetaldehyde. Overall, the use of this drug supports abstinence, which is often hard to realistically achieve.

Figure 13.2.26 : Disulfiram

At one time it was thought that before prescribing disulfiram to people, patients should experience mixing the drug with alcohol in a supervised setting. Researchers felt it was important for the individuals to have full knowledge of what would happen if they mixed disulfiram and alcohol. This practice is no longer used, but every person must be educated on the reactions of combining alcohol with disulfiram before a prescription is written.

Along with these effects, alcohol may contribute to cancer growth in other, unknown ways.

and other mammals, various ADHs are used to oxidatively metabolize ethanol to acetaldehyde (also toxic) which is converted to acetate by the enzyme aldehyde dehydrogenase which catalyzes the following reaction:

$$\mathbf{CH_3CHO + NAD^+ + H_2O \rightarrow acetate + NADH + H^+}$$

They can also convert methanol to the very toxic and reactive formaldehyde, which makes methanol poisoning so dangerous if not fatal. Mammalian enzymes are dimers of up to nine different monomers.

13.3: GLUCONEOGENESIS

13.3.1: INTRODUCTION

Gluconeogenesis is a metabolic pathway that results in the generation of glucose from non-carbohydrate carbon substrates such as lactate, glycerol, and glucogenic amino acids. It is one of the two main mechanisms humans and many other animals use to keep blood glucose levels from dropping too low (hypoglycemia). The other means of maintaining blood glucose levels is through the degradation of glycogen (glycogenolysis). Gluconeogenesis is a ubiquitous process, present in plants, animals, fungi, bacteria, and other microorganisms. In animals, gluconeogenesis takes place mainly in the liver and, to a lesser extent, in the cortex of the kidneys. This process occurs during periods of fasting, starvation, low-carbohydrate diets, or intense exercise and is highly endergonic. For example, the pathway leading from phosphoenolpyruvate to glucose-6-phosphate requires 6 molecules of ATP. Gluconeogenesis is often associated with ketosis. Gluconeogenesis is also a target of therapy for type II diabetes, such as metformin, which inhibits glucose formation and stimulates glucose uptake by cells.

Lactate is transported back to the liver where it is converted into pyruvate by the Cori cycle using the enzyme lactate dehydrogenase. Pyruvate, the first designated substrate of the gluconeogenic pathway, can then be used to generate glucose (Figure 13.3.1). All citric acid cycle intermediates, through conversion to oxaloacetate, amino acids other than lysine or leucine, and glycerol can also function as substrates for gluconeogenesis. Transamination or deamination of amino acids facilitates the entry of their carbon skeleton into the cycle directly (as pyruvate or oxaloacetate), or indirectly via the citric acid cycle. Glycerol, which is a part of the triacylglycerol molecule, can be used in gluconeogenesis.

Gluconeogenesis is a pathway consisting of eleven enzyme-catalyzed reactions. The pathway can begin in the mitochondria or cytoplasm, depending on the substrate being used. Many of the reactions are the reversible steps found in glycolysis (Figure 13.3.1). In humans, gluconeogenesis is restricted to the liver and a lesser extent the kidney.

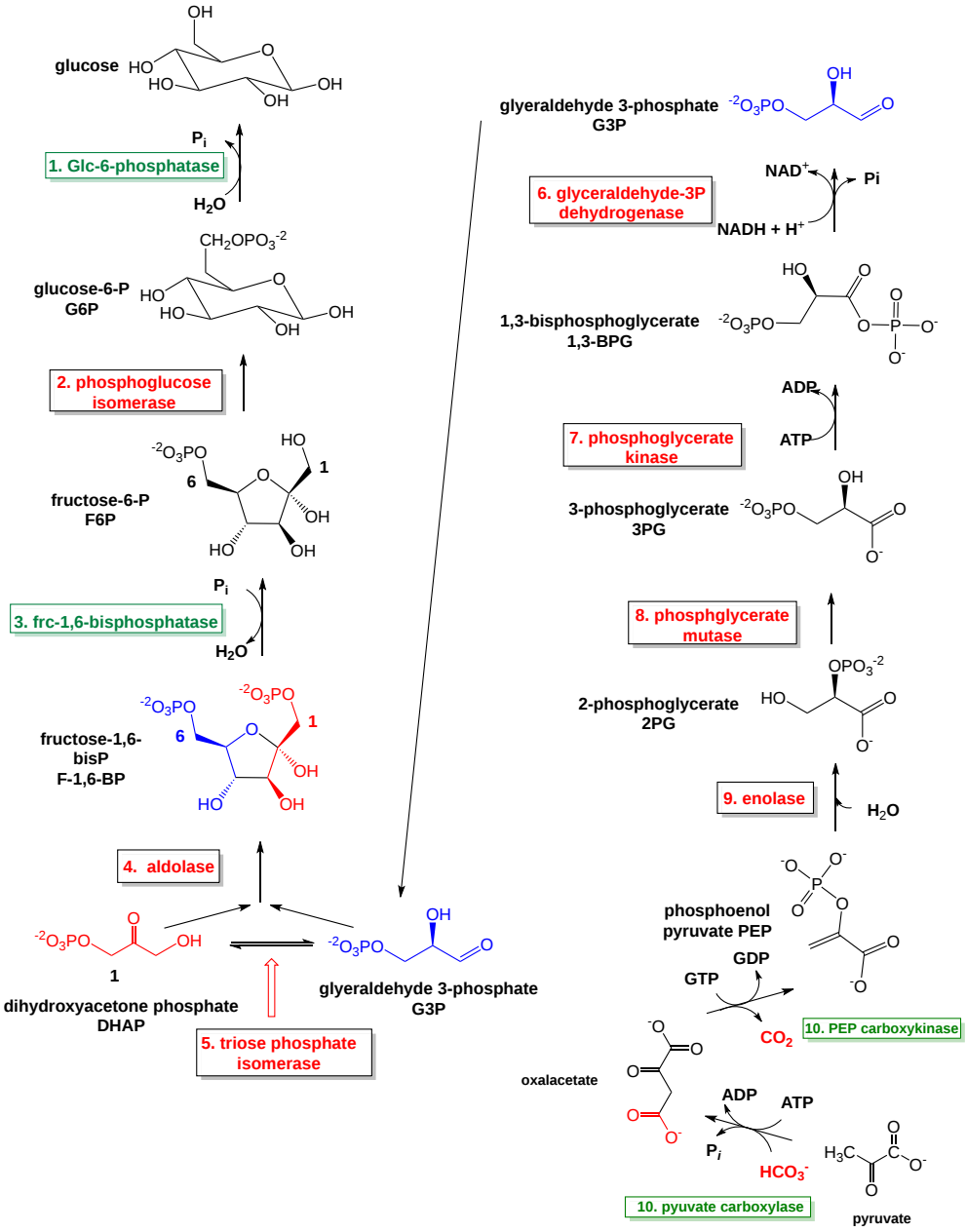


Figure13.3.1 : Gluconeognesis

13.3.2: BYPASS I (REVERSE OF STEP 10 IN GLYCOLYSIS): PYRUVATE TO PHOSPHOENOLPYRUVATE

The conversion of pyruvate into phosphoenolpyruvate requires two enzymatic steps and the formation of oxaloacetate as the intermediate. In all species, the formation of oxaloacetate from pyruvate and any other TCA cycle intermediates is restricted to the mitochondrion, and the enzymes that convert PEP to glucose are found in the cytosol. The location of the enzyme that links these two parts of gluconeogenesis by converting oxaloacetate to PEP, PEP carboxykinase, is variable by species: it can be found entirely within the mitochondria, entirely within the cytosol, or dispersed evenly between the two, as it is in humans. Transport of PEP across the mitochondrial membrane is accomplished by dedicated transport proteins; however, no such proteins exist for oxaloacetate. Therefore in species that lack intra-mitochondrial PEP, oxaloacetate must be converted into malate or aspartate, exported from the mitochondrion, and converted back into oxaloacetate to allow gluconeogenesis to continue.

Here is the net reaction: Pyruvate + ATP + HCO_3^- + GTP → PEP + ADP + GDP + CO_2 $\Delta G^{o'}$ = +0.2 kcal/mol (0.8 kJ/mol)

Figure13.3.2 shows the overall reaction for the conversion of pyruvate to phosphoenol pyruvate.

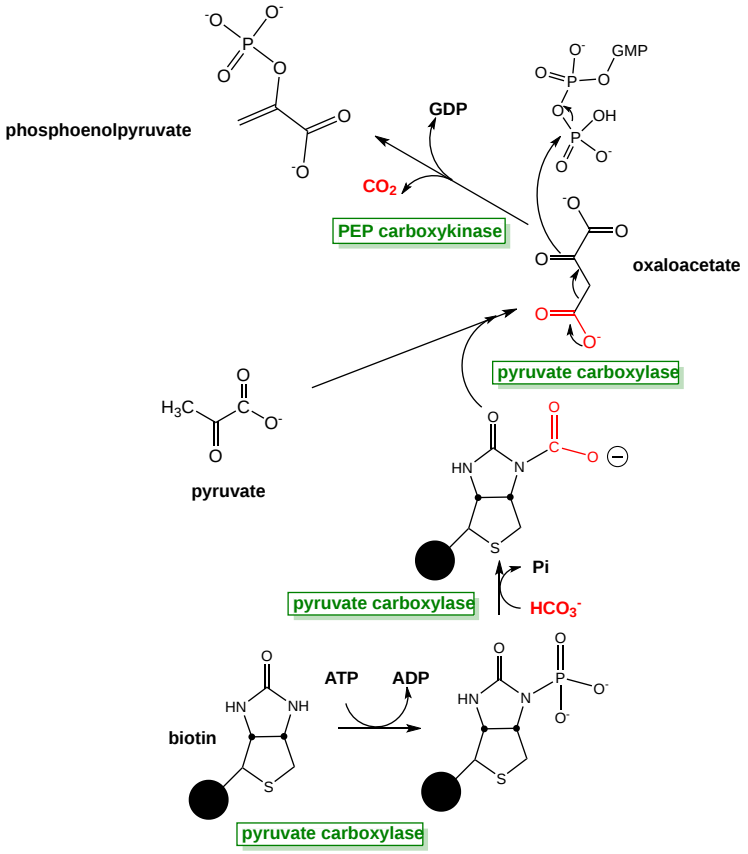


Figure 13.3.2 : Overall reaction for the conversion of pyruvate to phosphoenol pyruvate

In systems that produce oxaloacetate in the mitochondria and then need to transport it to the cytosol to be converted into phosphoenolpyruvate, three enzymes are needed for the process: pyruvate carboxylase (PC) located in the mitochondrial matrix, aspartate aminotransferase (AAT), located in the matrix and the cytosol, and the phosphoenolpyruvate carboxykinase (PCK) located in the cytosol. Oxaloacetate and aspartate are intermediate compounds formed in the process.

In the first reaction, pyruvate carboxylase converts pyruvate into oxaloacetate, based on the following reaction:

Pyruvate carboxylase: pyruvate + HCO_3^- + ATP → oxaloacetate + ADP + P_i

The pyruvate carboxylase enzyme requires biotin as a cofactor and has two major enzymatic functions: 1) carbon fixation from carbon dioxide, and 2) carbon transferase activity, placing the carbon dioxide that has been fixed, onto the molecule of pyruvate to create oxaloacetate. The biotin cofactor is shown in Figure (\PageIndex{3}\). Biotin is a water-soluble vitamin (known as D-biotin or vitamin B7) and is a cofactor in many enzymatic reactions, especially those involving carboxylation (or carbon fixation reactions). It is part of the vitamin B2 complex and is an essential vitamin for mammals. Deficiency results in dermatitis, loss of hair, and neurologic symptoms.

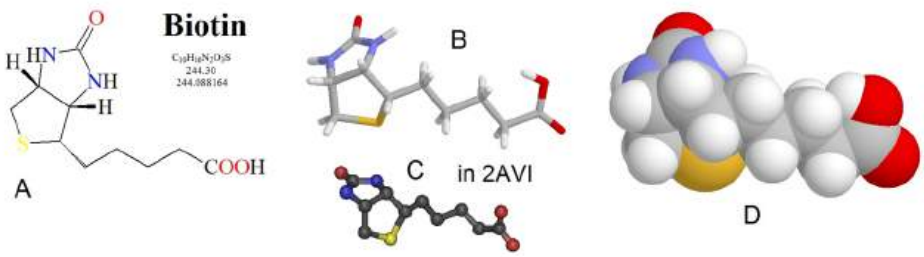


Figure (\PageIndex{3}\): Chemical Structure of Biotin. Biotin is shown in several different formats, (A) line structure, (B) stick model, (C) ball and stick model, and (D) spacefilling model.

Figure from: Biosynthesis

The biotin cofactor forms an amide linkage with the pyruvate carboxylase enzyme at a lysine residue (Figure (\PageIndex{4}\)). This creates a flexible linker region within the pyruvate carboxylase that is capable of dipping the biotin cofactor into the different catalytic domains of the enzyme to obtain the biological activity of the protein. Attaching the biotin to the enzyme requires the energy of ATP.

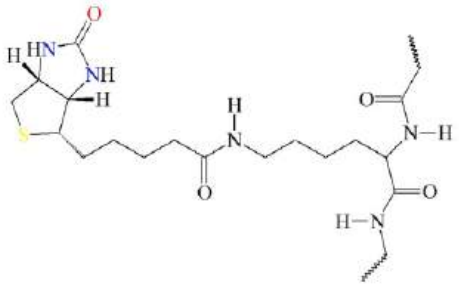


Figure (\PageIndex{4}\) Attachment of Biotin Cofactor to Lysine Residue in Pyruvate Carboxylase.

Image from: Biosynthesis

The pyruvate carboxylase enzyme is a tetramer that contains four functional protein subunits as shown in Figure 13.3.5. The BCCP flexible arm is capable of extending into the biotin carboxylation (BC) domain where biotin is first carboxylated, and then the BCCP shifts over to the carboxyl transferase (CT) domain where the fixed molecule of carbon dioxide can be transferred to pyruvate forming oxaloacetate.

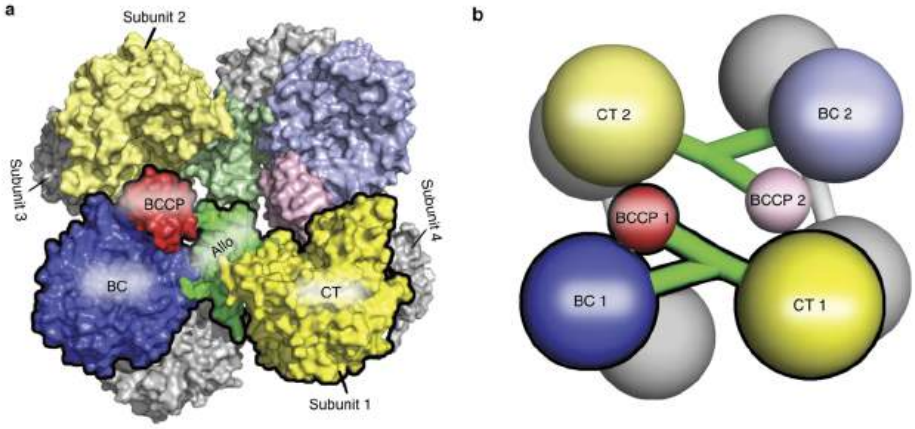


Figure 13.3.5 Structure of the Pyruvate Carboxylase Enzyme. The overall structure of the Pyruvate Carboxylase is a tetramer that contains four functional protein subunits. Two subunits are shown in color, with the other two indicated in gray at the back of the structure. (a) shows the space-filling model while (b) shows the major domains as a cartoon graphic. Each subunit contains a biotin carboxylase (BC) domain shown in blue, a carboxytransferase (CT) domain shown in yellow, and the biotin-carboxyl carrier protein (BCCP) domain shown in red and green.

Figure modified from Liu, Y., et al (2018) Nat Commun 9:1384

In the first part of the reaction, the biotin cofactor is carboxylated using bicarbonate as a substrate with ATP to drive the formation of a molecule with high energy with respect to its hydrolysis product intermediate. (**Remember, there is no such thing as a "high energy" bond.**) Thus, biotin is acting as an intermediate carrier of the carboxy group that will be added to pyruvate during the formation of oxaloacetate. This occurs in the biotin carboxylase domain (BC) with biotin bound to a biotin carboxyl carrier protein (BCCP) domain. The carboxyl group is then transferred to pyruvate to form oxaloacetate in the carboxyl transferase (CT) domain. Let's break up the mechanism of the enzyme from *Rhizobium etli* into three figures.

Figure13.3.6 below shows the first steps in the carboxylation of biotin in the biotin carboxylase domain, using bicarbonate as a substrate. In Figure13.3.6, you can see that an oxygen from bicarbonate mediates nucleophilic attack on the outer phosphate group of the ATP molecule forming the carbonyl-phosphate intermediate + ADP. The carbonyl-phosphate decomposes to release carbon dioxide and phosphate. The negatively charged phosphate activates the biotin cofactor to enable a nucleophilic attack on the carbon of the carbon dioxide producing the carboxybiotin intermediate, shown in the upper part of Figure13.3.7.

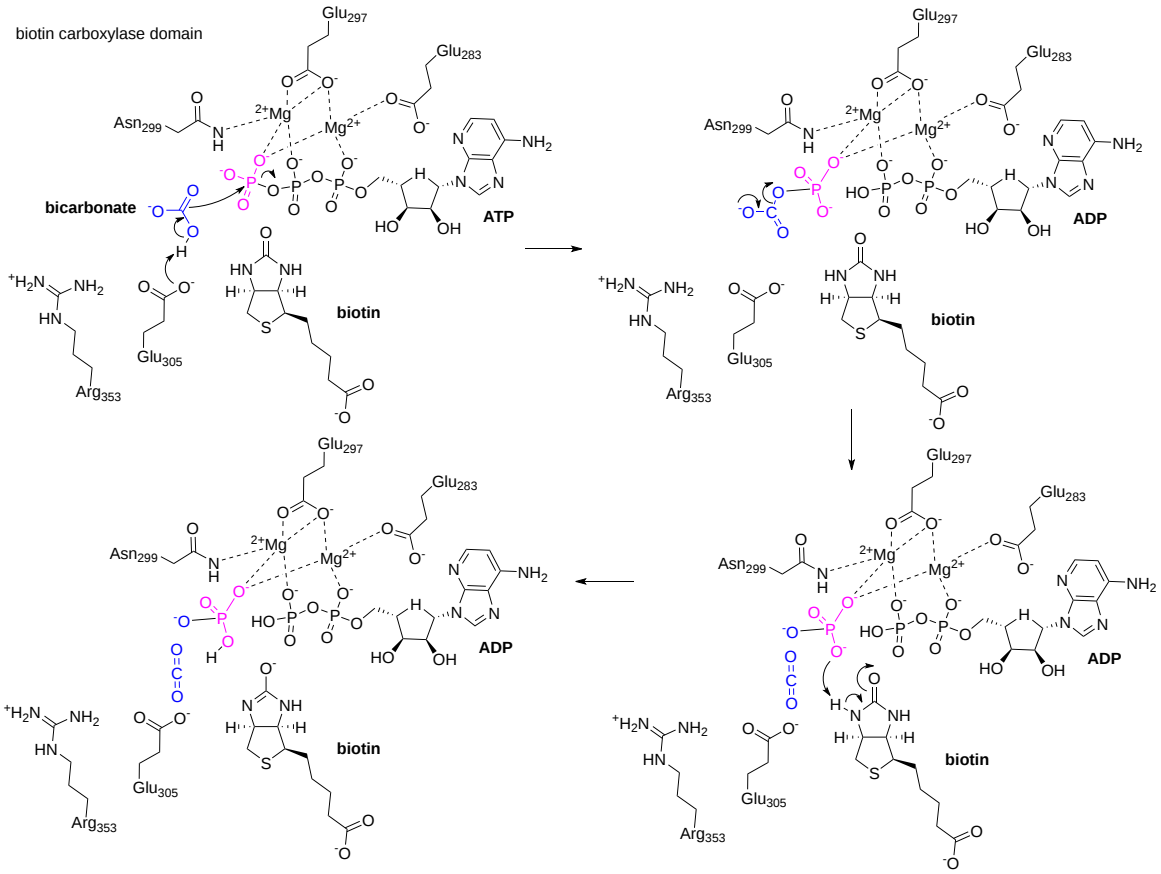


Figure 13.3.6 : The Carboxylation of Biotin during the Pyruvate Carboxylase Reaction Mechanism. Bicarbonate, shown in blue, enters the active site of the enzyme and reacts with the terminal phosphate of ATP forming a phosphorylated intermediate. This reaction is stabilized by two Mg^{2+} cofactors. The carboxyphosphate intermediate decomposes to release carbon dioxide and phosphate.

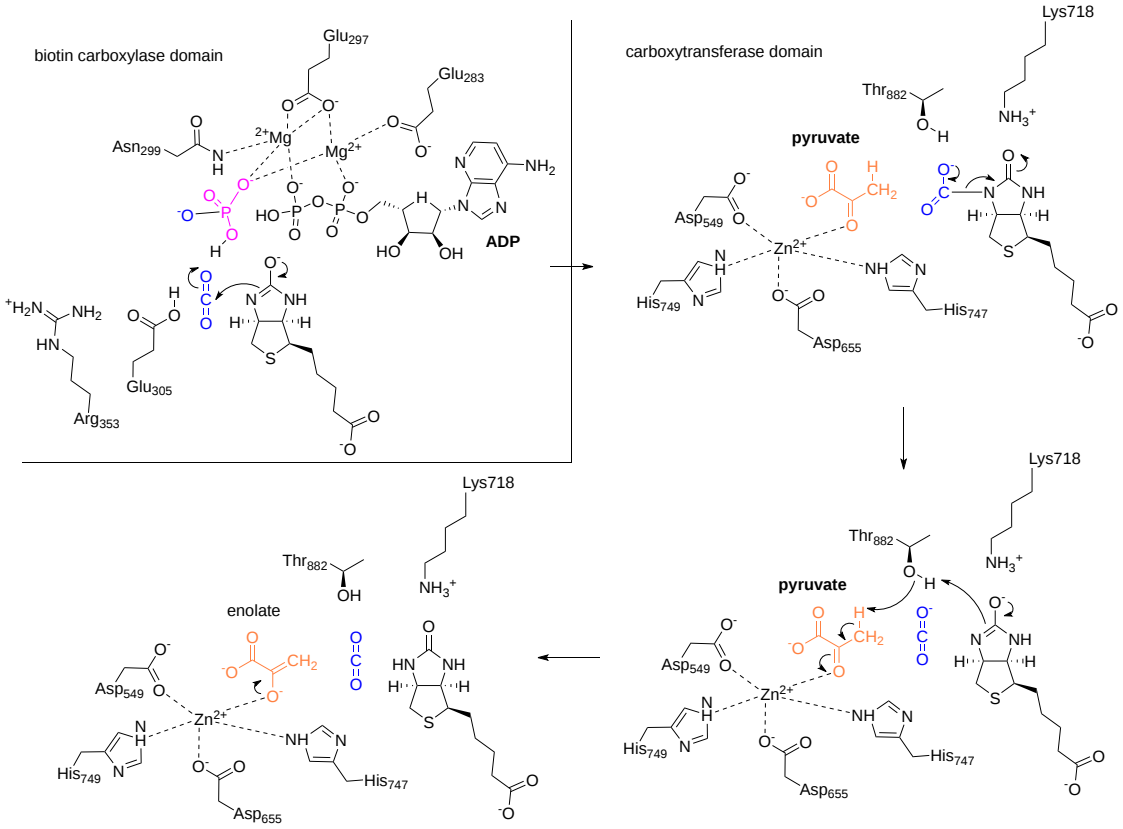


Figure13.3.7 : Carboxylation of biotin in pyruvate carboxylase. The upper diagram shows the actual carboxylation of biotin, which occurs in the BC domain. The carboxybiotin intermediate then shifts on the flexible arm and enters the carboxytransferase (CT) domain. Pyruvate enters the active site and this begins the steps involved in the carboxylation of pyruvate to form oxaloacetate. Ribeiro, ibid

The carboxyltransferase activity is housed in the CT domain of the enzyme and requires the formation of a carbanion intermediate to occur on the pyruvate molecule to make it reactive enough to cause the addition of the carboxyl group. A base from the enzyme eliminates a proton from the C3 position of pyruvate causing the formation of an enol reaction intermediate. As the electrons on the oxygen from the pyruvate molecule reform the carbonyl group, the electrons from the carbon-carbon double bond shift onto the carbon, creating a reactive carbanion intermediate. At this point, the carboxybiotinyl complex swings into the CT domain, providing the carboxyl group that will be incorporated into pyruvate. The carbanion intermediate mediates a nucleophilic attack onto the carboxyl carbon from the BCCP complex, and carbon dioxide is transferred to the pyruvate, creating oxaloacetate. The biotin cofactor is restored by the donation of the proton from the basic residue in the enzyme. This also restores the Pyruvate Carboxylase enzyme for another round of activity. Figure13.3.8 below shows the carboxylation of the enolate form of pyruvate to form oxaloacetate.

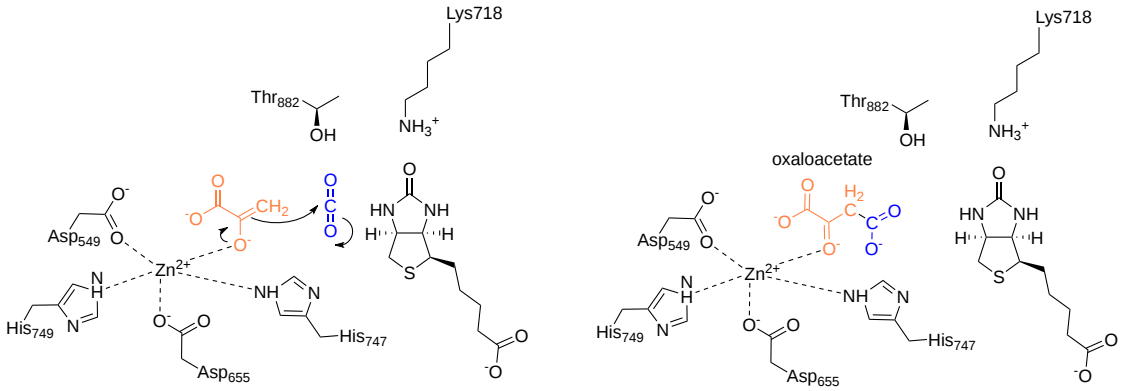
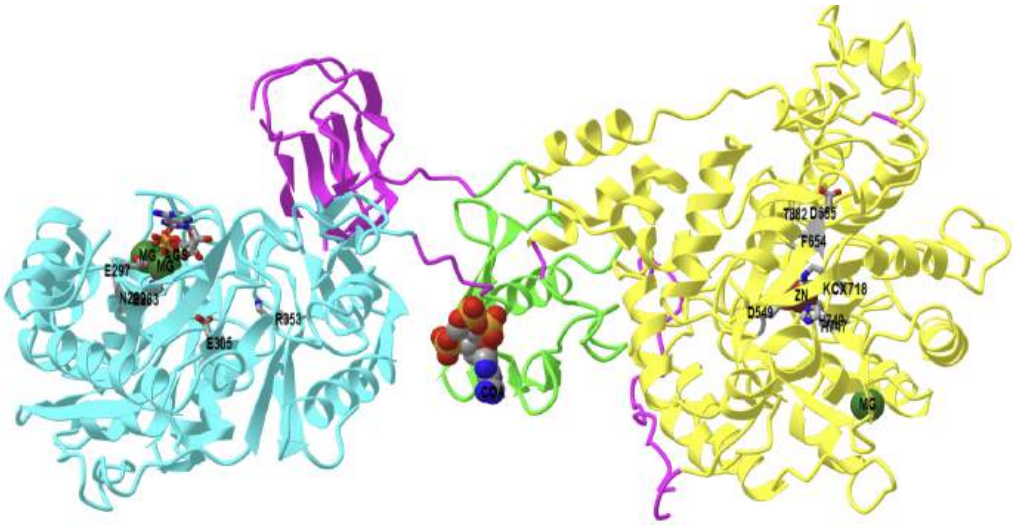


Figure13.3.8 : Carboxylation of the enolate form of pyruvate to form oxaloacetate by pyruvate carboxylase. Ribeiro et al, ibid

Figure13.3.9 below shows an interactive iCn3D model of the biotin-dependent multifunctional enzyme pyruvate carboxylase from Rhizobium etli (2QF7)



NCBI iCn3D **Figure13.3.9 : Biotin-dependent multifunctional enzyme pyruvate carboxylase from Rhizobium etli (2QF7).** The enzyme is a homotetramer with C2 symmetry but only one subunit chain is shown for clarity. The biotin carboxylase domain (BC) is shown in cyan, the carboxytransferase domain (CT) in yellow, a C-terminal biotin carboxyl carrier protein (BCCP) domain (disconnected in parts) in magenta, and the allosteric domain in green. AGS in the cyan domain is a phosphothiophosphoric acid-adenylate ester, an ATP analog. It is shown interacting with the 2 Mg^{2+} ion and it is represented in sticks and labeled. Acetyl-CoA and a nonhydrolyzable analog, ethyl-CoA, are allosteric activators. CoASH (ethyl group not shown) is illustrated in spacefill bound to the green allosteric domain. Acetyl-CoA binding leads to a conformational change that decreases the distance between the two active sites on the BT (cyan) and CT (yellow) domains. Biotin, not shown, is tethered, in the BCCP domain, which transfers biotin between the two catalytic domains not within a single monomer (as shown above), but between different monomers in the homo 4-mer.

(Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...7g91JmzJMad4r9

Figure13.3.10 shows the flexible nature of the biotin flexible arm of the enzyme during the catalytic process.

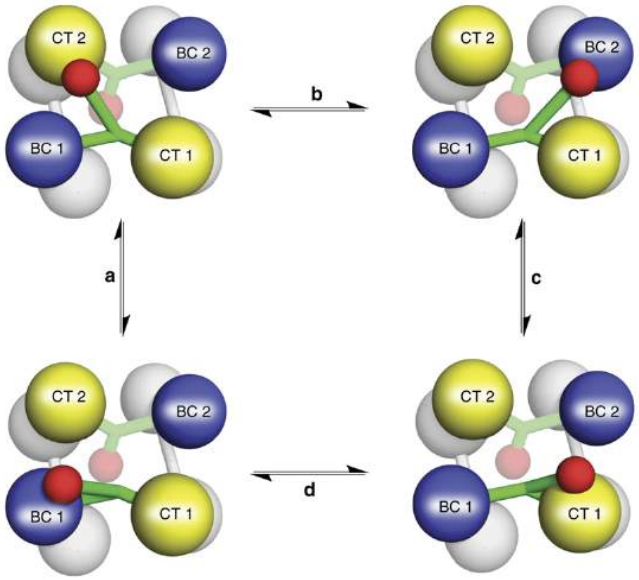


Figure13.3.10 The Catalytic Activity of Pyruvate Carboxylase. The Biotin-Carboxyl Carrier Protein Domain (BCCP) is a flexible arm of the protein structure that can swing from one catalytic domain to the other during the reaction process. As shown in the diagram, the BCCP domain of subunit 1 will be fixed with carbon dioxide at the biotin carboxylase domain from subunit 1 (BC1). The BCCP will then translocate to the carboxyltransferase domain from subunit 2 (CT2) where it will create one molecule of oxaloacetate and reset the biotin cofactor. The BCCP will then swing into the BC2 domain where carbon fixation will occur. BCCP then shifts to the CT1 domain where it creates a second oxaloacetate molecule and resets the biotin. While this is going on with the BCCP from subunit 1, the other BCCP arms from subunits 2, 3, and 4 are also active. Thus, one full sequence of the enzyme activity will produce 8 molecules of oxaloacetate.

Image from: Liu, Y., et al (2018) Nat Commun 9:1384

The next reaction in the process requires the conversion of oxaloacetate to phosphoenolpyruvate by phosphoenolpyruvate carboxykinase (PEPCK). However, the major human PEPCK enzyme, PCK1, resides in the cytoplasm of the cell. Thus the oxaloacetate produced in the matrix of the mitochondria by pyruvate carboxylase must be transported into the cytosol. However, there are no oxaloacetate transporters that can mediate the transfer directly. To leave the mitochondrial matrix, oxaloacetate has to first be converted to aspartate by the aspartate aminotransferase enzyme (AAT). It can then be transported to the intermembrane space of the mitochondria through an antiporter that transports one molecule of aspartate out of the matrix and one molecule of glutamate into the matrix Figure13.3.11. Once in the intermembrane space, the aspartate can pass freely through a pore in the outer mitochondrial membrane. When it reaches the cytoplasm, aspartate is reconverted to oxaloacetate using the cytoplasmic aspartate aminotransferase enzyme (AAT). Oxaloacetate can then be used as a substrate by the phosphoenolpyruvate carboxykinase enzyme (PEPCK).

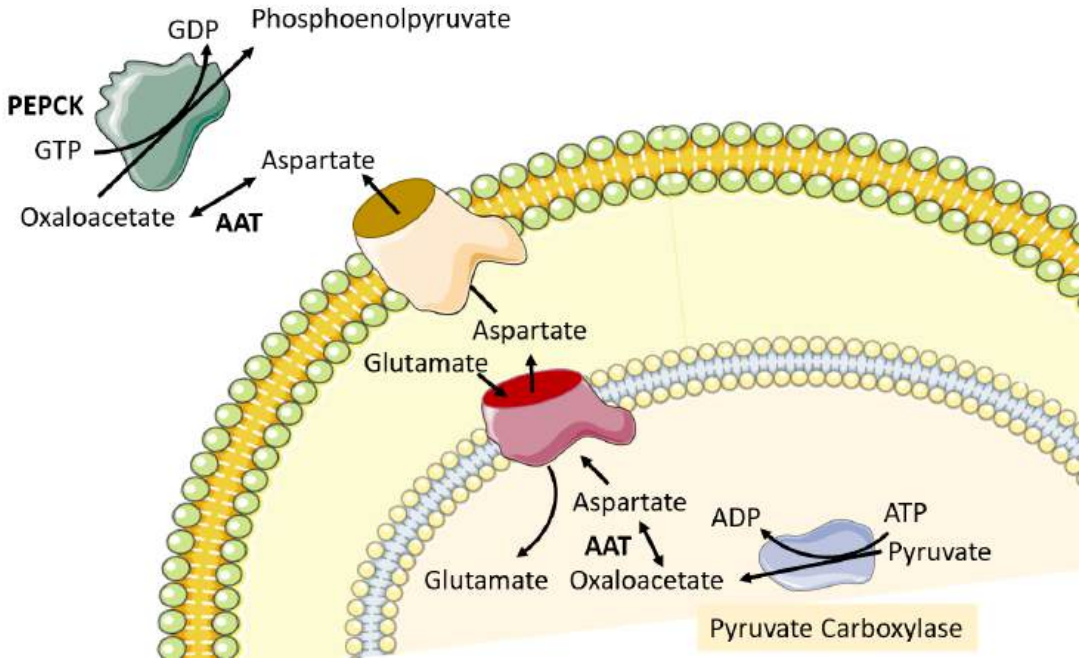


Figure13.3.11 Conversion of Oxaloacetate to Aspartate and Transfer to the Cytoplasm.

Image modified from SMART Servier Medical Art

The glutamate/aspartate transporter has some additional complexities associated with it. It cannot function on its own. It requires the coordinated functioning of the malate/alpha-ketoglutarate antiporter. Together, these antiporters are known as the Malate-Aspartate Shuttle System.

The Malate-Aspartate Shuttle System is dependent on the functioning of two enzymatic processes. The first is the aspartate aminotransferase that was indicated more simplistically in Figure13.3.11. The aspartate aminotransferase enzyme can utilize glutamate as an amine donor to generate aspartate from oxaloacetate. Alpha-ketoglutarate is also formed in this process. Depending on substrate concentrations and other regulatory mechanisms, this enzyme can also work in the reverse reaction to produce glutamate and oxaloacetate. In a different reaction using Malate Dehydrogenase, oxaloacetate can be reduced to form malate using a molecule of NADH as the electron donor.

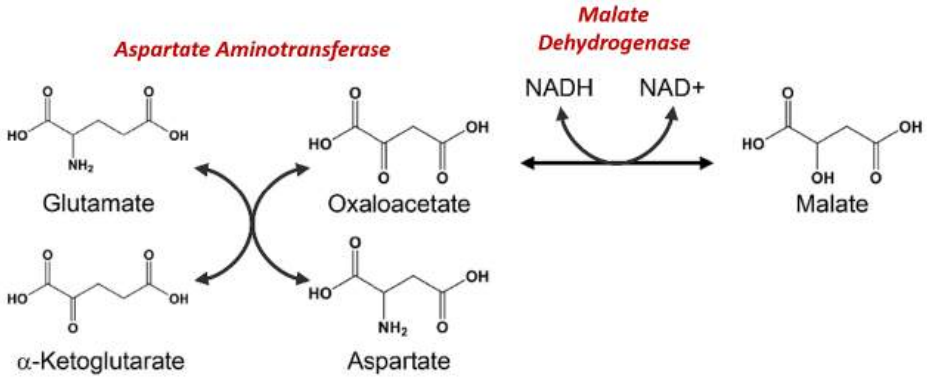


Figure13.3.12 Chemical Reactions of the Malate-Aspartate Shuttle System

Figure modified from Son, H.F. and Kim, J-J. (2016) PLoS One 10.1371

The malate dehydrogenase enzyme is expressed to high levels in the cytoplasm of liver cells as well as in the matrix of the mitochondria. This enzyme is a component of the Kreb Cycle where it mediates the formation of oxaloacetate in the last step of the cycle. Within the cytoplasm, it predominantly converts oxaloacetate to malate. The malate can then be shuttled into the matrix of the mitochondria through the malate/alpha-ketoglutarate antiporter (Figure13.3.13). In this antiporter, malate moves into the mitochondrial matrix while alpha-ketoglutarate moves into the intermembrane space (and subsequently into the cytosol of the cell). Together with aspartate, it can be used by cytoplasmic aspartate aminotransferase to produce glutamate and oxaloacetate. This is the oxaloacetate pool that is then utilized in the gluconeogenic pathway when PEPCK is active. The glutamate generated from this reaction is transported back into the matrix of the mitochondria through the aspartate/glutamate antiporter. The pool of aspartate in the matrix of the mitochondria is supplied by the reaction of the aspartate aminotransferase enzyme, which completes the reverse reaction from the one seen in the cytoplasm. In the matrix, aspartate aminotransferase uses glutamate and oxaloacetate as substrates to generate aspartate and alpha-ketoglutarate. This enables the transport of aspartate and glutamate through their specific antiporter. Within the gluconeogenic pathway, heightened levels of oxaloacetate are produced in the matrix of the mitochondria. Oxaloacetate is then converted to aspartate and transported across the inner membrane, where it can subsequently be converted back into oxaloacetate and used for the production of glucose.

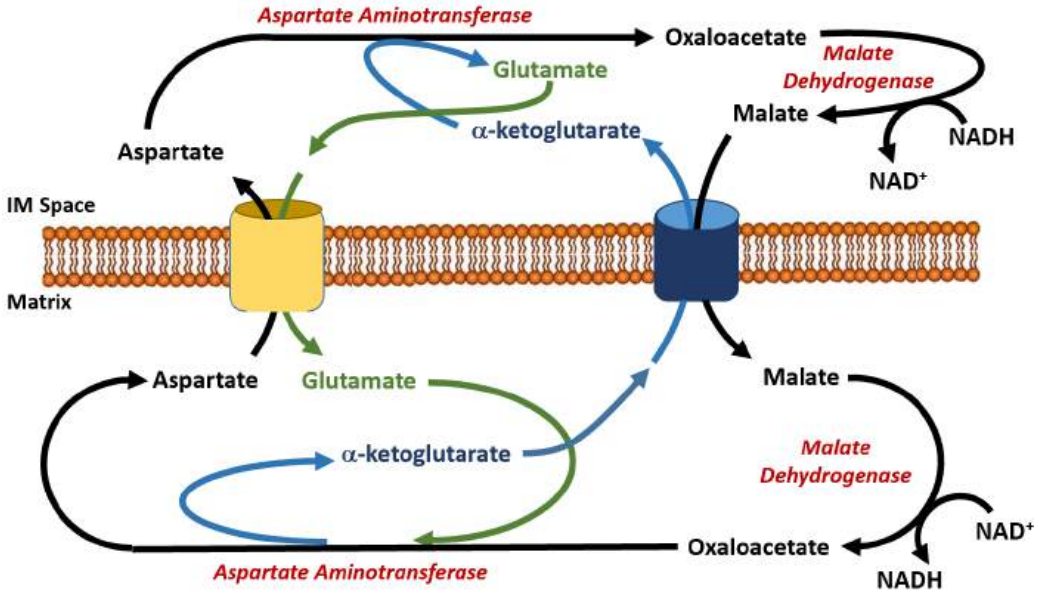


Figure13.3.13 : The Malate-Aspartate Shuttle System.

Once oxaloacetate has been effectively transported via the Malate-Aspartate Shuttle into the cytoplasm, it is converted to phosphoenolpyruvate (PEP) by the PEP carboxykinase (PEPCK). The overall reaction mediated by PEPCK is:

PEP carboxykinase (PEPCK): oxaloacetate + GTP → phosphoenol pyruvate + GDP

PEPCK is classified as a lyase enzyme and exists in two isozymes, a cytoplasmic PEPCK, and a mitochondrial PEPCK. The cytoplasmic form is the one that is predominantly used in the gluconeogenic pathway and requires the action of the Malate-Aspartate Shuttle. However, small amounts of PEP can be made directly by the mitochondrial PEPCK and then transported across the mitochondrial membrane. The conversion of oxaloacetate to phosphoenolpyruvate by PEPCK mediates the removal of carbon dioxide and the addition of a phosphate group. A molecule of GTP is used during this process as the phosphate donor and magnesium ions serve as a cofactor. Figure 13.3.14 shows the reaction mechanism for the human PEP carboxykinase.

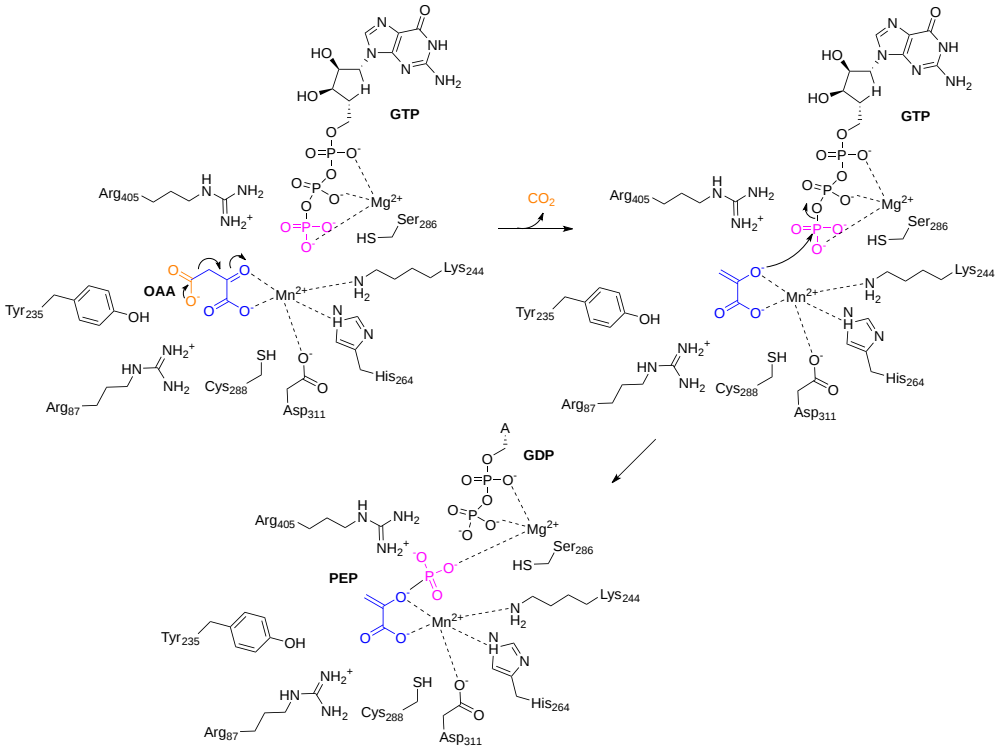
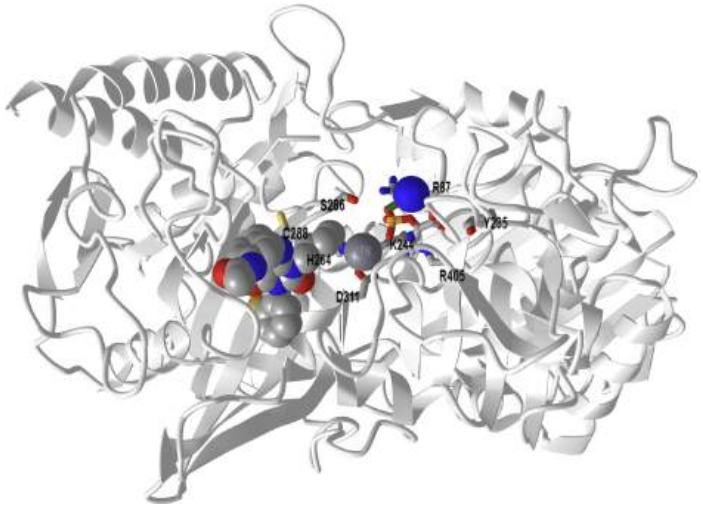


Figure 13.3.14 : Reaction mechanism for human PEP carboxykinase (PEPCK)

Figure13.3.15 below shows an interactive iCn3D model of phosphoenolpyruvate carboxykinase (PEPK) with a bound PEP and GTP-competitive inhibitor (1NHX).



NCBI iCn3D **Figure 13.3.15 : Phosphoenolpyruvate carboxykinase (PEPK) with a bound PEP and GTP-competitive inhibitor (1NHX)**. PEP is shown in colored sticks. The competitive inhibitor is shown in colored spacefill. The active site residues from the human PEPCK mechanism are shown as labeled sticks. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...QFcTssCJbuW2K7

13.3.3: PHOSPHATASE REACTIONS OF GLUCONEOGENESIS (REVERSE OF STEPS 1 AND 3 IN GLYCOLYSIS):

The last two unique enzymes of the gluconeogenic pathway are both phosphatase enzymes. Fructose 1,6-bisphosphatase converts fructose 1,6-bisphosphate to fructose 6-phosphate. The same phosphoglucoisomerase used in glycolysis can convert fructose 6-phosphate back into glucose 6-phosphate. A unique glucose 6-phosphatase enzyme will then convert glucose 6-phosphate to free glucose. The unique phosphatase reactions are shown in Figure13.3.16.

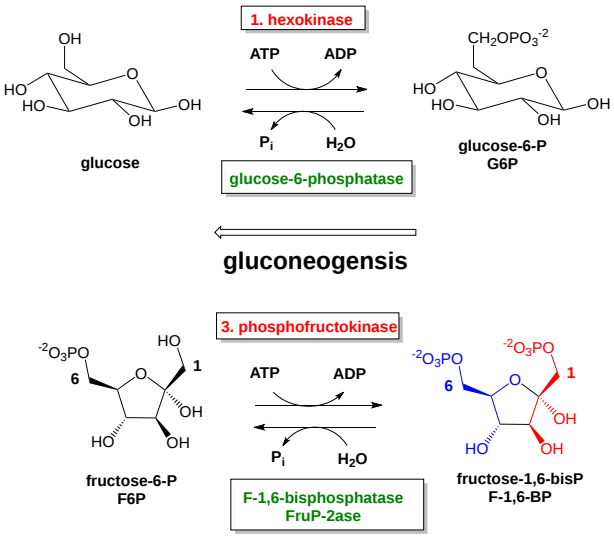


Figure 13.3.16 : Phosphorylase Enzymes utilized during Gluconeogenesis

First, let's consider the dephosphorylation of fructose 1,6-bisphosphate. The gluconeogenic enzyme is named fructose-1,6-bisphosphatase, (FBP or FBPase-1). FBP requires a metal cofactor and is competitively and allosterically regulated. Fructose 2,6-bisphosphate and low energy load (AMP and ADP) inhibit the enzyme. The overall reaction is shown below:

Fructose-1,6-bisphosphatase: F1,6BP + H_2O → F6P + P_i

Figure 13.3.17 shows the mechanism of mouse fructose,1-6-bisphosphatase (FBPase-1)

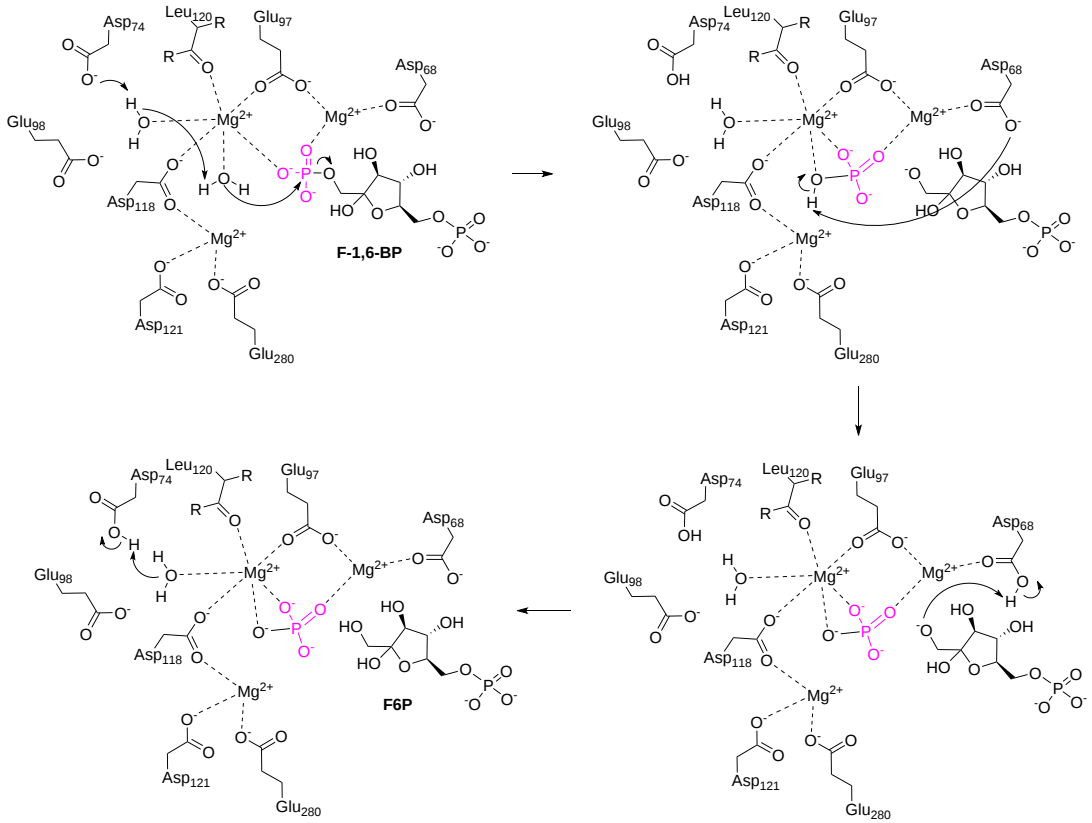


Figure 13.3.17 : Mechanism of mouse fructose,1-6-bisphosphatase. Mg^{2+} ions coordinate the positioning of F-1,6-BP near an activated water molecule that mediates nucleophilic attack on the 1-position phosphate group. This leads to the cleavage of the phosphate group from the sugar molecule, F6P, which can then leave the active site.

Figure13.3.18 below shows an interactive iCn3D model of pig fructose-1,6-bisphosphatase with bound Mg^{2+}, fructose-6-phosphate, and phosphate in the R state (1EYI).

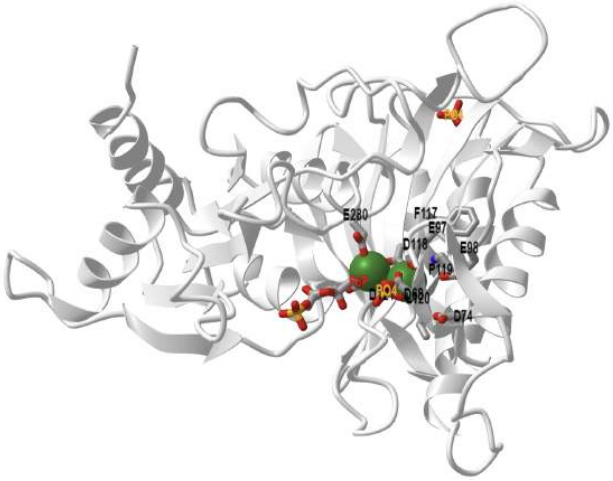


Figure 13.3.18 : Fructose-1,6-bisphosphatase with bound Mg, fructose-6-phosphate and phosphate in the R state (1EYI). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...DbWMHBRTdPGhJ8

FBP is an allosteric homotetramer and can exist in a tense (T) and relaxed (R) state. When the enzyme is in the R state, it has the highest catalytic activity. Figure13.3.18 shows only one monomer of the R state with bound ligands and substrates for clarity.

Figure 13.3.19 shows an interactive iCn3D model of tetrameric FBP enzyme in the T state with bound Mg^{2+}, fructose-6-phosphate, phosphate, and AMP, an allosteric inhibitor, in the T state (1EYJ).

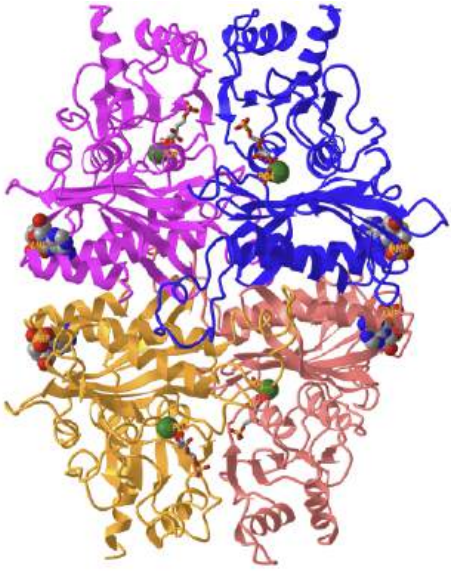


Figure 13.3.19 : Tetrameric fructose-1,6-bisphosphatase in the T state with bound Mg^{2+}, fructose-6-phosphate, phosphate and AMP in the T state (1EYJ). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...UAbjNtkPsoBdA9

FBPase is inhibited by AMP and fructose-2,6-bisphosphate, and activated by ATP. Allosteric changes control the catalytic activity of the enzyme and the transition from the T to the R state, or vice versa. As shown in Figure13.3.17, FBP has 3 metal binding sites. The mechanism shown in Figure 13.3.17 In the model, all of the binding sites are filled with Mg^{2+}, however, Zn^{2+} can also play a role. In the presence of AMP, only one site is occupied by Mg^{2+}, as shown in the model of the T state in Figure13.3.19. The T state form of FBP has a loop (52-72) that is disordered. In the R state (no AMP), the loop interacts with the active site and 1 Zn^{2+} and 2 Mg^{2+} ions are bound in the three sites. The binding of AMP leads to the dissociation of two of the bound metal cofactors and causes a decrease in enzyme activity.

The resulting fructose 6-phosphate is isomerized to glucose 6-phosphate. Glucose 6-phosphate is then converted by the last unique gluconeogenic enzyme, glucose 6-phosphatase, to generate free glucose. The overall reaction is shown below:

Glucose-6-phosphatase: G6P + H_2O → glucose + P_i

The dephosphorylation of glucose only occurs appreciably in liver cells, as this is the primary location for the regulation of blood glucose levels. This serves as the final step in the gluconeogenic pathway. The glucose 6-phosphatase enzyme is a transmembrane protein that resides in the inner membrane of the endoplasmic reticulum. Thus, for glucose 6-phosphate to be dephosphorylated, it must first be transported from the cytoplasm into the lumen of the endoplasmic reticulum (ER) through transporter 1 (T1) (Figure13.3.20). The glucose 6-phosphatase (G-6-Pase) then cleaves the phosphate from the substrate, releasing inorganic phosphate (P) and glucose (red molecule). The inorganic phosphate is then transported back into the cytoplasm through transporter 2 (T2) and glucose is transported through Transporter 3 (T3). Free glucose is then transported back into the bloodstream through a glucose (GLUT) transporter (not shown in Figure13.3.20).

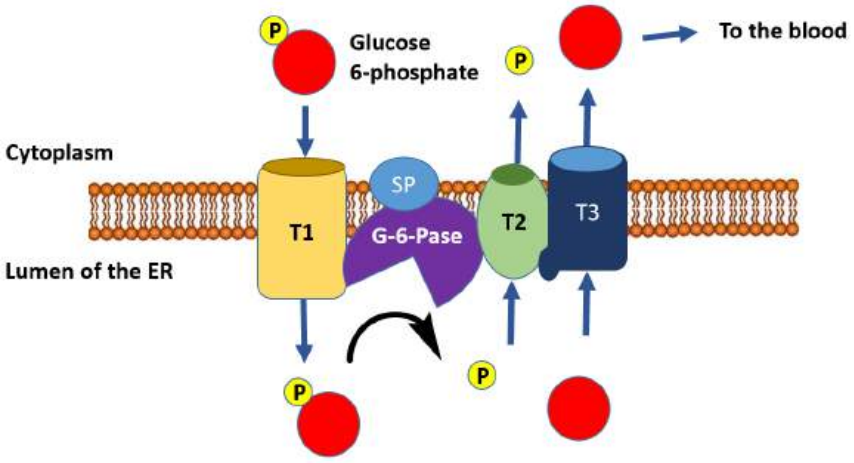


Figure13.3.20 Dephosphorylation of Glucose 6-Phosphate in the Lumen of the Endoplasmic Reticulum.

The catalytic site of this hydrolase enzyme is comprised of Lys76, Arg83, His119, Arg170, and His176. His 176 acts as a nucleophile that attacks the phosphorous of G6P in an S_N2-like reaction to form a His-PO_3^{2-} intermediate. Hydrolysis follows restoring the enzyme to its original state. Figure 13.3.21 shows an abbreviated mechanism for the reaction.

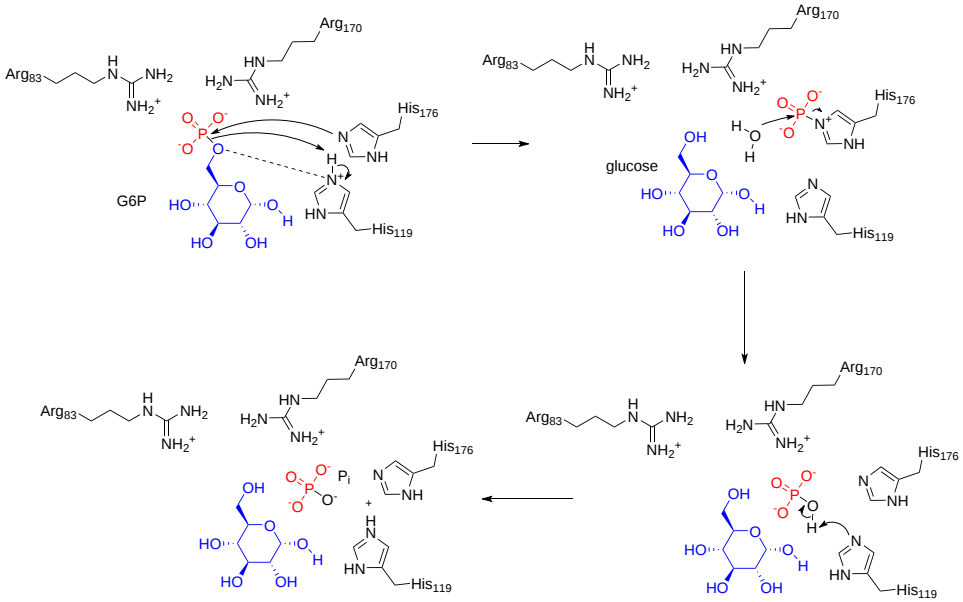


Figure 13.3.21 : Abbreviate mechanism for human glucose-6-phosphatase

Figure 13.3.21 shows an interactive iCn3D model of the AlphaFold predicted structure of human glucose-6-phosphatase (P35575). The enzyme contains nine transmembrane helices that dock the protein in the inner membrane of the ER. The active site faces the ER lumen.

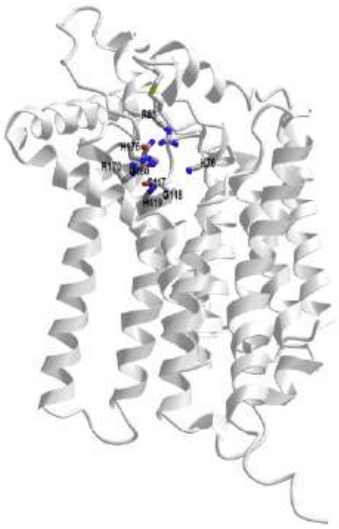


Figure 13.3.14: AlphaFold predicted structure of human glucose-6-phosphatase (**P35575)**(Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...jmcQAo49eZ8p16

13.3.4: SUMMARY:

It is noteworthy that the process of gluconeogenesis is energy intensive. For example, the phosphatase hydrolysis reactions that liberate glucose are not reverse kinase reactions and they do not produce substrate-level phosphorylation the way reactions occur in the second half of the glycolytic pathway. Overall, there are approximately 6 molar equivalents of ATP required to make 1 mole of glucose. Over 3 times as much energy is consumed during glucose formation than is generated from glycolysis during the breakdown of glucose. Per glucose molecule made, 2 ATP are used in the pyruvate carboxylase step, 2 GTP are used in the phosphoenolpyruvate carboxykinase step, and 2 ATP are used in the kinase reaction to create 1,3-bisphosphoglycerate from 3-phosphoglycerate. There is also the loss of 2 moles of NADH during the reverse reaction to create glyceraldehyde 3-phosphate from 1,3-bisphosphoglycerate, which reduces energy potential through oxidative phosphorylation in the mitochondria. Plus there are energy costs to changing the metabolite pools in the matrix of the mitochondria and for transporting molecules across the mitochondrial membrane that we have not considered here.

13.3.5: REFERENCES

Principles of Biochemistry/Gluconeogenesis and Glycogenesis. (2019, February 8). *Wikibooks, The Free Textbook Project*. Retrieved 22:01, August 11, 2022, from https://en.wikibooks.org/w/index.php?title=Principles_of_Biochemistry/Gluconeogenesis_and_Glycogenesis&oldid=3515124.

13.4: PENTOSE PHOSPHATE PATHWAY OF GLUCOSE OXIDATION

13.4.1: INTRODUCTION

The pentose phosphate pathway (PPP), also known as the pentose phosphate shunt, is an important part of glucose metabolism. The PPP branches after the first step of glycolysis and consumes the intermediate glucose 6-phosphate (G6P) to generate fructose 6-phosphate (F6P) and glyceraldehyde 3-phosphate (G3P) through the oxidative and non-oxidative branches of the PPP. Unlike glycolysis and glucose aerobic oxidation, the PPP does not provide adenosine 5′-triphosphate (ATP) to meet the energy demands of cells. Instead, it supplies NADPH and ribose 5-phosphate (R5P). These two metabolites are vital for the survival and proliferation of cells. R5P is a building block for nucleic acid synthesis. NADPH is the reducing power required for the synthesis of fatty acids, sterols, nucleotides, and non-essential amino acids. Moreover, NADPH-derived conversion of oxidized glutathione (GSSG) to reduced glutathione (GSH) *via* glutathione reductase is important for cellular antioxidant defenses. Interestingly, NADPH also serves as the substrate of NADPH oxidases (NOXs) which produce reactive oxygen species (ROS).

Both the oxidative branch and non-oxidative branch of the PPP take place in the cytosol (Figure 13.4.1). Glucose 6-phosphate dehydrogenase (G6PD) is the rate-limiting enzyme of the oxidative PPP, determining the flux of G6P directed into the pathway. G6PD catalyzes the conversion of G6P to 6-phosphogluconolactone, accompanied by NADPH production. 6-phosphogluconolactonase (6PGL) is the enzyme that hydrolyzes 6-phosphogluconolactone to produce 6-phosphogluconate (6PG). 6-phosphogluconate dehydrogenase (6PGD) converts 6-PG to ribulose 5-phosphate (Ru5P) and generates NAPDH (Figure 13.4.1). The largest contributor to cytosolic NADPH is the oxidative PPP in mammalian cells.

The non-oxidative branch is composed of a series of reversible transfer reactions of chemical groups. Ribose 5-phosphate isomerase (RPI) and ribulose 5-phosphate epimerase (RPE) catalyze reversible reactions converting Ru5P to R5P and xylulose 5-phosphate (Xu5P), respectively. TKT catalyzes two reversible reactions. One is the conversion of Xu5P and R5P to G3P and sedoheptulose 7-phosphate (S7P). The other is the conversion of Xu5P and erythrose 4-phosphate (E4P) to G3P and F6P. Therefore, TKT can bi-directionally regulate the carbon flux between the non-oxidative PPP and glycolysis or gluconeogenesis. Transaldolase (TALDO) reversibly converts G3P and S7P to E4P and F6P. The non-oxidative branch not only replenishes metabolites of the oxidative branch (by their reversal), but also regulates the flux of glycolysis or gluconeogenesis by providing F6P and G3P (Figure 13.4.1).

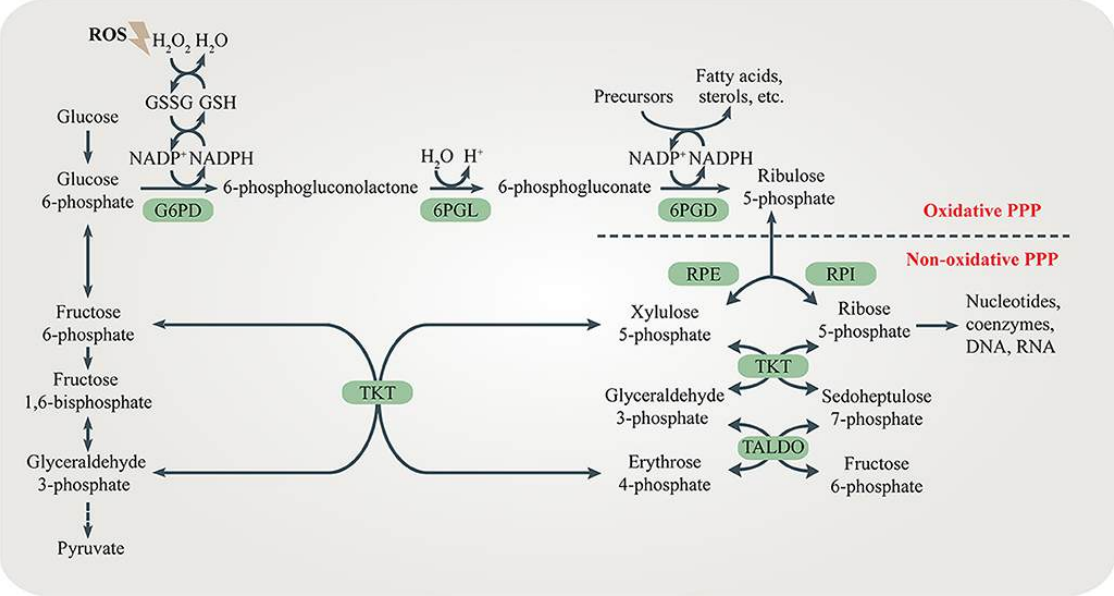


Figure 13.4.1 : The pentose phosphate pathway (PPP). The PPP branches after the first step of glycolysis and goes back to fructose 6-phosphate and glyceraldehyde 3-phosphate in the glycolytic and gluconeogenic pathway. The PPP produces R5P and NADPH for biosynthesis and redox regulation. Enzymes in the oxidative and non-oxidative PPP are shaded in green. Figure from: Ge, T., et al. (2020) Cellular Endocrinology DOI:10.3389

13.4.2: OXIDATIVE BRANCH

The oxidative branch of PPP (ox-PPP) is a non-reversible metabolic pathway where glucose-6-phosphate (G6P) is transformed into 6-phosphoglucono-δ-lactone by glucose-6-phosphate dehydrogenase (G6PD) and, subsequently, to ribulose-5-phosphate by 6-phosphogluconate dehydrogenase (6PGD) with the concomitant production of nicotinamide adenine dinucleotide phosphate (NADPH). The resulting ribulose-5-phosphate is then converted to ribose-5-phosphate and used for the biosynthesis of nucleotides (Figure13.4.2).

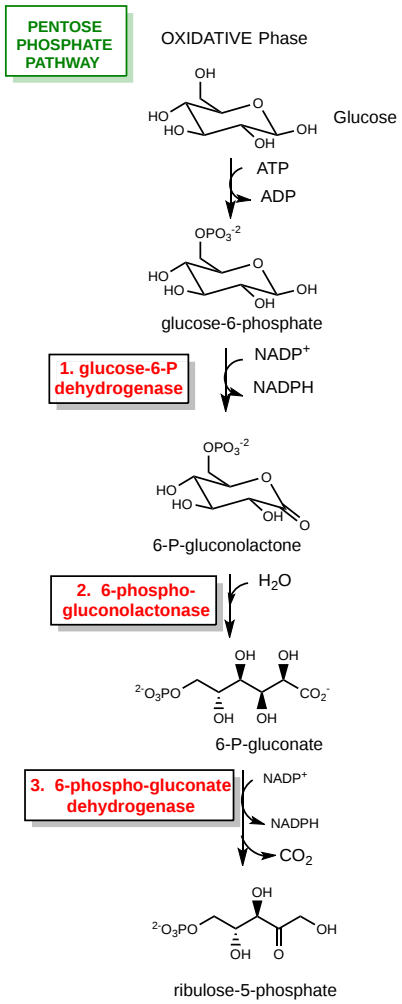


Figure 13.4.2 : Summary of the oxidative branch of phosphopentose pathway

The first enzyme in the oxidative branch is the glucose-6-phosphate dehydrogenase enzyme. It is also the first committed enzyme in the pathway and is involved in the regulation. PPP metabolizes from 5 to 30% of glucose depending on the tissue type. The enzymatic reaction is:

1. Glucose-6-phosphate dehydrogenase: Glc6P + $NADP^+$ → 6-phosphogluconolactone + NADPH

The oxidative branch produces NADPH for reductive biosynthesis and it also maintains the reducing condition of the cell to protect it against oxidative stress, which is especially important in erythrocytes (Figure13.4.3). It also starts the pathway to produce 5-carbon sugars for nucleotide biosynthesis.

G6PDH AND ROS

G6PDH is very important for the protection of free radicals in red blood cells since they don't have mitochondria that could provide another source of NADPH for protection. NADPH is involved in protection against ROS through the three cycles shown in Figure 13.4.3.

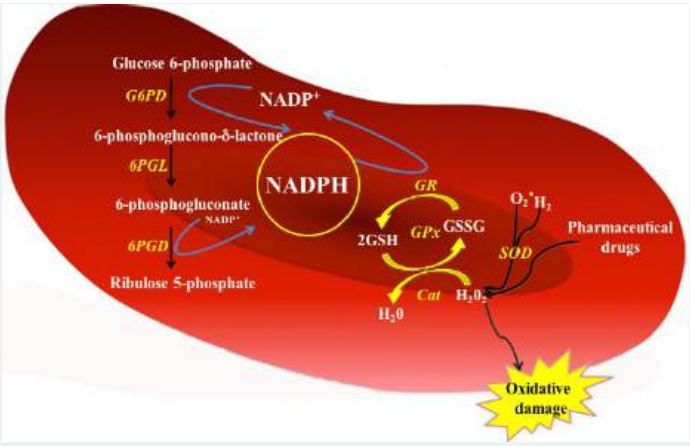


Figure 13.4.3 : Function of G6PD enzyme in the PPP from red blood cells. In G6PD-normal red cells, the NADPH is produced by the action of glucose 6-phosphate dehydrogenase (G6PD) and 6-phosphogluconate dehydrogenase (6PGD) enzymes. The NADPH serves as a proton donor to regenerate the GSSG oxidized. Cat = Catalase; GPx = Glutathione peroxidase; GR = Glutathione reductase; G6PD = glucose 6-phosphate dehydrogenase; 6PGL = 6-phosphogluconolactonase; 6GPD = 6-phosphogluconate dehydrogenase; SOD = Superoxide dismutase; GSH = Reduced glutathione; GSSG = Oxidized glutathione; H_2O_2 = Peroxide; O_2^- = Superoxide. Gomez-Manzo et al. Int J Mol Sci. 2016 Dec; 17(12): 2069. doi: 10.3390/ijms17122069. Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).

NADPH can keep glutathione in its reduced form, which as a substrate for catalase can help rid the cell of hydrogen peroxide and indirectly other ROS.

The enzyme is active as a dimer or tetramer. Each monomer in the complex has a substrate binding site that binds to G6P, and a catalytic coenzyme binding site that binds to $NADP^+$/NADPH. For some higher organisms, such as humans, G6PD contains an additional $NADP^+$ binding site, called the $NADP^+$ structural site, that does not seem to participate directly in the reaction catalyzed by G6PD. The evolutionary purpose of the $NADP^+$ structural site is unknown, however, it does play a role in the overall stability of the enzyme.

Mutations that cause dysfunction or deficiencies are very common, especially in males and in Africa, Asia, the Mediterranean, and the Middle East, in a geographic distribution that parallels the incidence of malaria. Interestingly, many of the mutations occur near the $NADP^+$ structural site. Glucose-6-phosphate dehydrogenase deficiency is very common worldwide and causes acute hemolytic anemia in the presence of simple infection, ingestion of fava beans, or reaction with certain medicines.

An abbreviate mechanism for glucose-6-phosphate dehydrogenase from *Leuconostoc mesenteroides* is shown in Figure 13.4.4.

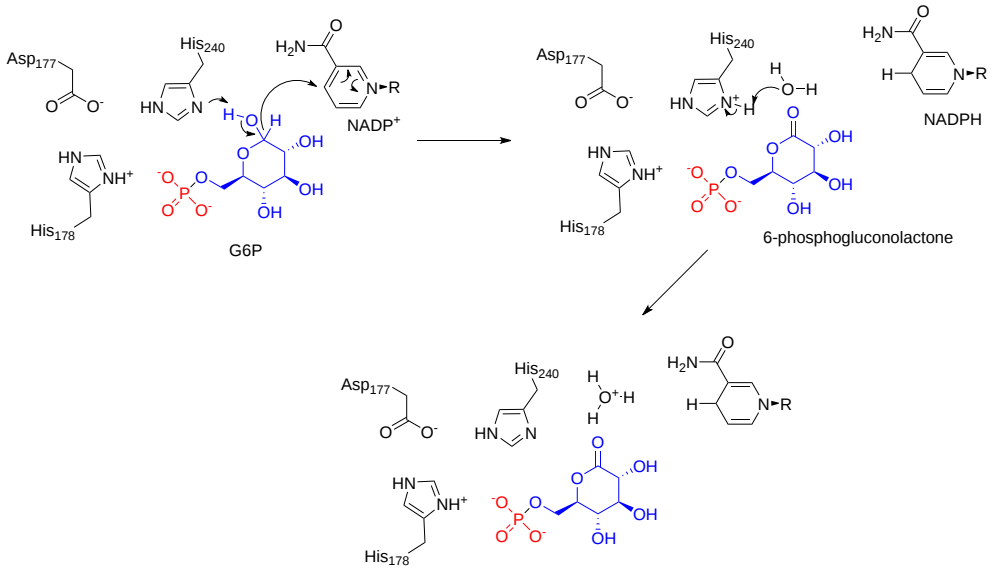


Figure 13.4.4 : Mechanism of *Leuconostoc mesenteroides* glucose-6-phosphate dehydrogenase. Ribeiro AJM *et al.* (2017), *Nucleic Acids Res*, **46**, D618-D623. Mechanism and Catalytic Site Atlas (M-CSA): a database of enzyme reaction mechanisms and active sites.

DOI:10.1093/nar/gkx1012. PMID:29106569 https://www.ebi.ac.uk/thornton-srv/m-csa/entry/843/.

Figure 13.4.5 shows an interactive iCn3D model of the glucose 6-phosphate dehydrogenase from Leuconostoc mesenteroides (1DPG).

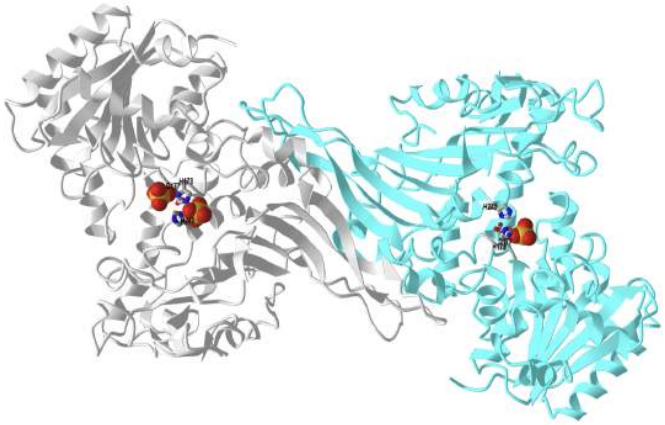


Figure 13.4.5 : Glucose 6-phosphate dehydrogenase from Leuconostoc mesenteroides (1DPG). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...m8NCXiBuQbTQb7

Interestingly, this bacteria has an incomplete glycolytic pathway and can use this enzyme with either $NADP^+$ for anabolism or NAD^+ for catabolism.

Each monomer can bind two $NADP^+$, one at a site that promotes structure integrity and stability and the other at the catalytic site where $NADP^+$ serves as a substrate (or cofactor). The other substrate, glucose-6-phosphate bind between the two.

Figure 13.4.6 shows an interactive iCn3D model of the human glucose 6-phosphate dehydrogenase with bound structural and substrate $NADP^+$ (2BH9).

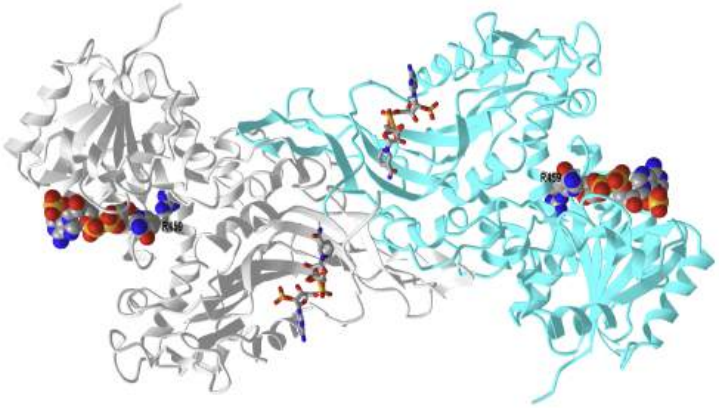


Figure 13.4.6 : Human glucose 6-phosphate dehydrogenase with bound structural and substrate NADP (2BH9). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?NFJnS7UtUt3Uwu2D7

The biological unit shown is a dimer. The structural $NADP^+$ is shown in spacefill bound in the N-terminal end of each monomer, while the substrate (cofactor) $NADP^+$ is shown in sticks. Arg 459 is shown in spacefill.

A common disease-causing mutation of G6PDH is the Canton R459L mutation. It's common in China and Southeast Asia. The activity of the enzyme is significantly decreased and the enzyme is less able to form tetramers. It unfolds at a lower temperature. This suggests that the mutation causes a significant conformation change. The Arg459 is shown in the above model. Its contribution to interhelical noncovalent attractions between it and D181 and N185. D181 is shown in proximity to Arg459 in the above model.

Figure 13.4.7s shows a static image of glucose 6-phosphate dehydrogenase with bound structural and substrate $NADP^+$ as well as glucose-6-phosphate.

Figure 13.4.7 : Glucose 6-phosphate dehydrogenase with bound structural and substrate $NADP^+$ as well as glucose-6-phosphate (2BHL and 2BH9). Structural $NADP^+$ (**blue** molecular surface), catalytic $NADP^+$ (dark **purple** molecular surface), and G6P substrate (**yellow** molecular surface) in the dimer are shown. The two monomers are shown in **cyan** and **green**. Right inset, close-up of the dimer interface and both structural $NADP^+$ molecules. Gomez-Manzo et al., ibid.

The second reaction of the oxidative branch is mediated by the phosphogluconolactonase (6PGL, PGLS) enzyme. The overall reaction is shown here:

2. 6-phosphogluconolactamase: 6-phosphogluconolactone + H_2O → 6-phosphogluconate

6PGL is a cytosolic enzyme found in all organisms that catalyzes the hydrolysis of 6-phosphogluconolactone to 6-phosphogluconic acid in the oxidative phase of the pentose phosphate pathway. 6PGL hydrolysis of 6-phosphogluconolactone to 6-phosphogluconic acid has been proposed to proceed via proton transfer to the O5 ring oxygen atom (Figure 13.4.8). The reaction initiates via the attack of a hydroxide ion at the C5 ester. A tetrahedral intermediate forms and the elimination of the ester linkage follows, aided by the donation of a proton from an active site histidine residue.

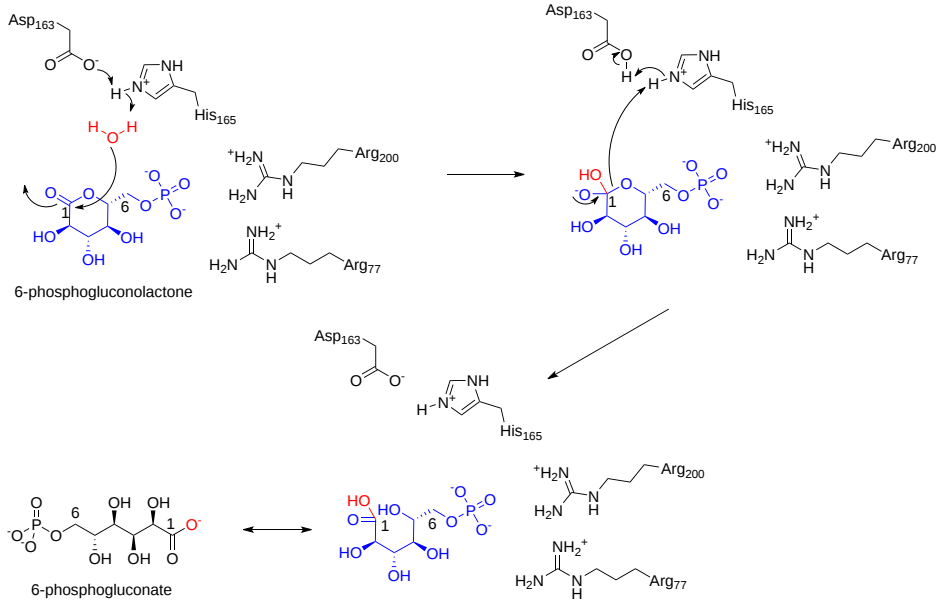


Figure 13.4.8 : Mechanism for6-phosphogluconolactamase His 165 and Asp 163 appear to be involved in a proton relay scheme. His 163 is conserved in lactonase. Arg 77 and 200 are involved in binding the substrate.

Figure 13.4.9 shows an interactive iCn3D model of the 6-phosphogluconolactonase from Trypanosoma brucei complexed with 6-phosphogluconic acid (3E7F). This enzyme is a target for developing drug treatment strategies for African sleeping sickness.

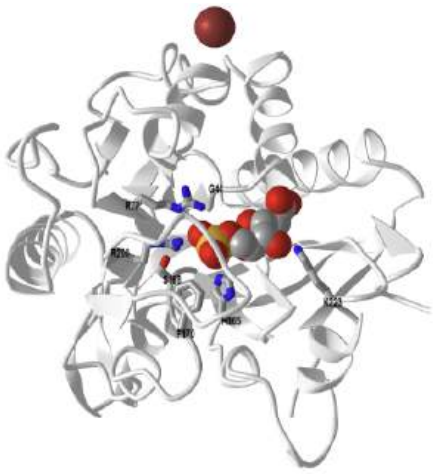
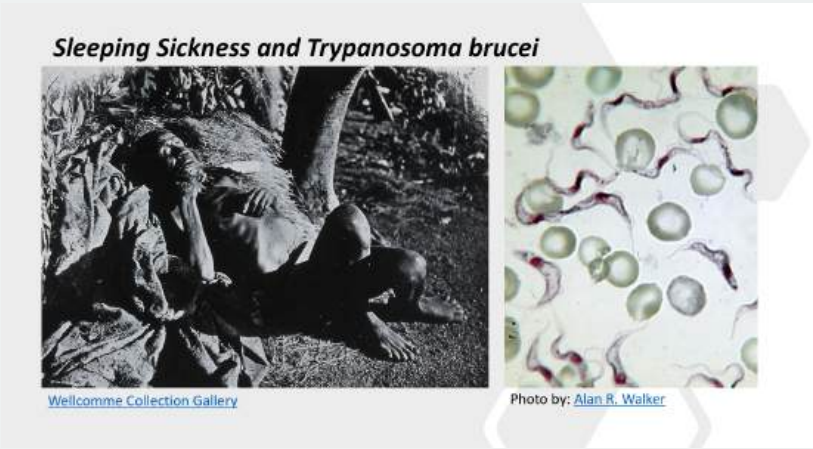


Figure 13.4.9 : 6-phosphogluconolactonase from Trypanosoma brucei complexed with 6-phosphogluconic acid (3E7F). The key residues involved in binding and the charge relay system are shown in sticks and labeled. The product, 6-phosphogluconic acid, is shown in spacefill. The single sphere is a Zn^{2+} which does not appear to be involved in catalysis. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?jRgi1nfrSGGgTYWo8

GLYOXOSOMES AND GLYCOSOMES



Trypanosomiasis is a disease usually referring to African human trypanosomiasis or African sleeping sickness. This infectious disease is caused by the parasites *Trypanosoma brucei gambiense or Trypanosoma brucei rhodesiense* cause this infectious disease, and the tsetse fly transmits the disease.

The vector tsetse fly, *Glossina*, carries the trypanosome within the midgut after a blood meal. These protozoa then migrate to the salivary glands of the fly whereby they can be transmitted during the next feeding. After inoculation within the host, the parasite can live freely within the bloodstream and evade mammalian host defenses through variable surface glycoproteins (VSG). The clinical disease has 2 stages. These are characterized by an early/first hemolymphatic stage and a late /second meningoencephalitis stage with an invasion of the central nervous system (CNS). In stage 1, systemic symptoms develop including intermittent fever, headache, pruritus, and lymphadenopathy. Undulating fevers reflect parasites multiplying within the blood. Less frequent hepatosplenomegaly may occur in the early stage. In the late/second stage, CNS symptoms manifest as sleep disturbances or neuropsychiatric disorders. A sleep disorder is the most common symptom of the second stage, and it is from this that the term “African sleeping sickness” was ascribed.

13.4.3: NEW DRUG TARGETS:

Enzymes involved in the pentose phosphate pathway provide a new set of drug targets to help combat parasitic infections such as those caused by trypanosomes. In this context, a focus for new therapeutics has been on the disruption of glycosomes within these organisms. Glyoxosomes are specialized peroxisomes found in plants and some fungi. They all have enzymes for the glyoxylate shunt and can hence create some glucogenic intermediates. In trypanosomes and Leishmania, another internal organelle called the glycosome is

found. This houses most of the glycolytic enzymes and also the oxidative enzymes of the pentose phosphate pathway. This makes the enzymes in the glycosome unique drug targets.

The third enzyme utilized in the oxidative branch of the PPP is 6-phosphogluconate dehydrogenase (6PGDH). The overall reaction is shown here:

3. 6-phosphogluconate dehydrogenase: 6-phosphogluconate + $NADP^+$ ↔ ribulose-5-phosphate + NADPH + CO_2

6PGDH catalyzes a reversible oxidative decarboxylation reaction, as shown in Figure 13.4.10. Oxidation, followed by decarboxylation forms an endiol intermediate followed by conversion to product. Notably, $NADP^+$ serves as an electron acceptor in the reaction, leading to the production of a second molecule of NADPH. Essentially, an active center lysine abstracts a proton from the substrate, 6-phosphogluconate, and NADP+ is reduced, causing the formation of a ketone-intermediate. Decarboxylation leads to the formation of the enediol intermediate. The active site lysine abstracts a proton in the second half of the reaction causing the formation of ribulose 5-phosphate. The abstracted proton is transferred to the active site Glu_{190} resetting the enzyme for another catalytic cycle.

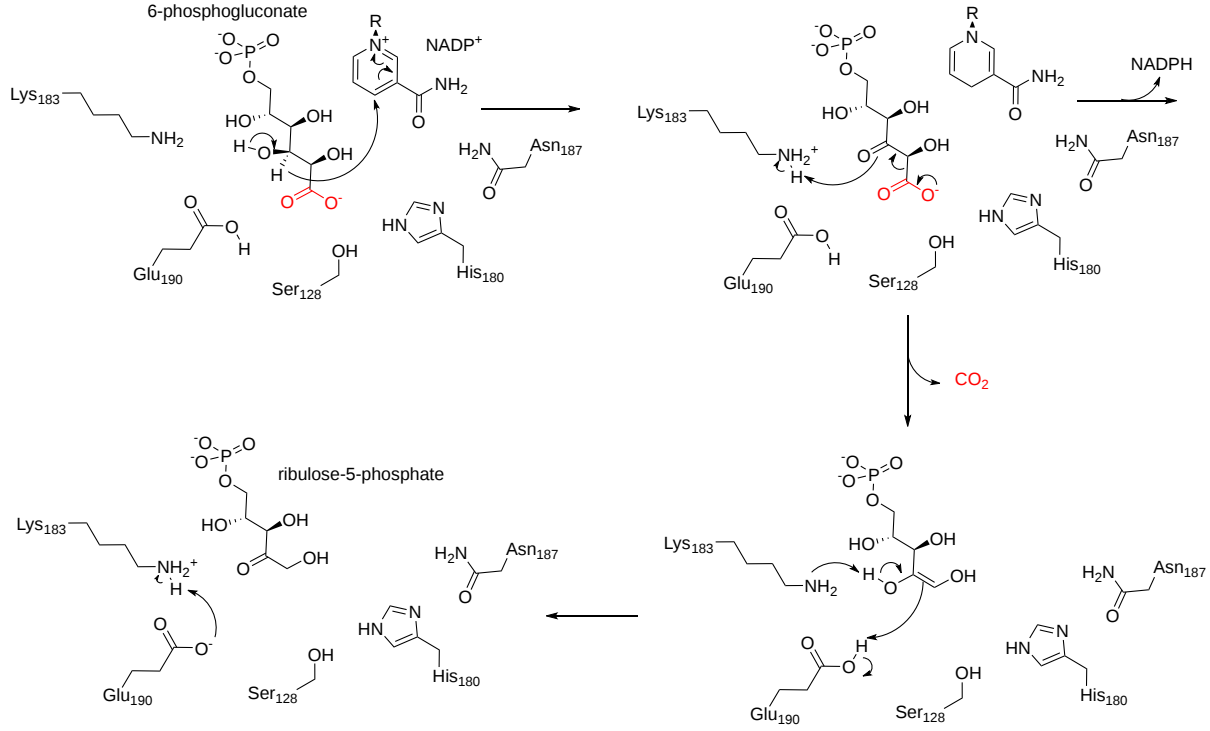
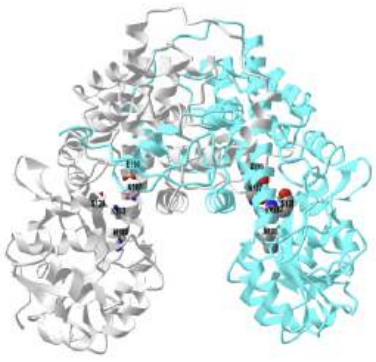


Figure 13.4.10 : Mechanism of 6-phosphogluconate dehydrogenase.

Figure 13.4.11 shows an interactive iCn3D model of sheep 6-phosphogluconate dehydrogenase (2PGD).



NCBI iCn3D **Figure 13.4.11 : Sheep 6-phosphogluconate dehydrogenase (2PGD).** (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...DqosWmFg1DqK16

The protein is a homodimer in which the monomers act independently: each contains a large, mainly alpha-helical domain and a smaller beta-alpha-beta domain, containing a mixed parallel and anti-parallel 6-stranded beta-sheet. $NADP^+$ is bound in a cleft in the small domain,

the substrate binding in an adjacent pocket.

13.4.4: NONOXIDATIVE BRANCH

The non-oxidative branch of the pentose phosphate pathway (nonox-PPP) is responsible for the formation of simple sugars within the cell and maintaining these building blocks in appropriate concentrations. The enzymes utilized in this pathway mediate a set of reversible reactions that lead to the production of ribose-5-phosphate, xylulose-5-phosphate, sedoheptulose 7-phosphate, and erythrose 4-phosphate, as well as intermediates utilized in the glycolytic pathway, including glyceraldehyde-3-phosphate and fructose-6-phosphate Major enzymes utilized include transketolase (TKT) and transaldolase (TALDO), as well as important isomerase and epimerase enzymes.

Ribose 5-phosphate is an important building block for the biosynthesis of nucleotides, and erythrose 4-phosphate is used in the synthesis of aromatic amino acids. When these simple sugars are in excess, they can also be converted into glycolytic intermediates and utilized for energy production.

The nonoxidative branch of the phosphopentose pathway is shown in Figure 13.4.12.

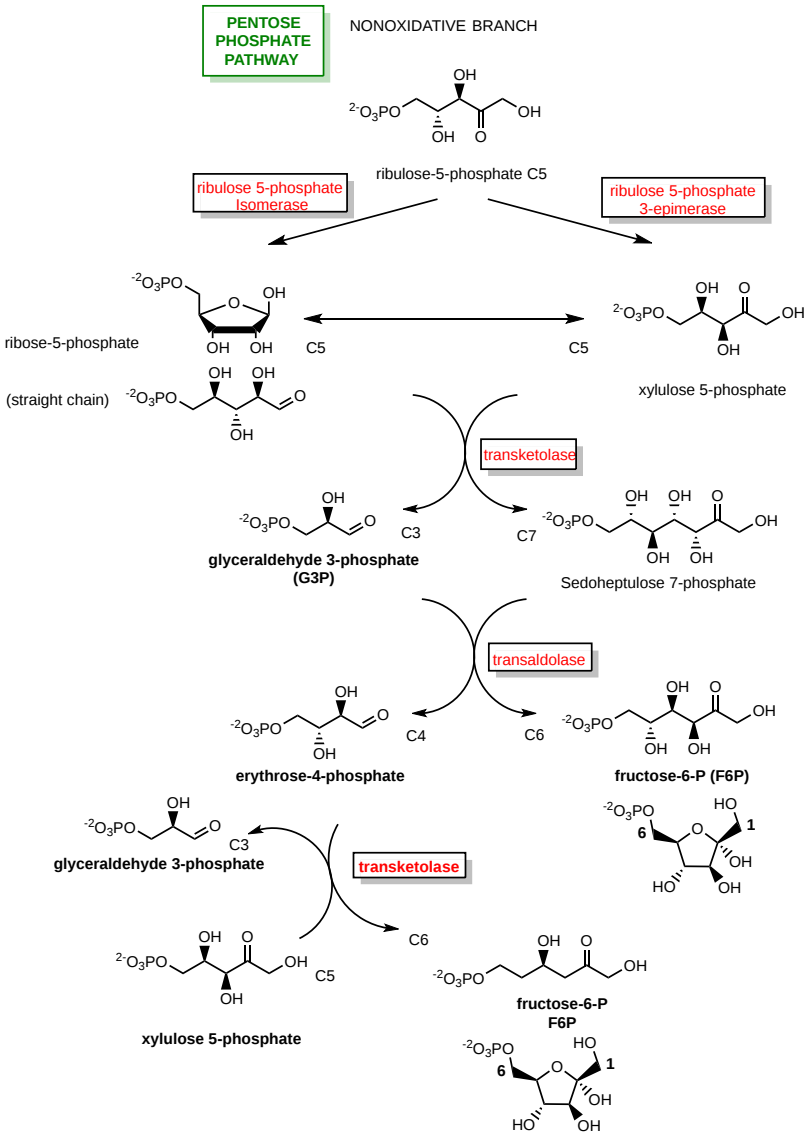


Figure 13.4.12 :Nonoxidative branch of the phosphopentose pathway

Now let's examine the individual enzymes. Ribulose 5-phosphate is the starting place for the non-oxidative portion of the PPP. It can be converted down two different pathways, either to ribose-5-phosphate or to xylulose-5-phosphate. The ribose 5-phosphate isomerase is involved in the ketose-aldose conversion and will be the first enzyme discussed.

Ribose-5-phosphate isomerase (Rpi)

This enzyme is fully reversible and is usually referred to as ribose-5-phosphate isomerase, although it can also be known as ribulose 5-phosphate isomerase, as it mediates the transition between this aldose-ketose pair. In addition to the PPP, this enzymatic reaction is also required in the Calvin cycle in photosynthesis.

There are two different Rpi enzymes, RpiA and RpiB, which have little sequence or structural similarities as well as different mechanisms. RpiA is found in all three kingdoms of life and is highly conserved due to its role in the PPP and the Calvin Cycle of photosynthesis. RpiB, on the other hand, is only found in some bacteria and protozoans. Thus, RpiB is a potential therapeutic target for the treatment of diseases such as African Sleeping Sickness, Chagas disease, and leishmaniasis. While RpiA and RpiB are structurally very different, both enzymes catalyze the isomerization reaction through an enediol intermediate using the linear form of the sugar. The reaction mechanism of RpiB is detailed below.

Figure 13.4.13 shows a mechanism for *Escherichia coli* RpiB.

Figure 13.4.13 : Mechanism for *Escherichia coli* RpiB. https://www.ebi.ac.uk/thornton-srv/m-csa/entry/680/. Creative Commons Attribution 4.0 International (CC BY 4.0) License.

RpiB is a homodimer that can form a tetramer. The monomer subunits are coordinated in a head-to-tail format, such that there are two active sites per dimer where side chains from each monomer contribute to each active site (Figure 13.4.14). The reaction mechanism requires two major steps: (1) ring opening of the furanose form of the sugar, and (2) isomerization through an enediol intermediate (Figure 13.4.13). Figure 13.4.14 shows an interactive iCn3D model of Ribose-5-phosphate isomerase (RPIB_AlsB) from Escherichia coli (1NN4).

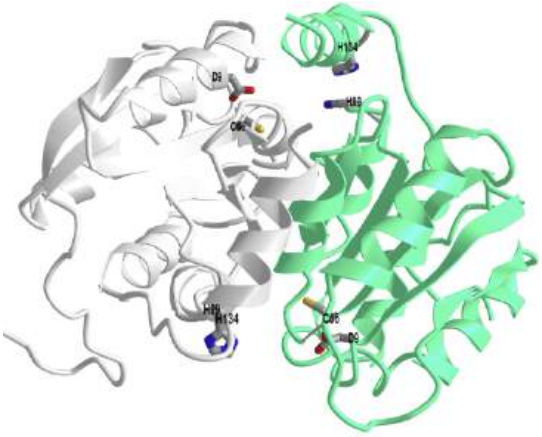


Figure 13.4.14 : Ribose-5-phosphate isomerase (RPIB_AlsB) from Escherichia coli (1NN4). Active site residues are indicated in each of the monomer subunits. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...ByuLUqhhs9B8c9

The other major pathway for ribulose 5-phosphate conversion involves the ribulose 5-phosphate (3) epimerase (RPE) and the conversion to xylulose 5 phosphate.

Ribulose-5-phosphate (3) epimerase (RPE)

This enzyme is also called pentose-5-phosphate 3-epimerase or PPE. Sugar epimerase enzymes alter the stereochemistry of the sugar at one of the alcohol positions creating a different epimer. The RPE enzyme has a Zn^{2+} cofactor, however, its requirement for enzyme activity is not essential. A Zn^{2+} independent form can still function and stabilize an oxyanion intermediate with adjacent and conserved methionines. A mechanism of the Zn^{2+} dependent form of the enzyme is shown in Figure 13.4.15.

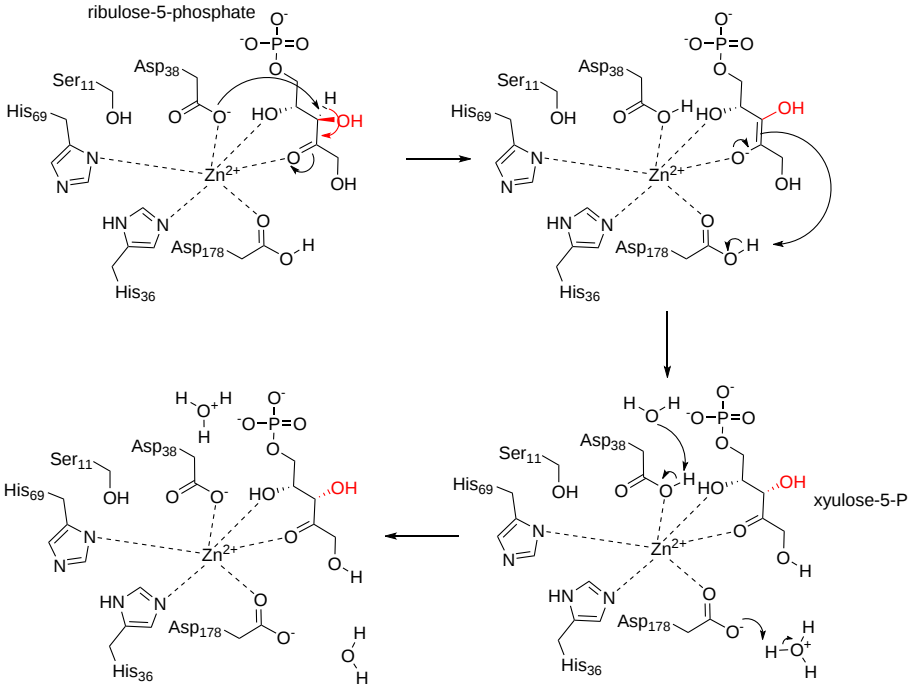
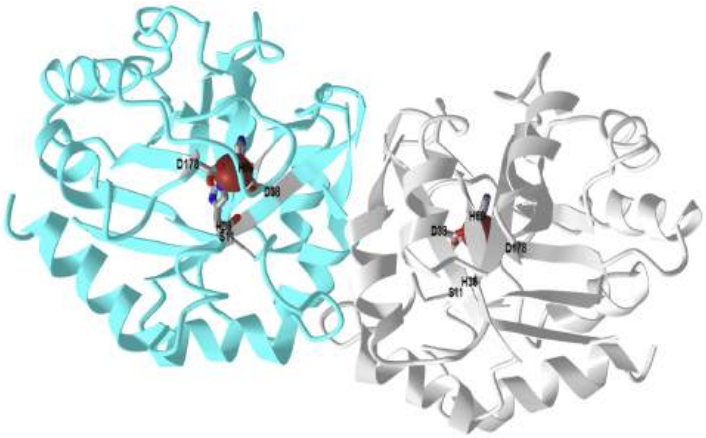


Figure 13.4.15 : Mechanism for rice ribulose-5-phosphate epimerase

The enzyme utilizes an acid/base catalytic mechanism that mediates the formation of a trans-2,3-enediol phosphate intermediate. Key aspartic acid residues act as proton donors and acceptors during the reaction. A zinc metal cofactor helps stabilize charges during the reaction.

Figure 13.4.16 shows an interactive iCn3D model of cytosolic D-ribulose-5-phosphate 3-epimerase from rice (1h1z).



NCBI iCn3D **Figure 13.4.16 : Cytosolic D-ribulose-5-phosphate 3-epimerase from rice (1h1z).** (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...AH7uS2T8hsnpXA

Transketolase (TK)

The pool of ribose 5-phosphate that is created can be utilized for nucleotide production, or some of it can undergo a reaction with the xylulose 5-phosphate created in the RPE reaction. The transketolase enzyme converts two five-carbon sugars (ribose 5-phosphate and

xylulose 5-phosphate) into a 3-carbon and a 7-carbon sugar (glyceraldehyde 3-phosphate and sedoheptulose 7-phosphate). Essentially, transketolase enzymes transfer ketone functional groups from ketoses to aldoses, effectively creating a new ketose that is two carbons larger. The ketose donor then becomes an aldose with two fewer carbons. The basic transketolase reaction shown in Figure 13.4.17

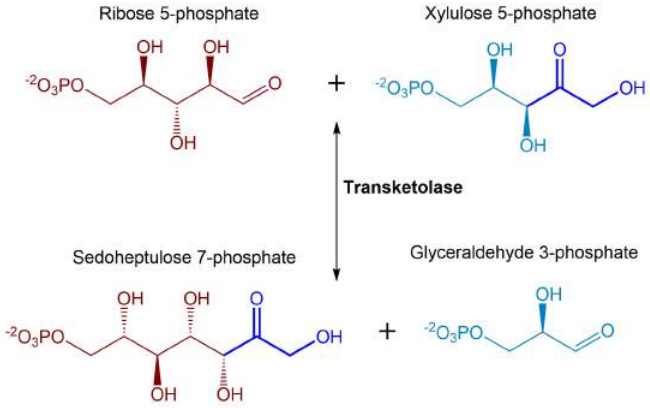


Figure 13.4.17 : Transketolase reaction. The reaction shows the reversible conversion of ribose 5-phosphate and xylulose 5-phosphate to sedoheptulose 7-phosphate and glyceraldehyde 3-phosphate.

Note the number of carbons in the reactants and products: 5C + 5C ↔ 3C + 7C. In this reversible reaction, the enzyme uses the cofactor thiamine pyrophosphate (TPP) and a divalent cation. The enzyme transfers a 2C ketol group from xylulose-5-phosphate to ribose-5-phosphate, an aldose. The product glyceraldehyde-3-phosphate is a glycolytic intermediate and can be used in the glycolytic pathway.

The mechanism for the reverse reaction (yeast numbering system) is shown in Figure 13.4.18. In this reaction, sedoheptulose 7-phosphate binds with the enzyme and the TPP cofactor is activated to form a carbanion. The carbanion mediates nucleophilic attack on the carbonyl carbon of the substrate forming a covalent intermediate. His_{263} serves as a base and abstracts a proton, enabling bond cleavage and the formation of ribose 5-phosphate. Ribose 5-phosphate leaves the active site and glyceraldehyde 3-phosphate enters. The two carbon intermediate covalently bound to the TPP mediates nucleophilic attack on the glyceraldehyde 3-phosphate enabling the formation of xylulose 5-phosphate and the restoration of the enzyme.

Figure 13.4.18 : Mechanism of the transketolase reaction during the pentose phosphate pathway.
https://www.ebi.ac.uk/thornton-srv/m-csa/entry/219/

Figure 13.4.19 shows an interactive iCn3D model of Human transketolase in a covalent complex with donor ketose D-xylulose-5-phosphate (4kxv).

Figure 13.4.19 : Human transketolase in covalent complex with donor ketose D-xylulose-5-phosphate (4kxv). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...WtGtQ7mKA6UYx7

The enzyme is a dimer (gray and cyan coloring). Numbering of key residues compared to yeast mechanism:

Human	Yeast
E366	E418
Q428	H481
H258	H262
H38	H30

Also shown are human active site residues H474, R318, and S345. The covalent xylulose-5-phosphate adduct is shown in spacefill and labeled DX5 (xylitol-5-phosphate). TPP is shown in sticks and labeled. The metal ion is Ca^{2+}. Structural analyzes show a 20^0 distortion in the planarity of the cofactor-substrate bond and a lengthening of the C-C bond in the substrate which breaks.

Transaldolase

In addition to being used as a resource in the glycolytic pathway, glyceraldehyde 3-phosphate can also be utilized as a substrate in a transaldolase reaction along with the sedoheptulose 7-phosphate that is produced in the previous reaction. This results in the formation of erythrose 4-phosphate and fructose 6-phosphate. A summary of the reaction is shown in Figure 13.4.20.

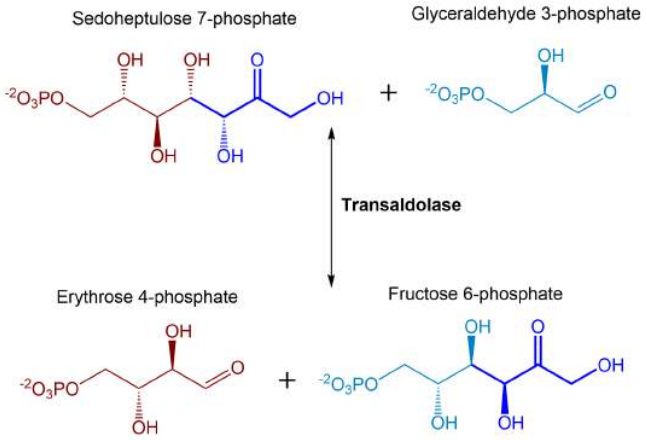


Figure 13.4.20 : Summary of transaldolase reaction

As with transketolase, the transaldolase enzyme is reversible. Again, note the number of carbons in the reactants and products: 3C + 7C ↔ 4C + 6C. Unlike the transketolase used in the last reaction, the transaldolase enzyme does NOT use TPP as a cofactor. Instead, it forms a Schiff base intermediate similar to that of the aldolase enzyme in the glycolytic pathway. The enzyme removes a 3C ketol group

(dihydroxyacetone) from sedoheptulose 7-phosphate and transfers it to glyceraldehyde 3-phosphate forming fructose 6-phosphate. Erythrose 4-phosphate is left from the original sedoheptulose 7-phosphate.

We will explore the mechanism for the *E. Coli* enzyme in the reverse direction. The first part of the mechanism of transaldolase is shown in Figure 13.4.21.

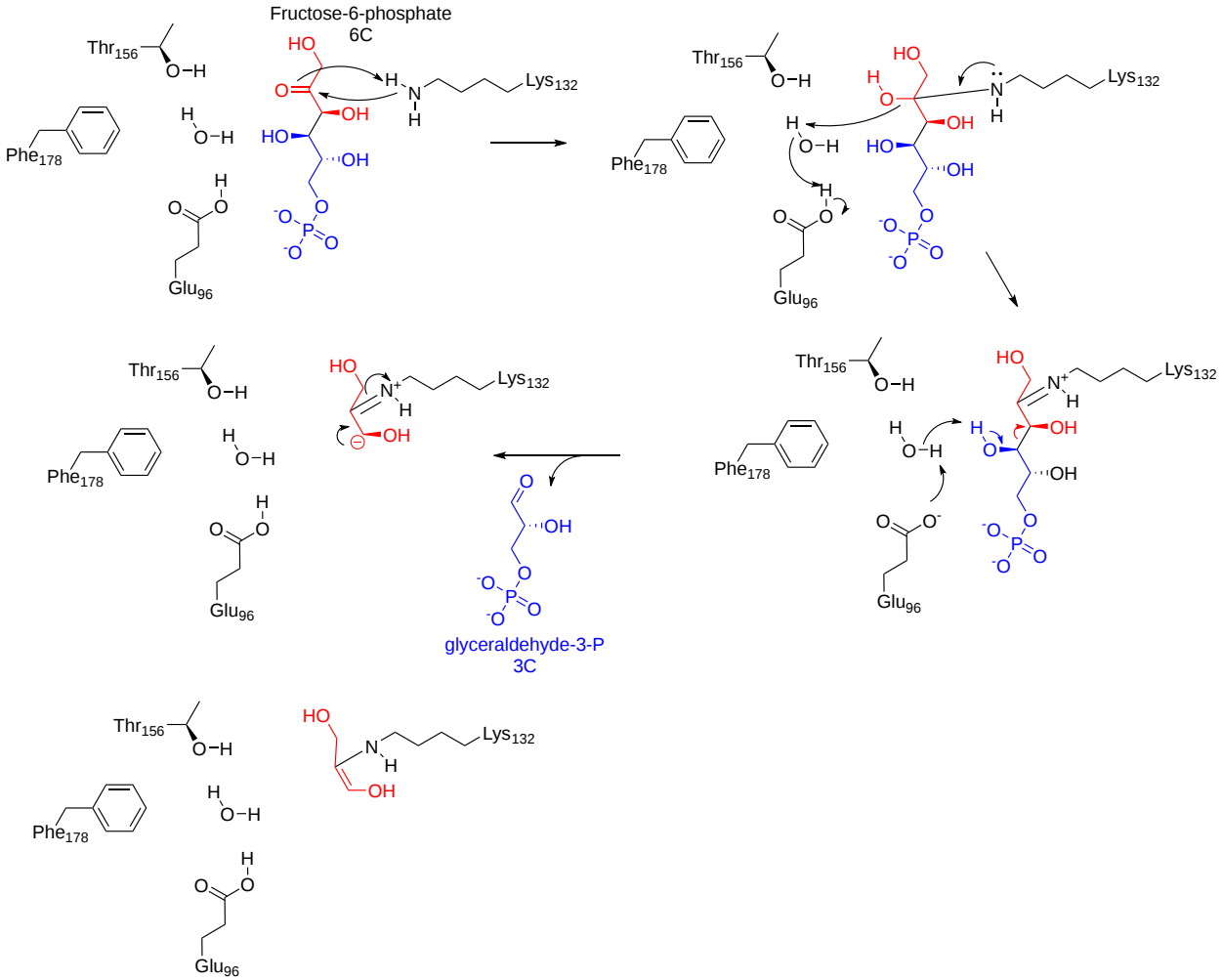


Figure 13.4.21 : Mechanism for the first half of the transaldolase reaction. https://www.ebi.ac.uk/thornton-srv/m-csa/entry/148/

In the first part of the reaction, fructose 6-phosphate binds to the active site of the enzyme, where an active site lysine residue mediates nucleophilic attack on the carbonyl carbon and forms a covalent intermediate with the enzyme. Formation of the Schiff base leads to dehydration of the intermediate. The Schiff base nitrogen becomes protonated and this leads to the oxidation of the C4 hydroxyl leads and subsequent bond cleavage releasing glyceraldehyde 3-phosphate. The remaining ES-complex rearranges to form an enol intermediate.

The mechanism for the second half of the transaldolase reaction is shown in Figure 13.4.22.

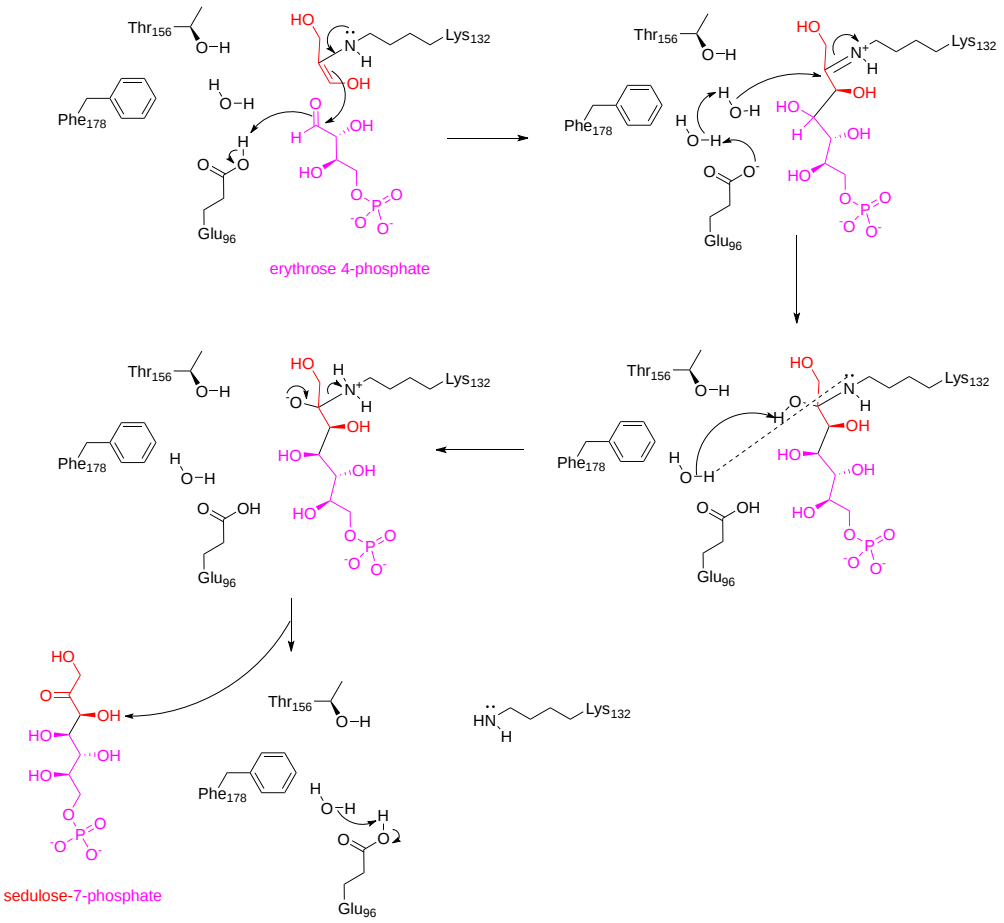


Figure 13.4.22 : Mechanism for the second half of the transaldolase reaction. https://www.ebi.ac.uk/thornton-srv/m-csa/entry/148/

Once glyceraldehyde 3-phosphate has left the active site of the enzyme, erythrose 4-phosphate can bind. The enol from the ES intermediate mediates nucleophilic attack on the aldehyde carbonyl group of erythrose 4-phosphate. This results in the formation of a Schiff base intermediate. Hydration at the Schiff base carbon atom ensues followed by the oxidation of the newly incorporated alcohol to form a ketone functional group. Formation of the ketone causes bond cleavage between the enzyme and the newly formed ketose, sedoheptulose 7-phosphate.

Figure 13.4.23 shows an interactive iCn3D model of transaldolase B from Escherichia coli (1ONR).

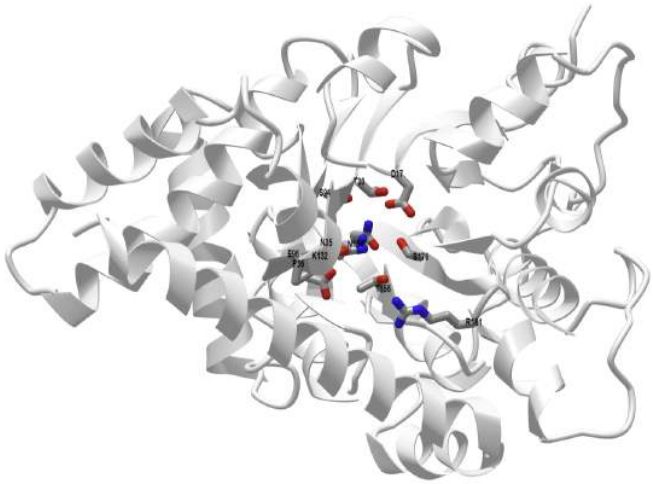


Figure 13.4.23 : 3D structure of Transaldolase B from E. coli (1ONR). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...7swLZVpNPe3eU9

Transketolase

As seen above, fructose 6-phosphate is a product of the transaldolase enzyme during the nonoxidative branch of the PPP. This sugar can feed back into the glycolytic cycle. The formation of fructose 6-phosphate can also occur using the transketolase enzyme when it transfers a 2C ketol group from xylulose 5-phosphate to the aldose, erythrose 4-phosphate. This reaction is summarized in Figure 13.4.24.

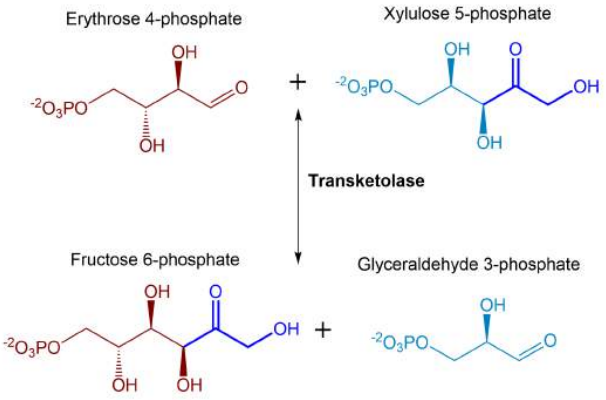


Figure 13.4.24 Summary of second transketolase reaction during the pentose phosphate pathway.

Both products of this final transketolase reaction can be utilized in energy formation through glycolysis. Thus, there are clear metabolic ties of the PPP with the energy-producing glycolytic pathway and intermediates from both pathways can easily be exchanged through the activity of these reversible enzymes. The PPP provides key intermediates, including ribose 5-phosphate that is used in the biosynthesis of nucleotides and ultimately nucleic acids, as well as, other important metabolic molecules such as FAD, NAD^+, $NADP^+$, and Coenzyme A. In addition to being a biosynthetic source for these molecules, the PPP also is the primary source for reducing $NADP^+$ to NADPH, during the oxidative reactions. NADPH is utilized in many biosynthetic processes as an electron donor. For example, the biosynthesis of many lipid molecules requires NADPH, including the production of triacylglycerols, phospholipids, and steroids. NADPH is also required for the biosynthesis of some amino acids (such as glutamate and proline) and is also involved in the reduction of ribonucleotides and deoxyribonucleotides during the synthesis of RNA and DNA. NADPH is also utilized by a number of oxidoreductases involved in detoxification reactions within the body. Within the immune system, NADPH oxidases or NOX enzymes are involved in the production of superoxide and utilized to damage invading pathogens nonspecifically. $NADP^+$/NADPH ratios also can play a regulatory role in cellular metabolic processes and are utilized as allosteric effectors for several enzymes and oxidation sensor proteins.

13.4.5: REGULATION OF THE PENTOSE PHOSPHATE PATHWAY

The cellular demand for the two major products of the PPP (ribose 5-phosphate and reduced NADPH) can be different depending on the cell type or the current cellular environment, such as times of increased metabolic demand or oxidative stress. Thus, the two major products may need to be produced in different quantities and independently of one another. For example, we can imagine times when the needs for ribose 5-phosphate and NADPH are in balance with the standard PPP reactions. However, we can also imagine times when the demand for ribose 5-phosphate may be much higher than the demand for NADPH, or vice versa, the demand for NADPH may be much greater than the demand for ribose 5-phosphate. Thus, there are regulatory strategies in place that enable the ability to regulate the production of these different pools independently of one another and adapt to cellular needs. This ability is largely dependent on the production of metabolic intermediates that can be easily interchanged within the glycolytic pathway.

For example, the ratio of $NADP^+$/NADPH serves as a key regulator of the oxidative branch of the PPP. The first enzymatic step of the pathway mediated by the glucose 6-phosphate dehydrogenase (G6PD) is regulated in this fashion and helps control the pool of glucose 6-phosphate that will be utilized within the PPP to produce reduced NADPH. Low levels of $NADP^+$ inhibit the G6PD enzyme. The G6PD reaction is essentially irreversible and serves as the committed step for glucose to enter into the oxidative portion of the PPP. Thus the regulation of this enzymatic step is key in the regulation of NADPH levels within the cell. When the ratio of $NADP^+$/NADPH increases, G6PD becomes more active and the reduction of $NADP^+$ to NADPH increases.

When ribose 5-phosphate is also in high demand, the pool of ribose 5-phosphate will be low and increase the activity of the ribose 5-phosphate isomerase in the forward direction to convert the ribulose 5-phosphate produced in the oxidative branch to generate more ribose 5-phosphate.

However, if more ribose 5-phosphate is required than can be supplied from the activity of the oxidative branch, the pool of ribose 5-phosphate can be fed from the conversion of intermediates from the glycolytic pathway in the nonoxidative branch of the PPP. The opposite occurs as well when there is a high need for NADPH in the cell but a low need for ribose 5-phosphate. In this situation, ribose 5-phosphate is converted to fructose and glyceraldehyde 3-phosphate and can be incorporated into the glycolytic pathway.

In addition, the activity of the PPP varies depending on the tissue type and location within the body. For example, skeletal muscle has very low PPP activity, as this tissue requires more energy production and activity of the glycolytic pathway. On the other hand, the PPP is highly

active in adipose tissue due to the heightened requirement for intermediates needed for lipid biosynthesis.

13.4.6: REFERENCES:

Edwards et al.: Structural characterization of a ribose5-phosphate isomerase B from the pathogenic fungus Coccidioides immitis. BMC Structural Biology 2011 11:39. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3212906/

Ge T, Yang J, Zhou S, Wang Y, Li Y and Tong X (2020) The Role of the Pentose Phosphate Pathway in Diabetes and Cancer. *Front. Endocrinol.* 11:365. doi: 10.3389/fendo.2020.00365

Polat, I.H., Tarrado-Castellarnau, M., Bharat, R., Perarnau, J., Benito, A., Cortes, R., Sabatier, P., and Cascante, M. (2021) Oxidative Pentose Phosphate Pathway Enzyme 6-Phosphogluconate Dehydrogenase Plays a Key Role in Breast Cancer Metabolism. *Biology* 10(2):85. **https://doi.org/10.3390/biology10020085**

Wikipedia contributors. (2022, March 6). Glucose-6-phosphate dehydrogenase. In *Wikipedia*, *The Free Encyclopedia*. Retrieved 22:01, August 12, 2022, from https://en.wikipedia.org/w/index.php?title=Glucose-6-phosphate_dehydrogenase&oldid=1075517394

Wikipedia contributors. (2021, August 29). 6-phosphogluconolactonase. In *Wikipedia*, *The Free Encyclopedia*. Retrieved 23:05, August 12, 2022, from https://en.wikipedia.org/w/index.php?title=6-phosphogluconolactonase&oldid=1041297267

Wikipedia contributors. (2022, March 28). 6-Phosphogluconate dehydrogenase. In *Wikipedia*, *The Free Encyclopedia*. Retrieved 05:50, August 14, 2022, from https://en.wikipedia.org/w/index.php?title=6-Phosphogluconate_dehydrogenase&oldid=1079702040

Wikipedia contributors. (2022, July 4). Ribose-5-phosphate isomerase. In *Wikipedia*, *The Free Encyclopedia*. Retrieved 05:33, August 15, 2022, from https://en.wikipedia.org/w/index.php?title=Ribose-5-phosphate_isomerase&oldid=1096406233

CHAPTER OVERVIEW

14: PRINCIPLES OF METABOLIC REGULATION

14.1: Regulation of Metabolic Pathways
14.2: Basic Principles of Metabolic Control Analysis (MCA)
14.3: The Flux Control Coefficient
14.4: Concentration Control and Elasticity Coefficients
14.5: Metabolism and Signaling: The Steady State, Adaptation and Homeostasis

Thumbnail: Overview of regulatory interactions involved in metabolic regulatory networks. The function of metabolic networks are governed by constraints. The regulation of a metabolic network involves a tight interplay between different cellular networks such as signalling and gene networks and by interactions with its environment. The enzyme capacity is the net result of the amount of enzyme expressed and its activity as dictated by post-translational modification and allosteric regulation. Metabolite pools and fluxes are considered as the outputs of metabolic reaction networks and can be involved in various regulatory feedback loops to other networks within the metabolic reaction networks as indicated by the dashed arrows. (CC BY 3.0; Jan Berkhout, Frank J. Bruggeman, and Bas Teusink via MDPI)

14.1: REGULATION OF METABOLIC PATHWAYS

Exquisite mechanisms have evolved that control the flux of metabolites through metabolic pathways to insure that the output of the pathways meets biological demand and that energy in the form of ATP is not wasted by having opposing pathways run concomitantly in the same cell.

Enzymes can be regulated by changing the activity of a preexisting enzyme or changing the amount of an enzyme.

14.1.1: CHANGING THE ACTIVITY OF A PRE-EXISTING ENZYME

The quickest way to modulate the activity of an enzyme is to alter the activity of an enzyme that already exists in the cell. The list below, illustrated in the following figure, gives common ways to regulate enzyme activity

1. Substrate availability: Substrates (reactants) bind to enzymes with a characteristic affinity (characterized by a dissociation constant) and a kinetic parameter called Km (units of molarity). If the actual concentration of a substrate in a cell is much less than the Km, the activity of the enzyme is very low. If the substrate concentration is much greater than Km, the enzyme's active site is saturated with substrate and the enzyme is maximally active.
2. Product inhibition: A product of an enzyme-catalyzed reaction often resembles a starting reactant, so it should be clear that the product should also bind to the activity site, albeit probably with lower affinity. Under conditions in which the product of a reaction is present in high concentration, it would be energetically advantageous to the cell if no more product was synthesized. Product inhibition is hence commonly observed. Likewise, it is energetically advantageous to a cell if the end product of an entire pathway could likewise bind to the initial enzyme in the pathways and inhibit it, allowing the whole pathway to be inhibited. This type of feedback inhibition is commonly observed. Figure 14.1.1 shows product and end-product inhibition.

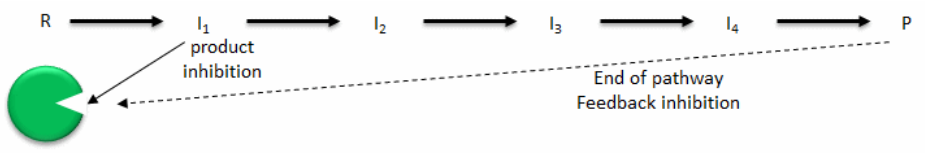


Figure 14.1.1 : Product and end product inhibition of an enzyme

3. Allosteric regulation: As many pathways are interconnected, it would be optimal if the molecules of one pathway affected the activity of enzymes in another interconnected pathway, even if the molecules in the first pathway are structurally dissimilar to reactants or products in a second pathway. Molecules that bind to sites on target enzymes other than the active site (allosteric sites) can regulate the activity of the target enzyme. These molecules can be structurally dissimilar to those that bind at the active site. They do so by conformational changes which can either activate or inhibit the target enzyme's activity.
4. pH and enzyme conformation: Changes in pH that can accompany metabolic processes such as respiration (aerobic glycolysis for example) can alter the conformation of an enzyme and hence enzyme activity. The initial changes are covalent (change in the protonation state of the protein) which can lead to an alteration in the delicate balance of forces that affect protein structure.
5. pH and active site protonation state: Changes in pH can affect the protonation state of key amino acid side chains in the active site of proteins without affecting the local or global conformation of the protein. Catalysis may be affected if the mechanism of catalysis involves an active site nucleophile (for example), that must be deprotonated for activity.
6. Covalent modification: Many if not most proteins are subjected to post-translational modifications which can affect enzyme activity through local or global shape changes, by promoting or inhibiting binding interaction of substrates and allosteric regulators, and even by changing the location of the protein within the cell. Proteins may be phosphorylated, acetylated, methylated, sulfated, glycosylated, amidated, hydroxylated, prenylated, or myristoylated, often in a reversible fashion. Some of these modifications are reversible. Regulation by phosphorylation through the action of kinases, and dephosphorylations by phosphates are extremely common. Control of the phosphorylation state is mediated through signal transduction processes starting at the cell membrane, leading to the activation or inhibition of protein kinases and phosphatases within the cell.

Figure 14.1.2 shows ways to regulate the activity of pre-existing enzymes.

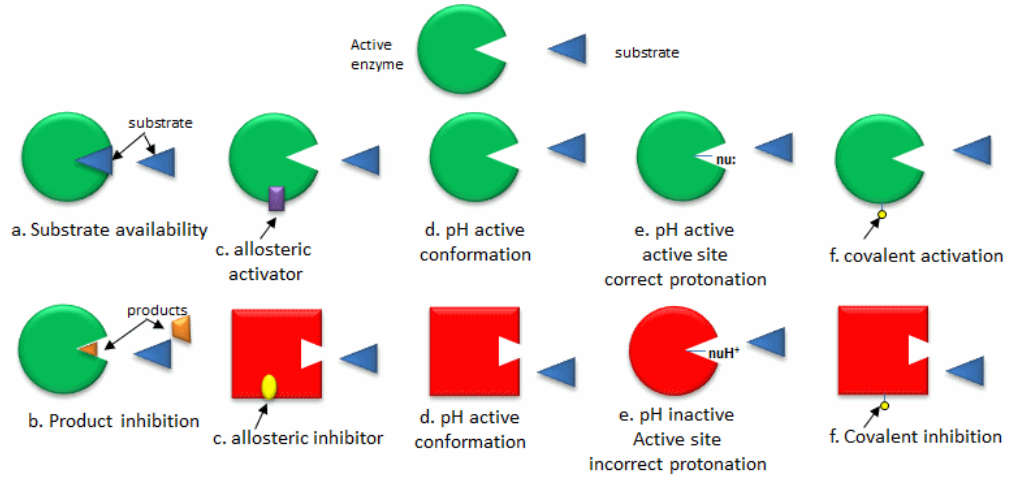
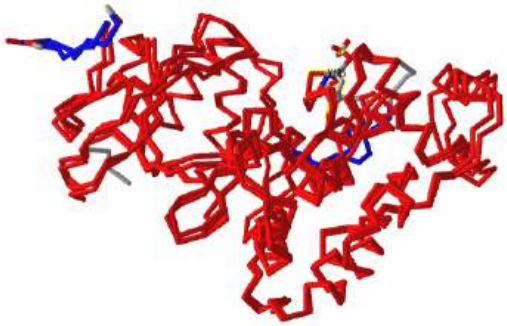


Figure 14.1.2: Ways to regulate the activity of pre-existing enzymes

Extracellular regulated kinase 2 (ERK2), also known as mitogen-activated protein kinase 2 (MAPK2) is a protein that plays a vital role in cell signaling across the cell membrane. Phosphorylation of ERK2 on Threonine 183 (Thr153) and Tyrosine 185 (Tyr185) leads to a structural change in the protein and the regulation of its activity.

Figure 14.1.3 shows an interactive iCn3D model showing the structural alignment of ERK2 in the dephosphorylated (5UMO) and phosphorylated (pY185) forms (2ERK). Toggle back and forth between the two structures with the "a" key.



NCBI iCn3D Figure 14.1.3: Structural alignment of ERK2 in the dephosphorylated (5UMO) and phosphorylated (pY185) forms (2ERK). (Copyright; author via source). Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...8iMQUVBbmMTRJ6

The residues that change significantly in conformation on phosphorylation are shown in blue. The side chain of tyrosine 185 in the unphosphorylated form is shown in CPK-colored sticks and labeled.

14.1.2: REGULATION OF SINGLE ENZYMES OR ENTIRE PATHWAYS: ENZYME CONDENSATES

Single enzymes or all the enzymes of a given pathway can be coordinately regulated to maximize end-product output by organizing the enzymes in one large complex built from soluble enzymes that produce a "condensate" through a process similar to phase separation. Such condensates are shown for a series of enzymes in Figure 14.1.4.

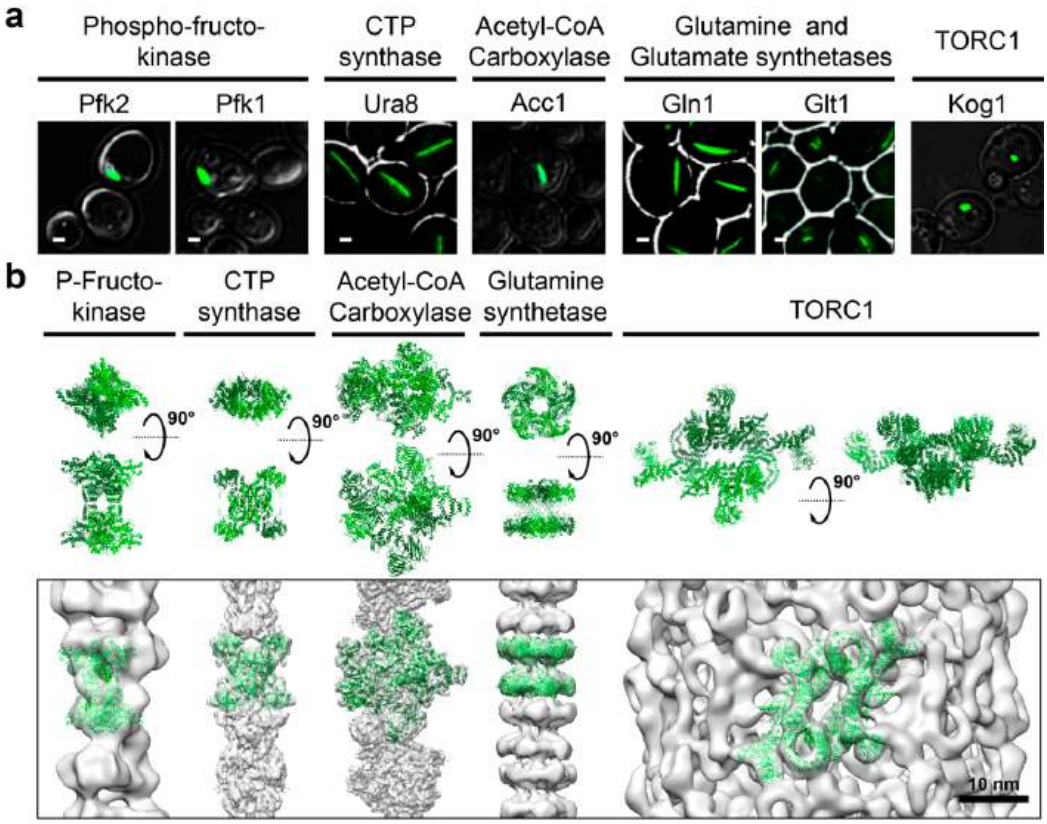


Figure 14.1.4: Supramolecular assembly of enzyme condensates. **Prouteau and Loewith.** *Biomolecules* **2018**, *8*(4), 160; **https://doi.org/10.3390/biom8040160.** Creative Commons Attribution (CC BY) license (**http://creativecommons.org/licenses/by/4.0/**).

The figure shows metabolism-related enzymes that form polymers in various organisms.

Panel (**a**) shows examples of metabolic enzymes observed to coalesce into cytosolic condensates.

Panel (**b**) top shows structures of metabolic enzymes that polymerize into filaments. The protomer of the polymer is shown above and placed into the filament below. These include P-Fructo-Kinase (4XYJ), cytidine triphosphate synthase (5U03), acetyl-CoA carboxylase (6G2D), glutamine synthetase (3FKY), mTORC1 (5FLC) from PDB files.

Panel (**b**) bottom shows the same structures from the Electron Microscopy Data Bank, including P-Fructo-Kinase filament (emd-8542), human CTP synthase filament (and-8474), human acetyl-CoA carboxylase with citrate (emd-4342), and the yeast glutamine synthetase filament.

14.1.3: CHANGING THE AMOUNT OF AN ENZYME

Another longer-duration method to modulate the activity of an enzyme is to alter the activity of an enzyme that already exists in the cell. Figure 14.1.5 shows ways in which enzyme concentrations are regulated.

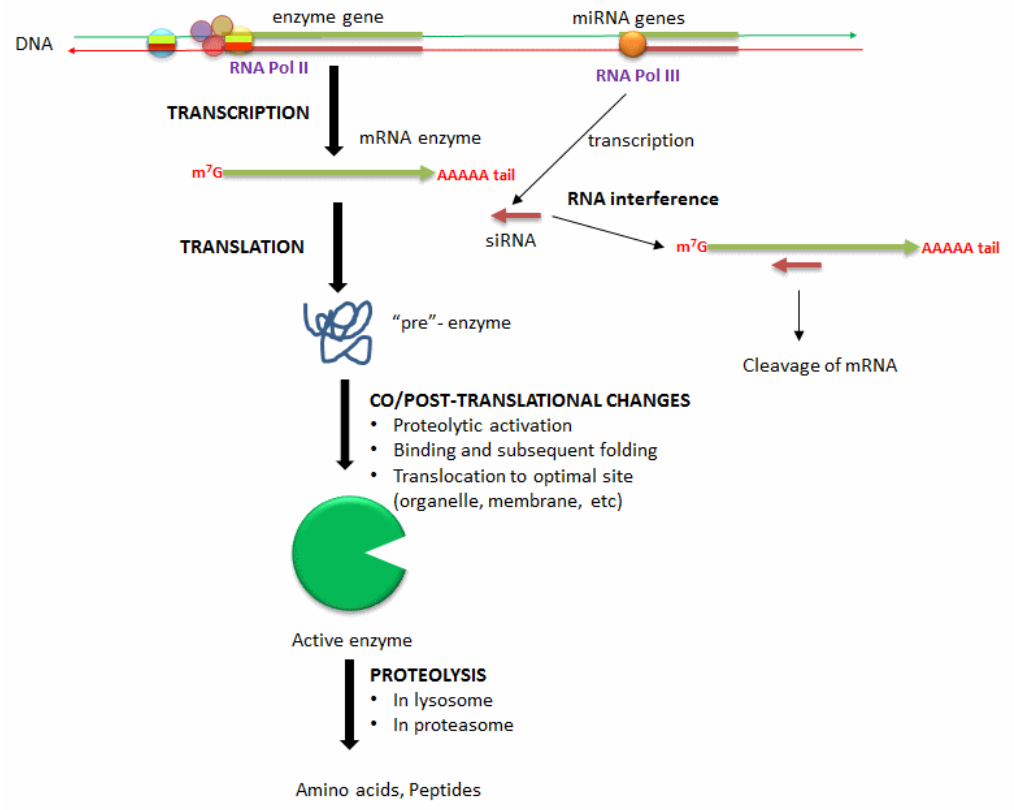


Figure 14.1.5: Ways to regulate the concentration of an enzyme

Methods include:

1. Alterations in the transcription of enzyme's gene: Extracellular signals (hormones, neurotransmitters, etc) can lead to signal transduction responses and ultimate activation or inhibition of the transcription of the gene for a protein enzyme. These changes result from the recruitment of transcription factors (proteins) to DNA sequences that regulate the transcription of the enzyme gene.
2. Degradation of messenger RNA for the enzyme: The levels of messenger RNA for a protein will directly determine the amount of that protein synthesized. Small inhibitor RNAs, derived from microRNA molecules transcribed from cellular DNA, can bind to specific sequences in the mRNA of a target enzyme. The resulting double-stranded RNA complex recruits an enzyme (Dicer) that cleaves the complex with the effect of decreasing the translation of the protein enzyme from its mRNA.
3. Co/Post-translational changes: Once a protein enzyme is translated from its mRNA, it can undergo many changes that regulate its activity. Some proteins are synthesized in a "pre" form which must be cleaved in a targeted and limited fashion by proteases to activate the protein enzyme. Some proteins are not fully folded and must bind to other factors in the cell to adopt a catalytically active form. Finally, fully active protein can be fully proteolyzed by the proteasome, a complex within cells, or in lysosomes, which are organelles within cells containing proteolytic enzymes.

All proteins are ultimately regulated, if only by modulating the rates of their synthesis and degradation. However, some enzymes positioned at key points in metabolic pathways are ideal candidates for regulation, as their activity can affect the output of entire pathways. These enzymes typically have two common characteristics, they catalyze reactions far from equilibrium and they catalyze early committed steps in pathways.

14.1.4: WHICH ENZYMES TO REGULATE: REACTIONS NOT AT EQUILIBRIUM

The optimal enzymes for regulation are those at the beginning of pathways that carry out thermodynamically favored reactions. Why is the latter so important? These enzymes control the flux of metabolites through pathways, so to understand their regulation we can use the analogy of flow (or flux) of water from one container to another as illustrated in Figure 14.1.6.

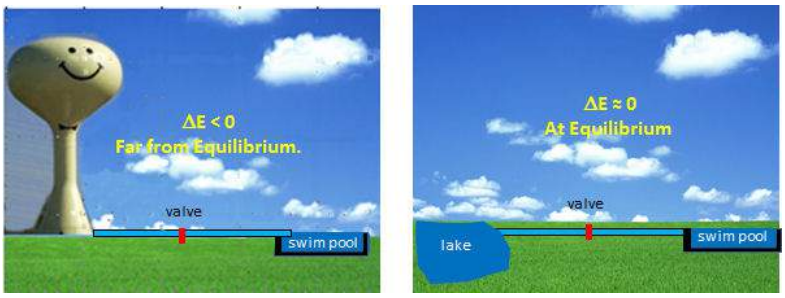


Figure 14.1.6: Regulation of water flow in pipe

Let's say you wish to fill a swimming pool at any desired height you wish and you have two ways to do so (see figure below). You could open a valve that controls the flow from your town's water tower to the pool. In this, the reaction (flow of water) is energetically (thermodynamically) favored given the difference in height of the water levels and the potential energy difference between the two. Even though flow (or flux) is cleared flavored, you can regulate it, from no flow to maximal flow, by opening and closing the valve (analogous to activating and inhibiting an enzyme). Your choices in the other scenario, filling the pool from a lake, are not so great. It would be hard to fill the water to the desired level (especially if it was an above-ground pool). It would be hard to regulate the flow.

By analogy, the best candidates for regulation are those enzymes whose reactions are thermodynamically favored (not at equilibrium) but which can be controlled by the mechanisms discussed in the previous section.

Which reactions are commonly not at equilibrium (i.e. $\Delta G<0$ and usually also $\Delta G^0<0$ if the ratio of products to reactants is not too high)? The answer is those that have reactants that are thermodynamically unstable compared to their reaction products. Several types of reactions often fit these criteria:

Hydrolysis (or similar reactions) of anhydride or analogous motifs: The figure below shows molecules with similar "anhydride" motifs and the ΔG^0 for hydrolysis of the molecules. Those with more negative ΔG^0 values can transfer their phosphate group to ADP to make ATP, which is necessary to drive unfavorable biological reactions. Metabolic reactions that involve hydrolysis (or other types of transfer reaction of these groups) usually proceed with a negative ΔG^0 and ΔG, making them prime candidates for pathway regulation. Many textbooks label these types of molecules as having "high energy" bonds. This is confusing to many students as bonds between atoms lower the energy compare to when the atoms are not bonded. It takes energy to break the "high" energy phosphoanhydride covalent bond. What make hydrolysis of the molecules below so exergonic is that more energy is released on bond formation within the new products than was required to break the bonds in the reactants. In addition, other effects such as preferential hydration of the products, lower charge density in the products, and less competing resonances in the products all contribute to the thermodynamically favorable hydrolysis of the reactants.

Figure 14.1.7 shows thermodynamically unstable molecules (compared to their reaction products in aqueous solutions).

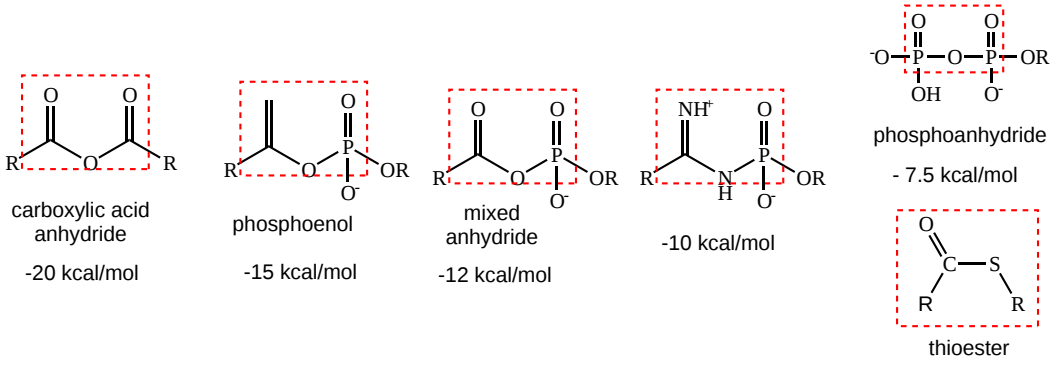


Figure 14.1.7: Molecules that are thermodynamically unstable compared to their hydrolysis products

Thioesters (such as Acetyl-SCoA) are also included as they have the same negative ΔG^0 of hydrolysis as ATP, even though they lack an "anhydride" motif. Thioesters are destabilized compared to their hydrolysis products and in comparison to esters made with alcohol since the C-S bond is weaker. Why?

Redox reactions: Everyone knows that redox reactions are thermodynamically favored if the oxidizing agent deployed is strong enough. The oxidation reactions of hydrocarbons, sugars, and fats by dioxygen are clearly exergonic (we do call these combustion reactions after all). What about redox reactions with less powerful oxidants? NAD^+ is used frequently as a biological oxidizing agent. Are all these reactions as favored as combustion? Hardly so. Remember that in every redox reaction, an oxidizing and reducing agent react to form another oxidizing and reducing agent. Consider the following reaction:

$$\text{Pyruvate} + \text{NADH} \leftrightarrow \text{Lactate} + \text{NAD}^+.$$

This reaction can go either way and is reversible. The above form is written in the favored direction in anaerobic metabolism when both Pyr and NADH levels are high. Although the ΔG^0 favors the oxidation of lactate, given the high concentration of Pyr and NADH, the reaction is driven in the opposite direction and proceeds as shown. To determine if a redox reaction is favored and likely to occur (and possibly be regulated), the ΔG^0 for a redox reaction should be calculated from standard reduction potentials, using the formula $\Delta G^0 = -nFE^0$.

14.1.5: WHICH ENZYMES TO REGULATE: THOSE CATALYZING COMMITTED STEPS IN PATHWAYS

The best enzymes to regulate are those that catalyze the first committed step in the reaction pathway. The committed step proceeds with a $\Delta G^0 < 0$ and is essentially irreversible. These reactions often occur from key metabolic intermediates that are immediately before or proximal to branches in reaction pathways. Two examples of key intermediates at branch points of metabolic pathways are shown in Figure 14.1.8 shows the reactions for the production and use of the intermediate glucose-6-phosphate.

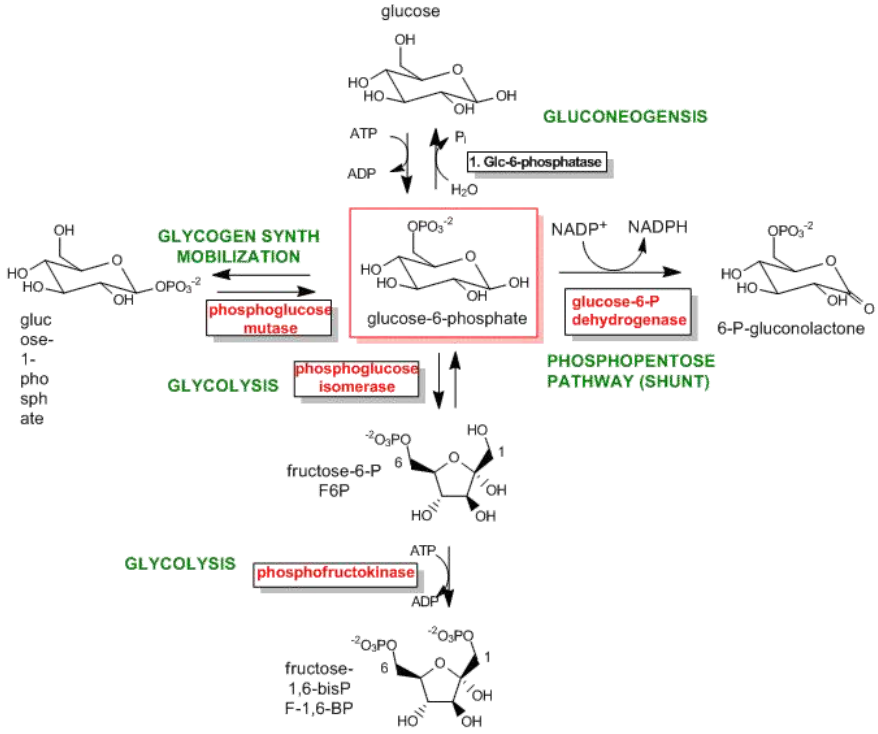


Figure 14.1.8: Reactions for the production and use of the intermediate glucose-6-phosphate

Figure 14.1.10 shows reactions for the production and use of the intermediate acetyl-CoA.

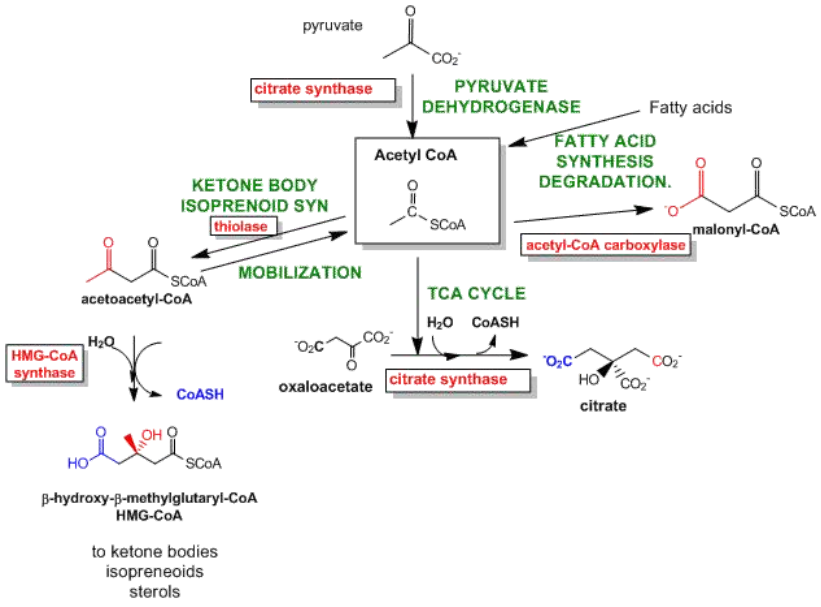


Figure 14.1.10: Reactions for the production and use of the intermediate acetyl-CoA

In reality, metabolic regulation is more complex and is distributed to many steps in a reaction pathway in ways that might not be evident without details mathematical analyses. We will discuss that in the next sections on metabolic control analysis.

14.2: BASIC PRINCIPLES OF METABOLIC CONTROL ANALYSIS (MCA)

14.2.1: INTRODUCTION TO METABOLIC CONTROL ANALYSIS

Enzyme kinetics may seem difficult given the complicated mathematical derivations, the number of chemical species involved (an enzyme and all its substrates and products), the number of steps in the mechanism, and a large number of rate, kinetic, and dissociation constants. An example of such a “complicated” reaction explored earlier is shown in Figure 14.2.1.

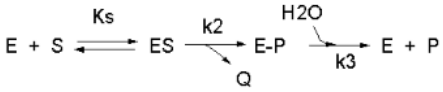


Figure 14.2.1: Chemical Equations for an enzyme-catalyzed reaction with a covalent E-intermediate

$$v = \frac{\frac{k_2 k_3}{k_2+k_3}(E_0)(S)}{\left[\frac{k_3}{k_2+k_3}\right]K_S + S} = \frac{k_{cat}(E_0)(S)}{K_M + S} \tag{14.2.1}$$

But single enzymes rarely act in isolation. They are components of complex pathways which have a multitude of steps, many of which are regulated. To fully understand a reaction, it is important to study the concentrations of all species in the entire pathway as a function of time. Imagine deriving the equations and determining all the relevant concentrations and constants of a pathway such as glycolysis!

To study enzyme kinetics in the lab, you have to spend much time developing assays to measure how the concentration of species changes as a function of time to be able to measure the initial velocities of an enzyme-catalyzed reaction. However, in networks of connected metabolic reactions, the concentration of some species in the system may not change. How can this happen? Two simple examples might help explain how.

Example 1: There is no change in input or output from a given reaction. This would occur in a closed system for a reversible reaction at equilibrium.

For a reversible reaction of reactant R going to product P (R ↔ P) with forward and reverse first-order rate constants, the following equation can be written at equilibrium:

$$\begin{aligned} v_f &= k_1 R = k_2 P \\ K_{eq} &= \frac{P_{eq}}{R_{eq}} = \frac{k_1}{k_2} \end{aligned} \tag{14.2.2}$$

At equilibrium, R and P don’t change.

Example 2: Consider the reaction as part of a pathway of reactions (like an open system). Now imagine a nonzero input to form reactant R and a nonzero output that consumes product P as shown in Figure 14.2.2.

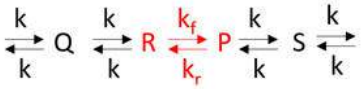


Figure 14.2.2: Reactant R conversion to Product P with preceding and following steps

If the input and output rates are the same, the concentrations of R and P would not change with time. That is the rate of formation of a reactant R for a given reaction is equal to the rate at which the product P of the given reaction is used. This would lead to a **steady state** but not equilibrium concentrations of the species.

14.2.1.1: THE STEADY STATE - A THRESHOLD BIOCHEMISTRY CONCEPT

Students have a difficult time understanding the steady state and recognizing when it may prevail under a set of conditions.

IN VITRO VS. *IN VIVO*

Cornish-Bowden in his book Fundamentals of Enzyme Kinetics talks about key differences between the enzyme kinetics we do in test tubes (*in vitro*) and what happens in cells (*in vivo*). We've discussed this previously but it is important to reconsider it now. The conditions under which the enzymes are studied (*in vitro*) and operate (*in vivo*) are very different.

- *In vitro* (in the lab), the enzyme is held at a constant concentration while the substrate is varied (i.e the substrate concentration is the independent variable). The velocity is determined by the *substrate* concentration. When inhibition is studied, the substrate is varied while the inhibitor is held constant at several different fixed concentrations.

- *In vivo* (in the cell), the velocity might be held at a relatively fixed level in a pathway with the substrate determined by the *velocity*. To avoid a bottleneck in flux, substrate can't build up at the enzyme, so the enzyme processes it in a steady-state fashion to produce the product as determined by the Michael-Menten equation

These differences are vastly underappreciated and not understood by students and instructors alike.

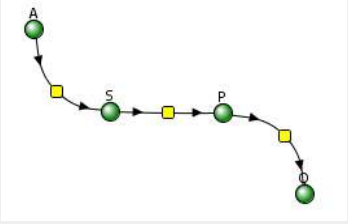
Additionally, we confound our efforts in helping students understand the steady state when almost **all** of our efforts are focused on presenting initial rate v_0 vs [S] curves when the substrate concentration is changed. Then instructors expect students to magically understand the steady state when substrate levels in pathways don't change. It's a big leap out of this box. We can help students better understand the steady state by shifting to progress curves (concentrations vs time), which are easily constructed by using programs such as VCell and Copasi.

Let's use Vcell to analyze a very simple reversible enzyme in isolation, S ↔ P. Then we will place that same enzyme into a "mini" pathway, A ↔ S ↔ P ↔ Q in which there is one preceding reactant A and one following product Q that help control S and P levels. We have to start the simulation at some specified concentration so the default for the simulations is set such that [S] and [A] at time t = 0, S_0 and A_0, are 5, and the rest are set at 0. Hence as the simulation starts, there will be a readjustment of concentration until equilibrium or steady state concentration is reached. The reaction diagrams and their parameters are shown in Figure 14.2.3.

A reversible enzyme in isolation, S ↔ P

J	reaction rate	$\frac{(\frac{S \cdot VmaxFwd}{KmFwd} - \frac{P \cdot VmaxRev}{KmRev})}{(1.0 + \frac{S}{KmFwd} + \frac{P}{KmRev})}$
KmFwd	Km forward	10.0
VmaxFwd	max forward rate	10.0
KmRev	Km reverese	20.0
VmaxRev	max reverse rate	5.0
Size	size	50000.0
initConc	initial concentration for S	5.0

A reversible enzyme for S ↔ P in a mini-pathway A ↔ S ↔ P ↔ Q



J	reaction rate	$\frac{(\frac{S \cdot VmaxFwd}{KmFwd} - \frac{P \cdot VmaxRev}{KmRev})}{(1.0 + \frac{S}{KmFwd} + \frac{P}{KmRev})}$
KmFwd	Km forward	10.0
VmaxFwd	max forward rate	10.0
KmRev	Km reverese	20.0
VmaxRev	max reverse rate	5.0
J	reaction rate	$(Kf \cdot P - Kr \cdot Q)$
Kf	forward rate constant	20.0
Kr	reverse rate constant	2.0
J	reaction rate	$(Kf \cdot A - Kr \cdot S)$
Kf	forward rate constant	2.0
Kr	reverse rate constant	2.0
Size	size	50000.0
initConc	initial concentration for S	5.0
initConc	initial concentration for P	0.0
initConc	initial concentration for Q	0.0
initConc	initial concentration for A	5.0

Figure 14.2.3: Reactions schemes and parameters for simple enzyme-catalyzed reversible reaction and the same reaction embedded in a "mini" pathway.

Vcell simulation for the isolated enzyme

Remember that an enzyme does not change the thermodynamics of a given reaction and hence doesn't alter K_{eq}. It just speeds up both the forward and the reverse reactions. Hence you can calculate the K_{eq} for the reaction condition from the Vcell model time course graph:

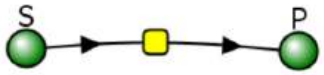
$$K_{eq} = \frac{[P]_{eq}}{[S]_{eq}} = \frac{4}{1} = 1 \tag{14.2.3}$$

Now you observe that in this reaction, S does change from its initial value, S_0 = 5 (since P_0 = 0). Soon, however, the reaction comes to a real dynamic equilibrium in that both S and P don't change with time.

You could choose a different initial concentration of S and P, rerun the simulation and calculate K_{EQ} for the new conditions using the csv-downloaded spreadsheet data fro each run you make. They should be the same.

VCell MODEL

Reversible reaction E + S ↔ ES ↔ EP ↔ E + P



Initial values

Name	Description	Global	
J	reaction rate	☐	$\frac{(\frac{S \cdot VmaxFwd}{KmFwd} - \frac{P \cdot VmaxRev}{KmRev})}{(1.0 + \frac{S}{KmFwd} + \frac{P}{KmRev})}$
KmFwd	Km forward	☐	10.0
VmaxFwd	max forward rate	☐	10.0
KmRev	Km reverese	☐	20.0
VmaxRev	max reverse rate	☐	5.0
S	Species Concentration	☑	Variable
P	Species Concentration	☑	Variable

Select Load [model name] below

Load EnzymeRevIsolated

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

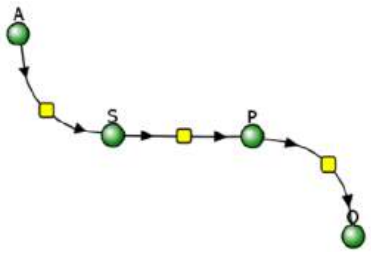
Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

Vcell simulation for the the enzyme in a "mini-pathway"

Now compare this same reaction but in which S and P are part of the "mini-pathway" A ↔ S ↔ P ↔ Q, as shown in the Vcell model below. The enzyme kinetic parameters for S ↔ P ($K_{M\ forward}$, $V_{M\ forward}$, $K_{M\ reverse}$, $V_{M\ reverse}$) in the A ↔ S ↔ P ↔ O pathway were made the same as for the simple S ↔ P reaction since its the same enzyme. How do the A ↔ S and P ↔ Q reactions affect the apparent K_{EQ}? Run the Vcell model with the defaults automatically set to the values in the table above!

VCell MODEL



Initial Values

Name	Description	Global	
J	reaction rate	☐	$(Kf \cdot A - Kr \cdot S)$
Kf	forward rate constant	☐	2.0
Kr	reverse rate constant	☐	2.0
A	Species Concentration	☑	Variable
S	Species Concentration	☑	Variable

Name	Description	Global	
J	reaction rate	☐	$\frac{\left(\frac{S \cdot VmaxFwd}{KmFwd} - \frac{P \cdot VmaxRev}{KmRev}\right)}{\left(1.0 + \frac{S}{KmFwd} + \frac{P}{KmRev}\right)}$
KmFwd	Km forward	☐	10.0
VmaxFwd	max forward rate	☐	10.0
KmRev	Km reverese	☐	20.0
VmaxRev	max reverse rate	☐	5.0
S	Species Concentration	☑	Variable
P	Species Concentration	☑	Variable

Name	Description	Global	
J	reaction rate	☐	$(Kf \cdot P - Kr \cdot Q)$
Kf	forward rate constant	☐	20.0
Kr	reverse rate constant	☐	2.0
P	Species Concentration	☑	Variable
Q	Species Concentration	☑	Variable

Select Load [model name] below

Load EnzymeRev_Constant_S_Premoved

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

As in the first simulation, constant values for S and P are soon reached. Both are significantly lower than in the simple reaction of S ↔ P since P is being pulled toward Q faster than Q is converted back to S. Note also that Q reaches a higher concentration than either A or S but remember that the sum of the initial values of A and S is 10.

Run the simulation again only this time set k_r=2000 for the P ↔ Q reaction.

Now calculate the "$K_{EQ\ apparent}$" for the different k_f values of the conversion of P → Q from this equation. Again use the csv-downloaded spreadsheet data for each run you make.

$$K_{\text{eq apparent}} = \frac{[\text{P}]_{\text{steady state}}}{[\text{S}]_{\text{steady state}}} \tag{14.2.4}$$

Do they equal 4 as in case 1? No, they do not. You should see that the "$K_{EQ\ apparent}$" value deviates most from 4 (lower number) when the rate constant for removal of P is highest (k_f=2000).

14.2.2: ANIMATIONS

Now let's look at some animation for these reactions as an additional way to understand the dynamics of the reactions. Click the image in Figure 14.2.4 to view (in a new window) the animation of the chemical species and an inserted graph showing concentration vs time for the reversible enzyme-Catalyzed Reaction S ↔ P

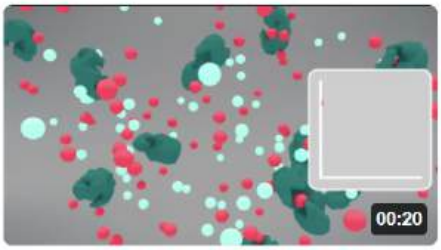
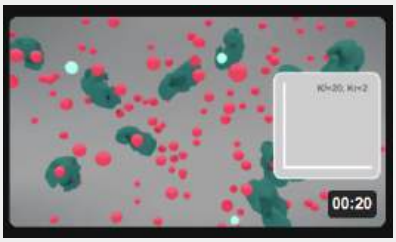


Figure 14.2.4: Animation of the reversible enzyme-Catalyzed Reaction: S ↔ P

Pay attention to the disappearance of S (red) and the appearance of P (blue) species after interacting with the enzyme (green).

Now here are two animations for the S ↔ P reaction when its embedded in the "minipathway" A ↔ S ↔ P ↔ Q when k_f for the reaction P ↔ Q is **20** (left) and **2**00 (right). Click the images in Figure 14.2.5 to view (in a new window) an animation of the chemical species and inserted graphs showing concentrations vs time.

"mini-pathway" A ↔ S ↔ P ↔ Q, k_f for the reaction P ↔ Q = **20**



"mini-pathway" A ↔ S ↔ P ↔ Q, k_f for the reaction P ↔ Q = **200**

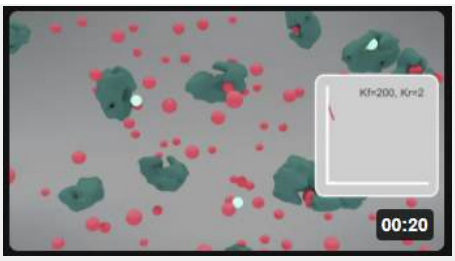


Figure 14.2.5: Animation of the reversible enzyme-Catalyzed Reaction embedded in a "minipathway" A ↔ S ↔ P ↔ Q when k_f for the reaction P ↔ Q is **200.**

These animations should reinforce your understanding of the differences between equilibrium and steady-state conditions, although you can calculate the actual $K_{EQ\ apparent}$ from the insert graphs without numerical data.

Understanding a pure enzyme in vitro and in vivo requires different approaches. Biochemists like to isolate and purify to homogeneity an enzyme found in some tissue and study its mechanism of action. In doing thermodynamic measurements to measure equilibrium constants (K_{eq}) or dissociation constants (K_D), from which ΔG^0 can be calculated, a protein concentration is usually held constant as the binding ligand concentration is varied (independent variable). A dependent variable signal (often spectroscopic) is measured. Measurements are made when equilibrium is reached.

For enzyme kinetic measurements *in vitro*, the enzyme concentration is usually held constant while substrate and modifiers are varied (independent variables) to determine how velocity (dependent variable) changes. The velocity is determined by the substrate concentration. When inhibition is studied, the substrate is varied while the inhibitor is held constant at several different fixed concentrations.

In vivo, the substrate concentration and even the enzyme concentration are determined by the velocity. Again compare this to in vitro kinetics when concentrations determine the velocity. For sets of reactions in pathways, it is better to use the term flux, J. In the steady state, the in and out fluxes for a given reaction are identical. Flux J is used to describe the rate of the system whereas rate or velocity v is used to describe the rate of an individual enzyme in a system.

Computer programs can find steady-state concentrations by finding the roots of the ordinary differential equations (ODE) when set to zero ($v_f = v_r$). To model a process at very low concentrations, programs can also use probabilistic or stochastic simulations to model probability distributions for species and their change with time for a finite number of particles. In such simulations, concentrations (mM) are placed with the number of particles. ODEs don't work well to describe these conditions since changes in concentrations are not continuous.

Now back to our earlier rhetorical question of deriving the equations and determining all the relevant concentrations and constants of a pathway such as glycolysis! It has been done by Teusink et al for glycolysis in yeast. Many such complicated metabolic and signal transduction pathways have been mathematically modeled in the hopes of better understanding cellular and organismal responses. Quantitatively modeling and predicting input, outputs, and concentrations of all species in complex pathways is the basis of systems biology.

14.2.2.1: BASIC PRINCIPLES OF METABOLIC CONTROL ANALYSIS - GLYCOLYSIS

Let's use a more complex example, glycolysis, to illustrate the powers of computational modeling of entire pathways. The reference for this model is shown below.

Can yeast glycolysis be understood in terms of in vitro kinetics of the constituent enzymes? Testing biochemistry, Bas Teusink, Jutta Passarge, Corinne A. Reijenga, Eugenia Esgalhado, Coen C. van der Weijden, Mike Schepper, Michael C. Walsh, Barbara M. Bakker, Karel van Dam, Hans V. Westerhoff, and Jacky L. Snoep, 2000, *European Journal of Biochemistry*, 267, 5313-5329. PubMed ID: 10951190.

Databases containing curated models with all the above information have been developed for many pathways. The yeast glycolysis model described in the reference above and the material below is found in the Biomodels Database. The model, BIOMD0000000064, can be downloaded as a systems biology markup language (SBML) file and imported into any of the programs described above.

A variety of inputs are required for such computational analyses:

a. **Defined pathways.** These are available in many databases. An example from the KEGG pathways for yeast glycolysis is shown in the left panel of Figure 14.2.6 below. The right panel shows a more familiar representation, with glucose on the outside of the cell (Glc_o) entering the cell.

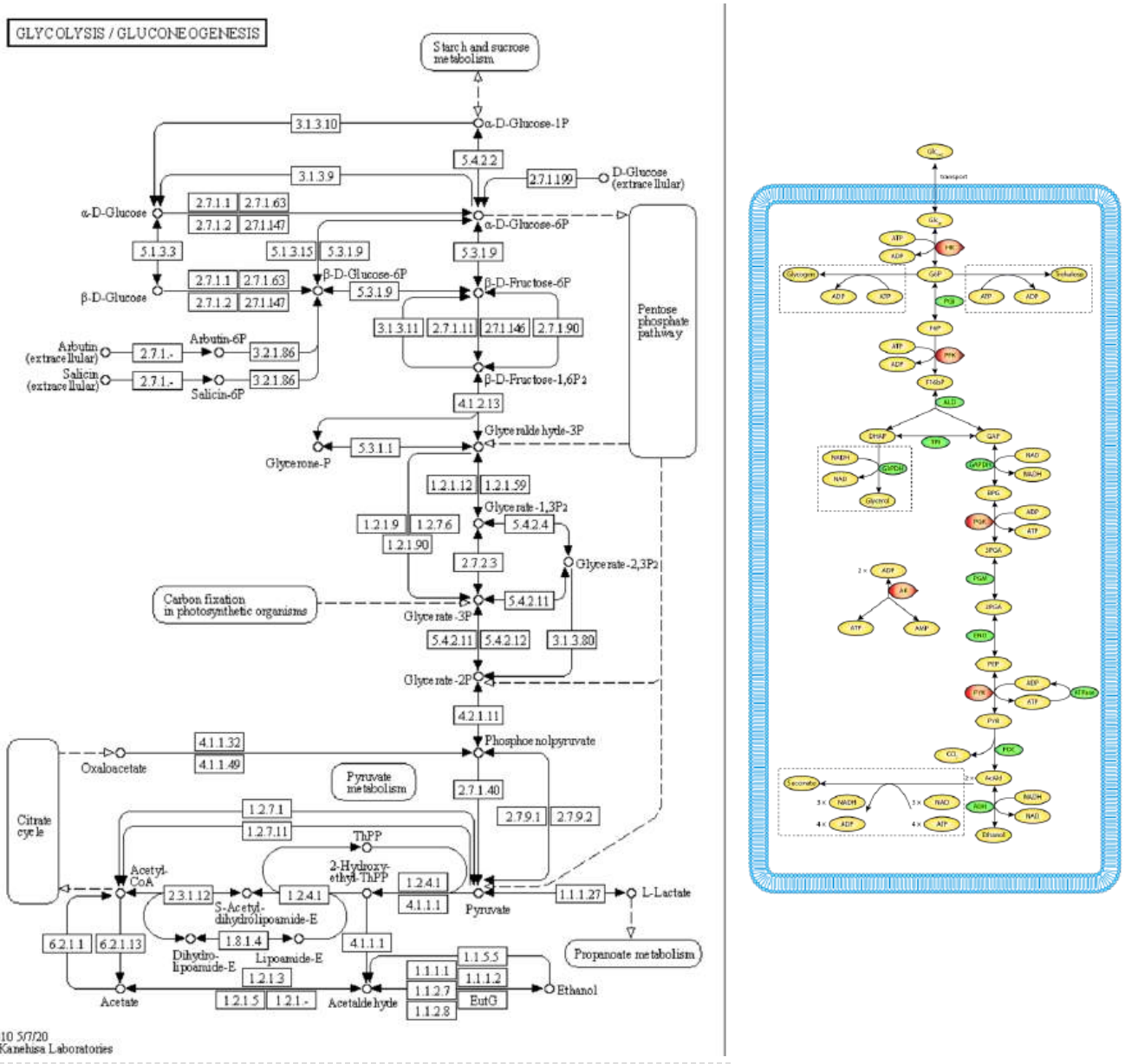


Figure 14.2.6: Left - Glycolysis Pathway from KEGG. http://www.genome.jp/dbget-bin/www_bget?pathway:sce00010; Right - Teusink Glycolysis Model based on BIOMD0000000064. https://models.physiomeproject.org/e/72/. 6:43am 23st April 2023. CellML author(s): Catherine Lloyd. Creative Commons Attribution 3.0 Unported License.

b. A **computational modeling program** to input all parameters, equations, and models and calculate concentrations for all species as a function of time. We have been using Vcell throughout this book. A **reaction diagram** is constructed that connects all of the species. Two are shown in Figure 14.2.7 below.

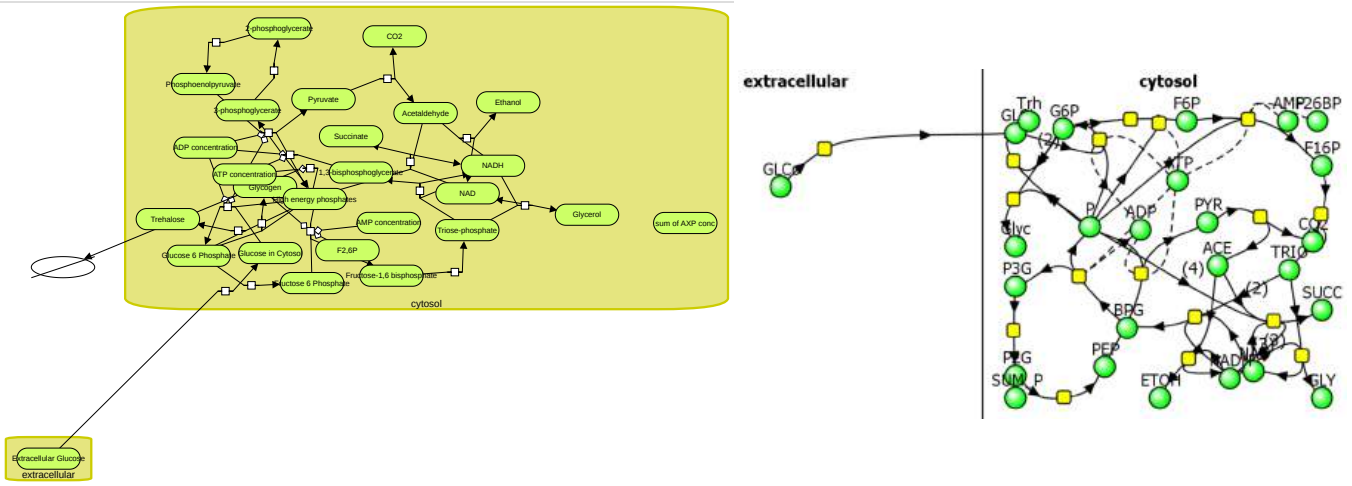


Figure 14.2.7: Reaction Scheme for Yeast Glycolysis. Left made using NEWT Pathway Viewer using Systems Biological Graphical Notation (SBGN) for BIOMD0000000006. Right is a reaction diagram using Vcell. P represents ATP and ACE acetaldehyde.

The computational results from the analyzes for yeast glycolysis were able to fit experimental data only if corrected by the addition of several branching reactions (shown in the figures above and Figure 14.2.8) and in the dotted boxes in Figure 6 (Right)

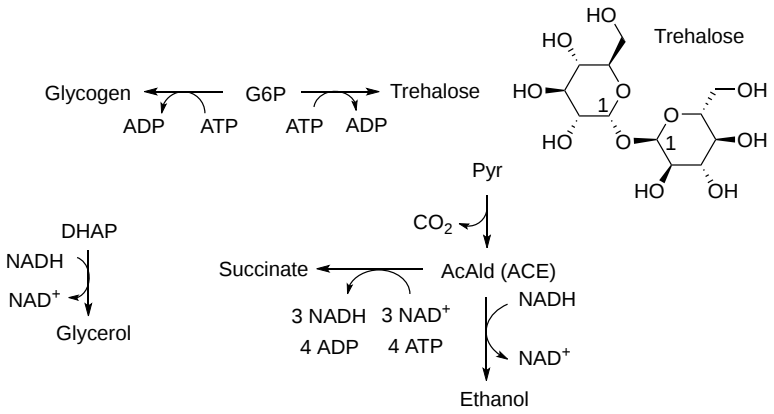


Figure 14.2.8: Branching reactions for Yeast Glycolysis model (Teusink)

Trehalose is the non-reducing disaccharide Glc-(α 1,1)-Glc. It is a reserve carbohydrate that also protects yeast against the effects of stress (desiccation, dehydration, temperature extremes) as well as lethal levels of ethanol. These effects arise from its effects on protein stability which likely arise by its preferential exclusion from the hydration sphere of protein. A similar mechanism accounts for the stabilizing effect of glycerol on protein stability, as described in Chapter 4.9.

c. A l**ist of all reactions** as shown in Figure 14.2.10 for glycolysis and branching reactions (taken from COPASI).

Reaction	Name	Structure
ACE + NADH -> NAD + ETOH	vADH	cytosol
F16P -> 2TRIO	vALD	cytosol
P ->	vATP	cytosol
P2G -> PEP	vENO	cytosol
TRIO + NADH -> NAD + GLY	vG3PDH	cytosol
NAD + TRIO -> NADH + BPG	vGAPDH	cytosol
P + GLCi -> G6P	vGLK	cytosol
GLCo -> GLCi	vGLT	extracellular
G6P + P -> Glyc	vGLYCO	cytosol
PYR -> ACE + CO2	vPDC	cytosol
P + F6P -> F16P	vPFK	cytosol
G6P -> F6P	vPGI	cytosol
BPG -> P + P3G	vPGK	cytosol
P3G -> P2G	vPGM	cytosol
PEP -> P + PYR	vPYK	cytosol
3NAD + 4P + 2ACE -> 3NADH + SUCC	vSUC	cytosol
2G6P + P -> Trh	vTreha	cytosol

Figure 14.2.10: List of reactants and concentrations for yeast glycolysis. P is ATP, ACE is acetaldehyde, a reactant/product of the enzyme alcohol dehydrogenase.

d. **Parameters** (concentrations, rate, enzyme kinetic, and equilibria constants) of all species. Figure 14.2.11 below shows the initial concentrations for the glycolysis model.

Species	Structure	Depiction	Clamped	Rules	
GLCo	extracellular		☑		50.0 [mmol.litre⁻¹]
GLCi	cytosol		☐		0.087 [mmol.litre⁻¹]
G6P	cytosol		☐		2.45 [mmol.litre⁻¹]
F6P	cytosol		☐		0.62 [mmol.litre⁻¹]
F16P	cytosol		☐		5.51 [mmol.litre⁻¹]
TRIO	cytosol		☐		0.96 [mmol.litre⁻¹]
BPG	cytosol		☐		0.0 [mmol.litre⁻¹]
P3G	cytosol		☐		0.9 [mmol.litre⁻¹]
P2G	cytosol		☐		0.12 [mmol.litre⁻¹]
PEP	cytosol		☐		0.07 [mmol.litre⁻¹]
PYR	cytosol		☐		1.85 [mmol.litre⁻¹]
ACE	cytosol		☐		0.17 [mmol.litre⁻¹]
P	cytosol		☐		6.31 [mmol.litre⁻¹]
NAD	cytosol		☐		1.2 [mmol.litre⁻¹]
NADH	cytosol		☐		0.39 [mmol.litre⁻¹]
Glyc	cytosol		☑		0.0 [mmol.litre⁻¹]
Trh	cytosol		☑		0.0 [mmol.litre⁻¹]
CO2	cytosol		☑		1.0 [mmol.litre⁻¹]

Figure 14.2.11: Parameters for hexokinase reactions

e. **Equations** that can be used to compute the change in concentrations of all species with time. These are usually ordinary differential equations (ODE) as described in Chapter 6B - Kinetics of Simple and Enzyme-Catalyzed Reactions, Sections B1: Single Step Reactions.) ODEs are easy to write but require a computer to solve as the number of interacting species increases.

For a quick review, consider the relatively simple reaction scheme shown in Figure 14.2.12 below. The set of ODEs for each species is shown below that.

$$A + B \underset{k_2}{\overset{k_1}{\rightleftharpoons}} C \xrightarrow{k_3} D$$

Figure 14.2.12: Chemical reactions for ODEs

This set of ODEs for each species can be written.

$$\begin{aligned}
\frac{d[A]}{dt} &= -k_1[A][B] + k_2[C] \\
\frac{d[B]}{dt} &= -k_1[A][B] + k_2[C] \\
\frac{d[C]}{dt} &= +k_1[A][B] - k_2[C] - k_3[C] = +k_1[A][B] - (k_2 + k_3)[C] \\
\frac{d[D]}{dt} &= +k_3[C]
\end{aligned} \tag{14.2.5}$$

If a reaction removes species X, the right side of the ODE for the disappearance of that species has a – sign for that term. Likewise, if a reaction increases species X, the right-hand side has a + sign. The above examples show unimolecular (C to D, C to A + B) and bimolecular (A+B to C) reactions.

Now let's look at the equations just for the change in the concentration of glucose-6-phosphate. One reaction forms it and 3 remove it, as shown in Figure 14.2.13 below.

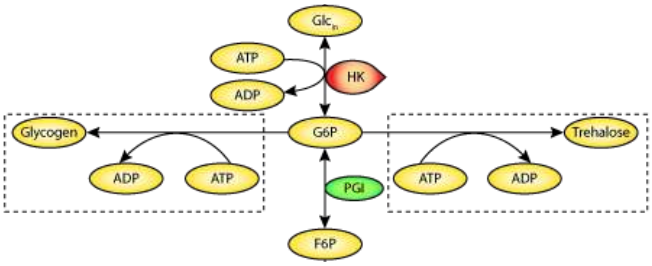


Figure 14.2.13: The reactions that change [G6P] in the Teusink yeast glycolysis model

An ODE can be written for the change in concentration of glucose-6-phosphate.

$$\mathrm{d[G6}P\mathrm{]/dt} = v_{\mathrm{HK}} - v_{\mathrm{PGI}} - 2v_{\mathrm{trehalose}} \ - v_{\mathrm{glycogen}} \ ) \tag{14.2.6}$$

where HK is hexokinase and PGI is phosphoglucoisomerase.

Let's look at the terms more closely and how the terms are written in Vcell.

Formation of G6P

Only one reaction goes towards G6P. Its the reaction catalyzed by hexokinase (G + ATP → G6P + ADP). In the Vcell reaction diagram, ATP is represented by P. Flux J =

$$\frac{VmGLK \cdot \left(GLCi \cdot ATP - \frac{G6P \cdot ADP}{KeqGLK}\right)}{KmGLKGLCi \cdot KmGLKATP \cdot \left(1.0 + \frac{GLCi}{KmGLKGLCi} + \frac{G6P}{KmGLKG6P}\right) \cdot \left(1.0 + \frac{ATP}{KmGLKATP} + \frac{ADP}{KmGLKADP}\right)} \tag{14.2.7}$$

Removal of G6P

3 reactions go away from G6P.

i. G6P → F6P. Flux J =

$$\frac{VmPGI\_2 \cdot \left(G6P - \frac{F6P}{KeqPGI\_2}\right)}{KmPGIG6P\_2 \cdot \left(1.0 + \frac{G6P}{KmPGIG6P\_2} + \frac{F6P}{KmPGIF6P\_2}\right)} \tag{14.2.8}$$

ii. G6P → Glycogen

Flux J = KGLYCOGEN_3

iii. G6P → Trehalose

Flux J = KTREHALOSE

Now let's run a simulation of the model of yeast glycolysis (*Teusink,B et al.: Eur J Biochem 2000 Sep;267(17):5313-29*). The actual model is available from Biomodels (BIOMD0000000064).

Yeast Glycolysis

Yeast Glycolysis

Teusink et al., 2000. BIOMD0000000064

Species conditions

Parameters

SUBMIT OMEX

Simulation ran using VCell & Tellurium and powered by https://run.biosimulations.org

For Figure 14.2.14 below, the model above model was run and the CSV file was used to plot the concentration vs time plots for each species. Two of the species, EtOH (50 mM) and GLCo - glucose outside (50 mM), have fixed concentrations much larger than the rest, so they are omitted from the graph for clarity. When you run the simulation above, do the same when you download the csv file and plot the progress curves. Also, note that when you run the simulation, the y-axis values are not shown correctly. Rescale them in the spreadsheet based on the correct values of EtOH and GLC0, which were held constant (clamped) throughout the time course at 50 mM.

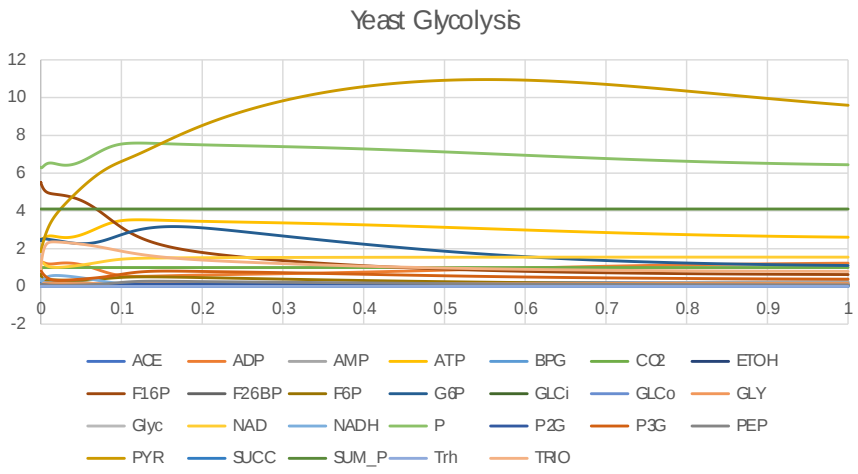


Figure 14.2.14: Time course graphs of all glycolytic species vs time

How close does simulation replicate fluxes found experimentally in vivo? As mentioned above, without the additions of the branch reaction, steady state was not achieved in the simulation. Fluxes in the model were within a factor of two for only about half of the enzymes so the model needs improvement. Other discrepancies could be accounted for generically by differences in kinetics in vivo compared to in vitro measurements. Additional products (glycogen, trehalose, glycerol and succinate) were derived from glucose, which mainly entered glycolysis and was converted to ethanol, showing the complexity in modeling even a relatively "simple" pathway as glycolysis.

14.2.2.2: METABOLIC CONTROL ANALYSIS AND SIMPLE ENZYME INHIBITION

Biochemists model complex enzyme-catalyzed reactions in the presence and absence of modifiers (either activators or inhibitors) to develop mechanisms for the reactions. The following kinetic parameters are experimentally determined by fitting initial velocity (vo) vs substrate concentration ([S]) through nonlinear fitting algorithms.

- K_m – the Michaelis Constant, the concentration of substrate at half-maximal velocity;
- V_m – the velocity at saturating substrate concentration;
- k_{cat} - the turnover number for conversion of bound reactant to product;
- K_{ix} – inhibition dissociation constants.

An example of competitive inhibition, shown in Figure 14.2.15, illustrates a common type of analysis for such reactions.

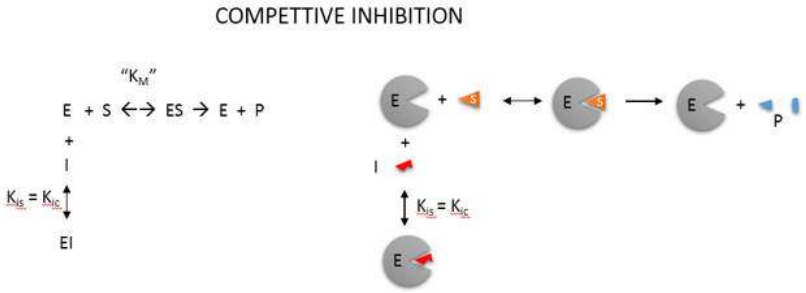


Figure 14.2.15: Chemical reactions for competitive inhibition

One modern pictorial depiction of a simple irreversible, enzyme-catalyzed reaction of substrate S going to product P with inhibition by the product and by an added inhibitor is shown in Figure 14.2.16. The square, a "node" in the reaction diagram, represents the enzyme.

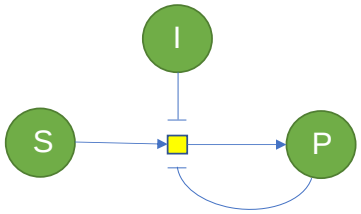


Figure 14.2.16: Pictorial Description for product and separate competitive inhibitor

Consider the simple enzyme-catalyzed reaction for a reversible conversion of substrate S to product P that has 3 reversible steps., as shown in Figure 14.2.17

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} EP \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} E + P$$

Figure 14.2.17: Three step reversible enzyme-catalyzed reaction

If the forward (f) and reverse (r) chemical reaction steps were irreversible and written separately, simple Michaelis-Menten equations could be written for each.

$$\begin{aligned} v_f &= \frac{V_f S}{K_{MS} + S} = \frac{\frac{V_f S}{K_{MS}}}{1 + \frac{S}{K_{MS}}} \\ v_r &= \frac{V_r P}{K_{MP} + P} = \frac{\frac{V_r P}{K_{MP}}}{1 + \frac{P}{K_{MP}}} \end{aligned} \tag{14.2.9}$$

For the actual reversible reactions, the net forward rate cannot be found by simple subtraction of the two equations above as the differential equations describing the simple forward and reverse rates don't account for the reverse steps

$$v \neq \left[\frac{\frac{V_f S}{K_{MS}}}{1 + \frac{S}{K_{MS}}} - \frac{\frac{V_r P}{K_{MP}}}{1 + \frac{P}{K_{MP}}} \right] \tag{14.2.10}$$

A simple derivation (assuming rapid equilibrium for both forward and reverse steps) can be made for the net forward reaction. Again consider the following enzyme-catalyzed reaction (Figure 17 above):

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} EP \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} E + P$$

? A DERIVATION OF THE RAPID EQUILIBRIUM, FULLY REVERSIBLE MICHAELIS-MENTEN EQUATION CONVERSION OF SUBSTRATE S TO PRODUCT P

Here it is!

Derivation

You may remember that for the isolated E + S ↔ ES and for the E + P ↔ EP reactions, the simple dissociation constants, K_S and K_P are given by

$$\begin{aligned} K_S &= \frac{E_{eq}S_{eq}}{ES_{eq}} = \frac{k_{-1}}{k_1} \text{ or } ES = \frac{[E][S]}{K_S} \\ K_P &= \frac{E_{eq}P_{eq}}{EP_{eq}} = \frac{k_3}{k_{-3}} \text{ or } EP = \frac{[E][P]}{K_P} \end{aligned} \tag{14.2.11}$$

The rapid equilibrium assumption states that the rate of dissociation of ES and EP, which are both physical steps, is much faster than the rate of the chemical conversion steps for each complex. Hence $k_{-1} \gg k_2$ and $k_3 \gg k_{-2}$, so the relative amounts of ES and EP can be determined from the dissociation constants as shown above.

Mass conservation of enzymes gives

$$E_0 = E + ES + EP = E + \frac{[E][S]}{K_S} + \frac{[E][P]}{K_p} = E\left[1 + \frac{[S]}{K_S} + \frac{[P]}{K_p}\right] \tag{14.2.12}$$

From this, we can get the fractional amount of both ES and EP

$$\begin{aligned} \frac{ES}{E_0} &= \frac{\frac{[E][S]}{K_S}}{E\left[1 + \frac{[S]}{K_S} + \frac{[P]}{K_p}\right]} \text{ or } ES = \frac{E_0\frac{[S]}{K_S}}{\left[1 + \frac{[S]}{K_S} + \frac{[P]}{K_p}\right]} \\ \frac{EP}{E_0} &= \frac{[E][P]}{E\left[1 + \frac{[S]}{K_S} + \frac{[P]}{K_P}\right]} \text{ or } EP = \frac{E_0\frac{[P]}{K_P}}{\left[1 + \frac{[S]}{K_S} + \frac{[P]}{K_P}\right]} \end{aligned} \tag{14.2.13}$$

Now we can derive the rate equation for the net forward reaction for the rapid equilibrium case:

$$v = k_2[ES] - k_{-2}[EP] = k_2\frac{E_0\frac{[S]}{K_S}}{\left[1 + \frac{[S]}{K_S} + \frac{[P]}{K_p}\right]} - k_{-2}\frac{E_0\frac{[P]}{K_p}}{\left[1 + \frac{[S]}{K_S} + \frac{[P]}{K_p}\right]} \tag{14.2.14}$$

Knowing that k_2E_0 and $k_{-2}E_0$ represent the maximal velocities, V_f and V_r, respectively, the equation becomes:

$$v = k_2[ES] - k_{-2}[EP] = \frac{V_f\frac{[S]}{K_S}}{\left[1 + \frac{[S]}{K_S} + \frac{[P]}{K_p}\right]} - \frac{V_r\frac{[P]}{K_p}}{\left[1 + \frac{[S]}{K_S} + \frac{[P]}{K_p}\right]} = \frac{V_f\frac{[S]}{K_S} - V_r\frac{[P]}{K_p}}{\left[1 + \frac{[S]}{K_S} + \frac{[P]}{K_p}\right]} \tag{14.2.15}$$

Here is the derived equation.

$$v = k_2[ES] - k_{-2}[EP] = \frac{V_f\frac{[S]}{K_S}}{\left[1 + \frac{[S]}{K_S} + \frac{[P]}{K_p}\right]} - \frac{V_r\frac{[P]}{K_p}}{\left[1 + \frac{[S]}{K_S} + \frac{[P]}{K_p}\right]} = \frac{V_f\frac{[S]}{K_S} - V_r\frac{[P]}{K_p}}{\left[1 + \frac{[S]}{K_S} + \frac{[P]}{K_p}\right]} \tag{14.2.16}$$

An equation of a similar form can be derived from the steady state assumption. Hence the net forward rate can't be determined by simple substration of the backward rate from the forward rate as shown in Equation 10 above.

14.2.3: COMMON FORMS OF KINETIC EQUATIONS

Programs like COPASI and VCell have many built-in equations for the velocities of many enzyme-catalyzed reactions that have similar forms. Two are shown below:

Reversible Michaelis-Menten:

$$\frac{\frac{V_f[\text{substrate}]}{K_{Ms}} - \frac{V_r[\text{product}]}{K_{Mp}}}{1 + \frac{[\text{substrate}]}{K_{Ms}} + \frac{[\text{product}]}{K_{Mp}}} \tag{14.2.17}$$

Competitive Inhibition Reversible:

$$\frac{\frac{V_f[\text{substrate}]}{K_{Ms}} - \frac{V_r[\text{product}]}{K_{Mp}}}{1 + \frac{[\text{substrate}]}{K_{Ms}} + \frac{[\text{product}]}{K_{Mp}} + \frac{[\text{inhibitor}]}{K_I}} \tag{14.2.18}$$

What is most important for readers to understand is not detailed derivations or how to solve the differential equations on their own. However, you should be able to:

- write the differential equation of a given reaction;
- recognize the common equations used for nonenzyme-catalyzed reactions (mass action) and enzyme-catalyze ones;
- change parameters in programs that use numerical methods to solve systems of linked differential equations for a system and see how the time course graphs are affected.

14.3: THE FLUX CONTROL COEFFICIENT

14.3.1: INTRODUCTION

Metabolic control analysis (MCA) is one method used to address the complexity of dynamic changes of species in a complex metabolic system. As such, an understanding of MCA would apply to complex signal transduction pathways as well as to other emerging areas in systems biology. In MCA, external inputs (source) and outputs (exits), and pools and reservoirs exist which are connected to the internal metabolic enzymes, reactants, and products of the pathway connecting the two external reservoirs.

In studying complex metabolic systems, what is interesting to know is not just the K_m or V_m of particular enzymes (such as those far from equilibrium or catalyzing "rate limiting" steps based on in vitro studies), but rather the control parameters, known as **control coefficients**, for the system. These coefficients are properties of the whole pathway as an emerging system with properties different from those of the individual steps. For another example of an emergent property, consider consciousness. It must emerge from the incredible assembly of around 90 billion neurons in the brain forming around 10^{15} synapses. It cannot simply be predicted from the properties of isolated neurons. Three especially relevant parameters are used in metabolic control analyses, the Flux Control Coefficient, the Concentration Control Coefficient, and the Elasticity Coefficient

To understand how a pathway is regulated and how scientists modify it to meet their demands, one needs to understand these coefficients. In particular, we would like to know which:

- parameters have the greatest effect (increase or decrease) on a preferred outcome;
- properties of the system make it most resilient or fragile;
- parameters should be measured most accurately

14.3.2: THE FLUX CONTROL COEFFICIENT

The **Flux Control Coefficient** (C^J_{Ei}) gives the relative fractional change in pathway flux J (dJ/J) (a system variable) with fractional change in concentration or activity of an enzyme Ei (du_i/u_i). The change in the concentration of an enzyme can arise from increased synthesis or degradation of the enzyme, a change in a modifier (an inhibitor or activator), a covalent modification, the presence of inhibitor RNA, mutations, etc. Again, the word flux J is used to describe the rate of the system whereas rate or velocity v is used to describe an individual enzyme in the system. Think of u as "units" of activity".

The **Flux Control Coefficient** (C^J_{Ei}) can also be written as

$$C^J_{Ei} = \frac{dJ}{du}\frac{u}{J} \tag{14.3.1}$$

The dJ/du in Equation 14.3.1 suggests that the C^J_{Ei} is equal to the slope of a plot of J vs u multiplied by u/J at that point. Alternatively, it is the actual slope of an ln-ln plot of $\ln J$ vs $\ln u$. A more formal definition would be obtained using partial derivatives (a normal derivative like dy/dx with everyother variable held constant - in this case all other enzyme concentrations or activities other than the one under study :

$$C^J_{Ei} = \frac{\dfrac{\partial J}{J}}{\dfrac{\partial u}{u}} = \frac{\partial \ln J}{\partial \ln u_i} \tag{14.3.2}$$

C^J_{Ei} clearly shows how the system flux changes with changes in the activity of a single enzyme. The variable ui is a property of an enzyme (a local property) but J and concentrations (for the concentration coefficient – see below) are systems variables. Every enzyme in a pathway hence has a flux control coefficient. How are they related? It makes sense that the sum of all of the individual $C^J_{Ei} = 1$ as the fractional change in flux, $\partial J/J$, is equal in magnitude to the fractional change in all enzyme activities. This relationship, shown below, is called the **Summation Theorem**.

$$\sum_{i=1}^{n} C^J_{Ei} = 1 \tag{14.3.3}$$

This equation would predict that if there were one magical enzyme X that was truly rate limiting and as such completely controlled that rate, then C^J_{Ex} = 1 for that enzyme and 0 for all the rest of the enzymes. Yet this is counterintuitive since if all the enzymes are linked their responses should also be linked. Hence the control of flux is shared by all enzymes.

Another way to think about this is to consider a linear set of connected reactions. For these enzymes, C^J_{Ei} would vary between 0 and 1. If one enzyme was completely rate limiting, its C^J_{Ei} would be 1 indicating that the fractional change in flux, $\partial J/J$, is equal to the fractional change in enzyme concentration (or activity) u, $\partial u/u$, for enzyme X. This implies that all change in flux is accounted for by the change in the concentration of enzyme X. This is a bit difficult to believe so it would make sense to think of the control of a pathway as being

distributed over all the enzymes in the pathway. Plots of flux J vs enzyme activity u are similar to rectangular hyperbolas. Hence the flux coefficient changes with enzyme activity and with flux so at different fluxes, different enzyme flux coefficients would change as well. “Control” gets redistributed in the pathway.

Textbooks often describe that ideally the first step of a linear pathway would be the committed step and be rate limiting. But think of this. If the end product of a pathway feeds back and inhibits the first enzyme in the pathway, then the last enzyme influences the first even though the last is not considered “rate limiting”. Now if another enzyme removes the last product and hence first enzyme's feedback inhibitor of the pathway, that enzyme determines flux as well. This again shows that flux control is distributed throughout the system. The classical “rate limiting” enzyme might still have the largest C_{Ei}^{J} of the pathways, but if the system is in a steady state, all enzymes have the same rate. Hence the term “rate-limiting” is a bit useless in a pathway. Analysis of flux control coefficients gives a truly quantitative analysis of the system pathway.

Phosphofructokinase is said to be a “key” enzyme in glycolysis as it is one of three enzymes in the pathway that is characterized by a large negative free energy difference. Figure 14.3.1:

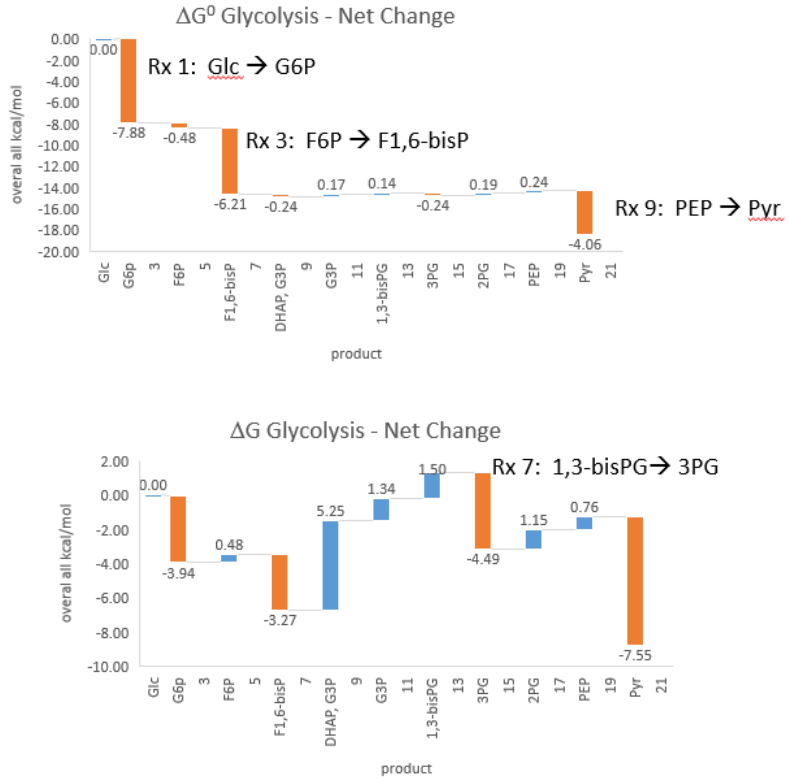


Figure 14.3.1: ΔG^0 and ΔG values for enzymes in glycolysis

It is also highly regulated. However, Heinisch et al (Molecular and General Genetics, 202, 75-82) found that increasing its concentration 3.5 fold in fermenting yeast did not affect the flux of ethanol production.

Here are some results of "sensitivity analyses" of yeast glycolysis showing the effect of small changes (up to 5%) on the flux control coefficients of glycolytic enzymes. The changes in the enzymes for these analyses are 1%. The numbers are scaled values that represent a relative change. A value of + 0.5 means that if the parameter is increased by 10%, the target value will increase by 5%. A column shows the rate of a reaction that is changed and the row indicates the flux of the reaction that has been affected. Green represents positive values and red negative with the intensity of the color correlating with the extent of the change. The table is split with column 1 of the first table applying to tables 2 and 3 as well.

	(Hexokinase)	(Glucose-6-phosphate is..	(Glycogen synthesis)	(Trehalose 6-phosphate ...	(Phosphofructokinase)	(Aldolase)
(Hexokinase)	0.155598	0.00218158	-0.0213366	-0.0147151	0.0016614	0.000351121
(Glucose-6-phos..	0.177323	0.00248618	-0.101886	-0.0788257	0.00189337	0.000400147
(Glycogen synth...	0	0	1	0	0	0
(Trehalose 6-pho..	0	0	0	1	0	0
(Phosphofructok..	0.177323	0.00248618	-0.101886	-0.0788257	0.00189337	0.000400147
(Aldolase)	0.177323	0.00248618	-0.101886	-0.0788257	0.00189337	0.000400147
(Glyceraldehyde ..	0.188586	0.0026441	-0.107526	-0.0833339	0.00201363	0.000425563
(Phosphoglycera..	0.188586	0.0026441	-0.107526	-0.0833339	0.00201363	0.000425563
(Phosphoglycera..	0.188586	0.0026441	-0.107526	-0.0833339	0.00201363	0.000425563
(Enolase)	0.188586	0.0026441	-0.107526	-0.0833339	0.00201363	0.000425563
(Pyruvate kinase)	0.188586	0.0026441	-0.107526	-0.0833339	0.00201363	0.000425563
(Pyruvate decarb..	0.188586	0.0026441	-0.107526	-0.0833339	0.00201363	0.000425563
(Succinate synth...	0.0928625	0.00130199	-0.0595875	-0.0450188	0.00099154	0.000209553
(Glucose transpo..	0.155598	0.00218158	-0.0213366	-0.0147151	0.0016614	0.000351121
(Alcohol dehydr...	0.19398	0.00271972	-0.110228	-0.0854928	0.00207122	0.000437735
(Glycerol 3-phos..	0.0928625	0.00130199	-0.0595875	-0.0450188	0.00099154	0.000209553
(ATPase activity)	0.268522	0.00376484	-0.292486	-0.202286	0.00286714	0.000605945

Table, part 2, continued:

	(Glyceraldehyde 3-phosp..	(Phosphoglycerate kinase)	(Phosphoglycerate muta...	(Enolase)	(Pyruvate kinase)
(Hexokinase)	0.0132581	0.000377931	0.00030863	0.000840434	0.000428631
(Glucose-6-phosphate isomerase)	0.0151093	0.0004307	0.000351723	0.000957781	0.000488479
(Glycogen synthesis)	0	0	0	0	0
(Trehalose 6-phosphate synthase)	0	0	0	0	0
(Phosphofructokinase)	0.0151093	0.0004307	0.000351723	0.000957781	0.000488479
(Aldolase)	0.0151093	0.0004307	0.000351723	0.000957781	0.000488479
(Glyceraldehyde 3-phosphate dehy...	0.0476129	0.00135724	0.00110836	0.00301819	0.00153931
(Phosphoglycerate kinase)	0.0476129	0.00135724	0.00110836	0.00301819	0.00153931
(Phosphoglycerate mutase)	0.0476129	0.00135724	0.00110836	0.00301819	0.00153931
(Enolase)	0.0476129	0.00135724	0.00110836	0.00301819	0.00153931
(Pyruvate kinase)	0.0476129	0.00135724	0.00110836	0.00301819	0.00153931
(Pyruvate decarboxylase)	0.0476129	0.00135724	0.00110836	0.00301819	0.00153931
(Succinate synthesis)	-0.228632	-0.0065173	-0.00532223	-0.014493	-0.00739161
(Glucose transport)	0.0132581	0.000377931	0.00030863	0.000840434	0.000428631
(Alcohol dehydrogenase)	0.0631784	0.00180094	0.0014707	0.00400489	0.00204254
(Glycerol 3-phosphate dehydrogen...	-0.228632	-0.0065173	-0.00532223	-0.014493	-0.00739161
(ATPase activity)	0.165497	0.0047176	0.00385254	0.0104909	0.00535047

table, part 3, continued

	Pyruvate decarboxylase)	(Succinate synthesis)	(Glucose transport)	(Alcohol dehydrogenase)	(Glycerol 3-phosphate d...	(ATPase activity
(Hexokinase)	.000256363	-0.0178963	0.984936	0.00560057	-0.0289301	-0.08292
(Glucose-6-phosphate isomerase)	.000292158	-0.0203951	1.12246	0.00638256	-0.0329695	-0.0944978
(Glycogen synthesis)		0	0	0	0	0
(Trehalose 6-phosphate synthase)		0	0	0	0	0
(Phosphofructokinase)	.000292158	-0.0203951	1.12246	0.00638256	-0.0329695	-0.0944978
(Aldolase)	.000292158	-0.0203951	1.12246	0.00638256	-0.0329695	-0.0944978
(Glyceraldehyde 3-phosphate dehy...	.00092066	-0.0681563	1.19375	0.0213292	-0.110646	-0.0946468
(Phosphoglycerate kinase)	.00092066	-0.0681563	1.19375	0.0213292	-0.110646	-0.0946468
(Phosphoglycerate mutase)	.00092066	-0.0681563	1.19375	0.0213292	-0.110646	-0.0946468
(Enolase)	.00092066	-0.0681563	1.19375	0.0213292	-0.110646	-0.0946468
(Pyruvate kinase)	.00092066	-0.0681563	1.19375	0.0213292	-0.110646	-0.0946468
(Pyruvate decarboxylase)	.00092066	-0.0681563	1.19375	0.0213292	-0.110646	-0.0946468
(Succinate synthesis)	0.00442091	0.337761	0.587821	-0.1057	0.549517	-0.0933806
(Glucose transport)	.000256363	-0.0178963	0.984936	0.00560057	-0.0289301	-0.08292
(Alcohol dehydrogenase)	.00122164	-0.0910284	1.2279	0.0284869	-0.147844	-0.0947182
(Glycerol 3-phosphate dehydrogen...	0.00442091	0.337761	0.587821	-0.1057	0.549517	-0.0933806
(ATPase activity)	.00320011	-0.240966	1.69975	0.0754091	-0.39165	-0.116635

Some things to note:

- almost everything has changed in each column, which shows that a change in one enzyme affects all enzymes;
- 2 adjacent rows often have identical numbers since they describe two connected reactions (without branches) so their steady state fluxes will be the same;
- the sum of the C^{J}_{Ei} values across the same row in both tables is 1 (summation theorem);
- perturbations of the "big 3" enzymes for steps 1,3, and 10 in glycolysis (hexokinase, phosphofructokinase, and pyruvate kinase), all of which proceed with a significant $-\Delta G^0$ and $-\Delta G$, have **minor effects** on the flux.

It should be clear from these results that you can't predict the enzymes that control the flux without this kind of mathematical analysis. Don't rely on just your "biochemical intuition"

14.3.3: FLUX CONTROL IN AEROBIC GLYCOLYSIS

Let's look at another pathway for which we can run a full simulation using Vcell - aerobic glycolysis. We have mostly encountered glycolysis as a central anaerobic pathway in almost all organisms. In some cases, it can also run aerobically. One prime example is in cancer cells in which there is a need for energy in the form of ATP but also for anabolic building blocks so cells can proliferate - a hallmark of cancer cells. Aerobic glycolysis was noted by Warbug and this effect now has his name.

Aerobic glycolysis has been modeled in yeast cells. The pathway is shown in Figure 14.3.2 below.

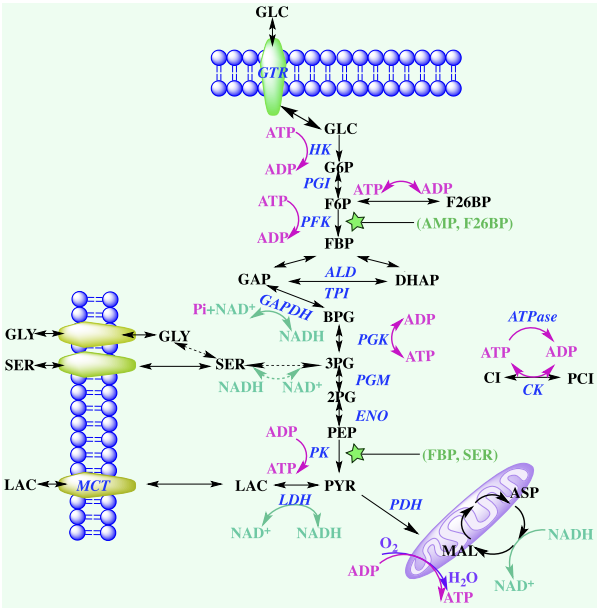


Figure 14.3.2: Schematic of the glycolysis model with chemical reactions and allosteric points of regulation described. Alexander A Shestov et al. (2014) Quantitative determinants of aerobic glycolysis identify flux through the enzyme GAPDH as a limiting step. eLife 3:e03342. https://doi.org/10.7554/eLife.03342. Creative Commons Attribution License

Abbreviations: GLC—glucose, G6P—glucose-6-phosphate, F6P—fructose-6-phosphate, FBP—fructose-1,6,-bisphosphate, F26BP—fructose-2,6,-bisphosphate, GAP—glcyceraldehyde-3-phosphate, DHAP—dihydroxyacetone phosphate, BPG—1,3 bisphosphoglycerate, 3PG—3-phosphoglycerate, 2PG—2-phosphoglycerate, PEP—phosphoenolpyruvate, PYR—pyruvate, SER—Serine, GLY—glycine, Lac—lactate, MAL—malate, ASP—aspartate, Pi—inorganic phosphate, CI—creatine, PCI—phosphophocreatine, GTR—glucose transporter, HK—hexokinase, PGI—phosphoglucoisomerase, PFK—phosphofructokinase, ALD—aldolase, TPI—triosephosphoisomerase, GAPDH—glyceraldehyde-phosphate dehydrogenase, PGK—phosphoglycerate kinase, PGM—phosphoglycerate mutase, ENO—enolase, PK—pyruvate kinase, LDH—lactate dehydrogenase, MCT—monocarboxylate transporter, PDH—pyruvate dehydrogenase, CK—creatine kinase.

The Warburg Effect **W** can be quantitatively described by the ratio of the flux of pyruvate to lactate (J_{Lac}) compared to the flux of pyruvate entry into the mitochondria and subsequent consumption of O_2 by mitochondrial oxidative phosphorylation (J_{ox}), so $\mathbf{W} = J_{Lac}/J_{ox}$).

Usually in normal conditions for healthy cells, W <0.10, which means that less than 10% of glucose is converted (or diverted) to lactate. On average for different tissues, W < 0.3. Hence some lactate is always being made. High values of W (> 1) mean that most glucose is diverted to lactate synthesis. The W values is determined by lots of factors that determine the energy state of cells, including rates of glucose uptake, ATP hydrolysis, and biosynthesis, as well as the balance of cytoplasmic and mitochondrial NAD^+/NADH, hence the redox state of the cell. Key to the balance leading to aerobic glycolysis is the need to reoxidize cytosolic NADH back to NAD^+ under aerobic conditions so glycolysis can continue under conditions when there is a high growth rate.

How does a cell know which path to take, pyruvate to lactate (anaerobic), pyruvate to CO_2 with ATP synthesis (aerobic), or some of both? In terms of metabolic control analysis, what controls the flux of pyruvate through these paths? This is where computational modeling informed by experimental data helps. We can determine the flux control coefficient (C^J_{Ei}), more easily written as **FCC**, for each enzyme in glycolysis to help us understand what controls the flux. Figure 14.3.3 metaphorically describes these different paths. In a way, the flux of pyruvate through these 3 choices, anaerobic, aerobic, and both (as reflected by the Warburg effect W) is then perhaps the best example to use to discuss flux control since the readers are likely very familiar with the concept of anaerobic and aerobic metabolism and the "switch" between them when running short sprints vs marathons.

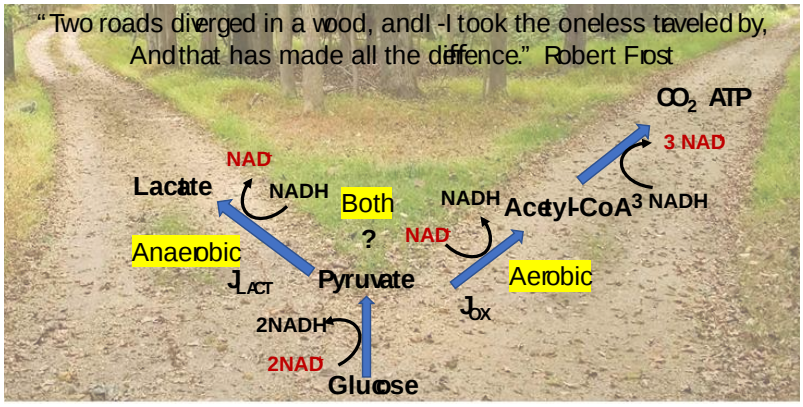


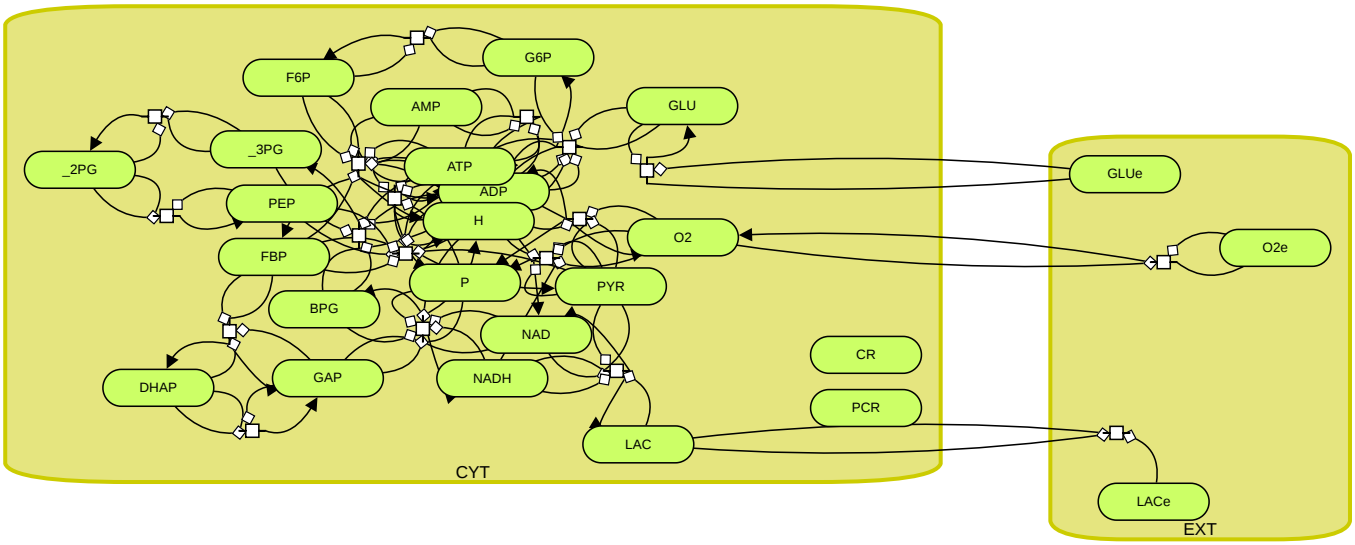
Figure 14.3.3: Fluxs of pyruvate through multiple paths

Now let's run the Vcell model for aerobic glycolysis in yeast.

VCell MODEL

AerobicGlycolysis

Quantitative determinants of aerobic glycolysis identify flux through the enzyme GAPDH as a limiting step. .Shestov AA, Liu X, Ser Z, Cluntun AA, Hung YP, Huang L, Kim D, Le A, Yellen G, Albeck JG, Locasale JW. *eLife* , 7/ 2014 , Volume 3 , PubMed ID: 25009227. Biomodel MODEL1504010000



Select Load [model name] below

Load AerobicGlycolysis

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

Where did all the parameters in the model come from? They are derived experimentally and the goal of modeling is to produce a computational model that is consistent with the experimental findings which are fed into the model.

The result of the model suggests that the Warburg effect is determined by the growth rate of the cells as well as the activity of mitochondria. The model does **not** support the idea that aerobic glycolysis results as a balance between energy needs and new biosynthesis as cells proliferate. Rather, the cells that proliferate the most have the most active mitochondria and hence lower levels of lactate production. Redox balance appears to be key in pushing cells toward lactate synthesis and aerobic glycolysis.

A key step is catalyzed by glyceraldehyde-3-phosphate dehydrogenase (**GAPDH**), the glycolytic enzyme that separates the top and bottom halves of glycolysis. Experimental data shows that in fact, flux through this enzyme is rate-limiting with the levels of F1,6-BP also being very important. Some key steps that were thought to be rate-limited in fact had negative fluxes through them. Negative flux control coefficients were found and confirmed for several steps thought to be rate-limiting in glycolysis.

The flux control coefficients, given by the formula

$$C_{Ei}^{J} = \frac{\frac{\partial J}{J}}{\frac{\partial u}{u}} = \frac{\partial \ln J}{\partial \ln u_i} \tag{14.3.4}$$

were calculated for each enzyme in the glycolytic pathway. The authors of the study used this equation for the Flux Control Coefficient:

$$FCC = \frac{d\ln[J_{LACT}]}{d\ln[E_i]} \tag{14.3.5}$$

The fluxes were measured for lactate production by using ^{13}C-labeled lactate to get the absolute concentration of lactate, and then measuring the fluxes by measuring the changes in ^{13}C-lactate production with time. Using these data and the selective inhibition by small drugs of each step in glycolysis to effectively alter the "concentration/activity of the enzyme", the FCC of each enzyme could be calculated. The results are shown in Figure 14.3.4 below.

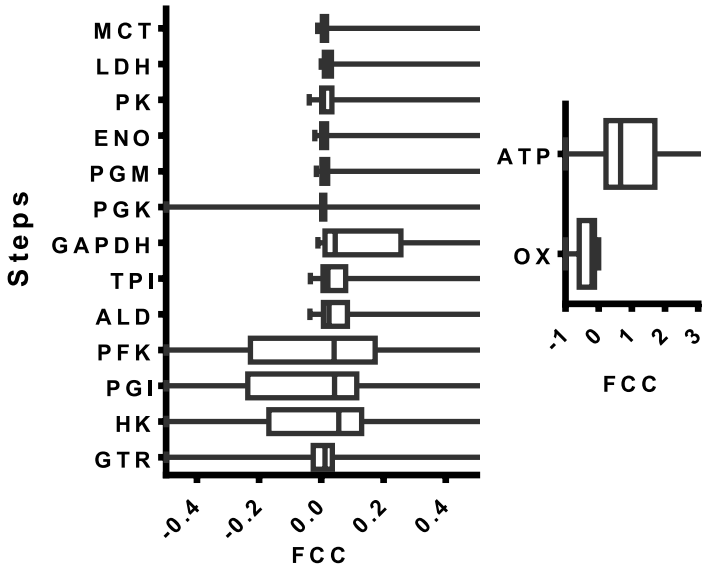


Figure 14.3.4: (left) Box plots of flux control coefficient (FCC) for lactate production for each enzymatic step in glycolysis (FCC = $dlnJ_{lac}/dln\ E_i$) where J_{lac} is the rate of pyruvate conversion to lactate, and E_i is the i^{th} enzyme in glycolysis for each step of glycolysis. (right) Box plots of flux control coefficient (FCC) for lactate production for Oxygen consumption (OxPhos) and ATP consumption (ATP). Shestov et al., ibid

The left-hand side of Figure 4 shows that two of the key enzymes that are thermodynamically favored in glycolysis, PFK (phosphofructokinase) and HK (hexokinase), both driven by ATP hydrolysis, have essentially *negative* flux coefficients, so they are **NOT** key in driving flux toward lactate. Only one enzyme, GAPDH, is characterized by a positive flux coefficient. (Remember we are talking about flux coefficients, not the Warburg effect factor W.)

Also important, as mentioned above, is the concentration of F1,6BP (FBP in the figure below). The model and data show that high levels of FBP and associated higher concentrations of the reactants in the first half of glycolysis (denoted by ↑↑ and large font size), and a corresponding depletion of substrates in the bottom half (denoted by ↓↓) actually cause a bottleneck in glycolysis and inhibit flux through GADPH. Hence GADPH determines the flux through glycolysis. This "bottleneck" is associated with the redox and energy states of the cell. This outcomes are illustrated in Figure 14.3.1 below.

Figure 14.3.5:

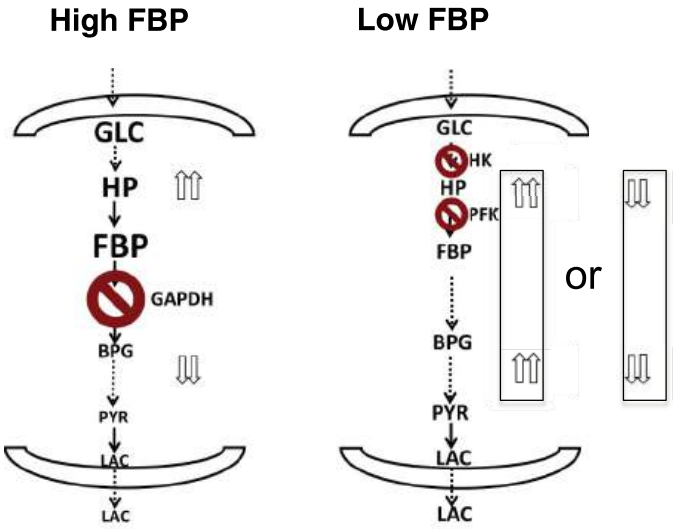


Figure 14.3.5: **A unified model of aerobic glycolysis.** Shestov et al., ibid

When there is a balance of metabolites in the top and bottom halves of glycolysis (right hand size), including both high levels in both top and bottom (denoted by ↑↑ and font size) or low levels in both (denoted by ↓↓ and font size), no bottleneck at GAPDH exists and flux through glycolysis is determined by the usual "suspects", hexokinase and phosphofructokinase.

The results described above are those predicted by computational modeling. Do experimental data support the model? The only way to know is to test the model. This was done by using drugs to inhibit each of the enzymes in glycolysis. This offers an effective way to change the "effective" concentration or activity in the enzyme. From this data, flux changes due to changes at each step could be determined. These data supported the computation results - GADPH regulates aerobic metabolism in normal as well as. When the top half substrates were elevated, the blockade of flux was determined by GADPH. In Chapter 13, we studied all of the glycolytic enzymes and the myriad of small molecules that control the activity of the 3 major enzymes that have the biggest thermodynamic push, HK, PFK and PK. GADPDH appears to be equally important. It is known to be regulated by many mechanisms, including post-translational modifications, such as nitrosylation, and modification of its active site cysteine with ROS. Perhaps this helps regulate the balance between anaerobic and aerobic glycolysis. Also, there are high levels of expression of GADPH in cells engaged in aerobic glycolysis.

14.4: CONCENTRATION CONTROL AND ELASTICITY COEFFICIENTS

14.4.1: CONCENTRATION CONTROL COEFFICIENTS

The **Concentration control coefficient** ($C^S_{E_i}$), a global property of the system, gives the relative fractional change in metabolite concentration $S_j(dS_j/S_j)$, where S_j is the concentration of any metabolite in the system and as such is a system variable, with fractional change in concentration or activity of enzyme $E_i(du_i/u_i)$. Similar equations as section 15.3 apply.

$$C^S_{E_i} = \frac{\frac{\partial S_i}{S_i}}{\frac{\partial u_i}{u_i}} = \frac{\partial \ln S_i}{\partial \ln u_i} = \frac{\partial S_i}{\partial u_i}\frac{u_i}{S_i} \tag{14.4.1}$$

It can be shown that the sum of all of the individual $C^S_{E_i} = 0$ (another summation theorem), is not 1 as in the case of flux control coefficients. This again would make sense in the steady state. Flux coefficients usually vary from 0 to 1, but concentration coefficients can vary from negative to positive and small to large.

$$\sum_{i=1}^{n} C^S_{Ei} = 0 \tag{14.4.2}$$

The concentration control coefficients can have large values as seen in a simple example. For a given enzyme, at low [S], for example, when [S] << Km,

$$v = \frac{V_m S}{K_M + S} = \frac{V_m S}{K_M} = \frac{k_{cat} E_{tot} S}{K_M} \text{ when } S \ll Km \tag{14.4.3}$$

If the enzyme had only 0.1x of its normal activity (due to a mutation for example), then to maintain constant flux, the [S] would have to increase 10-fold.

The tables below show the C^S_{Ei} values for incremental changes in the substrate (1%).

	(Hexokinase)	(Glucose-6-phosphate is...	(Glycogen synthesis)	(Trehalose 6-phosphate s...	(Phosphofructokinase)	(Aldolase)
High energy pho..	0.149291	0.00209315	-0.162615	-0.112466	0.00159405	0.000336889
Glucose 6 Phosp..	0.718017	-0.834548	-0.581964	-0.420826	-0.635556	-0.134319
Triose-phosphate	0.290279	0.0040699	-0.192953	-0.144737	0.00309946	0.000655043
NAD	0.000899872	1.26168e-05	-0.000702168	-0.000511095	9.60839e-06	2.03065e-06
Acetaldehyde	0.0928625	0.00130199	-0.0595875	-0.0450188	0.00099154	0.000209553
2-phosphoglycer..	0.579779	0.00812887	-0.417532	-0.308373	0.00619059	0.00130833
1,3-bisphosphog..	1.00165	0.0140438	-0.889801	-0.63383	0.0106951	0.00226032
Fructose 6 Phos...	0.946321	0.013268	-0.825856	-0.589942	-1.06108	-0.224249
Glucose in Cytoso	-2.15363	-0.0301953	0.295321	0.203672	-0.0229954	-0.00485988
Phosphoenolpyr...	0.751848	0.0105414	-0.614078	-0.44347	0.00802786	0.00169662
Pyruvate	0.458649	0.00643055	-0.261508	-0.202671	0.00489723	0.00103499
Fructose-1,6 bis...	0.567903	0.00796236	-0.369069	-0.278109	0.00606379	-0.525832
3-phosphoglycer..	0.535653	0.0075102	-0.37565	-0.27884	0.00571944	0.00120875
NADH	-0.0312961	-0.000438791	0.0244203	0.017775	-0.000334164	-7.06227e-05

table, part 2:

	(Glyceraldehyde 3-phosp...	(Phosphoglycerate kinase)	(Phosphoglycerate mutas...	(Enolase)	(Pyruvate kinase)	(Pyruvate decarboxylase)
High energy pho..	0.092012	0.00262286	0.00214191	0.00583266	0.00297472	0.00177918
Glucose 6 Phosp..	-0.194066	-0.00553197	-0.00451758	-0.0123019	-0.00627409	-0.00375253
Triose-phosphate	-0.968588	-0.0276102	-0.0225473	-0.061399	-0.0313141	-0.0187289
NAD	-0.00695116	-0.000198147	-0.000161813	-0.000440635	-0.000224729	-0.00013441
Acetaldehyde	-0.228632	-0.0065173	-0.00532223	-0.014493	-0.00739161	-0.00442091
2-phosphoglycer..	0.207334	0.00591018	0.00482644	-1.19124	-0.607546	-0.363372
1,3-bisphosphog..	0.47628	-0.337666	-0.275748	-0.750894	-0.382964	-0.22905
Fructose 6 Phos...	-0.344437	-0.00981839	-0.008018	-0.0218339	-0.0111355	-0.00666015
Glucose in Cytoso	-0.183506	-0.00523095	-0.00427175	-0.0116325	-0.00593269	-0.00354833
Phosphoenolpyr...	0.319778	0.00911547	0.00744398	0.0202708	-1.29924	-0.777074
Pyruvate	0.115796	0.00330085	0.00269558	0.00734036	0.00374367	-2.4298
Fructose-1,6 bis...	-1.5751	-0.0448992	-0.0366661	-0.0998459	-0.0509225	-0.0304567
3-phosphoglycer..	0.184471	0.00525847	-0.352935	-0.961081	-0.490162	-0.293165
NADH	0.24175	0.00689123	0.0056276	0.0153246	0.0078157	0.00467456

table, part 3

	e kinase)	(Pyruvate decarboxylase)	(Succinate synthesis)	(Glucose transport)	(Alcohol dehydrogenase)	(Glycerol 3-phosphate d...	(ATPase activity)
High energy pho..		0.00177918	-0.133971	0.945013	0.0419254	-0.217747	-0.620819
Glucose 6 Phosp..		-0.00375253	-0.338999	4.54506	0.106088	-0.620486	-1.57602
Triose-phosphate		-0.0187289	-0.0705359	1.83747	0.0220739	-0.280221	-0.339013
NAD		-0.00013441	-0.0177345	0.0056962	0.00554992	0.016672	-0.00178363
Acetaldehyde		-0.00442091	-0.662239	0.587821	-0.1057	0.549517	-0.0933806
2-phosphoglycer..		-0.363372	-0.299326	3.67001	0.0936725	-0.486218	-0.903552
1,3-bisphosphog..		-0.22905	-0.691066	6.34045	0.216266	-1.12294	-2.74768
Fructose 6 Phos...		-0.00666015	-0.53586	5.99023	0.167695	-0.986943	-2.49168
Glucose in Cytoso		-0.00354833	0.247704	0.208505	-0.0775176	0.400422	1.1477
Phosphoenolpyr...		-0.777074	-0.463155	4.75921	0.144942	-0.752507	-1.68334
Pyruvate		-2.4298	-0.165759	2.90326	0.0518733	-0.269095	-0.230185
Fructose-1,6 bis...		-0.0304567	-0.126035	3.59483	0.039442	-0.475374	-0.603896
3-phosphoglycer..		-0.293165	-0.266111	3.39069	0.0832782	-0.432242	-0.763605
NADH		0.00467456	0.616777	-0.198105	-0.193017	-0.579826	0.0620318

14.4.2: ELASTICITY COEFFICIENT

The **Elasticity Coefficient**, in contrast to the flux and concentration control coefficients, which are properties of the system, is a local property and can be measured using isolated enzymes and substrates. It makes sense that some kinetic property of the isolated enzyme would affect its propensity to affect system flux. The elasticity coefficient gives a measure of how much a substrate S (or other substance) can change the reaction rate (v) of an isolated enzyme. (Note we use v and not flux J, which describes a system property.) Hence

$$\varepsilon_S^v = \frac{\frac{\partial v}{v}}{\frac{\partial S}{S}} = \frac{\partial \ln v}{\partial \ln S} = \frac{\partial v}{\partial S}\frac{S}{v} \tag{14.4.4}$$

PRICE ELASTICITY

The term elasticity is also used in economics and is especially useful in times when inflation is high. If the price of your favorite product, such as a Starbucks Latte coffee goes up, consumers might either buy them less frequently or buy another cheaper latte from a competitor. If a small rise for a Starbucks latte leads to a large drop in demand, the Starbucks latte is characterized by a high elasticity. If however, people don't change their latte buying behavior when there is a big price rise on the latte, the product is said to be inelastic. If you are the CEO of a company, it is good to know the elasticity of your products to maximize your profits.

Hence the elasticity coefficient can be determined using basic enzyme kinetics of the isolated enzyme. Note that the coefficient at each S concentration is the slope of the v vs S curve multiplied by the S/v at that tangent point. The elasticity coefficient must be evaluated at the same concentration of enzyme and substrate as found in vivo in the steady state. Velocity (v), not flux, is used. In the above case, S is the substrate, but it could be a product or modifier. There is a different elasticity coefficient for each parameter. There is no summation theory for elasticities. Values can be positive for species that increase the velocity or negative for those that decrease it. Hence there can be multiple elasticities.

The tables below show the elasticity coefficients (relative or scaled) for yeast glycolysis determined using COPASI.

	High energy phos...	Glucose 6 Phosphate	Triose-phosphate	NAD	Acetaldehyde	2-phosphoglycerate	1,3-bisphosphogly...	Fructose 6 Phosph...
(Hexokinase)	1.02118	-0.0166	0	0	0	0	0	0
(Glucose-6-..	0	1.18397	0	0	0	0	0	-0.710951
(Glycogen s..	0	0	0	0	0	0	0	0
(Trehalose 6..	0	0	0	0	0	0	0	0
(Phosphofr...	-3.20462	0	0	0	0	0	0	0.93355
(Aldolase)	0	0	-3.10067	0	0	0	0	0
(Glyceralde...	0	0	0.919332	0.145002	0	0	-0.0811565	0
(Phosphogl...	-9.63644	0	0	0	0	0	2.84704	0
(Phosphogl...	0	0	0	0	0	-2.261	0	0
(Enolase)	0	0	0	0	0	0.8303	0	0
(Pyruvate ki...	-1.82687	0	0	0	0	0	0	0
(Pyruvate d...	0	0	0	0	0	0	0	0
(Succinate s..	0	0	0	0	1	0	0	0
(Glucose tra..	0	0	0	0	0	0	0	0
(Alcohol de..	0	0	0	-3.03483	3.19545	0	0	0
(Glycerol 3-..	0	0	0.38368	-0.36338	0	0	0	0
(ATPase act..	1.79865	0	0	0	0	0	0	0

table, part 2

	phoglycerate	1,3-bisphosphogly...	Fructose 6 Phosph...	Glucose in Cytosol	Phosphoenolpyruv...	Pyruvate	Fructose-1,6 bisph...	3-phosphoglycerate	NADH
(Hexokinase)		0	0	0.457337	0	0	0	0	0
(Glucose-6-..		0	-0.710951	0	0	0	0	0	0
(Glycogen s..		0	0	0	0	0	0	0	0
(Trehalose 6..		0	0	0	0	0	0	0	0
(Phosphofr...		0	0.93355	0	0	0	-0.400941	0	0
(Aldolase)		0	0	0	0	0	1.89712	0	0
(Glyceralde...		-0.0811565	0	0	0	0	0	0	-0.0921294
(Phosphogl...		2.84704	0	0	0	0	0	-2.28602	0
(Phosphogl...		0	0	0	0	0	0	2.79933	0
(Enolase)		0	0	0	-0.389446	0	0	0	0
(Pyruvate ki...		0	0	0	0.763604	-0.245925	0	0	0
(Pyruvate d...		0	0	0	0	0.411178	0	0	0
(Succinate s..		0	0	0	0	0	0	0	0
(Glucose tra..		0	0	-0.072249	0	0	0	0	0
(Alcohol de..		0	0	0	0	0	0	0	3.19614
(Glycerol 3-..		0	0	0	0	0	0	0	0.581052
(ATPase act..		0	0	0	0	0	0	0	0

Things to note:

- the columns show substrates not enzymes;
- green cells (with positive elasticities) are generally substrates for their target enzymes (for example, glucose-6-phosphate for phosphoglucomutase);

Other "generic" sensitivities can be determined as well. For example, incremental changes in K_m, V_m, or K_{ix} for specific enzymes could affect fluxes in a pathway.

A link between system control coefficients and local coefficients:

You would think that there should be some relationship between a system variable such as the flux control coefficient and a local variable such as the elasticity coefficient. There is and it is expressed by the Connectivity Theorem (below). If a substrate S is acted upon by many different enzymes ($i \ldots\ldots n$), which is very likely, especially for branch points in metabolic pathways, then it can be shown that

$$\sum_{i=1}^{n} C_{vi}^{J} e_{s}^{vi} = 0 \tag{14.4.5}$$

14.5: METABOLISM AND SIGNALING: THE STEADY STATE, ADAPTATION AND HOMEOSTASIS

14.5.1: INTRODUCTION

We have studied binding interactions in Chapter 5, kinetics in Chapter 6, and principles of metabolic control in this chapter. We've learned the following:

Binding Reactions

- for simple binding of a ligand to a macromolecule, graphs of fractional saturation of the macromolecule vs free ligand concentration are hyperbolic and demonstrate saturation binding. In the initial part of the binding curve, when [L] << K_D, the fractional saturation shows a linear dependence on free ligand concentration. Figure 14.5.1 shows [ML] vs L, which is the same basic equation as a plot of Y vs L.

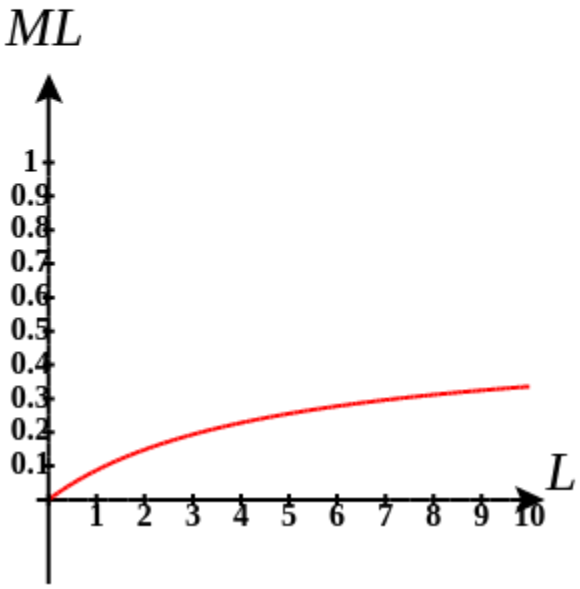


Figure 14.5.1

- for allosteric binding of a ligand to a multimeric protein, graphs of fractional saturation vs free ligand concentration are sigmoidal and also display saturation binding. In the first parts of the binding curve, the fractional saturation is much more sensitive to ligand concentration than in simple binding of a ligand to a macromolecule with one binding site. Figure 14.5.2 below shows graphs for the allosteric binding of a ligand to a macromolecule using the Hill Equation (instead of the MWC equation we used to model O_2 binding to tetrameric hemoglobin).

Figure 14.5.2

In these two plots, the system (in this case a single macromolecule) displays different sensitivities to ligand concentration, allowing the system to have different responses to changes in physiological conditions.

Binding and Chemical Reactions

As with the case for binding interactions, we have seen hyperbolic and sigmoidal plots of initial velocity (v_0) vs [substrate] for enzyme-catalyzed reactions. These also allow appropriate responses to a single substrate in a physiological setting.

But what if you put the same macromolecule and ligand into a larger metabolic or signal transduction pathway in vivo? What kinds of responses would they make to a change in input? As we have just seen in our discussion of the steady state, the ligand or substrate concentration might not change at all as flux continues through the pathway. One could imagine a lot of scenarios with different inputs and different optimal outputs. For example, what if the input (a reactant or small signaling molecule) comes in pulses? Ultimately a system should return to its basal state since a prolonged response (such as cell proliferation) could be detrimental to the health of the organism.

Let's look at some simple examples and see how different inputs lead to specific outputs. We'll just construct some very simple reaction diagrams in Vcell and see how varying them leads to different outputs. Here are two simple cases for **isolated** chemical species and reactions, analogs to the simple binding reactions described above.

Linear Response: A Signal S and a Response R; S → R

If no enzyme is involved, the rate doubles as the signal (substrate) doubles since dR/dt = k[S] for the first-order reaction. If S is the stimulus and R is the response, a plot of R vs S is linear. Hence the system responds linearly with increasing S. Here is the simple chemical equation

$$S \underset{k_2}{\overset{k_1}{\rightleftarrows}} R \tag{14.5.1}$$

As a concrete example, consider the synthesis and degradation of a protein, characterized by the following equation derived from mass action.

$$\frac{dR}{dt} = k_0 + k_1 S - k_2 R \tag{14.5.2}$$

where S is the signal (ex. concentration of mRNA) and R is the response (ex concentration of the transcribed protein). A constant k_0 has been added to account for any basal rate of the reaction. (This is a vastly oversimplified way to model a complex process like mRNA translation to a protein as it omits 100s of steps.)

Here is the simplified derivation under steady state (SS) conditions typically found for enzymes embedded in a pathway.

$$\begin{aligned} \frac{dR_{SS}}{dt} &= k_0 + k_1 S - k_2 R = 0 \\ R_{SS} &= \frac{k_0 + k_1 S}{k_2} \end{aligned} \tag{14.5.3}$$

The equation is a linear function of S.

Hyperbolic Response: E+S ↔ ES → E + R

In a simple enzyme-catalyzed reaction with a fixed concentration of enzyme, as S increases the initial velocity saturates. Hence there is a limit on the response, so the response R is a hyperbolic function of S. Increasing S ever more after saturation won't lead to more R (in a given amount of time).

As a concrete example of this consider the phosphorylation/dephosphorylation of a protein R. R_P represents the phosphorylated and active form of the protein R with concentration $[R_P]$. The reaction is simply written as R ↔ R_P, where R_P is the response. Mass action shows that the total amount of R, $R_T = R + R_P$. A simple mass action equation can be derived.

Here is the chemical equation

$$\mathrm{R} + \mathrm{S} \underset{k_2}{\overset{k_1}{\rightleftarrows}} \mathrm{R_P} \tag{14.5.4}$$

Here is the math equation, again for the steady state (SS), when $dR_P/dt = 0$. (We derived the same equation for the steady-state version of the Michaelis-Menten equation in Chapter 6.

$$\frac{dR_P}{dt} = k_1 S\left(R_T - R_P\right) - k_2 R_P \tag{14.5.5}$$

? DERIVATION: STEADY STATE HYPERBOLIC RESPONSE TO A STIMULUS S

Click below to see the derivation

Derivation

$$\frac{dR_P}{dt} = k_1 R[S] - k_2 R_P \tag{14.5.6}$$

then in the stead state:

$$\begin{aligned} \frac{dR_P}{dt} &= k_1 S\left(R_T - R_P\right) - k_2 R_P = 0 \\ k_2 R_{P,SS} &= k_1 S\left(R_T\right) - k_1 S\left(R_{P,SS}\right) \\ k_2 R_{P,SS} + k_1 S\left(R_{P,SS}\right) &= k_1 S\left(R_T\right) \\ R_{P,SS}\left(k_2 + k_1 S\right) &= k_1 S\left(R_T\right) \end{aligned} \tag{14.5.7}$$

Finally, we get

$$R_{P,Ss} = \frac{k_1 S\left(R_T\right)}{\left(k_2 + k_1 S\right)} = \frac{\left(R_T\right) S}{\left(\frac{k_2}{k_1} + S\right)} \tag{14.5.8}$$

In the steady state, $dR_P/dt = 0$, and the steady state equation can be written as:

$$R_{P,ss} = \frac{k_1 S\left(R_T\right)}{\left(k_2 + k_1 S\right)} = \frac{\left(R_T\right) S}{\left(\frac{k_2}{k_1} + S\right)} \tag{14.5.9}$$

Sigmoidal Response

Consider this simple reaction for a homotetramer in which each monomer can bind a substrate S: $nS + E_n \leftrightarrow E_nS_n \rightarrow E_n + nR$: If E_n is a multimeric allosteric enzyme, as S increases the initial velocity also saturates but the response R is a sigmoidal function of S (in analogy to the above example). The equation is too complicated to derive there, but the result reproduces a sigmoidal curve for the steady state, much as the Hill equation does for cooperative binding.

14.5.2: ADAPTATION AND HOMEOSTASIS

The above examples show that the response of proteins or enzymes to increasing levels of a stimulus like a ligand or a substrate can be linear, hyperbolic, or sigmoidal, with quite a varied set of outcomes. However, in many biological conditions, an ever-increasing or increasing and plateauing response might be too much. The cell needs a way to turn off the response and settle back to a basal state, even in the presence of constant or changing stimuli. This allows the **adaption** of a system to a stimulus and the maintenance of **homeostasis**. Every system needs to be able to respond and return to a homeostatic basal level. The maintenance of homeostasis is critical to life.

HOMEOSTASIS - ASBMB

The American Association for Biochemistry and Molecular Biology (ASBMB) describes both homeostasis and evolution as key underlying concepts for all biology. Homeostasis shapes both form and function from the molecular to organismal levels. Homeostasis is needed to maintain biological balance. The steady state at the molecular to organismal levels in metabolic and signaling pathways is a hallmark of homeostasis. Here are the learning goals for homeostasis designated by the ASBMB

1. Biological need for homeostasis

Biological homeostasis is the ability to maintain relative stability and function as changes occur in the internal or external environment. Organisms are viable under a relatively narrow set of conditions. As such, there is a need to tightly regulate the concentrations of metabolites and small molecules at the cellular level to ensure survival. To optimize resource use, and to maintain conditions, the organism may sacrifice efficiency for robustness. The breakdown of homeostatic regulation can contribute to the cause or progression of disease or lead to cell death.

2. Link steady-state processes and homeostasis

A system that is in a steady state remains constant over time, but that constant state requires continual work. A system in a steady state has a higher level of energy than its surroundings. Biochemical systems maintain homeostasis via the regulation of gene expression, metabolic flux, and energy transformation but are never at equilibrium.

3. Quantifying homeostasis

Multiple reactions with intricate networks of activators and inhibitors are involved in biological homeostasis. Modifications of such networks can lead to the activation of previously latent metabolic pathways or even to unpredicted interactions between components of these networks. These pathways and networks can be mathematically modeled and correlated with metabolomics data and kinetic and thermodynamic parameters of individual components to quantify the effects of changing conditions related to either normal or disease states.

4. Control mechanisms

Homeostasis is maintained by a series of control mechanisms functioning at the organ, tissue, or cellular level. These control mechanisms include substrate supply, activation or inhibition of individual enzymes and receptors, synthesis and degradation of enzymes, and compartmentalization. The primary components responsible for the maintenance of homeostasis can be categorized as stimulus, receptor, control center, effector, and feedback mechanism.

5. Cellular and organismal homeostasis

Homeostasis in an organism or colony of single-celled organisms is regulated by secreted proteins and small molecules often functioning as signals. Homeostasis in the cell is maintained by regulation and by the exchange of materials and energy with its surroundings.

In the rest of the chapter section, we will describe chemically and mathematically simple circuits/motifs that are employed that allow perfect or near-perfect adaptation to a stimulus, a hallmark of homeostasis. We will define adaptation as a complete or almost complete return to a basal state after the introduction of a stimulus. In all the cases below we will consider not a single application of a stimulus but a pulse application (a repetitive step wave function). The pulsed stimuli could be of constant magnitude or an increasing/decreasing pulse of a signal such as a substrate. All responses must be transient to avoid uncontrolled responses such as proliferation (a hallmark of tumor cells) or cell death.

Adaptation is commonly found in sensory systems like vision, hearing, pressure, taste, etc. Think of eating your favorite cookie. The first bite is delicious but by the tenth bite, there is significant attenuation in the positive sensory response, which helps keep most from adding significant weight continually.

Ma et al. conducted simulations on three component/nodes (proteins, enzyme) systems to see which might display the potential for perfect or near-perfect adaption. The simple 3-component motifs or circuits were modeled using simple mass action kinetic equations, ordinary differential equations (which we learned to write in Chapter 6.2), or a combination of both. The systems that displayed adaption had to conform to three criteria:

1. The stimulus had to initially induce a response of high magnitude
2. The system had to return to a basal or near basal state.
3. The return to a basal state had to be mostly parameter-independent. That is, the return to the basal state must occur for many different combinations of parameters.

The possible 3-component components (nodes) and the links among the nodes are shown in Figure 14.5.3 below.

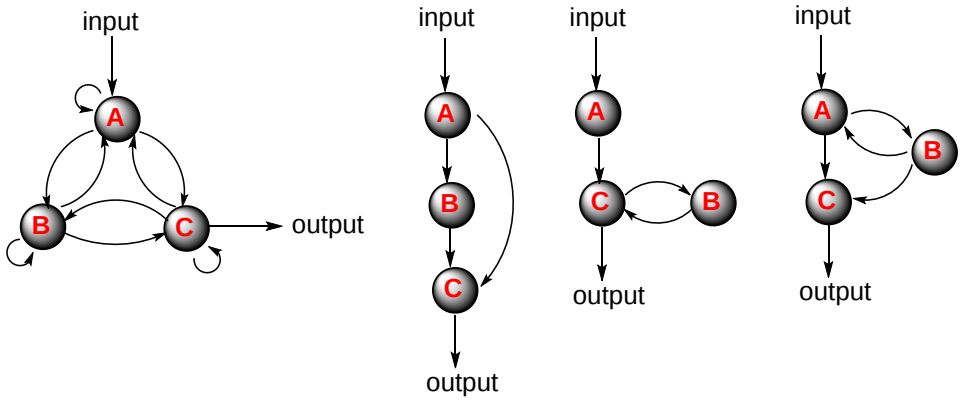


Figure 14.5.3: Possible 3-component components (nodes) and the links among the nodes. After Ma et al. Cell Theory,138, 760-773 (2009) https://www.cell.com/fulltext/S0092-8674(09)00712-0. DOI:https://doi.org/10.1016/j.cell.2009.06.013.

Out of over 16,000 models, several hundred were found that met the criteria. Most were variations of simple motifs that we will show below. The most common motifs were the **negative feedback loop** and the **incoherent feedforward system**.

Much of the discussions, models, and equations used below are from two articles:

- John J Tyson, Katherine C Chen, Bela Novak, Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell, Current Opinion in Cell Biology, Volume 15, Issue 2, 2003, Pages 221-231, https://doi.org/10.1016/S0955-0674(03)00017-6.
- James E. Ferrell, Perfect and Near-Perfect Adaptation in Cell Signaling, Cell Systems, Volume 2, Issue 2, 2016, Pages 62-67, https://doi.org/10.1016/j.cels.2016.02.006.

By adding a third component to form a mini pathway, we can now change the response R to a stimulus S from linear, or hyperbolic/sigmoidal in the steady state, to one that exhibits perfect or near-perfect adaptation. Again we see this kind of response in signaling pathways in sensation and also in responses like chemotaxis, in which a cell moves toward a stimulus (a chemoattractant molecule).

14.5.2.1: SIMPLE 3-NODE MOTIF/CIRCUIT FOR PERFECT ADAPTATION

Figure 14.5.4 below shows our first example of a 3-component system that displays perfect or near-perfect adaption. The right-hand side shows a Vcell reaction diagram. In this example, a stimulus S (could be a reactant, neurotransmitter, mRNA, etc) leads to the synthesis of X and also of R, a response molecule. Both X and R get degraded. The yellow squares represent the nodes through which the flux of S to X and R proceeds. Each node has an equation for the flux, J, through the node. The left part of Figure 4 shows the periodic pulse of stimuli S that increases the concentration of S from an initial value of S_0 = 1 uM to S + 0.2 uM for each step. Note that the flux equations for J are very simple and are based on mass action, and are not derived through Michaelis-Menten kinetic equations.

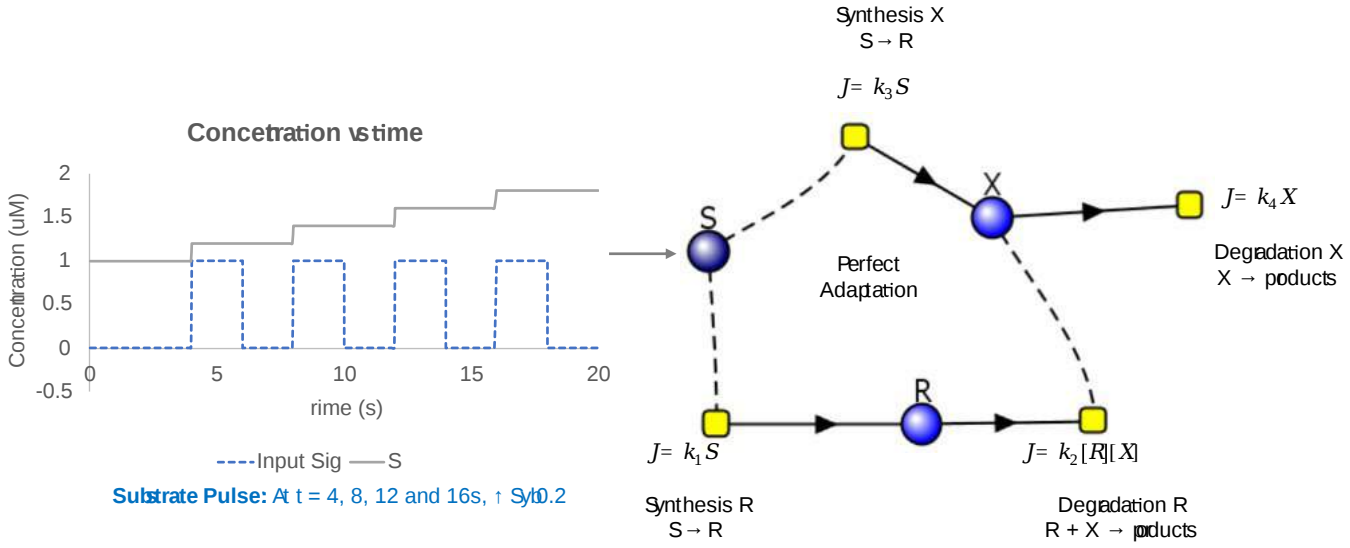


Figure 14.5.4: Simple 3-component system that displays perfect or near-perfect adaption.

Note that S, the stimulus (or substrate for example) is a square wave step function varying from 0 to 1 over the time interval shown in the graph. The dotted blue line simply shows when the pulse is delivered. The initial S concentration is 1 uM and increases by 0.2 for each step (as shown in the gray line). Hence S increases in a stepwise fashion.

Figure 14.5.5 below is a time course graph that shows the stepwise (=0.2 uM) increase in S from 1 uM and the concentration of R (the response) over 20 seconds. Even though S continues to increase in a stepwise fashion, R rises substantially only from the initial input of S (1 uM) and subsequent increases in S with each increment of S are damped out!

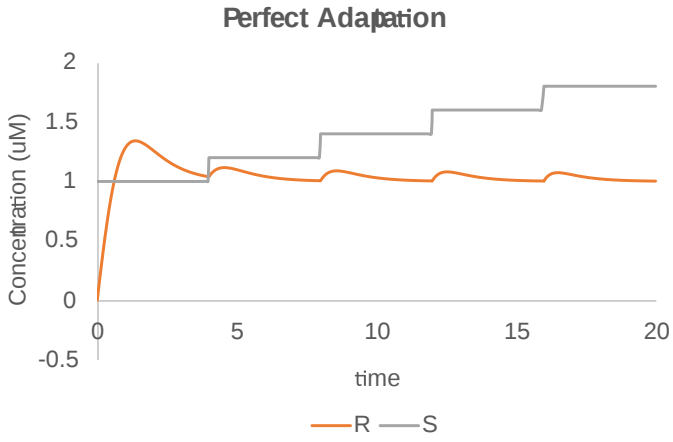


Figure 14.5.5: Time course for a 3-Component Perfect Response system. Model by ModeBrick from VCell: CM-PM12648679_MB4:Perfect_Adaptation; Biomodel 188456707

RUNNING ADAPTIVE VELL MODELS IN THIS BOOK

The present version of Vcell release (as of 4/28/23) does not yet allow the export of a file compatible with the software used to run simulations with this book. The Vcell model includes an "event" which allows for the production of stepwise changes in stimuli. A future release will allow users to run the simulations within this book (as is the case for the other Vcell simulations throughout the book).

14.5.2.2: NEGATIVE FEEDBACK LOOP

The negative feedback loop is one of the simplest circuits/motifs to generate perfect or near/perfect adaptation. It has only two nodes (yellow dots) and two proteins. An example is bacterial chemotaxis. Figure 14.5.6 below shows a Vcell reaction diagram (left), and another representation (middle) and the time course graphs for all species. This model works especially well with certain parameters assigned.

Figure 14.5.6

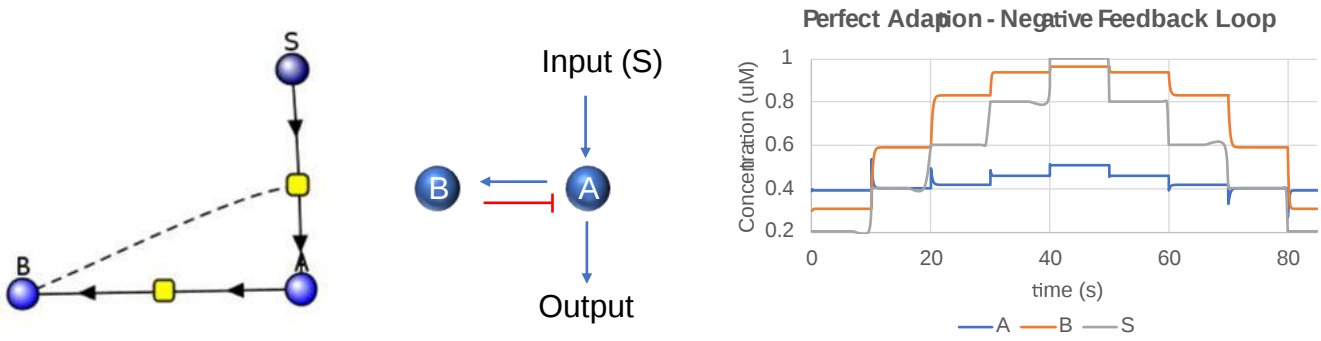


Figure 14.5.6: Near-Perfect Adaptation from Negative Feedback. Adapted from Ferrell (ibid)

The gray line in the graph is the stimulus S (substrate). The blue line is the response, designated in this model as A. B acts as an inhibitor (note the dotted line to the input node in the left diagram and the blunt-ended red bar in the middle diagram. Note that the stimulus goes from 0.2 uM (initial concentration) at t=0 to 1 uM (a 5-fold increase) at 40 seconds, but the response A increases at most from 0.4 (initial condition) to 0.5 (a 1.25-fold increase).

If we say the [A] is the output, then the differential equation for dA/dt is given by

$$\frac{dA}{dt} = k_1 \, \mathrm{S} \cdot (1 - A) - k_2 A \cdot B \tag{14.5.10}$$

dB/dt is given by

$$\frac{dB}{dt} = k_3 A \frac{1 - B}{K_3 + 1 - B} - k_4 \frac{B}{K_4 + B} \tag{14.5.11}$$

The constants for the graph (right) produced by the Vcell model are:

- $k_1 = k_2 = 200$
- $k_3 = 10$; $k_4 = 4$
- $K_3 = K_4 = 0.01$

14.5.2.3: INCOHERENT FEEDFORWARD SYSTEMS

In this circuit/motif, the stimulus S increases the concentration of A (the output) but also forms a negative modulator, B, which with a bit of a time lag decreases the concentration of A through inhibition. There is no feedback inhibition from A in this simple system. If you're reading carefully, you'll see that the reaction scheme and inhibition are the same as the first circuit/motif we introduced. Here we simplify the diagram and give it an official name. The word incoherent in the name makes sense since the stimulus S is converted both to the output A, and to the inhibitor B, which on the surface seems like a crazy thing to do.

Figure 14.5.7 below shows the Vcell reaction diagram (left) and more classical reaction diagram (middle) and progress curves showing S, the stimulus, A the output or response, and B, the inhibitor. The dashed line in the left diagram from B to the reaction node for the S → A reaction shows that B affects the rate of that reaction. The equations used account for the inhibitory effect of B.

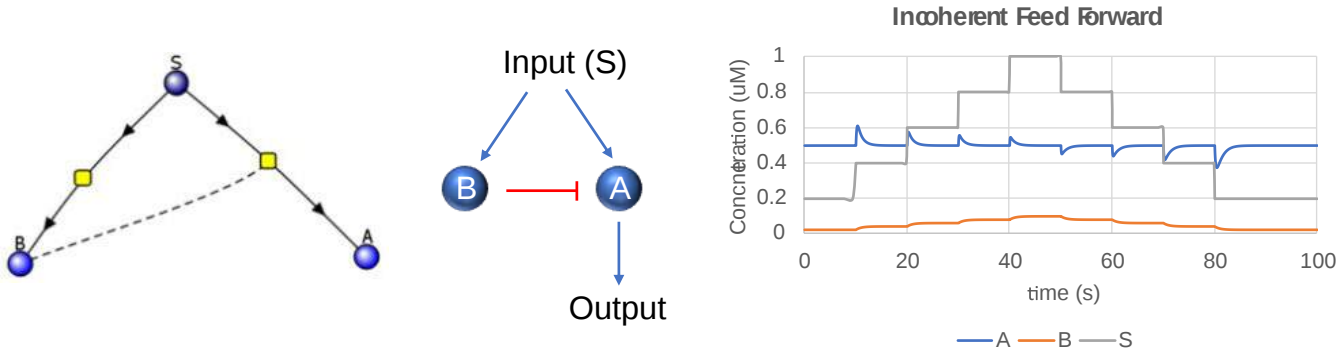


Figure 14.5.7: Near-Perfect Adaptation from an Incoherent Feedforward System. Adapted from Ferrell (ibid)

Note that the response A goes up or down a bit with each new step in concentration of S but to a very minimal degree. The system is certainly almost perfectly adapted.

The differential equation for dA/dt (where A is the response) is

$$\frac{dA}{dt} = k_1 \, \mathrm{S} \cdot (1 - A) - k_2 A \cdot B \tag{14.5.12}$$

The equation for dB/dt (the inhibitor generated from A) is

$$\frac{dB}{dt} = k_3 \, \mathrm{S} \, \frac{1 - B}{K_3 + 1 - B} - k_4 B \tag{14.5.13}$$

The constants for the graph (right) produced by the Vcell model are:

- $k_1 = 10$; $k_2 = 100$
- $k_3 = 0.1$; $k_4 = 1$
- $K_3 = 0.001$

14.5.3: STATE-DEPENDENT INACTIVATIONS SYSTEMS.

There are two simple circuits/motifs in this system that were found after the initial analyses that showed all possible interactions in a 3-component system (see Figure 3). The motif was patterned after the inhibition of proteins in neuron stimulation, specifically in ion channels in neural cell membranes that open up on a change in the transmembrane potential but then close again quickly to avoid constant neuronal stimulation (or inhibition). In the Na^+ ion channel, there are both fast (1-2 ms) and slow (100 ms) inactivation mechanisms. The fast one allows for repetitive firing, the development of action potentials, and the control of the excitation of neurons, and at the neuromuscular junction. Neuronal signaling is discussed in Chapter 28.9. Figure 14.5.8 below shows a simplified model for one type of inactivation of the Na^+ ion channel

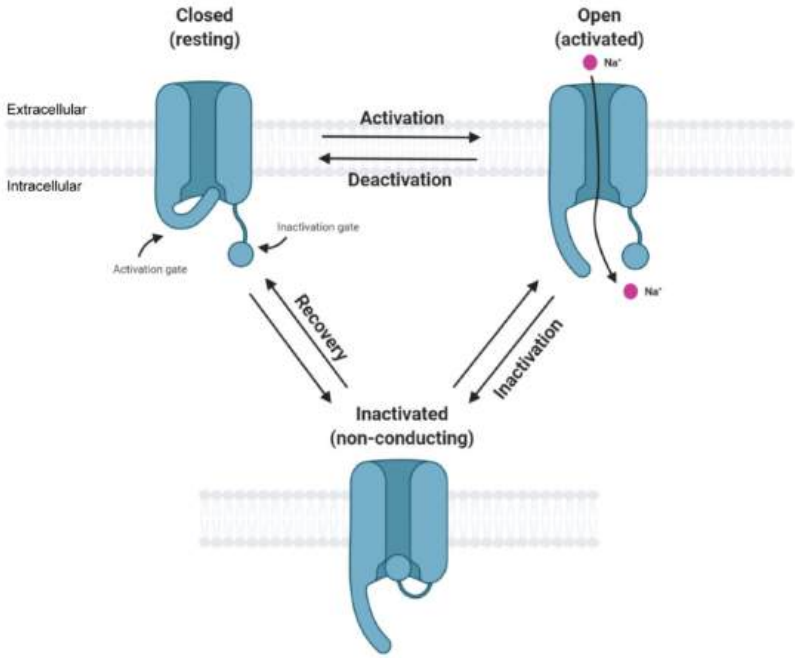


Figure 14.5.7: Simplified state transition model of voltage-gated sodium channels featuring closed, open, and inactivated states. Zybura, A. et al. *Cells* **2021**, *10*, 1595. https://doi.org/10.3390/cells10071595. Creative Commons Attribution (CC BY) license (**https://creativecommons.org/licenses/by/4.0/**).

The figure implies that there are at least 3 conformational states of the channel so the inactivation for the channel and the circuit/motif for adaptation we will now discuss are called **state-dependent inactivations**. The slow return to the original state is observed in many ion channels as well as in the return of G protein-coupled receptors to the normal state after their desensitization. Also, some protein kinases (kinases that use ATP to phosphorylate protein substrates) can be inactivated by internalizing the membrane kinase into vesicles where they can be reactivated and returned to the plasma membrane in a slow process.

For the construction of a perfect or near-perfect adaption state, we will assume the protein A exists in an off state (A_{off}) which binds the stimulus (B or S), an on state (A_{on}) which is viewed as the response (or A produces the response), and an inactivated state (A_{in}) which slowly reverts to the A_{off} state which can be activated again. The inactive state can be produced by conformational transitions with the

protein itself or another molecule produced downstream of it in a metabolic or signaling pathway. For example, a GPRC could be phosphorylated or bind to another species to produce an inactive state.

There are two different circuits/motifs that can produce state-dependent inactivation. We'll refer to these as Type A and Type B

Type A

Figure 14.5.9 shows the Vcell reaction diagram (top left), a classical reaction diagram (bottom left) and time course graphs for Type A state-dependent inactivation.

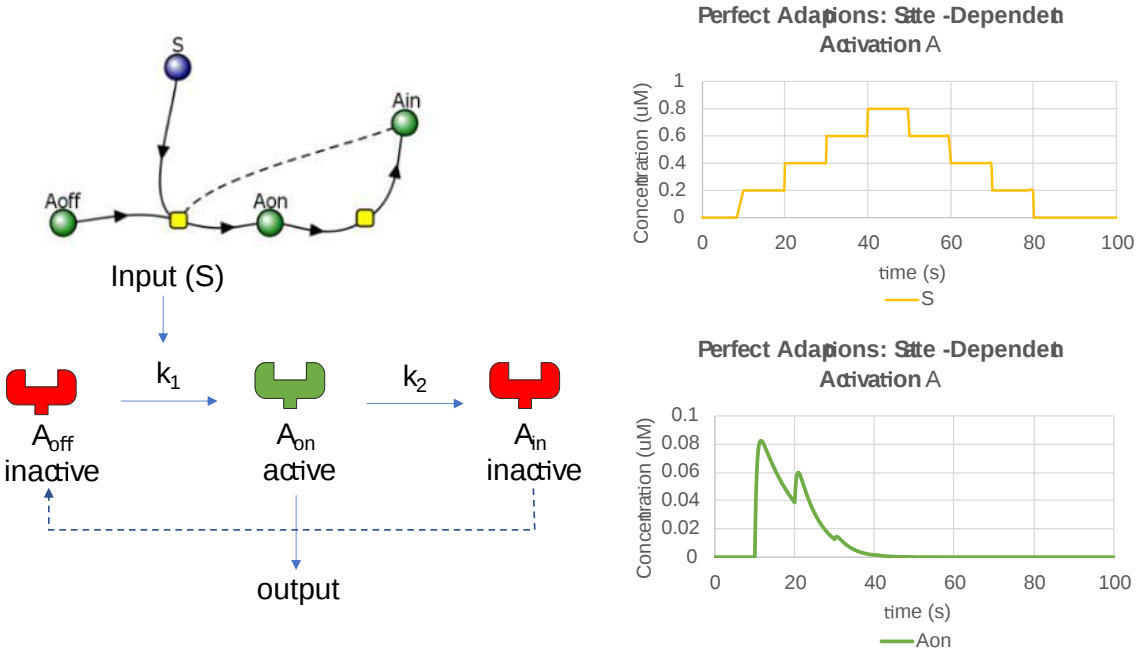


Figure 14.5.9: Perfect Adaptation for Type A State-Dependent Inactivation. Adapted from Ferrell (ibid).

A_{on} represents the active state of the protein. This mechanism applies well to the Na^+ channel. The differential equations for dA_{on}/dt and dA_{off}/dt are shown below.

For dA_{on}/dt

$$\frac{dA_{on}}{dt} = k_1 \,\text{Input} \cdot (1 - A_{on} - A_{in}) - k_2 A_{on} \tag{14.5.14}$$

dA_{in}/dt

$$\frac{dA_{in}}{dt} = k_2 A_{on} \tag{14.5.15}$$

with constants $k_1 = k_2 = 1$.

Again, as with the other cases, the stimulus S is pulsed. The different colors in the bottom left reaction diagram imply an off and inactive red state and a green active state, each of different conformations. The graphs were produced using Vcell. There is a slight anomaly in the graph of A_{on} which shows two additional small peaks as the system returns to the basal state. This contrasts to just 1 peak which returns to the basal state in a simple exponential fashion as described in the Ferrell paper. We are uncertain as to the source of the discrepancy.

Type B

In this case, the periodic stimulus, abbreviated as B, is a binding partner for A_{off} which produces an active complex B-A_{on}. Figure 14.5.10 below shows the Vcell reaction diagram (top left), a classical reaction diagram (bottom left), and time course graphs for Type B state-dependent inactivation

Figure 14.5.10

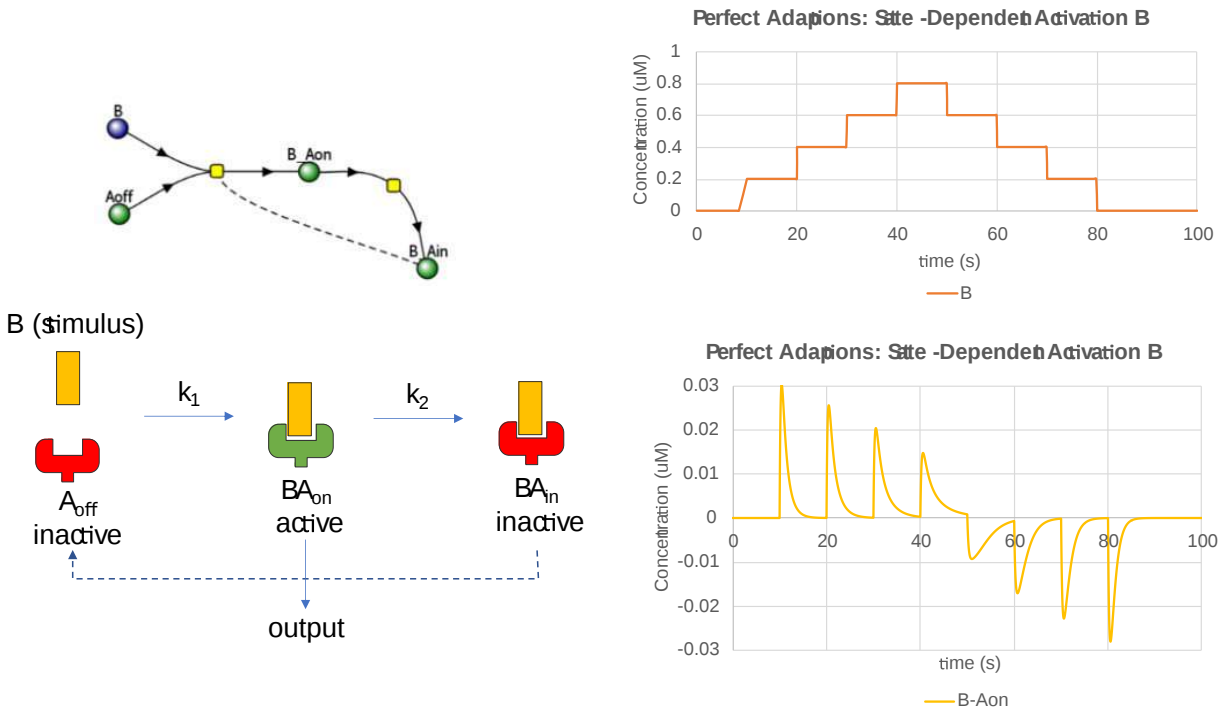


Figure 14.5.10: Perfect Adaptation for Type B State-Dependent Inactivation. Adapted from Ferrell (ibid)

BA_{on} represents the active state of the protein bound to B while BA_{in} represents the inactive complex.

The equation of dBA_{on}/dt for the formation of the active state is

$$\frac{dBA_{on}}{dt} = k_1 \left(B_{tot} - BA_{on} - BA_{in}\right) * \left(1 - BA_{on} - BA_{in}\right) - k_2 BA_{on} \tag{14.5.16}$$

and the equation for dBA_{in}/dt for the formation of the inactive state is

$$\frac{dA_{in}}{dt} = k_2 AB_{on} \tag{14.5.17}$$

with constants $k_1 = k_2 = 4$.

The graphs (note the different time concentration scales on the left) show a fairly quick return to the basal state after each pulse of stimuli (B).

CHAPTER OVERVIEW

15: GLUCOSE, GLYCOGEN, AND THEIR METABOLIC REGULATION

15.1: Insulin Signaling in the Liver
15.2: Glycogenesis
15.3: Glycogenolyis and its Regulation by Glucagon and Epinephrine Signaling
15.4: Regulation of Glycolysis
15.5: Regulation of Gluconeogenesis

15.1: INSULIN SIGNALING IN THE LIVER

15.1.1: INTRODUCTION

In this section, we will discuss insulin signaling and glycogen synthesis. Insulin is released in the fed state, and leads to glucose uptake where it can be stored, if not needed, for glycogen synthesis. Recall that glycogen is a large polymer of glucose residues connected in the main chain by alpha 1 → 4 linkages and with branching side chains about every 12 – 15 residues at the alpha 1 → 6 positions. The reducing ends of the carbohydrate (two for each polymer) are connected to the glycogenin dimeric protein at the center of the macromolecule (Figure 15.1.1). With the branching nature of the polymer, many non-reducing ends of the molecule are present, allowing easy access and fast release of glucose for energy utilization.

Most of the body's pools of glycogen are stored in the liver, with 10% of the liver biomass in glycogen granules, and in the skeletal muscle, with glycogen comprising 2% of the biomass of the muscle. Each glycogen polymer may have upwards of 30,000 glucose residues, making the glycogen polymer visible using standard microscopic techniques. Storage of glycogen within muscle tissue is used by the muscle cells as a source of energy to fuel muscle contraction. In the liver, the purpose of glycogen storage is different. Glycogen stored at this location is used to maintain the homeostatic balance of blood glucose levels. The liver is the primary organ that can actively transport glucose into the bloodstream. Our only other major source of glucose within the blood is our diet.

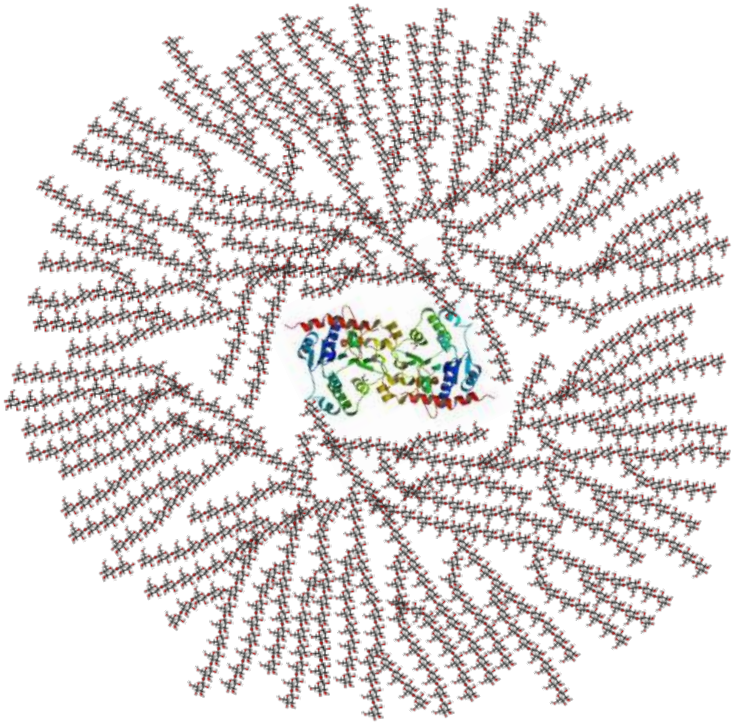


Figure 15.1.1: The Glycogen Polymer. The glycogen-reducing ends of the glycogen polymer are connected with the dimeric glycogenin protein shown at the center of the diagram using the ribbon projection. The polymer is constructed predominantly of glucose monomers connected through alpha 1 --> 4 linkages. Alpha 1 --> 6 branches are apparent about every 12 - 15 residues. (Public Domain; Mikael Häggström via Wikipedia.

Blood glucose homeostasis is critical for brain function (Figure 15.1.2). The brain has a huge energy demand, but nearly zero storage of key energy molecules required for ATP production. Furthermore, glucose and ketone bodies are the only energy sources that can pass the blood-brain barrier and be utilized by the brain for ATP production. Note that ketone bodies are only produced during starvation, or disease states such as diabetes, and are not a regular source of energy for the brain. Thus, glucose is critical for brain function. Nearly 10% of the whole body's energy is used for nerve impulse transmission by the brain. If blood flow to the brain carrying critical oxygen and glucose is impeded, people will lose consciousness within approximately 20 seconds! And brain death/permanent damage occurs within 4 minutes of blood flow cessation. This exemplifies the importance of the liver in maintaining blood glucose levels, as well as the importance of oxygen maintenance.

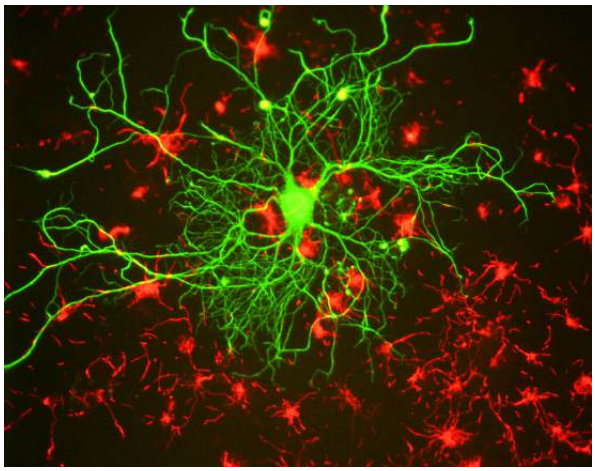


Figure 15.1.2: Cortical neuron stained with antibody to neurofilament subunit NF-L in green. In red are neuronal stem cells stained with antibodies to alpha-internexin. Image created using antibodies from EnCor Biotechnology Inc. (CC BY-SA 3.0 Unported; GerryShaw via Wikipedia)

Glycogen in the liver or muscle can be broken down into glucose 1-phosphate (Figure 15.1.3). This can be interconverted to glucose 6-phosphate which is then readily used in many cellular processes. The process of glycolysis (or the breakdown of glucose into pyruvate) occurs in all cells and produces a small amount of ATP in the process. Further processing of pyruvate can occur anaerobically (or in the absence of oxygen) to produce lactate, or the process can continue to occur in the aerobic pathway to complete oxidation to carbon dioxide and water in the Kreb cycle. Note that oxygen from breathing is used to create the water within this pathway. This fuels the process of oxidative phosphorylation within the mitochondria and produces large quantities of ATP (from 30-36 molecules/glucose). Within the liver, glucose can be freed from glycogen and released back into the bloodstream to maintain homeostatic levels. Glucose 6-phosphate can also be utilized as a precursor for other major macromolecules such as ribose and deoxyribose, as well as the hexosamine compounds commonly found cushioning joints or attached to proteins of the plasma membrane.

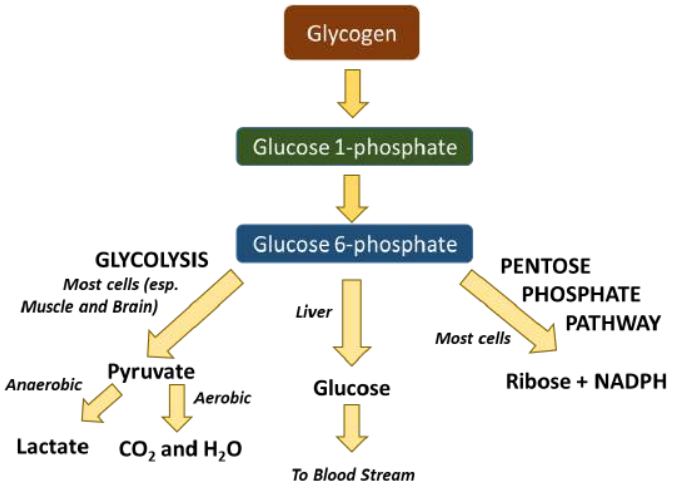


Figure 15.1.3: Pathways of Glycogen Utilization.

In healthy individuals, hormone signaling is critical to maintaining blood glucose homeostasis. Within this system, the hormones glucagon and insulin work together to maintain normal plasma glucose levels (Figure 15.1.4). During hyperglycemia, pancreatic beta (β) cells release insulin, which stimulates glucose uptake by energy-consuming cells and the formation of glycogen in the liver. During hypoglycemia, pancreatic alpha (α) cells release glucagon, which stimulates gluconeogenesis and glycogenolysis in the liver and the release of glucose to the plasma.

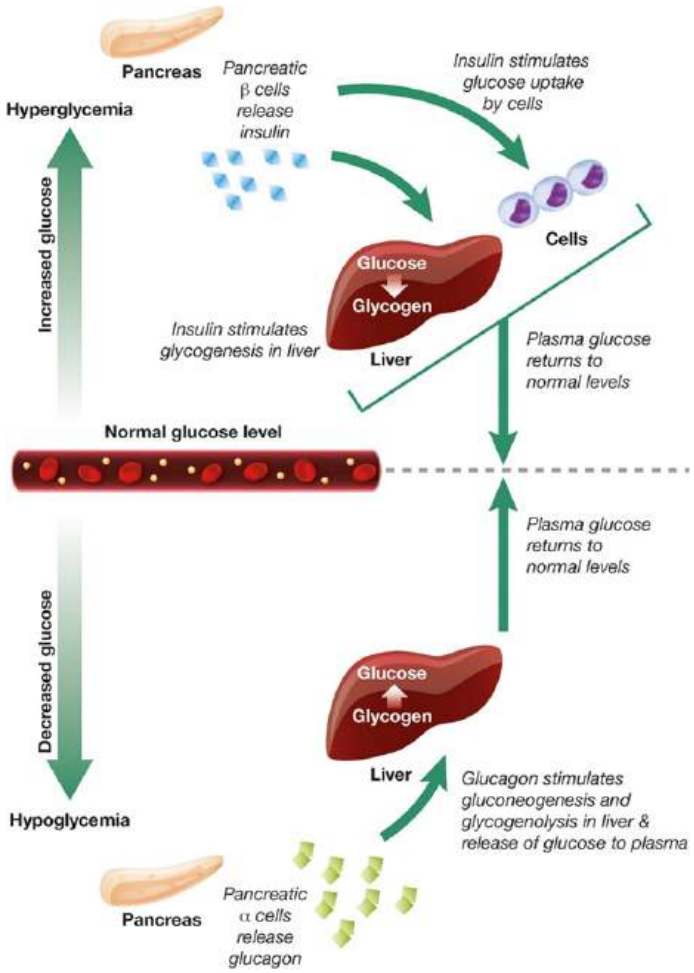


Figure 15.1.4: Hormone Signaling Involved in Blood Glucose Homeostasis. The hormones, glucagon, and insulin, released by the pancreas are critical for regulating the release or the uptake of glucose from the bloodstream by the liver. In the *upper diagram*, when blood glucose levels rise, beta cells within the pancreas secrete the insulin hormone. Insulin signaling causes the liver to increase the uptake of glucose from the bloodstream and promotes its storage as glycogen. Alternatively, when blood glucose levels are low, as shown in the *lower diagram*, alpha cells in the pancreas release the hormone, glucagon. Glucagon signaling in the liver leads to the breakdown of glycogen and the release of stored glucose into the bloodstream. Hædersdal, S., et al (2018) Mayo Clinic Proceedings 93(2):217-239

The first area we will focus our attention on will be the mechanism utilized by insulin to reduce blood glucose levels. Figure 15.1.5 shows the structure of the pancreas and its anatomical relationship with the liver and the stomach. The pancreas is the sensor organ that detects blood glucose levels. It is responsible for signaling to the liver to either remove or release glucose in response to changing levels. Notably, the pancreas also produces most of the digestive enzymes utilized by the body, including proteases, amylases, and lipases.

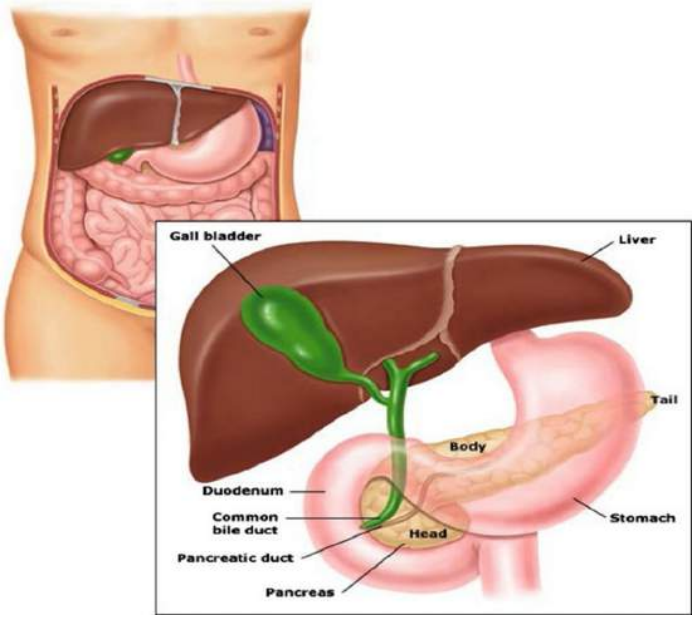


Figure 15.1.5: The Anatomy of the Liver and Pancreas. Akinlade, A., et al (2014) Int. Archives of Med 7(50):28

Figure 15.1.6: shows a light microscope image of the pancreatic islet cells. They are responsible for the production of glucagon and insulin. The islets are distinguished from the surrounding tissue by a continuous connective tissue capsule and extensive vascularity.

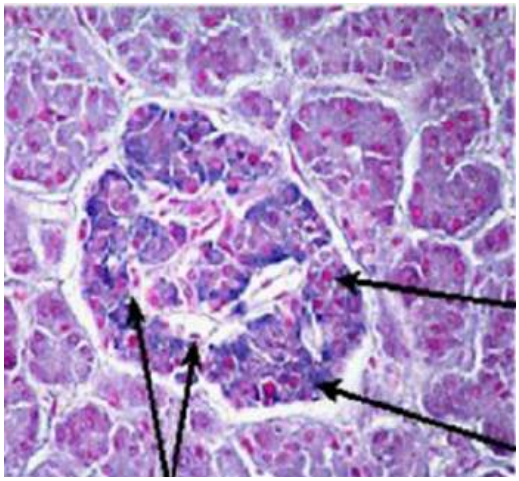


Figure 15.1.6: Pancreatic Islet Cells. In this diagram, the islet cells that secrete glucagon (known as alpha cells) are stained in red, while the beta cells that produce insulin are stained in blue. Akinlade, A., et al (2014) Int. Archives of Med 7(50):28

Insulin is a peptide hormone composed of 51 amino acids as shown in Figure 15.1.7. It is initially synthesized as preproinsulin, which is converted to proinsulin after the signal peptide is removed. Two disulfides are made in the ER (catalyzed by protein disulfide isomerase) along with selective proteolytic cleavage to form insulin. Mature insulin consists of A and B chains that were connected in proinsulin by the C-peptide.

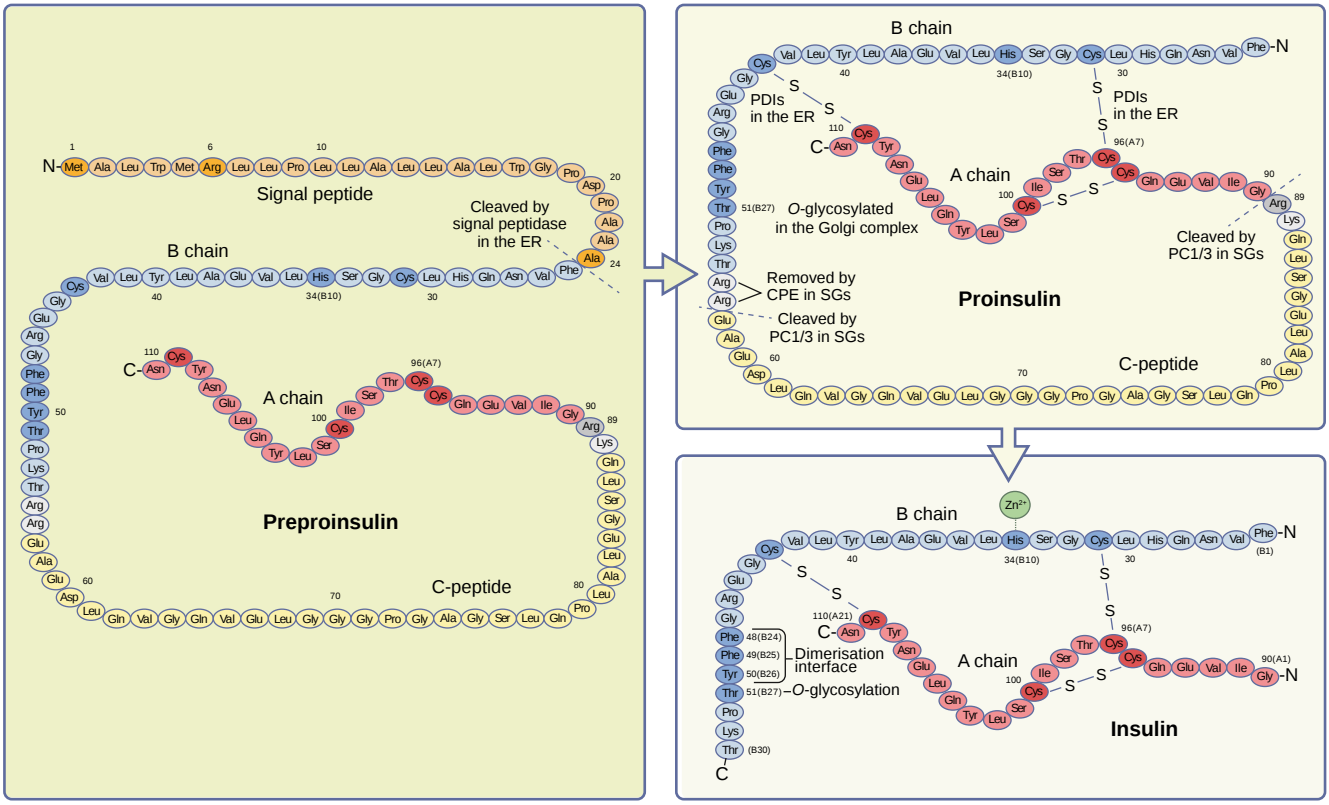
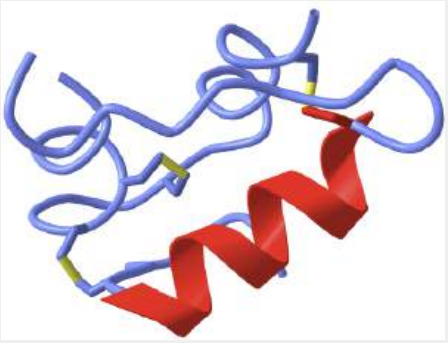


Figure 15.1.7: Conversion of preproinsulin to mature insulin. Vasiljević, J. *et al. Diabetologia* **63**, 1981–1989 (2020). https://doi.org/10.1007/s00125-020-05192-7. http://creativecommons.org/licenses/by/4.0/.

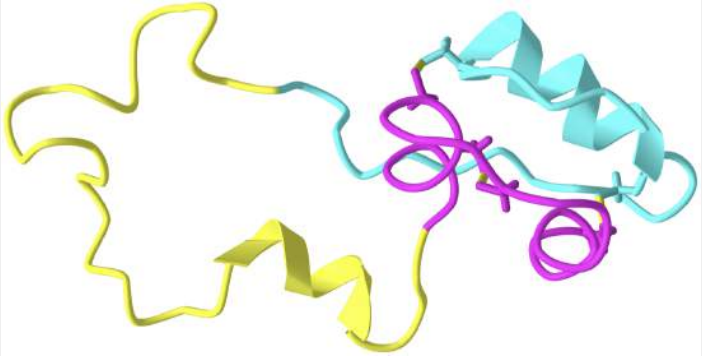
Figure 15.1.8 shows interactive iCn3D models of human insulin (3I40) and human proinsulin (2KQP). (Copyright; author via source). Click the image for a popup or use the external link provided:

Human insulin (3I40)



Colored code to show secondary structure
External link: https://structure.ncbi.nlm.nih.gov/i...A3QZdZpfEMxR17

Human Proinsulin (2KQP)



The "future" A chain in mature insulin is shown in magenta, the C-peptide connecting the A and B chains is shown in yellow, and the future B chain is in cyan.
External link: https://structure.ncbi.nlm.nih.gov/i...nQJXhrEfHD1GH9

The maturation of preproinsulin to insulin is shown in more detail in Figure 15.1.9. The peptide is first translated on ribosomes linked to the rough endoplasmic reticulum (ER), where a signal peptide docks the peptide to the ER membrane. The proinsulin is folded and the signal peptide is cleaved. It is transported to the Golgi where it is further packaged into secretory vesicles. Within the secretory vesicles, the proinsulin is cleaved to release the C-peptide. The A and B peptides are held together by disulfide bridges and form the active insulin component.

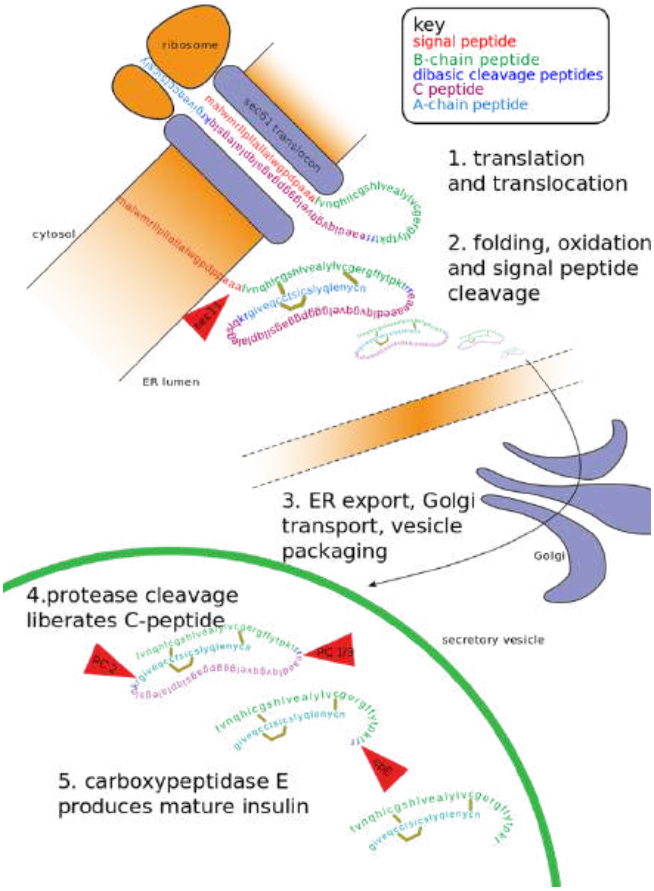


Figure 15.1.9: Insulin Production. The process of insulin production includes the translation of the insulin mRNA into the rough endoplasmic reticulum. The peptide is transported through the Golgi and packaged into secretory vesicles where protease cleavage by carboxypeptidase E produces the mature insulin peptide and peptide C. Image from Fred the Oyster

The C-peptide is a bioactive peptide secreted at the same time and in equimolar amounts to the insulin hormone. It also has a longer half-life than insulin and is excreted by the kidneys into the urine, making detection easy. Furthermore, it allows for the detection of patient-produced insulin, even if they are receiving insulin injections. Thus, C-peptide detection is often utilized to help distinguish between patients with type 1 diabetes from patients with type 2 diabetes (or Maturity onset diabetes). Details about the different forms of diabetes will be discussed in greater detail later.

Once insulin is released from the pancreas, it travels throughout the body and binds with cellular targets that contain the insulin receptor. The Insulin Receptor is a tyrosine kinase receptor that dimerizes upon insulin binding, as shown in Figure 15.1.10. Insulin receptors are located on most cell types throughout the body causing pleiotropic effects during insulin response. Primary targets of insulin action are the liver, where it promotes the uptake of glucose and the production of the glycogen storage molecule, as well as skeletal muscle and fat. The tyrosine kinase portion of the receptor located on the internal side of the plasma membrane is quite flexible.

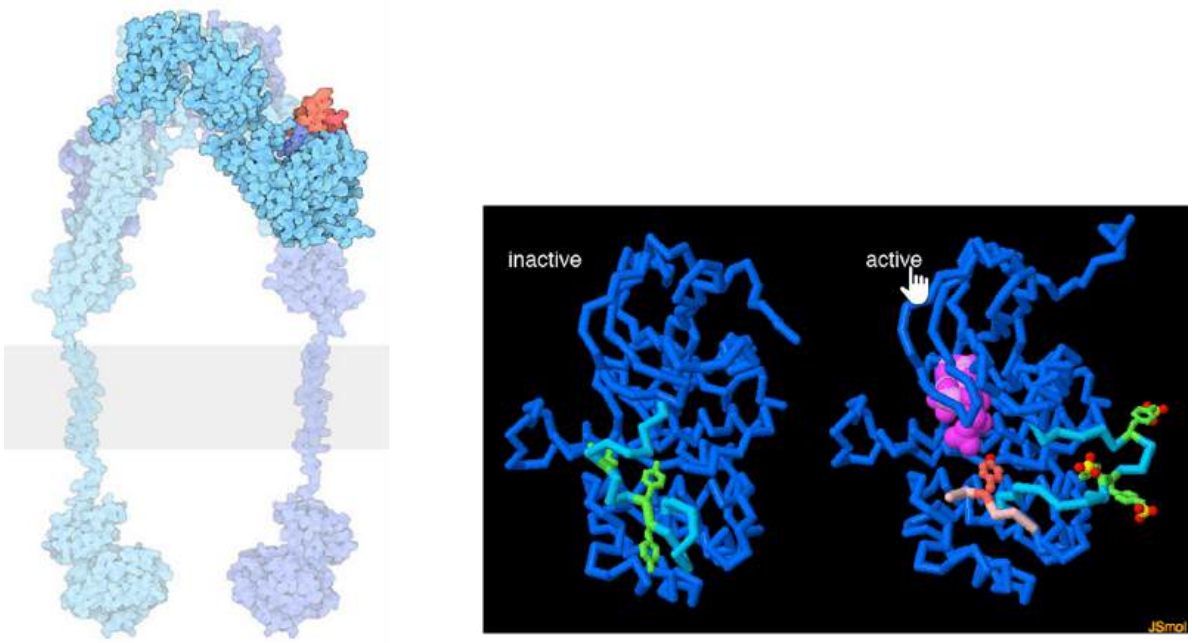


Figure 15.1.10: The Insulin Receptor. Images from Goodsell, D., et al (2015) RCSB PDB-101 'Molecule of the Month'

The *lefthand diagram* shows a space-filling model of the activated insulin receptor dimer embedded into the plasma membrane (shown as the gray bar). The tyrosine kinase portion of the receptor is shown on the inside of the cell whereas the insulin binding domain is present on the external side of the plasma membrane. The *middle and right-hand diagrams* show the inactive *(middle)* and active forms *(far right)* of the tyrosine kinase domain of an insulin receptor monomer. When activated the tyrosine kinase domain binds to ATP (hot pink) and phosphorylates downstream targets, including several of its tyrosine residues (green). In the inactive state (middle), a mobile loop (turquoise) binds in the ATP binding site and prevents ATP association. When the insulin receptor is activated, the mobile loop opens, allowing for the binding of ATP and self-phosphorylation of tyrosine residues, as well as other signaling proteins (a small peptide from one is shown in light pink).

Figure 15.1.11 below which shows an interactive iCn3D model of the **Full-length mouse insulin receptor bound to four insulins (7SL7)**.

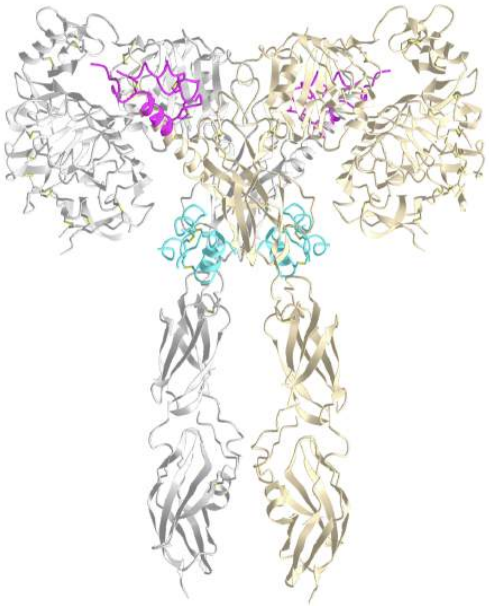


Figure 15.1.11 : Full-length mouse insulin receptor bound to four insulins (7SL7). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...gkTrtvaqQF6o2A

The receptor is a dimer (one monomer gray and the other light brown). Four insulins are bound in maximally activated insulin receptors. Two insulins (magenta) are bound at the same respective place in each monomer (site-1, the primary site) and the two others are bound at a second parallel site (site-2). The full active state is a symmetric T-shape. Less active receptors have fewer bound insulins with receptor geometry more asymmetric (one insulin bound at site 1 gives a Γ-shaped conformation, while two produce a T-shaped conformation as the second insulin binds). When 4 insulins are bound at both sites, the asymmetric conformation can't be formed. Although the structure is described as full-length, both monomers end at amino acid 910. The single membrane-spanning alpha-helix membrane occurs at amino acids 947-967 and is NOT shown in the model.

Activation of the insulin receptor in the liver when insulin is present initiates a phosphorylation signaling cascade, as shown in Figure 15.1.12. One function of the signaling cascade results in the activation of the Rab10 protein. Rab 10 promotes the fusion of GLUT4-containing secretory vesicles (GSVs) with the plasma membrane allowing for increased surface expression of GLUT4. GLUT4 is a glucose transporter protein. Thus, an increased concentration of the protein in the plasma membrane results in the upregulation of glucose import into the cell. Having GLUT4 proteins stored within secretory vesicles makes it available more readily than having to activate gene transcription pathways and production of the protein *de novo*. This allows a faster response to help lower blood glucose levels. The result is increased glucose uptake from the bloodstream into liver cells and other cellular targets, reducing blood glucose levels.

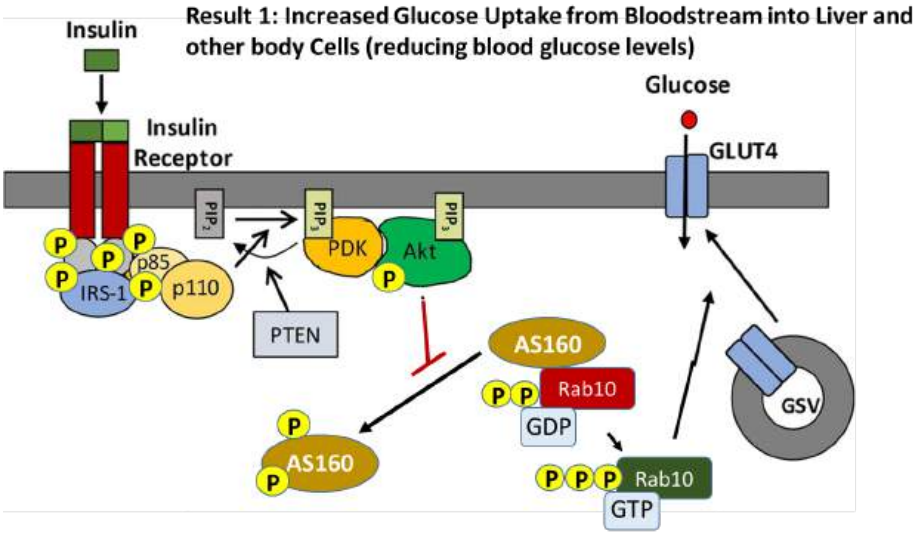


Figure 15.1.12 Insulin Activation of Liver Cells and Glucose Uptake. Image from Carmichael, R.E., et al (2019) Scientific Reports 9:6477

The insulin receptor is a receptor tyrosine kinase, which undergoes dimerization and autophosphorylation of Tyr residues upon insulin binding. The phosphorylated receptor also recruits and phosphorylates the insulin receptor substrate 1 (IRS-1) on tyrosine residues, which then recruits dimeric Phosphoinositol (PI)3-kinase (p85/p110 in the diagram above) and phosphorylates the p85 regulatory subunit. The PI3 kinase catalyzes the phosphorylation of phosphatidylinositol bisphosphate (PIP2) within the plasma membrane to form phosphoinositol, 3,4,5-triphosphate (PIP3). PIP3 then recruits PIP3-dependent kinase (PDK) which phosphorylates and activates Akt. Once activated, Akt dissociates from the membrane into the cytosol where one of its downstream targets is AS160. AS160 is a GTPase that normally binds with Rab10 (a G-protein) causing the cleavage of GTP to GDP. Thus, AS160 downregulates the activity of Rab10. In the phosphorylated state, AS160 cannot bind or inhibit Rab10, enabling Rab10 to release GDP and bind with a molecule of GTP. In the activated state, Rab10 helps promote the fusion of GLUT4-containing secretory vesicles (GSVs) secretory vesicles with the plasma membrane.

Figure 15.1.13 provides a deeper look at some of the initial activation steps in the insulin signaling pathway. This step shows the phosphorylation of Phosphoinositol 4,5-bisphosphate (PIP2) to Phosphatidylinositol 3,4,5-triphosphate (PIP3). PIP2 is a common phospholipid within the lipid bilayer structure. In future lectures, we will see the utilization of this phospholipid in other signaling pathways as well.

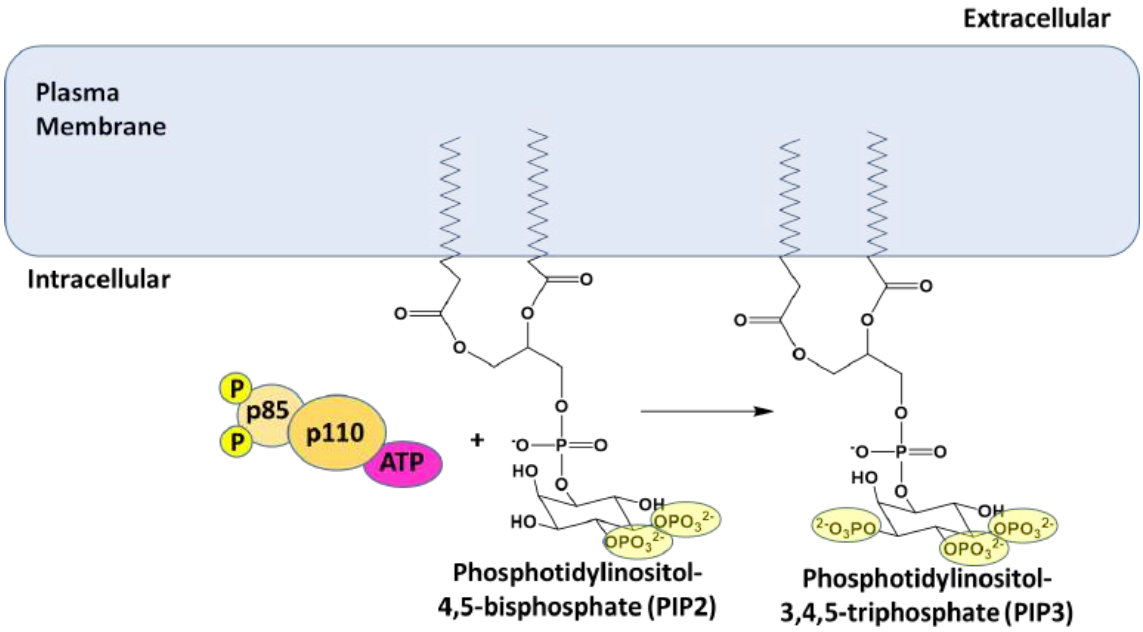


Figure 15.1.13 Phosphorylation of phosphotidylinositol-4,5-bisphosphate During Insulin Signaling.

Once glucose enters a cell, it is rapidly converted to glucose 6-phosphate via the enzyme hexokinase, as shown in Figure 15.1.14 below. This enzyme is covered in more detail in our section on glycolysis. Importantly, phosphorylation traps the glucose inside the cell and does not allow it to be redistributed back into the bloodstream. This helps to maintain the homeostasis of glucose within the bloodstream. In addition, glucose 6-phosphate is the first step in many pathways utilizing glucose, including energy utilization and the formation of building blocks such as ribose and deoxyribose used in RNA and DNA synthesis.

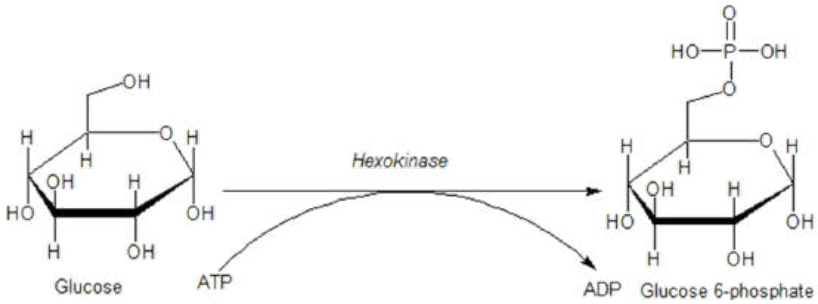


Figure 15.1.14 Conversion of Glucose to Glucose-6-Phosphate by Hexokinase. Image from Jmun7616

Insulin signaling increases the number of GLUT4 transporters in the plasma membrane causing an increased uptake of glucose into the cell. Within liver and muscle tissue, if glucose is not required for energy or other metabolic intermediates, it is then converted to glycogen for storage. The major enzyme required for glycogen synthesis is also activated via insulin signaling, as shown in Figure 15.1.15. In addition to phosphorylating the AS160 protein, Activated Akt also phosphorylates the Glycogen Synthase Kinase enzyme (GSK-3) which inactivates this protein. This allows protein phosphorylase 1 (PP1) to dephosphorylate the Glycogen Synthase enzyme shifting it into a more active state, and causing glycogen synthesis to commence.

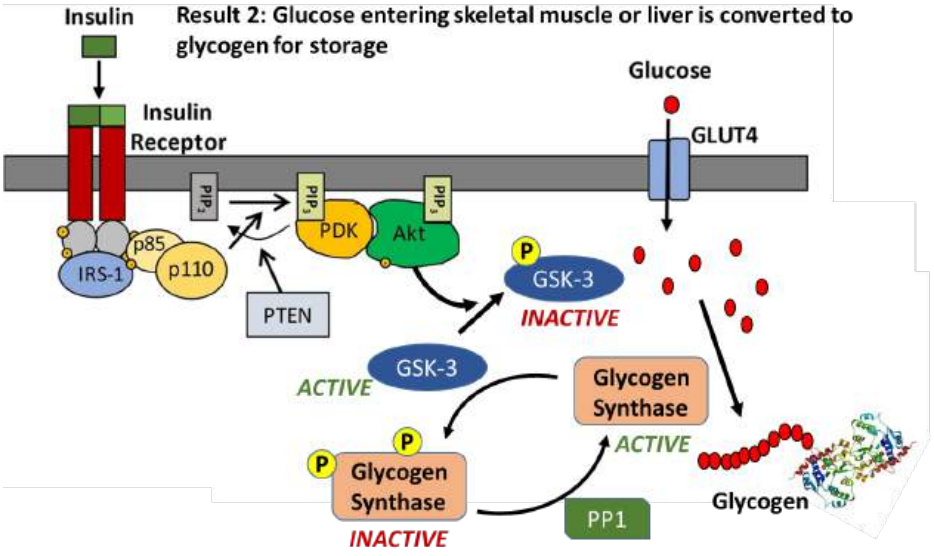


Figure 15.1.15 Activation of Glycogen Synthase During Insulin Signaling. Image modified from Carmichael, R.E., et al (2019) Scientific Reports 9:6477

In this section, we have covered two pathways activated during cellular response to insulin signaling, and I am sure that you are feeling a bit overwhelmed with the complexity. However, biological processes are incredibly complex and signaling pathways have multiple pleiotropic downstream effects. For insulin signaling, we have only touched the tip of the iceberg, as evidenced in Figure 15.1.16. This figure gives a more complete representation of the chemical changes induced within a liver cell in response to insulin signaling. For our purposes, we will restrict coverage to the two downstream effects: an increase in GLUT4 transporters in the plasma membrane and increased activity of glycogen synthase.

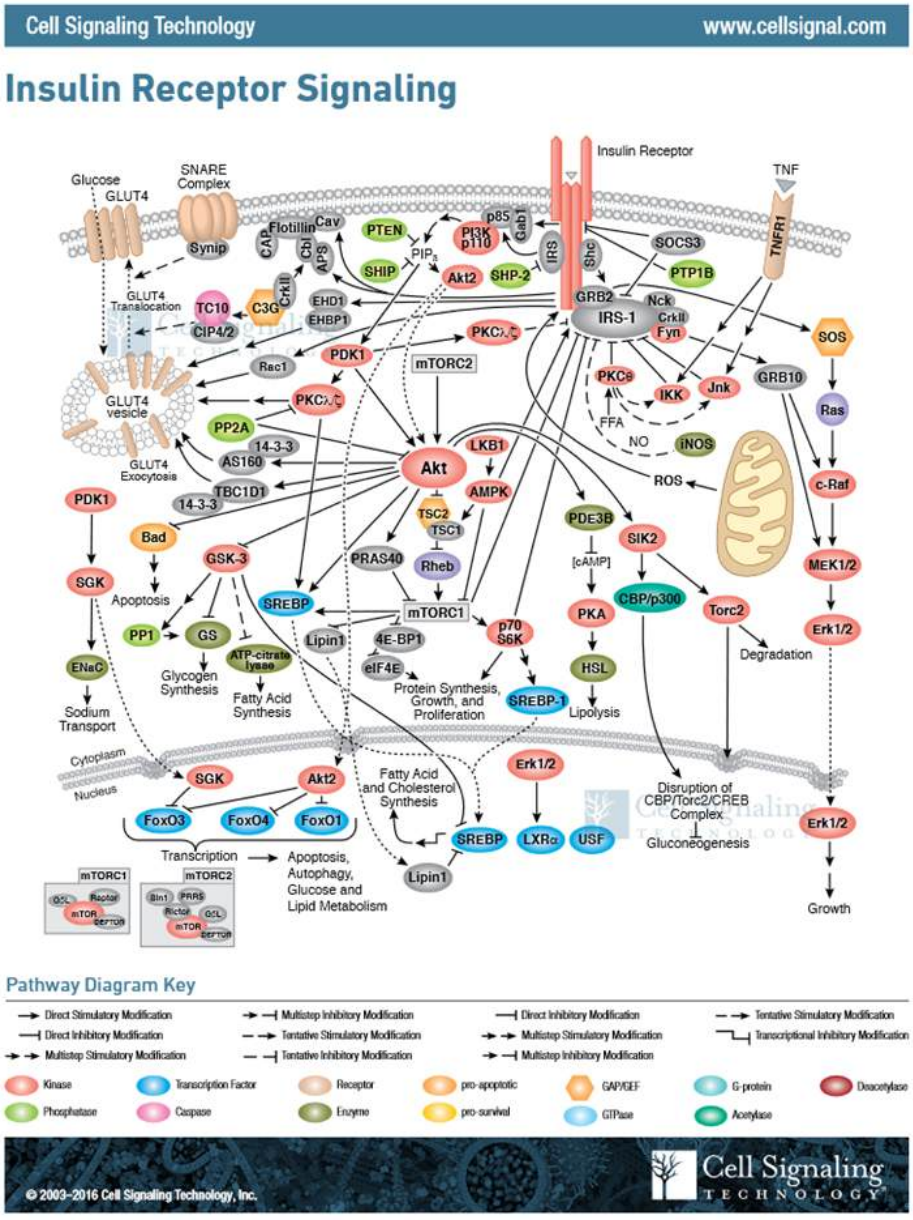


Figure 15.1.16: Insulin Signaling Pathway. Image by Cell Signaling Technology

15.2: GLYCOGENESIS

15.2.1: INTRODUCTION

The process of forming glycogen is called **glycogenesis** and it requires the activity of six enzymes as illustrated in Figure 15.2.1. We have already discussed several including hexokinase which phosphorylates the 6'-OH of glucose and phosphoglucomutase which converts glucose-6-phosphate to the glucose-1-phosphate isomer. In this section, we will discuss the remaining four enzymes and their role in glycogen biosynthesis. They are Glycogen Synthase, UDP-Glucose Pyrophosphorylase (preferred name UTP-glucose-1-phosphate uridylyltransferase), Glycogenin, and Glycogen Branching Enzyme.

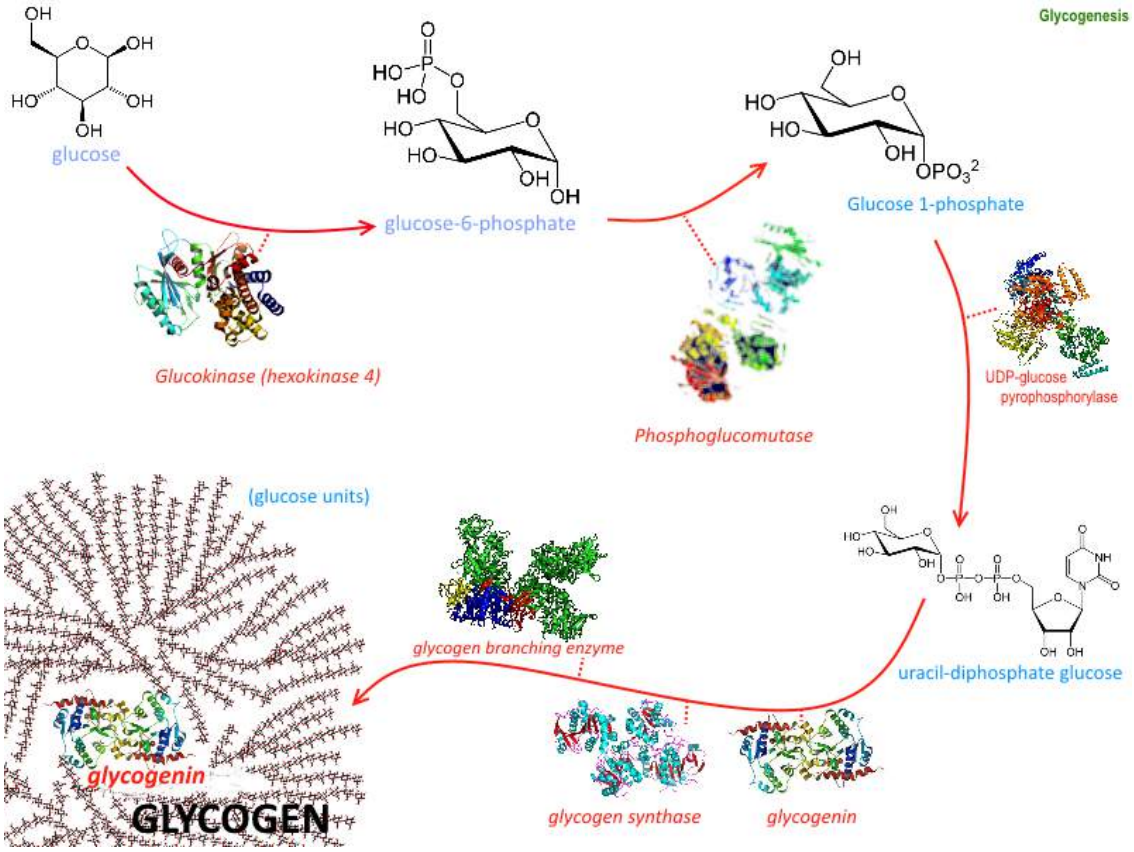


Figure 15.2.1: Enzymes involved in Glycogenesis. Image from Mark Cidade

Given the importance of these enzymes in the synthesis of the second main energy storage molecule, we must probe the enzymes in detail.

Glycogen synthase (GS) is a key enzyme and its activity is highly regulated. In Chapter 15.1, we have already explored how insulin signaling upregulates the activity of this enzyme by inhibiting phosphorylation by GSK-3. Other effectors include the allosteric binding of glucose-6-phosphate, which also increases the activity of the GS. In a later section, we will also see that the hormone glucagon can also regulate the activity of the GS through protein kinase A (PKA) in a fashion that decreases glycogen synthesis and increases glycogen breakdown.

In the glycogenesis pathway, GS is responsible for building the majority of the main alpha 1 → 4 chain glucose acetal linkages. The GS does require a primer of 4 -6 glucose residues linked together by alpha 1 → 4 bonds to begin synthesis. Since GS can only form alpha 1 → 4 linkages in the main chain, it CANNOT create the alpha 1 --> 6 branches inherent to the core structure of glycogen.

To build the glycogen main chain, GS uses the glycogen primer and glucose that has been activated through covalent attachment to uridine diphosphate (UDP) at the 1-position. Upon completion of one round of synthesis, the 1 position of the incoming UDP-glucose is covalently attached to the 4-position of the nascent glycogen molecule, releasing the UDP as a leaving group.

$$\text{Glycogen}_{(n)} + \text{UDP-glucose} \rightarrow \text{Glycogen}_{(n+1)} + \text{UDP}$$

15.2.2: **UTP--GLUCOSE-1-PHOSPHATE URIDYLYLTRANSFERASE (OR** UDP-GLUCOSE PYROPHOSPHORYLASE)

The formation of the UDP-glucose required for the synthesis of the main chain of glycogen is mediated by **UTP-glucose-1-phosphate uridylyltransferase (preferred name), which is also called UDP-glucose pyrophosphorylase** (GalU or UGPase; EC 2.7.7.9). UGPase catalyzes the reversible reaction of glucose 1-phosphate and UTP into UDP-glucose and inorganic pyrophosphate (PP_i) (Figure 15.2.2). Enzymes of the UGPase family are ubiquitous and can be found in the tree of life.

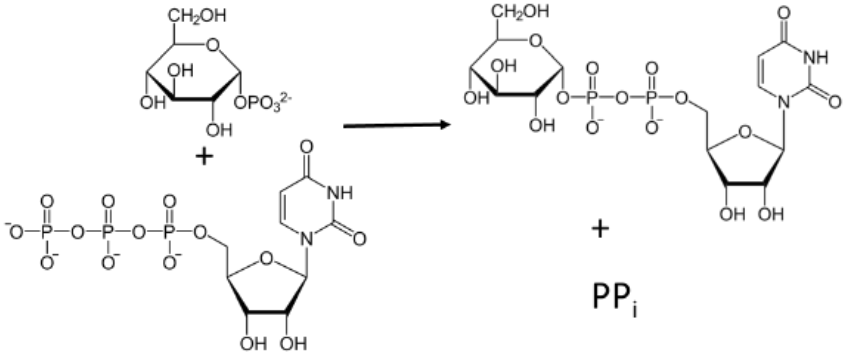


Figure 15.2.2: Formation of UDP-Glucose

UDP-glucose is an activated form of glucose used in the synthesis of other glycans including sucrose, cellulose, start, and glycogen, and the glycan parts of glycoproteins, glycolipids, and proteoglycans. Hence it is a key metabolite, which gives more importance to understanding UGPase.

Like many other nucleotidyl transferases, UGPase requires divalent cations to promote the reaction (Figure 15.2.3). In most cases, magnesium ions are employed. The reaction mechanism follows a sequential bi-bi-mechanism starting with the binding of UTP to the active site, in presence of a magnesium ion, followed by the binding of glucose 1-phosphate. The octahedral coordination sphere of the magnesium positions the substrates in the right way and enables the nucleophilic attack of glucose 1-phosphate on UTP. A lysine, an aspartate, and several water molecules within the active site help to stabilize the position of the substrates and cofactor for the proper nucleophilic attack of the phosphoryl oxygen of glucose 1-phosphate towards the α-phosphorous atom of UTP. Finally, PP_i is released from the UGPase/Mg^{2+}/UDP-glucose complex. UDP-Glucose then dissociates from the complex restoring the active site of the enzyme for another round of synthesis.

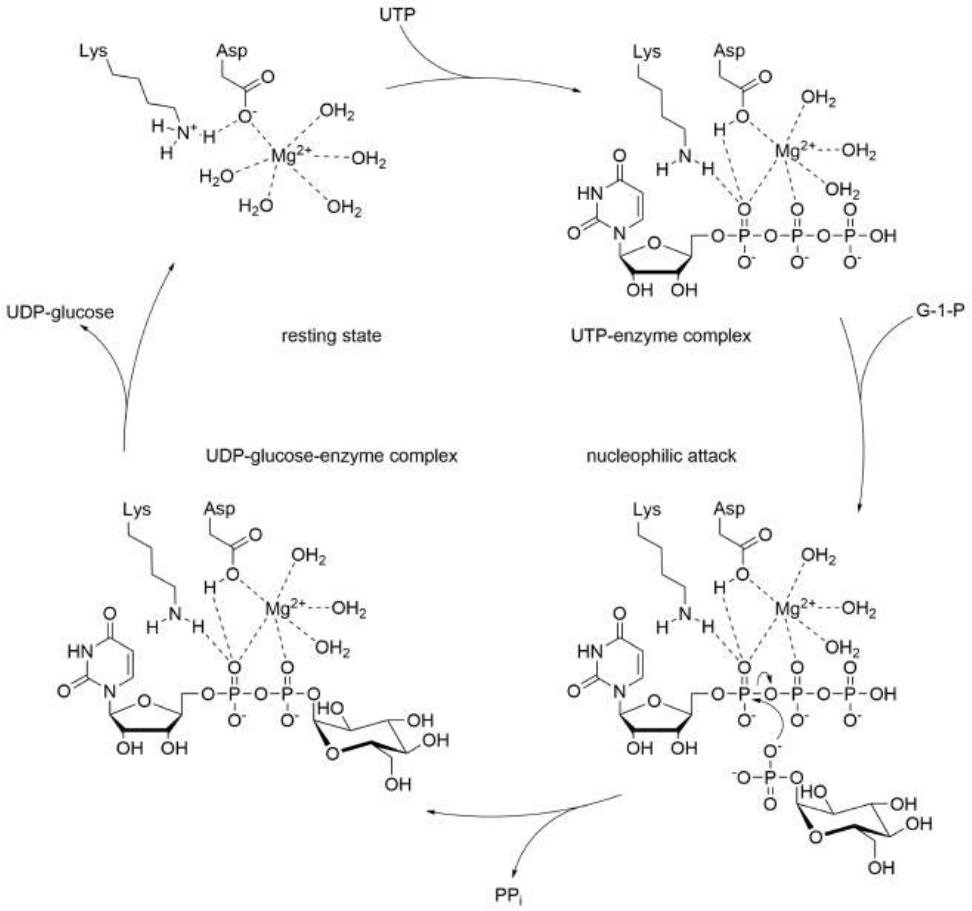
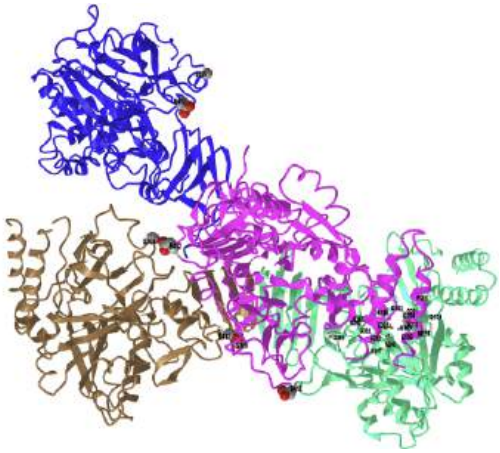


Figure 15.2.3: Proposed Reaction Mechanism of UTP--glucose-1-phosphate uridylyltransferase (UDP-Glucose Pyrophosphorylase). Image from Kumpf, A., et al. (2019) *Int. J. Mol. Sci. 20(22)* 5809; **https://doi.org/10.3390/ijms20225809**

Figure 15.2.4 shows an interactive iCn3D model of the human UDP-glucose pyrophosphorylase tetramer (3R2W).



NCBI iCn3D Figure 15.2.4: Human UDP-glucose pyrophosphorylase tetramer (3R2W). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?9yq9STPGvq87Wnbt8

Substrate binding and active site residues are shown as CPK-colored sticks and labeled in one of the subunits. Three key loops whose correct positioning is required for catalysis are colored as follows:

- Latch Loop: 406-416, yellow, which contains Glu 412 (shown in spacefill)
- SB Loop: 275-282, red
- 309 Loop: 309-311, cyan, which contains Ser 309 (shown in spacefill).

Only four of the eight subunits are shown for clarity. Note the proximity of the interacting loops positioned between the brown and magenta subunits. The SB and 309 loops at that location are part of the brown subunit, while the yellow latch loop is part of the magenta subunit. The latch loop is between the SB and 309 loops. The cyan 309 loop hence is prevented from interacting with the substrate and with the movements of the SB and 309 loops needed for catalysis. In other species, it has been observed the SB loop moves down UDP-glucose on binding to the active site. The structure above is for the apo-enzyme without bound substrate, and hence the inactive or **closed** form of the protein.

Mutations in the 309 loop (S309N/S311R) still had 84% of normal activity. Glu412 in the latch loop is highly conserved in vertebrates (but not present in yeast). Mutations in Glu 412 didn't affect the formation of oligomers of the enzyme but did affect activity.

- replacing E412 with a short aspartate (E412D) significantly increase activity (176%)
- E412Q, which eliminates the charge while retaining the approximate size of the side chain showed just a marginal increase in activity (19%)
- E412K, which flips the charge and increases the length of the side chain decreased activity to 22%.

These mutations generally suggest that steric effects in the region of subunit interaction are most important in activity.

15.2.3: GLYCOGEN SYNTHASE

UDP-glucose is then utilized by **glycogen synthase** (GS) to extend the main chain of glycogen by one glucose residue. In this reaction, the 4'-OH group of the glycogen main chain attacks the anomeric carbon of UDP-glucose (Figure 15.2.5). The UDP functional group serves as a good leaving group allowing for the formation of the alpha 1 --> 4 bond.

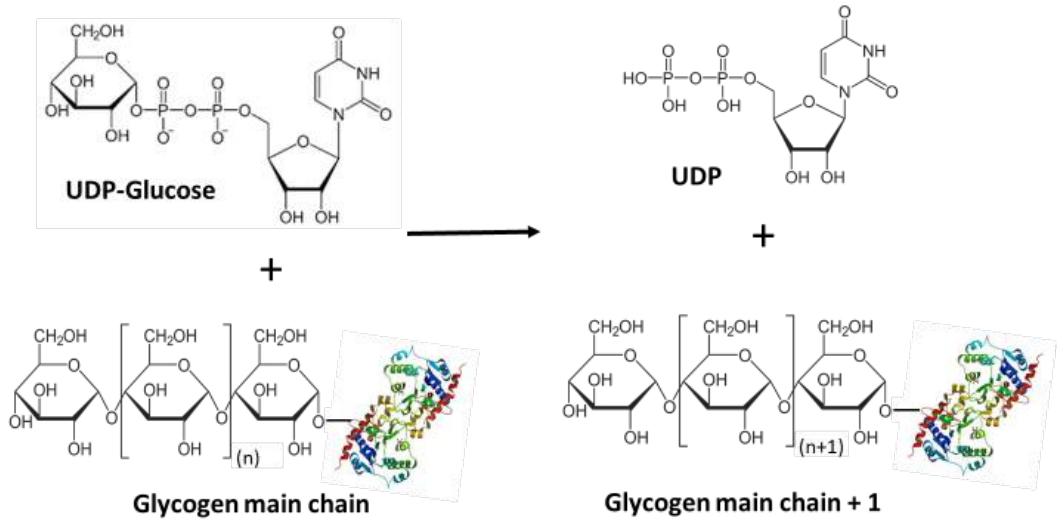


Figure 15.2.4: Formation of the Glycogen Mainchain by Glycogen Synthase. Image modified from Mikael Häggström

Glycogen synthase has two isoforms, GS1 expressed in tissue where glycogen is present (for example muscle) and GS2 expressed only in the liver. We will explore its mechanism more fully below.

15.2.4: GLYCOGENIN

Previously, we mentioned that GS requires a glycogen primer of 4 – 6 glucose residues to begin adding new residues to the main chain. This primer is provided by the small docking protein, **Glycogenin (GN or GYG)**. This protein is a homodimer that self-catalyzes its own glycosylation at amino acid Tyr-194. In this reaction, UDP-glucose is coordinated by a Mn^{2+} metal cofactor and critical aspartate residues (Figure 15.2.6). The –OH group of Tyr-194 then mediates nucleophilic attack on the anomeric carbon of UDP-glucose. Thus, glycogenin is tethered to the reducing end of the glycogen molecule.

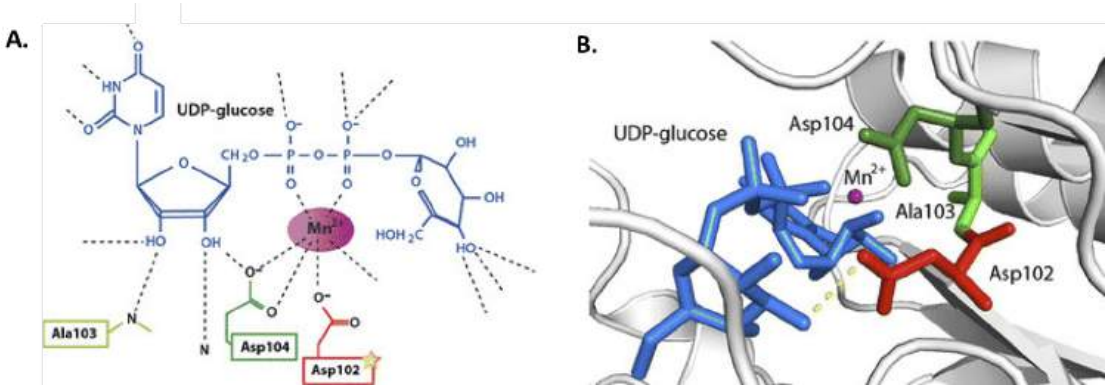


Figure 15.2.6: Coordination of UDP-Glucose by Glycogenin. Critical aspartic acid residues and a manganese ion cofactor are required for the coordination of UDP-glucose by glycogenin. Image from Hedberg Oldfors, C., Glamuzina, E., Ruygrok, P. and Anderson, L. (2016) *J. Inherit. Met. Dis.* 40(1) DOI: 10.1007/s10545-016-9978-1

As with glycogen synthase, glycogenin (GN or GYG) has two isoforms, with GN2 (or GYG2) expressed mostly in the liver, pancreas, and heart.

15.2.5: MECHANISMS FOR GLYCOGEN SYNTHESIS BY GLYCOGENIN AND GLYCOGEN SYNTHASE

How do glycogenin and glycogen synthase cooperate in the synthesis of glycogen? Structures of the complex of glycogenin-1 (GYG1), which seeds the molecule by starting glycogen synthesis by autoglucosylation, and glycogen synthase-1 (GS1 or GYS1), which extends the molecule, show allosteric transitions between three main states, the closed/inactive, partially open, and open active complex, as shown in Figure 15.2.7:

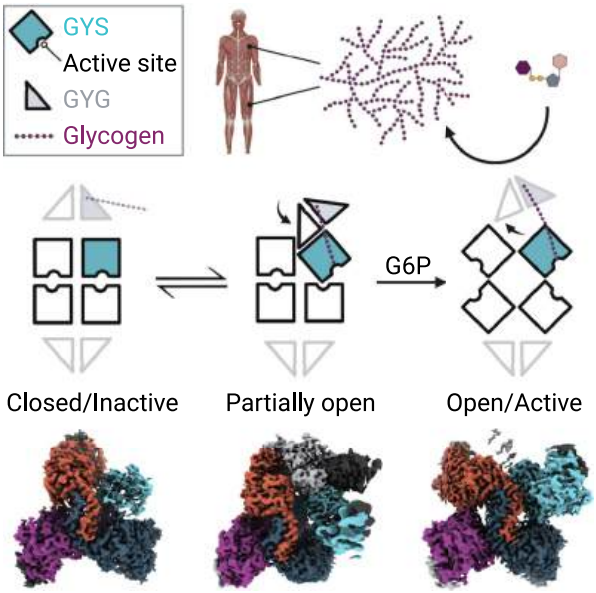


Figure 15.2.7: GYS1 chains A, B, C, and D are colored orange, turquoise, purple, and navy respectively. GYG1 globular domains and the GYG1-tail fragment are colored in gray shades for all chains. The left bottom structure shows the apo GYS1:GYG1 mobile complex. The bottom middle structure is the apo GYS1:GYG1 ordered complex. The bottom right structure is the +G6P GYG1:GYS1 complex. Fastman et al., 2022, Cell Reports 40, 111041 July 5, 2022, 2022. https://doi.org/10.1016/j.celrep.2022.111041. Creative Commons Attribution (CC BY 4.0)

It makes sense that both enzymes bind to each other and cooperate in the synthesis of glycogen. In the closed state, GS (GYS) is a tetramer, which like hemoglobin can be described as a T or closed/inactive state. In a slight difference, one of the GS (GYS) subunits appears to display an asymmetric conformation which leads to close interactions with GN (GYG) allowing the glycogen seed polymer on GN (GYG) to move to GS (GYS) for elongation (a partially open state). Multiple conformations of the complex have been resolved. Further conformational changes lead to the open/active state and a more open binding groove for GN (GYG).

Figure 15.2.8 shows an interactive iCn3D model of the Human glycogenin-1 and glycogen synthase-1 complex in the presence of glucose-6-phosphate (8CVX). Glucose-6-phosphate is an allosteric activator of glycogen synthase.

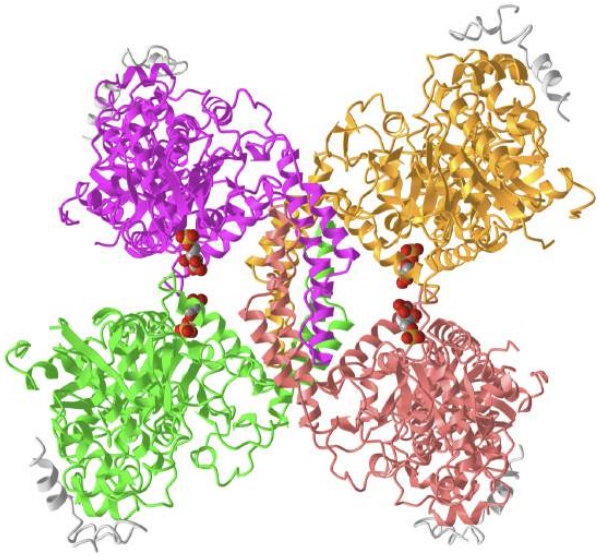


Figure 15.2.8: Human glycogenin-1 and glycogen synthase-1 complex in the presence of glucose-6-phosphate (8CVX). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...ivh4gZCupGcXZ7

The larger glycogen synthase subunits in the tetramer are shown in various colors while the glycogenin fragments bound to each monomer of GS (GYS) are shown in gray.

Figure 15.2.9 illustrate in greater detail the conformational changes induced on the binding glucose-6-phosphate to glycogen synthase.

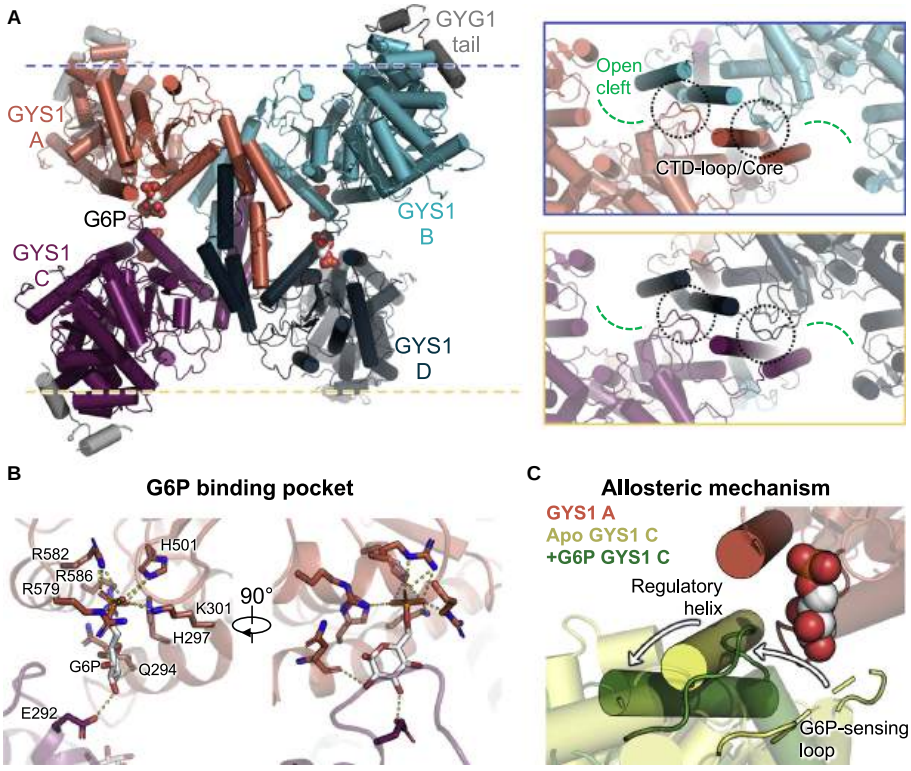


Figure 15.2.9: G6P binding induces a conformational change across the GYS1 tetramer leading to an open conformation for all four active sites. Fastman et al, ibid.

Panel (A) shows a cartoon representation of the +G6P GYS1:GYG1 complex (left) with GYS1 and GYG1 chains colored as in the previous figure. G6P is shown as spheres colored by heteroatom. Blue and yellow lines and boxes (right) indicate relative positions for perpendicular views. A black dotted circle indicates the oligomeric interface between the CTD-loop region and tetramerization core domains. The CTD-

loop, residues 484–488, forms cross-protomer interactions on the other side of the active site. Green dashed semicircles indicate open active sites.

Panel (B) shows G6P-binding site interactions. G6P (white) and proximal interacting residues are shown as sticks and colored by heteroatom. Polar interactions are indicated by yellow dotted lines.

Panel (C) shows changes at the G6P-binding site across conformations. A G6P-bound protomer is highlighted in orange with G6P shown as spheres. The adjacent active state protomer (chain C) is shown as a green cartoon. A basal state protomer (chain C) is modeled relative to the orange G6P-bound protomer and shown in yellow. Key changes to the G6P-sensing loop and a regulatory helix are highlighted by outline and non-transparent representation with arrows indicating the relevant motions. The rest of each chain is shown in a transparent cartoon representation.

We will return to the mechanism after exploring the last enzyme in the pathway.

15.2.6: GLYCOGEN BRANCHING ENZYME

The final enzyme, the **glycogen branching enzyme** (GBE), catalyzes the hydrolytic cleavage of an α(1→4) glycosidic linkage and subsequent inter- or intra-chain transfer of the non-reducing terminal fragment to the C6 hydroxyl position of an α-glucan (Figure 15.2.10). In this example, an inter-chain transfer is occurring. At the top of the scheme, above the arrow, you can see that the GBE enzyme transiently removes several glucose residues (usually around 7) from one linear glycogen chain and then attaches it as an alpha 1→ 6 branch to the other chain. In this process, an additional non-reducing end is created which can act as a primer site for Glycogen Phosphorylase (the main enzyme that breaks down glycogen). Thus, glucose residues can be released very quickly when needed.

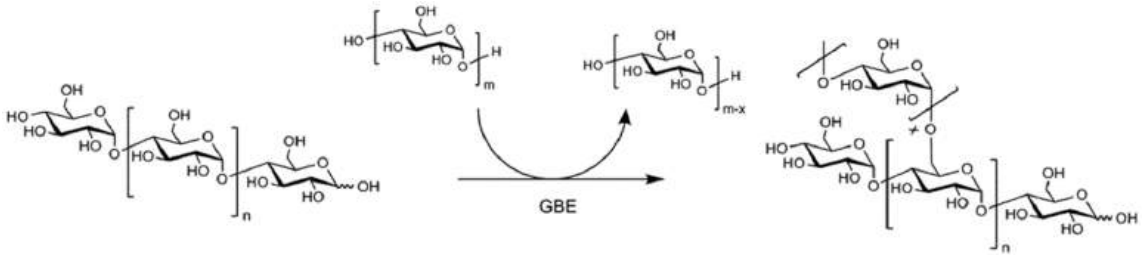


Figure 15.2.10: Formation of Glycogen alpha 1 --> 6 Branches. Image from van der Vlist, J., et al (2012) Polymers 4(1) 674-690. DOI: 10.3390/polym4010674

Details of the structure and domain organization of the human glycogen branching enzyme are shown in Figure 15.2.11. The four domains include N1, CBM48, the catalytic domain, and the C-terminal domain.

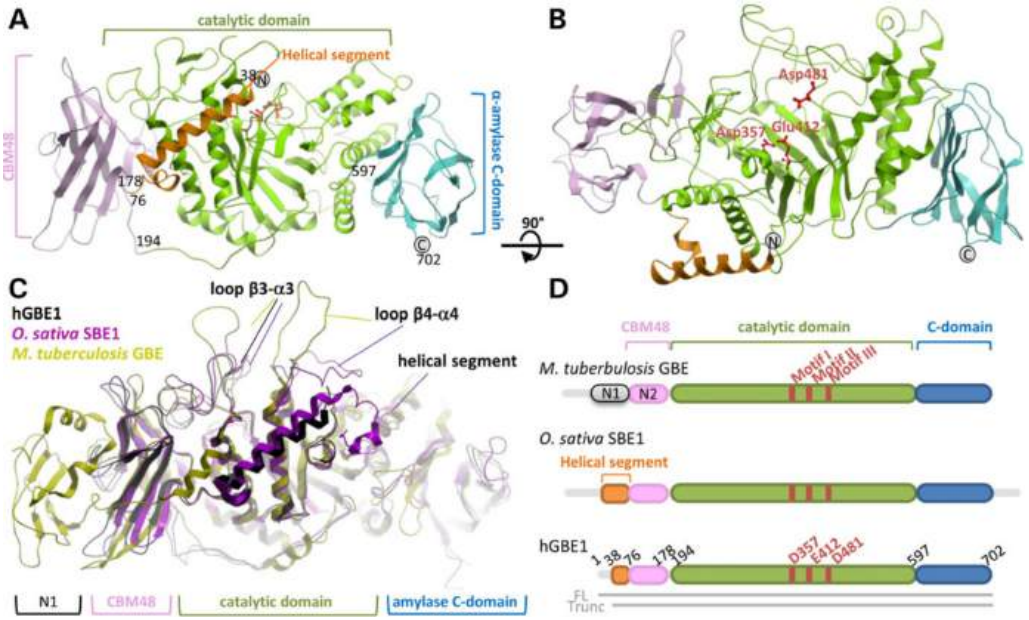


Figure 15.2.11: Crystal structure of hGBE1. Froese et al. *Human Molecular Genetics*, Volume 24, Issue 20, 15 October 2015, Pages 5667–5676, https://doi.org/10.1093/hmg/ddv280. Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/)

Panels (**A** and **B**) show orthogonal (perpendicular) views of hGBE1 showing the N-terminal helical segment (orange), CBM48 (pink), central catalytic domain (green), and C-terminal domain (blue). The catalytic triad Asp357-Glu412-Asp481 is shown as red sticks. Numbers refer to domain boundaries. N- and C-termini are labeled as grey spheres.

Panel (**C**) shows the superposition of branching enzyme structures from human (hGBE1, this study), *O. sativa* SBE1, and *M. tuberculosis* GBE, highlighting the conserved domain architecture and three regions of structural variation.

Panel (**D**) shows the domain organization of hGBE1, *O. sativa* SBE1, and *M. tuberculosis* GBE revealing differences in the N-terminus between prokaryotic and eukaryotic polypeptides. Prokaryotic GBEs contain two N-terminal carbohydrate-binding domains (N1, N2) whereas eukaryotes contain only one (CBM48) and replace the prokaryotic N1 domain with a helical extension.

Figure 15.2.12 shows an interactive iCn3D model of the **human glycogen branching enzyme (GBE1)**

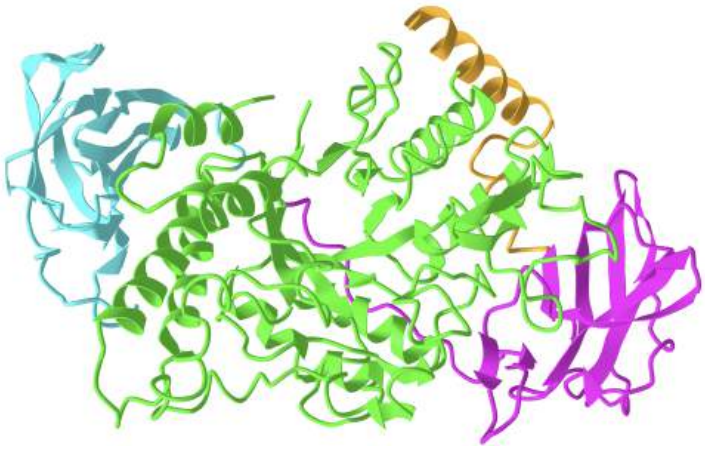


Figure 15.2.12: **human glycogen branching enzyme (GBE1)**. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...47aqw5qto8hkU7

The domains in the iCn3D module are colored similarly to that of the previous figure.

The enzyme core is similar to amylase with the conserved active site. Diseases of glycogen storage often result from mutations in the amylase core domain. For example, late-onset adult polyglucosan body disease (APBD) arises from a common mutation, Y329S. The effect of this mutation may arise from misfolding. A tetrapeptide, Leu-Thr-Lys-Glu, given to patients, increased activity twofold, probably by acting as a chaperone to facilitate proper folding.

15.2.7: PUTTING IT ALL TOGETHER: GLYCOGEN SYNTHESIS AND ITS REGULATION

Figure 15.2.13 shows the steps involved in the addition of glucose from the donor UDP-glucose to glycogenin and through glycogen synthase to the growing glycogen polymer which becomes branched through the activity of the glycogen branching enzyme.

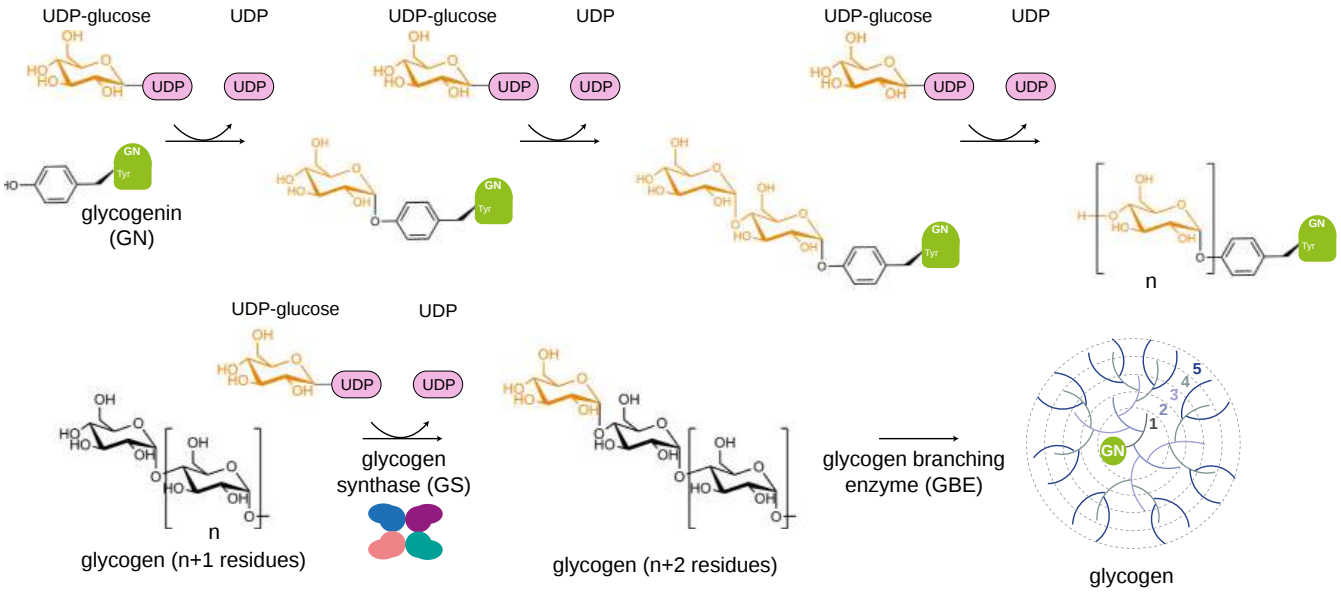


Figure 15.2.13: Summary of glycogen synthesis. Marr, L., Biswas, D., Daly, L.A. *et al. Nat Commun* **13**, 3372 (2022). https://doi.org/10.1038/s41467-022-31109-6. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/

The top reaction shows the step catalyzed by glycogenin (GN), while the bottom reactions are those catalyzed by glycogen synthase and the glycogen branching enzyme.

Regulation of glycogen synthesis occurs in part through the phosphorylation of key residues in both GS (GYS) and GN (GYG). Those sites are shown in Figure 15.2.14.

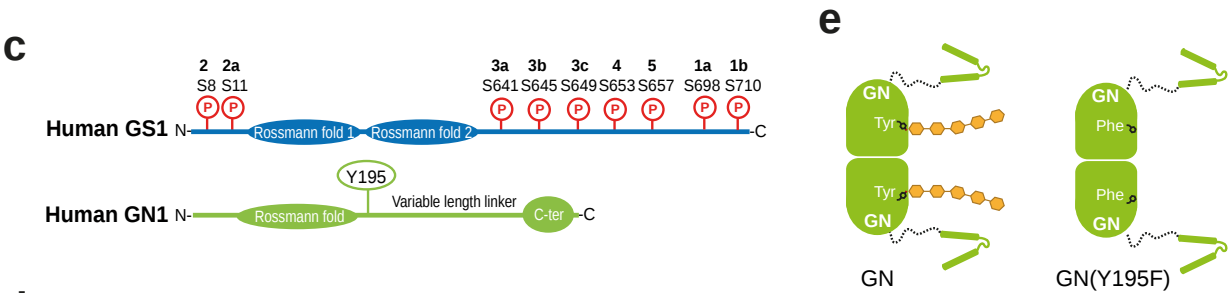


Figure 15.2.14: Domain and phosphorylation sites for glycogen synthase 1 (GS1) and glycogenin 1 (GN1). Marr et al, ibid

Panel c shows the domain architecture of human GS (top) and GN (bottom). Known in vivo phosphorylation sites of GS are shown in red and are labeled with residue number and classical nomenclature (in bold). GN tyrosine 195 that becomes auto-glucosylated and was mutated to phenylalanine (Y195F) in this study is indicated. Not to scale. **Panel e** shows the cartoon representation of GN WT and Y195F.

As we showed above, glycogen synthase is allosterically activated on the binding of glucose-6-phosphate, which can be thought to activate the enzyme by a T to R state change. The enzyme is inactivated by phosphorylation at multiple sites as shown above. Activation hence can also occur through dephosphorylation. Phosphorylation of a site can provide a binding site that leads to more phosphorylations. This can lead to a flexible "spike" of hyperphosphorylated residues forming from two of the monomers. Particularly important is pSer641 (site 3A) which interacts with a series of arginine residues in a regulatory helix of glycogen synthase. This arginine cluster has been called the arginine cradle. The interaction sites are illustrated in Figure 15.2.15.

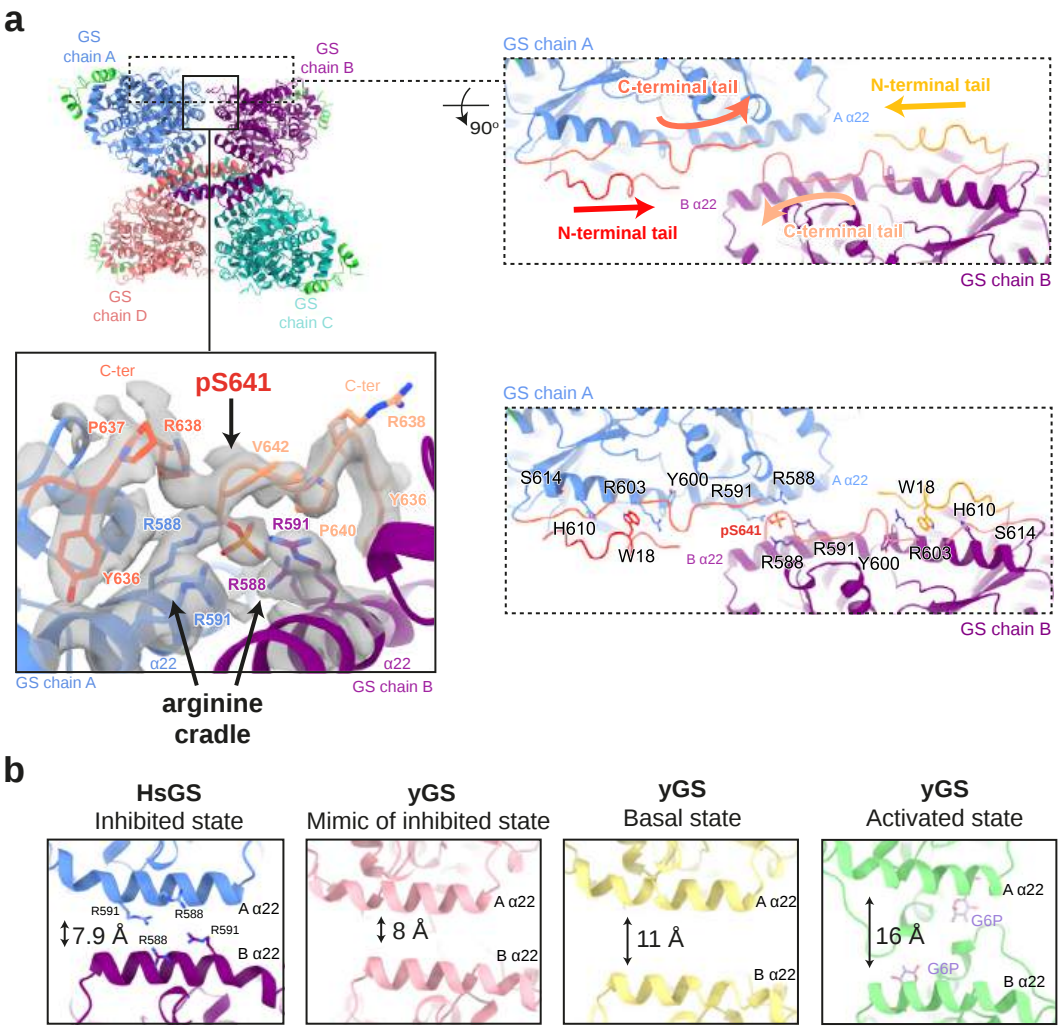


Figure 15.2.15: **The phosphoregulatory region of human GS.** Marr et al, ibid

Panel a shows the human (Hs)GS-GN[34] structure in ribbons (top left). The N- and C- terminal tails of one GS protomer (chain A) lie next to one another and move towards the adjacent protomer, meeting the N- and C-terminal tails from chain B. Arrows indicate a continuation of cryo-EM density (top right). Electron density (C1 symmetry) for phosphorylated S641 (pS641) interacting with R588 and R591 on the regulatory helices α22 (bottom left). Residues that are interacting with the N- and C-terminal tails that are mutated in this study are shown (bottom right).

Panel b shows a comparison of distances between regulatory helices of adjacent monomers of HsGS (reported here), low activity inhibited mimic (PDB ID 5SUL), basal state (PDB ID 3NAZ), and G6P activated (PDB ID 5SUK) yeast GS (yGS) crystal structures. Quoted distances were measured from Cα of Arg591 (chain A) and -Cα of Arg580 (chain B) of HsGS and corresponding yeast residues.

The strong electrostatic arginine-pSer interactions lock the tetramer into the inactive T state.

A cartoon model illustrating the regulation of glycogen synthase by phosphorylation/dephosphorylation and interconversion between T and R state is shown in Figure 15.2.16.

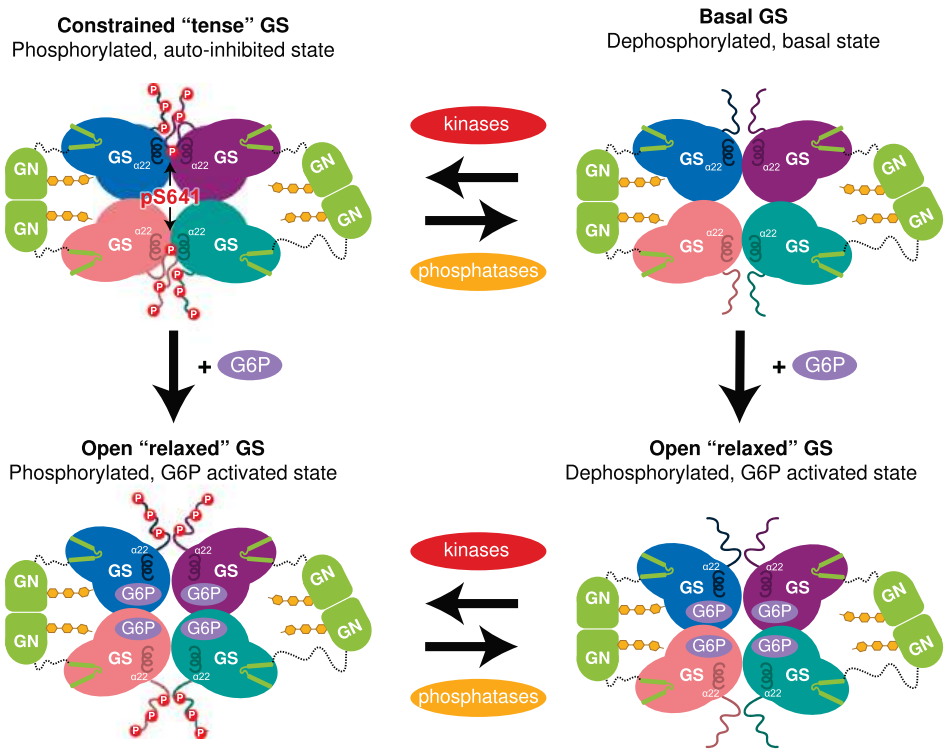


Figure 15.2.16: **GS and GN cooperate to synthesize glycogen.** Marr et al, ibid

The inhibition by phosphorylation can be relieved by binding the allosteric effector glucose-6-phosphate and does not require phosphatases,

15.2.8: REFERENCES

1. Kumpf, A., Partzsch, A., Pollender, A., Bento, I., and Tischler, D. (2019) Two Homologous Enzymes of the GalU Family in *Rhodococcus opacus* 1CP-*Ro*GalU1 and *Ro*GalU2. *Int. J. Mol. Sci.* 20(22), 5809. https://doi.org/10.3390/ijms20225809

15.3: GLYCOGENOLYIS AND ITS REGULATION BY GLUCAGON AND EPINEPHRINE SIGNALING

In the previous section, you learned that glucagon signaling down-regulates glycogen synthesis. Now let's look at glycogen breakdown, called glycogenolysis, and its control by two hormones, glucagon, and epinephrine. Only two enzymes are required for the breakdown of glycogen, the glycogen phosphorylase enzyme, and the glycogen debranching enzyme.

15.3.1: GLYCOGENOLYSIS: AN OVERVIEW

Two key enzymes are required for the stepwise catabolism of glycogen, glycogen phosphorylase, and glycogen debranching enzyme. In the liver, the ultimate end product is glucose-1-phosphate, which is dephosphorylated in the liver to enable the export of free glucose into the circulation. In contrast, in the respiring skeletal muscle, it is converted to glucose-6-phosphate for use in glycolysis. Glycogenolysis is also activated by the hormones glucagon and epinephrine.

Glycogen Phosphorylase

Glycogen phosphorylase (GP) catalyzes the release of glucose 1-phosphate from the alpha 1→ 4 non-reducing ends of glycogen. An overview of this reaction is shown in Figure 15.3.1.

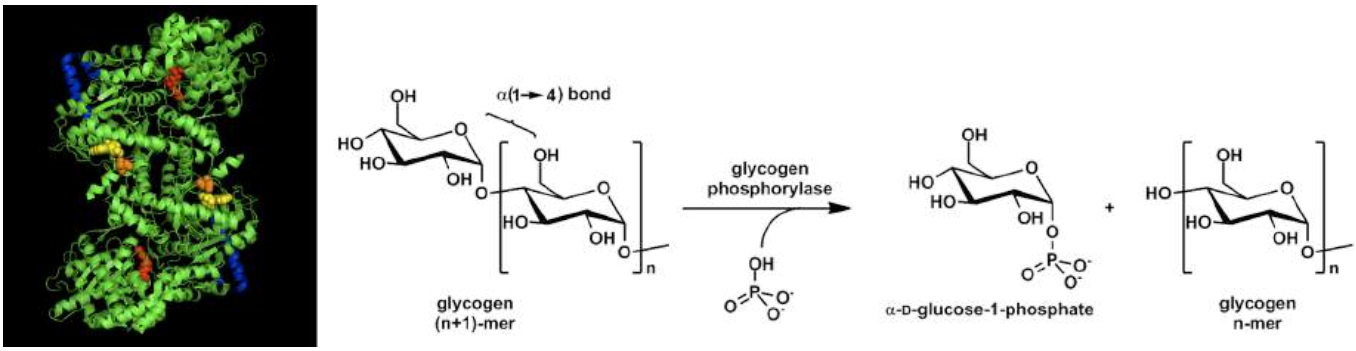


Figure \(\PageIndex{1}\): Overview of Glycogen Phosphorylase Reaction. Images from Ascherer730 and Michal Sobkowski

Glycogen phosphorylase is a homodimer with two active sites. It also requires a cofactor, pyridoxal phosphate (PLP) to be functional (Figure 15.3.2). The PLP is derived from Vitamin B6. You may have heard previously that low levels of B vitamins are associated with lethargy or a lack of energy. We will continue to see that the B vitamins provide essential cofactors for enzymes involved in the production of ATP. Thus, if you lack B vitamins, you are not efficiently producing ATP. The PLP cofactor of GP is attached covalently to the enzyme through a Schiff-base linkage with a Lysine (K) residue.

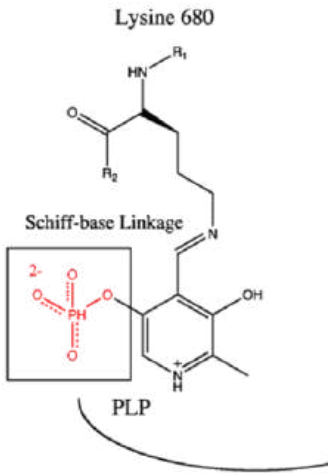


Figure 15.3.2: Pyridoxyl Phosphase Cofactor Associated with Glycogen Phosphorylase Image modified from Ascherer730

The reaction mechanism of glycogen phosphorylase is detailed in Figure 15.3.3. When glycogen phosphorylase binds with glycogen a free inorganic phosphate anion is positioned by the PLP and the enzyme active site in proximity with the anomeric carbon position of the non-reducing end residue of the glycogen molecule. The oxygen involved in the glycosidic bond attacks the partially charged hydrogen

associated with the phosphate ion, leading to the cleavage of the glycosidic bond. The cleaved glycogen chain leaves the active site and one of the phosphate oxygens attacks the carbocation intermediate created during the cleavage. This results in the release of the terminal glucose residue as glucose 1-phosphate

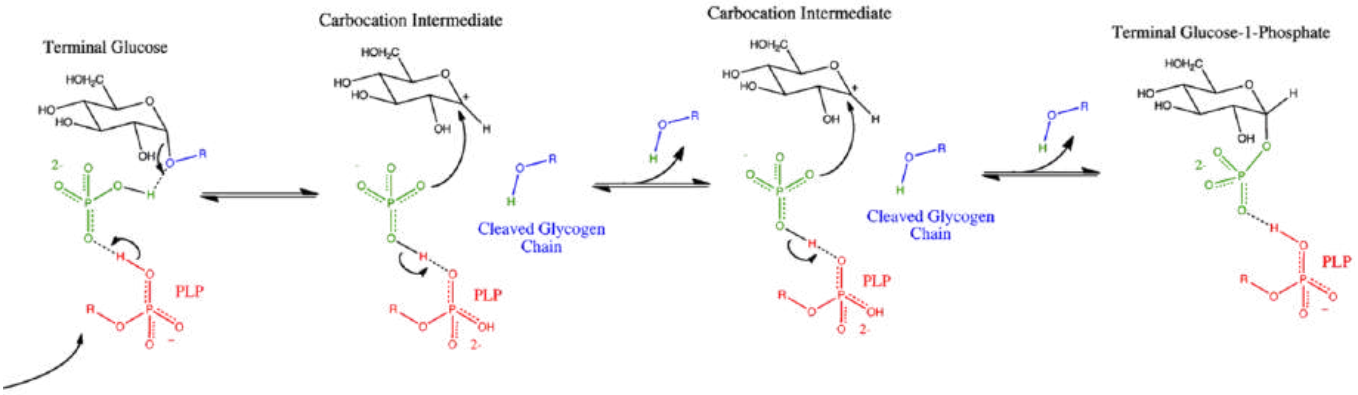
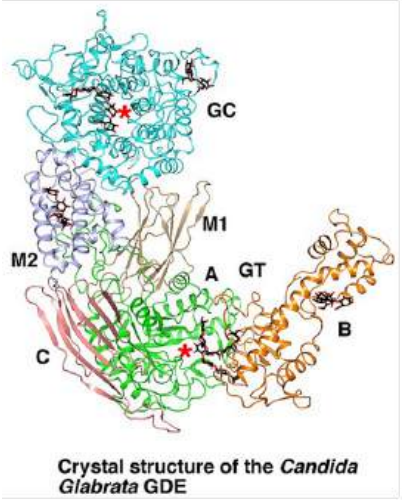


Figure \(\PageIndex{3}\): Glycogen Phosphorylase Reaction Mechanism. Image from Ascherer730

Glycogen Debranching Enzyme

Glycogen phosphorylase cannot cleave the alpha 1 → 6 linkages, and it also cannot cleave alpha 1 --> 4 linkages that are within 4 residues of an alpha 1→ 6 linkage (the glycogen chain will no longer fit into the active site of the enzyme) The Glycogen Debranching Enzyme (GDE) has two catalytic activities that enable it to deal with this problem. The first catalytic activity is a Glycosyl Transferase (GT) activity. In this process, the three remaining alpha 1 → 4 extended units on the branch site (colored in green) are clipped off of the branch site and attached to a straight chain of alpha 1 → 4 extended glucose residues. The second part of the reaction requires the Glucosidase (GC) activity that mediates the hydrolysis of the alpha 1 → 6 glycosidic bond and the release of free glucose in the process. Glycogen Phosphorylase can then resume the breakdown of the remaining alpha 1 → 4 chain. An overview of glycogen breakdown is shown in Figure 15.3.4.



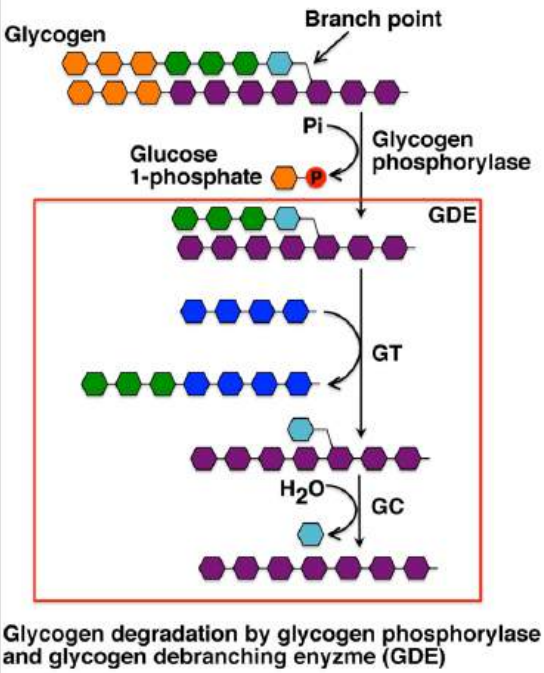


Figure \(\PageIndex{4}\): Biological Activity of the Glycogen Debranching Enzyme. Image modified from XiangSong

Dephosphorylation of Glucose 1-Phosphate

Following the activation of glycogenolysis, the liver cell has now released large quantities of glucose 1-phosphate from glycogen, as well as a smaller amount of free glucose from the clipped branch residues. The free glucose can be transported to the bloodstream straight away, but the glucose 1-phosphate must be dephosphorylated prior to release (Figure 15.3.5).

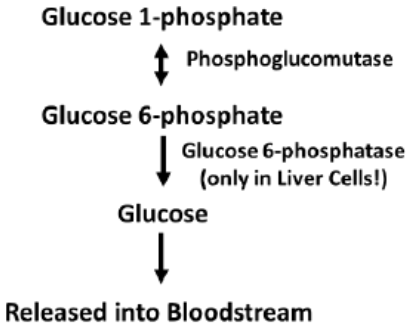


Figure 15.3.5: Process of Glucose Dephosphorylation in Liver Cells

The dephosphorylation of glucose only occurs in liver cells, as this is the primary location for the regulation of blood glucose levels. Free glucose can exit the cell while phosphorylated forms are trapped inside the cell. Figure 15.3.6 outlines the process of glucose dephosphorylation in the liver. To mediate the dephosphorylation of glucose, glucose 6-phosphate is transported from the cytoplasm into the lumen of the endoplasmic reticulum (ER) through transporter 1 (T1). The glucose 6-phosphatase (G-6-Pase) then cleaves the phosphate from the substrate, releasing inorganic phosphate (P) and glucose (red molecule). The inorganic phosphate is then transported back into the cytoplasm through transporter 2 (T2) and glucose is transported through Transporter 3 (T3). Free glucose is then transported back into the bloodstream through a glucose (GLUT) transporter located in the plasma membrane.

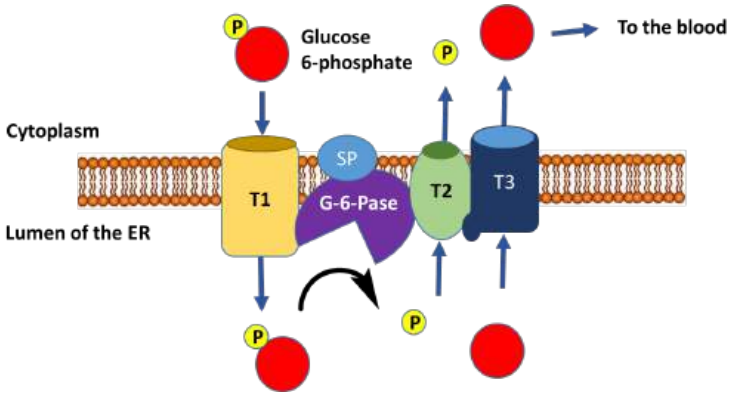


Figure 15.3.6: Dephosphorylation of Glucose 6-phosphate in Liver Cells.

15.3.2: HORMONAL CONTROL OF GLYCOGEN BREAKDOWN

In the previous sections, we've discussed insulin signaling and the process of building glycogen (glycogenesis) in detail. Now let's take a look at the other side of the homeostatic balance which begins with glucagon signaling. During hypoglycemia (or low blood glucose levels), pancreatic alpha (α) cells release the hormone peptide, glucagon, which stimulates gluconeogenesis (the formation of glucose) and glycogenolysis (the breakdown of glycogen) in the liver, resulting in the release of glucose to the plasma, and the raising of blood glucose levels, as shown in Figure 15.3.7.

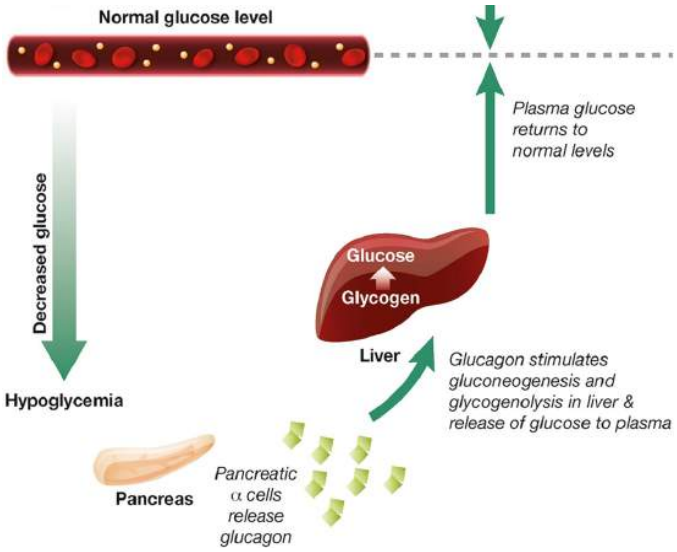


Figure 15.3.7: Summary of Glucagon Signaling During Low Blood Glucose Levels. Image from Hædersdal, S., et al (2018) Mayo Clinic Proceedings 93(2):217-239

Let's review a few terms before we begin. In the last section, we were introduced to glycogenesis, the synthesis of glycogen. We saw that this pathway was activated during insulin signaling. In glucagon signaling, this pathway is inhibited and the opposite pathway, glycogenolysis (glycogen breakdown) is activated. Glucagon signaling in the liver also down-regulates glycolysis (the utilization of glucose for energy production), as the liver is trying to use glucose to maintain blood glucose levels. It doesn't utilize it for its own energy needs during this time. Instead, lipids can be used by liver cells to generate ATP, and in fact, glucagon signaling increases lipolysis or the breakdown of lipids. Finally, glucagon also up-regulates the process of gluconeogenesis or the generation of glucose from non-sugar metabolites. We will address the mechanisms of glycolysis and gluconeogenesis regulation in a later section. Here we will only take a cursory look at these pathways and will focus more on the process of glycogenolysis.

Overview of Glucagon Signaling

Glucagon signaling begins when the hormone binds with its receptor on liver cells as shown in Figure 15.3.8. Glucagon receptors are not widespread within the body as insulin receptors. Since the purpose of this hormone is to cause the release of glucose back into the bloodstream, this process is highly controlled and only the liver can deliver glucose back into the bloodstream to maintain homeostasis. Thus, other target tissues such as skeletal muscle do not need to have these receptors expressed and are not sensitive to glucagon signaling.

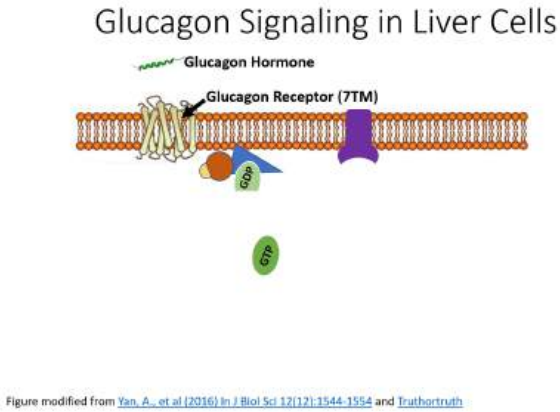


Figure 15.3.8: Overview of Glucagon Signaling Cascade

The glucagon receptor is a G-protein-coupled receptor and is also referred to as a 7TM receptor (as it contains 7 transmembrane domains that span the plasma membrane). This family of receptors is widespread throughout the body and responsible for many of the pharmaceutical mechanisms of action seen in our treatment of different disease conditions. With regards to this pathway, once glucagon binds to the receptor, the receptor moves laterally in the plasma membrane and binds with a G-protein that is stationed as a peripheral protein to the plasma membrane. The G-protein contains three major domains, the alpha, the beta, and the gamma domain. The alpha domain is capable of binding to the GDP/GTP cofactor. When the G-protein is inactive, all three subunits stay together and the alpha subunit remains inactive and bound to GDP.

When the G-protein associates with an activated receptor, the alpha subunit exchanges GTP for the bound GDP cofactor, and the gamma and beta subunits dissociate into the cytoplasm. The activated alpha subunit moves laterally on the periphery of the plasma membrane until it contacts the adenylyl cyclase enzyme (also called adenylate cyclase). This activates the adenylyl cyclase that converts ATP into cyclic AMP (cAMP). cAMP production is an amplification step within this pathway. That means that more cAMP is produced than G-proteins are activated.

After some time a G-protein hydrolase will cause the hydrolysis of the GTP to GDP and inactivate the G-protein. At this point, the G-protein will associate with the gamma and beta subunits reforming its inactive state. Another glucagon signaling event will be required to reactivate the process. The cyclic AMP produced in the process serves as a second messenger in the process and activates a myriad of downstream targets. We will focus on two of the major targets.

The first is Protein Kinase A, which becomes activated upon binding with cAMP. The second target is a cAMP Response Element-Binding *Protein* (CREB). The CREB protein is also activated when bound to the cAMP molecule. This causes the CREB protein to translocate from the cytosol into the mitochondria and the nucleus. In both of these locations, the activated CREB binds to specific response element sequences in the DNA and activates the transcription of genes that are involved in gluconeogenesis. These genes and their encoded proteins have been discussed in more detail in chapter 14. What is important to note now is that glucagon signaling in the liver results in the upregulation of glucose production *de novo* from non-carbohydrate precursors. This is NOT a favored pathway in the body. It is expensive energetically for the liver to manufacture glucose. In fact, more expensive in the cost of ATP than can be produced from the newly formed molecule. However, organs like the brain can only utilize free glucose as an energy resource. Thus, the liver will engage in this energy deficit to build glucose for use by the brain and other cellular targets.

Glucagon signaling also leads to the downregulation of glycolysis, which we will cover in more depth in section 15.4 and glycogenesis. It also leads to an increase in glycogenolysis or the breakdown of glycogen. Let's take a further look at the regulation of both of these processes.

15.3.3: REGULATION OF GLYCOGENESIS -

Since glycogen synthase (GS) is the primary enzyme required for glycogenesis, it is also the primary target for the regulation of this pathway. Recall that GS is active in the dephosphorylated state. Thus, PKA down-regulates the activity of this enzyme through the phosphorylation of GS, as shown in Figure 15.3.9. Phosphorylation of GS causes it to shift into its inactive conformation and inhibits glycogenesis.

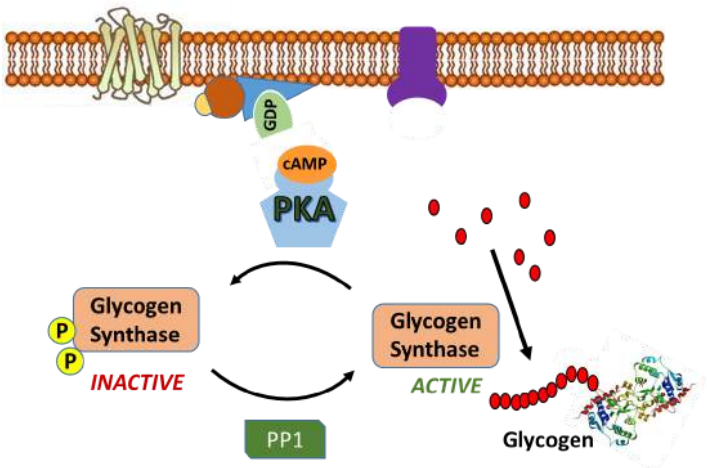


Figure 15.3.9: Inactivation of Glycogen Synthase through Phosphorylation. When bound to cAMP, protein kinase A (PKA) phosphorylates and inactivates the glycogen synthase enzyme. The activity of protein phosphorylase 1 (PP1) is required to dephosphorylate GS and restore its activity. Image modified from Yan, A., et al (2016) In J Biol Sci 12(12):1544-1554 and Servier Medical Art

In addition, activated PKA also phosphorylates the protein phosphatase 1 (PP1) enzyme leading to the inactivation of the phosphatase. PP1 normally dephosphorylates GS, helping to retain the active conformation of GS. Thus, phosphorylation of PP1 by PKA helps to maintain the GS in the phosphorylated, inactive state. The inhibition of PP1, is quite complicated, as shown in Figure 15.3.10. PP1 contains a regulatory domain and a catalytic domain. Normally the regulatory domain of PP1 binds with glycogen, keeping the molecule close to the location where GS will be present. Thus, when GS is near its substrate it can also bind with PP1 and be dephosphorylated into its active state. This is more efficient than diffusing in the cell and trying to find the PP1 randomly. When PKA phosphorylates the regulatory domain of PP1, it dissociates from the catalytic domain, causing the catalytic domain to float away from the glycogen molecule. This makes PP1 less efficient at dephosphorylating GS because it is harder for the molecules to randomly come into contact with one another. Thus, PP1 is less active. PKA reduces this activity even further, by phosphorylating an allosteric inhibitor (I) of PP1. In the phosphorylated state, the inhibitor can bind to PP1 fully inactivating the phosphatase. Both phosphorylation events need to be reversed to regain full PP1 activity.

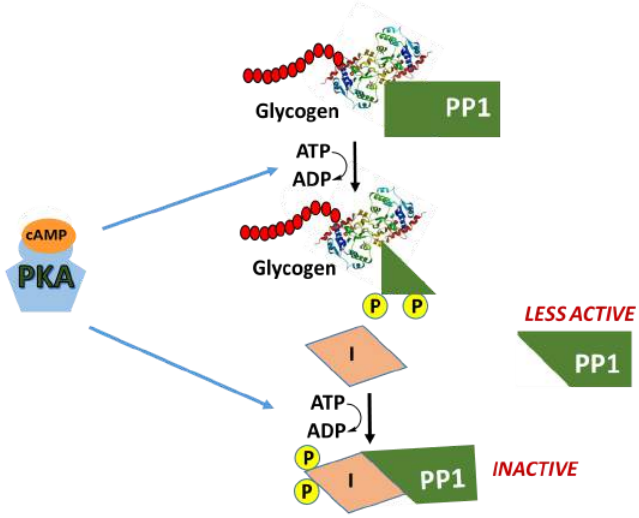


Figure 15.3.10: Inactivation of Phosphorylase 1 (PP1) by Protein Kinase A (PKA). PKA phosphorylates PP1 causing it to dissociate from glycogen and become less active. PKA also phosphorylates an allosteric inhibitor (I) of PP1 which increases its binding affinity for PP1. The phosphorylated inhibitor maintains PP1 in an inactive conformation. Image Modified from Yan, A., et al (2016) In J Biol Sci 12(12):1544-1554 and Servier Medical Art

In summary, glucagon signaling in the liver downregulates glycogenesis through the activation of PKA. PKA phosphorylates GS directly, inactivating the enzyme, and maintains it in the inactive state by also inhibiting the PP1 responsible for dephosphorylating GS.

15.3.4: ACTIVATION OF GLYCOGENOLYSIS

In addition to phosphorylating GS and PP1 during the inactivation of glycogenesis, PKA also phosphorylates the enzyme **Phosphorylase Kinase**, which is upstream of glycogen phosphorylase, the primary enzyme involved in glycogen breakdown. As its name implies, phosphorylase kinase is a protein kinase that phosphorylates the enzyme to activate it. Figure 15.3.11 offers a first view of the phosphorylation cascade required for glycogen phosphorylase activation.

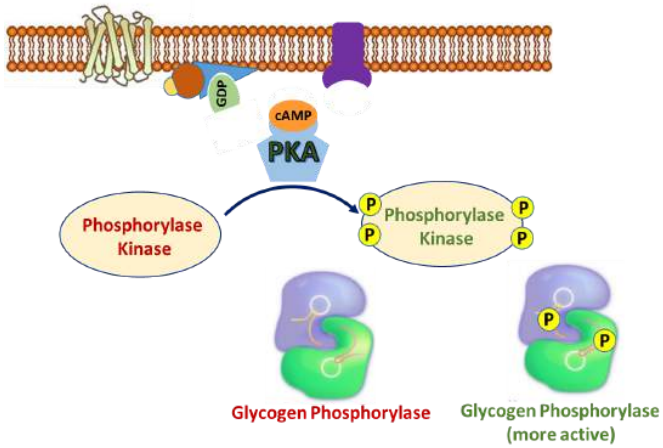


Figure 15.3.11: Activation of Glycogen Phosphorylase by Protein Kinase A (PKA) Signaling. PKA upregulates the activity of Phosphorylase Kinase through direct phosphorylation. The activated kinase enzyme phosphorylates its downstream target, Glycogen Phosphorylase (GP). In the phosphorylated state, GP is more active. Image modified from Yan, A., et al (2016) In J Biol Sci 12(12):1544-1554 and Llavero, F., et al (2019) Int. J. Mol. Sci. 20(23):5919

The phosphorylase kinase enzyme is a complex enzyme that is a tetramer of a tetrameric complex, αβγδ, so the full holoenzyme has an $(\alpha\beta\gamma\delta)_4$ structure. It is large with a molecular weight of around $1.3x10^6$. As you would expect, it is highly regulated in multiple ways, including phosphorylation by PKA, ADP (an allosteric effector), divalent cations like Ca^{2+}, and pH. The α and β are regulatory subunits that affect activity through their phosphorylation. The δ is calmodulin, a calcium-binding protein we discussed in Chapter 12.7, and its binding of calcium affects the holoenzyme activity. The γ subunit has kinase activity and has an N-terminal catalytic domain and a C-terminal calmodulin-binding domain. This is primarily regulated by phosphorylation through the PKA pathway as shown in Figure 15.3.5.

"PHK is one of the largest of the protein kinases and is composed of four types of subunit, with stoichiometry $(\alpha\beta\gamma\delta)_4$, and a total mol. wt of 1.3×10^6 Da. Activity is regulated by cyclic AMP-dependent protein kinase phosphorylation, autophosphorylation, allosteric effectors (e.g. ADP), metal ion concentration (Ca^{2+} and Mg^{2+}), proteolysis and pH (Pickett-Gies and Walsh, 1986). The α and β subunits are regulatory

and are the targets for control by phosphorylation. The δ subunit is essentially identical to calmodulin and confers Ca^{2+} sensitivity. The 386 amino acid γ subunit is the catalytic subunit which comprises an N-terminal kinase domain (residues 1–298) and a regulatory calmodulin-binding domain (residues 299–386)"

Calcium, an allosteric regulator, may be present within cellular targets due to nerve impulse firing, muscle contraction, or through hormone signaling. The presence of calcium in the cell generally indicates that there is high energy demand on the cell at that time and that energy production is needed. Thus, calcium binding to the phosphorylase kinase is a positive effector of the enzyme and upregulates activity. Maximal activity of the enzyme is achieved through combined phosphorylation and calcium binding. Thus, phosphorylase kinase can exist in 4 different states of activity as shown in linked equilibria in Figure 15.3.12.

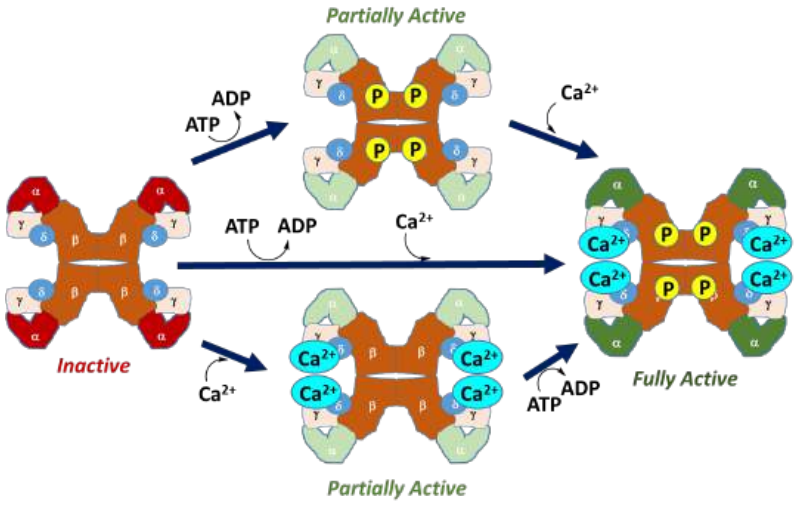


Figure 15.3.12 Activation States of Phosphorylase Kinase (PK). In the left-hand diagram, PK is in the inactive state, with the kinase-containing alpha domains shown in red. The upper diagram shows the activation of PK through phosphorylation by Protein Kinase A during hormone signaling. This leads to a partially active enzyme. Similarly, calcium binding, shown in the lower diagram, also results in a partially active enzyme. Calcium plays a particularly important role in the activation of this enzyme in skeletal muscle. The process of muscle contraction causes the release of high levels of calcium into the cytoplasm. Thus, the presence of calcium within the cytoplasm of muscle cells indicates high energy demand, as the muscle is being called into action. This activates PK and stimulates the breakdown of glycogen within muscle tissue to help meet energy demands. Maximal activity is obtained with both calcium binding and phosphorylation.

This diagram by now should be quite familiar. It is yet another example of a tetrameric enzyme (where the "enzyme" composition is αβγδ) existing in two major states, an inactive T state and an active R state, with the interconversion regulated by allosteric effectors (Ca^{2+}) and post-translationally phosphorylation.

A mechanism for the phosphorylation of a Ser in a substrate target protein (i.e glycogen phosphorylase) by the catalytic domain of the γ subunit of phosphorylase kinase is shown in Figure 15.3.13.

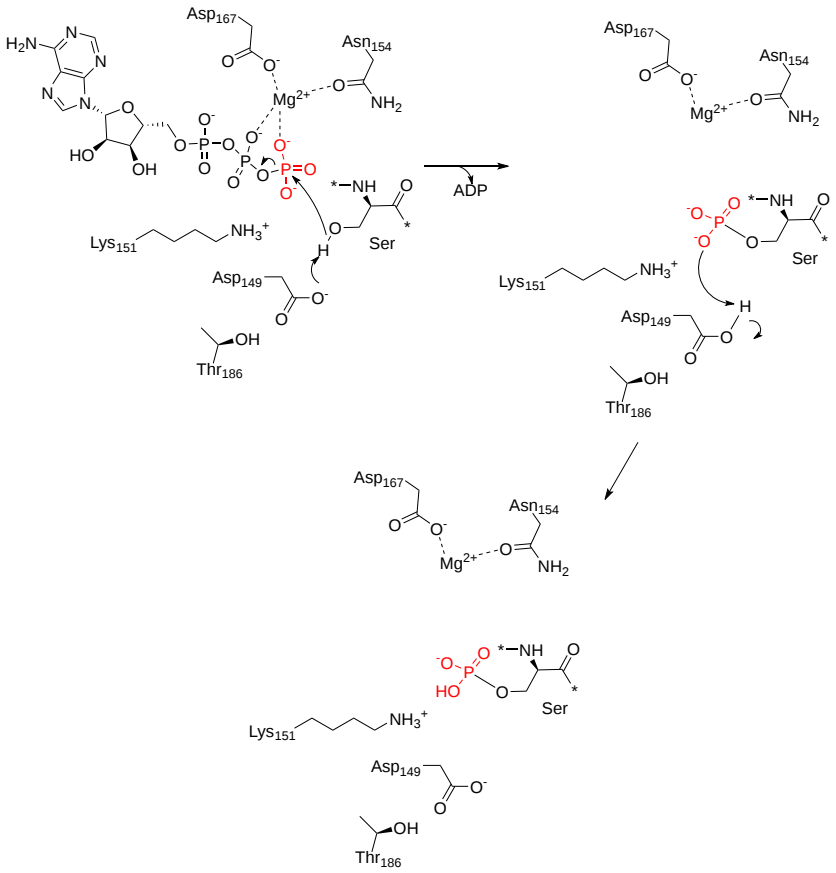


Figure 15.3.13: A mechanism for the phosphorylation of a Ser in a substrate target protein (i.e glycogen phosphorylase) by the catalytic domain of the γ subunit of phosphorylase kinase. https://www.ebi.ac.uk/thornton-srv/m-csa/entry/35/. Creative Commons Attribution 4.0 International (CC BY 4.0) License.

The * in the mechanism denotes the serine of the target protein.

Figure 15.3.14 below which shows an interactive iCn3D model of a truncated form of the rabbit phosphorylase kinase gamma subunit dimer bound to a peptide substrate complex (2PHK).

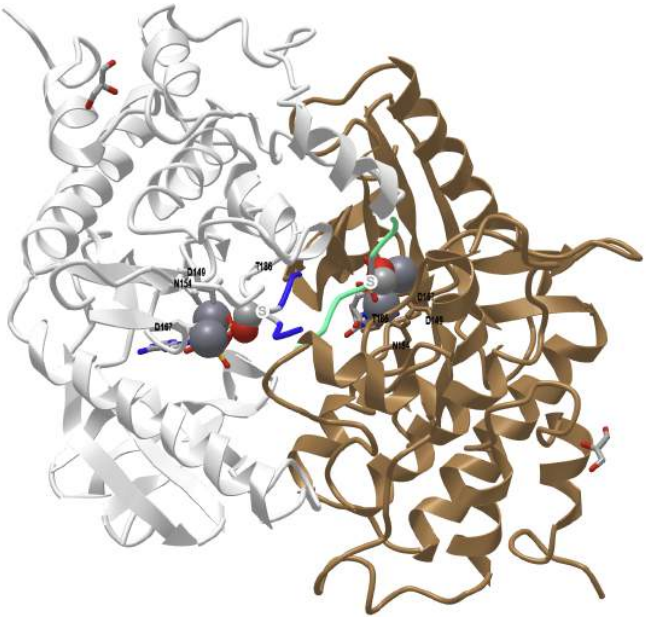


Figure 15.3.14 : Rabbit phosphorylase kinase gamma subunit dimer bound to a peptide substrate complex (2PHK). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...3wHapyHWEy7qi9

The phosphorylase kinase dimer is shown (gray and brown subunits). A non-hydrolyzable ATP analog (adenylyl imidodiphosphate, AMPPNP) in each subunit is shown as CPK-colored sticks. The backbone of two identical peptide substrates (blue and cyan) with the sequence RQMSFRL (similar to the target sequence in glycogen phosphorylase and an ideal peptide substrate) is shown as a backbone trace, with the central Ser which gets phosphorylated shown as CPK-colored spheres. Active site residues are shown as CPK-colored sticks and labeled in each subunit.

We have been primarily discussing the regulation of glycogenolysis in the liver. However, in considering the activity of PK and its reactivity with Ca^{2+} ions, we should also consider the activation of glycogenolysis in skeletal muscle, as well. Of note, glycogenolysis in liver tissue and skeletal muscle has many differences. First, the G-protein coupled pathway is activated by different hormones. Liver tissue is responsive to Glucagon stimulation, as well as stimulation through the Epinephrine hormone signaling pathway. Glycogenolysis in skeletal muscle tissue, on the other hand, is only activated by the Epinephrine signaling pathway, but not by glucagon. This is because the liver is the primary organ responsible for regulating blood glucose levels. Thus, pancreatic signaling due to low blood glucose levels primarily targets glycogenolysis within the liver tissue. Both systems are responsive to Epinephrine, which is described in more detail below.

15.3.5: EPINEPHRINE SIGNALING

Epinephrine is a small amino acid-derived hormone (can you guess the amino acid?? Yes it is Tyrosine!!), as shown in Figure 15.3.15. It is also called adrenaline, as it is secreted from the adrenal glands located just above the kidneys, during the flight or fight response. It is also secreted during heavy or sustained exercise. Epinephrine has pleiotropic responses in the body, which include the activation of glycogenolysis in the liver and skeletal muscles. Epinephrine also promotes fat breakdown in adipose tissue, which releases this energy reserve into the bloodstream for utilization by muscle tissue. It also causes the relaxation of smooth muscles in the lungs and respiratory tract enabling better oxygen absorption. Cardiac contractility is also increased to increase blood flow to skeletal muscles. This supports the generation of ATP from glucose and fatty acids for sustained muscle utilization. It also reduces blood flow to the skin and causes the contraction of smooth muscles in the skin causing goosebumps.

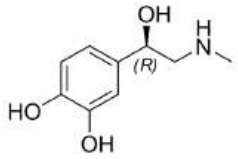


Figure 15.3.15: Structure of Epinephrine

Within the liver and skeletal muscle, the epinephrine signaling pathway overlaps with the glucagon signaling pathway seen in the liver. The epinephrine receptor is also a G-protein coupled receptor related to the glucagon receptor. However, it is specific for epinephrine and cannot bind with glucagon. It does activate the same G-protein pathway leading to Protein Kinase A activation, as shown in Figure 15.3.16. The body is very efficient at reusing machinery in different parts of the body, in this case, it does so under different regulatory parameters.

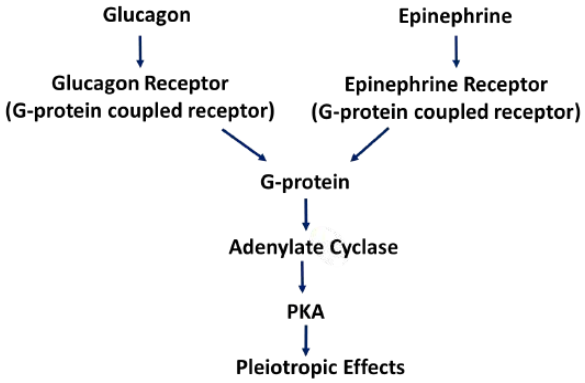


Figure 15.3.16: Similarities of Glucagon and Epinephrine Signaling Pathways

The Glycogen Phosphorylase enzymes are also encoded by different genes within the Liver and Skeletal Muscle. These are known as isozymes. Recall, that isozymes have the same biological function, but since they are expressed from different genes, they have different enzyme kinetics and they are regulated in different and unique ways within each tissue.

The liver and skeletal muscle forms of Glycogen Phosphorylase share approximately 90% sequence identity. Yet again, both isozymes can exist in two major conformations, the a-form, and the b-form. The protein adopts the a-form when it is phosphorylated at Ser 12 by

phosphorylase kinase, as shown in Figure 15.3.17. The Glycogen Phosphorylase enzyme can also be in two different states, the relaxed, flexible state which is the active form of the enzyme, and the tense or rigid state which is inactive. When the protein is in the a-conformation, it favors the relaxed and active state of the protein. Therefore, phosphorylation of Glycogen Phosphorylase leads to an increase in the activity of the enzyme. This is depicted in the following diagram.

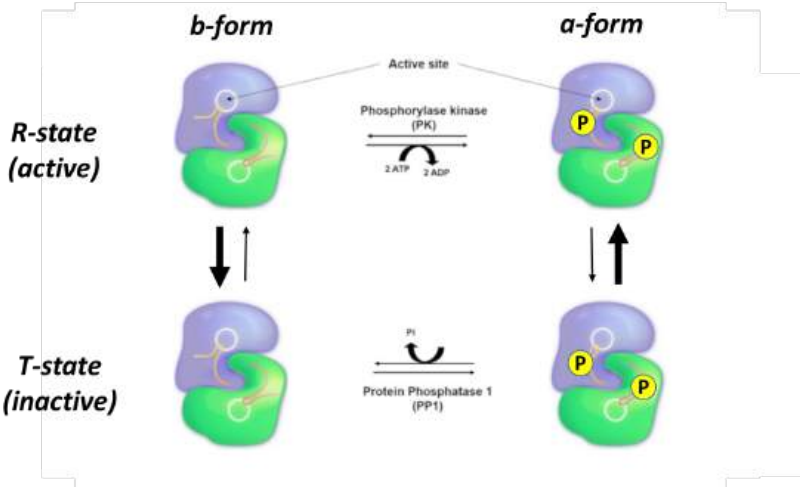


Figure 15.3.17: Enzymatic States of Liver and Skeletal Muscle Glycogen Phosphorylase. Both isozymes of Glycogen Phosphorylase are responsive to phosphorylation by phosphorylase kinase. Phosphorylation of Glycogen Phosphorylase causes it to shift from the b-form to the a-form of the protein. The a-form of the protein favors the relaxed, active state of the protein, whereas the b-form of the protein favors the tense and inactive state of the protein. Image modified from Llavero, F., et al (2019) Int. J. Mol. Sci. 20(23):5919

Another way to think about these four different enzyme states is that they exist in a dynamic equilibrium, with the population of each state determined by the phosphorylation state AND the presence of allosteric inhibitors and activators which we will explore below. In Figure 17 above, the R state is on top and the T state is on the bottom. The equilibrium between just those two states is indicated by the vertical arrows. The thickness of the arrows indicates the preferred direction of the reversible reaction. In the absence of phosphorylation (left-hand vertical states), the equilibrium favors the T or inactive form. When phosphorylated (right-hand vertical states) on Ser12 by the enzyme phosphorylase kinase, the R or active form is favored. The horizontal equilibria show the phosphorylation of the Ser12 by phosphorylase kinase (top, shown reversibly but the enzyme is not acting physiologically to remove phosphate) and dephosphorylation of Ser 12 by the enzyme protein phosphatase 1 (which acts physiologically only as a phosphatase).

The different isozymes of the Glycogen Phosphorylase enzyme are also regulated by different, tissue-specific allosteric effectors. Within the liver, glucose is a negative regulator of Glycogen Phosphorylase, which makes sense, as the role of this pathway in liver tissue is to promote the release of glucose into the bloodstream. The presence of free glucose in the cytoplasm of the liver would indicate either the fed-state when blood glucose levels are high, or that high levels of glycogenolysis have released substantial glucose. Within liver tissue, the presence of free glucose will cause the a-form of Glycogen Phosphorylase to shift to the Tense state, reducing the activity of the enzyme. This is shown in Figure 15.3.18.

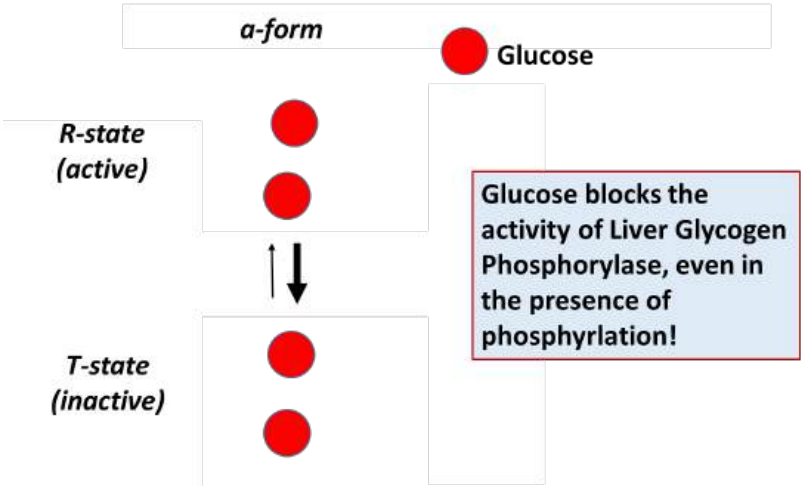


Figure 15.3.18: Regulation of Liver Glycogen Phosphorylase by Free Glucose. When the liver isoform of Glycogen Phosphorylase is in the a-form (phosphorylated), it can be shifted into the Tense state in the presence of high levels of free glucose. This blocks the glycogen binding site of the enzyme, essentially serving as a competitive inhibitor of the enzyme. Image modified from Llavero, F., et al (2019) Int. J. Mol. Sci. 20(23):5919

Skeletal muscle glycogen phosphorylase (or Myophosphorylase, as it is sometimes called), is more responsive to allosteric effectors that indicate the energy state of the cell. This makes sense, as the main purpose of glycogen breakdown in muscle tissue is to fuel the energy demand for the muscle tissue. Thus, the energy housed in glucose will be used to produce ATP within these cells. The presence of either glucose 6-phosphate or ATP within skeletal muscle indicates high levels of energy are present. Thus, glycogen breakdown will be inhibited. The presence of AMP, on the other hand, indicates a low energy state and is an activator of Glycogen Phosphorylase. This is shown in Figure 15.3.19 to Figure 15.3.21 below.

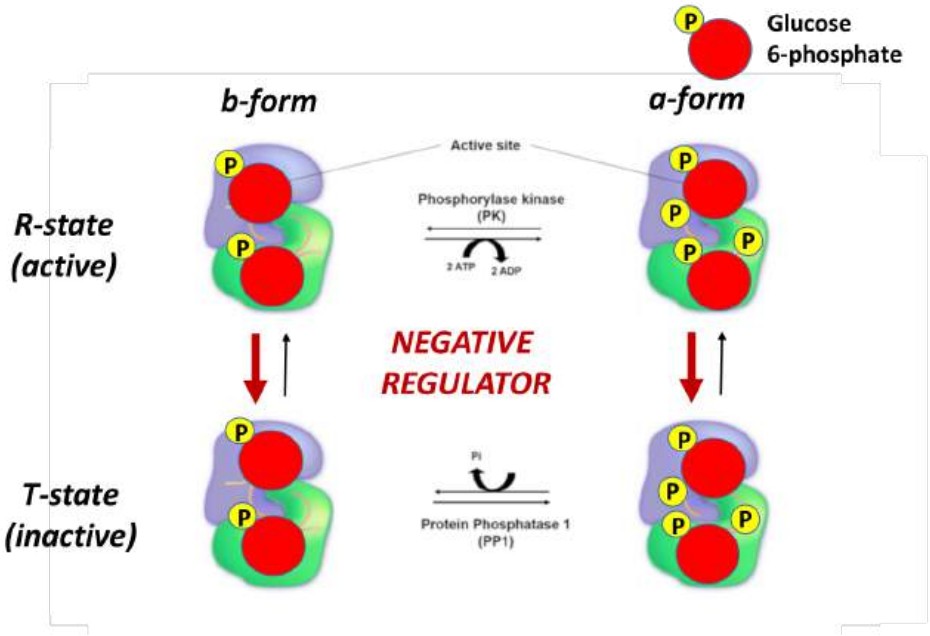


Figure 15.3.19: Regulation of Skeletal Muscle Glycogen Phosphorylase by Glucose 6-phosphate. Muscle Glycogen Phosphorylase is negatively regulated by the presence of high levels of Glucose 6-phosphate. Regardless of the form (a or b) the enzyme is shifted into the Tense state. Image modified from Llavero, F., et al (2019) Int. J. Mol. Sci. 20(23):5919

Again, note that the bold red arrows point downward, showing that on the addition of glucose-6-phosphate, the T state (inactive) is favored even if glycogen phosphorylase has been phosphorylated.

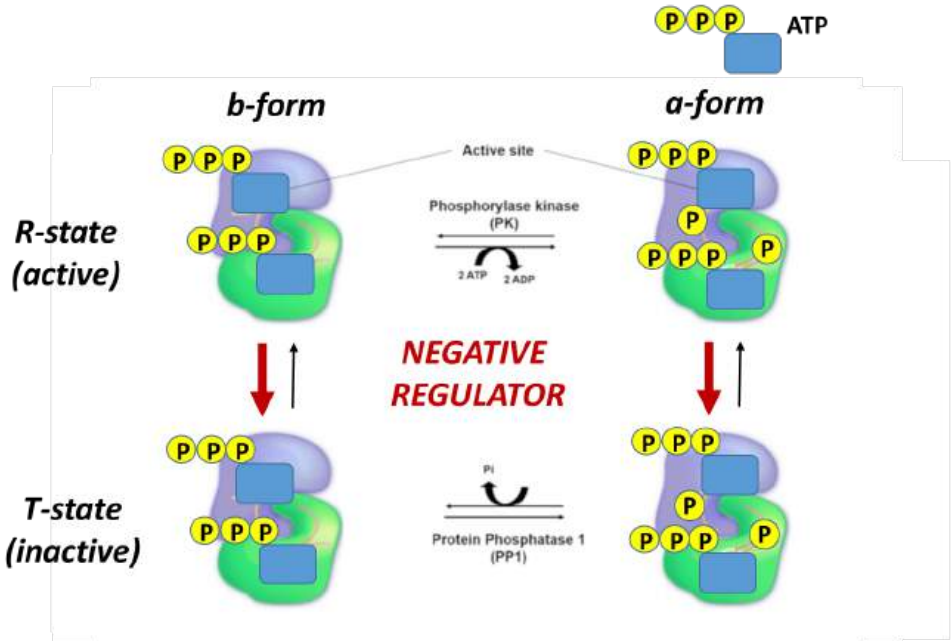


Figure 15.3.20: **Regulation of Skeletal Muscle Glycogen Phosphorylase by ATP.** In the presence of high levels of ATP, GP is converted to the Tense state, showing the inhibition of the enzyme in the presence of high levels of energy. Image modified from Llavero, F., et al (2019) Int. J. Mol. Sci. 20(23):5919

Again, note that the bold red arrows point downward, showing that when the ATP levels are high, the T state (inactive) is favored even if glycogen phosphorylase has been phosphorylated. Again, under conditions of a high energy state (as reflected by high ATP), there is no need to cleave glycogen to enter glycolysis.

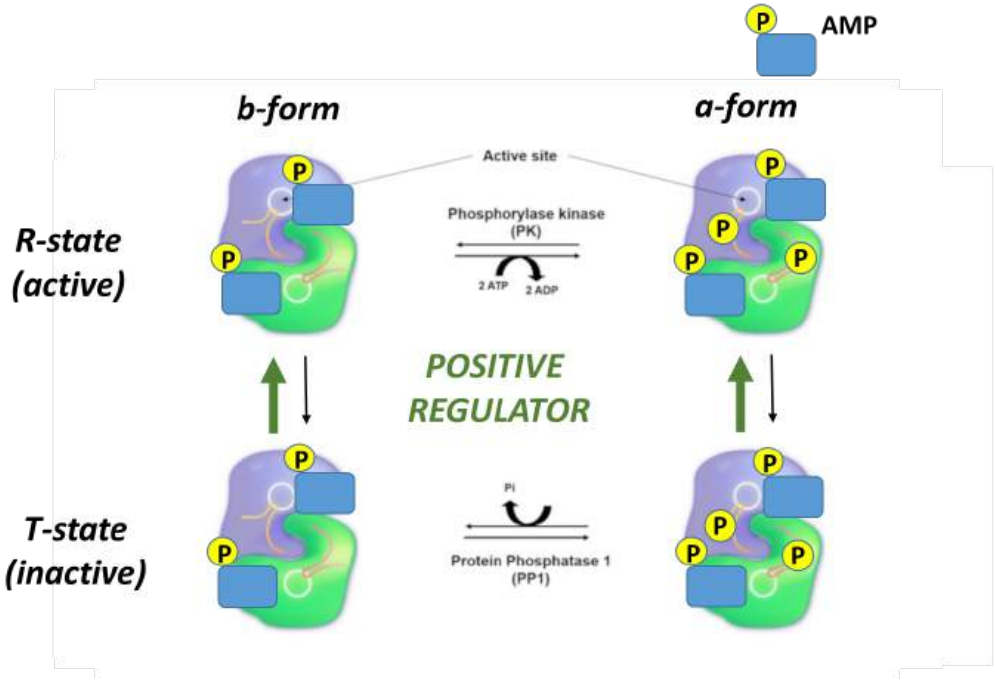


Figure 15.3.21: Regulation of Skeletal Muscle Glycogen Phosphorylase by AMP. In the presence of low-energy indicators, such as AMP, the glycogen phosphorylase enzyme is activated, even in the absence of phosphorylation. Thus, when AMP is bound, the b-form of glycogen phosphorylase is converted into the relaxed and active state. Image modified from Llavero, F., et al (2019) Int. J. Mol. Sci. 20(23):5919

In contrast, note that the bold green arrows point upward showing that when the AMP levels are relatively high, the R state (active). Higher levels of AMP reflect a need to activate glycogen breakdown to increase ATP production.

A mechanism for the phosphorolysis of Glc_{n+1} to Glc_n and glucose-1-phosphate is shown in Figure 15.3.22. Note the unusual presence of a molecule of pyridoxal phosphate covalently attached through a Schiff base linkage to Lys 568 (rabbit phosphorylase).

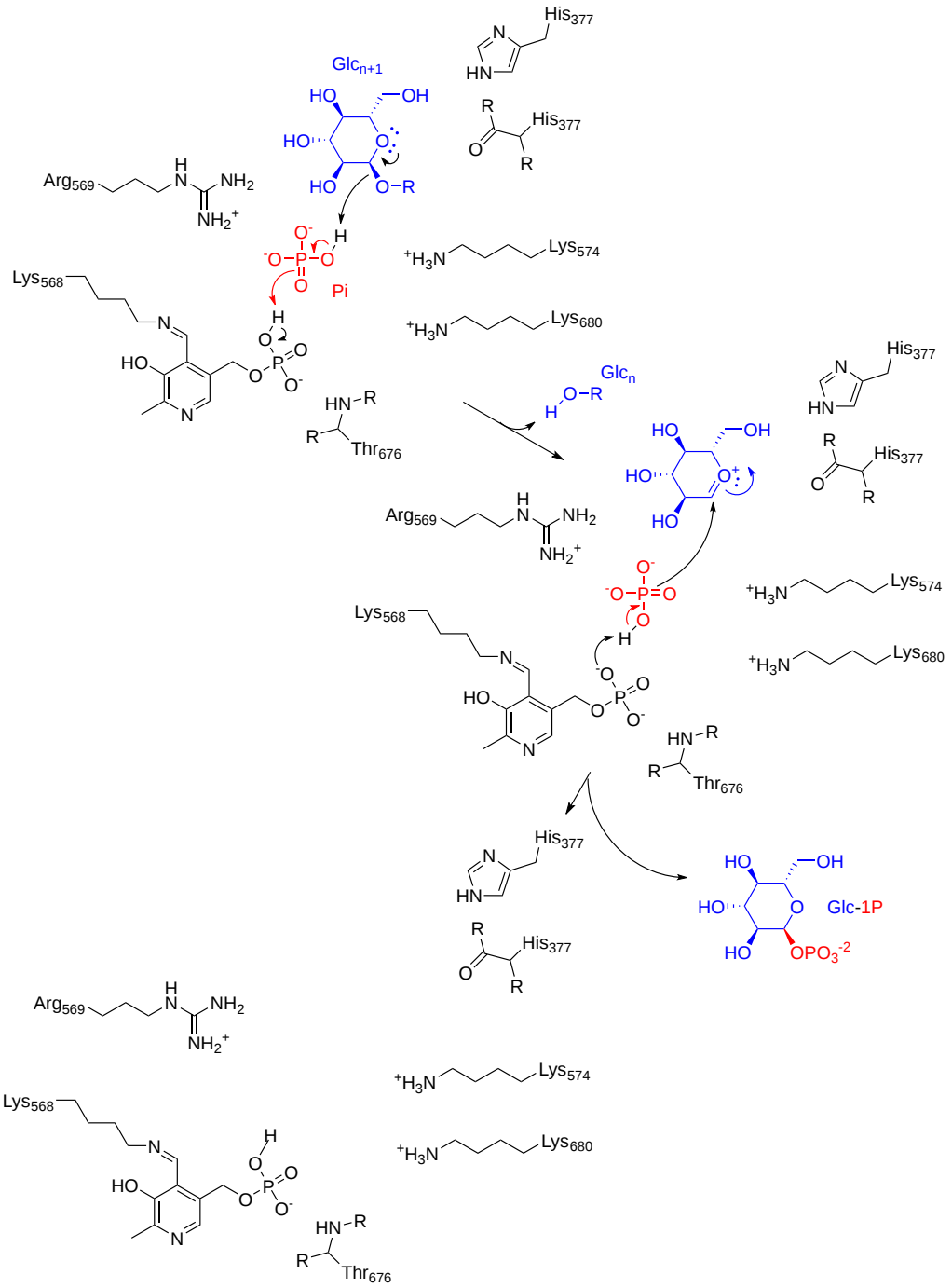


Figure 15.3.22: A mechanism for the phosphorolysis of Glc_{n+1} to Glc_n and glucose-1-phosphate. https://www.ebi.ac.uk/thornton-srv/m-csa/entry/205/. Creative Commons Attribution 4.0 International (CC BY 4.0) License

PLP, which is covalently attached through a Schiff base, functions in this enzyme as a general acid and base, and not as a cofactor that facilitates covalent bond cleavage in amino acid substrates that are covalently attached to it (Chapter 6.8: Cofactors and Catalysis - A Little Help From My Friends).

It is important to note that the reaction is a phosphorolysis, not a hydrolysis, which would leave free glucose, which would more readily leave the cell and hence be less available for cellular energy needs and less available for glycolysis.

Figure 15.3.23 below which shows an interactive iCn3D model of a dimer of rabbit glycogen phosphorylase (1GDB).

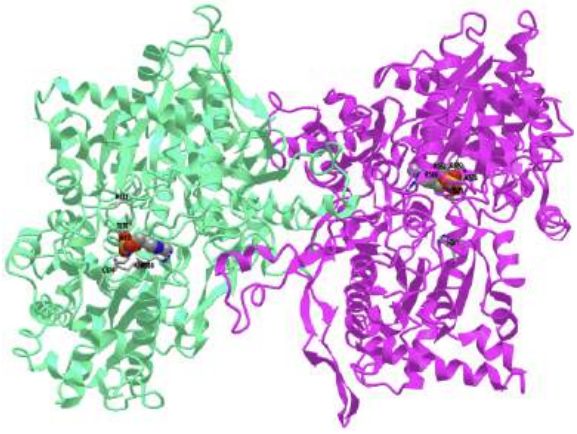


Figure 15.3.23 : Dimer of rabbit glycogen phosphorylase (1GDB). (Copyright; author via source). Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...s1LosR5TJLjru9

The subunits in the dimeric form are shown in different colors. PLP is shown in spacefill. Active site residues are shown in both subunits as CPK-colored sticks and labeled.

Let's look at a monomer (from the tetramer) of 2 nonphosphorylated states of glycogen phosphorylase. Since they are both unphosphorylated, they both represent the **b state**. One (**2GPB**) has glucose bound, so it represents the inactive (T state). The other (**3E3N**) has AMP bound, so it represents the active (the R state). Figure 15.3.24 shows the conformation differences between the monomeric states

Figure 15.3.24: Conformational differences between the monomeric unphosphorylated b state of glucose-bound GP (T state, inactive) and AMP-bound GP (R state, active)

Now let's look at the conformations of two different GP activated by different means. In one, the nonphosphorylated form of GP (the b state, and inactive T state,) binds AMP, an allosteric activator (pdb 8GPA), and converts to the active, b state (unphosphorylated R state). Let's compare its active conformation to a form of GP activated by phosphorylation (the phosphorylated **a** state, and active R state) to which just SO_4^{2-} is also bound. In the 1st case, GP-b T state is driven to the active Gp-b R state by binding of the allosteric activator AMP. In the second case, GP-a is already in the R active state since it is phosphorylated. Figure 15.3.25 compares their conformations.

Figure 15.3.25: Comparison of the conformations of two active forms of GP - phosphorylase b bound to the allosteric activator AMP (R state) and phosphorylase a activated by phosphorylation of Ser 14.

The cyan monomer is glycogen phosphorylase b which is not phosphorylated but driven into the **R state** on the binding of AMP (note that two are bound at the periphery). The dark blue monomer is glycogen phosphorylase a which is phosphorylated at Ser14 (a different number in this crystal file which is shown in spacefill and labeled SEP-14) which is also in the R state. It also has two SO_4^{2-} bound that help stabilize the state. Look carefully! the conformations are very similar to each other, in contrast to the different conformations for the T and R states shown in Figure 24.

Figure 15.3.26 below which shows an interactive iCn3D model of unphosphorylated (b state) rabbit glycogen phosphorylase with bound glucose (inactive T state, **2GPB).**

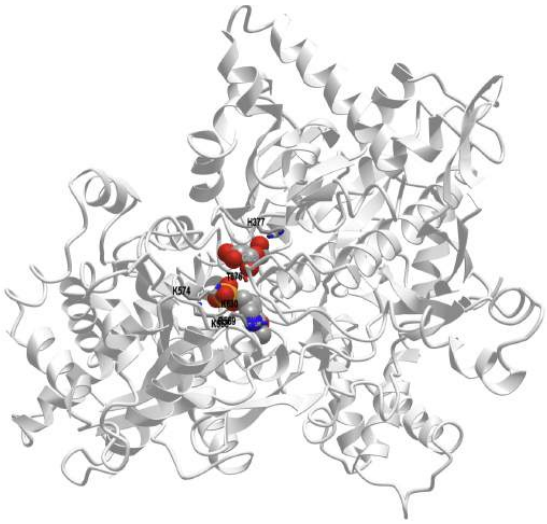


Figure 15.3.26 : Unphosphorylated (b state) rabbit glycogen phosphorylase with bound glucose (inactive T state, 2GPB). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...eXQciAPczmQpLA

Only 1 monomer of the tetramer is shown. PLP and glucose are shown in spacefill, CPK colors, and labeled. Key active site residues are shown as colored sticks and labeled.

Figure 15.3.27 below which shows an interactive iCn3D model of unphosphorylated (b state) rabbit glycogen phosphorylase with bound AMP (active R state, 3E3N).

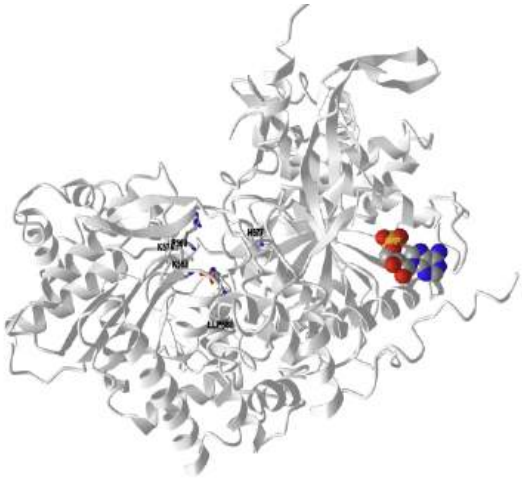


Figure 15.3.27 : Unphosphorylated (b state) rabbit glycogen phosphorylase with bound AMP (active R state, 3E3N). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...WDZJc1CcpFqEG7

Note the different locations for the binding site of the allosteric activator AMP compared to the inhibitor glucose in the previous model.

MCARDLE'S DISEASE

McArdle's Disease, also referred to as myophosphorylase deficiency or type V glycogen storage disease, is a recessive inherited disorder characterized by an inability to metabolize glycogen due to the absence of a functional myophosphorylase (PYGM). In Figure 15.3.28 shown below, the normal functional pathway is shown in blue, on the left, while the mutant pathway is shown on the right in red. Patients with this disease lack sufficient glucose-1-phosphate (G1P) monomers needed for glycolysis and the hexosamine biosynthetic pathway (HBP). This results in lower ATP and, consequently, lower muscle contraction, as well as in lower post-translational modifications by O-GlcNAcylation in comparison to normal conditions. This is especially pronounced during extended or heavy workouts, where people with McArdle's Disease will sustain painful cramping of their muscle tissue during workouts, can have dark red/brown urine, and can easily tire during activity. Some patients also note a second-wind phenomena occur during workouts as

the body shifts from carbohydrates to lipids as a primary energy source. The dark red/brown color in the urine happens if muscle tissue is damaged during the workout. The damaged muscle releases the protein myoglobin into the bloodstream. This is filtered out by the kidneys and excreted in the urine, causing the color change. The severe uncontrolled disease can cause life-threatening kidney problems.

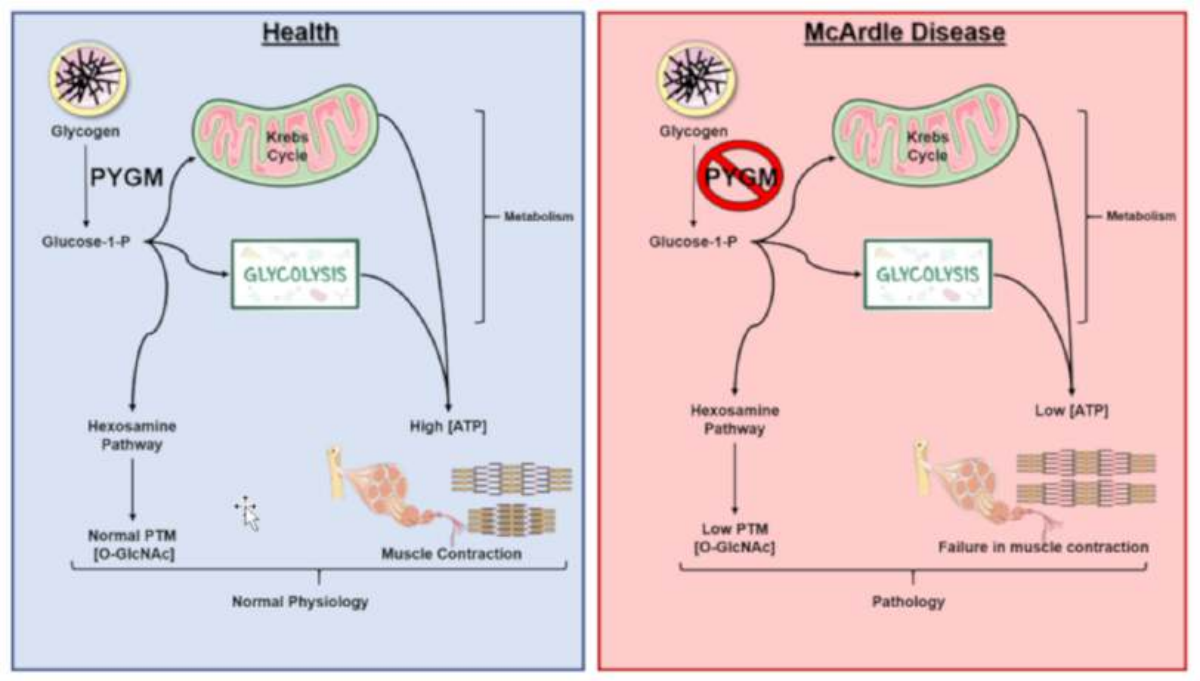


Figure 15.3.28 McArdle's Disease. *Left hand side (blue)* represents the normal pathway, whereas the *right hand side (red)* notes the deficiency of muscle glycogen phosphorylase. Image from Llavero, F., et al (2019) Int. J. Mol. Sci. 20(23):5919

Figure 15.3.29 presents an overview of glucose metabolism in skeletal muscle. Both glucose-1-phosphate (G1P) released from the intracellular glycogen stores by glycogen phosphorylase (GP), as well as the glucose introduced into the cell through glucose transporters (GLUT) are converted to glucose-6-phosphate (G6P) by phosphoglucomutase (PGM) and hexokinase (HK), respectively. The G6P can be directed to different destinations. One of them is the pentose phosphate pathway for the formation of nucleic acid building blocks (ribose and deoxyribose). Another is in the formation of ATP. Here G6P enters the metabolic pathway of glycolysis. The glycolytic reactions culminate in the production of pyruvate and adenosine triphosphate (ATP). Pyruvate can be fermented to lactate by the catalysis of the lactate dehydrogenase (LDH), as happens during anaerobic muscle exercise. On the other hand, pyruvate can be used to obtain ATP through full oxidation in the Kreb Cycle. In total, oxidative phosphorylation produces between 30-36 molecules of ATP (depending on the organism and tissue), 6 molecules of carbon dioxide (CO_2), and 6 molecules of water (H_2O) from 1 glucose molecule. Glycolysis alone only produces two net ATP molecules per glucose. Glucose, in addition to being the main fuel of the cell's energy metabolism, is also used by the cellular machinery as a vitally important substrate for the production of key intermediaries of the hexosamine biosynthetic pathway (HBP) forming O-GlcNAc, β-linked N-acetylglucosamine. And finally, in times of plenty, glucose will be utilized by glycogen synthase (GS) to make glycogen.

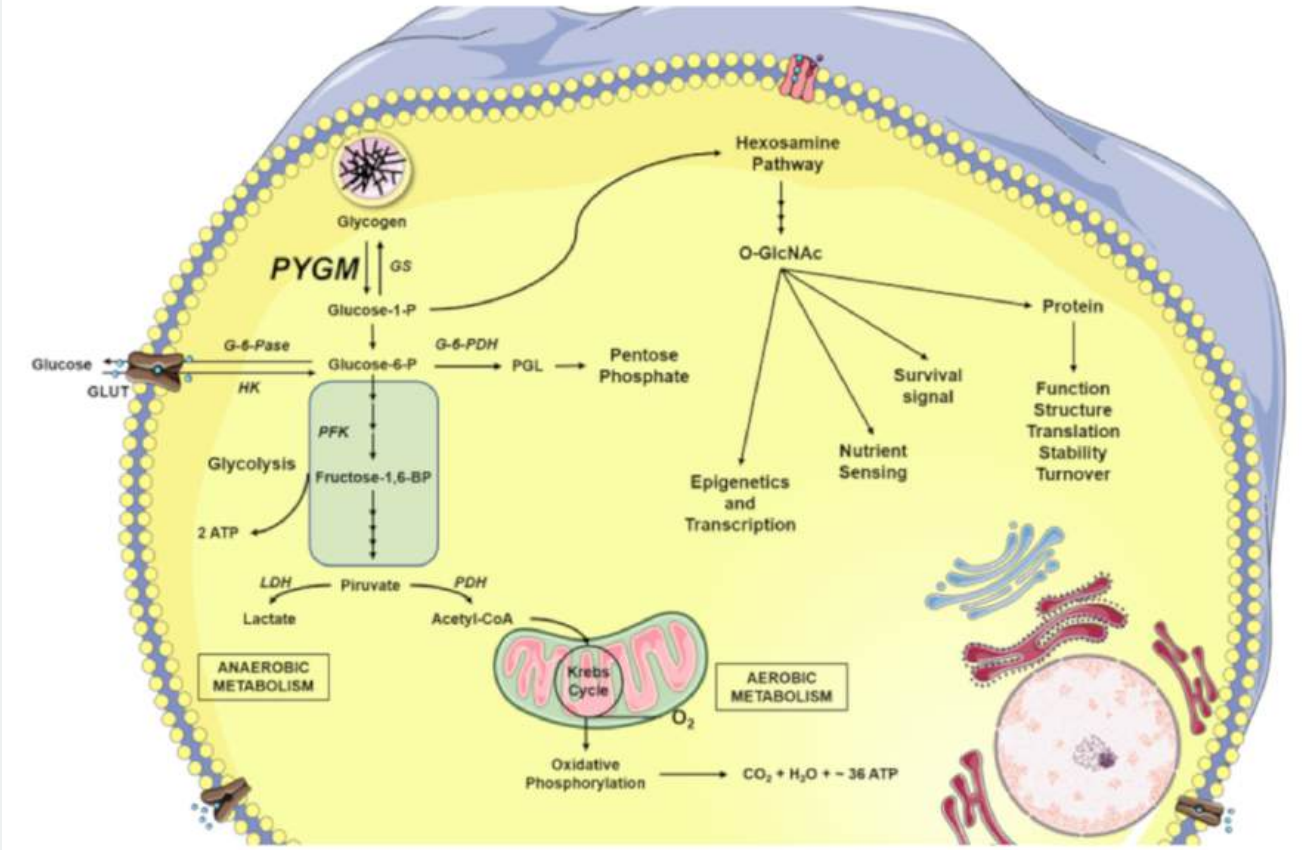


Figure 15.3.29 Summary of Glucose Metabolism in Skeletal Muscle. Image from Llavero, F., et al (2019) Int. J. Mol. Sci. 20(23):5919

15.4: REGULATION OF GLYCOLYSIS

There are three major enzymatic control points within the glycolytic pathway. These include hexokinase, phosphofructokinase, and pyruvate kinase reactions. Key drivers for regulating the pathway are energy demand within the cell as determined by local indicators such as ATP and AMP, as well as energy demand within the organism as a whole, which can be influenced by hormone signaling pathways. We will also see that the regulation of the pathway can vary depending on cell type and cellular needs.

15.4.1: HEXOKINASE REGULATION

One of the primary mechanisms that control the regulation of the hexokinase step in glycolysis is the presence of different hexokinase enzymes in different cellular types. Essentially, these are proteins that are encoded by different genes but perform the same function within the cell. They are known as isozymes. Isozymes can have different enzyme kinetics, different expression patterns in different tissues, different post-translational modifications, and bind with different allosteric effectors. This affords the body to have differential control over the same processes in different locations within the body. Within vertebrates, four important hexokinase isozymes vary in subcellular locations and kinetics parameters. This allows the differential phosphorylation of hexoses depending on local conditions, and physiological function. They are designated hexokinases I, II, III, and IV. All of the Hexokinases can use multiple hexoses as substrates, in addition to glucose. These include mannose, fructose, and 2-deoxyglucose. Hexokinase IV is also often referred to as Glucokinase and is specific to the Liver and Pancreas.

Recall that hexokinase enzymes mediate the first step in the glycolytic pathway with the formation of glucose 6-phosphate, as shown in Figure 15.4.1 below. They also require ATP as a cofactor in the process.

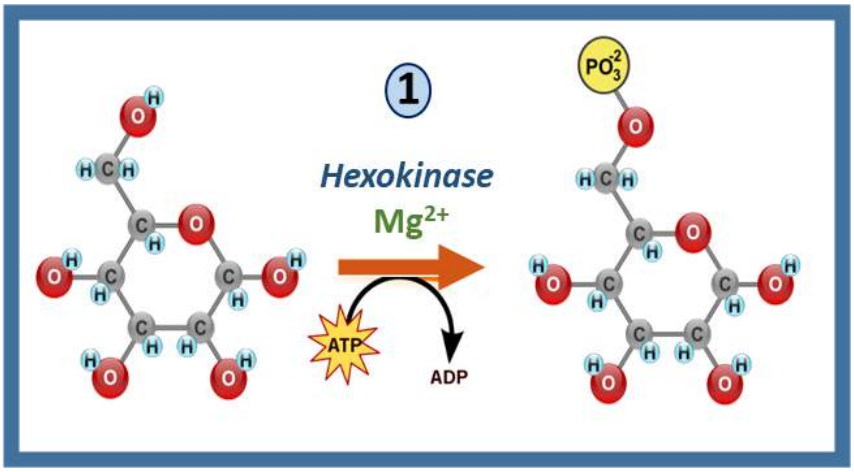


Figure 15.4.1: Glucose Phosphorylation by Hexokinase Enzymes. Figure modified from YassineMrabet

Recall that glucose-6-phosphate (G6P) has several potential fates within the body, as shown in Figure 15.4.2. It can be used as an energy source through the pathways of glycolysis and aerobic respiration. Short bursts of anaerobic respiration can also be sustained in animals that convert pyruvate into lactate. G6P can also be dephosphorylated in the liver and released back into the bloodstream to maintain homeostatic balance. The pancreas uses G6P as a sensor to determine when to secrete insulin and glucagon. The G6P can also serve as a building block for anabolic processes. It can be converted to ribose through the Pentose Phosphate Pathway where it will be used in the construction of nucleotide monomers. It can also be used for the formation of hexosamines used in proteoglycan and glycoprotein formation.

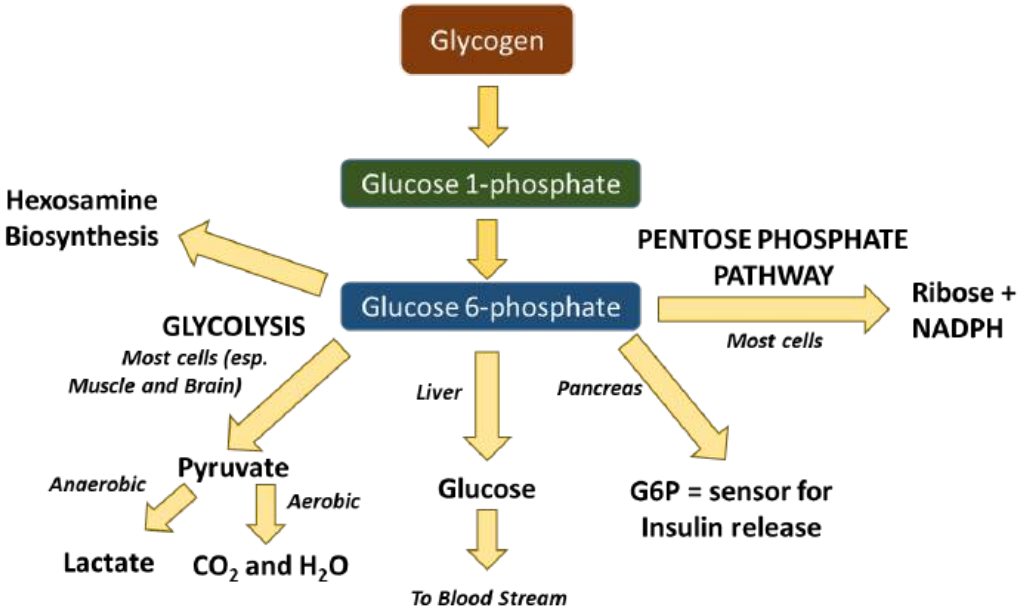


Figure 15.4.2: Cellular Fates of Glucose 6-Phosphate

Now let's take a look at the different isozymes of hexokinase in a little more detail. Hexokinase I (HKI) is found widely distributed throughout the body and is the main form expressed in brain tissue and red blood cells, as shown in Figure 15.4.3. In brain cells, this protein is localized to the mitochondria. This colocalization aids in the efficient coupling of glycolysis and the Krebs' cycle and oxidative phosphorylation pathways inside the mitochondria. It also ties the activity of HKI with oxidative phosphorylation and energy load, as HKI preferentially uses mitochondrially-derived ATP in its reaction mechanism. HK association with the mitochondria also has a cellular protective effect, reducing the potential for programmed cell death or apoptosis to occur. Red Blood Cells (RBCs), on the other hand, are highly differentiated cells with a very short lifespan. They are replaced in humans approximately every two weeks. RBCs are enucleated and do not have mitochondria, and thus, only generate ATP through the process of glycolysis. the HKI protein is free floating in the cytoplasm in this system. HKI has a low Km, meaning that it has a high affinity for glucose and is active at low substrate concentrations. It is also inhibited by the product Glucose-6-Phosphate (G6P) in the process of ***negative feedback inhibition.*** Essentially, you do not want to waste time and energy making more than you need. Low to moderate levels of free inorganic phosphate can overcome this negative feedback inhibition.

Figure 15.4.3: Localization of Hexokinase I. Hexokinase I (HKI) is found widely throughout the body and is the primary hexokinase found in brain tissue and red blood cells. Within the brain tissue, HKI is localized to the mitochondria. Images from Smart Servier Medical Art and Jessica Polka

HKI and HKII are expressed in Skeletal Muscle, Heart Muscle, and insulin-sensitive tissues. While it is thought that HKI is providing a predominantly catabolic role for the use of G6P in energy production, HKII may play a more pertinent role in anabolic processes, providing G6P for conversion to G1P and subsequent utilization in Glycogenesis. Both HKI and HKII are localized to the outer membrane of the mitochondria. However, while 95% of HKI is associated with mitochondria, only about 70% of HKII is associated, with the remaining HKII fractionating with the cytosolic proteins. This could help to explain the heightened role of HKII in anabolic glycogenesis processes within skeletal muscle, and why it is not found in brain tissue. HKII is also often overexpressed in tumor cells, where it is associated with higher mortality rates. It has also been linked with the processes of metastasis and with the development of drug resistance. Similar to HKI, HKII also has a low Km and is inhibited by G6P, although this inhibition is not released by the presence of inorganic phosphate.

Not a lot is known about the functions of HKIII. It may be an inactive gene duplication or remnant. Under basal conditions, it is not expressed to appreciable levels in any major tissues, and studies on its biological activity it is inhibited by glucose at physiological concentrations. However, some studies suggest that it may be expressed during cellular stress responses, such as hypoxia, although its function in these types of responses is not currently understood.

Hexokinase IV or Glucokinase is specifically expressed within the liver and pancreas. HKIV is cytoplasmic and not tethered to the mitochondria. Activity within the pancreas serves as a sensor for the release of insulin, and in the liver for the production of G6P that will fuel glycogen production. HKIV has a higher Km than HKI and HKII, thus it does not work efficiently at low concentrations of glucose. However, it is NOT inhibited by the product, G6P. Thus, it will continue to make G6P, even when levels are high. This helps to explain the high levels of glycogen that are stored within liver tissue, but not elsewhere in the body. This also ensures that the sensor system in the pancreas will accurately read blood glucose levels.

The four isozymes of HK share high homology with one another and appear to have arisen from gene duplication events, as shown in Figure 15.4.4. The left-hand panel of Figure 15.5.4 shows the linear protein domains of the different HK isozymes. Both HKI and HKII contain an N-terminal domain that localizes the protein to the mitochondrial membrane. HKI, II, and III all contain two repeating catalytic domains in the N- and C-terminals. However, mutations in the N-terminal domain in HKI and HKIII render them inactive. Both catalytic domains in HKII retain activity. HKIV (Glucokinase) is the most truncated isozyme, only containing the C-terminal catalytic domain. The lower right diagram shows the ribbon diagram of HKIV. Upstream promoter regions of HKIV (not shown in this diagram) also differ, allowing for controlled expression in the liver and pancreas. Feedback inhibition in HKI and HKII occurs through the N-terminal catalytic domain. The upper diagram on the right shows an HKI dimer complex with an ATP analog, glucose, G6P, and Mg2+ ion. HKI dimerizes when concentrations of the inhibitor, G6P, are high enough. Dimerization reduces the biological activity of the enzyme in brain tissue. Dimerization of HKI can be reversed in the presence of low levels of inorganic phosphate.

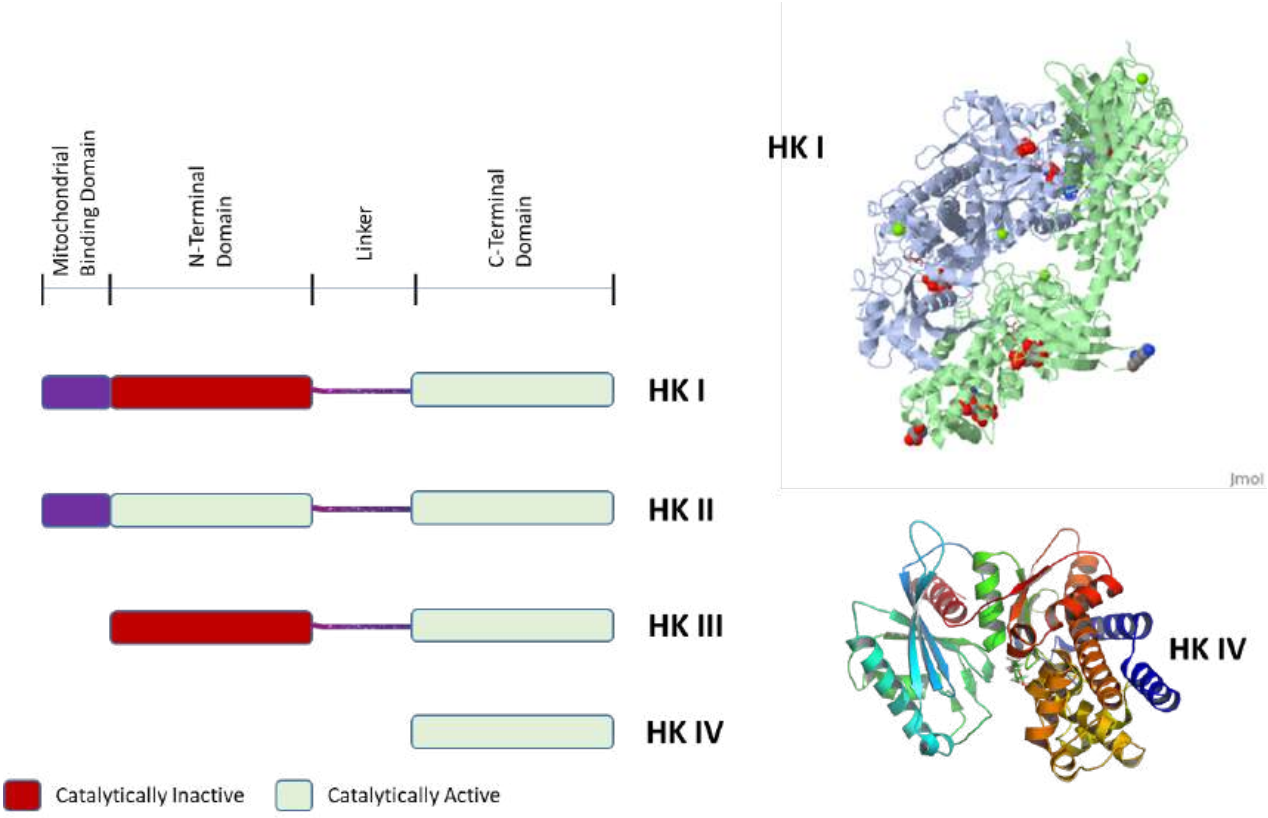


Figure 15.4.4: Protein Comparison of the Four Major Hexokinases in Vertebrates. The left-hand panel shows the linear protein domains of the hexokinases I - IV. The upper right diagram shows a dimer of HKI complexed with an ATP analog, glucose, glucose 6-phosphate, and Mg^{2+}. The lower panel depicts a ribbon diagram of HKIV. Figures modified from Proteopedia and JAG123

HKIV is also a good model for understanding enzyme conformation change during the reaction. HKs change shape by induced fit upon substrate binding. HKIV has a large induced fit motion that closes over the substrates ATP and xylose, as shown in Figure 15.4.5. The binding sites are shown in blue, substrates in black, and the Mg^{2+} cofactor in yellow (PDBs:2E2N,2E2Q).

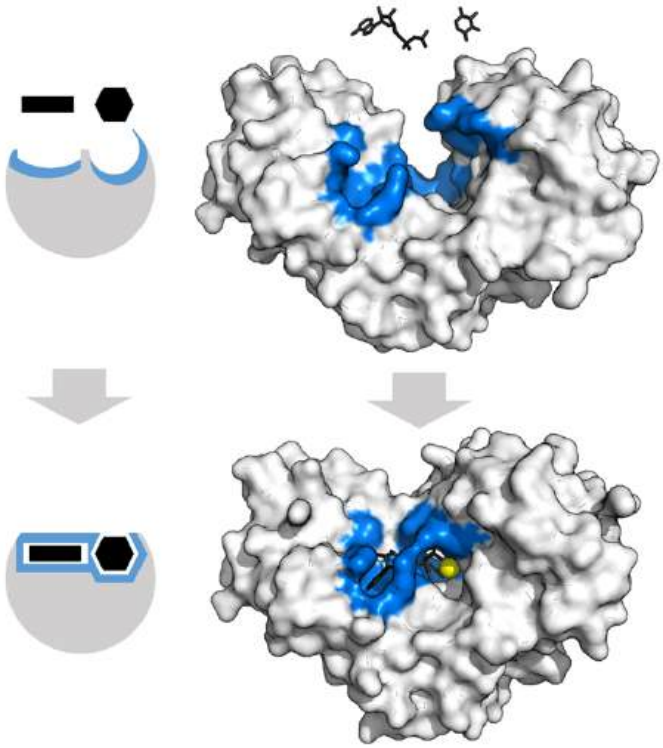


Figure 15.4.5: HKIV Conformational Change During Substrate Binding. The upper diagram shows the enzyme before substrate binding whereas the lower diagram shows the conformational change upon binding with xylose, ATP, and the Mg^{2+} cofactor. Mg^{2+} is shown in yellow and the substrate binding pocket in the enzyme in blue. Xylose and ATP are shown in black. Figure from Thomas Shafee

In summary, the isozyme expression patterns of HKs differentially regulate the fate of glucose within those tissues. Within brain tissue and red blood cells where only HKI is present, glucose is predominantly used in the glycolytic pathway for energy production. In muscle tissue, the presence of HKII allows for increased use of glucose for the formation of glycogen. HKIV expression in the pancreas and liver allows for the homeostatic regulation of blood glucose levels and stockpiling of glucose in the form of glycogen.

15.4.2: PHOSPHOFRUCTOKINASE-1 REGULATION

Recall that phosphofructokinase-1 (PFK1) mediates the third step in the glycolytic pathway with the conversion of fructose 6-phosphate to fructose 1,6-bisphosphate, as shown in Figure 15.4.6. The PFK1 reaction is the first irreversible reaction of glycolysis. It also represents the committed step within the pathway. The phosphorylation of fructose-6-phosphate (F6P) to fructose-1,6-bisphosphate (F1,6BP) commits the F1,6BP to continue through the glycolytic pathway. It cannot be utilized for any other purpose at that point. F6P, on the other hand, could be converted back into glucose-6-phosphate and used for many different purposes (ie glycogen synthesis, nucleotide synthesis, or hexosamine synthesis). Because of the committed nature of this step, PFK1 is one of the most important control points in the glycolytic pathway.

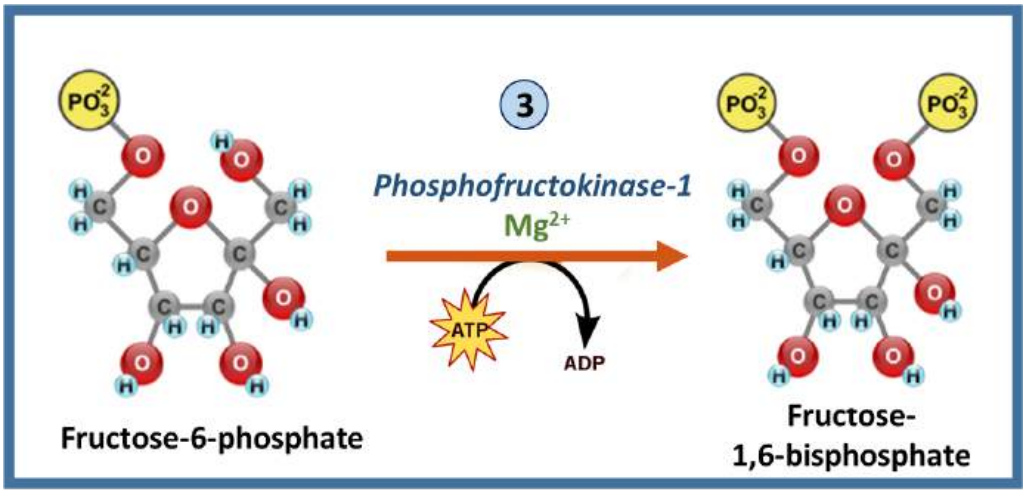


Figure 15.4.6: Phosphofructokinase-1 Activity During Glycolysis. Image modified from YassineMrabet

The PFK1 enzyme is composed of a tetramer that can contain different combinations of three types of subunits: muscle (M), liver (L), and platelet (P). The composition of the PFK1 tetramer differs according to the tissue type it is present in. For example, mature muscle expresses only the M isozyme, therefore, the muscle PFK1 is composed solely of homotetramers of M4. The liver and kidneys express predominantly the L isoform. In erythrocytes, both M and L subunits randomly tetramerize to form M4, L4, and the three hybrid forms of the enzyme (ML3, M2L2, M3L). As a result, the kinetic and regulatory properties of the various isoenzymes pools are dependent on subunit composition. Tissue-specific changes in PFK1 activity and isoenzymes content contribute significantly to the diversities of glycolytic and gluconeogenic rates which have been observed for different tissues

PFK1 is an allosteric enzyme and has a structure similar to that of hemoglobin in so far as it is a dimer of a dimer, as shown in Figure 15.4.7. One half of each dimer contains the ATP binding site, whereas the other half the substrate (fructose-6-phosphate or (F6P)) binding site, as well as a separate allosteric binding site, that can bind with ADP or AMP. All isoforms of PFK1 are activated by the allosteric binding of ADP or AMP. This indicates a low energy state within the cell and the need for glycolysis and energy generation. Allosteric inhibitors include high levels of ATP and Citrate. Note that ATP is a substrate of this enzyme and has the normal substrate binding site. When there is enough ATP present that it can also bind allosterically to the enzyme, it will act as an inhibitor. Citrate, the first molecule in the Kreb's Cycle (Citric Acid Cycle), can also act as an allosteric inhibitor of PFK-1. High levels of citrate that no more pyruvate is needed to generate ATP through oxidative phosphorylation. We are also going to see that Fructose 2,6-bisphosphate is predominantly a regulator of PFK1 in Liver Cells, where it serves as an activator.

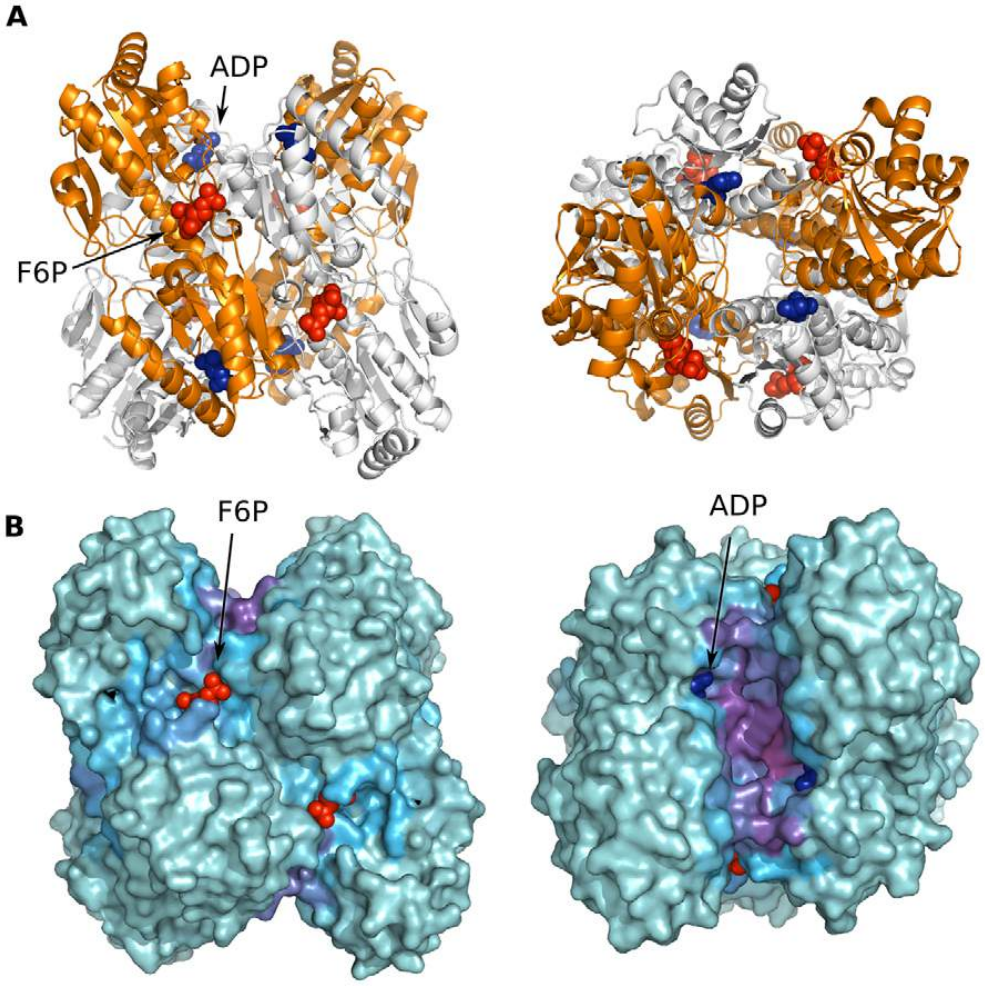


Figure 15.4.7: Structure of Phosphofructokinase-1. Lefthand diagrams show the ribbon and space-filling models from a side view whereas the right-hand diagrams has been rotated 90 degrees to show the top-down view. The ATP binding sites are shown in dark blue and the fructose 6-phosphate binding sites are shown in red. A separate allosteric binding site for ADP is also shown. Image modified from Mitternacht, S., and Berezovsky, I.N. (2011)

Before we discuss the formation and use of Fructose 2,6-bisphosphate and its role in the regulation of PFK1, let's review opposing glucose utilization/production pathways within the liver, as shown in Figure 15.4.8. Two opposing pathways within the liver are glycolysis (the breakdown of glucose) and gluconeogenesis (the formation of glucose). It would be unproductive to have both of these pathways operating at the same time. Thus, if one pathway is needed, then the other one needs to be turned off. We previously saw this same type of control for

glycogenesis (the formation of glycogen) and glycogenolysis (the breakdown of glycogen). When one of these pathways is upregulated, the other needs to be downregulated.

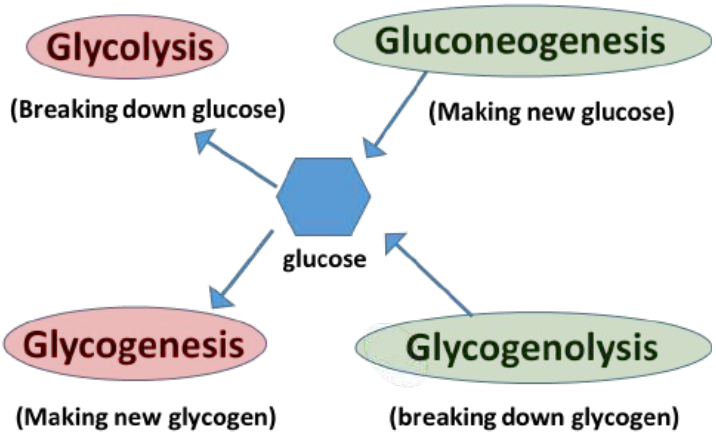


Figure 15.4.8: Opposing Pathways of Glucose in the Liver. Image modified from Yan, A., et al (2016) In J Biol Sci 12(12):1544-1554 and Servier Medical Art

In the resting state of the liver, when blood glucose levels are in the homeostatic range, gluconeogenesis will be shut off, as it is very expensive to make glucose and the liver will only invest in making glucose if blood glucose levels fall critically low (ie extreme exercise or long term fasting). In the resting state, the liver will use the glycolytic pathways to supply normal levels of ATP energy for the maintenance of housekeeping processes. Glycogenesis and glycogenolysis will be in balance or equilibrium as needed to augment the supply of glucose entering the system from the bloodstream.

If glucose in the blood drops below homeostatic levels, the pancreas will release glucagon and begin this hormone-signaling pathway that causes the liver to release glucose into the bloodstream. Thus, glucagon signaling leads to the downregulation of glycolysis and glycogenesis, so it can shunt glucose pools to the bloodstream. It also leads to an increase in glycogenolysis or the breakdown of glycogen. During this time, liver cells are predominantly generating ATP from lipids, rather than carbohydrates. Thus, glycolysis can be inhibited to promote the release of glucose into the bloodstream.

If cellular demand for glucose is high, liver cells will also turn on the gluconeogenesis pathway and make glucose new from non-carbohydrate precursors (which it then exports into the bloodstream for delivery to tissues). This is an energy-intensive pathway. The major site of gluconeogenesis is the liver, with a small amount also taking place in the kidney. Little gluconeogenesis takes place in the brain, skeletal muscle, or heart muscle. It just does not make energetic sense to do that! It costs more energy for cells to make glucose than they can get from breaking it down in oxidative phosphorylation. This cost can be dealt with by doing it in the liver and then releasing it into the bloodstream to fuel activities in the brain, heart, and skeletal muscles.

To do this, glucagon stimulates that lovely signaling pathway that you are all now familiar with (reviewed in Figure 15.4.9)! Within the liver, it activates the G-protein coupled receptor which in turn activates the downstream G-protein. Adenylate Cyclase is activated and produces the second messenger, cAMP. cAMP binds with the CREB protein and activates the transcription of proteins involved in gluconeogenesis. The cAMP also binds with Protein Kinase A and upregulates the activity of glycogen phosphorylase resulting in the breakdown of glycogen. It also down-regulates the activity of glycogen synthase to inhibit glycogenesis. In this section, we will also see how PKA activation will down-regulate the activity of the glycolytic pathway as well.

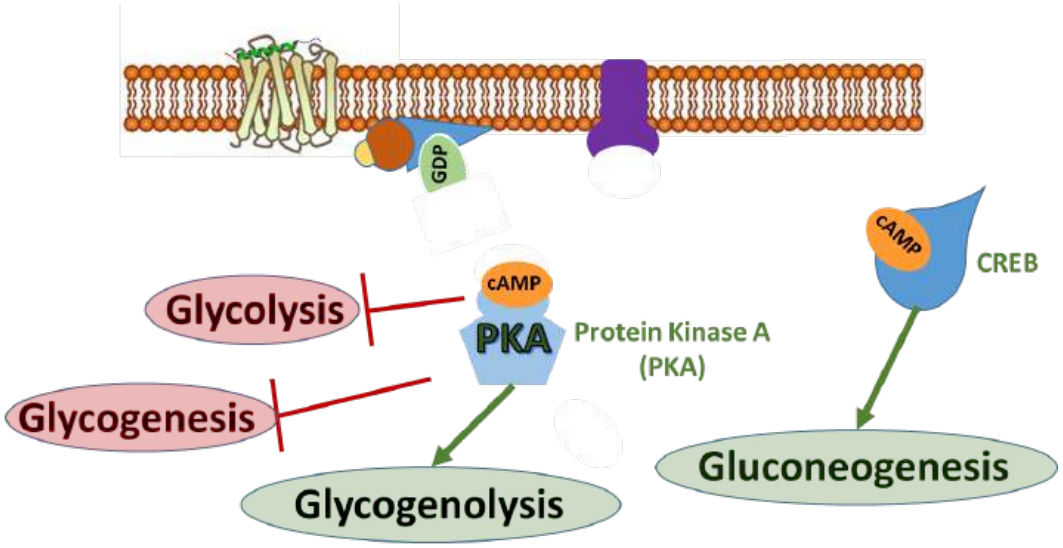


Figure 15.4.9: Overview of Glucagon Signaling Pathways within the Liver. Image modified from Yan, A., et al (2016) In J Biol Sci 12(12):1544-1554 and Servier Medical Art

Regulation of the glycolytic pathway in response to this signaling occurs through the regulation of the PFK-2/FBPase-2 Enzyme. The activity of this enzyme is controlled through the PKA signaling cascade. This enzyme is responsible for phosphorylating fructose 6-phosphate to the fructose 2,6-bisphosphate form. Note that this bisphosphate form of fructose is ***DIFFERENT*** than the bisphosphate form utilized in the glycolytic pathway. The glycolytic pathway requires fructose 1,6-bisphosphate that is formed from the PFK1 enzyme (or Step 3 of the glycolytic pathway). The PFK-2/FBPase-2 is a separate enzyme altogether and not involved directly in the glycolytic pathway. However, we noted previously that fructose 2,6-bisphosphate can serve as an allosteric activator of the PFK1 enzyme.

The PFK-2/FBPase-2 Enzyme is a dual-purpose enzyme, as shown in Figure 15.4.10. Half of the protein has kinase activity and can phosphorylate fructose-6-phosphate to fructose-2,6-bisphosphate. The other half of the enzyme contains a phosphatase that can cleave off the phosphate group from the 2-position and restore fructose-6-phosphate. Note that both activities are not functional at the same time! The enzyme activity is determined by the phosphorylation state. When the protein is dephosphorylated, the PFK-2 enzyme is active, leading to the production of fructose-2,6-bisphosphate. This molecule can then bind with PFK-1 in the glycolytic pathway and increase its activity. If the protein is phosphorylated by PKA during glucagon signaling, the FBPase is activated and the kinase activity is inhibited. This leads to the dephosphorylation of fructose at the 2-position and the release of fructose-6-phosphate.

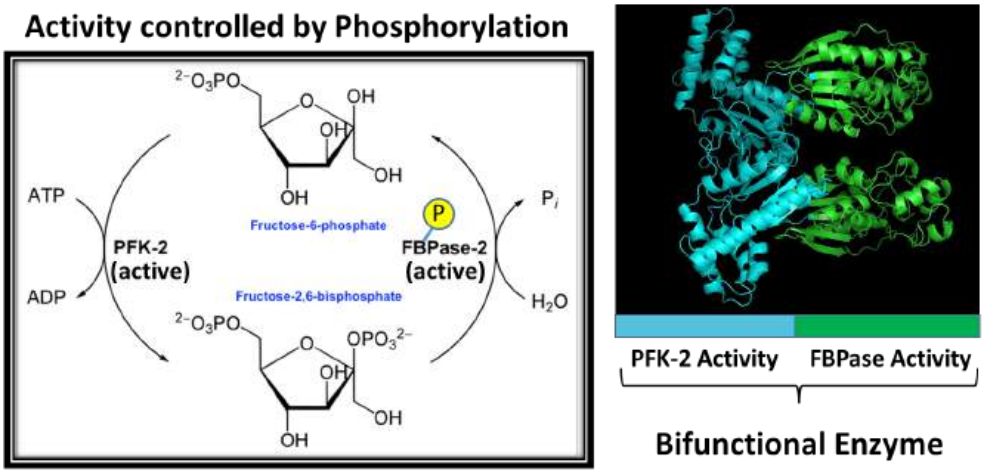


Figure 15.4.10: Biological Activity of the PFK-2/FBPase-2 Enzyme. The PFK-2/FBPase-2 is a dual functional enzyme that is capable of phosphorylating the 2'-position of fructose 6-phosphate using the kinase activity, whereas the FBPase activity can be used to remove the phosphoryl group from position 2' of fructose 2,6-bisphosphate, yielding fructose 6-phosphate. The PFK-2 component is active when the protein is in the dephosphorylated state while the FBPase remains inactive. However, following phosphorylation, the FBPase component becomes active and the kinase domain is inhibited. Image modified from Kedrosolan and Hyunsuky

As shown in Figure 15.5.10, the PFK-2/FBPase-2 Enzyme is responsible for phosphorylating fructose 6-phosphate to the fructose 2,6 bisphosphate form. Note again that this bisphosphate form of fructose is ***DIFFERENT*** than the bisphosphate form utilized in the glycolytic pathway. The glycolytic pathway requires fructose-1,6-bisphosphate that is formed from the PFK1 enzyme (or Step 3 of the glycolytic

pathway). If there is a lot of fructose-6-phosphate around (ie you just drank high fructose corn syrup in your sugary energy drink) it can support both the formation of fructose 1,6-bisphosphate by PFK1 and the production of fructose 2,6-bisphosphate by PFK-2. Fructose 2,6-bisphosphate will then bind with PFK1 and increase its activity converting fructose 6-phosphate into fructose-1,6-bisphosphate, as shown in Figure 15.4.11.

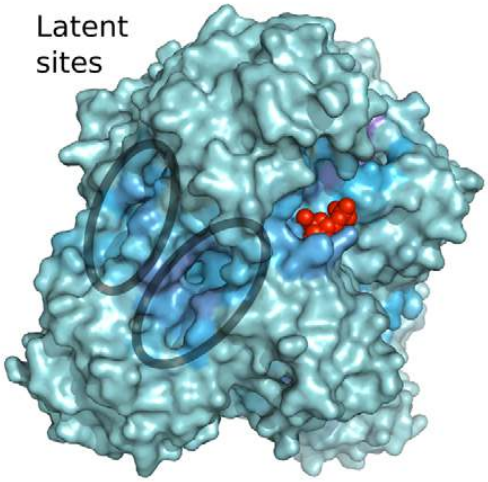


Figure 15.4.11: Fructose 2,6-bisphosphate is a positive allosteric effector of the PFK1 Enzyme. The upper diagram shows a space filling model of PFK1 with the fructose 2,6-bisphosphate binding sites indicated. Mitternacht, S., and Berezovsky, I.N. (2011), ibid

However, during glucagon signaling, you need to shut down this fast-track upregulation of PFK1 by fructose-2,6-bisphosphate and turn down the glycolytic pathway. To do this, Protein Kinase A will phosphorylate the PFK-2/FBPase-2 enzyme and alter its activity. The kinase activity is inhibited and the phosphorylase activity is turned on. Figure 15.4.12 provides a summary of this pathway control. In the presence of glucagon, PKA will phosphorylate the PFK-2/FBPase-2 enzyme causing the kinase activity to be switched off and the phosphatase activity to be switched on. Dephosphorylation of fructose-2,6-bisphosphate recovers fructose-6-phosphate (F6P). F6P can then go through the reverse isomerase reaction and recover glucose-6-phosphate (G6P). G6P will then be transported to the rER where it will be dephosphorylated and then free glucose can be released back into the blood.

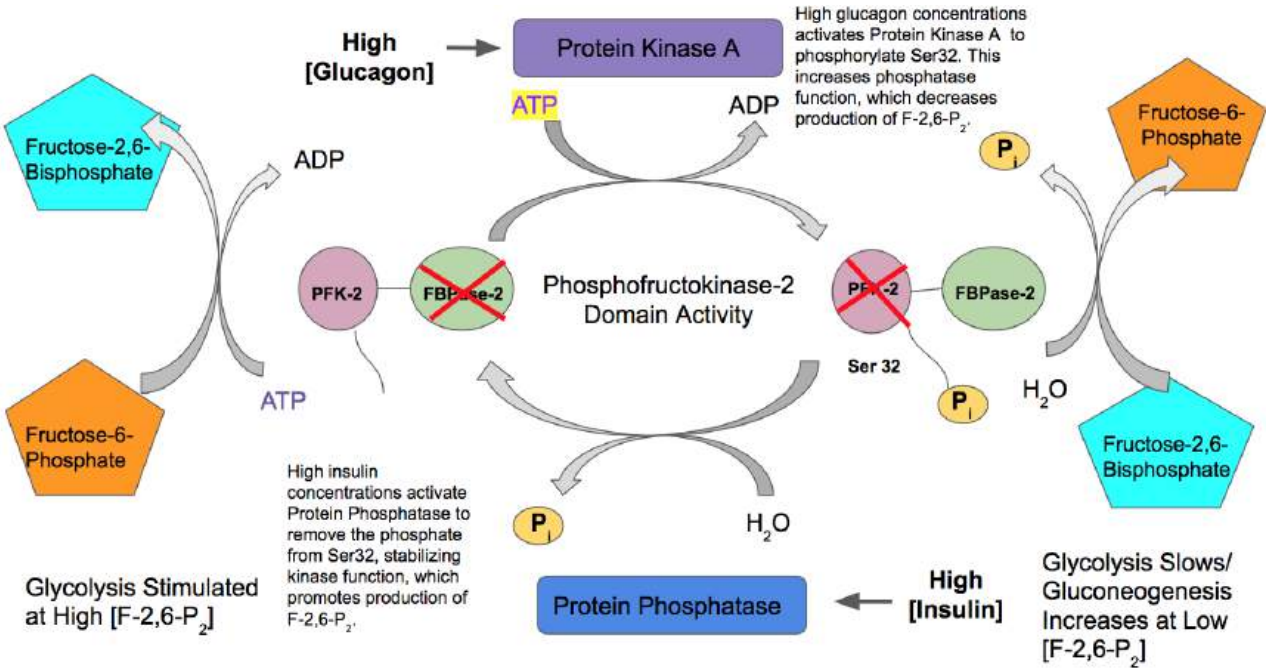


Figure 15.4.12: Allosteric Regulation of PFK1 in the Liver in Response to Glucagon Signaling. Image from Kedrosolan

The opposite holds for insulin signaling. High insulin concentrations result in the activation of protein phosphatase and the dephosphorylation of the PFK-2/FBPase-2 enzyme. In the dephosphorylated state, PFK-2 activity is high and the FBPase-2 activity is low, which will stimulate PFK1 and the glycolytic pathway.

15.4.3: PYRUVATE KINASE

Recall that pyruvate kinase mediates the final reaction during glycolysis resulting in the production of pyruvate and ATP as shown in Figure 15.4.13. Similar to PFK1 this enzyme is also a key regulatory component within the pathway.

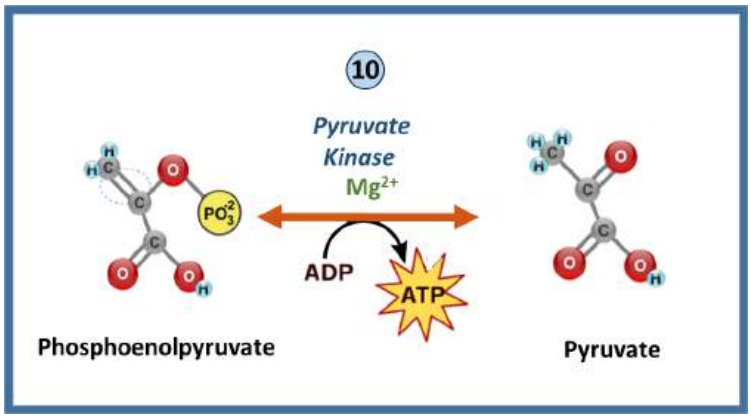


Figure 15.4.13: Review of the Pyruvate Kinase Reaction. Image modified from YassineMrabet

The pyruvate kinase enzyme exists as a tetramer, that is built from the combination of different isozymes expressed in different tissues. There are three major isozymes of pyruvate kinase, the L form that is predominantly found in the liver, the R form that is predominantly found in erythrocytes, and the M1 form in muscle and brain, and the M2 form that is expressed in fetal tissue and at some level in most adult tissues. The L and R forms are splice variants that arise from the same gene locus, and the M1 and M2 forms are also splice variants that arise from the same gene locus.

We will focus on some of the general regulatory mechanisms common to most of the isozymes of pyruvate kinase, starting with the activator, fructose 1,6-bisphosphate (FBP). Because FBP is an earlier product within the same metabolic cascade, the activation of pyruvate kinase enzymes by FBP is known as ***feedforward stimulation***. All of the isozymes, except for the M1 form are stimulated by the binding of FBP to the enzyme. Similarly, all of the pyruvate kinase isozymes are inhibited by the product of the reaction, ATP (or high energy load), and high levels of alanine. Alanine can be converted to pyruvate in one enzymatic step. Thus, pyruvate serves as a metabolic intermediate in the formation of alanine. If high levels of alanine are present, this indicates that there is a high energy load within the cell (ie that the cell is full of building blocks to make new macromolecules and is not in the need of more energy). Thus, high levels of alanine serve as a negative regulator of the pyruvate kinase family of enzymes.

The liver isozyme of pyruvate kinase is also regulated through protein phosphorylation, as shown in Figure 15.4.14. Similar to the PFK-2 activity of the PFK-2/FBPase-2 enzyme, the liver isozyme of Pyruvate Kinase is also downregulated during glucagon signaling. Protein kinase A phosphorylates Pyruvate Kinase inhibiting its activity and preventing the conversion of phosphoenolpyruvate to pyruvate. Dual regulation of the glycolytic pathway during glucagon signaling helps to ensure that glucose resources will be diverted away from cellular use by the liver and released into the blood stream to restore homeostatic blood glucose levels.

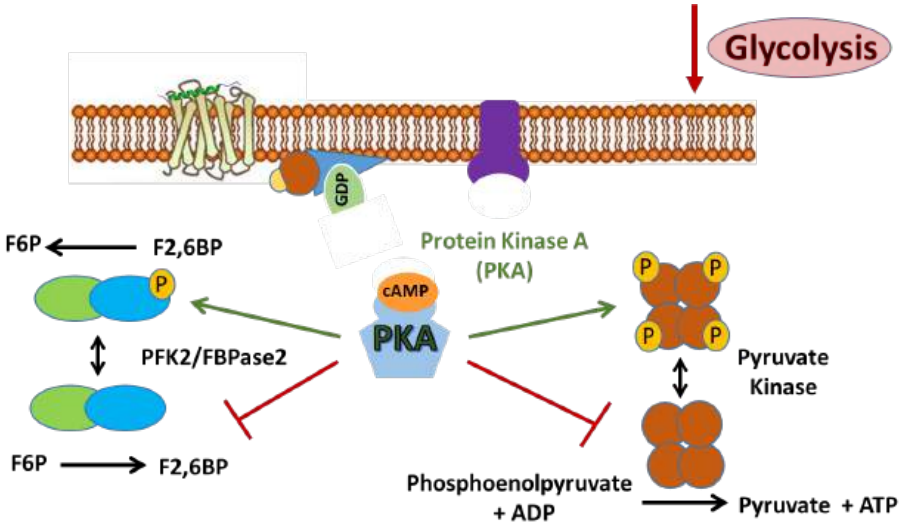


Figure 15.4.14: Glucagon Signaling in Liver Cells Down Regulates PFK2 Activity and Pyruvate Kinase Activity. Image modified from Yan, A., et al (2016)

15.4.4: FRUCTOSE REGULATORY BYPASS

Other sugars from the diet can also enter into the glycolytic pathway, as shown in Figure 15.4.15. Galactose is converted in a four-step process to Glucose-6-phosphate and mannose can be converted to fructose-6-phosphate. Within most of the body's tissues, fructose can also be converted into fructose-6-phosphate by hexokinase. However, in the liver and kidneys, there is an alternative route that fructose from the diet can take to enter into the glycolytic pathway. This pathway is concerning because it bypasses two of the major regulatory steps of the glycolytic pathway, the hexokinase step and the PFK1 step. Within the liver and kidneys, fructose can also be converted into fructose-1-phosphate by the enzyme fructokinase. The other isozyme of Aldolase, Aldolase B, can cleave the fructose-1-phosphate into 2 three carbon units, dihydroxyacetone phosphate, and glyceraldehyde. Dihydroxyacetone phosphate can be converted to glyceraldehyde 3-phosphate by Triose Isomerase and then continue into the glycolytic cascade. Glyceraldehyde can be phosphorylated to Glyceraldehyde 3-phosphate by Triokinase. This is an unregulated system that can flood the Kreb Cycle with high levels of pyruvate if high levels of fructose enter the cell (i.e. from high fructose corn syrup, sucrose, and other sweeteners common to the Westernized diet). The excess pyruvate can then be shunted into fatty acid biosynthesis for long-term storage in the form of triglycerides. If the pathway is overutilized by consuming too much sucrose and high fructose corn syrup, this can lead to the development of Hypertriglyceridemia (or the heightened increase of body fat). With this, we will end our discussions of glycolysis. In the next section, we will look at the complementary and opposite pathway, gluconeogenesis.

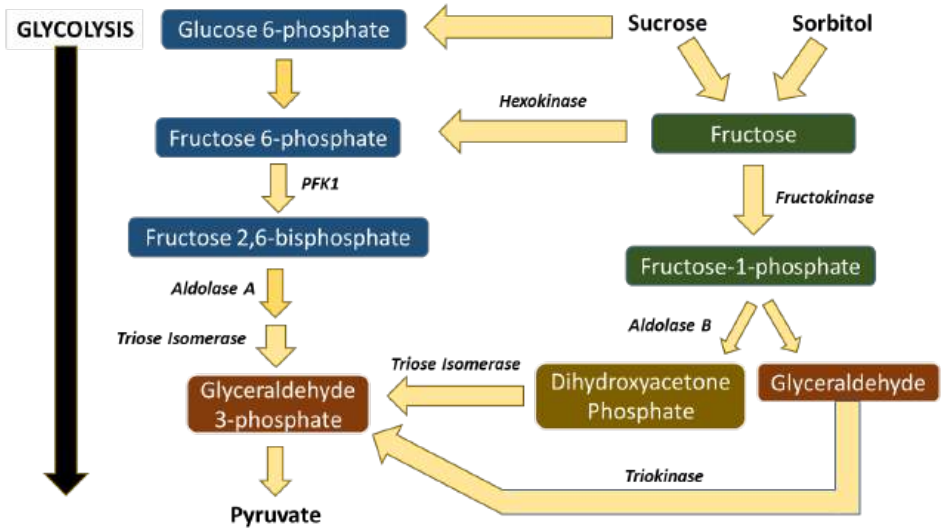


Figure 15.4.15: Alternate Sugar Sources for Glycolysis

15.5: REGULATION OF GLUCONEOGENESIS

Within the regulation of the gluconeogenic pathway, three of the major enzymatic steps are regulated. The first two are the pyruvate carboxykinase enzyme and the phosphoenolpyruvate carboxykinase (PEPCK). Recall that these two enzymes are required to convert pyruvate back into phosphoenolpyruvate via an oxaloacetate intermediate as shown in Figure 15.5.1. The third enzyme regulated in this pathway is Fructose 1,6-Bisphosphatase which converts fructose 1,6-bisphosphate into fructose 6-phosphate. We will explore the regulation of these three enzymes in more detail.

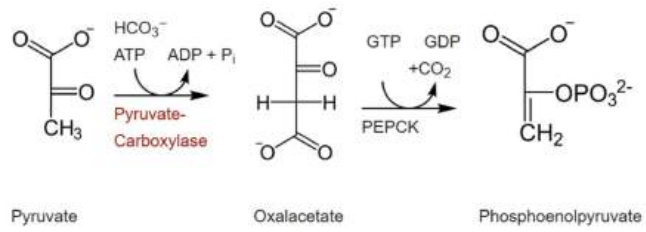


Figure 15.5.1: Conversion of Pyruvate to Phosphoenolpyruvate during Gluconeogenesis. Image modified from Principles of Biochemistry (2019) Wikibooks

15.5.1: PYRUVATE CARBOXYKINASE

Pyruvatecarboxykinase is one of the primary regulation points. It is primarily regulated by two allosteric effectors, Acetyl-CoA and ADP, as shown in Figure 15.5.2. When pyruvate enters into the Kreb Cycle, it is first converted to Acetyl-CoA. If abundant pyruvate is present, an ample supply of acetyl-CoA will also be available, indicating a high energy load for the cell. Acetyl-CoA, can bind with pyruvate carboxylase and act as an activator of the protein, stimulating the production of oxaloacetate. ADP, on the other hand, is a low-energy indicator and an inhibitor of the enzyme. In the next section, we will discover how oxaloacetate moves into the cytoplasm.

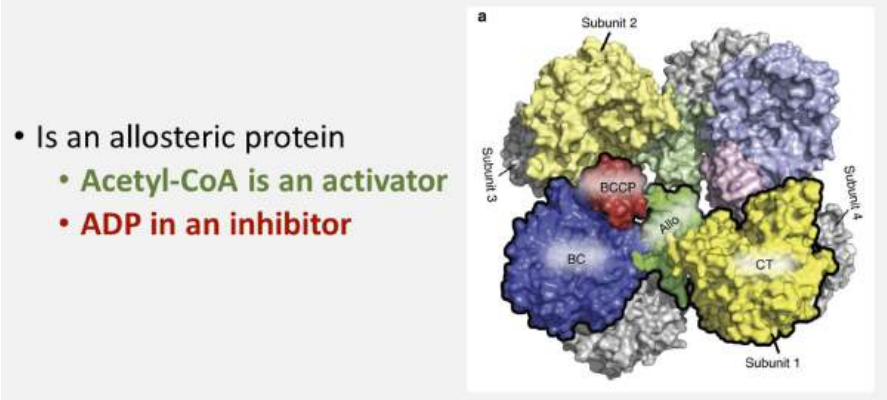


Figure 15.5.2: Allosteric Regulation of Pyruvate Carboxykinase. Figure modified from Liu, Y., et al (2018) Nat Commun 9:1384

15.5.2: PHOSPHOENOLPYRUVATE CARBOXYKINASE

Cytoplasmic PEPCK is largely regulated at the transcriptional level. Increases in gene expression are seen in response to elevated cAMP levels, increased glucocorticoids, and increased thyroid hormone levels, as shown in Figure 15.5.3. The activated CREB transcription factor plays a role in this response. Alternatively, decreased gene expression is caused by insulin signaling. ADP also acts as an allosteric effector of the protein, causing it to have lower activity. This indicates that when energy is low, the cell cannot afford to use its reserves to remake glucose and inhibits the pathway.

Figure 15.5.3: Regulation of Phosphoenolpyruvate Carboxykinase at the Transcriptional and Allosteric Levels. Image from ProteinBoxBot

15.5.3: FRUCTOSE 1,6-BISPHOSPHATASE

Fructose 1,6-bisphosphatase is both competitively and allosterically regulated, as shown in Figure 15.5.4. Fructose 2,6-bisphosphate serves as a competitive inhibitor of the enzyme reducing the overall activity of the enzyme for fructose 1,6-bisphosphate. Competitive inhibitors bind within the active site and compete for binding with the regular substrate. Thus, they lower the overall Km of the reaction and make the enzyme less effective at lower substrate concentrations. However, the Vmax of the enzyme is not affected during the process.

In addition to competitive inhibition, low energy load (AMP and ADP) also inhibits the enzyme. ADP and AMP will bind allosterically with the enzyme and inhibit its activity.

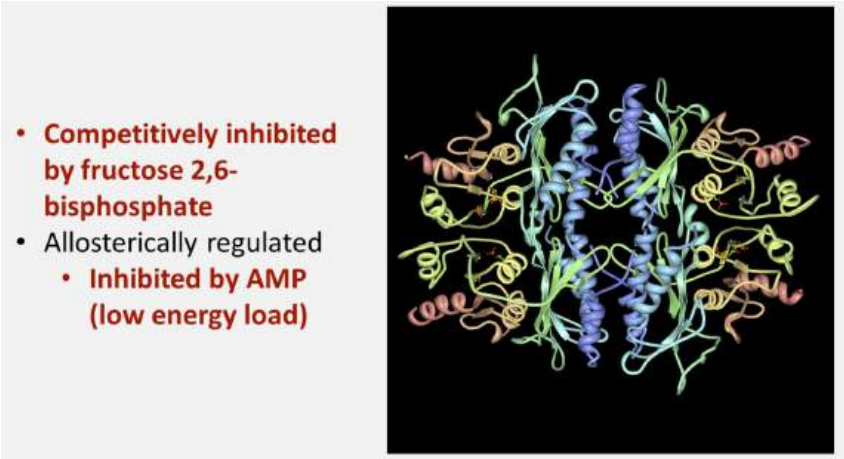


Figure 15.5.14 Regulation of Fructose 1,6-Bisphosphatase by Competitive Inhibition and Allosteric Effectors. Image from Jslipscomb

CHAPTER OVERVIEW

16: THE CITRIC ACID CYCLE

The citric acid cycle – also known as the TCA cycle or the Krebs cycle – is a series of chemical reactions to release stored energy through the oxidation of acetyl-CoA derived from carbohydrates, fats, and proteins.

16.1: Production of Acetyl-CoA (Activated Acetate)

16.2: Reactions of the Citric Acid Cycle

16.3: Regulation of the Citric Acid Cycle

16.4: Variants of the Citric Acid Cycle

16.1: PRODUCTION OF ACETYL-COA (ACTIVATED ACETATE)

16.1.1: INTRODUCTION

Let's do a short review of the metabolic processes for the extraction of energy from the oxidation of glucose (i.e. glycolysis). Glycolysis is a universal pathway used to anaerobically extract energy from glucose, a six-carbon. In this linear pathway, 2- 3 carbon molecules of pyruvate are formed as glucose is cleaved and converted to two molecules of glyceraldehyde-3-phosphate, form through an oxidation reaction using the oxidizing agent NAD^+. As glycolysis continues, NADH builds up. Using lactate dehydrogenase, pyruvate, the end product of glycolysis, can be converted to lactate, regenerating NAD^+ so the pathway can continue. The reactions are illustrated in Figure 16.1.1.

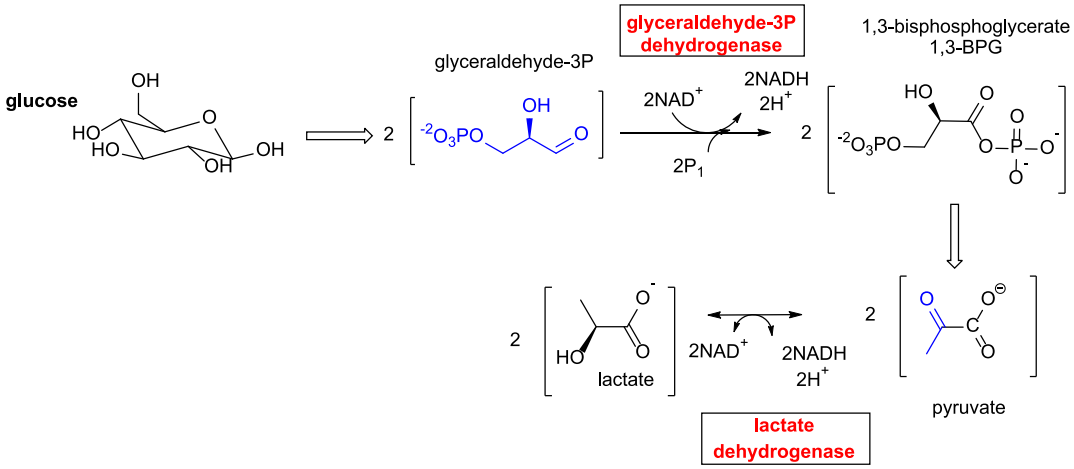


Figure 16.1.1: Anaerobic production of pyruvate and lactate

A careful glance at the structure of the 3-carbon pyruvate molecule shows that much more energy could be extracted from it, presumably through oxidative decarboxylation reactions, converting the carbons to 3 CO_2 molecules. A problem arises immediately when examining pyruvate. It is an α-ketoacid and there is no easy route to decarboxylate it as an electron "sink" is not available to receive the electrons and in the process stabilize the transition state and intermediate in the reaction. This stands in contrast to the decarboxylation of β-keto acids, which have a built-in electron "sink", an electronegative carbonyl carbon, to receive the electrons. This is illustrated in Figure 16.1.2.

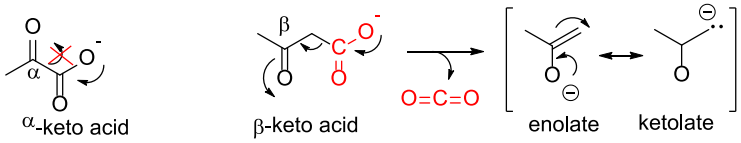


Figure 16.1.2: Comparison of the decarboxylation of α and β keto acids

To begin the process of complete oxidation of the remnants of glucose, pyruvate enters the mitochondria and starts the process of oxidative decarboxylations by interacting with the pyruvate dehydrogenase complex (PHC). This catalyzes a complicated reaction to attach an electron "sink" beta to the carboxylate, which is subsequently released as CO_2. The end products of the PHC oxidative decarboxylation reaction are the two-carbon acetyl-CoAs, NADH, and CO_2. The acetyl-CoA then enters a cyclic, non-linear pathway called the citric acid cycle, tricarboxylic acid (TCA) cycle, or Kreb's cycle, named after Hans Krebs who discovered it. We'll talk about that in section 16.2.

A glance reveals that we have taken glucose a small fraction along the way of oxidizing every carbon in it to CO_2 and H_2O. Complete oxidation happens under aerobic conditions when the glycolytic pathway is followed by the Kreb's cycle. Pyruvate formed in glycolysis enters the mitochondrial matrix where it gets oxidatively decarboxylated while reacting with a small thiol, Coenzyme A (CoASH) to form a 2C "activated acetate" acetyl group connected through a thioester link to CoASH, forming acetyl-CoA.

The third carbon from pyruvate is released as CO_2. The reaction is catalyzed by the enzyme **pyruvate dehydrogenase complex (PDC)**.

16.1.1.1: PYRUVATE DEHYDROGENASE MECHANISM

This enzyme complex is enormous. The E. Coli complex has a molecular weight of almost 4 million with at least 16 chains each of three different enzymes catalyzing part of the reaction. The components are pyruvate dehydrogenase (E1), dihydrolipoamide dehydrogenase (E2), and dihydrolipoamide dehydrogenase (E3). The molecular weight of the bovine complex is almost 8 million, and it has 22 E_1, 60 E_2, and 6 E_3 subunits. Nature often uses the same solution for identical problems. For example, many proteases have an active site nucleophilic

serine, which works with the assistance of histidine and aspartate to cleave peptide bonds. There are three α-ketoacid dehydrogenase complexes in many organisms. Each has a common E3 but specific E1 and E2 enzymes. Figure 16.1.3 shows an image of the structure so you can get an overview before we dive into the activity of each of the substrates.

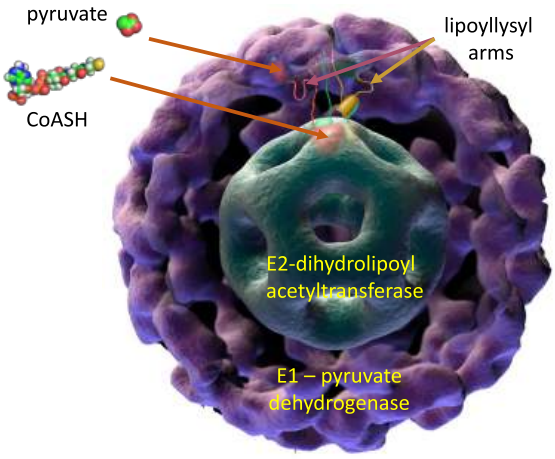


Figure 16.1.3: View of pyruvate dehydrogenase. https://electron.med.ubc.ca/2018/07/...dehydrogenase/

The E3 subunit is not readily seen in the image below. Why has nature produced such a monstrous enzyme complex to simply catalyze the oxidative decarboxylation of a small three-carbon molecule? We will explore that at the end of this section.

The complex also employs collectively 5 substrates/cofactors derived from vitamins.

- Thiamine in the form of thiamine pyrophosphate (TPP), which is covalently attached to E1
- lipoic acid, in the form of lipoamide, which is covalently attached to a lysine side chain in E2
- riboflavin the in the form of flavin adenine dinucleotide ($FAD/FADH_2$), which is bound very tightly (and not released) to E3
- pantothenic acid, incorporated into the structure of CoASH/Acetyl-CoA, a substrate/product pair for the reaction
- niacin, nicotinic acid, in the form of NAD^+/NADH, a substrate/product pair for the reaction

The structures for the five are shown in Figure 16.1.4, along with some additional descriptions that summarize some of the chemistry of these molecules.

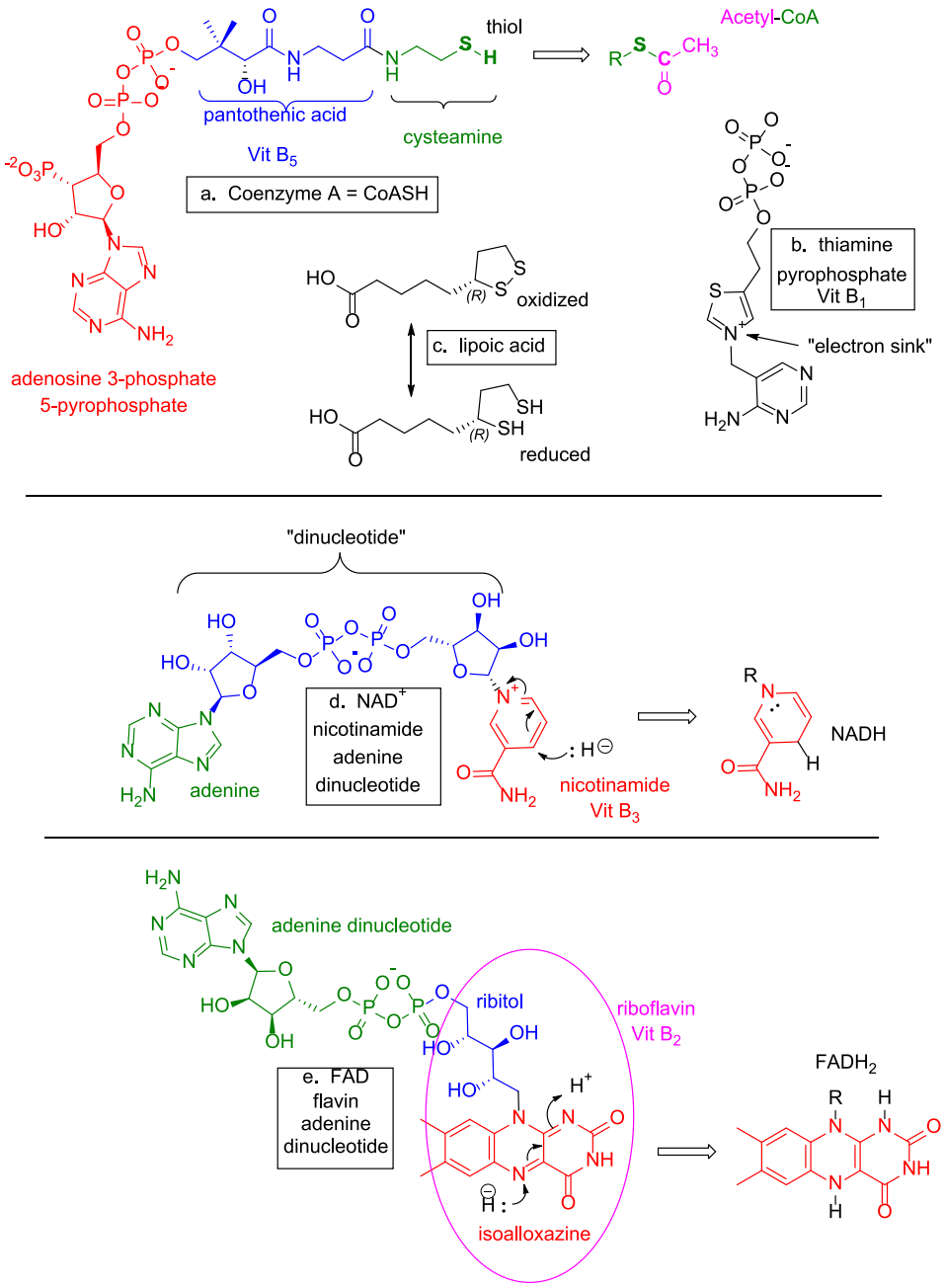


Figure 16.1.4: Structure of the cofactors in pyruvate dehydrogenase

Figure 16.1.5 shows a schematic of the overall reaction.

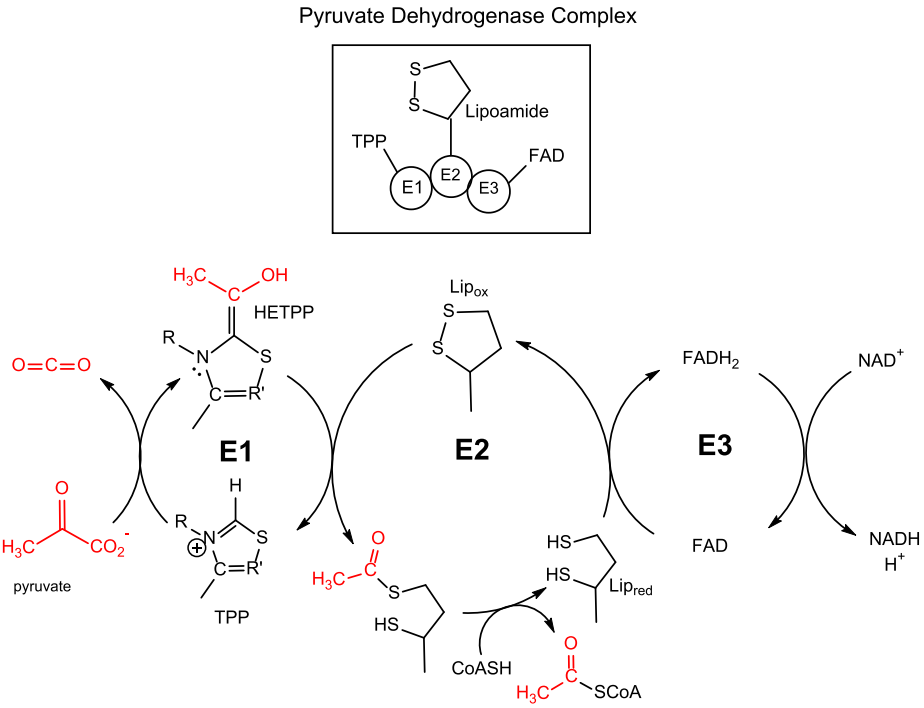


Figure 16.1.5: Overall reactions catalyzed by pyruvate dehydrogenase

The net reaction is

$$\text{pyruvate} + \text{CoASH} + \text{NAD}^+ \longrightarrow \text{Acetyl}-\text{CoA} + \text{CO}_2 + \text{NADH} + \text{H}^+$$

16.1.2: PART 1: OXIDATIVE DECARBOXYLATION - PYRUVATE DEHYDROGENASE (E1P)

So let it begin. We need to get rid of one carbon as CO_2 and transfer the other two carbons of pyruvate to CoASH to form acetyl-CoA, the thioester of CoASH. Thioesters are "high energy" with respect to their hydrolysis products as the thioester is destabilized compared to a normal carboxylic acid ester. (**Remember, there is no such thing as a "high energy" bond**). Since the sulfur atom is larger than the O in the C-S and C-O bond in their respective esters, the thioester as a reactant can not be stabilized well as the C-S single bond length is longer, as shown in Table 16.1.1 below.

bond	length (Angstroms)
C-O	1.43
C=O	1.21
C-S	1.82
C=S	1.56

Table 16.1.1: Bond lengths of carbon-oxygen and carbon-sulfur single and double bonds

This minimizes resonance stabilization compared to the carboxylic acid ester, as shown in the figure below. The products of hydrolysis of both a carboxylic and thiol ester are of comparable energy. Hence only the thioester is relatively destabilized compared to its hydrolysis product, with the $\Delta G^0_{\text{hydrolysis}}$ = -7.5 kcal/mol (-31 kJ/mol), the same as for the hydrolysis of a phosphoanhydride bond of ATP. Additionally, a resonance structure shows a positive charge on the carbonyl C and a negative on the oxygen allowing the carbonyl carbon to be more electrophilic. Another more sophisticated reason for the relative destabilization of the thiol ester is that the overlap between the carbonyl C p orbital is the larger S p orbital is less, hindering the delocalization of electrons needed to stabilize the thiol ester.

Figure 16.1.6 illustrates these points.

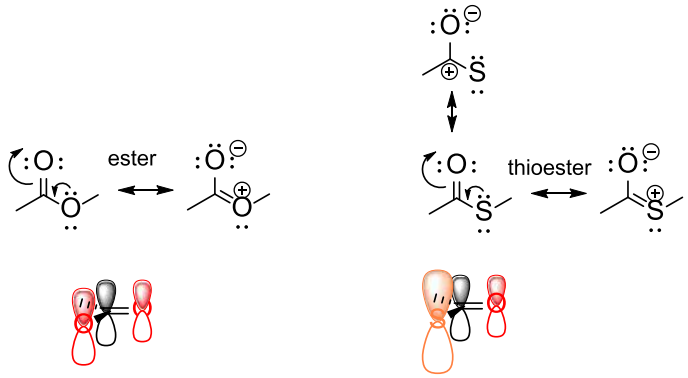


Figure 16.1.6: Comparison of resonance stabilization of carboxylic esters and thioesters

Now we can explore the mechanism of CO_2 release and acetyl-CoA production by E_1. The carbon atom directly between the N and S in the ring has a reduced pKa, so it can be deprotonated to form a carbanion. The negative charge can't be stabilized by resonance but it is adjacent to the positively-charge N, which stabilizes it. This zwitterion is called an **ylide**, which is a net neutral species with a positive charge (usually on a N, P, or S) and a negative charge (usually on a C) on adjacent atoms.

The carbanion on the ylid attacks the electrophilic C=O of pyruvate, forming a TPP intermediate with a wonderful electron sink (N^+) beta to the carboxyl carbonyl C. This is the essence of the entire reaction as this enables the decarboxylation event. The rest of the reactions catalyzed by E2 and E3 allow the release of the other 2 Cs of pyruvate as acetyl-CoA (E2) and the return of the enzyme to its original state (E2 and E3).

Figure 16.1.7 shows the reaction mechanism of E1.

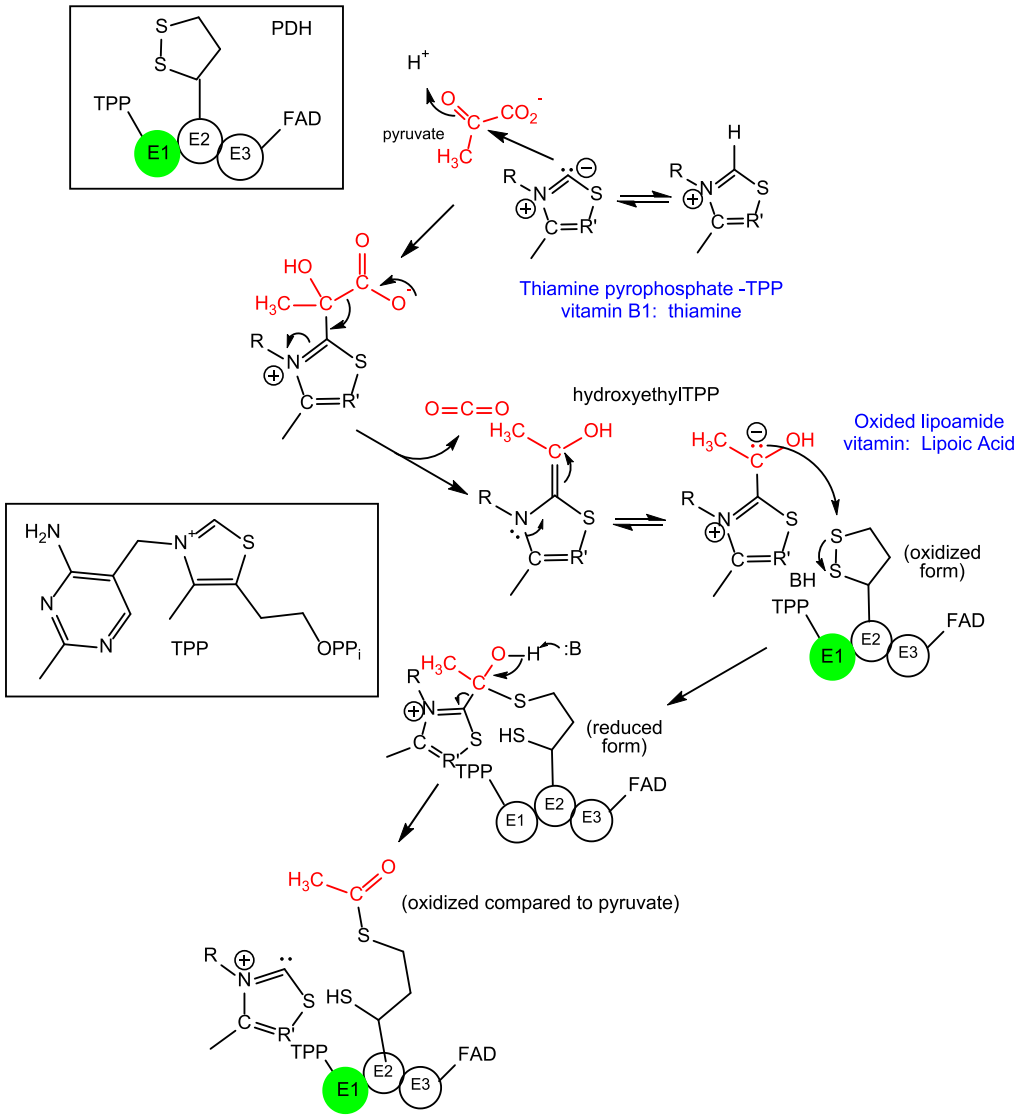
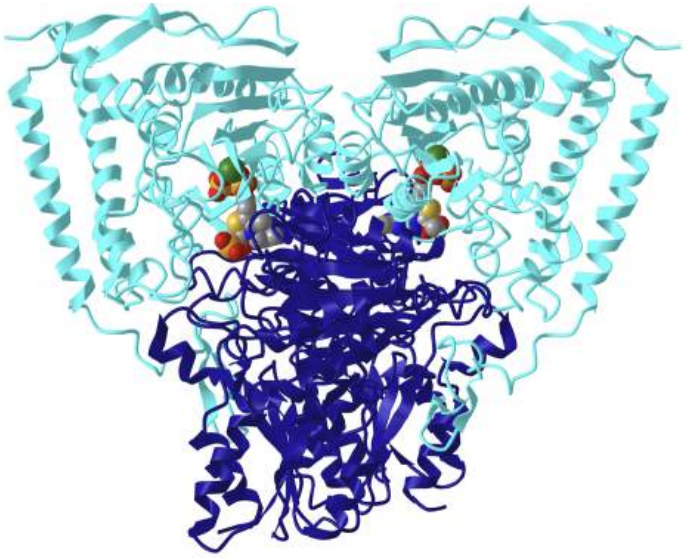


Figure 16.1.7: Reaction mechanism of E1 of pyruvate dehydrogenase

Note that the carbonyl C in pyruvate has two single bonds to two other carbon atoms, while in the final covalently attached form it has one bond to carbon and one to sulfur. Sulfur is under oxygen in the periodic table so by analogy, the replacement of one C-C bond with a C-S bond is an oxidation reaction, which requires an oxidizing agent. The covalently attached ring of the lipoamide with an S-S bond similar to that of a disulfide bond, an oxidized form of sulfur, is the oxidizing agent. On the formation of acetyl-lipoamide, the S-S bond is cleaved and a thioester is formed. Other sulfur is a free reduced thiol.

There are 22 E1 subunits in the bovine PDC. Here is an iCn3D model of one human pyruvate dehydrogenase E1 component complex that has TPnP (TDP acetyl phosphonate, a TPP analog, covalently attached (PDB ID: 6CFO). One E1 subunit is an $\alpha_2\beta_2$ heterodimer. The two alpha chains are shown in cyan while the beta chains are in dark blue. Orient the model to view along the C2 rotational symmetry axes shown.

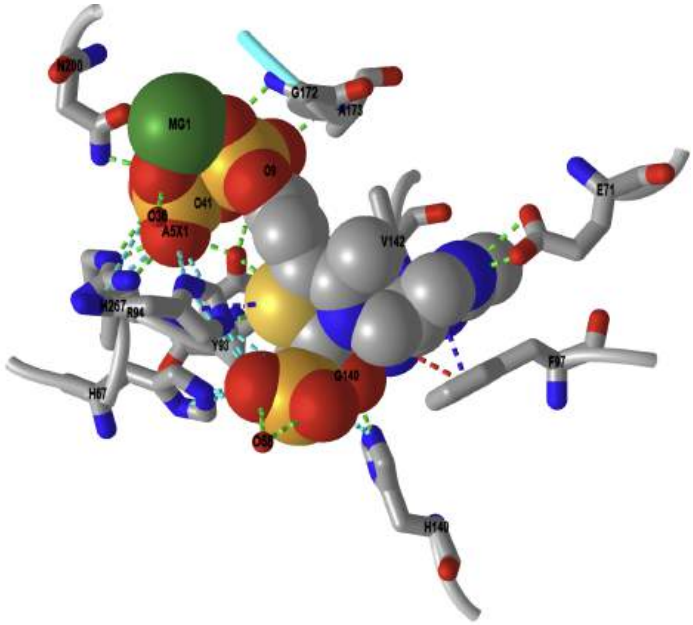
Figure 16.1.8 shows an interactive iCn3D model of the human pyruvate dehydrogenase E1 covalently bound to TDP acetyl phosphonate (TpnP), a TPP analog (6CFO)



NCBI iCn3D Figure 16.1.8: Human pyruvate dehydrogenase E1 covalently bound to TDP acetyl phosphonate (TpnP), a TPP analog (6CFO). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...7FJ2GE47tvS6g7

A heterotetramer containing 2 α (cyan) and 2 β chains (dark blue) is shown.

Figure 16.1.9 shows an interactive iCn3D model of the TPP analog covalently bound to E1 of pyruvate dehydrogenase (6CFO)



NCBI iCn3D Figure 16.1.9: TPP analog covalently bound to E1 of pyruvate dehydrogenase (6CFO) . (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...SNf5sC7SqmQkc8

The dotted lines show the interactions between the TPP analog, color-coded as shown in the legend below.

Hydrogen Bonds		3.8 A	Salt bridge/ion-ion		6 A	Contacts/interactions	4 A
Halogen Bonds		3.8 A	π-Cation		6 A	π-stacking	5.5 A

16.1.2.1:

16.1.3: PARTS 2 AND 3: FORMATION OF ACETYL-COA (E2) AND REGENERATION OF THE ACTIVE COMPLEX.

The next part of the reaction produces acetyl-CoA (E2), but after that, the enzyme is "dead" as it no longer has an oxidized form of lipoamide to serve as an oxidizing agent (which gets reduced) in another round of catalysis. To regenerate enzyme activity, the reduced

lipoamide, after the release of the attached acetyl group, must be reoxidized by another oxidizing agent. That oxidizing agent is FAD, which is covalently attached to E3, and is converted to $FADH_2$. It must be reoxidized back to FAD to restore activity to the enzyme complex. The final oxidizing agent used for that is solution-phase NAD^+, which is released by the enzyme as a product. So it's a bit complicated. Three oxidizing agents are used in the PDH, two of which are covalently attached to the enzyme (oxidized lipoamide on E2 and oxidized FAD on E3).

Figure 16.1.10 shows the transacetylation reaction and formation of reduced lipoamide

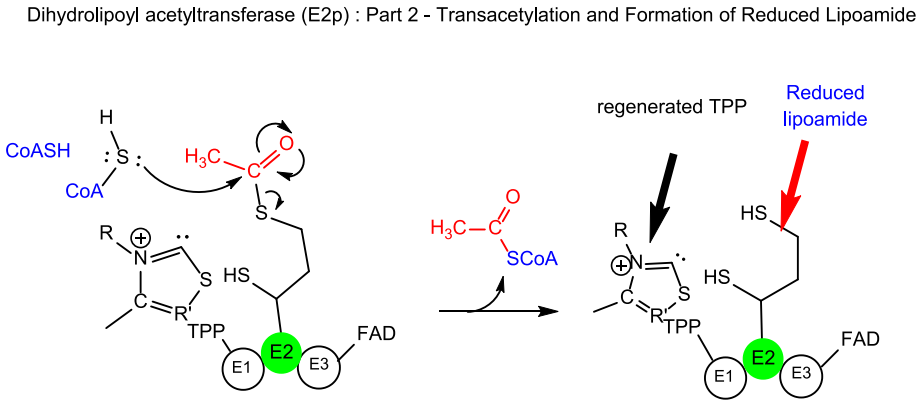


Figure 16.1.10: Transacetylation reaction and formation of reduced lipoamide by pyruvate dehydrogenase E2

The reaction of E3 follows to restore the fully catalytic enzyme, as shown in Figure 16.1.11.

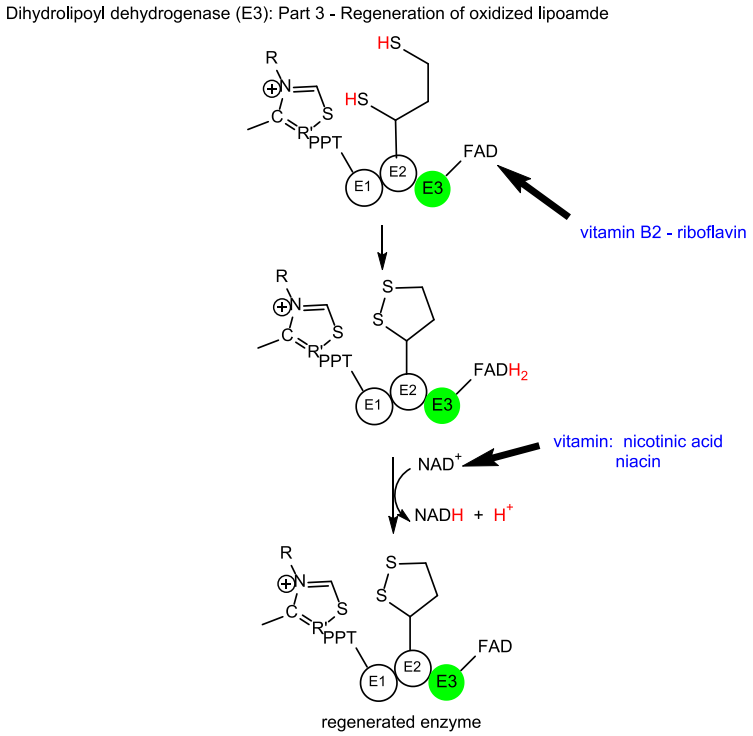


Figure 16.1.11: Regeneration of oxidized lipoamide by pyruvate dehydrogenase E3.

Let's look in greater detail at the structures of both E2 and E3.

16.1.4: E2: DIHYDROLIPOYL ACETYLTRANSFERASE -

In the mammalian complex, 60 E2 subunits arranges into a pentagonal dodecahedron. Most gram-negative bacteria E2 subunits arrange into a cubic of 24 monomers. Figure 16.1.12 shows a simple view of a pentagonal dodecahedron, which has 12 equivalent faces.

Figure 16.1.12: Pentagonal dodecahedron. 1. Rotating dodecahedron: https://commons.wikimedia.org/wiki/F...decahedron.gif. User Cyp on en.Wikipedia, CC BY-SA 3.0 <http://creativecommons.org/licenses/by-sa/3.0/>, via Wikimedia Commons

First let's consider one single E2 monomer. It has a longer disulfide redox domain followed by a smaller dimerization domain which allows the assembly of multiple subunits into the dodecahedron. In greater detail, the monomer has two lipoyl domains, a small domain that allows binding to E1 and a C terminal catalytic domain.

Figure 16.1.12 shows an interactive iCn3D model of E2 inner core 60-mer of human pyruvate dehydrogenase (pdb 6CT0). Symmetry axes are not shown

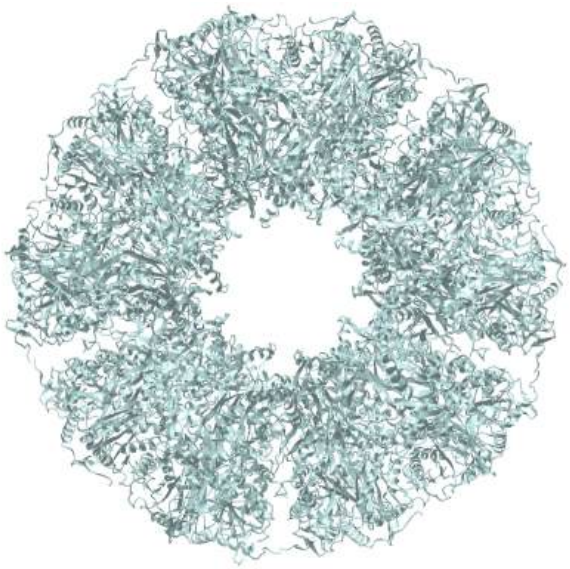


Figure 16.1.12: E2 inner core 60-mer of human pyruvate dehydrogenase (pdb 6CT0). (Copyright; author via source). Click the image for a popup or use this external: https://structure.ncbi.nlm.nih.gov/i...zs6pwfKNuJ2VS6

Each of the 60 subunits is shown in light cyan. To see the C symmetry axes:

- select the menu ☰
- Choose/Check in order: **Analysis, Symmetry, From PDB, 1(global), apply**
- When you see just a single monomeric chain choose **Cl**ear

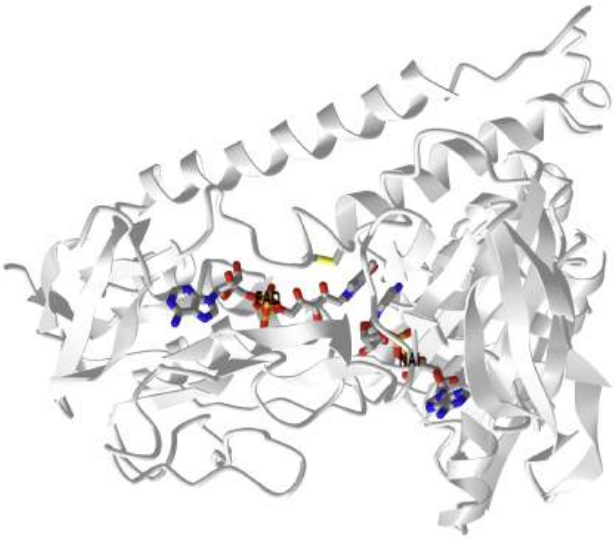
The symmetry axes will then appear.

16.1.5: E3: DIHYDROLIPOYL DEHYDROGENASE

The sole function of this subunit is reoxidation of the now reduced lipoamide with the free sulfhydryl to the cyclic disulfide form so the enzyme can engage in further catalysis. FAD covalently bound to the E3 subunit is the oxidizing agent. This is our first encounter with FAD. Similarly to NAD^+, this dinucleotide gains a hydride ($:H^-$) but also in contrast to NAD^+ also a proton to form $FADH_2$.

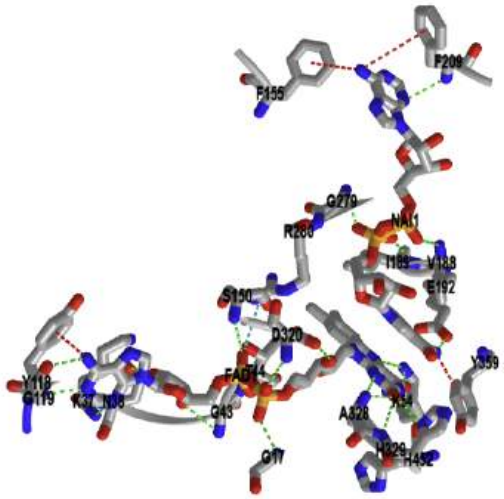
Another way that the $FAD/FADH_2$ differs from $NAD^+/NADH$ is that the $FAD/FADH_2$ or their mononucleotide analog ($FMN/FMNH_2$) pairs are either covalently attached (in about 10% of flavoproteins) or bound with such a low K_D (often in the nanomolar range) that they don't dissociate from the enzyme during catalysis. Hence after oxidizing a bound substrate, the reduced $FADH_2$ must be reoxidized by another oxidizing agent, often NAD^+ which can diffuse into the active site to do its job and then dissociate from the complex in the form of NADH, leaving the enzyme competent for another round of catalysis. (DOI: 10.1002/chem.201704622)

Figure 16.1.13 shows an interactive iCn3D model of E3 bound to both FAD (noncovalently) and NADH (NAI) (1ZMD) in the B chain of E3. Symmetry axes are not shown



NCBI iCn3D Figure 16.1.13: E3 bound to both FAD (noncovalently) and NADH (NAI) (1ZMD). (Copyright; author via source). Click the image for a popup or use this external: https://structure.ncbi.nlm.nih.gov/i...LAMqAS3Tx9SPE8.

Figure 16.1.14 shows an interactive iCn3D model highlighting the noncovalent interactions stabilizing bound FAD and NADH (NAI) in the E3 subunit of pyruvate dehydrogenase (1ZMD).



NCBI iCn3D Figure 16.1.14: Noncovalent interactions stabilizing bound FAD and NADH (NAI) in the E3 subunit of pyruvate dehydrogenase (1ZMD). (Copyright; author via source). Click the image for a popup or use this external:https://structure.ncbi.nlm.nih.gov/i...gUED1ZrJVjKcD8

Let's put it all together! Figure 16.1.15 shows a video of the pyruvate dehydrogenase complex from the HHMI.

Figure 16.1.15: Video of the pyruvate dehydrogenase complex

16.2: REACTIONS OF THE CITRIC ACID CYCLE

16.2.1: INTRODUCTION

The acetyl-CoA formed by the pyruvate dehydrogenase complex (PDC) now enters a cyclic, non-linear pathway called the citric acid cycle, tricarboxylic acid (TCA) cycle, or the Kreb's cycle after Hans Krebs who discovered it. The cycle is shown in Figure 16.2.1 in wedge/dash form with stereochemistry included to give a more exact representation.

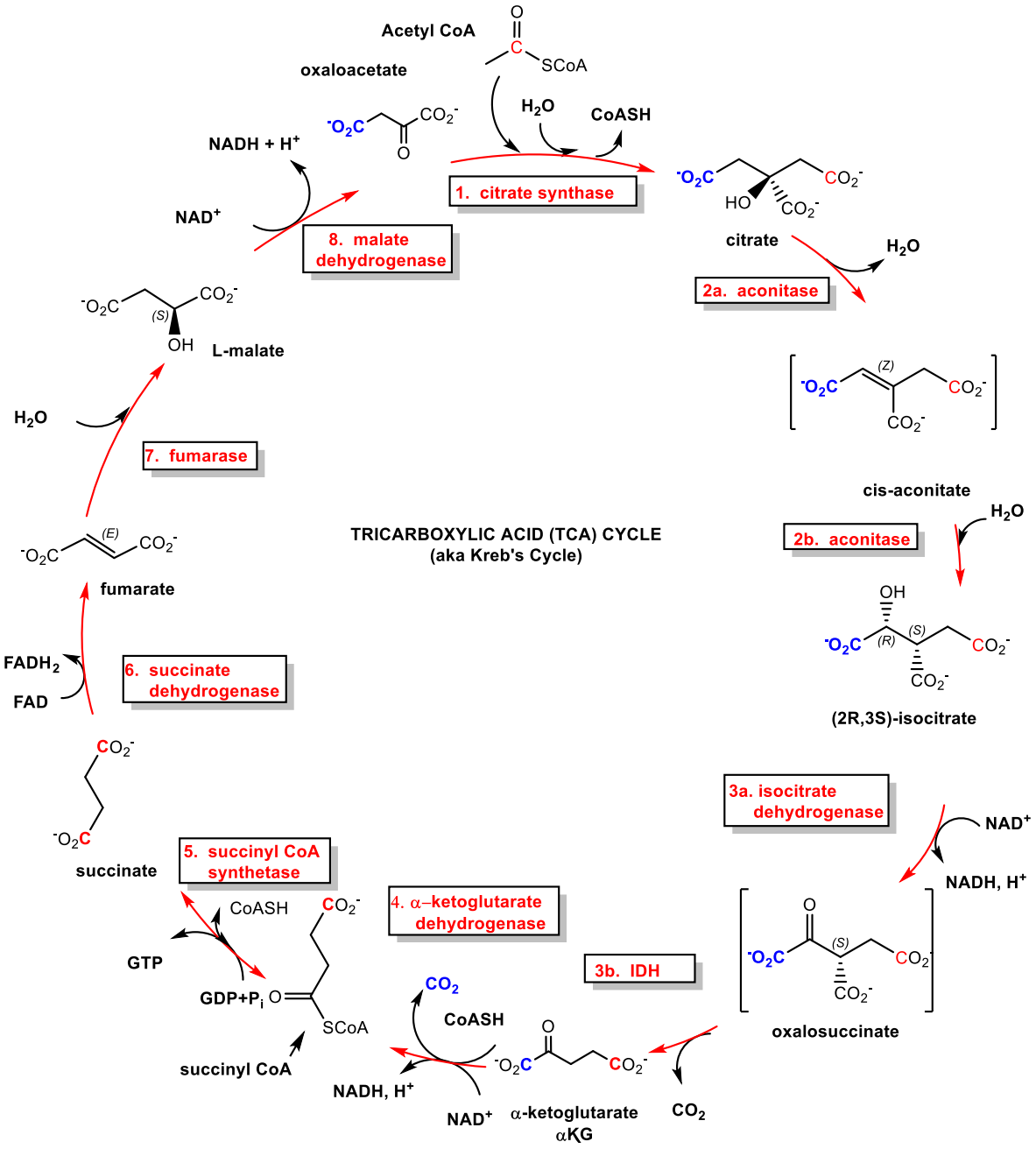


Figure 16.2.1: Citric Acid Cycle

Why is this pathway a cycle and not a linear pathway as we have seen for glycolysis? A simple answer is that it evolved that way, but why would that be advantageous? It turns out that some of the key "intermediates" in the pathway are pulled away for the biosynthesis of other biomolecules. If the citric acid cycle was linear, and intermediates pulled off for other reactions, the linear pathway would taper off, which would not be optimal for a key energy production pathway. Of course, the removal of intermediates from a cyclic pathway would also slow

it down but when this happens, enzymes outside of the cycle are used to produce key reaction intermediates of the cycle to it going. The replenishing reactions are called **anapleurotic reactions.**

Krebs, in his detailed analysis of the enzymes involved in "intermediary" metabolism, used radioisotope-labeled reactants to trace carbon atoms in respiring tissue. He found that when radiolabeled pyruvate and oxaloacetate were added to muscle tissue *in vitro*, radiolabeled citrate was formed.

$$\text{Pyruvate} + \text{Oxaloacetate} \rightarrow \text{citrate} + CO_2$$

This is correct but omits the initial conversion of pyruvate to acetyl-CoA. Hence the "end product" of the pathway (oxaloacetate) reforms the beginning reactant (citrate) so he surmised that the pathway was circular.

The next obvious question is why are eight reactions are required for the oxidation of just 2 Cs in pyruvate. Partly this is a matter of evolution again, as the early evolutionary pathway probably combined an oxidative (clockwise) set of reactions with a reductive (counterclockwise) set. Part of the chemistry in the cycle is devoted to producing either β-keto acids, which are easy to oxidatively decarboxylate, or converting α-ketoacid to molecules with better electron sinks β to the departing CO_2. After the net 2 carbon atoms added to the cycle are released as 2 CO_2s, the rest of the reactions are used to regenerate oxaloacetate, allowing the cycle to continue.

Of course, the ultimate goal of an energy-extractive oxidative pathway is not just to form CO_2 but to form ATP or its equivalent (i.e. GTP). Notice that 3 NAD^+s are used and converted to 3 NADH. In addition, a new, more potent oxidizing agent, FAD, is used and it is converted to $FADH_2$. NAD^+ and FAD are replenished by reoxidation of NADH and $FADH_2$ (reduced forms) back to NAD^+ and FAD, through mitochondrial electron transport (oxidation) reactions, in which electrons are passed to stronger and stronger oxidizing agents, the last being O_2. In this thermodynamically favored process, lots of ATPs are made. We will explore those reactions in the next section.

We will go through each of the steps in the citric acid cycle separately and show how the pathway is regulated (section 16.3). Why such detail? There are only 8 steps. It seems that we should be able to carefully examine each given that the citric acid cycle is a hub that along with glycolysis controls metabolic flow through many interconnected metabolic pathways.

At the same time, we can't explore each reaction in every pathway described in this text in great detail, otherwise, this book would become more of an encyclopedia. In this chapter and beyond, we will focus on mechanistic details only of enzymes that catalyze different types of reactions than those in glycolysis or the citric acid cycle, and those with interesting cofactors and mechanisms.

Other issues add complexity for learners. The PDC and citric acid cycle reaction occur in the mitochondrial matrix. Cytoplasmic pyruvate and NAD^+ must be transported into the matrix from the cytoplasm. In addition, some of the enzymes in the citric acid cycle have both cytoplasmic and mitochondrial variants. Some of these homologous pairs are differentiated by their use of NAD^+ or $NADP^+$ as an oxidizing agent. The ones in the cytoplasm are not part of the cycle. You would expect these enzyme pairs to have similar tertiary structures and active site chemistry. Prokaryotic forms of these enzymes are similar structurally to their eukaryotic forms so the interactive molecular models shown below will show enzymes from a variety of organisms. Finally, there are many variants, shunts, and bypasses of the citric acid pathway in different organisms. We will explore this topic in section 16.4

We will try to pair reaction mechanism diagrams that show the flow of electrons in bond-making and breaking with interactive molecular models of the active site. You should rotate the models to align and identify key amino acids and ligands (substrate, substrate analogs, inhibitors, activators) shown in the static 2D mechanism diagrams. Since the active sites are often conserved across prokaryotic and eukaryotic versions, the choice of PDB structures used depends on which best illustrates a conceptual point.

NOTE

There are many ways to write abbreviated chemical equations showing NAD^+/NADH and FAD/FAD_2 and hydrogen ions in metabolic pathway diagrams. To make sense of them, consider a simplified mechanism for the oxidation of ethanol by alcohol dehydrogenase, as shown in Figure 16.2.2. Note that there are 2 Hs on the oxidized substrate (ethanol) that are involved. One is a hydride and the other is a proton.

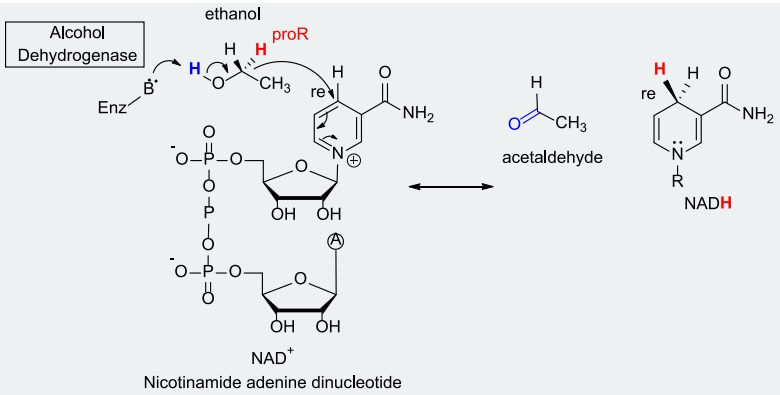


Figure 16.2.2: Oxidation of ethanol using NAD^+

Here is a list of different and seemingly contradictory ways to write a chemical equation to show changes in NAD^+/NADH and H ions:

1. $NAD^+ + :H^- \rightarrow NADH$. This chemical equation is charge balanced and shows just the changes to the NAD^+/NADH pair, but it doesn't show the proton (H^+) lost from the substrate.
2. $NAD^+ + 2e^- + H^+ \rightarrow NADH$. This is the same as equation (1) but with the hydride separated into an electron pair and a proton.
3. $NAD^+ + H^+ \rightarrow NADH$. This is balanced for Hs but not + charge as it doesn't explicitly show the electron pair from the hydride added to NAD^+.
4. $NAD^+ \rightarrow NADH + H^+$. This is balanced for charge but not for Hs. The extra H^+ is the proton from the oxidized substrate.

We will use example 4 above throughout this book. That equation is most useful when trying to account for the change in the number of protons in the individual reactions and entire pathways. We will also write simplified chemical equations involving FAD (in which 2H from the substrate are added) as FAD $\rightarrow$ $FADH_2$.

16.2.2: 1. CITRATE SYNTHASE (CS)

Oxaloacetate + Acetyl-CoASH + $H_2O \rightarrow$ Citrate + CoASH ΔG^o = -7.5 kcal/mol (-31 kJ/mol)

Acetyl CoA is a thioester. Hence it is "high energy" compared to its hydrolysis products. (**Remember, there is no such thing as a "high energy" bond.**) The free energy released in its hydrolysis is used to drive the reaction forward. This is important otherwise citrate would not be formed readily. This reaction feeds the end product of glycolysis into the citric acid cycle. It is summarized in Figure 16.2.3.

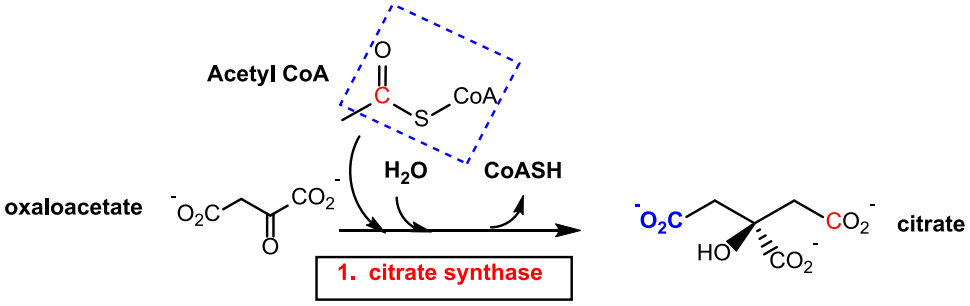


Figure 16.2.3: Summary reaction - Citrate Synthase

The enzyme exists in two major conformations, an open and closed form. When the open form, which has a binding site for oxaloacetate, binds the substrate, a shift to the closed conformation forms on the binding site of acetyl-CoA. These changes sequester the bound substrates and exclude water and prevent spurious hydrolysis of acetyl-CoA. The binding occurs sequentially so the kinetics follow a sequential ordered mechanism.

Figure 16.2.4 left shows an animated gif that shows the conformational changes between the citrate-bound version (open, green) and the citrate and CoASH-bound form (blue). The image below right shows a smoother transition between the open and closed form without bound ligands (1cts, 2cts)

Figure 16.2.4: Conformational changes in citrate synthase on binding substrate

For a more details view of the enzyme view the Regulation of Citrate Synthase.

The mechanism below is from https://chem.libretexts.org/Bookshel...trate_Synthase, with Contributors and Attributions from:

- Dr. Dietmar Kennepohl FCIC (Professor of Chemistry, Athabasca University)
- Prof. Steven Farmer (Sonoma State University)
- Organic Chemistry With a Biological Emphasis by Tim Soderberg (University of Minnesota, Morris)

In this reaction, a C-C bond must form between the substrates. One way to do that is to make a nucleophilic carbanion ion from the alpha carbon of acetyl CoA. Remember, this is not a decarboxylation reaction so we don't have to worry about an electron "sink" on the beta carbon. Forming the carbanion would be possible since the negative charge on the carbon can be withdrawn to the carbonyl oxygen to form an enolate. The enolate becomes even more stable if the negative oxygen is protonated. So this is reaction is an aldol condensation, the addition of an enolate to an aldehyde or ketone. The carboxylate group of aspartic acid 375 on citrate synthase removes the acidic alpha proton on acetyl CoA, while histidine 274 donates a proton to form the neutral enol, a much more stable molecule than the enolate anion. His 274 continues to stabilize the enol during the reaction. Bound oxaloacetate is stabilized in part by Arg 329. In the next part of the mechanism, a second histidine (320) protonates the carbonyl oxygen of oxaloacetate, activating the carbonyl carbon for nucleophilic attack by the enol in the next step to form (S)-citryl CoA. The hydrolysis of the CoASH occurs when His 320 deprotonates a water molecule, facilitating nucleophile attack on the carbonyl carbon bonded to -SCoA, forming citrate.

A plausible mechanism is shown in Figure 16.2.5.

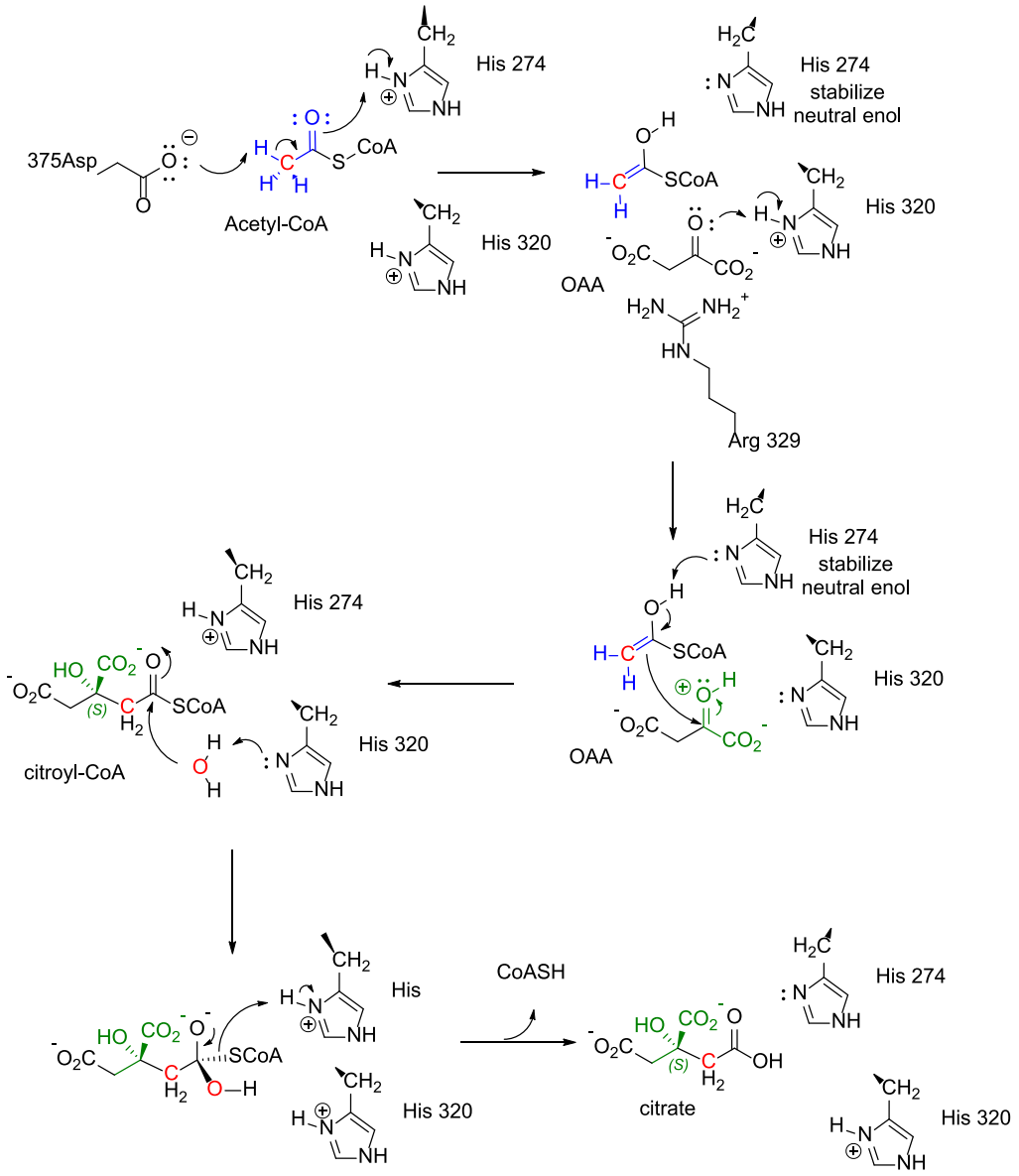
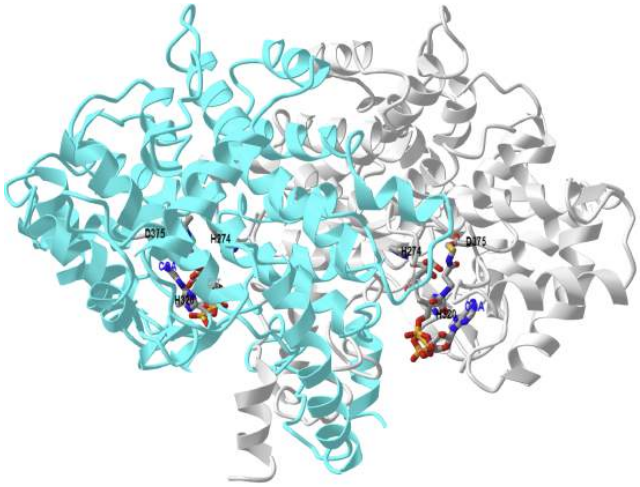


Figure 16.2.5: Citrate synthase mechanism

Figure 16.2.6 shows an interactive iCn3D model of the pig citrate synthase bound to CoASH and citrate (2CTS)



NCBI iCn3D Figure 16.2.6: pig citrate synthase bound to CoASH and citrate (2CTS). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...WXqcEsu336tcL7

The enzyme is a dimer with monomers shown in different colors. Citrate and CoASH are shown in sticks and labeled. The active site residues are shown as sticks and labeled in each subunit.

16.2.3: 2. ACONITASE

Citrate ↔ Isocitrate ΔG^o = +2 (rx 2a), -0.5 (rx 2b) kcal/mol; net ΔG^o = + 1.5 kcal/mol (+6.3 kJ/mol)

Thinking like a chess player, who must anticipate future moves, the chemical rationale for this reaction is to move an OH to a beta position, which in a subsequent reaction is converted to a beta C=O, so it can act as an electron sink to facilitate decarboxylation in a following reaction! The reaction is readily reversible (note the low ΔG^o value) since the reactant and product are simple isomers of each other. Figure 16.2.7 shows the summary reaction.

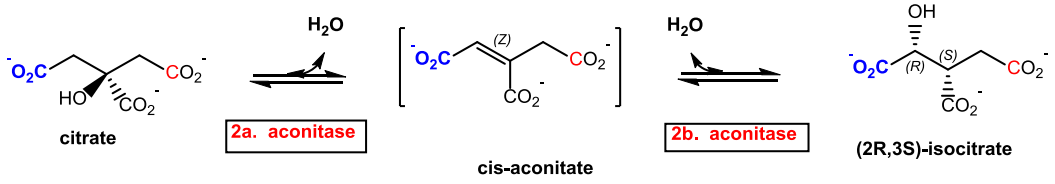


Figure 16.2.7: Summary reaction of aconitase

The enzyme has an inorganic Fe_4S_4 cluster. Each Fe in the cluster coordinates to 4 S^{2-} in a cubane structure, but when either citrate or isocitrate is bound, one of the Fe ions interacts with both the Os of a substrate carboxylate shown. The other two carboxylates of isocitrate are stabilized through ion-ion interactions by Arg 446 and 663.

Figure 16.2.8: below shows a plausible partial mechanism. An active site deprotonated serine abstract a proton at the S carbon. This is followed by the formation of the C-C double bond and a release of the resulting cis-aconitate from one bond to the FeS cluster.

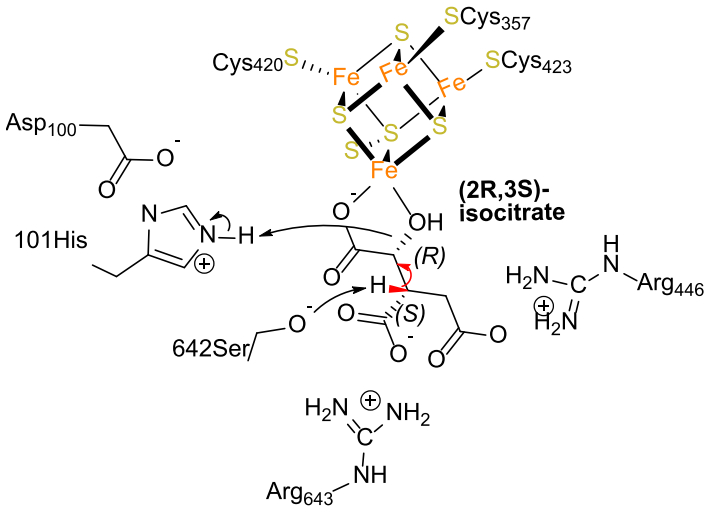
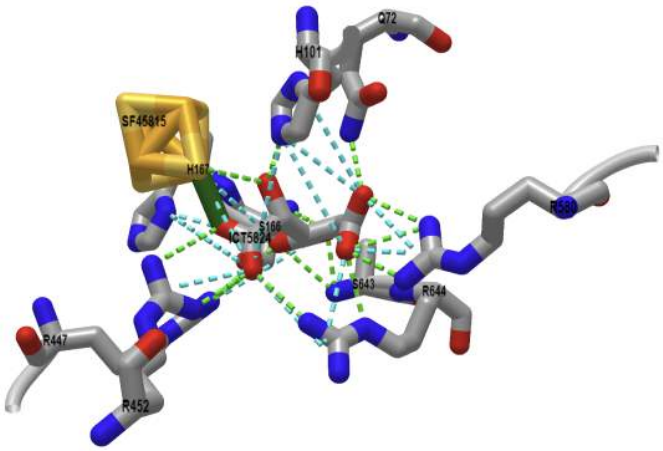


Figure 16.2.8: Mechanism of aconitase

Figure 16.2.9 shows an interactive iCn3D model of the bovine S642A aconitase with bound citrate (1C97).



NCBI iCn3D Figure 16.2.9: Bovine S642A aconitase with bound citrate (1C97). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...iJJRGaFoCDnhf7

This cis-aconitate intermediate in the interconversion of citrate and isocitrate must do an 180^0 flip around the C=C double bond. This is followed by rehydration to form the other isomer. The deprotonated His 101 abstracts a hydrogen from a water-bound to the FeS cluster, with the hydroxide acting as a nucleophile, which along with the redonation of a hydrogen ion on the protonated Ser 642 the alpha-carbon completes the rehydration step in the formation of the other isomer.

Exercise 16.2.1

Why was the S642A mutant used to produce the structure shown in the above iCn3D model?

Answer

It allows the binding of substrate/product, in this case, isocitrate, to an inactive enzyme as the active site serine was mutated to a non-nucleophilic alanine of similar size. Hence no bond-making/breaking occurs in the complex.

References:

https://collab.its.virginia.edu/acce...-_-/index.html

16.2.4: 3. ISOCITRATE DEHYDROGENASE (IDH)

Isocitrate + NAD^+ → α-ketoglutarate + NADH + H^+ ΔG^o = -2.0 kcal/mol (-8.4 kJ/mol)

The chemical rationale should be clear. In this step, through an oxidative decarboxylation, the CO_2 is removed as NADH is produced. The NADH will be reoxidized back to NAD^+ in the electron transport chain, leading to ATP production (see next section). The reaction is summarized in Figure 16.2.10.

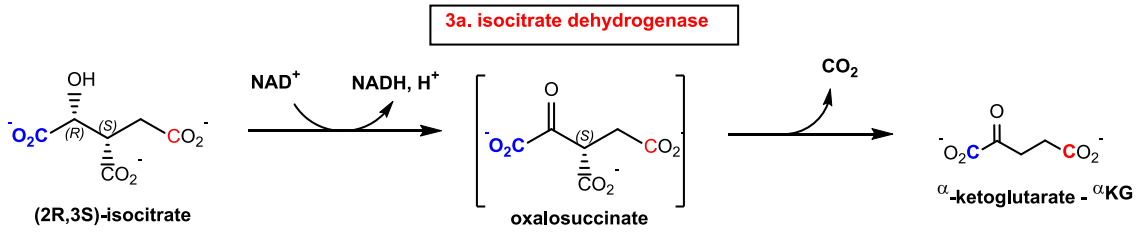


Figure 16.2.10: Summary reaction: isocitrate dehydrogenase

Exercise 16.2.1

Why must the oxidation reaction precede the decarboxylation reaction?.

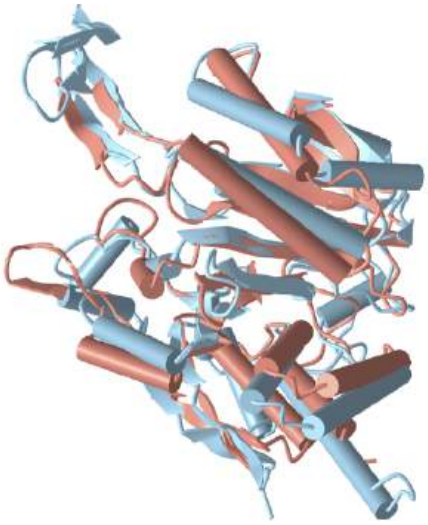
Answer

First, a beta-ketoacid intermediate must form, which allows easy decarboxylation of the intermediate as the beta carbonyl provides an electron "sink" to facilitate the decarboxylation.

There are two forms of this enzyme (IDH), a cytoplasmic ($NADP^+$) form and a mitochondrial (NAD^+) form. The cytoplasmic forms from various organisms are homodimers and have a common mechanism of catalysis. In contrast yeast mitochondrial IDH has two subunits, IDH1 (regulatory, binds the allosteric activator citrate and AMP) and IDH2 (catalytic, binds isocitrate and NAD^+). Mammalian IDHs are tetramers heterodimers ($\alpha\beta + \alpha\gamma$), which can also form a heterooctamer $(\alpha\beta + \alpha\gamma)_2$. The alpha chain is the catalytic subunit.

Mammalian NAD-IDHs are even more complex than yeast NAD-IDH. These enzymes are composed of three types of subunits, α, β, and γ, which share about 40–52% sequence identity. The α and β form an αβ dimer, while α and γ subunits form αγ. These then interact to form the $\alpha_2\beta\gamma$ heterotetramer, which effectively forms the holoenzyme. It can also form an active heterooctamer.e. The αγ heterodimer is regulated by citrate and ADP. On citrate binding to the allosteric site, a conformation change occurs to enhance isocitrate binding. ADP enhances the binding of the allosteric regulator citrate.

Figure 16.2.11 shows an interactive iCn3D model of that shows the superposition of the α chains of cytoplasmic IDH (NADP, sky blue, 4L03) and mitochondrial IDH (NAD) (6KDY, salmon).



NCBI iCn3D Figure 16.2.11: Superposition of the A chains of cytoplasmic IDH (4L03) and mitochondrial IDH (NAD) (6KDY) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...zy9yn4NaF9g1c9

Cytoplasmic IDH is sky blue and mitochondrial IDH is salmon. Click the 3 bar menu icon (top left) in the model in the window, scroll down to Alternate, and toggle back and forth between the two forms. The structures of the alpha chains, although not identical, align well.

Figure 16.2.12 shows a probable mechanism for the reaction based on the conserved catalytic site shown in the model above.

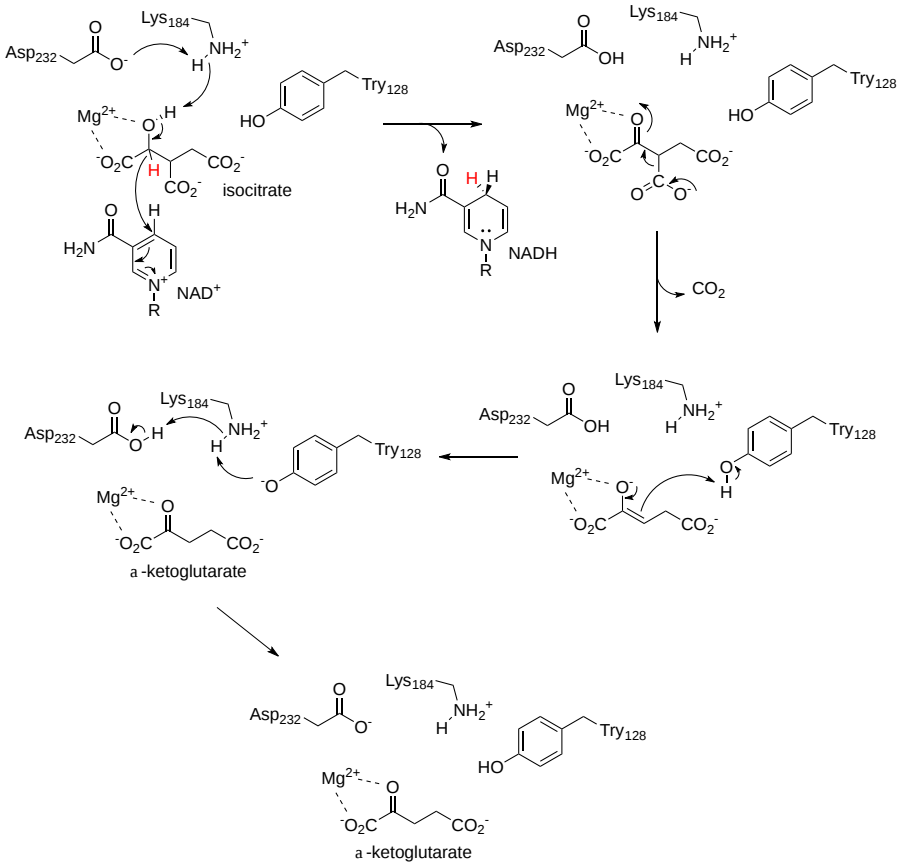
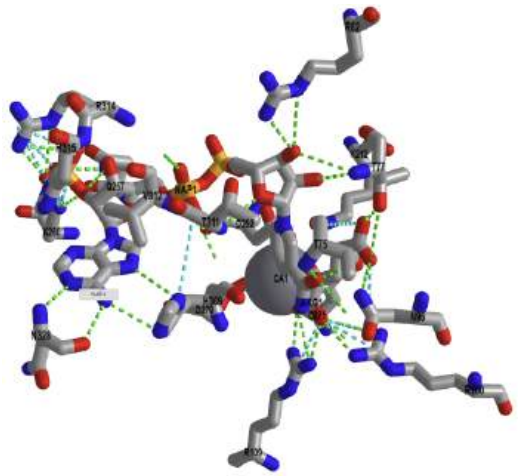


Figure 16.2.12: Mechanism of isocitrate dehydrogenase (after https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3706558/)

Figure 16.2.13 shows an interactive iCn3D model of active site of the cytoplasmic human IDH1 in complex with $NADP^+$ and Ca^{2+}/α-ketoglutarate (4L03).



NCBI iCn3D Figure 16.2.13: Active site of the cytoplasmic human IDH1 in complex with $NADP^+$ (NAP1) and Ca^{2+}/ alpha-ketoglutarate (4L03) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...P9DqLoPowEV2d9

Figure 16.2.14 shows an interactive iCn3D model of the active site of the αβ heterodimer of human IDH3 (6kdy) in complex with NAD^+.

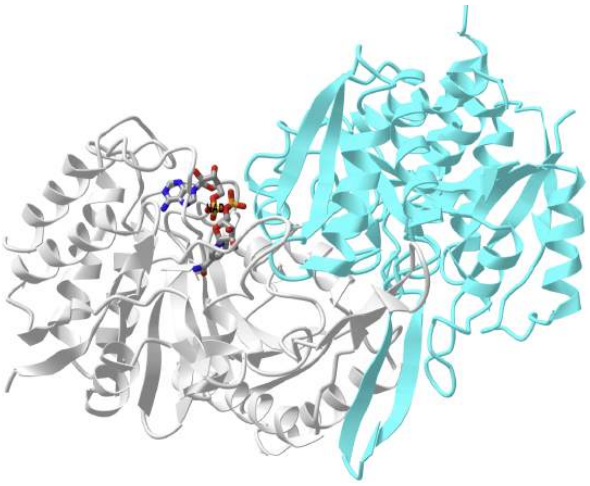


Figure 16.2.14: αβ heterodimer of human IDH3 (6kdy) in complex with NAD^+(6kdy) . (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...hyyJiMFEqwktB7

The α subunit is shown in gray and the β in cyan.

16.2.5: 4. A-KETOGLUTARATE DEHYDROGENASE

α-ketoglutarate + NAD^+ + CoASH → succinyl CoA + CO_2 + NADH + H^+ ΔG^o = -7.2 kcal/mol (-30 kJ/mol)

In the last reaction, α-ketoglutarate was formed. Oh no, you might say! It would have been nice to form a β-ketoacid, which could easily decarboxylation. No worries though. We spent all of section 16.1 explaining the biochemistry used to decarboxylate another α-ketoacid, pyruvate. The same chemistry is used to accomplish the oxidative decarboxlation of α-ketoglutarate. Hence we won't expand on the mechanism here. The reaction is shown in Figure 16.2.15.

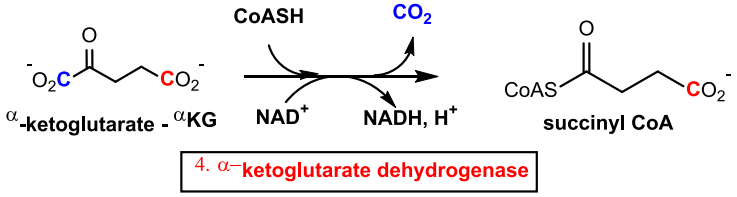


Figure 16.2.15: Summary reaction for α-ketoglutarate dehydrogenase

16.2.6: 5. SUCCINYL-COA SYNTHETASE (SCS)

succinyl CoA + GDP + P_i → succinate + CoASH + GTP ΔG^o = -0.8 kcal/mol (-3.3 kJ/mol)

This is the first step in which the energy change in the cycle is captured specifically in the form of a high energy (with respect to its hydrolysis product) phosphoanhydride bond in the form of GTP (and ATP in some organisms). From a chemical step, the cleavage of the thermodynamically unstable thioester is coupled to the endergonic synthesis of GTP. This can transfer its terminal phosphate to ADP to make ATP in reaction that has a ΔG^o of about 0 kcal/mol. The reaction is shown in Figure 16.2.16.

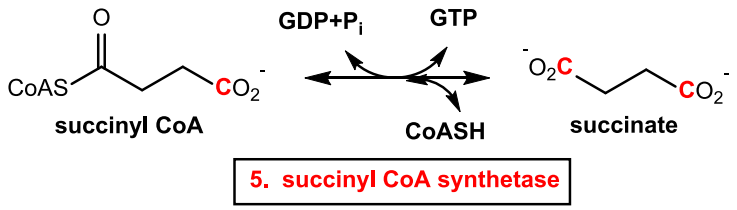


Figure 16.2.16: Summary reaction for succinyl-CoA synthetase

Succinyl-CoA synthases have two subunits, α and β. The enzyme in E. coli is a tetramer ($\alpha_2\beta_2$) with the catalysis occurring at the αβ interface. The alpha-subunits interact only with the beta-subunits, whereas the beta-subunits interact to form the dimer of alpha beta-dimers with CoA bound in each α subunit to a nucleotide-binding loop.

Two histidines, His 246 and His 142 are involved in the reaction, with His 246 becoming phosphorylated to form an intermediate in the reaction. A mutation of His 142 to an asparagine (H142N) essentially abolishes enzyme activity. Different SCSs have different specificities for purine nucleoside triphosphates. Organisms, including mammals, may have two different isoforms, one that binds ADP and

one that uses GDP (as shown in most diagrams of the citric acid cycle). In E. Coli, the α subunit binds CoASH and contains His 246, which gets phosphorylated. The β subunit determines the specificity for either GTP or ATP. In E.Coli the ATP binding site (Site II "in the ATP-grasp fold") is quite distant from the CoASH site (Site II) so phospho-His 246 must move between the sites in the dimer interface.

The three steps in the reaction are shown below, where E is the free enzyme, a · indicates a noncovalent complex, and a - represents a covalent bond (after Biochemistry 2002, 41, 537-546)

1. E + succinyl-CoA + P_i ↔ E · succinyl-PO_3 + CoASH
2. E · succinyl-PO_3 ↔ E-PO_3 + succinate
3. E-PO_3 + NDP ↔ E + NTP

Figure 16.2.17 shows an abbreviated mechanism that shows only the involvement of His 246.

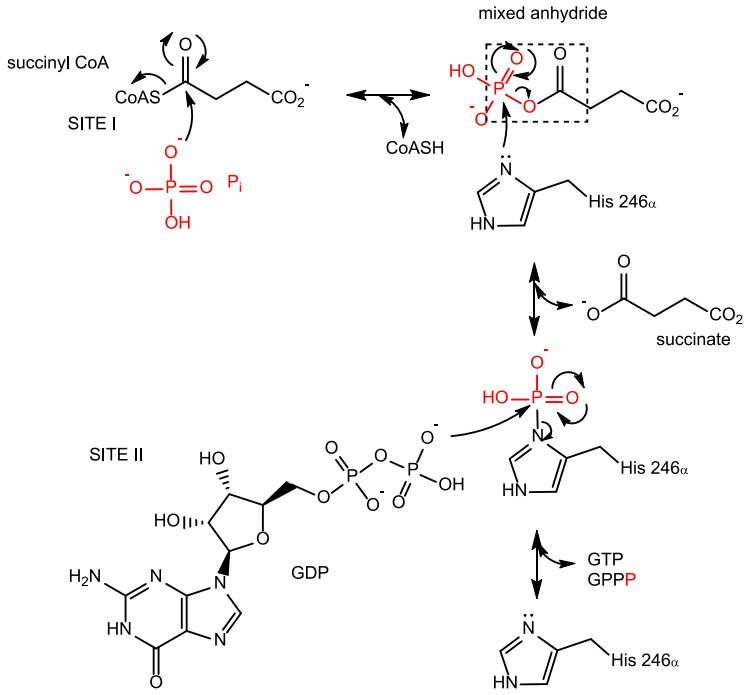


Figure 16.2.17: Abbreviated mechanism for succinyl-CoA synthase

Kinetic analysis suggests that the three substrates bind in a specific order, catalysis occurs, and then the three products leave. This type of reaction is called an ordered ter ter reaction, and is shown in Figure 16.2.18.

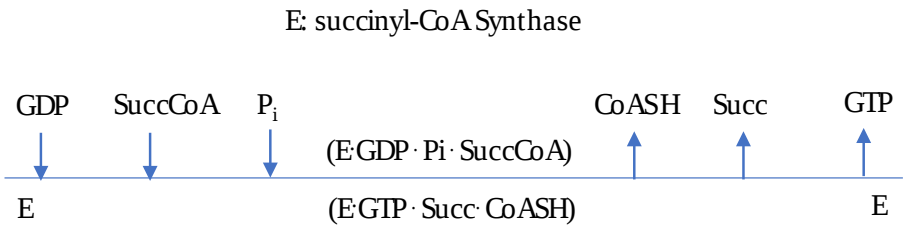


Figure 16.2.18: Ordered ter ter reaction for succinyl-CoA synthase

The iCn3D model below shows key alpha-chain residues in the active site, including the phosphorylated His 246 and bound CoASH. (use iCn3D to visualized 1CQJ, the nonphosphorylated form) I)

Figure 16.2.19 shows an interactive iCn3D model of the complex of ADP and Mg2+ with Dephosphorylated E. Coli Succinyl-CoA Synthetase (1CQI)

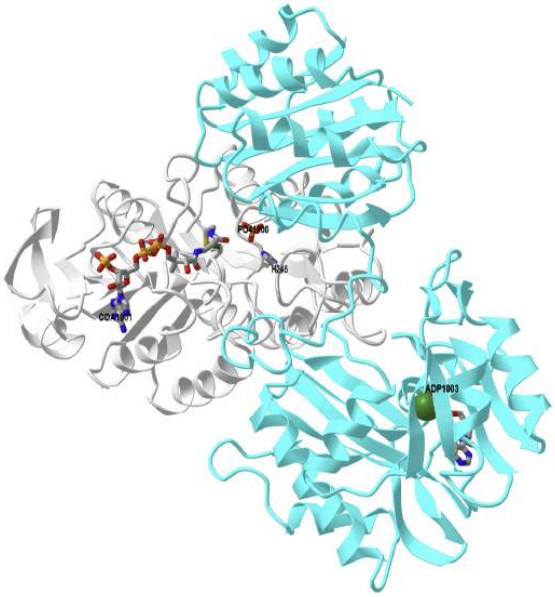


Figure 16.2.19: Complex of ADP and Mg2+ with Dephosphorylated E. Coli Succinyl-CoA Synthetase (1CQI). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...BssxsabV1S9cz6

Just one αβ dimer is shown. The α subunit is shown in gray and the β in cyan. Bound P_i and CoASH are labeled. The active site His246 in the α chain is shown in stick and labeled.

Figure 16.2.20 shows an interactive iCn3D model of the active site of pig GTP-specific succinyl-CoA synthetase in complex with succinate and CoASH (5CAE).

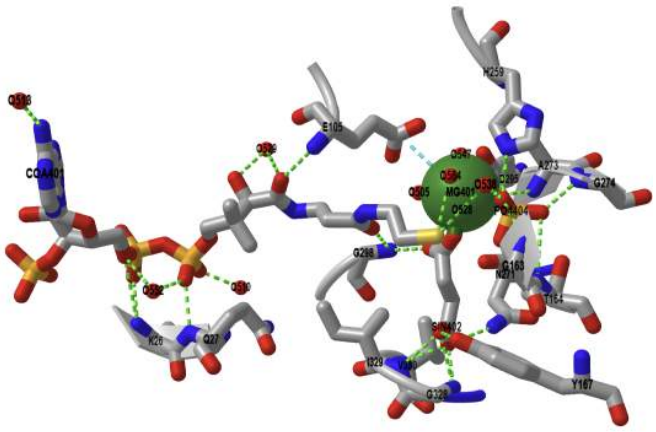


Figure 16.2.20: Active site of pig GTP-specific succinyl-CoA synthetase in complex with succinate and CoASH (5CAE). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...HjH9ER5wLqb1T7

16.2.7: 6. SUCCINATE DEHYDROGENASE

succinate + FAD ↔ fumarate + $FADH_2$ ΔG^o = 0 kcal/mol

The enzyme is yet another step in closing the cycle to reform oxaloacetate. It is also our first encounter with FAD as an oxidizing agent. Its reduction product, $FADH_2$, will be reoxidized in the electron transport chain (mitochondrial inner membrane for eukaryotes), producing energy for ATP. Hence it can be considered a proxy ATP-generating reaction. The reaction is shown in Figure 16.2.21.

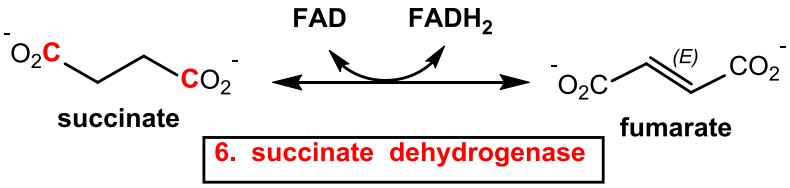
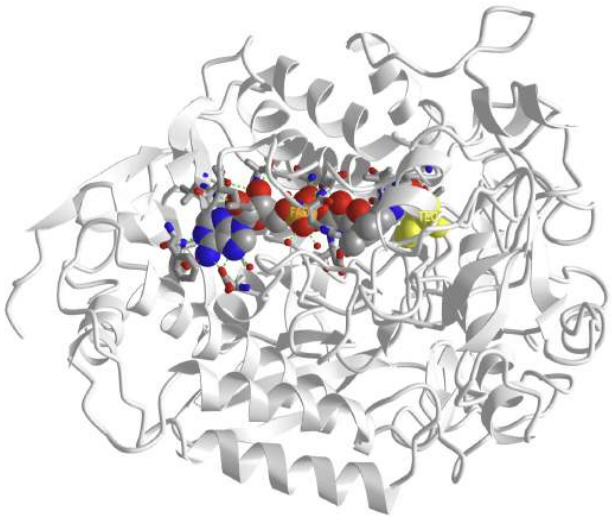


Figure 16.2.21: Summary reactoin for succinate dehydrogenase

The succinate dehydrogenase enzyme is part of the larger Complex II of the electron transport chain. Complex II has many cofactors involved in its overall activity. It also uses an iron/sulfur cluster cofactor, similar to aconitase, which also produces a C=C doubled bonded intermediate. We will discuss it in greater detail in the chapter on electron transport. For now, let's concentrate on this new cofactor and oxidizing agent, FAD. Many enzymes use FAD/$FADH_2$ in redox chemistry.

In contrast to NAD^+/NADH, the FAD/$FADH_2$ pair stays tightly bound to the enzyme and doesn't readily dissociate. This means that after one cycle of the enzyme (after FAD is converted to $FADH_2$), the enzyme is functionally "dead". Another oxidizing agent must bind to the enzyme and reoxidize $FADH_2$ back to FAD. The dissociation constants for FAD/$FADH_2$ and its protein binder in a flavoprotein are often in the nanomolar range. In around 10% of flavoproteins, FAD/$FADH_2$ are usally covalently bonded to the enzyme.

Figure 16.2.22 shows an interactive iCn3D model of the Avian respiratory complex II FAD binding subunit with FAD and a malate-like intermediate (1YQ3).



NCBI iCn3D Figure 16.2.22: Avian respiratory complex II FAD binding subunit with FAD and a malate-like intermediate (1YQ3). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...SGac3NuN4XBWk7

Note how buried the FAD is in the middle of the dehydrogenase subunit. The malate-like intermediate (TEO) is shown next to the FAD in yellow spacefill.

Exercise 16.2.1

Succinate dehydrogenase is irreversibly inhibited by the toxin 3-nitroproprionic acid (3np) made by some plants and fungi. Eating moldy sugar cane has led to reported deaths.

1. Draw the Lewis structures of succinate and 3-nitroproprionic acid. Compare them and the total number of valence electrons in each.

2. Here is a link to an iCn3D model showing the interaction of 3np with the enzyme. Explain the mode of action of the toxin.

https://structure.ncbi.nlm.nih.gov/i...cja9VKU2ZMCac9

Answer

These molecules are structurally similar and isoelectronic (the same number of electrons in their Lewis structures.

The inhibitor 3np form a covalent adduct through the guanidino group of Arg 297. This is a key catalytic residue that act as a general base that accepts a proton from succinate in the reaction.

Figure 16.2.23 (top) shows a very general and abbreviated mechanism for the enzyme for the immediate reaction of succinate with FAD. The bottom part of the image shows the amino acids surrounding succinate in the avian (bird) version of the enzyme (pdb 1yq4)

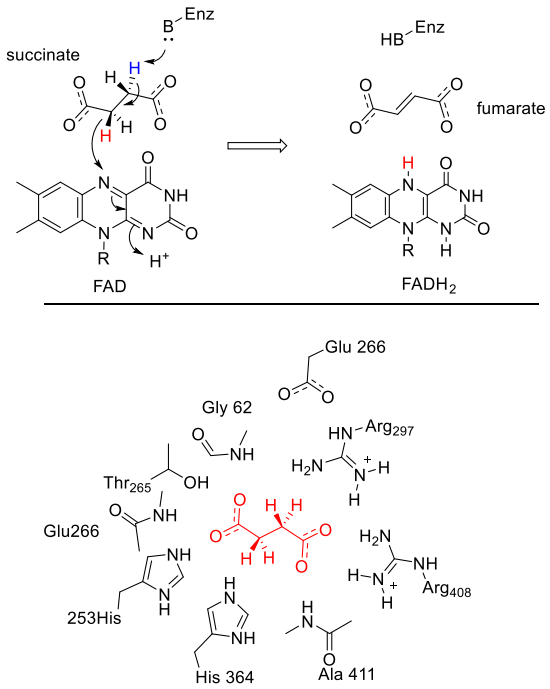


Figure 16.2.23: Top - a very general and abbreviated mechanism for the enzyme for the immediate reaction of succinate with FAD. Bottom - amino acids surrounding succinate in the avian (bird) version of the enzyme (1yq4)

Exercise 16.2.1

One of the amino acids surrounding succinate in the figure above acts as a general base and abstracts a protein from succinate as a hydride is transferred (from a plane above) to FAD. Go to this iCn3D of the active site bound to FAD. Which amino acid is the likely general base? (Note: the figure below shows general protonated states of side chains and not necessarily those involved in the proton abstraction. Go to **Analysis**, **Distance** and **Distance between 2 atoms** to find the likely general base.

Answer

Arg 297

16.2.8: 7. FUMARASE

fumarate + H_2O ↔ L-malate ΔG^o = -0.9 kcal/mol (-3.8 kJ/mol)

The chemical rationale for this reaction is clear - it is the penultimate step in the resynthesis of oxaloacetate, one of the reactants that starts the cycle, allowing the cycle to continue. This reaction introduces an O by hydration which can be oxidized in next step to produce NADH for e- transport/ATP production. The reaction is shown in Figure 16.2.24.

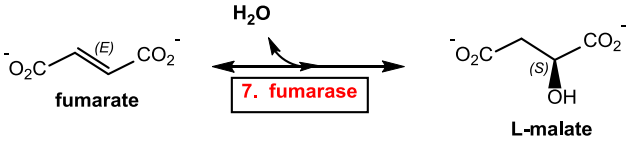


Figure 16.2.24: Summary reaction for fumarase

There are Class I (dimers containing an unstable FeS cluster, examples A and B) and Class II (tetramer, no bound iron, oxygen stable, example C) fumarases. Humans have both cytoplasmic and mitochondrial type II fumarases, resulting from alternative transcription of the fumarase genes. We will consider the type II, fumarase C from *E. Coli* in the following discussion.

The tetramer contains just alpha helices and random coils and has two distinct binding sites. Site A appears to be the active site and contains a buried water molecule. Site A, formed from three of the subunits, binds competitive inhibitors such as citrate and β-(trimethylsilyl)maleate, a cis substrate for fumarase, and is buried. 12 Angstroms away is site B, which is found in only one of the subunits

near a pi-helix (H129 through N135) and is more surface-exposed. Each site has a histidine, but mutation of only one H188N in the A site disrupts enzyme activity. Both sites bind multi-carboxylates. The role of site B is a bit unclear, but it is most likely an allosteric site involved in the transfer of product (malate) from the buried site to the surface for ultimate dissociation.

Figure 16.2.25 shows an interactive iCn3D model of the fumarase with beta-(trimethylsilyl)maleate and citrate (1fuq).

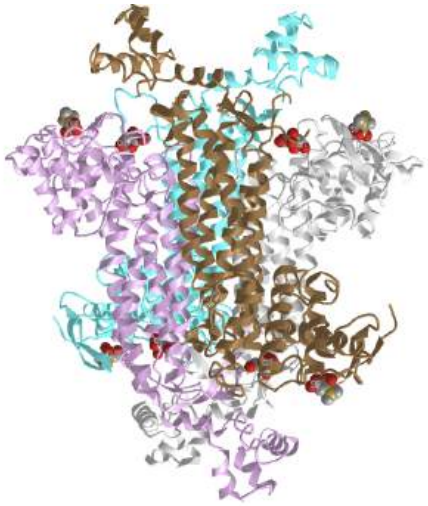


Figure 16.2.25: Fumarase with beta-(trimethylsilyl)maleate and citrate (1fuq). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...MewEQFebi4VL48

The monomers in the tetramer are shown in different colors. Citrate (Cit) and beta-(trimethylsilyl)maleate (SIF) are shown in spacefill.

Figure 16.2.26 shows an interactive iCn3D model of the binding site of citrate, a competitive inhibitor, of fumarase (1FUQ).

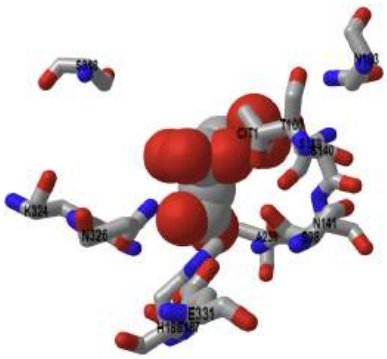


Figure 16.2.26: Binding site of the competitive inhibitor citrate in fumarase (1FUQ). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...yfPc5iPWbExr86

Figure 16.2.27 shows a plausible mechanism for the trans addition of water to fumarate.

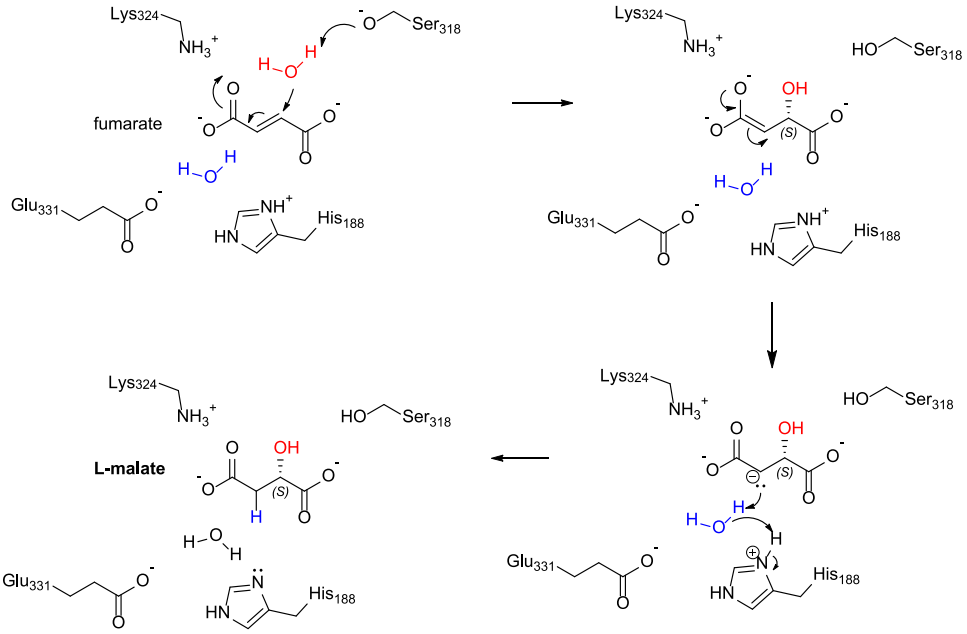


Figure 16.2.27: Mechanism for fumarase

16.2.9: 8. MALATE DEHYDROGENASE (MDH)

L-malate+ NAD^+ ↔ oxaloacetate + NADH + H^+ ΔG^o = +7.1 kcal/mol (30 kJ/mol)

We are finally there! This last reaction of the citric acid cycle produces oxaloacetate, the starting reactant, so the cycle can continue. It also produces NADH for mitochondrial e- transport/ATP production. Notice that is thermodynamically unfavorable (in the standard state) but the reaction is pulled to citrate formation by the first and next step of the cycle, citrate synthase. The reaction is shown in Figure 16.2.28.

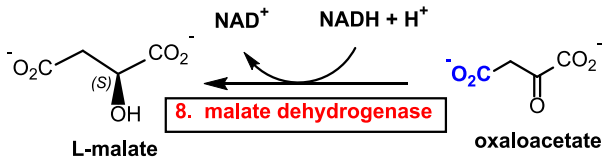


Figure 16.2.28: Summary reaction for malate dehydrogenase

Malate dehydrogenases are found in the cytoplasm, where it is part of the aspartate-malate shuttle that moves cytoplasmic malate (and through MDH indirectly NADH) into the mitochondria. It is also found in the mitochondria, where it is part of the citric acid cycle. There are also NAD^+ and $NADP^+$-dependent forms. Malate can undergo two different types of oxidation reactions, one producing oxaloacetate and using NAD^+, and one, an oxidative decarboxylation producing pyruvate and CO_2, using $NADP^+$. The latter is sometimes called malic enzyme.

Humans have two forms (MDH 1 and MDH 2) that use NAD^+. The enzyme is a homodimer in humans with binding sites on both. It activity is allosterically
regulated by citrate, and it is inhibited by many things, including ATP, ADP, AMP, fumarate, citrate, aspartate, and high concentrations of oxaloacetate.
The enzyme is similar to lactate dehydrogenase, which we encountered in the chapter of glycolysis. Kinetic analyses show that NAD^+ binds first followed by malate.

Figure 16.2.29 shows an interactive iCn3D model of NAD^+ and malate bound to human malate dehydrogenase 2 (4wlu).

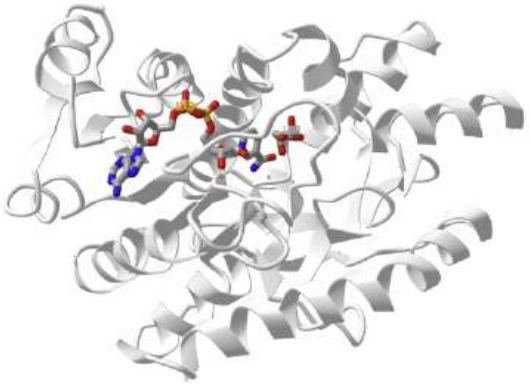


Figure 16.2.29: NAD^+ and malate bound to human malate dehydrogenase 2 (4wlu). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...DrAXBdu64GPzx8

Just one monomer is shown. NAD^+ and malate (LMR) are shown in sticks and labeled.

Figure 16.2.30 shows an abbreviated mechanism for malate dehydrogenase. The numbers refers to the *E. Coli* enzyme.

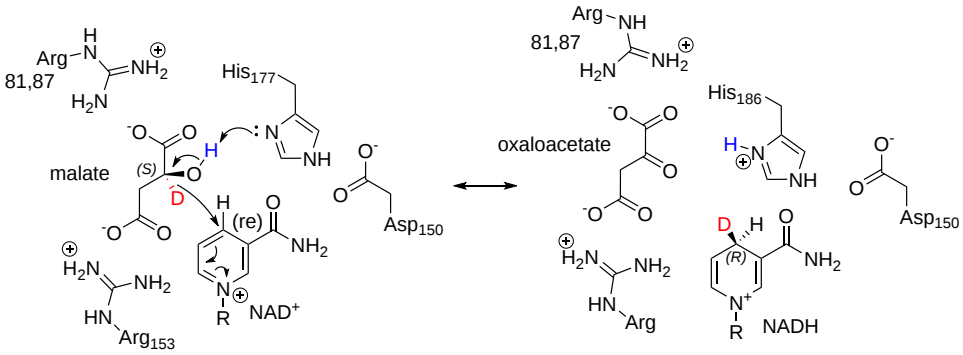


Figure 16.2.30: Abbreviated mechanism for malate dehydrogenase

Note that the hydride transferred from the malate is shown in red as a deuterium (D). It is transferred to the re face of NAD^+ to form NADH. The carbon with the transferred deuterium in NADH is prochiral. Think of that as that carbon being chiral if one of the 2 Hs could be arbitrarily assigned a higher priority in assigning R/S isomers. D has a higher priority than H in the Cahn/Ingold designation system. In the reverse reaction, the D atom, which is above the plane of the ring, occupies the proR position. The proR deuterium is transferred back in this reversible reaction. A D was used simply to indicate the stereochemistry and to assign it in NADH to the proR position.

Figure 16.2.31 shows an interactive iCn3D model of the active site of the E. Coli malate dehydrogenase with bound citrate and NAD^+ (1EMD).

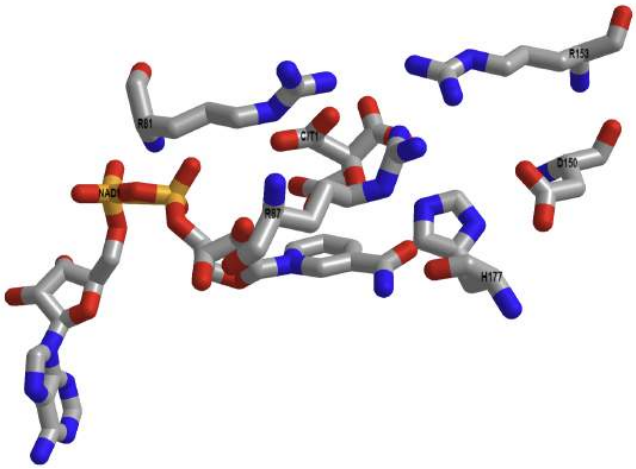


Figure 16.2.31: Active site of the E. Coli malate dehydrogenase with bound citrate and NAD^+ (1EMD). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...mHhmeFQyPpZj18

We have shown many renderings of the enzymes involved in the cycle. Yet another one is shown below. Figure 16.2.32 shows an interactive iCn3D model of the electrostatic surface potential of malate dehydrogenase (4WLU).

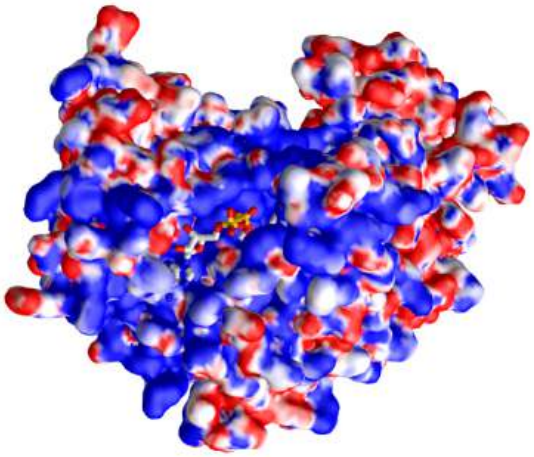


Figure 16.2.32: Electrostatic surface potential of malate dehydrogenase (4WLU)with bound NAD^+ (4WLU). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?izf2dsWXTw2bE8YD7

The display surface is the electrostatic surface potential map of the enzyme. Red shows the surfaces that are more anionic and with a negative electrostatic potential to which cationic molecules would be attracted, while blue represents more cationic surfaces to which the anion would be attracted. Note that the bound NAD^+, which has many oxygens which are slightly or fully negative, is bound in a blue, positive electrostatic potential region.

SUMMARY

Let's do some stoichiometry for the full cycle. Here is the net reaction (assuming that the GTP produced by succinyl-CoA synthetase is equivalent to 1 ATP).

$$\text{Acetyl-CoA} + 3NAD^+ + FAD + ADP + P_i + 2H_2O \rightarrow 2CO_2 + 3NADH + FADH_2 + ATP + 2H^+ + CoASH$$

This must seem like a lot of work to produce just 1 ATP, especially since the partial, anaerobic oxidation of glucose in glycolysis produced in net fashion 2 ATPs. The key, however, is to realize that 3 NADHs and 1 $FADH_2$ are produced, which when they are reoxidized in mitochondrial electron transport/oxidative phosphorylation, will produce multitudes of ATP.

As was true for glycolysis, this main energy-extracting pathway is highly regulated. We will this in the next section.

Figure 16.2.33 shows some key points about each reaction in the citric acid cycle.

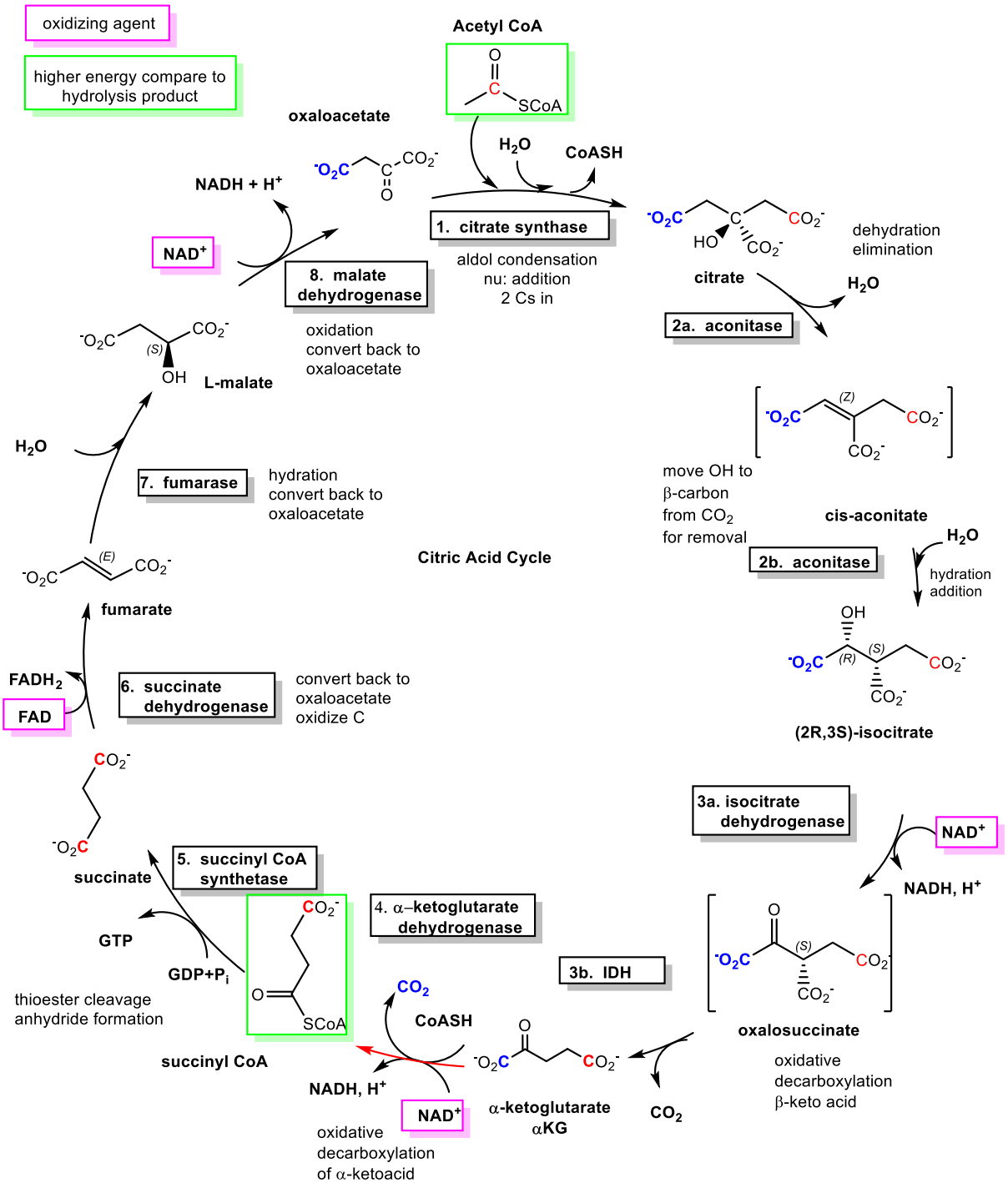


Figure 16.2.33: Summary of the citric acid cycle.

16.3: REGULATION OF THE CITRIC ACID CYCLE

16.3.1: OVERVIEW

Entry of pyruvate into the citric acid cycle leading to the aerobic production of energy and intermediates for biosynthesis is a key metabolic step. Hence both the pyruvate dehydrogenase complex and key enzymes in the cycle are targets for regulation. This occurs through substrate availability, product inhibition, allosteric effectors, and post-translational modifications of key enzymes in the pathway. A summary figure showing key regulators is shown in Figure 16.3.1.

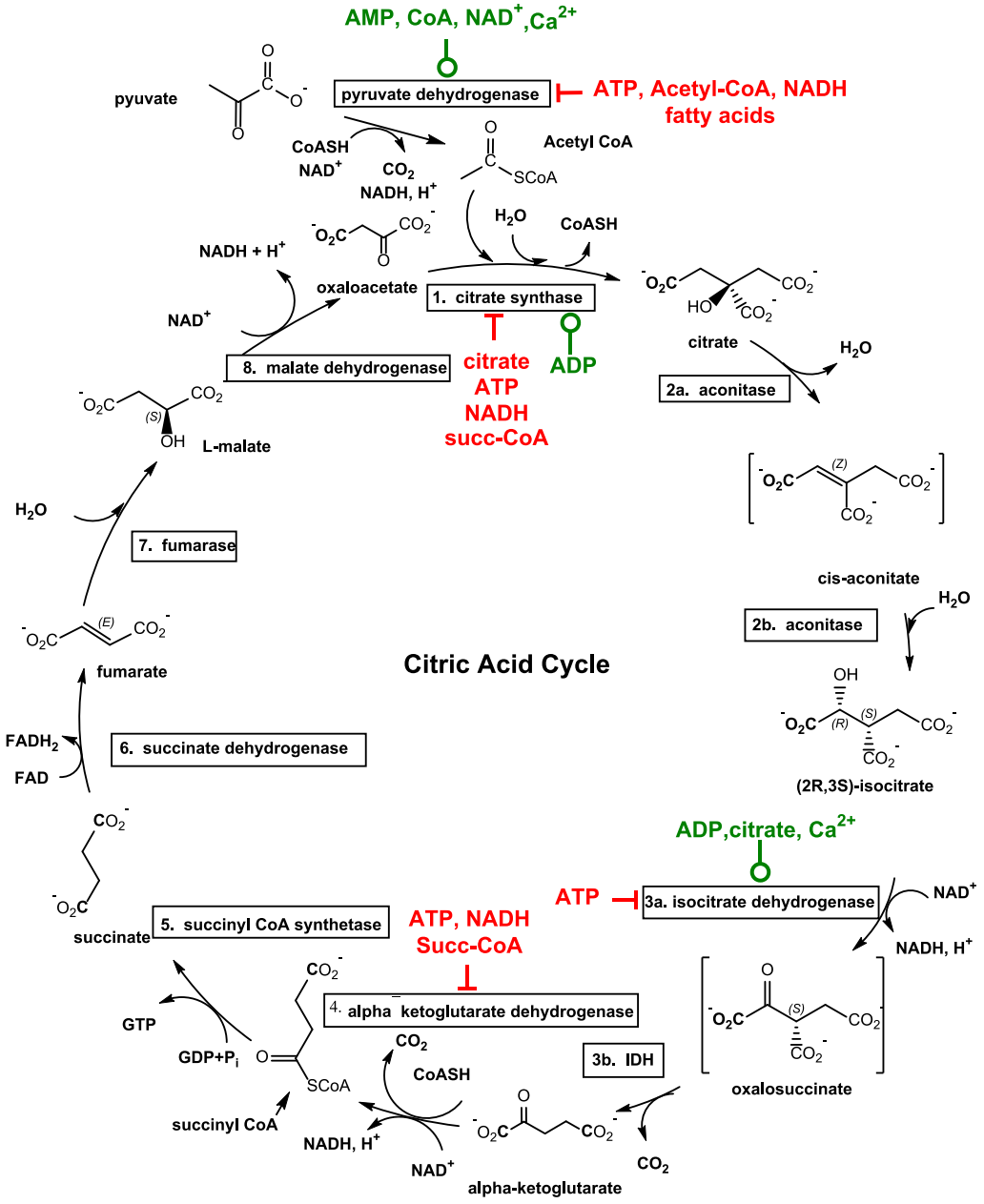


Figure 16.3.1: Regulation of the citric acid cycle

https://www.nature.com/articles/s41598-021-98314-z

Note that the key steps are regulated mainly by the ratio of mitochondrial NAD^+/NADH which is highly influenced by the ratio of ATP/ADP. High NADH inhibits the regulatory enzymes. High levels of acetyl CoA, derived from pyruvate dehydrogenase and also through fatty acids catabolism increases flux through the cycle in part by allosterically activating the first enzyme in the pathway, citrate synthase.

16.3.2: REGULATION OF THE PYRUVATE DEHYDROGENASE COMPLEX (PDC)

Under aerobic conditions, the pyruvate produced by glycolysis will be oxidized to acetyl CoA using the pyruvate dehydrogenase complex (PDC) in the mitochondria (note: its genes are encoded in the nucleus). As this enzyme is a key transition point (the gatekeeper) between cytosolic and mitochondrial metabolism and is highly exergonic ($\Delta G^{0'}$ = -7.9 kcal/mol, -38 kJ/mol), it is highly regulated by both covalent and allosteric regulation. Deficiencies of the PDC are X-linked and present with symptoms of lactic acidosis after consuming a meal high in carbohydrates. This metabolic deficiency can be overcome by delivering a ketogenic diet and bypassing glycolysis altogether.

The PDC is regulated by allosteric and covalent regulations. The complex itself can be allosterically activated by pyruvate and NAD^+. Elevation of the substrate (pyruvate) will enhance flux through this enzyme as will the indication of low energy states as triggered by high NAD^+ levels. The PDC is also inhibited by acetyl CoA and NADH directly. Product inhibition is a very common regulatory mechanism and high NADH would signal sufficient energy levels, therefore decreasing the activity of the PDC. Figure 16.3.2 summarizes the regulation. (Adapted from Marks' Medical Biochemistry)

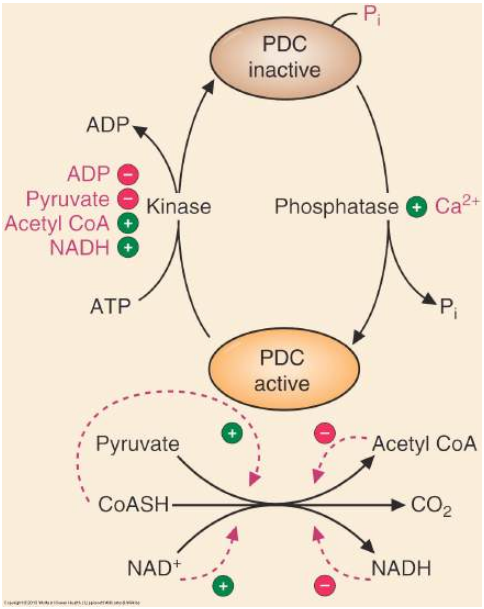


Figure 16.3.2: Regulation of pyruvate dehydrogenase.

The PDC is also regulated through covalent modification. Phosphorylation of the E1 subunits of the complex will decrease the activity of the enzyme.

The enzyme responsible for the phosphorylation of the PDC is **pyruvate dehydrogenase kinase**. The kinase is regulated inversely to the PDC, as shown in Figure 1 above. The kinase is most active when acetyl-CoA, NADH, and ATP are high. These compounds will stimulate the kinase to phosphorylate and inactivate the PDC. PDK is inhibited by dichloroacetate, TPP, Ca^{2+}, and pyruvate. The PDC can be dephosphorylated by a calcium-mediated phosphatase, PDP. Starvation and diabetes result in increased phosphorylation and inhibition of the complex, which impairs glucose oxidation.

Phosphorylation occurs on Serine 264 of the α subunit (site 1), Ser271 (site 2), and Ser203 (site 3) are located on a conserved phosphorylation loop. Sites 1 and 2 (in loop A) are involved in the stabilization of TPP in the active site, while Ser 203 in the adjacent loop B binds Mg^{2+} which stabilizes PP on bound TPP. All it takes for inhibition is the phosphorylation of just one of the Ser side chains. Phosphorylation prevents the ordering of the loop which occurs on TPP binding which hinders the binding of the lipoyl domains of the PDC core to E1p, which inhibits the flow of metabolites in the PDC.

Prevention of PDC phosphorylation by the specific PDK inhibitor dichloroacetate increases levels of reactive oxygen species in mitochondria, which promotes the expression of a mitochondria-K^+ channel axis, leading to cellular apoptosis and the inhibition of tumor growth

A summary of pathway regulation is shown in Figure 16.3.3.

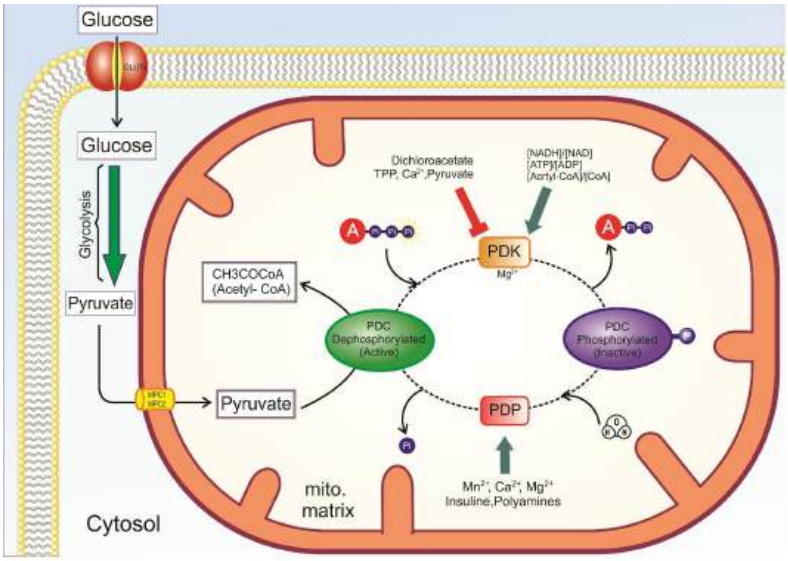


Figure 16.3.3: Summary of the regulation of pyruvate dehydrogenase

In general, PDC is activated through its substrates CoASH and NAD^+, and kinase inhibition or phosphatase activation (PDP) (dichloroacetate, TPP, Ca^{2+}, and pyruvate), is inhibited by its products acetyl CoA and NADH and activation of kinase (PDK). Abbreviations : PDC: pyruvate dehydrogenase complex, PDK: pyruvate dehydrogenase kinase, PDP: pyruvate dehydrogenase phosphatase TPP: thiamine pyrophosphate.

The complex is also acetylated and succinylated.

Table 16.3.1 below shows a summary of the regulation of pyruvate through glycolytic enzymes and pyruvate dehydrogenase

Metabolic Pathway	Major Regulatory Enzyme(s)	Allosteric Effectors	Post-translational modifications	Hormonal Effects
Glycolysis	hexokinase; glucokinase (liver)	Glucose 6P (-)		
	PFK-1	Fructose 2,6BP, AMP (+) Citrate (-)		↑ Insulin/Glucagon leads to dephosphorylation of PFK2 and increases producthe tion of F2,6BP
	Pyruvate Kinase	Fructose 1,6BP (+) ATP, Alanine (-)		↑ Insulin/Glucagon leads to dephosporylation
Pyruvate Dehydrogenase	PDC	Pyruvate, NAD^+ (+) Acetyl CoA, NADH, ATP (-)	dephosphorylation by PDP (+) phosphorylation by PDK (-)	↑ Insulin/Glucagon leads to dephosphorylation

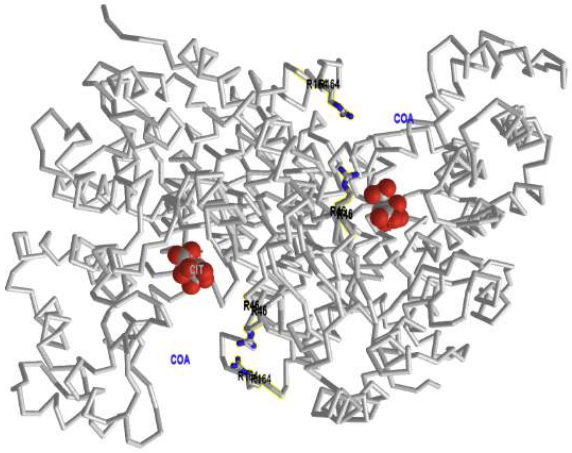
Table 16.3.1: Summary of the regulation of pyruvate through glycolytic enzymes and pyruvate dehydrogenase

16.3.3: REGULATION OF CITRATE SYNTHASE

Citrate synthase (ΔG^o = -7.5 kcal/mol, -31 kJ/mo), Isocitrate dehydrogenase (ΔG^o = -2.0 kcal/mol, -8.4 kJ/mol) and alpha-ketoglutarate dehydrogenase (ΔG^o = -7.2 kcal/mol, -30 kJ/mol) are all exergonic and likely candidates for regulation. Indeed they are. Let's start with citrate synthase.

Citrate synthase is regulated in part by the availability of substrate acetyl-CoA and oxaloacetate. It is inhibited by NADH and also citrate, a competitive inhibitor of oxaloacetate binding. Succinyl-CoA, a downstream product of the citric acid cycle, is a competitive inhibitor of acetyl-CoA binding.

Figure 16.3.4 shows an interactive iCn3D model of a structural comparison of pig citrate synthase with bound citrate (1CTS) and with bound citrate and CoASH (2CTS)



NCBI iCn3D Figure 16.3.4: Structural comparison of pig citrate synthase with bound citrate (1CTS) and with bound citrate and CoASH (2CTS). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...94UeRDsFyGTcdA

Toggle the "a" key back and forth to change from the open structure (gray) with bound citrate (1CTS) to the closed structure (cyan) after CoASH binds (2CTS).

Citrate and CoASH are the products of citrate synthase reaction, but seeing how they interact with the protein gives clues into catalysis. When both are bound, the enzyme is in closed conformation which would prevent spurious hydrolysis of the actual acetyl-CoA when the reaction proceeds to citrate formation. In just the presence of citrate, the enzyme is in open form, allowing the release of citrate as a product. The binding of the reactant oxaloacetate triggers the conversion to the closed form. NADH is reported to be an allosteric inhibitor of bacterial citrate synthases but no entries are available for the binding of NADH to mammalian enzymes.

The effects of binding acetyl-CoA on the structure of citrate synthase are shown in Figure 16.3.5:

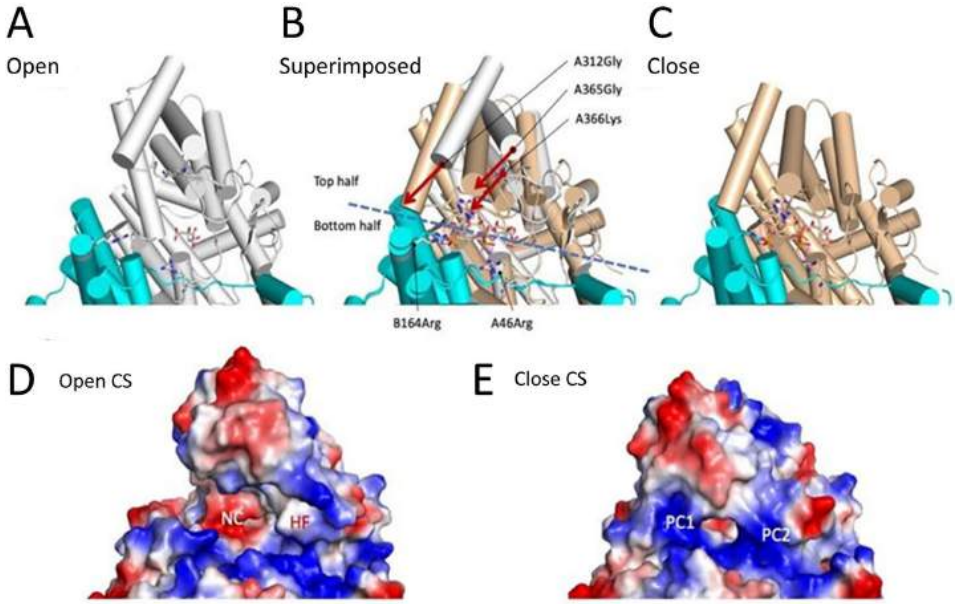


Figure 16.3.5: Effect of acetyl-CoA binding on CS structure. Omini, J., Wojciechowska, I., Skirycz, A. *et al.* The association of the malate dehydrogenase-citrate synthase metabolon is modulated by intermediates of the Krebs tricarboxylic acid cycle. *Sci Rep* **11,** 18770 (2021). https://doi.org/10.1038/s41598-021-98314-z. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Panel (**A**) shows the open format of CS (PDB ID, 1cts) in a cartoon model, with cylindrical α-helices containing a citrate molecule in a stick model. Subunits A and B are colored white and cyan, respectively. Key residues, A266Lys, A46Arg, and B164Arg, are shown in the sticks' side chains. The residues are shown in the order of the chain name, residue number, and amino acid. the dimer structure was generated using crystallographic symmetry. One active site domain composed mainly of the A-chain is shown.

Panel (**B**) shows the superimposed model between the open (1cts; white) and close (2cts; wheat) formats. The citrate molecule is at the same location with a slight rotation. The CoA molecule is present only in the closed format. The molecular domain shown can be divided into the movable upper half and the rigid bottom half. In the bottom domain, A45Arg and B146Arg are shown in the stick model. The locations of those two Cα in Arg are almost consistent between the open and closed formats. The top half domain is movable. The motion is visible as the rotation of α-helices represented by A312Gly and A365Gly, indicated by arrows. The A366Lys moves inward and forms a hydrogen bond network A366Lys (NZ):: A438COA(O8A):: B164Arg(NH1).

Panel (**C**) shows the closed format of CS (PDB ID, 2cts) with citrate and CoA molecules in stick models.

Panel (**D**) shows the surface electrostatic potential of the open format CS excluding ligands. Calculations were performed in the vacuum environment and ranged between − 71 and + 71. Red and blue represent the negative and positive potentials. The domain shown corresponds to panel A. Patches of negative charge (**NC**) and hydrophobic area (**HF**) are observed.

Panel (**E**) shows the surface electrostatic potential of the closed format CS excluding ligands. The domain shown corresponds to panel D. Patches of positive charge (PC1, PC2) are observed.

We will see how the electrostatic surface potential of citrate synthase allows it to interact with other citric cycle enzymes to form a metabolon at the end of this chapter section.

16.3.4: REGULATION OF ISOCITRATE DEHYDROGENASE

In the previous section, structures of the αγ and αβ heterodimer building blocks of the protein were described. The α subunits contain the catalytic site while the β and γ subunits were the regulatory subunits that bind allosteric effectors. Citrate and ADP allosterically activate both the $\alpha_2\beta\gamma$ heterotetramer and αγ heterodimer. They bind next to each other in the allosteric site, along with Mg^{2+}. Conformational changes in binding citrate lead to a change in and activation of the catalytic subunit α. The binding of ADP just enhances this effect.

The domain and cartoon structure of the αγ heterodimer of human NAD-IDH are shown in Figure 16.3.6.

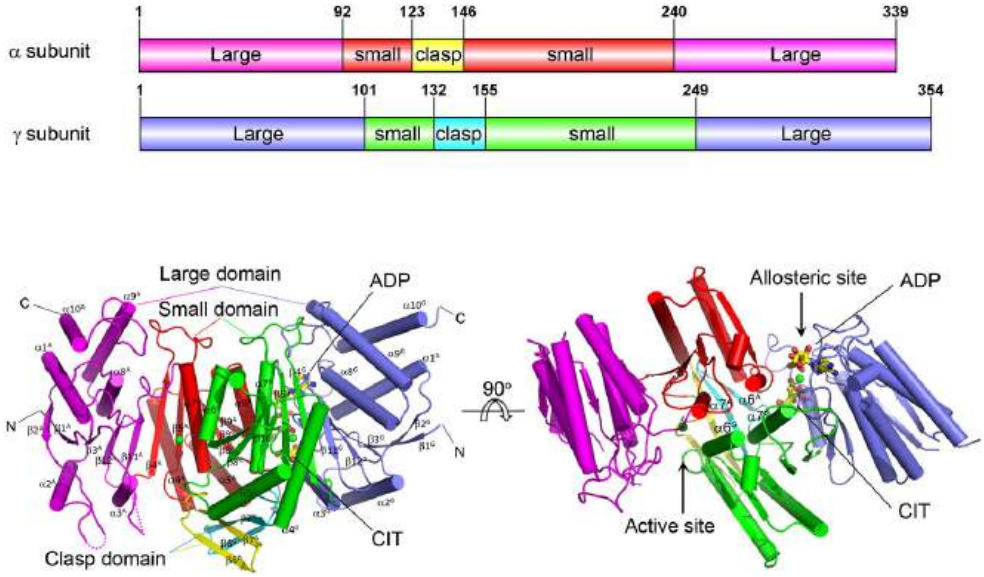


Figure 16.3.6: Domain and cartoon structure of the αγ heterodimer of human NAD-IDH. Ma, T., Peng, Y., Huang, W. *et al.* Molecular mechanism of the allosteric regulation of the αγ heterodimer of human NAD-dependent isocitrate dehydrogenase. *Sci Rep* **7,** 40921 (2017). https://doi.org/10.1038/srep40921. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/

The top panel shows the domain structure of the two monomers. The bottom panel shows two views of the dimer. The color coding is the same as in the top panel.

Figure 16.3.7 shows bound citrate and ADP in the allosteric binding site in the γ subunit of IDH.

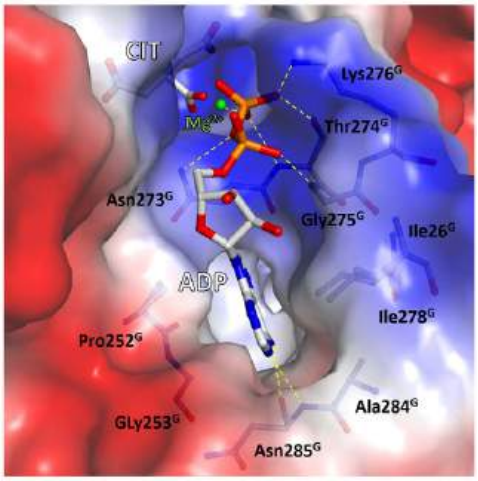


Figure 16.3.7: Bound citrate and ADP in the allosteric binding site in the γ subunit of IDH**.** Ma et al, ibid.

The color represents the electrostatic surface potential of the site with blue indicating more positive and red more negative. Note that both allosteric activators bind adjacent to each other. The binding of ADP does not change the conformation after the citrate is bound.

The proposed molecular mechanism for allosteric regulation of IDH is shown in Figure 16.3.8.

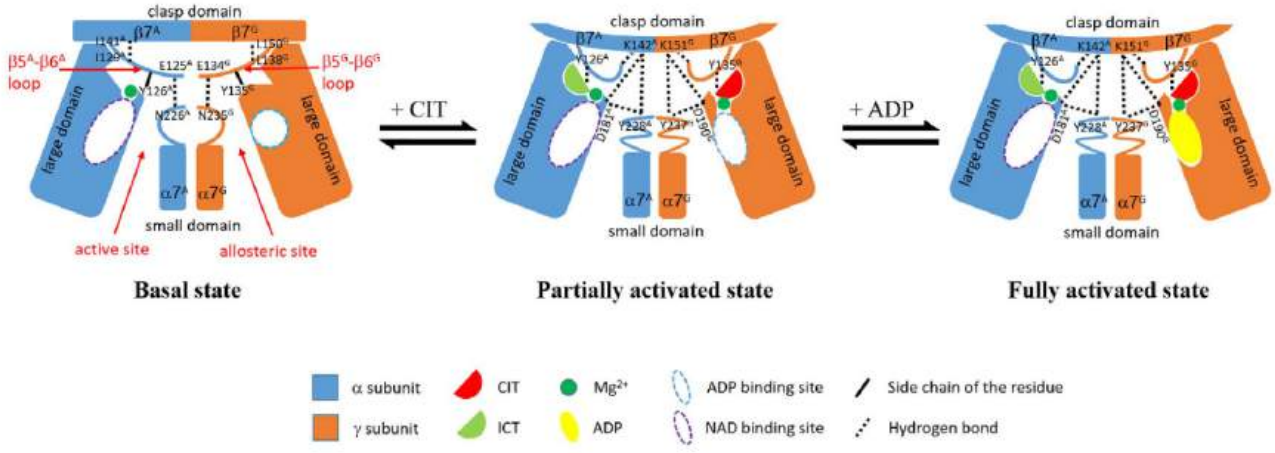


Figure 16.3.8: Mechanism mechanism of allosteric regulation of the αγ heterodimer of IDH. Ma et al, ibid.

Legend: In the absence of any activators, the active site adopts an inactive conformation unfavorable for the ICT binding, and the enzyme is in the basal state which has a high $S_{0.5,\mathrm{ICT}}$ with a low catalytic efficiency. The binding of CIT induces conformational changes at the allosteric site, which are transmitted to the active site through conformational changes of the structural elements at the heterodimer interface, including the β5–β6 loop, the α7 helix, and the β7-strand in both the α and γ subunits, leading to the conversion of the active site from the inactive conformation to the active conformation favorable for the ICT binding. Hence, the enzyme assumes the partially activated state which has a moderately decreased $S_{0.5,\mathrm{ICT}}$ (lower substrate concentration for half-maximal activity) with a moderately increased catalytic efficiency. The binding of ADP in the presence of CIT does not induce further conformational changes at the allosteric site and the active site but establishes a more extensive hydrogen-bonding network among CIT, ADP, and the surrounding residues through the metal ion, which conversely enhances or stabilizes the CIT binding. Hence, the binding of CIT and ADP together has a synergistic activation effect, and the enzyme assumes the fully activated state which has a substantially decreased $S_{0.5,\mathrm{ICT}}$ with a significantly increased catalytic efficiency.

16.3.5: REGULATION OF A-KETOGLUTARATE DEHYDROGENASE

α-ketoglutarate dehydrogenase and pyruvate dehydrogenase complex both catalyzed the oxidative decarboxylation of α-ketoacids. They use a common mechanism involving three enzymes, E1-E3, in a large complex. The regulation of α-ketoglutarate dehydrogenase activity is shown in Figure 16.3.9.

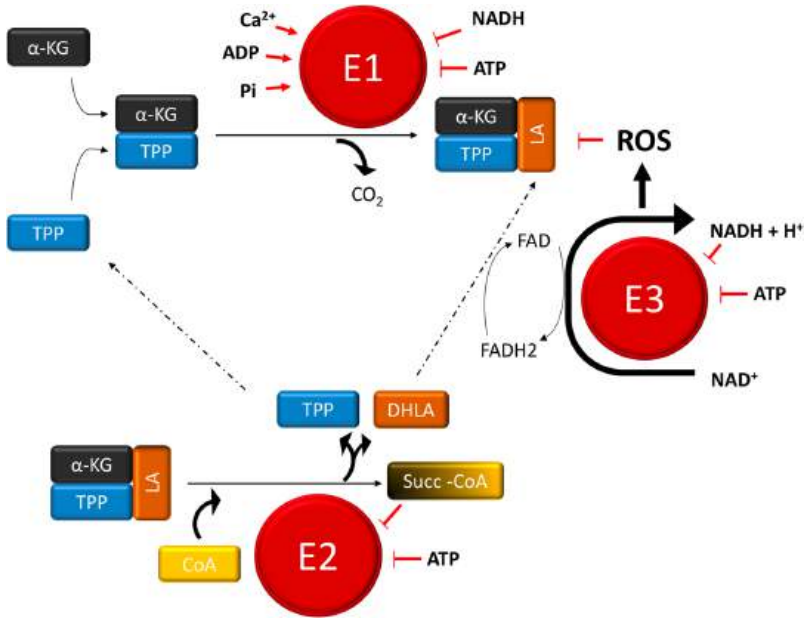


Figure 16.3.9: Regulation mechanisms of α-ketoglutarate dehydrogenase complex (α-KGDC). Vatrinet, R., Leone, G., De Luise, M. *et al.* The α-ketoglutarate dehydrogenase complex in cancer metabolic plasticity. LS and DHLA are lipoamide and dihydrolipoamide, respectively. TPP is thiamine pyrophosphate. *Cancer Metab* **5,** 3 (2017). https://doi.org/10.1186/s40170-017-0165-0. Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/),

As with the other key regulatory enzymes, α-KGDC is regulated by ATP/ADP and NADH/NAD^+ ratios. The product, succinyl-CoA inhibits the reaction at E2. The mitochondria are reservoirs for Ca^{2+} ions. These ions increase the activity of pyruvate, isocitrate, and α-ketoglutarate dehydrogenases with α-ketoglutarate dehydrogenases most affected. Calcium effects also depend on ATP/ADP and NADH/NAD^+ ratios.

16.3.6: REGULATION BY METABOLON FORMATION

Several citric acid cycle enzymes interact to form a metabolon, which allows enhance flux through pathways as substrate and products are channeled directly from one enzyme to another in the complex. This "facilitated" diffusion minimizes the dissociation of substrates/products and enhances catalysis. Three enzymes in the citric acid cycle, citrate synthase, malate dehydrogenase, and aconitase form a metabolon, as shown by chemical cross-linking and docking studies. The linkages between malate dehydrogenase, which produces oxaloacetate, and citrate synthase, which uses it, are important since concentrations of oxaloacetate are low and would produce reduced flux in the citric acid cycle if it were not channeled directly into citrate synthase. Malate dehydrogenase is also not favored to produce oxaloacetate based on standard free energy and K_{eq} values so pulling the reaction towards citrate synthase in the metabolon also helps the flux.

$$\text{L-malate} + NAD^+ \leftrightarrow \text{oxaloacetate} + NADH + H^+ \quad \Delta G^o = +7.1 \text{ kcal/mol } (+30 \text{ kJ/mol})$$

Figure 16.3.10 shows models of malate dehydrogenase (MDH) and the open and closed forms of citrate synthase (CS). Electrostatic interactions are key.

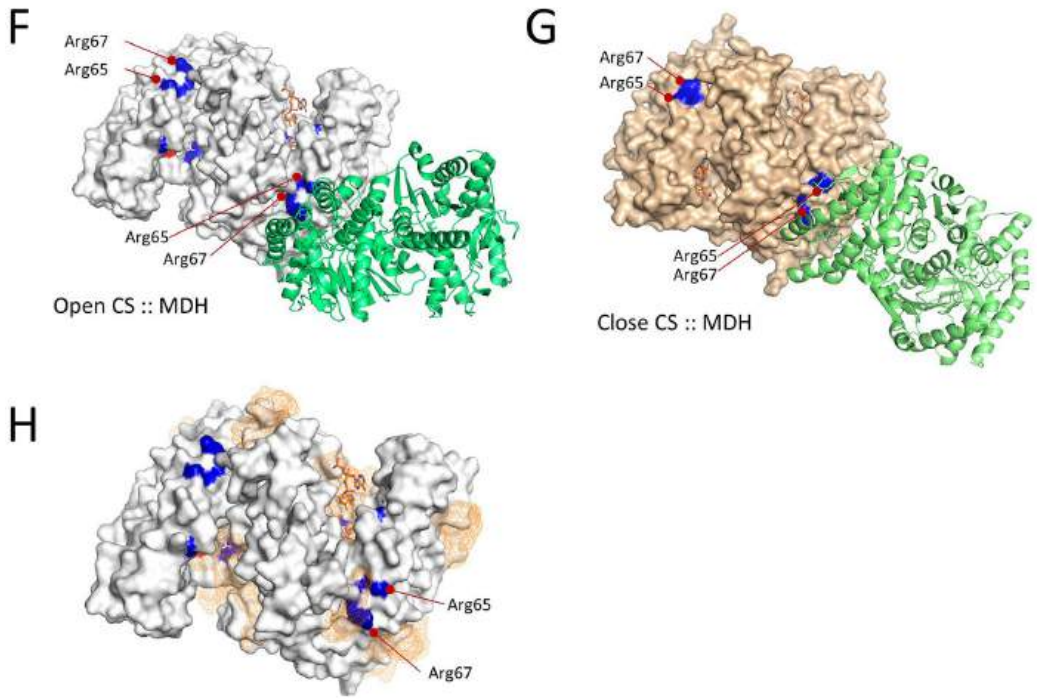


Figure 16.3.10: Models of malate dehydrogenase (MDH) and the open and closed forms of citrate synthase (CS). Omini, J., Wojciechowska, I., Skirycz, A. *et al.* Association of the malate dehydrogenase-citrate synthase metabolon is modulated by intermediates of the Krebs tricarboxylic acid cycle. *Sci Rep* **11,** 18770 (2021). https://doi.org/10.1038/s41598-021-98314-z. http://creativecommons.org/licenses/by/4.0/.Creative Commons Attribution 4.0 International License.

Panels (**F**) and (**G**) show simulated interactions between MDH (green) and CS in its open (white) form (panel F) and between MDH (green) and CS in its closed (wheat) format (panel G). The 65Arg and 67Arg residues that are involved in the MDH-CS interaction are highlighted in blue. Active site residues, His 274, His320 (blue), and Asp375 (red) were shown.

Panel (**H**) Predicted acetyl-CoA binding sites in CS apoenzyme. The white surface model of the CS apoenzyme in the open format is shown. The 274His, 320His, 65Arg, and 67Arg residues are highlighted in blue. The orange mesh indicates the positions of acetyl-CoA at the predicted binding sites. White and orange stick models indicate the citrate and CoA in the reported crystal structure, respectively.

Does the activity of citrate synthase change when it is complexed to malate dehydrogenase in a metabolon? The results of studies probing that question show that it is affected. The results of such studies are shows that it does as illustrated in Figure 16.3.11.

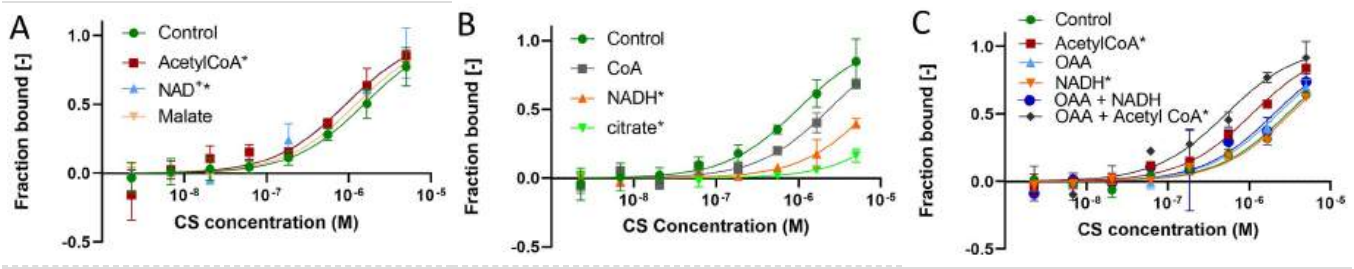


Figure 16.3.11: Effects of metabolites involved in the MDH and CS reactions on the affinity of the MDH-CS multi-enzyme complex. Omini et al, ibid.

Curves represent the response (fraction bound) against CS concentration (M). The interaction was assessed in the MST buffer (control, green) or those with 10 mM of metabolites. Error bars represent the standard deviations of three measurements. Asterisks indicate the conditions that showed significant K_d differences with no K_d confidence overlap with the control.

Panel (**A**) shows the effects of the reaction substrates of MDH and CS. The MDH-CS interaction was assessed in the presence of acetyl-CoA (red), NAD^+ (blue), or malate (brown).

Panel (**B**) shows the effects of the reaction on products of CS and MDH. The MDH-CS interaction was assessed in the presence of CoA (grey), NADH (orange), or citrate (green).

Panel (**C**) shows the effects of oxaloacetate (OAA) in combination with other CS substrates. The effects of sole substrates, acetyl-CoA (red), OAA (blue), and NADH (orange), as well as their combinations, OAA/acetyl-CoA (purple) and OAA/NADH (gray), were analyzed.

These curves clearly show how the ratio of NADH/NAD^+ affects the activity of citrate synthase when it is part of a metabolon. NAD^+ increases activity (Panel A) while NADH decreases it (Panel B).

16.4: VARIANTS OF THE CITRIC ACID CYCLE

16.4.1: CITRIC ACID CYCLE SHUNTS AND BYPASSES

Evolution has allowed variants in the citric acid cycle to produce new functionalities in organisms. Let's consider a few.

16.4.2: GLYOXYLATE SHUNT

What if you were a microorganism that has evolved to use acetate (2C) as a source (if not the sole source) of energy? Remember that the citric acid cycle is also used to generate metabolites for anaplerotic reactions (from the Greek ἀνά= 'up' and πληρόω= 'to fill'). In the full citric acid cycle, activated acetate (2C) reacts with oxaloacetate (4) to produce the 6C molecule citric acid. As the cycle continues back to oxaloacetate, 2 Cs are lost as CO_2, so in sum for 1 turn of the cycle:

oxaloacetate (4C) + acetyl-CoA (2C) → Citrate (6C) → → → succinate (4C) + 2 CO_2 (2C).

In the process, oxidation reactions occur leading to the formation of NADH and $FADH_2$ and the metabolites formed can be withdrawn for biosynthesis (OAA, αKG, and succinyl CoA).

An alternative "cycle" would be to eliminate the two decarboxylation reactions and generate a 4C molecule (succinate) and one 2C metabolite (glyoxylate), which could react with another activated acetate (acetyl-CoA) to form oxaloacetate (or more precisely malate which can be oxidized to it). Hence two acetates are required, but if you are organisms adapted to use this as a metabolic energy source, it's no problem. Here is the next reaction

oxaloacetate (4C) + acetyl-CoA (2C) → Citrate (6C) → → succinate (4C) + glyoxylate (2C);

glyoxylate (2C) + acetyl-CoA (2C) → malate (4C)

This reaction would act as a metabolic shunt, altering the flow of metabolites by bypassing part of the citric acid cycle. The shunt, which is found in many microorganisms and some plants is called the **glyoxylate shunt**, which is shown in Figure 16.4.1.

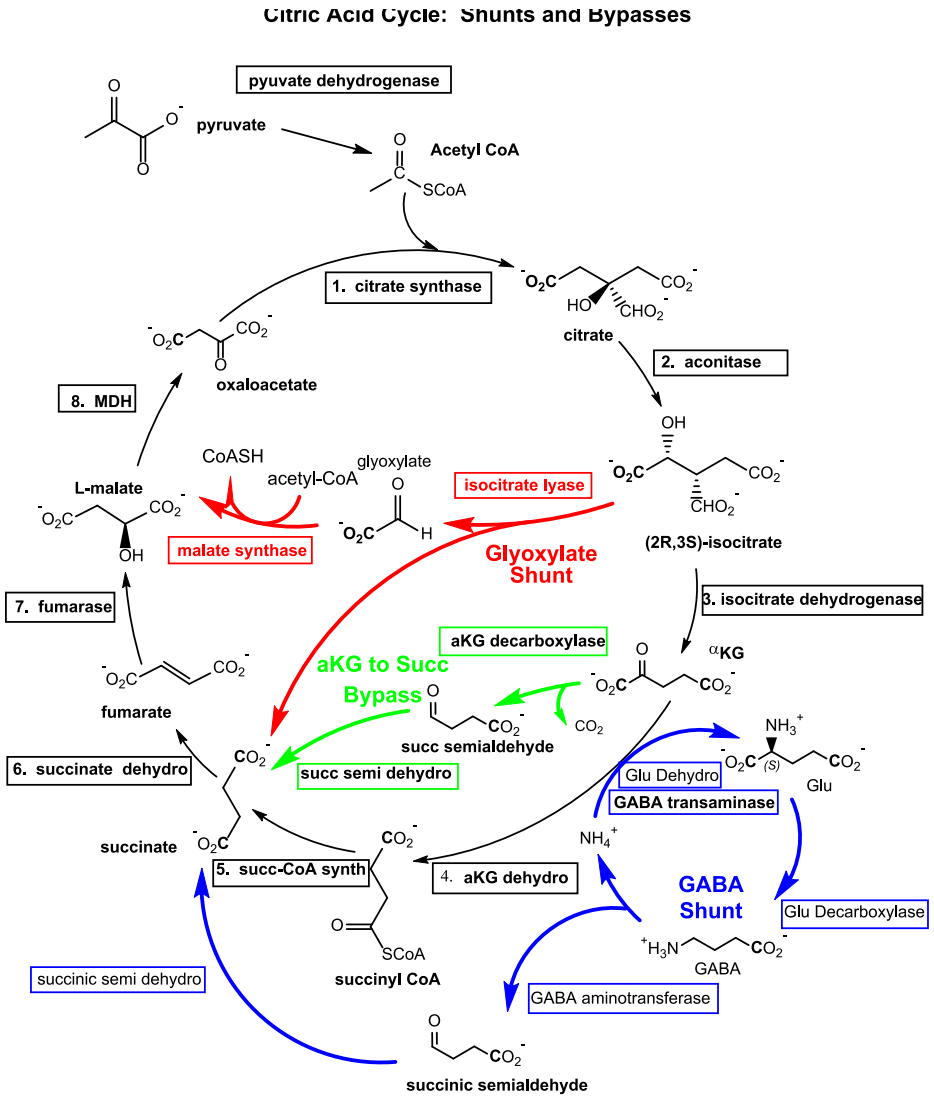


Figure 16.4.1: Citric Acid Cycle Shunts and Bypasses. (Copyright; author via source)

Why is this so cool? First, consider this observation about the citric acid cycle. One acetate, in the form of acetyl-CoA, is needed to form one oxaloacetate in one turn of the cycle, which in turn is needed in the next turn of the cycle. Hence acetyl-CoA cannot, in a net fashion, be used to synthesize oxaloacetate. We saw previously that oxalate is an intermediate in the synthesis of phosphoenolpyruvate from pyruvate in gluconeogenesis. Hence acetate can not be used to form glucose in a net fashion. (Note: pyruvate dehydrogenase is not reversible). Another way to think about this is 2Cs enter the cycle as acetyl-CoA and 2 leave as CO_2 so no net synthesis can occur. And, you are in trouble if you draw off oxaloacetate as well as α-ketoglutarate and succinate for anaplerotic reactions for biosynthesis as they are needed metabolites for a cycle that consumes and produces one of each of these will be depleted if withdrawn

Now if you are an organism (many bacteria and plants) that has the glyoxylate shunt, you have no worries. A mini cycle still occurs with the conversion of isocitrate to glyoxylate to malate and back around to oxaloacetate and citrate (for energy production) with the net production of 1 succinate (for biosynthesis). The only Cs "lost" from the mini cycle are in the form of succinate so they are not lost as they are as 2 CO_2s in the citric acid cycle since they are used for biosynthesis.

The next reaction for the glyoxylate cycle is:

$$2 \text{ Acetyl-CoA} + 2 \text{ NAD}^+ + 2 \text{ H}_2\text{O} \rightarrow 2 \text{ succinate} + 2 \text{ NADH} + 2 \text{ CoASH} + 2 \text{ H}^+ .$$

In plants, this reaction occurs in organelles called Glyoxysomes. The succinate can be converted through its continuation through the citric acid to oxaloacetate which can then form glucose and other carbohydrates through gluconeogenesis. Hence organisms that have the glyoxylate shunt (or cycle) can synthesize carbohydrates in a net fashion from acetate, which can derive from fatty acid degradation as we will see in a subsequent chapter.

Now to operate the glyoxylate cycle efficiently, you do need to activate the carboxyl end of acetate. Plants and other organisms that can grow on acetate can produce acetyl-CoA needed for the above pathways from the following reaction catalyzed by acetyl-CoA synthase (i.e.

not pyruvate dehydrogenase)

$$\text{Acetate} + \text{CoASH} + \text{ATP} \rightarrow \text{Acetyl-CoA} + \text{AMP} + PP_i$$

Note that one phosphoanhydride bond is cleaved and one thioester bond is formed. These molecules are both high energy compared to their hydrolysis products so this reaction alone is not thermodynamically favored in any significant way (remember that there is no such thing as a high energy bond). The hydrolysis of pyrophosphate (PP_i) drives this reaction forward. An analogous reaction drives the creation of peptide bonds in protein as that reaction also uses ATP to create an activated mixed anhydride from the free carboxyl end

16.4.3: GABA SHUNT

Now let's move to the brain where there is lots of chemistry happening as you read this page. Relatively high concentrations of key neurotransmitters are required for neural function and these need to be maintained against metabolic losses. Glutamate is the principal excitatory neurotransmitter and GABA (δ-aminobutryic acid) is the principal inhibitory neurotransmitter. Both are maintained at high concentrations and are principal players in the GABA shunt. Figure 16.4.2 shows a part of the large pathway shown above but in a more expanded form.

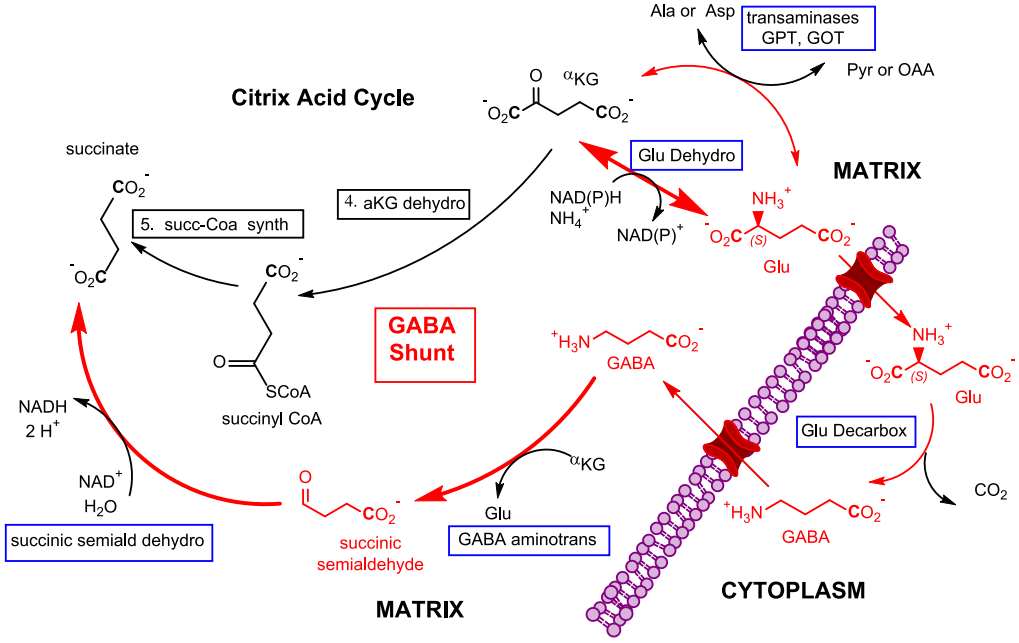


Figure 16.4.2: Expanded view of GABA shunt. (Copyright; author via source)

GABA is formed in the cytoplasm by decarboxylation of Glu by the enzyme glutamate decarboxylase, which appears to be expressed predominantly in neural tissue. It is metabolized by GABA transaminase but only if the compound from which the shunt starts, α-ketoglutarate, is present abundantly. This conserves the supply of GABA in the neuron. Stated in another way, the breakdown product of GABA, succinic semialdehyde, is formed only if GABA's precursor is present. Inhibitors of GABA aminotransferase are used to treat epilepsy. Another function of the shunt is that it effectively allows glutamate to loop into the cycle.

16.4.4: A-KETOGLUTARATE ↔ SUCCINATE BYPASS

In anticipation of the section below, one could ask the following question: What happens when a step in the cycle is impaired or missing? Cyanobacteria, a key player in atmospheric O_2 production and drawn down of natural and anthropogenic atmospheric CO_2, were thought to have an incomplete citric cycle as they lack α-ketoglutarate dehydrogenase. They work around this issue by converting α-ketoglutarate to succinate directly using two enzymes, alpha-ketoglutarate decarboxylase, and succinic semialdehyde dehydrogenase, as shown below. The first enzyme catalyzes a non-oxidative decarboxylation of the substrate to succinic semialdehyde. This bypass is shown in Figure 16.4.3. In addition, cyanobacteria also use the GABA shunt as a bypass for α-ketoglutarate dehydrogenase.

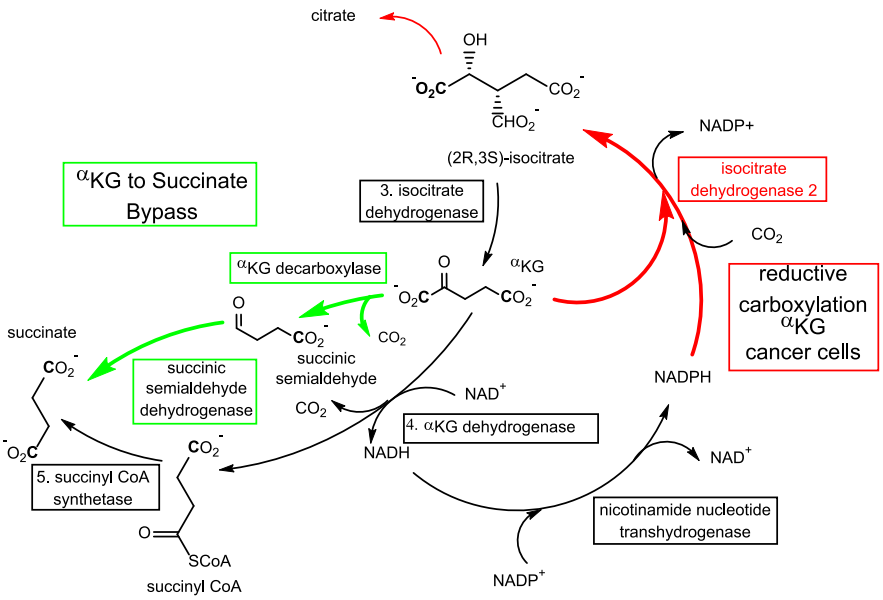


Figure 16.4.3 : α-ketoglutarate to succinate bypass. (Copyright; author via source)

Here is yet another example. Each member of the cycle is an important member. Let's consider the first one, citrate. It is formed in the key step coupling the output of anaerobic metabolism of glucose, pyruvate) with the formation of citrate in the citric acid cycle. What happens if aerobic metabolism is impaired, as it would in hypoxic and anoxic conditions, or if mitochondria function is compromised or in disease states such as cancer? Might there be another way to form citrate? Turns out that there is. and it centers on α-ketoglutarate again. It is illustrated in the above figure.

Lets's consider cancer which is characterized by a state of rapid proliferation of cells. This requires both energy and metabolic precursors for the biosynthesis of carbohydrates, lipids, and proteins. The citric acid cycle offers both. To accommodate this demand for the efflux of citric acid cycle intermediates for reductive biosynthesis, large fluxes of glutamate (derived from abundant glutamine) into the cycle are used in the form of α-ketoglutarate. In another anaplerotic reaction, citrate can be cleaved by citrate lyase to form acetyl-CoA which can be used for fatty acid synthesis needed by rapidly proliferating cells. To replenish the citrate, cancer cells convert α-ketoglutarate by reductive carboxylation to isocitrate by isocitrate dehydrogenase 2, which uses NADPH for the reduction reaction. That NADPH comes from the enzyme nicotinamide nucleotide transhydrogenase, which can interconvert matrix pools of NADH and NADPH using the collapse of a protein gradient across the mitochondria inner membrane (which we will study in a subsequent chapter).

$$H^+_{(in)} + NAD^+ + NADPH \leftrightarrow H^+_{(out)} + NADH + NADP^+$$

Note that both oxidative (α-ketoglutarate to succinyl CoA) and reductive reactions (α-ketoglutarate to isocitrate) occur in this process.

16.4.5: VARIANTS OF THE TCA CYCLE IN MICROORGANISMS

It's very easy to be anthropocentric in constructing a biochemistry text as many who take the course are interested in human medicine. Since a human is an ecosystem of organisms with an expansive microbiome on their skin and in their gut, even from a human health perspective, it is important to compare the same pathway in different organisms. It's also important to understand our role as a small part of a vast biosphere where our survival depends on other organisms, large and small.

In that light, let's consider the citric acid cycle of other organisms. It seems that most organisms have the anaerobic and universal glycolytic pathway. How about the aerobic citric acid cycle? These days of single-cell genomic analysis make it simple in principle to analyze the citric acid cycle genes of any organism. Variants of it are found in generally all aerobic organisms and even some anaerobic one. Some subtle differences exist between eukaryotic and prokaryotic organisms. NAD^+ is used as a substrate in the mammalian form of isocitrate dehydrogenase while prokaryotes use $NADP^+$. An NAD^+-dependent malate dehydrogenase is used in mammals while some prokaryotes use a different enzyme, a $NADP^+$-dependent malate-quinone oxidoreductase. Lastly different enzymes (and unfortunately with varying names) are used to convert succinyl-CoA to succinate. Plants and fungi use ADP as a substrate, mammals have two different enzymes often named different and for the reverse reaction as succinate-CoA ligase (ADP forming) and succinate-CoA ligase (GDP forming).

Analyzes shown that few bacteria have complete cycles. Of those with incomplete cycles, the early steps are least conserved, while the latter are most conserved. The cycle is used not only for the oxidative production of energy but also the generation of metabolites (α-ketoglutarate, oxaloacetate, and succinyl-CoA), which are pulled from the cycle for biosynthesis. Autotrophs that don't have a complete cycle can make those products from pyruvate. Consider α-ketoglutarate. Some make it through the clockwise oxidative reaction from citrate

to α-ketoglutarate while some methanogenic Archaea make it through counterclockwise reduction reactions. Comparative genomic analyses suggest that the citric acid cycle probably arose from a "linear" oxidative pathways leading to α-ketoglutarate and a reductive one leading to succinyl CoA. Knowing the pathways in individual microorganisms can assist in the rational drug design of new antibiotics. Figure 16.4.4 below, adapted from Huynen et al, shows a geometric analysis diagram of the citric acid cycle and variants in other organisms.

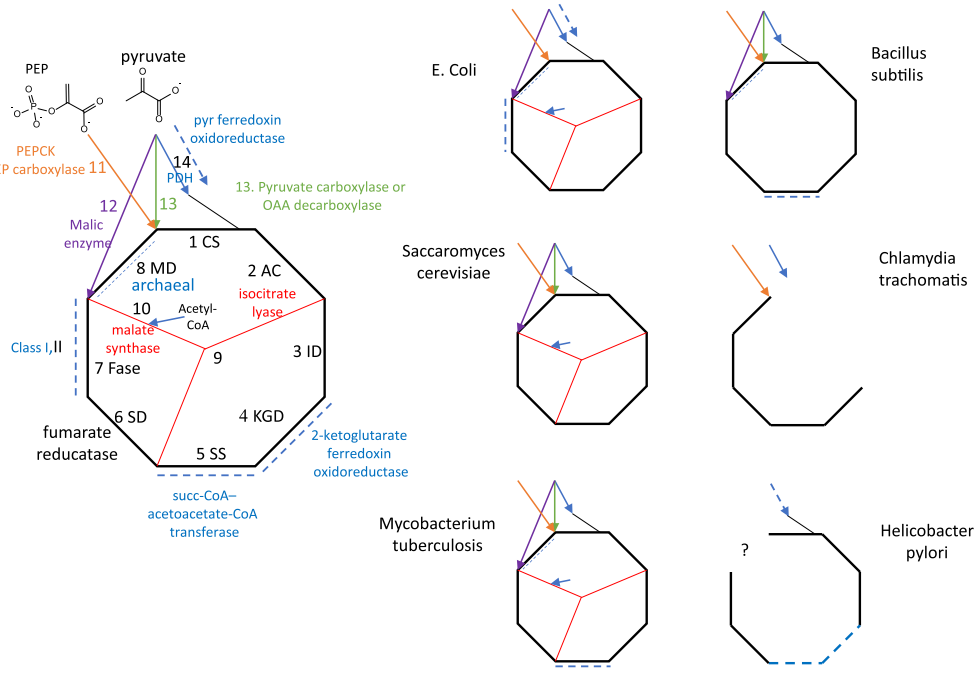


Figure 16.4.4: Geometric analysis diagram of the citric acid cycle and variants in other organism. (Copyright; author via source)

16.4.6: THE A-KETOACID PATHWAY - A PRIMORDIAL, PREBIOTIC ANABOLIC "TCA-LIKE" PATHWAY

You might be wondering after studying the complexity of pyruvate dehydrogenase and the citric acid cycle how the cycle might have originated. As discussed above, clues come from comparative genome analysis of genes encoding the enzymes for the reactions in a multitude of organisms. By looking at evolutionary changes in genes and functions for these reactions, one can obtain some ideas of the first enzymes that evolved in the pathway. But what about an abiotic pathway that might have arisen before the first biological one? Stubbs et al have shown that the very simple molecules glyoxylate and pyruvate can react in mild aqueous conditions (pH 7 in 0.5 M phosphate buffer heated to 50 °C) to form, in a single pot, α-ketoacids similar to present citric acid cycle intermediates. From a kinetic perspective, they were formed in the reverse, counterclockwise reductive direction compared to the citric acid cycle (clockwise, oxidative). No metals or "enzymes" were needed for the transformations. Glyoxylate acts as a reducing agent and is simultaneously the source of carbon atoms. Once these α-ketoacids form, they could be theoretically converted by transamination reactions to amino acids as they should be characterized by a $K_{eq} \approx 1$ and a $\Delta G^0 \approx 0$.

The citric acid cycle is filled are carboxylic acids. Only the α-ketoacids (α-ketoglutarate and oxaloacetate) can readily form carbanions at the alpha carbons as their Hs are slightly acidic due to resonance stabilization of the carbanion through and enolate tautomer. The others are electron-rich so they are less acidic at the alpha carbons, discouraging deprotonation in the absence of harsh catalysts. So perhaps alternative functional groups were present in abiotic precursors. Perhaps they contained α-ketoacids similar to α-ketoglutarate, oxaloacetate, and of course pyruvate. In addition, α-ketoacids have an electrophile in the form of the carbonyl carbon form C-C bond formation.

Figure 16.4.5 shows the reductive, counter-clockwise pathway they determined. Note that in both the glyoxylate "cycle" which has also been proposed as an abiotic pathway and the α-ketoacid reductive pathway, no CO_2 is lost in any step. That suggests that the actual decarboxylation steps, although often favored thermodynamically, required the development of catalysts. The two α-ketoacids substrates in this pathway, glyoxylate (2Cs) and pyruvate (3Cs) are the smallest and most likely formed acids available.

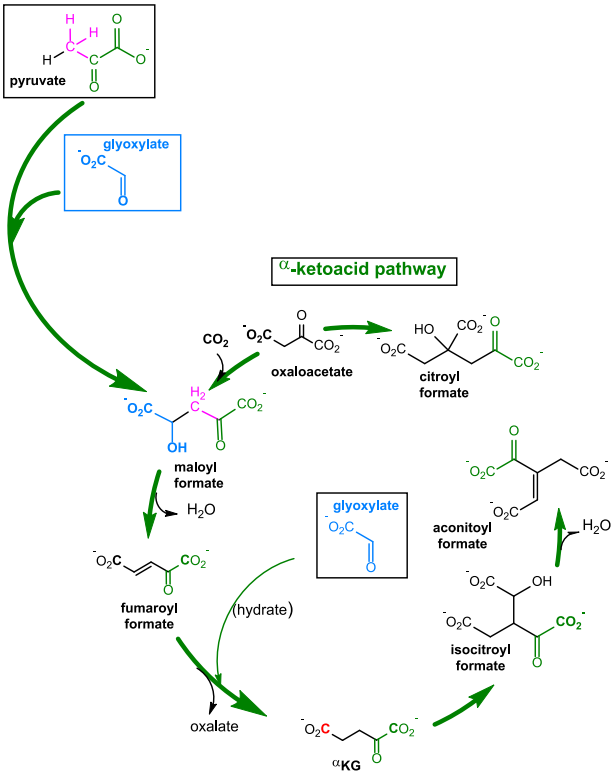


Figure 16.4.5: Proposed prebiotic alpha-keto acids pathway. (Copyright; author via source)

Figure 16.4.6 is an animation that shows the entire citric acid cycle (catabolic, oxidative, running clockwise), the glyoxylate shunt which bypasses many steps in the citric acid cycle, and the proposed prebiotic anabolic alpha-ketoacid pathway (bold green arrows) which essentially runs counterclockwise and is reductive.

Figure 16.4.6: Animation of the citric acid cycle (catabolic, oxidative, clockwise), the glyoxylate shunt (which bypasses many steps in the citric acid cycle), and the proposed prebiotic anabolic alpha-ketoacid pathway (bold green arrows). (Copyright; author via source)

CHAPTER OVERVIEW

17: FATTY ACID CATABOLISM

Fatty acid catabolism is the mechanism by which the body accesses energy stored as triglycerides.

17.1: Digestion, Mobilization, and Transport of Fats

17.2: Oxidation of Fatty Acids

17.3: Ketone Bodies

Thumbnail: Idealized representation of a molecule of a typical triglyceride, the main type of fat. Note the three fatty acid chains attached to the central glycerol portion of the molecule. (Public Domain; Benjah-bmm27 via Wikipedia)

17.1: DIGESTION, MOBILIZATION, AND TRANSPORT OF FATS

17.1.1: INTRODUCTION

In this chapter, we will discuss the breakdown of fats to produce ATP. Most of the available chemical energy stored in fats is in the form of highly reduced fatty acids. One form of fatty acid-containing lipids comes from our diet, which includes triacylglycerols (TAGs) and membrane lipids. Fatty acids, mostly in the form of TAGs, are moved in the circulation in the form of large lipid-carrying vesicles called **lipoproteins**. The lipids can be imported into cells for storage and energy use.

Another source of fatty acids comes from those synthesized within cells from the small molecule acetyl-CoA. Fatty acids are synthesized by an enzyme complex called **fatty acid synthase**. This enzyme is found most prevalently in adipose (fat) tissue and the liver. In addition, it is significantly expressed in the brain, lungs and mammary gland.

TAGs, stored in lipid droplets, are found in most cells. The major tissue used for TAG storage is adipose (fat) tissue, whose volume consists mostly of lipid droplet(s). Given the large mass of muscle tissue, there is also a considerable amount of TAGs stored as small lipid droplets in muscle cells. However, skeletal muscle cells don't synthesize fatty acids. They have the genes for fatty acid synthase but do not transcribe it into RNA so no enzyme is made. They can however import them for catabolism. Muscle TAGs can be oxidized for energy, especially during endurance excercise.

TAGs are also stored in the liver in lipid droplets. The liver also assembles lipoproteins, which are released by the liver. Excess TAGs are stored in the liver in various diseases including alcoholism and also in nonalcoholic fatty liver disease (NAFLD), which can progress into nonalcoholic steatohepatitis (NASH), a much worse disease.

WHITE AND BROWN ADIPOSE TISSUES

There are two major forms of triacylglycerol-storing fat tissues, **white adipose tissue (WAT)** and **brown adipose tissue (BAT)**. The more abundant WAT store triacylglycerols in one large lipid droplet in the cell and release fatty acid in processes controlled by the hormones insulin and epinephrine. This simple role can mask the fact that adipose tissue is a major player in the endocrine system and is involved in cell signaling and systematic control of metabolism. Adipose tissue releases the key hormones **leptin** and **adipisin**, which in analogy to the hormones and signaling agents released by immune cells (cytokines, lymphokines), can be called **adipokines**. They also secrete other adipokines including tumor necrosis factor α (TNF-α), adiponectin, and resistin.

In contrast, BAT is specialized not to store and release fatty acids. but rather to oxidize fatty acids in ways that maximize heat production, preventing hypothermia. They have multiple smaller lipid droplets, displaying a larger surface area for lipolysis, the hydrolytic cleavage of fatty acids from the TAGs. A particular mitochondrial protein, uncoupling protein 1 (UCP1), is expressed in brown but not white adipocytes, allowing a "futile" metabolic cycle leading to dissipation of heat instead of ATP synthesis. The relative abundance of white and brown adipocytes is critical in diseases like obesity and type 2 diabetes. BAT tissue is especially important in small animals (and in newborns) for thermoregulation. For smaller organisms, the surface area to volume ratio is greater than the ratio for larger animals, allowing more heat loss. The ratio of surface area AS per volume V for a sphere is given by:

$$\frac{\mathrm{SA_{sphere}}}{\mathrm{V_{sphere}}} = \frac{4\pi r^2}{\left(\frac{4}{3}\right)\pi r^3}$$

Let's assume an average large adipocyte is a sphere of diameter 100 uM. Compare this to a large sphere with a 100 times greater diameter (10,000 uM). The smaller sphere has a 1/100 of the diameter but a surface area/volume ratio 100 times greater than the large sphere.

An intermediate type of fat tissue consists of "bright" adipocytes. White adipocytes can be coaxed to differentiate into bright and brown cells, which could be an obesity treatment.

In this chapter section, we will follow the fate of fatty acids from dietary lipids which are cleaved from TAGs, loaded into chylomicrons, a lipoprotein assembled in the small intestine, secreted into the circulation, and taken up by the liver. The liver can store the incoming fatty acids in TAGs or release them back into the circulation in the form of another lipoprotein, very low-density lipoproteins (VLDL). Circulating VLDL can exchange lipids with other circulating lipoproteins. Lipoproteins deliver fatty acids to cells after interaction with the cell surface of target cells and either cleavage of TAGs by cell membrane-associated enzyme lipase, followed by fatty acid uptake, or by endocytosis of lipoproteins into the cells.

17.1.2: LIPOPROTEINS

Before we look in more detail at the individual steps in lipids processing, let's look at the different lipoproteins, the large vesicular structures that allow the transport of fats, very insoluble molecules, in the circulation. Unlike normal liposomes or vesicles that have a lipid bilayer

surrounding an interior aqueous compartment, lipoproteins have only a single monolayer of phospholipids encapsulating a nonaqueous interior filled with TAGs, cholesterol, and cholesterol esters. The protein part of the lipoprotein consists of one or several proteins bound on the outside of the particle. The proteins help solubilize the lipoprotein, confine its size, and prevent aggregation of the lipoproteins, which would be a health risk. The structure of a typical lipoprotein is shown in Figure 17.1.1.

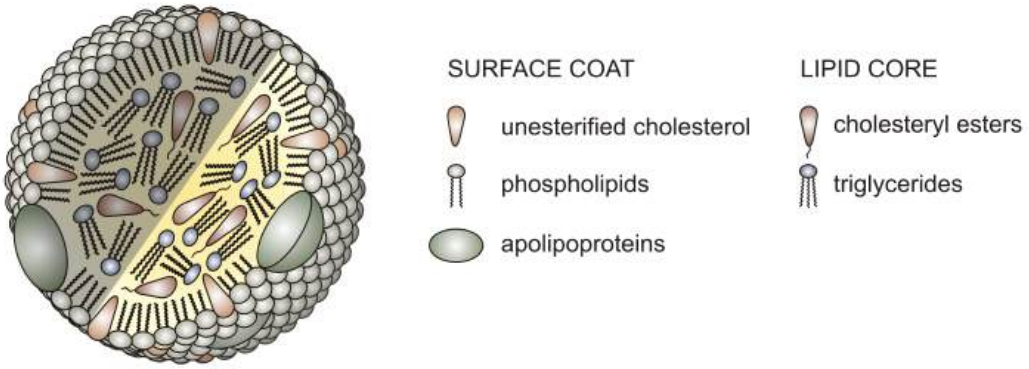


Figure 17.1.1: Lipoprotein particles are composed of a lipid core containing cholesteryl esters and triglycerides, and a surface coat of phospholipids, unesterified cholesterol, and apolipoproteins. (CC BY-SA 3.0 Unported; AntiSense via Wikipedia)

Lipoproteins are classified based on density. The lowest density chylomicrons are the largest with the most lipids (mostly TAGs) in their interior compartment. Very large density lipoproteins (VLDL), intermediate density (IDL), low density (LDL) and high density (HDL) have decreasing size, less encapsulated lipids, and increasing density. The relative sizes are shown in Figure 17.1.2.

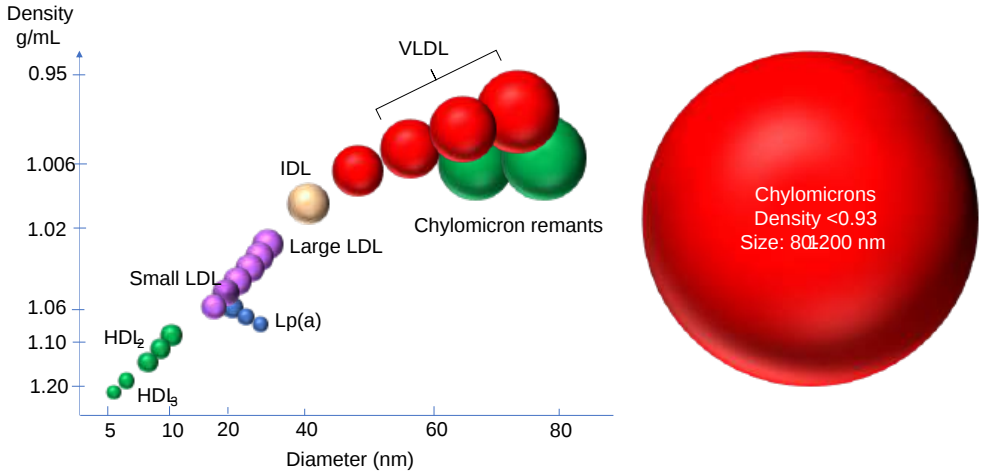


Figure 17.1.2: Modified and used with permission from Endotext.org.

Lipoproteins (except chylomicrons) could be classified as nanoparticles, which typically vary in size from 1-100 nm. Larger lipoproteins as well as chylomicrons form emulsions in the blood, much as milk (also cloudy) is an emulsion of lipid/protein particles. The serum of people with high levels of lipids (hyperlipidemia) can look milky white, especially after eating foods rich in TAGs, when levels of chylomicrons are very high. Figure 17.1.3 shows the blood of a patient with hyperlipidemia after the addition of EDTA (which binds Ca^{2+} and prevents clotting) that has settled (without centrifugation). The milky white plasma on top (lower density) most likely has high concentrations of chylomicrons and/or LDL. The lower layer contains mostly red blood cells.

Figure 17.1.3: This is a standard EDTA (cell count) tube, as may be used to obtain a full blood count or iron studies. The tube was set upright for four hours - it was not centrifuged. (Public Domain; Mark-shea via Wikipedia)

No x-ray structures of lipoproteins are available. However, a structure of a nascent HDL particle (3k2s) has been determined by small-angle neutron scattering. Figure 17.1.4 shows an interactive iCn3D model of it.

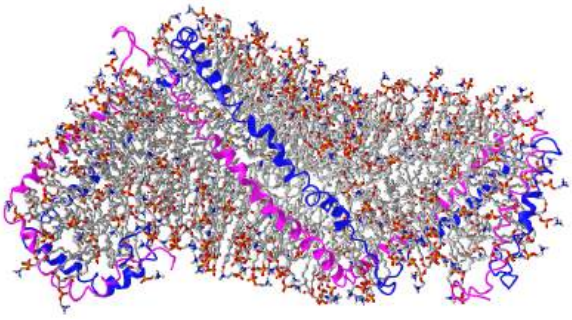


Figure 17.1.4: Nascent HDL (3k2s) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...fF85h11SpeYJg6

The major protein in HDL, a lipoprotein that protects against cardiovascular disease, is apolipoprotein A-I (apoA-I). Figure 17.1.4 shows that it adopts an antiparallel double superhelix as it wraps around the nascent HDL. The more hydrophobic surfaces of apo A-I are oriented inward allowing interactions with hydrophobic lipids in the core. It is probably prototypical for nascent lipoproteins. It will give you an idea of how proteins wrap around the outside of the particle. Mature lipoproteins are most likely spherical. This nascent HDL in the model contains 200 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholines (POPC) molecules, 20 cholesterol, and a single copy of apolipoprotein A-I (apoA-I).

Figure 17.1.5 below shows a cartoon image of VLDL, assembled from lipids synthesized/taken up by and released from the liver, and chylomicrons, assembled from dietary fats and released from enterocytes in the small intestine (size of the lipoproteins is not to scale). Note that VLDL has one copy of Apo B-100 while chylomicrons have one copy of Apo B-48.

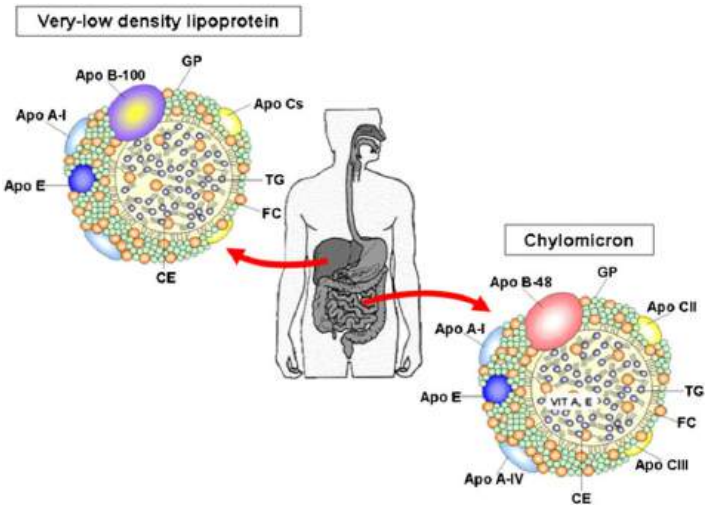


Figure 17.1.5: Emile Levy, Journal of Lipid Research. 56, 945-962, 5/2015. DOI:https://doi.org/10.1194/jlr.R052415. Creative Commons Attribution (CC BY 4.0)

All lipoproteins, except HDL, are members of the **B**eta-lipoprotein family as they contain an apo-**B** protein. The liver synthesizes apo B100, which becomes a permanent part of VLDL (i.e it is not exchangeable with other lipoproteins) and its metabolic derivatives so any lipoprotein containing apo B100 arose from the liver. Other proteins on lipoproteins are exchangeable. In contrast, enterocytes in the small intestine produce apo B-48 (48% the size of apo B100) so this protein marks the lipoproteins (chylomicrons and chylomicron remnants) that were assembled in the small intestine. Apo B100 has over 4500 amino acids and a molecular mass of 555K. The gene for the intestinal apo B48 is the same as for apo B100 except that it has a premature stop codon which leads to the shorter truncated apo B-48.

The apoproteins bind to specific receptors on cells which may allow the uptake of the lipoprotein. For example, the LDL receptors bind to the apo B-100 protein on a region removed from the apo B48 protein of chylomicrons. It also binds ApoE, which is found mostly predominately on HDL and VLDL but some are present in LDL. The LDL receptors have also been called the ApoB/ApoE receptor.

Over 90% of the apoB-containing particles in circulation are LDL. In addition, chylomicrons are present in circulation only after eating. Some apoproteins can act as cofactors and inhibitors for lipoprotein processing.

LIPOPROTEINS AND CARDIOVASCULAR RISK

High concentrations of LDL are associated with increased cardiovascular risk. Chylomicron levels, given their transient and lower concentration levels, do not pose a health risk unless the enzyme required to remove fatty acids from them, lipoprotein lipase, is missing or defective, or if another apoprotein component, apo CII, which mediates the interaction with lipoprotein lipase, is missing. LDL-C (a term used to describe the total cholesterol in LDL particles which is routinely measured in clinical labs) can be lowed by a healthy diet centered around plant food). Drugs like statins, which decrease endogenous cholesterol synthesis, also remarkably lower LDL levels and decrease cardiovascular risk.

However, another protein, lipoprotein, also called Lp(a) or LP little a is an independent cardiovascular risk factor. Its blood concentration is regulated by genetics and not by diet. These particles contain, in addition to apo B100, apo (a), a protein that has a very unique repeating structure (up to 40 times) called a **kringle**, which is also found in some proteins involved in the blood coagulation system. People whose genes encode apo (a) with the fewest number of kringles express lots of that protein and their Lp(a) particles are smaller. This confers a greater cardiovascular risk compared to those expressing proteins with a large number of kringle domains. Figure 17.1.6 (from Amgen) shows models of Lp(a) with different numbers of repeating kringle (kinked) domains.

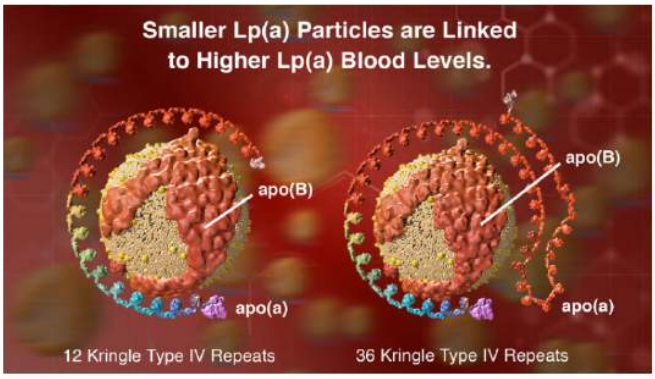


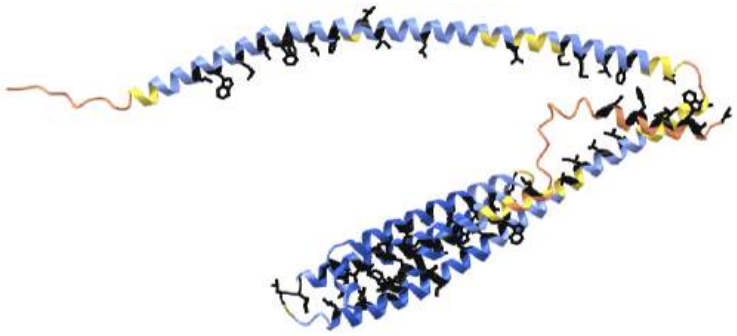
Figure 17.1.6: https://www.amgenscience.com/feature...-lipoproteina/

From a biochemical perspective, it is interesting to explore the differences in apolipoprotein binding to a single-leaflet encapsulated lipid nanoparticles compared to the interaction of peripheral and integral membrane proteins with intact bilayers (which we studied in Chapter 12.1). As mentioned above, the more nonpolar surfaces of apo AI in HDL are oriented inward toward the nonpolar lipid core. Presumably, apo B proteins in chylomicrons and LDL also wrap around the entire lipid surface.

The major organizing scaffolding protein of HDL is apoA-I (see iCn3D model above), It presumably plays a role similar to apoB in chylomicrons and LDL, but it is exchangeable. (Note: ApoA-I is also found in chylomicrons.) It is also a cofactor for the enzyme lecithin:cholesterol acyl transferase (LCAT), which effectively converts free cholesterol in the single bilayer into esterified cholesterol esters within HDL. In its apo-form, it also interacts with the cell surface transporter ATP-binding cassette A1 (ABCA1), which plays a role in the assembly of HDL particles. HDL also has apo C and apo E proteins, all of which are exchangeable.

It must be difficult to determine the structure of lipoproteins given their heterogeneity and size. The apoproteins have hydrophobic surfaces that promote self-association and aggregation. Apolipoproteins in the A, C, and E classes have repeating amphiphilic helices which imbed to some degree in the lipid particles. In addition, the proteins have a significant disorder and can adopt many bound conformations.

Figure 17.1.7 shows an interactive iCn3D model of the AlphaFold predicted structure of human ApoA (**P06727**)



NCBI iCn3D Figure 17.1.7: AlphaFold predicted structure of human Apo E (**P02649).** Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...5GtKt3QWaTZ519

The blue cartoon color represents high certainty in the AlphFold predicted structure while yellow to orange represents low certainty. The hydrophobic side chains are shown as sticks and help illustrate the amphiphilic nature of the structure.

Figure 17.1.8 shows an interactive iCn3D model of the AlphaFold predicted structure of human ApoE (**P02649**)

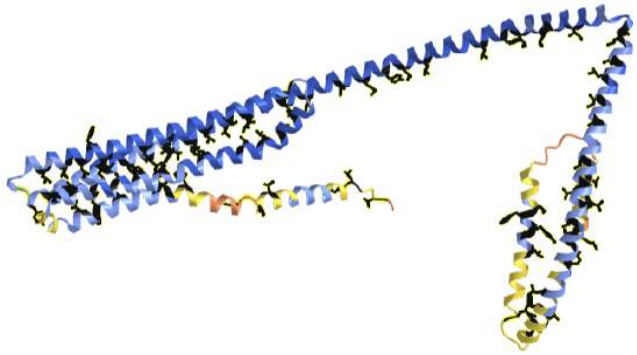


Figure 17.1.8: AlphaFold predicted structure of human Apo A4 (**P02649).** Copyright; author via source). Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...jSw8wSX1MtpcB8

The blue cartoon color represents high certainty in the AlphFold predicted structure while yellow to orange represents low certainty. The hydrophobic side chains are shown as sticks and help illustrate the amphiphilic nature of the structure.

The exchangeable apolipoproteins have similar genetic sequences (four exons and three introns), as well as similar amino acid sequences. They have 11-mer amino acid tandem repeats and some (A-I, A-IV) have 22-mer tandem repeats. These repeats form amphiphilic helices as determined by sequence analysis. The first amino acid in the amphiphilic helix is often positively charged and a negative one is often found in the middle. Proline, a helix breaker, is often, but not always found between the helices.

Figure 17.1.9s shows the primary sequence of apoA-I. An 11-mer repeat is shown in yellow highlight. The other highlighted stretches (different colors) are 22-mer repeats. Note that the repeats are not of identical sequences but rather of sequences that can form amphiphilic helices (i.e. secondary structure repeats).

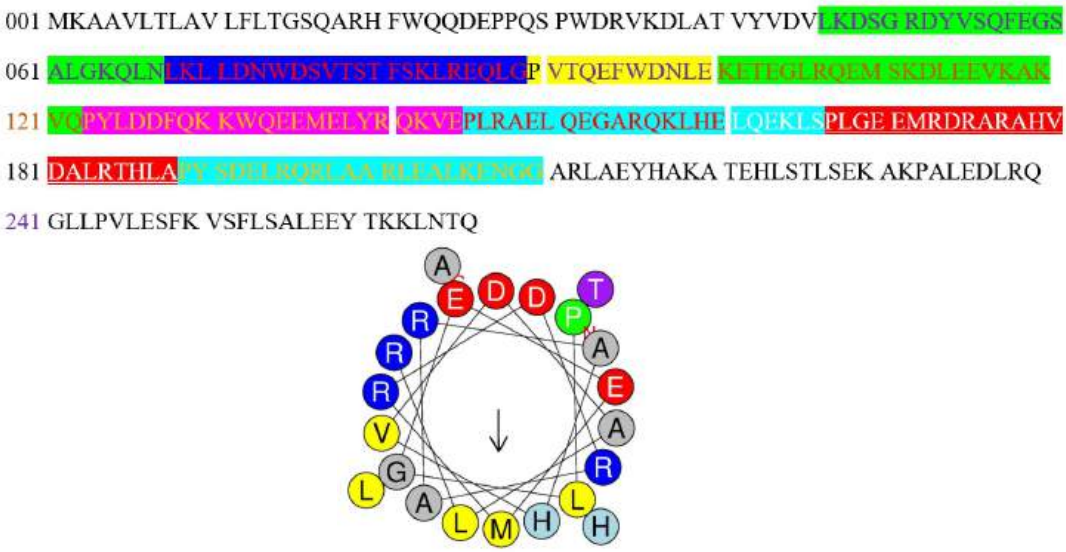


Figure 17.1.9.belows shows the primary sequence of apoA-I. An 11-mer repeat is shown in yellow highlight. The other highlighted stretches (different colors) are 22-mer repeats. Note that the repeats are not of identical sequences but rather of sequences that can form amphiphilic helices (i.e. secondary structure repeats).

The bottom part of Figure 17.1.9 shows a helical wheel projection (using Heliquest) of the red-highlighted 22-mer repeat. The arrow shows the hydrophobic moment with the arrowhead pointing to the more nonpolar face. The particular amphiphilic helix shown may or may not facilitate the binding of the bound conformation of the protein.

It follows that the relative areas of the hydrophilic and hydrophobic faces in the amphipathic helixes influence the lipid-associating properties of the exchangeable apolipoproteins. Another factor that might influence the lipid-binding ability of exchangeable apolipoproteins and which has not been studied in detail so far is the arrangement of tandem repeating amphipathic helixes with respect to one another.

Actual amphiphilic helices would bind to the membrane in a parallel fashion with the nonpolar face anchoring the protein to the lipid surface. Other experimental techniques are used to determine how a peptide or protein that can form amphiphilic helices interact with the

lipid surface. These include site-directed mutagenesis studies coupled with spectroscopic (CD, fluorescence) and binding assay methods (using liposomes).

The properties of a membrane-bound amphiphilic helix are affected by the exact size and distribution of the polar/charged and nonpolar side chains. On binding, they sense or cause membrane curvature, interact with specific lipids, and stabilize specific membrane conformations (such as spherical for lipoproteins). Figure 17.1.10 shows how different proteins with amphiphilic membranes interact with membrane surfaces.

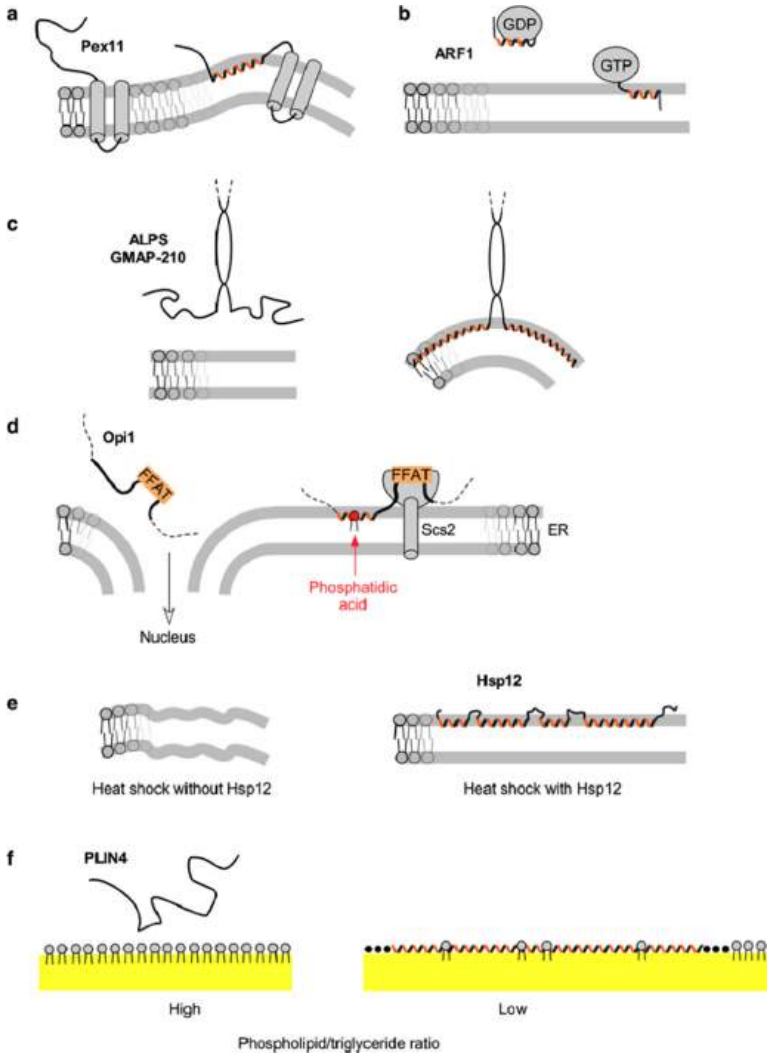


Figure 17.1.10 Interactions of amphiphilic helices with membranes. Manuel Giménez-Andrés et al. Biomolecules. 2018 Sep; 8(3): 45. doi: 10.3390/biom8030045. Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/)

A key point to note is the large conformational changes that occur as the protein or parts of it go from the free, more disordered state, to the bound state with lipid-associating amphiphilic helices. The following proteins are depicted in the figure.

a. The peroxisomal membrane protein Pex11 amphiphilic helix distorts the membrane;
b. ARF1 is a small G protein in which only the GTP form localizes and binds through an amphiphilic helix to the membrane;
c. The ALPS motif of the golgin GMAP-210 binds to only highly curved vesicles;
d. The yeast transcriptional repressor Opi1 binds to the endoplasmic reticulum (ER) membrane in part through an amphiphilic helix;
e. The heat shock protein Hsp12 has a long amphiphilic helix which helps stabilize the membrane;
f. The extremely long amphiphilic helix of perilipin 4 coats lipid droplets and stabilizes even if there is a lack of phospholipids.

17.1.3: DIETARY UPTAKE AND RELEASE INTO THE CIRCULATION

Now how are the lipid nanoparticles assembled? We'll start with dietary lipids in the form of TAGs, glycerophospholipids, and cholesterol esters. The figure below shows key steps which are described in Figure 17.1.11

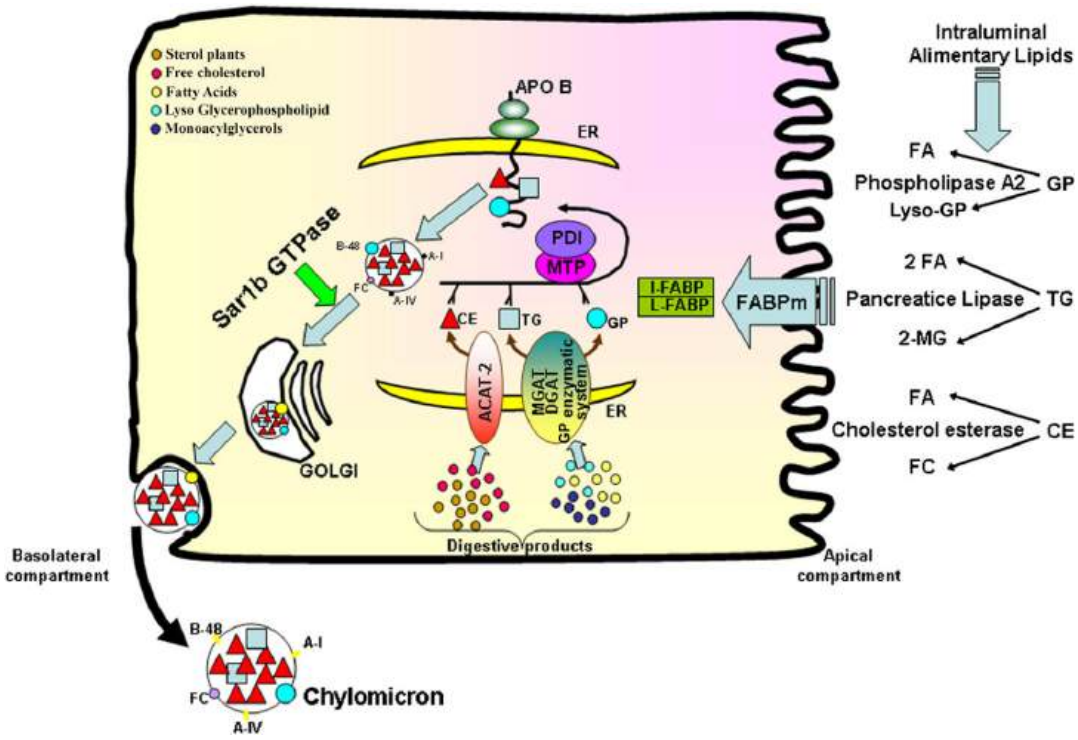


Figure 17.1.11: Emile Levy, Journal of Lipid Research. 56, 945-962, 5/2015. DOI:https://doi.org/10.1194/jlr.R052415. Creative Commons Attribution (CC BY 4.0)

Here are some key steps depicted in the figure:

- hydrolysis (lipolysis) of TAGs by pancreatic lipase, cholesterol esters (CE) by cholesterol esterase, and glycerophospholipids (GP)by phospholipase A2 in the lumen of the intestine. These enzymes interact at the interface of the lipid substrates and aqueous surroundings;
- the resulting products, which include free fatty acids (FA), 2-monoacylglycerol (MG), free cholesterol (FC), and lyso-glycerolphospholipids (lyso-GP), aggregate with the help of bile salts to form emulsions (like oil drops in water), which can be taken up by diffusion or possibly endocytosis when present in high amounts. Alternatively, membrane transporters (like FABPs and other proteins) can move them into the cell by facilitated diffusion;
- cytoplasmic transporters like fatty acid binding proteins move the lipolysis product to the ER where free fatty acids are reesterified. The enzymes involved include mono- and diacylglycerol acyltransferases (MGAT, DGAT) and sterol O-acyltransferase 2, also known as acyl-coenzyme A:cholesterol acyltransferase (ACAT-2). Multiple enzymes are involved in the resynthesis of glycerophospholipids);
- Apo B-48 is synthesized by ribosomes bound to the ER and interacts with a heterodimer of **microsomal triglyceride transfer protein large subunit two (MTP)** and protein disulfide isomerase (PDI). This facilitates the folding of apo B48 and loading of lipids using MTP into pre-chylomicrons;
- pre-chylomicron vesicles move to the Golgi with the help of Sar1b, a small G-protein (and GTPase) where the particle assembles to the full chylomicron, which is released from the cells as the mature large lipid nanoparticle.

An intriguing feature of lipases is that they work at the interface between the aqueous and nonaqueous (in this case lipid nanoparticle) environments. Let's briefly consider the mechanism of hydrolysis of TAGs by equine pancreatic mechanism. This enzyme utilizes the same mechanism we have seen earlier for the hydrolysis of a peptide bond by serine proteases. A catalytic triad of Asp 176, His 263 and Ser 152 as a nucleophilic catalyst is shown in the partial reaction displayed in Figure 17.1.12.

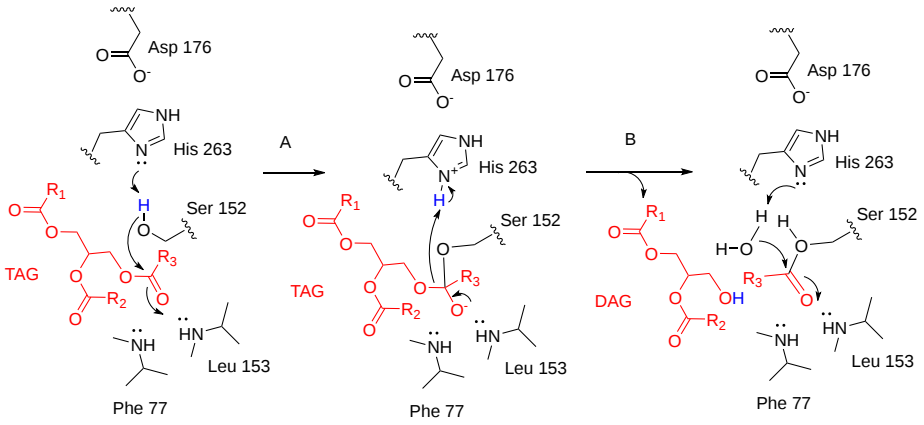


Figure 17.1.12

An acyl-Ser intermediate forms in step B (above), after the collapse of an oxyanion intermediate in step A, to form the product diacylglycerol. In the second half of the reaction (not shown completely), water, in a hydrolysis reaction, cleaves the acyl-Ser intermediate to reform the active enzyme as it releases the free fatty acid, R_3CO_2H. Other lipases also employ the same catalytic triad.

Figure 17.1.13 shows an expanded diagram showing the flow and fate of lipoproteins.

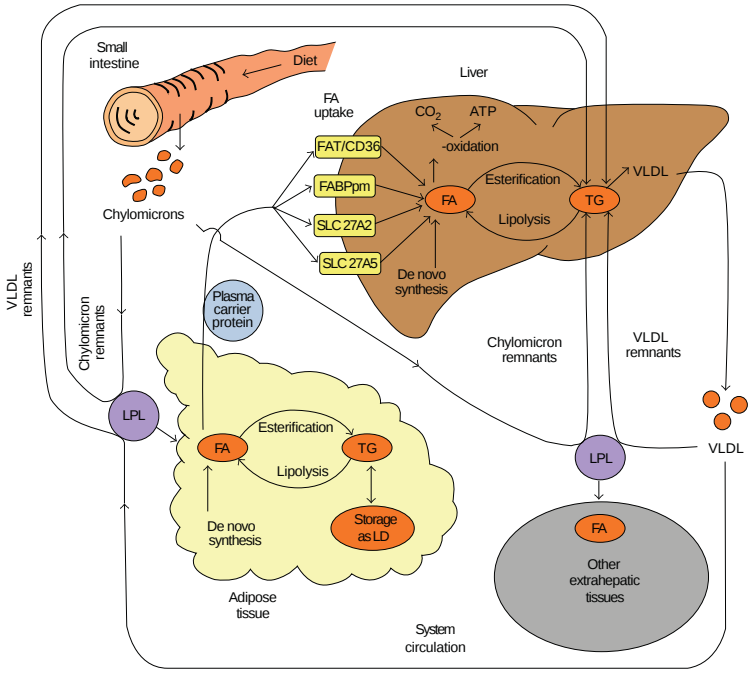


Figure 17.1.13: Merilin Al Sharif et al. PPAR Research. Volume 2014, Article ID 432647, http://dx.doi.org/10.1155/2014/432647. Creative Commons Attribution License.

Chylomicrons interact with lipoprotein lipase (LPL), which also uses an Asp-His-Ser catalytic triad, to cleave fatty acid esters, which allows the delivery of free fatty acids to adipose cells. The adipocytes can also undergo de novo fatty acid synthesis. Fatty acids (FA) can also be produced by lipase-mediated lipolysis of stored TGs. Any of these free fatty acids (FA) in the adipocyte has two fates. They can be reesterified to glycerol to form TAGs (TG in the figure) or be exported from the cell and bind to a plasma carrier protein and transported to the liver, where it can be taken up by a variety of membrane proteins importers shown in the figure in yellow boxes. There, as in adipose cells, they can be reesterified to form TAG stores, which can then be packaged into VLDL particles for export. The fatty acid delivered (or synthesized) could also be used for ATP production through the citric acid cycle and oxidative phosphorylation.

VLDL in circulation can undergo lipolysis by lipoprotein lipase to produce fatty acids for uptake in "extrahepatic" tissue (bottom right of the diagram). As fats are removed from VLDL, their density increases as it forms IDL and LDL, which could be considered VLDL "remnants". VLDL is very enriched in TAGs, but after metabolic processing, the resulting LDL is depleted in TAGs and enriched in cholesterol/cholesterol esters. LDL (not shown in the above figure) can be taken up (endocytosed) by the liver and other cells after binding to LDL receptors, which recognize apo-B100 and other apoproteins. This allows the delivery of predominately cholesterol and cholesterol esters to tissues.

How do adipocytes and hepatocytes determine if free fatty acids should be esterifed for storage or released for energy use by other tissue? We'll discuss that in a subsequent section but the short answer is that in healthy fasting and exercise states, hormones (glucagon, epinephrine) will activate lipolysis in the liver and adipose cells, while in the fed state, insulin will promote storage of fatty acids as triacylglycerols.

Adipose cells don't assemble and release lipoproteins. Instead they release free fatty acids in the circulation which are carried by albumin, the major serum/plasma protein in the blood. The iCn3D Figure 17.1.14 shows an interactive iCn3D model of the complex of human serum albumin (HSA) binding seven 20:4$^{\Delta 5,8,11,14}$ - arachidonic acids (1gnj).

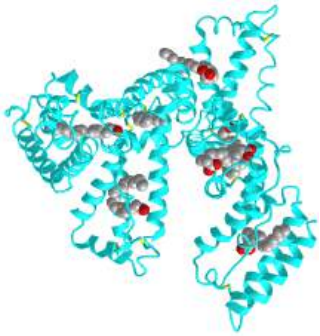


Figure 17.1.14 Human Serum Albumin/Arachidonic Acid complex (1gnj). Click the figure for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?WL7BxK4B23m2p3Z19 (Copyright; author via source)

Given the multiple binding sites for fatty acids in albumin, it should come as no surprise that albumin also binds a host of small drugs, including medicinal drugs and toxins such as warfarin (blood thinner), diazepam, ibuprofen, indomethacin, and amantadine. These appear to bind preferentially at two major drug binding sites. This binding is probably helpful in delivering drugs through the circulation but potentially not useful if they aren't delivered to appropriate target tissue.

We discussed the structure of micelles which are spherical assemblies of single-chain amphiphiles that act as detergents. Oil from your clothes can enter the nonpolar interior of the detergent micelle and effectively solubilize the nonpolar molecule in the micelle, which are effectively nanoparticles with a diameter of 5-15 nm. You should hence not be surprised to discover that lipoproteins can also carry fat-soluble vitamins, steroid-like endocrine-disrupting substances, and drugs.

17.1.4: LIPOPROTEIN LIPASE

The enzyme that breaks now TAGs in circulating chylomicrons and VLDL is **lipoprotein lipase (LPL)**. It is a soluble protein secreted by adipocytes and muscle cells but is made by many cell types. It works at the luminal side of blood vessel endothelial cells and is recruited to that membrane surface by binding to the **glycosylphosphatidylinositol-anchored high-density lipoprotein-binding protein 1 (GPIHBP1)** as well as the proteoglycan heparan sulfate at the cell surface.

What is so interesting is that GPIHBP1 is only synthesized by endothelial cells. When lipoprotein lipase is secreted from cells, it binds to the extracellular matrix heparan sulfate but dissociates on the cleavage of heparan sulfate by heparinases. GPIHBP1 is highly acidic with an intrinsically disordered N-terminal domain containing a sulfated tyrosine and is highly enriched in glutamates and aspartates, which are often sequential in the sequence. Here is the single-letter sequence for amino acids 25-50 of the human version of GPIHBP1: EEEEEDEDHGPDDYDEEDEDEVEEEE. This sequence would have similar electrostatic and binding properties to the highly negatively charged heparan sulfate to which it also binds.

LPL also binds Ca^{2+} which stabilizes the active dimeric form of the protein. Its enzymatic activity is activated by apoC-II. Like pancreatic lipase, it employs a Ser-132, Asp-156, and His-241 triad in its hydrolytic action on TAGs/

Figure 17.1.15l below shows an interactive iCn3D model of LPL in complex with GPIHBP1, shown in brown (6E7K). The calcium ion is shown (grey spacefill) as well as the catalytic triad (labeled, sticks, CPK colors). The highly negatively charged stretch of amino acids in GPIHBP1 was not present in the crystal structure.

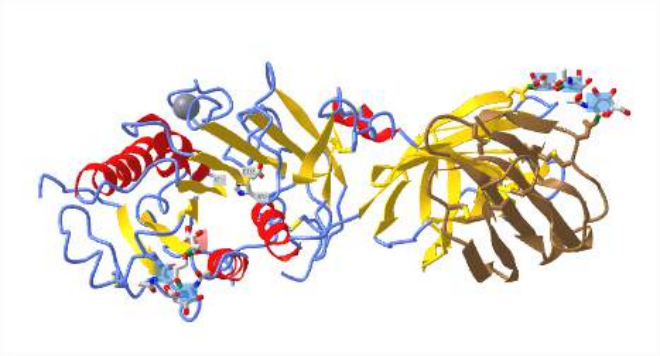


Figure 17.1.15: Lipoprotein Lipase/GPIHBP1 Complex (6E7K)). The calcium ion is shown (grey spacefill) as well as the catalytic triad (labeled, sticks, atom colors). The HPIHBP1 protein is shown in brown. The highly negatively charged stretch of amino acids in GPIHBP1 was not present in the crystal structure. Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...VQmAikgmf7nC67 (Copyright; author via source) Kumari, A.; Kristensen, K.K.; Ploug, M.; Winther, A.-M.L. The Importance of Lipoprotein Lipase Regulation in Atherosclerosis. Biomedicines 2021, 9, 782. https://doi.org/10.3390/biomedicines9070782. Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

17.1.5: LDL: RECEPTOR AND UPTAKE

Lipoproteins are taken up into cells through receptor-mediated processes. Let's focus generically on the LDL receptor, the major carrier of cholesterol, given its role in cardiovascular disease. It is found in the cell membranes in most tissues. It is has many domain repeats, as illustrated in the Figure 17.1.16 calculated by SMART.

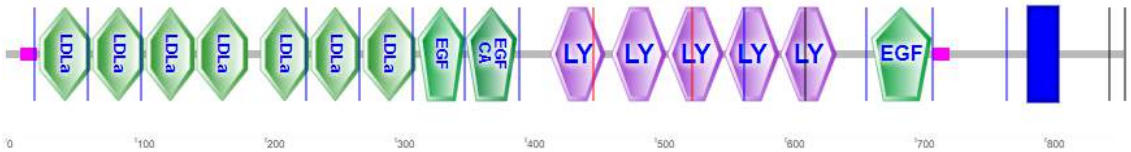


Figure 17.1.16: Domain strucuture of the LDL receptor

They include the N-terminal region cysteine-rich LDLa domains, which bind LDL, epidermal growth factor domains, LY (or LDLb) domains, and a transmembrane domain (blue rectangle).

Figure 17.1.17 shows an interactive iCn3D model of the extracellular domain of the LDL receptor (1n7d).

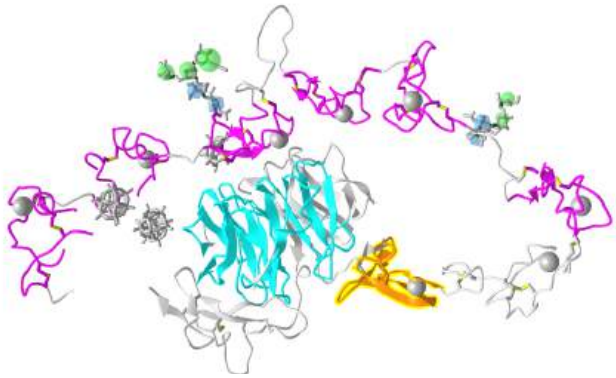


Figure 17.1.17: Click the figure for a popup model or use this external link: https://structure.ncbi.nlm.nih.gov/i...aT7M6ti5iDAbj8 (Copyright; author via source)

Four tandem LY (LDLb) domains are shown in cyan, LDLa domains are shown in magenta and the EGF domain is shown in dark orange. Glycans are shown in symbolic nomenclature for glycans. Zoom into the structure to see the two disulfide bonds in each LDLa domain as well as the Ca^{2+} ions that stabilize the domains.

LDL binds its receptor at a broad binding interface with multiple LDLa domains. This may account for the fact the lipid nanoparticles with apo B100 or apo E can bind to it. The binding triggers a series of signaling events that lead to internalization by endocytosis of the receptor in pits coated with the protein clathrin. These eventually fuse with lysosomes where they are degraded and cholesterol delivered to the cell. The steps are described in Figure 17.1.18.

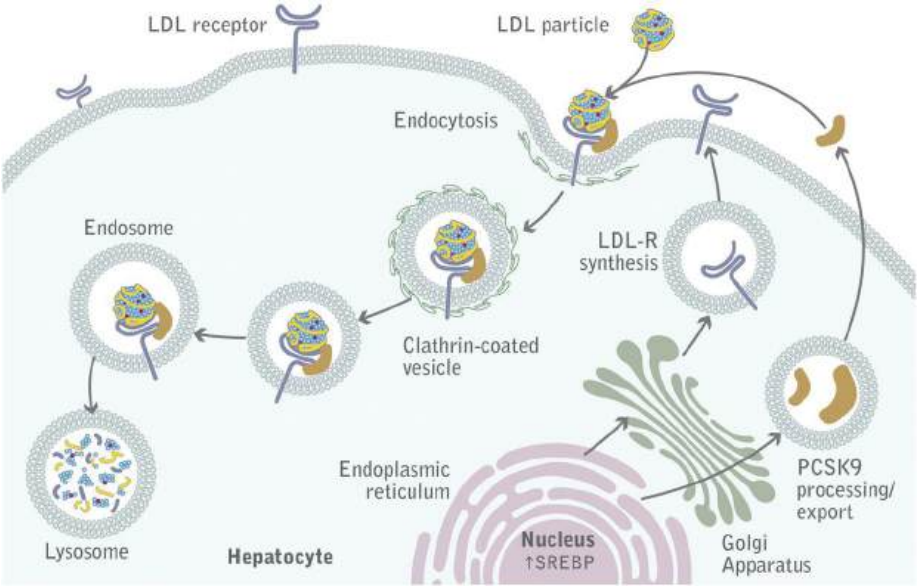


Figure 17.1.18 Gilles Lambert et al. J. Lipid Research, 53, 2515-2524 (2012) DOI:https://doi.org/10.1194/jlr.R026658. Creative Commons Attribution (CC BY 4.0)

The LDL receptor survives lysosomal degradation and along with newly made receptors is delivered to the plasma membrane continually. A key protein, **proprotein convertase subtilisin kexin type 9 (PCSK9)**, a serine protease secreted by the liver, promotes enzymatic degradation of the receptor and prevents its recycling to the membrane. It also binds to VLDLR and apolipoprotein E receptors and promotes their degradation as well. Its action reduces LDL clearance from the blood, increasing cardiovascular risk, so inhibitors of its action might be potent drugs to decrease circulating LDL.

The LDL receptor is just one member of the LDL receptor family. Other members of the family are illustrated in Figure 17.1.19.

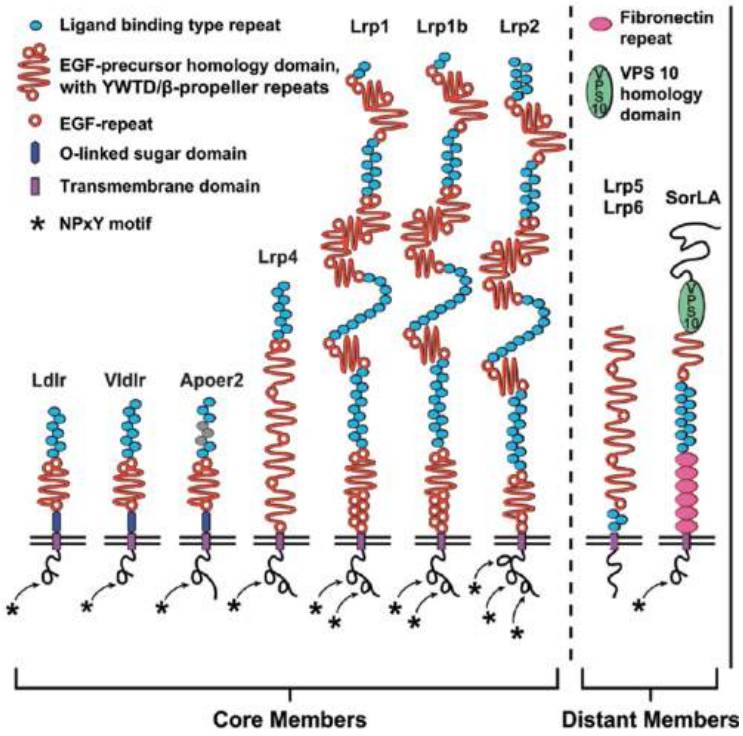


Figure 17.1.19: LDL Receptor Family. Adapted from Theresa Pohlkamp et al. Front. Mol. Neurosci., 01 March 2017 | https://doi.org/10.3389/fnmol.2017.00054. Creative Commons Attribution License (CC BY)

These include the LDL receptor (abbreviated Ldlr in Figure 17.1.17, as well as the VLDL receptor (Vldlr), apolipoprotein E2 receptor (Apoe2), and LDL receptor-related proteins (Lrp)1-4. These also have a NPxY-motif (asterisk in the cytoplasmic domain) and a YWTD/β-propeller domain. Given the similarity in domain structure for the LDL family of receptors, the conformational flexibility of the apolipoproteins (at least free in solution), and similar structures for the exchangeable apolipoproteins, it shouldn't be surprising that the LDL receptor would interact with different classes of lipoproteins, albeit with different affinities.

As mentioned previously, apoE is found most abundantly on HDL and VLDL/chylomicrons and their remnants. It serves as a ligand that binds to members of the LDL receptors family (remember that LDL generally binds the LDL receptor through apo B100).

APO E AND ALZHEIMER

Apolipoprotein E has three major variants (alleles) named ε2, ε3, and ε4 (also called ApoE2, 3, and 4). ApoE3 is the most prevalent. ApoE4 is found in only 15% of people but more than 50% with Alzheimer's Disease (AD), so it's a risk factor for this disease. AD affects the brain, which also contains up to 30% of the cholesterol in the body, so aberrations to cholesterol transport and uptake in the brain are not unexpected in neurodegenerative diseases like AD.

ApoE is secreted by brain microglia (immune) cells and astrocytes (specialized glial cells). It assembles lipids into lipoproteins (HDL-like) and becomes the major vehicle for binding to and importing into neurons in a process initiated by the apoE receptor. The major apoE receptor for clearance of lipoproteins in the brain is sortilin (SorLA in Figure 17.1.17).

AD is characterized by the accumulation of a toxic amyloid prion protein called amyloid beta (Aβ). It is derived from selective but abnormal proteolysis of the neural integral membrane protein amyloid precursor protein (APP). Aβ aggregates to form insoluble neurotoxic extracellular Aβ amyloid plaques. The process in normal and diseased cells is shown in Figure 17.1.20.

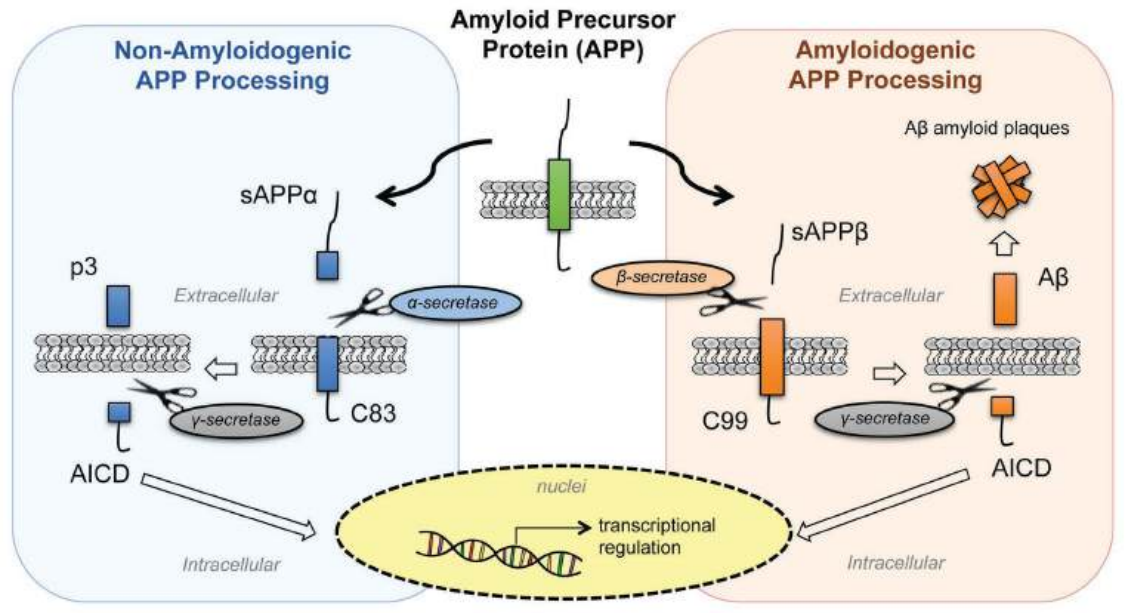


Figure 17.1.20 Processing of amyloid precursor protein (APP). Emma Ristori et al. Front. Physiol., 27 August 2020 | https://doi.org/10.3389/fphys.2020.01056. Creative Commons Attribution License (CC BY).

The figure shows normal (left) and aberrant processing of APP and the family of proteases (secretases) involved. While the LDLR doesn't appear to bind to APP or influence its proteolytic processing, it does bind Aβ. LRP1 is much bigger than LDLR, binds a multitude of ligands, and can be cleaved with the same enzymes as APP. Its expressed in the liver and especially in the brain and can regulate the removal of Aβ.

Immune cells in the brain, called microglial, remove Aβ plaques (which are extracellular) by phagocytosis. ApoE4 increases the inflammatory response (as measured by cytokine release) of the microglia (a good thing if the responses prevents infection or rids Aβ plaques) but also inhibits their ability to phagocytose the Aβ plaques and their metabolic activity.

An additional note: Having one allele ApoE4 allele appears to increase the risk of severe COVID-19 five times while being homozygous for E4 leads to a 17-fold increased risk of severe disease.

17.1.6: SCAVENGER RECEPTORS

Patients with homozygous familial hypercholesterolemia (FH) have very high levels of LDL derived from defects in binding and update. Patients display fatty acids streaks under vessel endothelial cells which morph into calcified plaques and lesions filled with fat. Monocytes/macrophages, which migrate to sites of vascular injury, take up LDL and eventually differentiate into foam cells filled with lipids. Somehow, they have receptors that can bind and internalize LDL when the "normal" LDL receptor can't. Brown and Goldstein found that a specific chemical modification of LDL, acetylation, was necessary for the rapid uptake of "modified" LDL into macrophage receptors. These receptors are now called scavenger receptors (SRs).

There is a large family of scavenger receptors. It consists of classes A-J proteins that share functional but not sequence homology. They are found on macrophages and endothelium. They bind to and help remove "damage" signals including damage-associated molecular patterns (DAMPs) and chemical species chemically modified by reactive oxygen species. The ligands are often polyanions, end-stage glycans, and extracellular matrix proteins. One such example is oxidized-LDL (produced in vivo or by chemical oxidative modification with malondialdehyde), which binds to the same scavenger receptor, **SR-A1**, also called **Macrophage scavenger receptor type I**, as acetylated-LDL. Figure 17.1.21 shows the domain structure of the SR family.

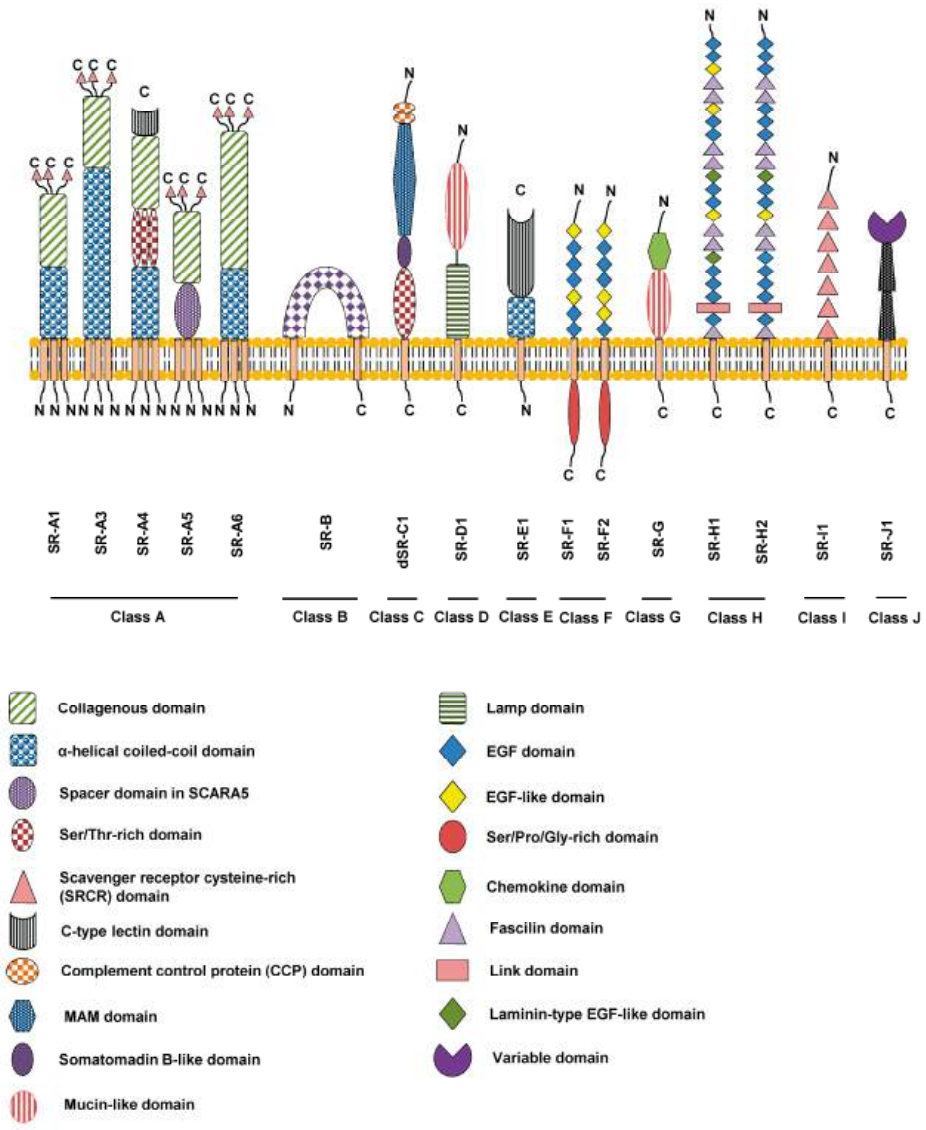


Figure 17.1.21: Domain Structure of the Scavenger Receptor Family. Izma Abdul Zani et al. Cells 2015, 4(2), 178-201; https://doi.org/10.3390/cells4020178. Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/)

SR-A1/MSR1 not only binds acetylated and oxidized LDL but also β-amyloid (42), heat shock proteins (43), and PAMPs from some bacteria and viruses.

It's very difficult, given the ever-increasing amount of "omic" data (genomic, proteomic, lipidomics, interactomics, metabolomics), for readers and authors alike, to conceptualize all of the possible combinations of interactions among biological molecules. For visual learners and perhaps everyone else, it's extremely useful to portray information on structures and interactions visually. An example using STRING, a database of known and predicted protein interactions, for the domain structure and protein:protein interactions of SR-A1/MSR1, is shown in Figure 17.1.22.

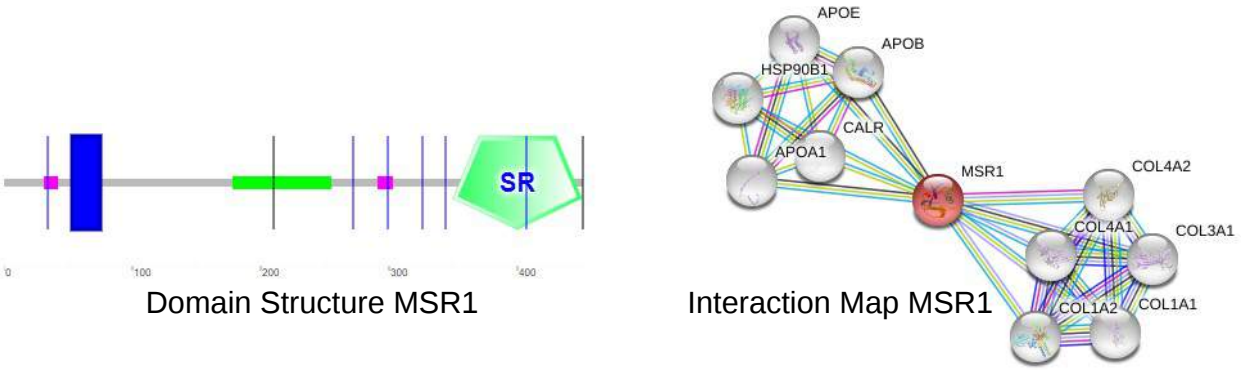


Figure 17.1.22 Visualization of Domain Sructure and Interaction Map of Macrophage Scavenger Receptor 1

Note the interactions with apolipoproteins, apoB, apoE, and apoA1. The right hand side of the figure also shows interactions with **collagen alpha-2(IV) chain** (COL4A2), which is found in the extracellular matrix.

17.1.7: HDL METABOLISM: THE GOOD CHOLESTEROL

High levels of LDL (and Lp(a)) pose a cardiovascular risk. In contrast, high levels of HDL and apo A-I are cardioprotective. HDL is involved in "reverse" cholesterol transport as it is taken up by the liver and sent to the intestines for elimination from the body. We have shown earlier in Figure 2 that HDL exists as many variants, reflecting the assembly and remodeling of HDL by enzymes and lipid transfer proteins. Figure 17.1.23 shows the lifecycle of HDL.

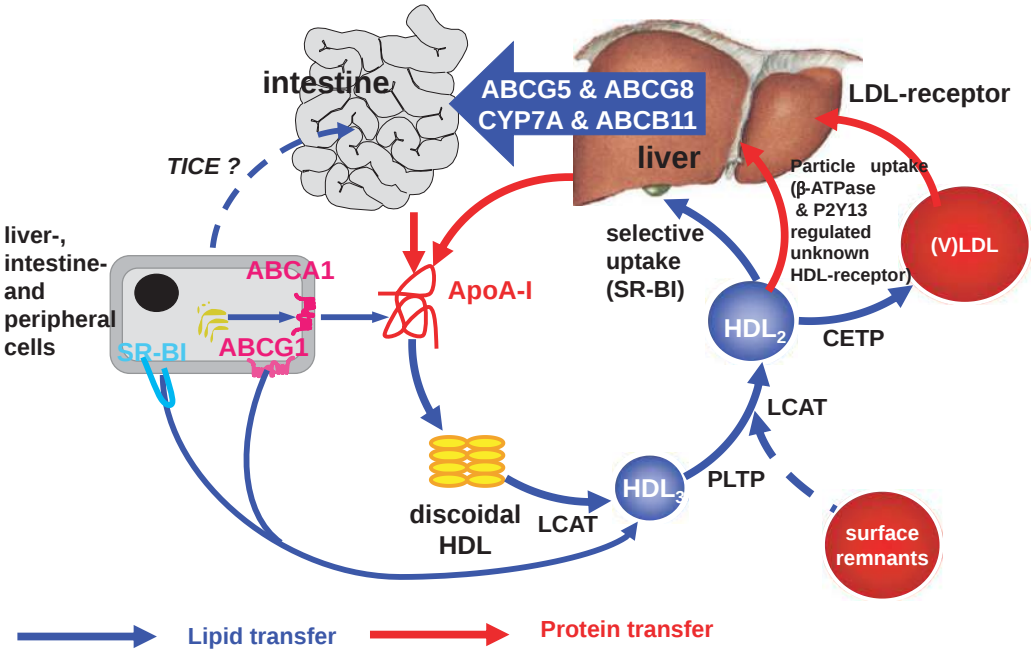


Figure 17.1.23 Arnold von Eckardstein. Tachometer for Reverse Cholesterol Transport? August 2012. Journal of the American Heart Association 1(4):e003723. DOI: 10.1161/JAHA.112.003723. License CC BY 2.5/. Open Access article under the terms of the creative Commons Attribution Noncommercial License

Secreted apo A-1 accretes lipids in the circulation through the transport and delivery of phospholipids and cholesterol from cell membranes by the ATP-binding cassette transporters (ABC) A1 and G1. Another protein, the **scavenger receptor BI (SR-BI)**, a polytopic integral membrane protein, is also involved. It acts as a receptor for a variety of "lipid" ligands including phospholipids, cholesterol esters, and phosphatidylserine (an outer membrane marker for cell apoptosis) as well as lipoproteins such as HDL.

Other proteins are involved as well in both the assembly but especially in the remodeling of HDL. The lipolytic enzyme lecithin-cholesterol acyltransferase (LCAT) removes a fatty acid from phospholipids and adds it to free cholesterol in the HDL to form cholesterol esters. Two major lipid transfer proteins, **phospholipid transfer protein (PLTP)** and **cholesterol ester transfer protein (CETP)**, move lipids between

HDLs and other lipoproteins. Cholesterol ester transport protein is made and secreted from the liver. It appears to exchange cholesterol esters from HDL for the return of TAGs from VLDL. Other enzymes (lipoprotein lipase and hepatic lipase are also involved in forming free fatty acids.

In the final step, HDL can deliver cholesterol ester and cholesterol to the liver through binding to liver scavenger receptor BI (SR-BI) mediated by apo A-I. The protein is expressed by the liver, adrenal gland, endothelial cells, macrophages, and many other tissues. It appears that HDL is **not** primarily taken up by the liver by endocytosis. In contrast, LDL is taken up by endocytosis mediated by the LDL receptor. SR-BI facilitates the transfer of cholesterol esters from bound HDL to the liver cell.

HDL AND CARDIOVASCULAR RISK

Unlike the widespread use of statins, which reduce LDL-C concentrations and clinically reduce cardiovascular disease risk, drugs (fibrates, niacin, inhibitors of cholesterol ester transfer protein - CETP), which raise HDL levels, don't seem to lead to significant decreases in cardiovascular risk. The cholesterol delivered in excess to macrophages can lead to the formation of foam cells under the endothelial layer. Foam cells are proinflammatory and convert in time to cholesterol plaques. In contrast, HDL-C does appear to have this direct atherogenic effect.

17.2: OXIDATION OF FATTY ACIDS

17.2.1: INTRODUCTION

Fatty acids, esterified to glycerol in triacylglycerols, are the major source of stored energy in organisms. As we burn fossil fuels to produce energy to drive our society, so can we "burn" fatty acids to indirectly produce for heat, ATP to drive biosynthetic reactions, and to do work. As discussed in an earlier chapter, fatty acids are highly reduced so their oxidation by dioxygen is highly favored both enthalpically (exothermic reaction) and entropically. Of course, the biological oxidation reactions occur in a stepwise fashion using not O_2 directly, but rather less potent oxidizing agents like NAD^+ and FAD. We'll focus first on fatty acid oxidation in animals (humans).

As we discussed in the previous section, fatty acids released from triglyceride stores on signaling by epinephrine and glucagon in exercise and between meals are used for energy when glycogen stores are low and without breaking down protein in muscles to produce energy. Some fatty acids are broken in the normal process of membrane turnover and removal of xenobiotic lipids.

Most fatty acids are oxidized in the mitochondria, where the oxidation reaction occurs at the beta-carbon of the acyl chain, as shown in Figure 17.2.1.

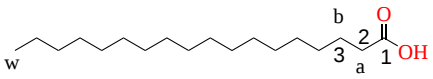


Figure 17.2.1: α and β carbons of fatty acids

This pathway is called **β-oxidation**. Fatty acids are oxidized in a step-wise fashion by this pathway. In each repetitive cycle of this pathway, acetyl-CoA and one CO_2 are released. In addition, oxidation can occur at both the alpha- and beta carbons when oxidized in an organelle called the peroxisome. The **α-oxidation** pathway is used for fatty acid branched at the beta carbon 3 releasing one CO_2 until the beta-oxidation pathway can be used. The peroxisome degrades fatty acids that can't be oxidized in the mitochondria. These include very long-chain fatty acids (VLCFAs) like 24:0 and 26:0, and in addition, branched-chain fatty acids (BRCHAs) including some fatty acids from dietary sources such as pristanic acid (an odd-chain 15:0 fatty acid methylated at carbons 2, 6, 10, and 14). The α-oxidation pathway can't be used to completely oxidized fatty acids in the peroxisome. At some point in the oxidative stepwise pathway, the resulting shorter fatty acids are exported to the mitochondria for β-oxidation. Also, enzymes in the endoplasmic reticulum have the **ω-oxidization** pathway which oxidizes fatty acids at the omega or terminal carbon. The enzyme used is the monooxygenase cytochrome P_{450} which uses one oxygen from O_2 to hydroxylate the ω-carbon.

PEROXISOMES - AN UNDERAPPRECIATED ORGANELLE

These organelles, initially called microbodies, are vital to cellular metabolism and health. In people with Zellweger syndrome spectrum, there is a severe disorder in the formation of peroxisomes, which is often lethal. They are important metabolically in lipid metabolism, synthesis of myelin sheath lipids, and metabolism of reactive oxygen species like peroxides. The enzymes catalase and urate oxidase are found in such high concentrations they often form crystal "bodies" in the matrix of the peroxisome. Additional roles include responses to pathogens and viruses. Effectively they are a protective organelle.

In contrast to mitochondria, peroxisomes, like most other organelles, have a single bilayer and no DNA, from which transcription of RNA and translation of proteins occur. All proteins are hence imported from the cytoplasm after synthesis on free ribosomes. Imported proteins have a peroxisome targeting sequence (PTS) of serine-lysine and leucine (SKL) near their C-terminus which facilitates the binding of the proteins to a PTS receptor in the peroxisome membrane. These organelles oxidize very long-chain fatty acids (VLFA), make and break down hydrogen peroxide (hence the name), and also synthesize plasmalogens. The enzymes involved in the stepwise cycle of reactions in the peroxisome β-oxidation pathway use enzymes different from those used in the mitochondria for β-oxidation.

For those more inclined towards chemistry than biology, yet another organelle with its structures and function may seem like one too many. However, this less-discussed organelle is critically important in its own right. Figure 17.2.2 shows features of peroxisomes and their proteins.

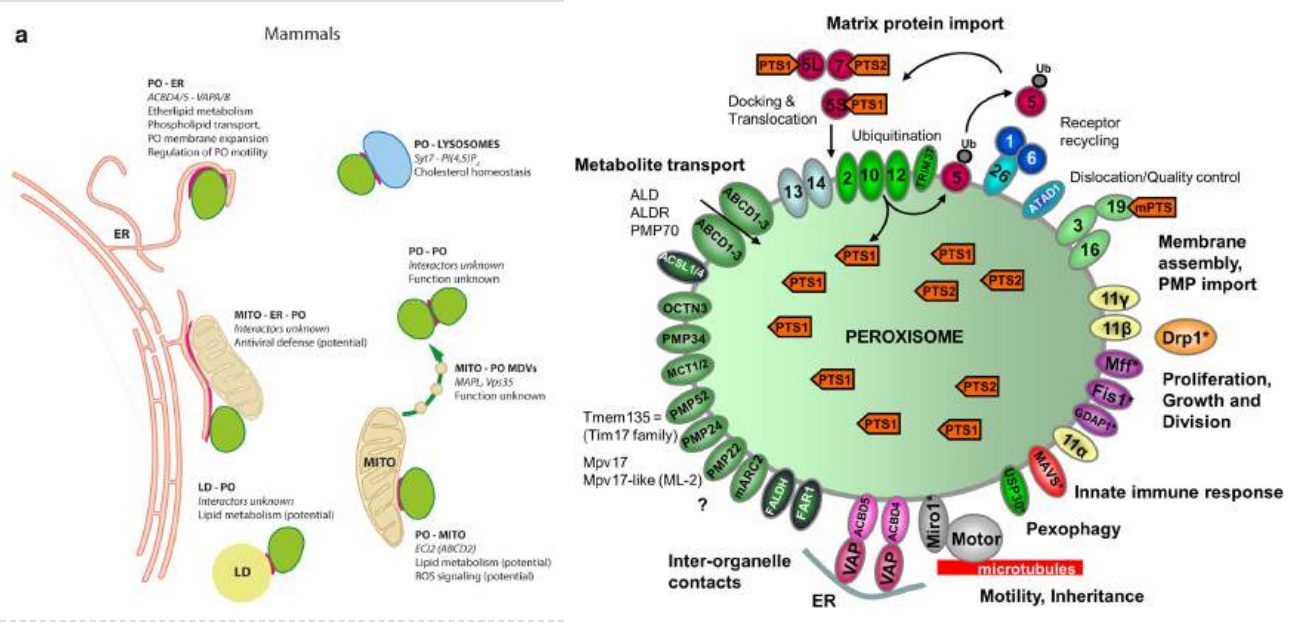


Figure 17.2.2: Features of peroxisomes and their proteins. Islinger, M., Voelkl, A., Fahimi, H.D. *et al.* The peroxisome: an update on mysteries 2.0. *Histochem Cell Biol* **150,** 443–471 (2018). https://doi.org/10.1007/s00418-018-1722-5. Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/)

One interesting feature is its relationship with different organelles in cells, as shown in the left panel of Figure 2. Some proteins involved in organelle functions are shown as well (right panel).

The peroxisome (PO, green) has binding interactions (red interfaces) with the endoplasmic reticulum (ER) lysosomes, mitochondria, lipid droplets (a pseudo-organelle), and also itself (left figure). Some of the key membrane proteins (which we have discussed previously) involved in peroxisome function include the ABC transporter proteins ABCD1-3 for fatty acids transport, OCTN3 for organic and cation/carnitine transport, and MCT1/2 for monocarboxylate transport. In addition, peroxisomes have receptors for protein import mediated by PTSs and for peroxisome movement along microtubules in the cell.

17.2.2: MITOCHONDRIAL B-OXIDATION

Mitochondrial β-oxidation in muscle generates acetyl-CoA, which enters the citric acid cycle for subsequent production of ATP through mitochondrial electron transport and oxidative phosphorylation. In the liver, the generated acetyl-CoA is used for ketone body production under fasting states. Figure 17.2.3 shows the β-oxidation pathway for palmitic acid (16:0), a saturated fatty acid, starting with its import from the cytoplasm.

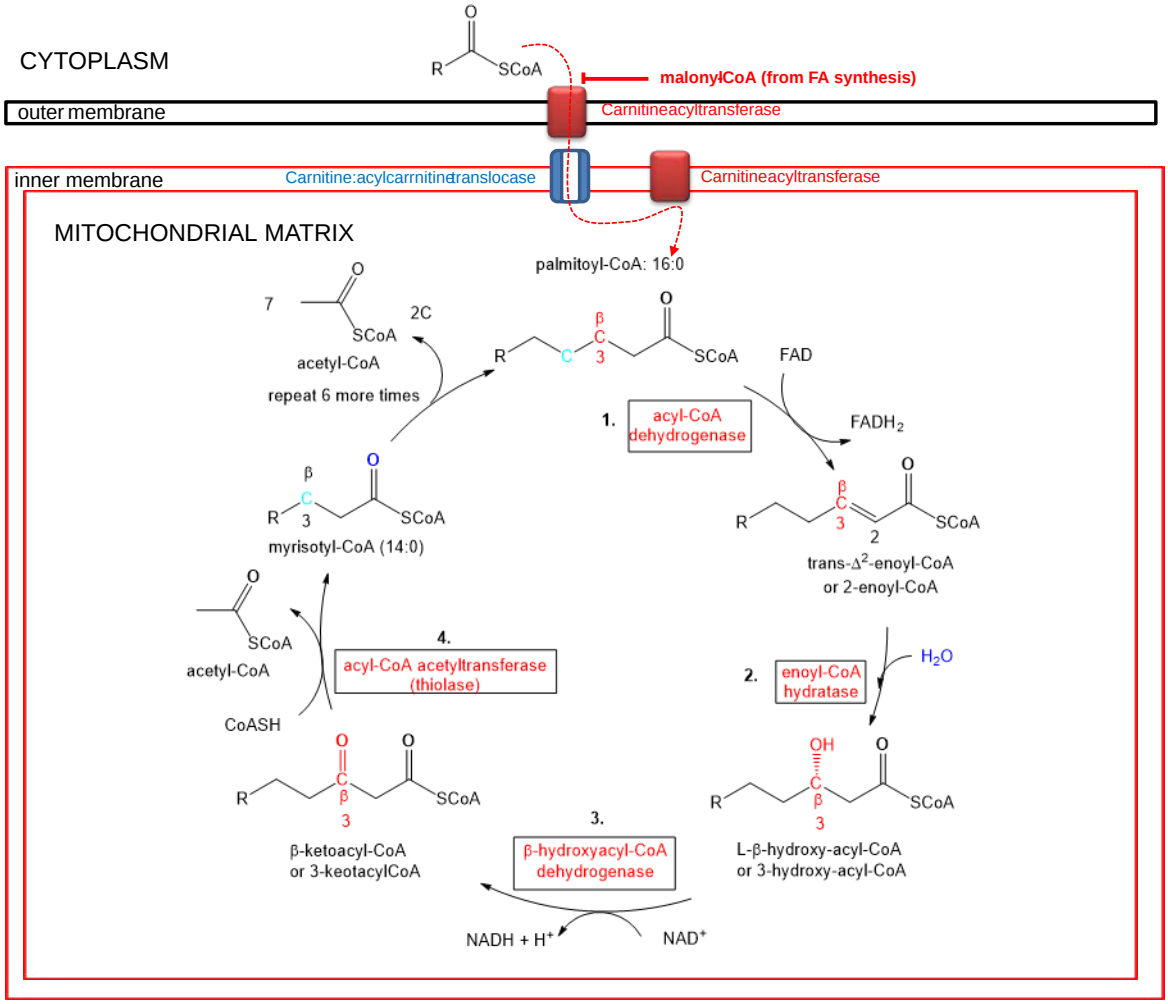


Figure 17.2.3: Mitochondrial β-oxidation of palmitic acid (16:0)

The pathway involved cyclic removal of 2C unit until 16:0 is cleaved 7 times producing 8 2C acetyl-CoAs. The net chemical equation of beta-oxidation of 16:0 is shown in the equation below.

$$C_{16} - CoA + 7NAD^+ + 7FAD + 7CoASH + 7H_2O \rightarrow 8Acetyl - CoA + 7NADH + 7FADH_2 + 7H^+ \quad (17.2.1)$$

Figure 17.2.4 shows an abbreviated comparison of the β-oxidation pathways in the mitochondria and peroxisomes. Peroxisomal beta-oxidation is used to metabolize very-long-chain fatty acids (VLCFAs), which are composed of 24-26 carbon units as well as branched-chain fatty acids (BRCHAs).

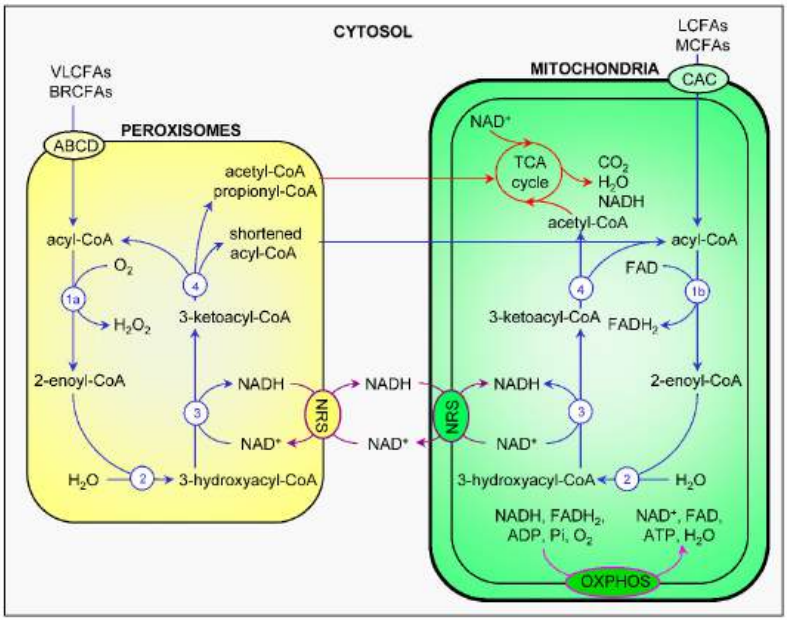


Figure 17.2.4: Comparison and interplay of peroxisomal and mitochondrial fatty acid β-oxidation. Fatty acid β-oxidation, the NAD(H) redox shuttles, the tricarboxylic acid cycle, and the electron transfer chain are respectively depicted in blue, purple, red, and pink. 1a, acyl-CoA oxidase; 1b, acyl-CoA dehydrogenase; 2, enoyl-CoA hydratase; 3, 3-hydroxyacyl-CoA dehydrogenase; 4, 3-ketoacyl-CoA thiolase. ABCD, ATP-binding cassette transporters of subfamily D; ADP, adenine dinucleotide phosphate; BRCFA, branched-chain fatty acid; CAC, carnitine-acylcarnitine carrier; FAD, flavin adenine dinucleotide; FADH2, reduced FAD; LCFA, long-chain fatty acid; MCFA, medium-chain fatty acid; NAD, nicotinamide adenine dinucleotide; NADH, reduced NAD; NRS, NAD(H) redox shuttles; OXPHOS, oxidative phosphorylation; TCA, tricarboxylic acid; VLCFA, very-long-chain fatty acid. Fransen et al. International Journal of Molecular Sciences. 18. 1126. 10.3390/ijms18061126. DOI: 10.3390/ijms18061126. Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

It should be noted that a likely NAD^+/NADH mitochondria transporter, a multi-pass inner mitochondrial membrane protein has just been identified The transporter, MCART1, is also called SLC25A51,

17.2.3: B-OXIDATION - MECHANISMS

Fatty acids are imported into the matrix from the cytoplasm through their acyl-CoA derivatives. Two different proteins are required for their import. One is **carnitine palmitoyltransferase-1(CPT-1),** which transfers the acyl group from CoASH to a carrier protein carnitine. The acylcarnitine is translocated through the inner membrane by the carrier protein **carnitine-acylcarnitine translocase (CACT)**. Once inside the matrix, the acyl group is transferred back to CoASH by **carnitine palmitoyltransferase-2 (CPT-2)**. This **carnitine cycle** is illustrated in Figure 17.2.5.

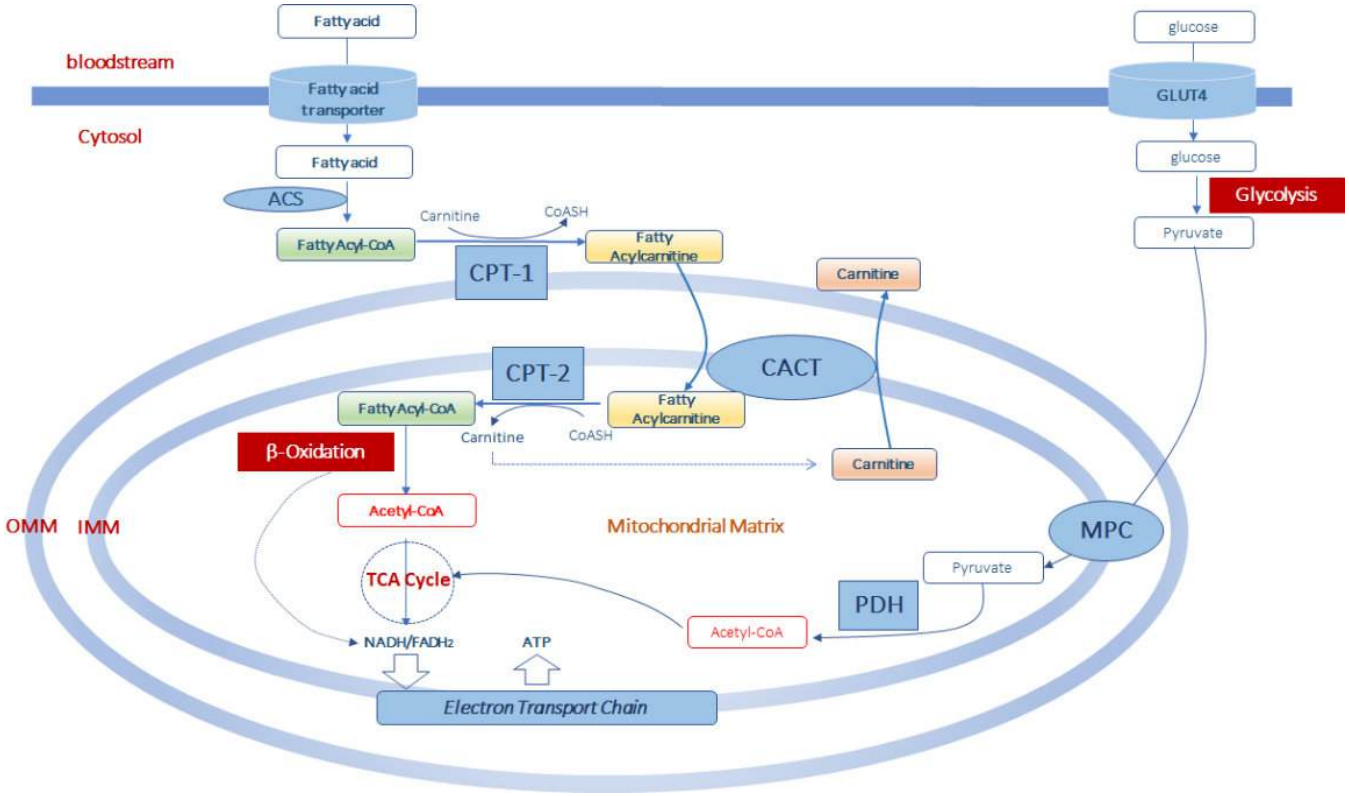


Figure 17.2.5: Carnitine cycle and the connections between major catabolic pathways.

At the outer mitochondrial membrane (OMM), fatty acyl-CoAs become linked to carnitine through carnitine palmitoyltransferase-1 (CPT-1). The complex is translocated across the inner mitochondrial membrane (IMM) via carnitine-acylcarnitine translocase (CACT). In the mitochondrial matrix, CPT-2 converts fatty acylcarnitines back to fatty acyl-CoAs, which enter the β-oxidation pathway. Free carnitine moves back into the cytoplasm through exchange with acyl-carnitines with CACT. β-oxidation in the matrix produces acetyl-CoA, which is also made from glycolytic pyruvate through pyruvate dehydrogenase. Hence acetyl-CoA links both glycolysis and fatty acid oxidation. The resulting acetyl-CoA can enter the TCA cycle when energy is needed.

The mitochondrial carnitine/acylcarnitine carrier protein helps transport acylcarnitines of different lengths across the mitochondrial inner membrane for β-oxidation into the mitochondrial matrix. Figure 17.2.6 shows an interactive iCn3D model of the human mitochondrial carnitine/acylcarnitine carrier protein AlphaFold model (O43772)

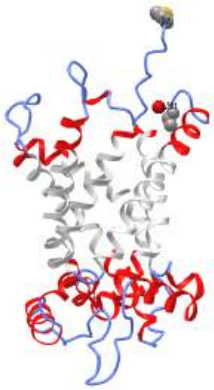


Figure 17.2.6: Mitochondrial carnitine/acylcarnitine carrier protein AlphaFold model (O43772). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...DdDNu9xVJGg2i9

The transmembrane helices are shown in gray. The N- (Met 1) and C-terminal (Leu 301) amino acids are shown in spacefill color CPK.

Malonyl-CoA produced in the first committed step in fatty acids synthesis inhibits CPT1. This should make biological sense since fatty acid oxidation should not occur as fatty acids are synthesized. Palmitoyltransferase II (CPT II), which serves to convert acylcarnitine to fatty acyl

CoA, traps the molecules within the mitochondrial matrix.

In contrast to this regulated transport mechanism, very long chain fatty acids (VLCFAs) and branched-chain fatty acids are transported into peroxisomes by the ABCD1-3 transporters through an ATP-dependent process

Mitochondrial β-oxidation of fatty acids has four steps that occur in the mitochondrial matrix. In those steps, a 16:0 fatty acid (for example) is converted to a (14):0 fatty acid and the 2C molecule acetyl-CoA. The (14):0 fatty acid undergoes 6 more rounds of the β-oxidation cycle until the entire 16:0 fatty acid is fully converted to 8 acetyl-CoAs.

Step 1: Acyl-CoA dehydrogenase

There are long- (LCAD), medium-(MCAD), and short-chain acyl-CoA dehydrogenases (SCAD) which catalyze the first oxidative step in the β-oxidation pathway. These enzymes catalyze the formation of a trans double bond between the α and β carbons (C2 and C3) on the acyl-CoA substrates. A stronger oxidizing agent than NAD^+ is required to form an alkene between the two methylene groups, so FAD is used. Eventually the reduced $FADH_2$ produced will lead to the production of 1.5 equivalents of ATP in the mitochondrial electron transport chain/oxidative phosphorylation.

Figure 17.2.7 shows the key oxidative step in the mechanism of acyl-CoA dehydrogenase to 2-enoyl-CoA by acyl-CoA dehydrogenases

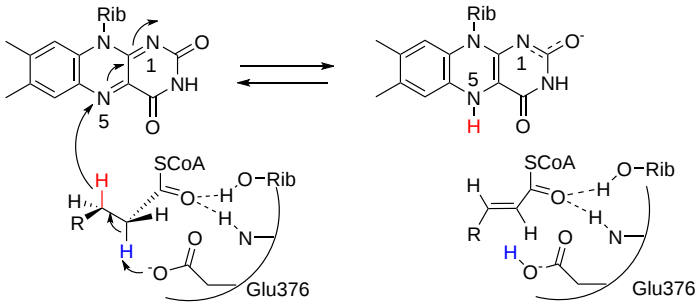
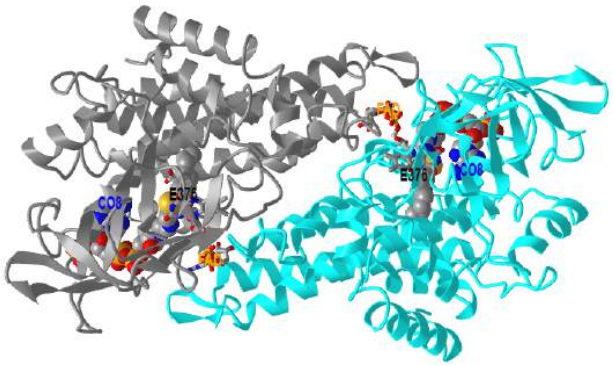


Figure 17.2.7: Mechanism for conversion of acyl-CoA to enoyl-CoA by acyl-CoA dehydrogenase. (after Ghisla and Thorpe. Eur. J. Biochem. 271, 494–508 (2004) doi:10.1046/j.1432-1033.2003.03946.x)

Figure 17.2.8 shows an interactive iCn3D model of the medium-chain acyl-CoA dehydrogenase from pig liver mitochondria with octanoyl-CoA, a substrate (3MDE)



NCBI iCn3D Figure 17.2.8: Medium-chain acyl-CoA dehydrogenase from pig liver mitochondria with octanoyl-CoA substrate (3MDE). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...SiAZExwXVV2TG6

Just two subunits of the biologically active tetramer are shown (dark gray and cyan). FAD is shown in each subunit (sticks, CPK colors, labeled). A bound substrate, octanoyl-CoA (spacefill, CPK colors, labeled CO8) is also shown in each subunit. The catalytic base, Glu 376, is shown in sticks, CPK colors, and labeled.

The structures of the unliganded and acyl-CoA forms of the enzymes are very similar, so there are no large conformational changes on binding octanoyl-CoA. The ligand binds to the enzyme at the rectus (re) face of the FAD with the acyl chain buried. The fatty acyl chain of the thioester substrate is buried inside of the polypeptide and the 3'-AMP moiety is close to the surface of the tetrameric enzyme molecule. The carbonyl oxygen of octanoyl-CoA interacts with the ribityl 2'-hydroxyl group of the FAD and the main-chain carbonyl oxygen of Glu-376. Glu-376 acts as a general base as it removed the alpha proton in the reaction.

Step 2. Enoyl CoA hydratase

This enzyme catalyzes a hydration step of the double bond between the α and β carbons (C2 and C3), adding an OH group to the β carbon, in a reaction that poses little energy barrier. A likely mechanism for enoyl-CoA hydratase is shown in Figure 17.2.9.

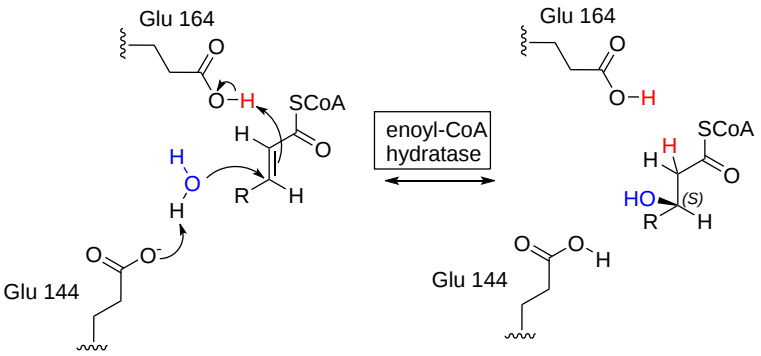
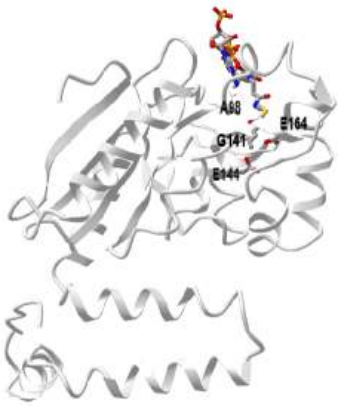


Figure 17.2.9: Mechanism for enoyl-CoA hydratase

Figure 17.2.10 shows an interactive iCn3D model of the rat enoyl-CoA hydratase in complex with hexadienoyl-CoA (1MJ3)



NCBI iCn3D Figure 17.2.10: Rat enoyl-CoA hydratase in complex with hexadienoyl-CoA (1MJ3). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...byx3eEKewEeEz9

Only one subunit of the biological hexamer is shown for clarity. Two glutamic acids (141 and 164) appear to activate a water molecule for the hydration reaction. Alanine 98 and Gly 141 appear also to be situated in an oxyanion hole which stabilizes the transition state and intermediate.

The addition of the water is syn since the proton and OH group are added to the same size of the double bond. The glycine amide NH provides a strong hydrogen bond to the carbonyl of the substrate, hexadienoyl-CoA, helping to polarize the ene-one. The substrate trans-2-crotonyl-CoA is converted to the 3(S) alcohol instead of the 3(R) alcohol by a huge factor. Both the cis and trans isomers of a substrate analog (hexadienoyl-CoA) can bind to the enzyme, but only the cis isomer is polarized. Since the transition state is polarized as well, it would appear the bound cis isomer is strained and destabilized, suggesting that its binding is an example of transition state binding catalysis.

Step 3. Beta-hydroxyl acyl CoA dehydrogenase

After the addition of the OH on C3 (beta) OH during the hydration reaction, the resulting ROH is oxidized to a ketone, β-ketoacyl-CoA, by the oxidizing agent NAD^+ using the enzyme β-hydroxyl acyl CoA dehydrogenase. The resulting NADH is reoxidized to NAD^+ through the mitochondrial electron transport chain, which leads to the formation of 2.5 molecules of ATP for each NADH. Figure 17.2.11 shows a plausible mechanism for the beta-hydroxyl acyl CoA dehydrogenase-catalyzed reaction.

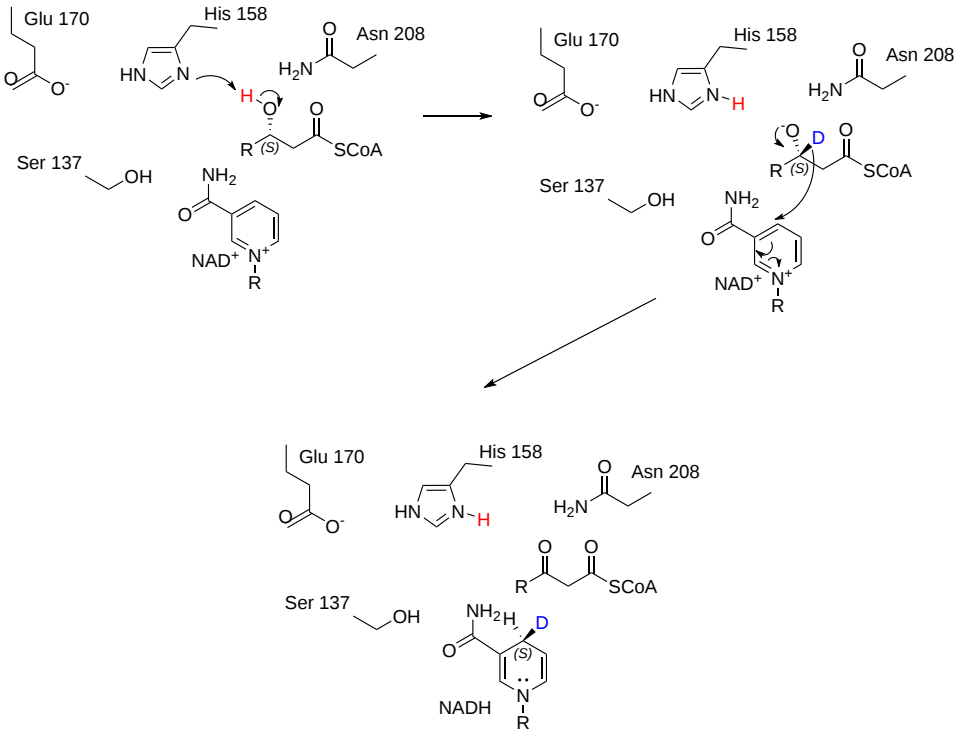


Figure 17.2.11: Mechanism for production of β-ketoacyl-CoASH by b**eta-hydroxyl acyl CoA dehydrogenase**

His 158 acts as a general base. Glu 170 increases the basicity of His 158. The other group is involved in H-bond and electronic stabilization interactions.

Step 4. **Acetyl-CoA acetyltransferase, mitochondrial - ACAT1** (also called 3-ketoacyl-CoA **thiolase**)_

The final step in the beta-oxidation pathway involves cleavage of the bond between the alpha and beta carbon by CoASH. This step is catalyzed by beta-keto thiolase and is a thiolytic (as opposed to cleavage by water - a hydrolysis) reaction. The reaction produces one molecule of acetyl CoA and a fatty acyl CoA that is two carbons shorter. The process repeats until the even chain fatty acid is completely converted into acetyl CoA. The activity of the enzyme is reversible and it can also catalyze the Claisen condensation of two acetyl-CoA molecules into acetoacetyl-CoA, as we will see in the synthesis of ketone bodies in the next chapter section.

The reaction starts with the acylation reaction of the nucleophilic Cys 89 with the carbonyl at the 3-oxoacyl-CoA, with the concomitant release of acetyl-CoA. This forms a Cys 89-acyl covalent intermediate. In the next step, Cys 378 acts as a general base to facilitate the nucleophilic attack of free CoASH on the acyl-intermediate. His 348 acting as a general acid protonates the thiolate leaving group. The amino acids (Cys89, Cys378, and His348) are generally conserved in thiolases (Bhaskar et al. 2020). Kinetically this mechanism is a ping-pong reaction. A reaction mechanism is shown in Figure 17.2.12.

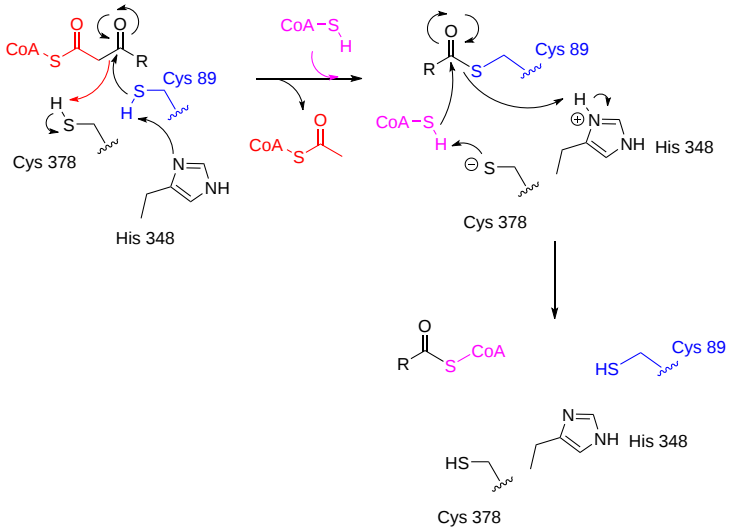
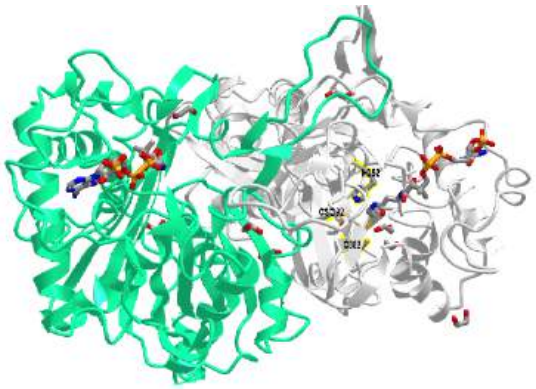


Figure 17.2.12 Reaction mechanism for mitochodrial acetyl-CoA acetyltransferase (ACAT1, thiolase T1)

Figure 17.2.13 shows an interactive iCn3D model of Human Mitochondrial 3-Ketoacyl-Coa Thiolase (T1) (4C2J)



NCBI iCn3D Figure 17.2.13: Human Mitochondrial 3-Ketoacyl-Coa Thiolase (T1) (4C2J). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...dzAHKEY77WpNi7

Two subunits in the biological function dimer are shown (cyan and gray). The active site is shown in the gray subunit as CPK-colored stick and labeled. The numbers are a bit different than shown in the mechanistic figure. CoASH is shown in each subunit as sticks. The fatty-acyl tail appears to bind in a tunnel.

17.2.4: A FEW MORE ENZYMES ARE NEEDED

Steps 1 through 4 outlined above apply to the beta-oxidation of a saturated fatty acid with an even-numbered carbon skeleton. Unsaturated fatty acids, such as oleate (18:1) and linoleate (18:2), contain cis double bonds that must be isomerized to the trans configuration by the enzyme **enoyl CoA isomerase** or reduced by the enzyme NADPH (2,4-dienoyl CoA reductase or 24DCR), using NADPH.

Enoyl CoA isomerase

This enzyme catalyzes the isomerization of cis double bounds to the trans form which mimics those formed by acyl-CoA dehydrogenase by FAD in step 1. Figure 17.2.14 shows a plausible mechanism for the conversion of cis double bonds to their trans isomer.

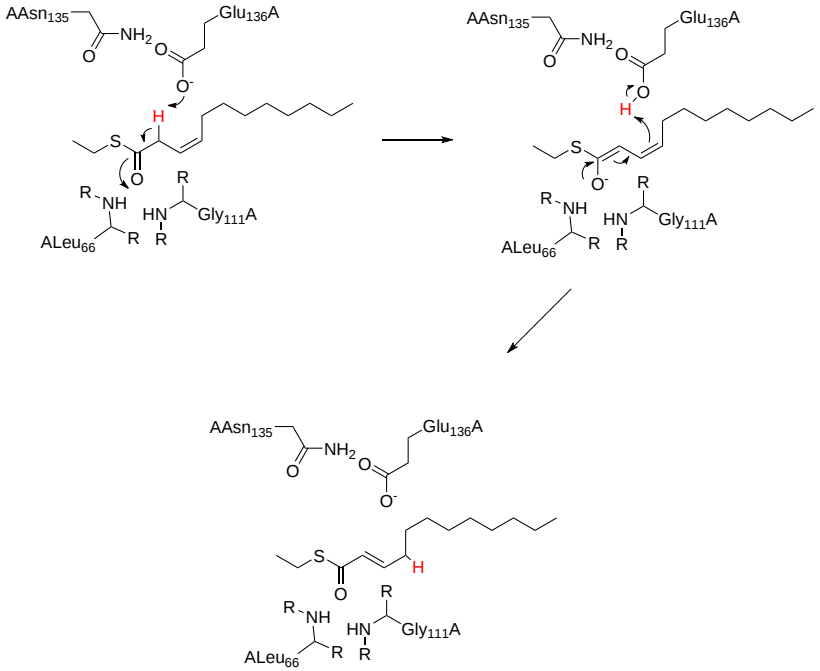


Figure 17.2.14 Cis to trans isomerization by enoyl-CoA isomerase (https://www.ebi.ac.uk/thornton-srv/m-csa/entry/341/)

Glu 136 acts as a general base, while amide Hs of Leu 66 and Gly 111 stabilized the intermediate oxyanion and hence the developing charge in the transition state. They are optimally situated in the oxyanion hole.

Figure 17.2.15 shows an interactive iCn3D model of the Human mitochondrial Δ^3-Δ^2-enoyl-CoA isomerase (1SG4)

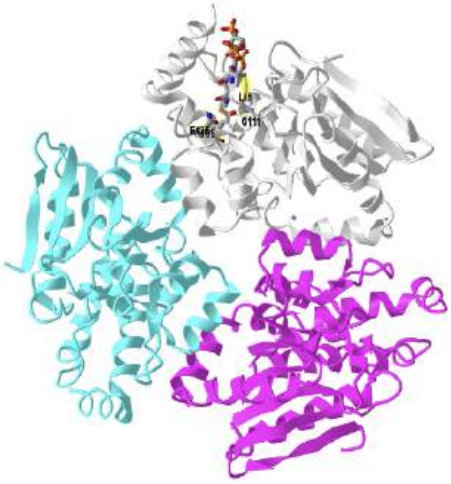


Figure 17.2.15: Human mitochondrial Δ^3-Δ^2-enoyl-CoA isomerase (1SG4) (Copyright; author via source). Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...H6LxEzBKajq8b7

The three subunits are shown in different colors. The substrate analog octanoyl-CoA is shown in spacefill CPK colors bound to the gray subunit. The catalytic residues are shown in sticks with CPK colors and labeled. The distal omega end binds in a hydrophobic tunnel.

2,4-dienoyl CoA reductase or **24DCR**

An alternative way to deal with the cis double bond is simply to reduce it, in this case, NADPH. This enzyme is used on all C=C at the even-number position and more at odd-numbered positions. A mechanism for the reduction is shown in Figure 17.2.16.

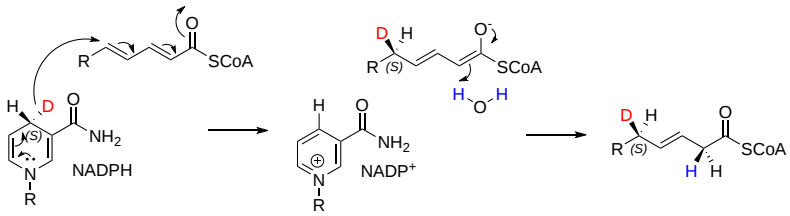


Figure 17.2.16: a mechanism for the reduction of double bonds in unsaturated fatty acids (after Fillgrove and Anderson, B*iochemistry* (2001) https://doi-org.ezproxy.csbsju.edu/10.1021/bi0111606

Figure 17.2.17 shows an interactive iCn3D model of a monomer of the homotetrameric human mitochondrial 2,4-dienoyl-Coa reductase (1W6U)

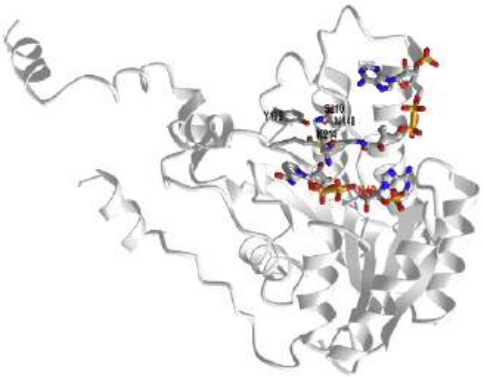


Figure 17.2.17: Monomer of the homotetrameric human Mitochondrial 2,4-Dienoyl-Coa Reductase (1W6U) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...pQueKc84sFqzt7

The model shows bound $NADP^+$ and the substrate trans-2,trans-4-dienoyl-CoA. The active site is open enough to accommodate fatty acids of different lengths. Tyr-199 and Asn-148 stabilize the enolate and the oxidized nicotinamide.

For odd-number chain fatty acids, propionyl-CoA to succinyl-CoA

Although most fatty acids of biological origin have even numbers of carbons, not all of them do. Oxidation of fatty acids with odd numbers of carbons ultimately produces an intermediate with three carbons, propionyl-CoA, which cannot be oxidized further in the beta-oxidation pathway. These additional steps are necessary:

1. carboxylation to make (S)-methylmalonyl-CoA;
2. isomerization to (R)-methylmalonyl-CoA;
3. rearrangement to form succinyl-CoA. The last step of the process utilizes the enzyme methylmalonyl-CoA mutase, which uses the B12 coenzyme in its catalytic cycle. Succinyl-CoA can then be metabolized in the citric acid cycle.

Figure 17.2.18 shows the pathway for the conversion of propionyl-CoA to succinyl-CoA.

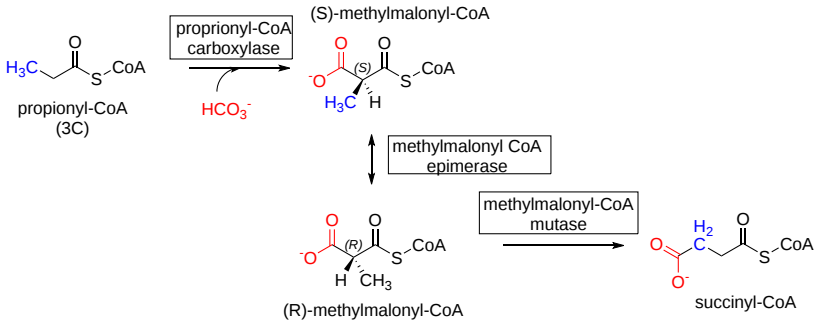


Figure 17.2.18 Conversion of propionyl-CoA to succinyl-CoA

Recent Updates: 7/24/23 Mechanisms for conversion of propionyl-CoA to succinyl-CoA.

We will look at each of the enzymes in turn.

Propionyl-CoA carboxylase

The net reaction for this enzyme is shown in Figure 17.2.19 below.

$$\text{Propionyl-CoA} + \text{ATP} + HCO_3^- \rightarrow \text{(S)-methylmalonyl-CoA} + \text{ADP} + P_i + H^+$$

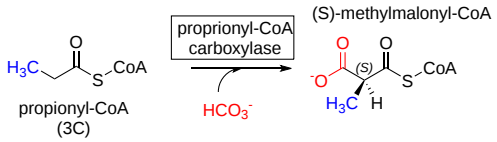


Figure 17.2.19: Propionyl-CoA carboxylase reaction

The enzyme has 3 subunits. The alpha subunits has biotin carboxyl carrier protein and biotin carboxylase domains. The beta subunit has carboxytransferase activity. The mechanism for the *Streptomyces coelicolor* propionyl-CoA carboxylase (PCC) is shown below in two parts.

In Part 1, biotin is carboxylated by HCO_3^- to form an activated bicarbonate derivative (similar to an anhydride) that is high energy in comparison to its hydrolysis product, as shown in Figure 17.2.20 below. This reaction takes place in the alpha subunit.

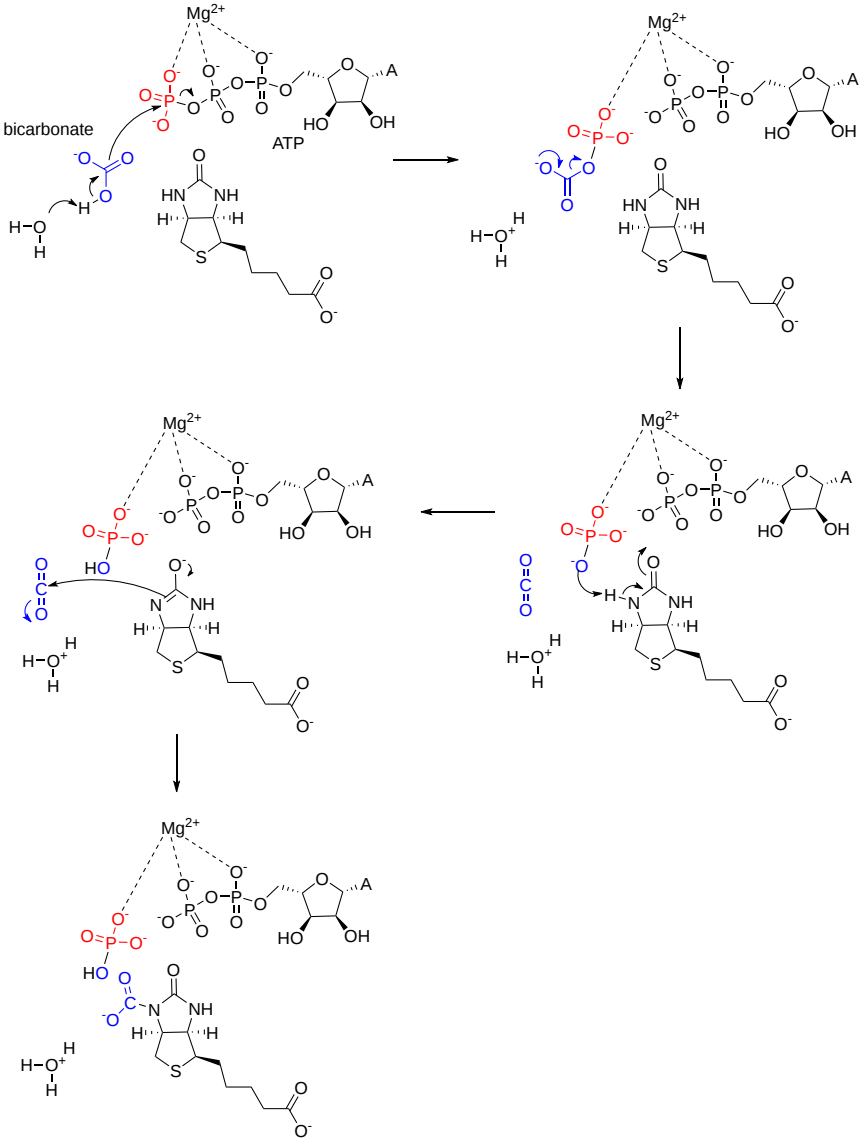


Figure 17.2.20: Formation of carboxylated biotin by propionyl-CoA carboxylase - Part 1. M-CSA. Gemma L. Holliday, Daniel E. Almonacid, Jonathan T. W. Ng. Creative Commons Attribution 4.0 International (CC BY 4.0) License

In Part 2, shown in Figure 17.2.21 below, the carboxyl group on biotin is transferred to propionyl-CoA to form methymalonyl-CoA. This reaction occurs in the beta subunit.

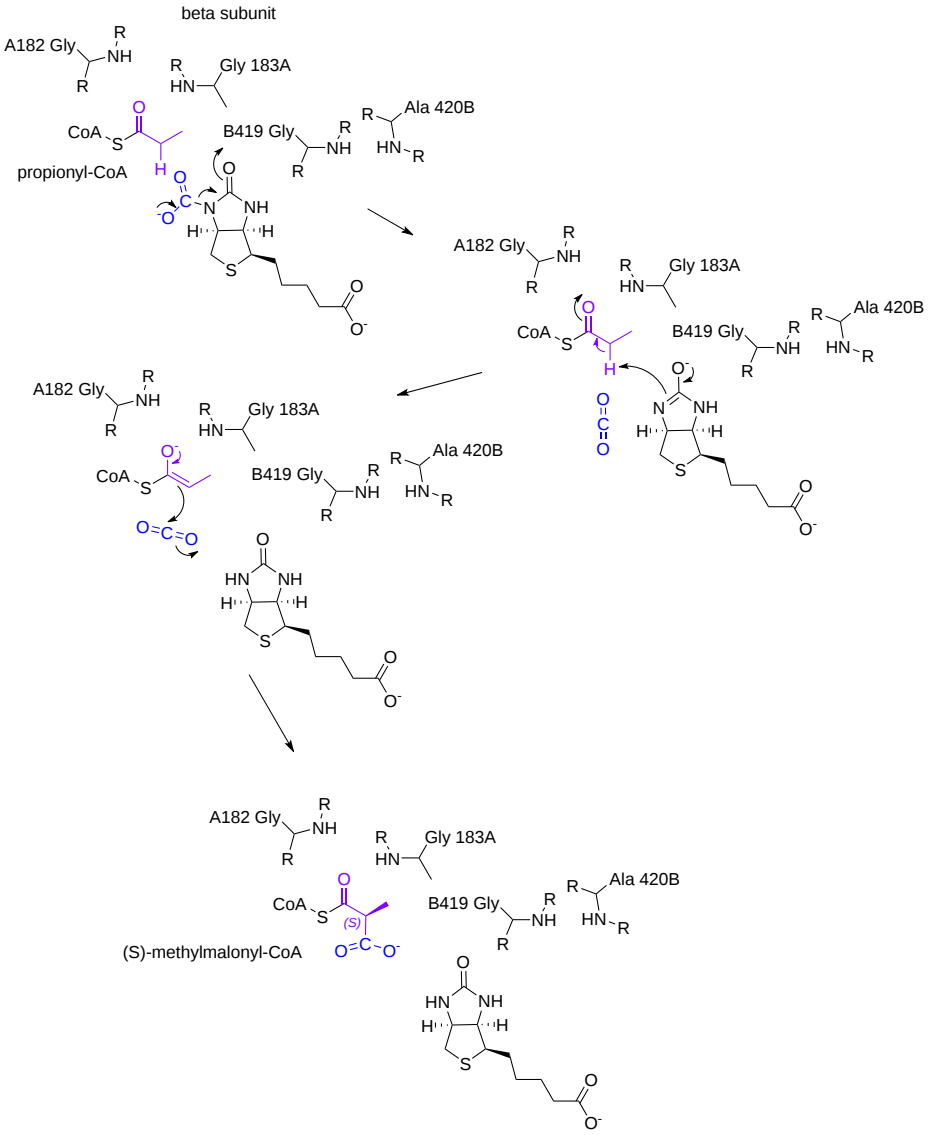
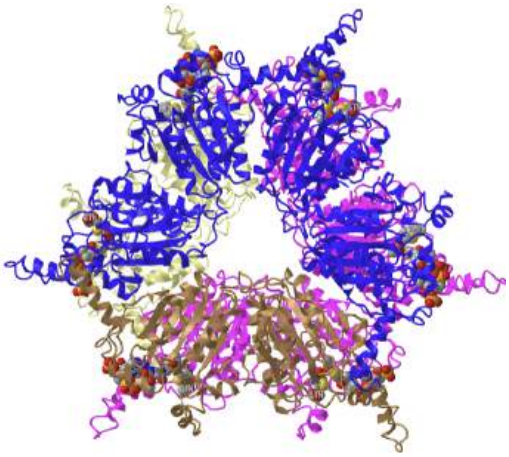


Figure 17.2.21: Part 2 - Formation of (S)-methylmalonyl-CoA by activated CO_2 transfer from carboxybiotin by propionyl-CoA carboxylase. Holliday et al., ibid.

The roles of the main chain atoms of Ala 182 and Gly 183 in the alpha subunit are to stabilize the propionyl-CoA while the backbone atoms of Gly 429 and Ala 430 in the beta subunit are to stabilize the oxyanion of carboxylated biotin.

Figure 17.2.22 shows an interactive iCn3D model of biotin and propionyl-CoA bound to Acyl-CoA Carboxylase Beta Subunit from S. coelicolor (1XNY)



NCBI iCn3D Figure 17.2.22: Monomer of the biotin and propionyl-CoA bound to Acyl-CoA Carboxylase Beta Subunit from S. coelicolor (1XNY). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...K46r5EdZezF4o6

Biotin (labeled BTN) and propionyl-CoA are shown in spacefill in CPK colors.

Methylmalonyl-CoA Epimerase

The net reaction is shown in Figure 17.2.23 below.

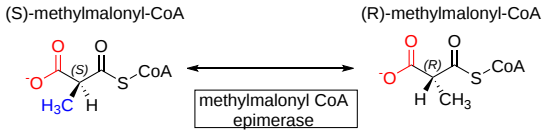


Figure 17.2.23: Methylmalonyl-CoA Epimerase reaction

This reaction proceeds through an enolate intermediate after an abstraction of a proton from the chiral center of the methylmalonyl-CoA, as shown in the reaction mechanism for the enzyme from propionibacterium freudenreichii subsp. shermanii is shown Figure 17.2.24 below. The reaction is readily reversible.

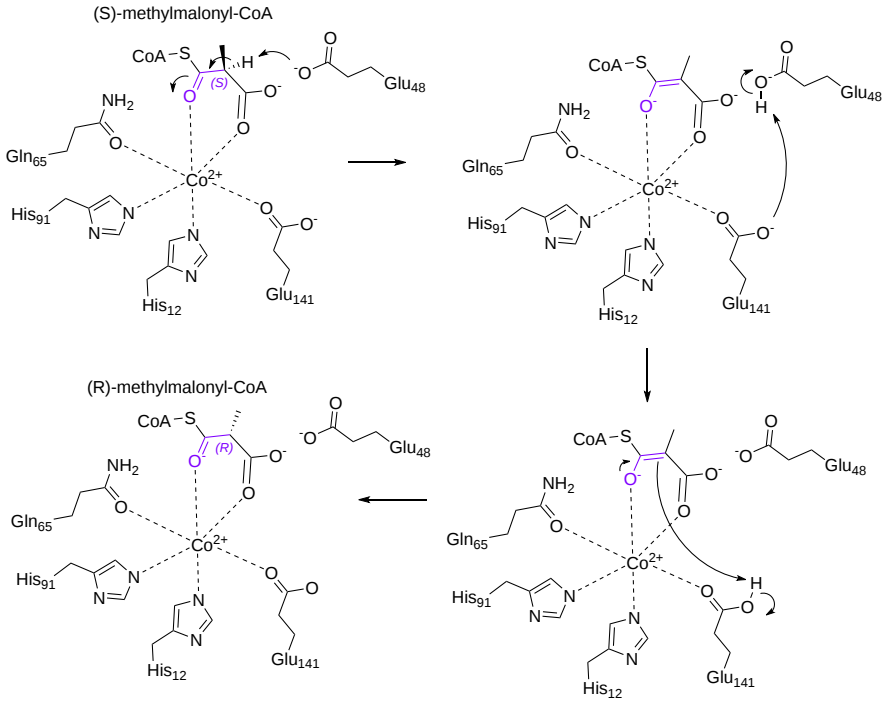
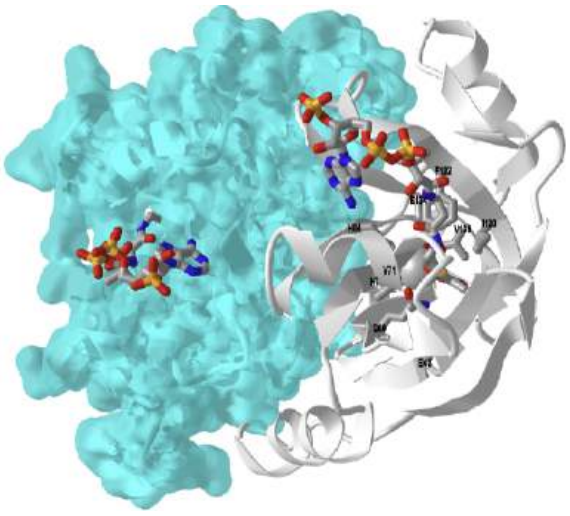


Figure 17.2.24: Mechanism for methylmalonyl-CoA Epimerase. Gemma L. Holliday et al. https://www.ebi.ac.uk/thornton-srv/m-csa/entry/33/. Creative Commons Attribution 4.0 International (CC BY 4.0) License

Note the presence of a Co^{2+} ion in the active site.

Figure 17.2.25 shows an interactive iCn3D model of Methylmalonyl-CoA epimerase in complex with 2-nitronate-propionyl-CoA from S. coelicolor (6WFI)



NCBI iCn3D Figure 17.2.25: Methylmalonyl-CoA epimerase in complex with 2-nitronate-propionyl-CoA from S. coelicolor (6WFI). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...Sqc98CCjAvzkp7

The enzyme is a dimer with two identical subunits and catalytic. One of the subunits is shown with a blue transparent surface with the bound 2-nitronate-propionyl-CoA inhibitor. The other is shown in gray cartoon with the key side chains involved in substrate binding and catalysis shown as color sticks and labeled.

Methylmalonyl-CoA Mutase

In this reaction, a methyl group is removed from (R)-methylmalonyl-CoA to form succinyl-CoA, a shown in Figure 17.2.26 below.

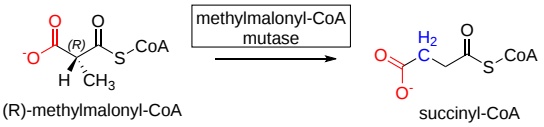


Figure 17.2.26: Methylmalonyl-CoA reaction

This reaction going in the opposite direction is an example of a methyltransferase. Another similar enzyme (which we will see in Chapter 18.4 - Amino Acid Degradation) catalyzes a methyl transfer from homocysteine to a new cofactor, **cobalamin**, which transfers it to cysteine to form methionine in a reaction catalyzed by methionine synthase. Cobalamine and its methylated form are derivatives of Vitamin B12. We will leave details of cobalamin biochemistry to the next chapter and will present a mechanism for methylmalonyl-CoA mutase here.

We present the reaction for the reverse reaction, the conversion of succinyl-CoA to (R)-methylmalonyl-CoA by the enzyme from *Propionibacterium freudenreichii subsp. shermanii* in two parts below. The enzyme is a heterodimer of an alpha and beta subunit and has a cofactor, adenosylcobalamin (coenzyme B12). In contrast, the human mutase, a homodimer, is very similar to the alpha subunit.

First, a free radical is formed from the adenosyl group on the cofactor. This promotes a free-radical rearrangement of succinyl-CoA to (R)-methylmalonyl-CoA (or the reverse for the pathway of interest here). The reaction is shown in two parts for more optimal viewing.

Figure 17.2.27 below the first parts of the reaction for the conversion of succinyl-CoA to (R)-methylmalonyl-CoA by the mutase from *Propionibacterium freudenreichii subsp. shermanii*.

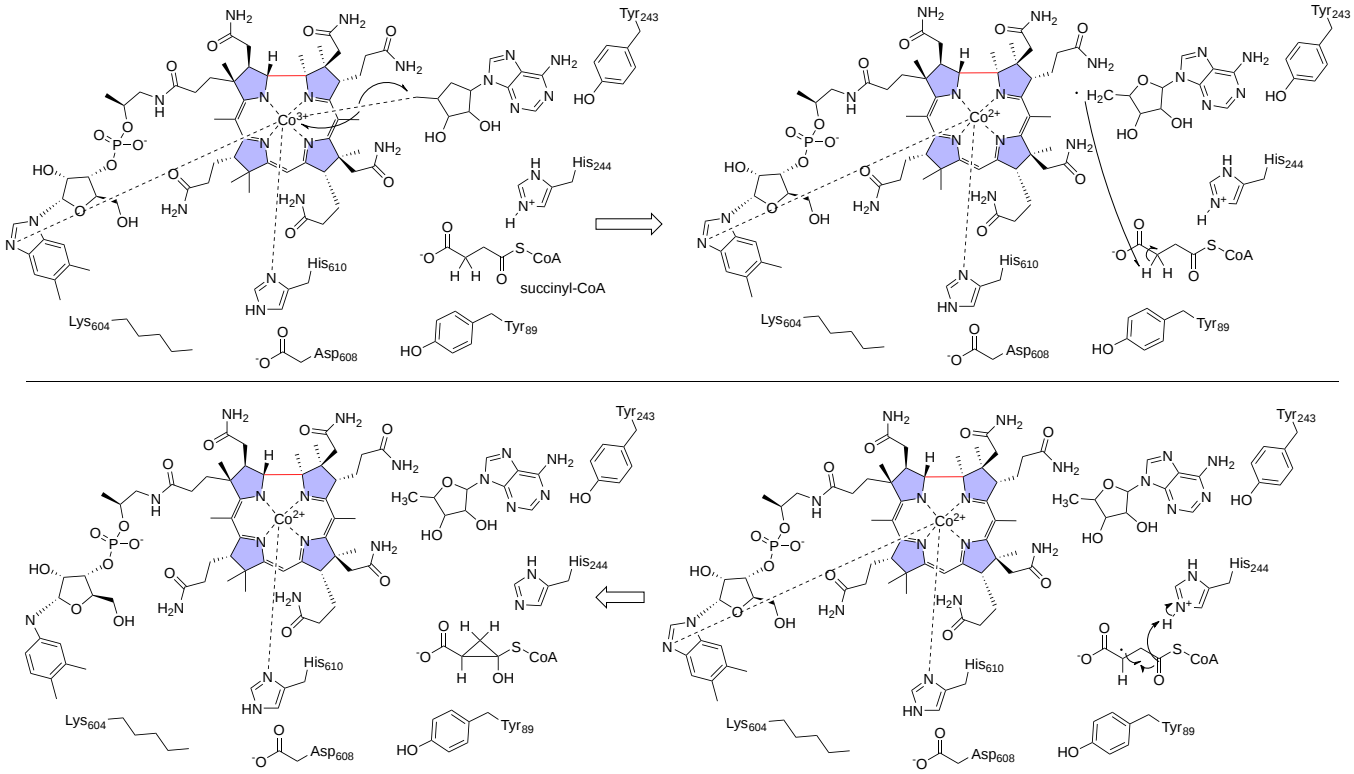


Figure 17.2.27: Part 1 of conversion of succinyl-CoA to (R)-methylmalonyl-CoA by methylmalonyl-CoA mutase. Gemma L. Holliday, Gail J. Bartlett, Daniel E. Almonacid. M-CSA. https://www.ebi.ac.uk/thornton-srv/m-csa/entry/62/. Creative Commons Attribution 4.0 International (CC BY 4.0) License

Figure 17.2.28 below shows the rest of the reaction for the conversion of succinyl-CoA to (R)-methylmalonyl-CoA.

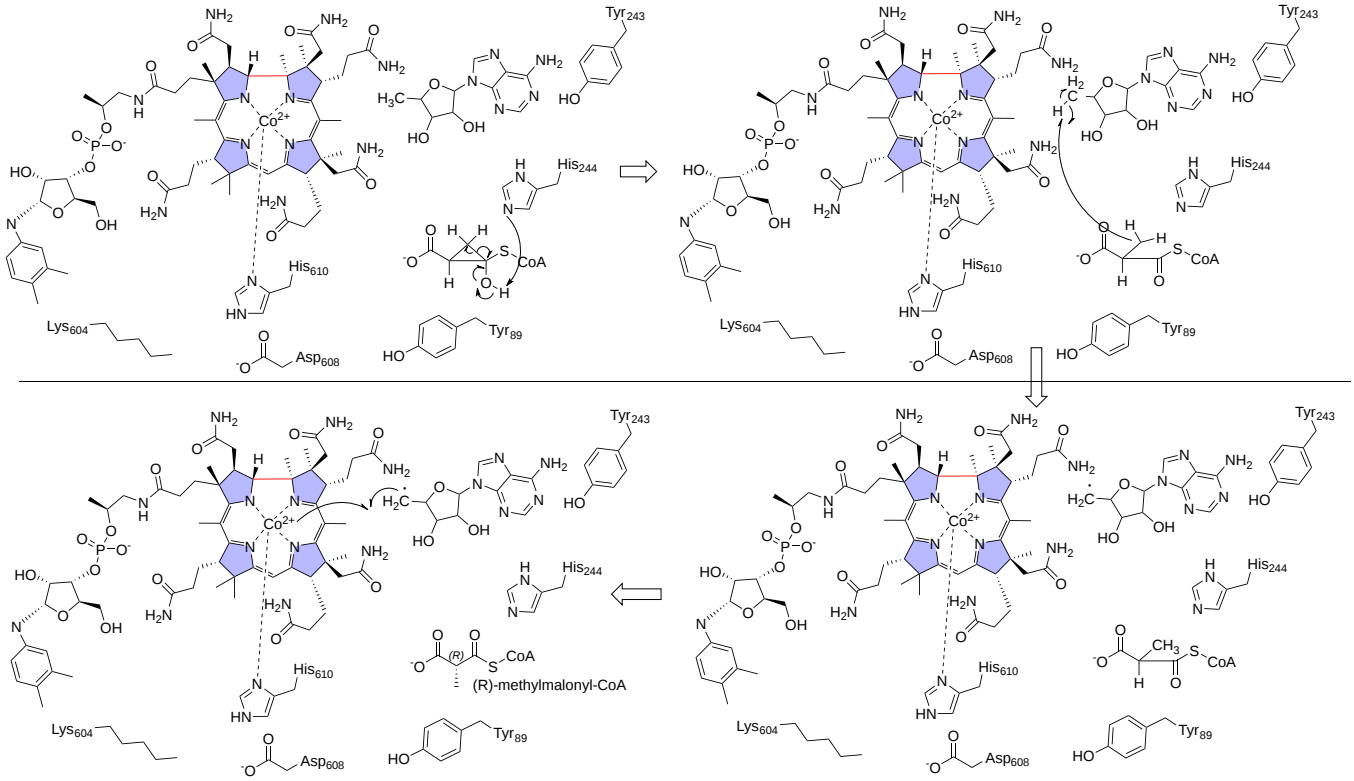
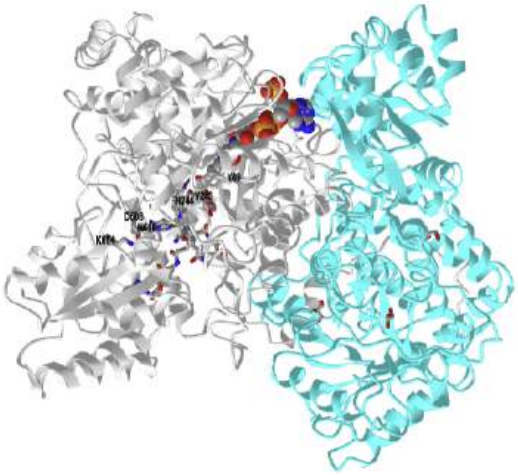


Figure 17.2.28: Part 2 of the conversion of succinyl-CoA to (R)-methylmalonyl-CoA by methylmalonyl-CoA mutase. Gemma L. Holliday et al., ibid.

The very last product from the bottom left reaction, the reformed active adenosylcobalamin cofactor, is not shown. The enzyme facilitates the hemolytic cleavage of the Co-C bond. This is followed by a free-radical rearrangement.

Figure 17.2.29 shows an interactive iCn3D model of the methylmalonyl-coenzyme A mutase from from *Propionibacterium freudenreichii subsp. shermanii* (1REQ)



NCBI iCn3D Figure 17.2.29: Methylmalonyl-coenzyme A mutase from from *Propionibacterium freudenreichii subsp. shermanii* (1REQ). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...qPYm2sXjraCuN8

An analog of CoASH, desulfo-CoA (DCA) , is shown in spacefill, CPK colors and labeled. Adenosylcobalamin is shown in colored sticks and labeled B12. The alpha subunit is shown in gray with the key amino acids from the enzyme mechanism presented above shown as colored sticks and labeled. The beta subunit is shown in blue.

Propionyl-CoA is also produced as a product of the oxidation of methionine, valine, isoleucine, and threonine. (See the amino acid metabolism chapter for more details on mechanisms.)

17.2.5: VERY LONG CHAIN OXIDATION

In contrast to the oxidation of short and medium-chain fatty acids, which under beta-oxidation require four different discrete enzymes, the oxidation of very long-chain fatty acids (VLCFs) is carried out by two proteins with the second protein expressing three enzyme activities. The first step, analogous to step 1 in beta-oxidation described above, is carried out by a very long chain acyl-CoA dehydrogenase (VLCAD). The next three reactions are carried out by a single trifunctional protein (TFP) with two subunits. The α-subunit carries out the hydration (2-enoyl-CoA hydratase, ECH) and next oxidation step (3-hydroxyl-CoA dehydrogenase (HAD), while the β-subunit has 3-ketothiolase (KT) activity. Deficiencies of TFP can cause significant disease and even death.

TFP is a heterotetramer of two α and two β subunits with the beta subunits forming a central homodimer. The two alpha units bind at each end and the whole complex forms an arc. There appears to be a "tunnel" allowing substrate transfer after each step. This minimizes premature intermediate release. Figure 17.2.30 shows an interactive iCn3D model of the human mitochondrial trifunctional protein, a fatty acid beta-oxidation metabolon (6DV2)

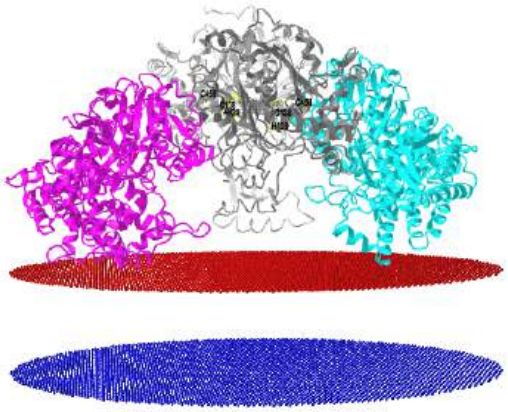


Figure 17.2.30: Human mitochondrial trifunctional protein fatty acid beta-oxidation metabolon (6DV2). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...XiUXru4vkRVh87

- Grays: two thiolase subunits: reversible thiolytic cleavage of 3-ketoacyl-CoA into acyl-CoA and acetyl-CoA, a 2-step reaction involving a covalent intermediate formed with a catalytic cysteine. The catalytic residue (C138, C458, and H428) are shown as sticks with CPK colors.
- Cyan and magenta subunits: enoyl-CoA hydratase (EC 4.2.1.17) and 3-hydroxyacyl-CoA dehydrogenase (EC 1.1.1.35)
- The red dots represent the inner leaflet of the inner membrane so the proteins reside in the mitochondrial matrix.

17.2.6: REGULATION OF BETA-OXIDATION

We saw that the metabolic decision to use carbohydrate energy reserves (glycogen) is a highly regulated event. Glycogen breakdown occurs during fasting and energy need. Fatty acids, our largest stores of energy, are released from triglyceride reserves in adipose cells through extracellular epinephrine and glucagon activation of pathways that activate intracellular hormone-sensitive lipase. Released fatty acids are bound to the serum protein albumin which transports them to tissue. Also as we mentioned above, malonyl-CoA inhibits the movement of fatty acids into the mitochondria. Malonyl-CoA is the first committed product of fatty acid biosynthesis. Each acyl-CoA product of each of the four enzymes engages in product inhibition for the enzyme that produced it. 3-ketoacyl-CoA also inhibits enoyl-CoA hydratase and acyl-CoA dehydrogenase [17]. The NADH/NAD^+ and acetyl-CoA/CoA ratios also influence the beta oxidation pathway through allosteric regulation. For example, the ratio of acetyl-CoA/CoA ratio affects the activity ketoacyl-CoA thiolase.

Fatty acids also bind to the transcription factors called peroxisome proliferator-activated receptors (PPARs) and also coactivator PGC-1α, which regulate the transcription of enzymes in the beta-oxidation pathways. PPARs and the retinoid X receptor form heterodimers that bind to the PPAR response element in key promoters site involved in fatty acid degradation. These include CPT1, long-chain acyl-CoA dehydrogenase (LCAD), medium-chain acyl-CoA dehydrogenase (MCAD), and acyl-CoA synthetase (ACS). The PPARs have tissue specificity. We will discuss PPARs in more detail in the chapter on fatty acid synthesis.

17.2.7: PEROXISOMAL **A**-OXIDATION

Alpha oxidation of fatty acids occurs in the peroxisome. It is used to metabolize phytanic acid (3,7,11,15-tetramethyl hexadecanoic acid), found in dairy products, animal fat, and some fish. Phytanic acid is produced in ruminants on the degradation of plant material and derives from phytol, an isoprenoid alcohol esterified to chlorophyll. Phytol is first converted to phytanic acid.

Fatty acid β-oxidation can also occur in peroxisomes. In animals, peroxisomes are believed to be important in the initial breakdown of very long-chain fatty acids and methyl-branched fatty acids [11]. The enzymes involved in fatty acid oxidation in peroxisomes are different from mitochondria. An important difference is acyl-CoA oxidase, the first enzyme in peroxisome β-oxidation, which transfers the hydrogen to oxygen producing H_2O_2 instead of producing $FADH_2$. The H_2O_2 is broken down to water by catalase. The fatty acyl-CoA intermediates formed during β-oxidation are the same in peroxisomes and mitochondria. Peroxisomes also contain the necessary enzymes for α-oxidation, which are necessary for the oxidation of some fatty acids with methyl branches.

Branched-chain fatty acids also require additional enzymatic modification to enter the alpha-oxidation pathway within peroxisomes. Phytanic acid (3,7,11,14-tetramethylhexadecanoic acid), requires additional peroxisomal enzymes to undergo beta-oxidation. Phytanic acid initially forms phytanyl CoA which is then hydroxylated at the alpha carbon by phytanyl CoA hydroxylase (alpha-hydroxylase), encoded by the PHYH gene. The alpha carbon-hydroxyl bond then undergoes two successive rounds of oxidation to pristanic acid. Pristanic acid undergoes beta-oxidation, which produces acetyl CoA and propionyl CoA in alternative rounds. As with peroxisomal beta-oxidation of VLCFAs, this process generally ends when the carbon chain length reaches 6-8 carbons, at which point the molecule is shuttled to the mitochondria by carnitine for complete oxidation to carbon dioxide and water. Figure 17.2.31 shows the steps in the catabolism of phytanic acid (3,7,11,14-tetramethylhexadecanoic acid).

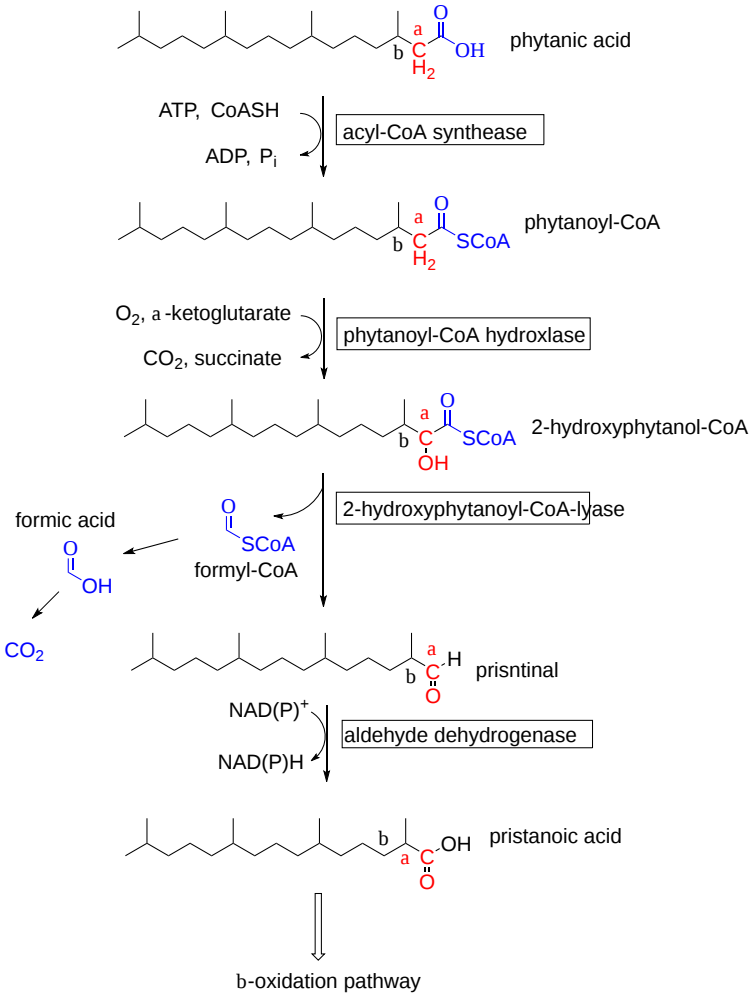


Figure 17.2.31: Alpha oxidation pathway of catabolism of phytanic acid (3,7,11,14-tetramethylhexadecanoic acid)

17.2.8: OMEGA-OXIDATION

The omega-oxidation pathway occurs in the endoplasmic reticulum and is used to metabolize larger fatty acids, which given their hydrophobicity might be damaging to cells in high concentrations. In this pathway the fatty acids are metabolized to dicarboxylic acids which increases their water solubility for excretion in the urine. The first step uses cytochrome P450 enzymes that are also used to modify xenobiotic compounds with dioxygen, making them more soluble as well. The omega oxidation pathway is shown in Figure 17.2.32.

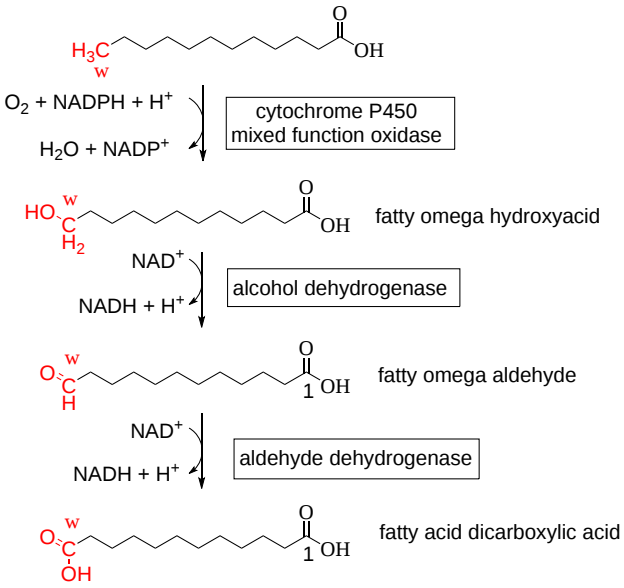


Figure 17.2.32 The ER omega oxidation pathways for catabolism of fatty acids

Three subfamily members of the cytochrome P450s (CYPs) show a preference for hydroxylation of short-chain fatty acids (C7-C10, CYP4B), medium-chain (C10-C16, CYP4A), and long-chain (C16-C26, CYP4F) fatty acids. which can be saturated, unsaturated and branched, fatty acids. Figure 17.2.33 shows a summary of the alpha, beta, and omega oxidation pathway

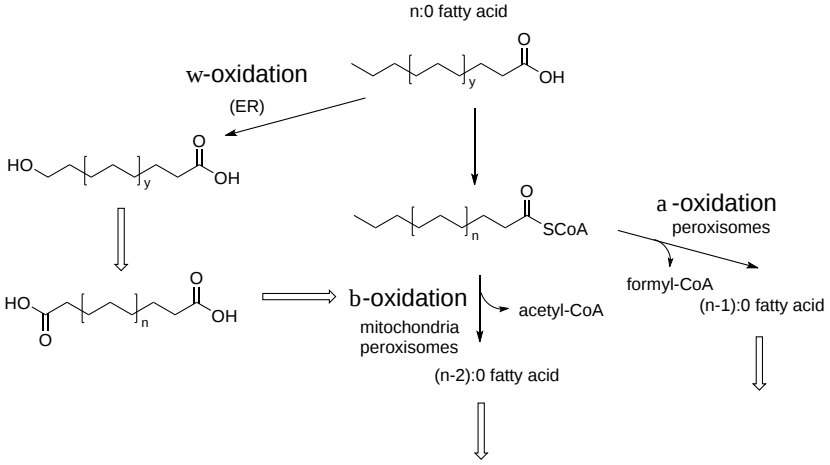


Figure 17.2.33 https://febs.onlinelibrary.wiley.com...8.2010.07947.x

17.2.9: DISEASES OF FATTY ACID METABOLISM

Listed below are a few select diseases that either directly involves defective fatty acid metabolism through intrinsic enzyme deficiencies or indirectly prevent the proper functioning of fatty acid metabolism through extrinsic enzyme deficiencies. Many, but not all, deficiencies of enzymes involved in fatty acid oxidation result in abnormal neurological development and or function early in life; a brief list of signs and symptoms appears under the selected diseases mentioned.

MCAD Deficiency

Medium-chain acyl dehydrogenase is the most common inherited defect of fatty acid oxidation in humans; as one would expect, medium-chain, 6-8 carbon molecules accumulate in this disease. Clinical manifestations of MCAD deficiency primarily present during fasting conditions and include lethargy, weakness, sweating, and hypoglycemia, most commonly in children under the age of 5. Serum measurements of octanoyl carnitine are usually elevated in these patients and can aid in the diagnosis. These abundant molecules then undergo oxidation by the cytochrome P450 system involved in omega-oxidation, resulting in a dicarboxylic acidemia and dicarboxylic aciduria.

Zellweger Syndrome

Zellweger syndrome results from autosomal recessive mutations in the PEX genes, which code for peroxin proteins needed for the assembly of peroxisomes. Almost 70% of all peroxisomal biogenesis disorders (PBDs) result from a PEX1 gene mutation. Many different fatty acid

compounds can accumulate without the oxidative machinery of peroxisomes, including VLCFAs and phytanic acid. Manifestations of this disease generally include the brain, kidneys, and skeleton.

X-Linked Adrenoleukodystrophy (X-ALD)

X-ALD is a genetic deficiency of the ABCD transporters in the membrane of peroxisomes, as mentioned previously, resulting in the pathological accumulation of phytanic acid and VLCFAs within cells and is most clinically significant when the ABCD1 transporter is absent. The disease presents with neurodegenerative and adrenal abnormalities.

Refsum Disease

Refsum disease results from a genetic deficiency of the enzyme phytanyl CoA 2-hydroxylase, which, as previously mentioned, is involved in the alpha-oxidation of phytanic acid, a breakdown product of chlorophyll. Notable clinical manifestations of Refsum disease include cardiac malfunction and defective functioning of the olfactory and auditory nerves due to the accumulation of phytanic acid.

17.3: KETONE BODIES

Much of the material below derives from Kolb et al. BMC Medicine (2021) 19:313 https://doi.org/10.1186/s12916-021-02185-0. Creative Commons Attribution 4.0 International License, http://creativecommons.org/licenses/by/4.0/.

17.3.1: INTRODUCTION

Ketone bodies are the name given to two molecules, acetoacetate (a ketone) and D-β-hydroxybutyrate (its reduction product) that derive from the condensation of acetate in the form of acetyl-CoA. Their structures are shown in Figure 17.3.1.

acetoacetate

R-b-hydroxybutyrate
D-b-hydroxybutyrate

Figure 17.3.1: Structure of ketone bodies

They are synthesized when acetyl-CoA concentrations are high. That would occur under a couple of conditions:

- fatty acids are being mobilized in adipocytes and broken down into acetyl-CoA through β-oxidation in the mitochondria;
- the acetyl-CoA produced can not enter the Kreb cycle leading to the buildup of acetyl-CoA in the matrix.

The latter condition arises when carbohydrate metabolism is compromised so the citric acid cycle has slowed. This occurs if the citric acid cycle intermediate oxaloacetate is pulled away from the cycle for the gluconeogenic synthesis of glucose. Pyruvate, another source of acetyl-CoA (through pyruvate dehydrogenase) is also depleted in gluconeogenesis. Of course, we are describing conditions in the liver, the major site of gluconeogenesis, and also of ketone body synthesis.

These conditions are also met in the underfed or starving state when glycogen supplies are exhausted and fatty acid reserves are being used for energy. Likewise, these conditions are met in diabetes when glucose might be abundant in the circulation but its entry into the cell is compromised by problems with insulin signaling.

Fatty acids released by adipocytes into the blood are bound to serum albumin for transport to tissue. Albumin, with bound fatty acids, does not, however, cross the blood-brain barrier. Ketone bodies are very small and water soluble so they can cross the blood-brain barrier, where they can deliver a soluble equivalent of fatty acids to the brain and other tissues in times of energy need. Ketone bodies can be considered solubilized and easily transportable fatty acid equivalents. They have low pKa values of 3.6 (acetoacetate) and 4.7 (D-β-hydroxybutyrate) so they can lead to "ketoacidosis" in diabetics. Acetone is a nonacidic and sweet-smelling molecule, which can be detected in the breath of diabetics, is produced spontaneously and through catalysis through the decarboxylation of the beta-keto acid acetoacetate. It is also considered a ketone body.

BLOOD-BRAIN BARRIER

The blood–brain barrier (BBB) protects the entire nervous system (CNS). Endothelial cells line the surface of all blood vessels. In the CNS there are additional structures that prevent the entry of many molecules. Tight and adherence junctions form between endothelial cells that prevent the movement of solutes between the cells. Additional protection is provided by the basement membrane, glial cells, and pericytes. In addition, drug efflux transporters (ABC transporters) move xenobiotics (foreign toxic molecules back through the barrier.

Nonpolar molecules like dioxygen molecules and low molecular weight can cross the BBB by simple diffusion. Other molecules like glucose and amino acids are carried across by transporters. Larger substances can move in through receptor-mediated endocytosis. Albumin, the major carrier of free fatty acids, like other proteins, does not readily pass through the barrier, and its drug-macromolecule complex, cannot cross.

Figure 17.3.2 shows a representation of the blood-brain barrier.

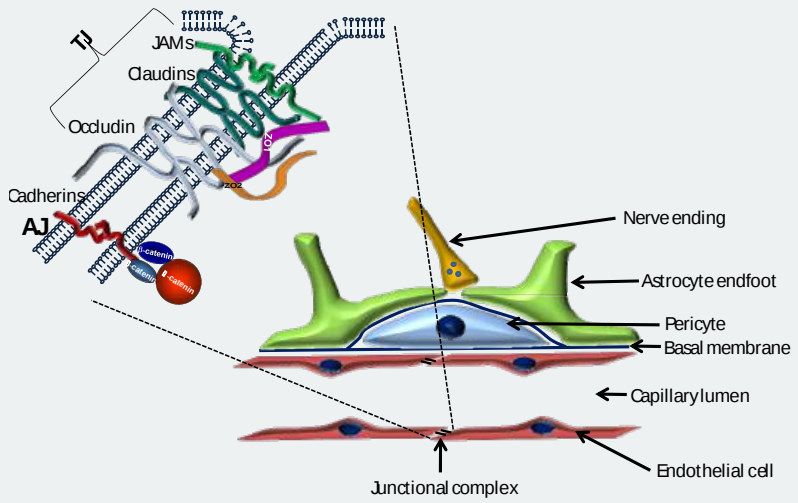


Figure 17.3.2: Structure of the blood-brain barrier. Wilhelm et al. Int. J. Mol. Sci. 2013, 14, 1383-1411; doi:10.3390/ijms14011383. Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0/).

In this section, we will explore the synthesis and utilization of ketone bodies. The synthesis pathway is called **ketogenesis**, and the degradation pathway is called **ketolysis**. Ketone body production and use are intimately tried to fatty acid metabolism so it is appropriate to discuss them right after fatty acid oxidation. An overview of both ketogenesis and ketolysis is shown in Figure 17.3.3.

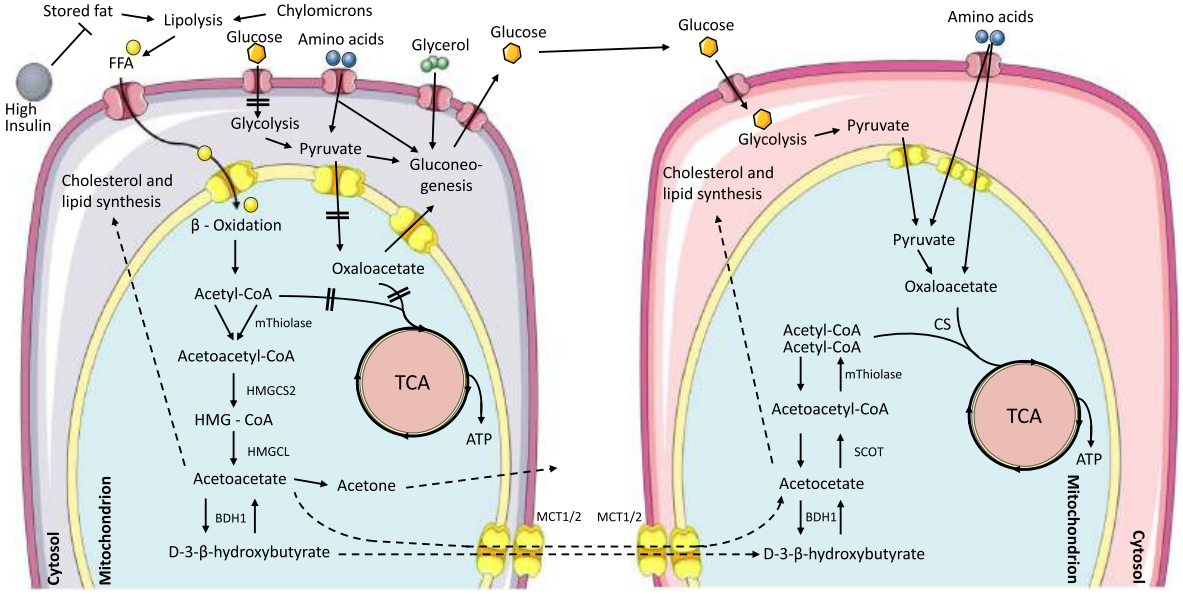


Figure 17.3.3: Overview of the ketogenesis (left) and ketolysis (right) pathways. FFA, free fatty acids; mThiolase, mitochondrial thiolase; HMGCS2, hydroxy methylglutaryl-CoA synthase; HMGCL, HMG-CoA lyase; BDH1, mitochondrial βOHB dehydrogenase; MCT1/2, monocarboxylate transporter 1 and 2; SCOT, succinyl-CoA:3-oxoacid-CoA transferase; CS, citrate synthase. Kolb et al. BMC Medicine (2021) 19:313 https://doi.org/10.1186/s12916-021-02185-0 Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Most acetoacetate is converted to its reduction product, β-hydroxybutyrate, for blood transport. It can spontaneously or through enzymatic catalysis, can be converted to acetone since it is a β-ketoacid with a built-in electron sink that stabilizes the negative charge in the transition state and intermediate. When ketone bodies enter cells, the β-hydroxybutyrate is converted back to acetoacetate and eventually to acetyl-CoA.

17.3.2: KETOGENESIS: THE DETAILS

Figure 17.3.4 shows the chemical structures of all intermediates in the synthesis of ketone bodies.

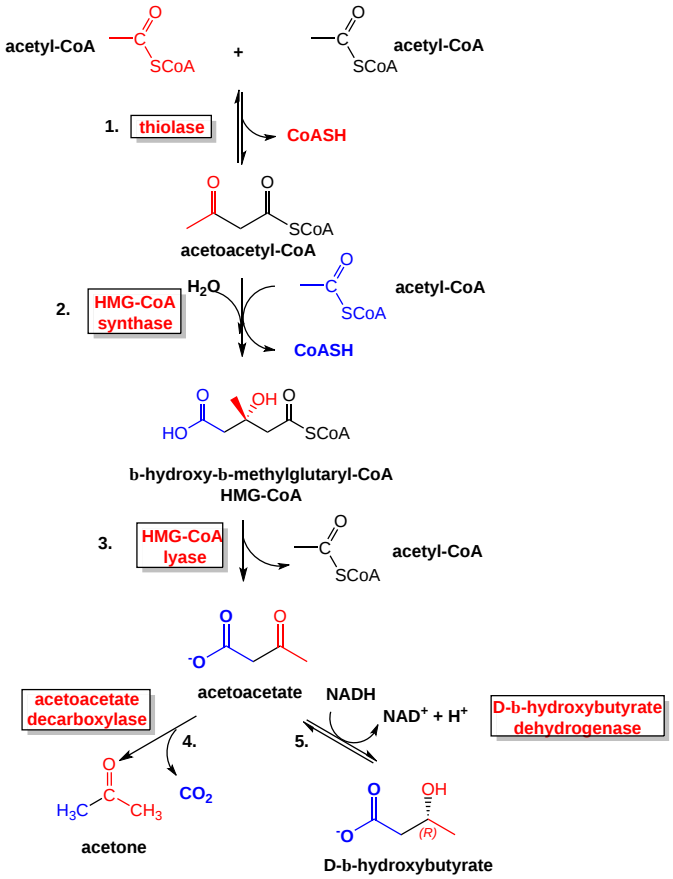


Figure 17.3.4: Ketogenesis

We'll study three of the reactions in ketogenesis in more detail below.

Reaction 1: Thiolase

We have presented the mechanism of thiolase in Step 4 of mitochondrial β-oxidation of fatty acid using the enzyme **acetyl-CoA acetyltransferase (ACAT1)** also called 3-ketoacyl-CoA **thiolase**. The final step in the beta-oxidation pathway involves cleavage of the bond between the alpha and beta carbon by CoASH. This step is catalyzed by beta-keto thiolase and is a thiolytic (as opposed to cleavage by water - a hydrolysis) reaction. The reaction produces one molecule of acetyl CoA and a fatty acyl CoA that is two carbons shorter. The process repeats until the even chain fatty acid is completely converted into acetyl CoA. The activity of the enzyme is reversible and it can also catalyze the Claisen condensation of two acetyl-CoA molecules into acetoacetyl-CoA, which occurs in the first step of ketogenesis. This step makes a C-C bond.

Reaction 2: HMG-CoA Synthase

This enzyme catalyzes the condensation of acetoacetyl–CoA (AcAc–CoA) and acetyl–CoA (Ac-CoA) to form 3-hydroxy-3-methylglutaryl (HMG)–CoA, through the formation of a C-C bond. Forming C-C bonds is critical to all biosynthesis. It's not a simple process. Cofactors are often used such as thiamine pyrophosphate, as we saw with pyruvate dehydrogenase. The reverse process, breaking a C-C bond, is also difficult unless the atoms are activated. One way to activate a carbon atom is to oxidize it through hydroxylation, often using dioxygen in reactions catalyzed by cytochrome P450. Likewise, C-C bond formation often occurs through the reaction of radicals in hemolytic reactions or through the rearrangement of key ionic intermediates. Radical reactions can be between a carbon free radical and heme derivatives. Radical reactions involve heme cofactors (as in P450), non-heme iron proteins (where the iron ion can cycle between different oxidation states), flavoproteins (in which the flavin can undergo both 1 and 2 electron redox steps), radical *S*-adenosylmethionine (SAM) enzymes, and cobalamins.

HMG-CoA synthase is somewhat unique in that it forms a C-C bond by activating the methyl group of acetyl-cysteine. The acetyl group comes from an acyl-CoA "donor". The enzyme catalyzes the first committed step in the formation of complex isoprenoids (like cholesterol) and ketone bodies. The product, 3-hydroxy-3-methylglutaryl HMG–CoA, can either be reduced by HMG-CoA reductase to form mevalonate, which leads to cholesterol synthesis, or cleaved by the enzyme HMG-CoA lyase, to produce acetoacetate, a ketone body.

Since it catalyzes the first step in cholesterol synthesis, the enzyme HMG-CoA reductase has been a primary focus of drug therapy to reduce high levels of serum cholesterol, which is generally associated with cardiovascular disease. Over 200 million people worldwide are on statins, the primary drugs used to block HMG-CoA reductase activity. Luckily, it doesn't affect HMG-CoA lyase. Gram-positive bacteria also require the mevalonate pathway.

HMG-CoA synthase catalyzes a bisubstrate reaction that displays ping-pong kinetics, characteristic of a covalent enzyme intermediate. The first substrate binds to the enzyme and transfers an acetyl group to a nucleophilic Cys 111 in the active site to form an acetyl-Cys 111 intermediae. Free CoA departs. Next the second substrate, acetoacetyl-CoA binds, and condenses with the acetyl group donated by acetyl-Cys 111. This condensation involves an enolate. A plausible reaction mechanism is shown in Figure 17.3.5.

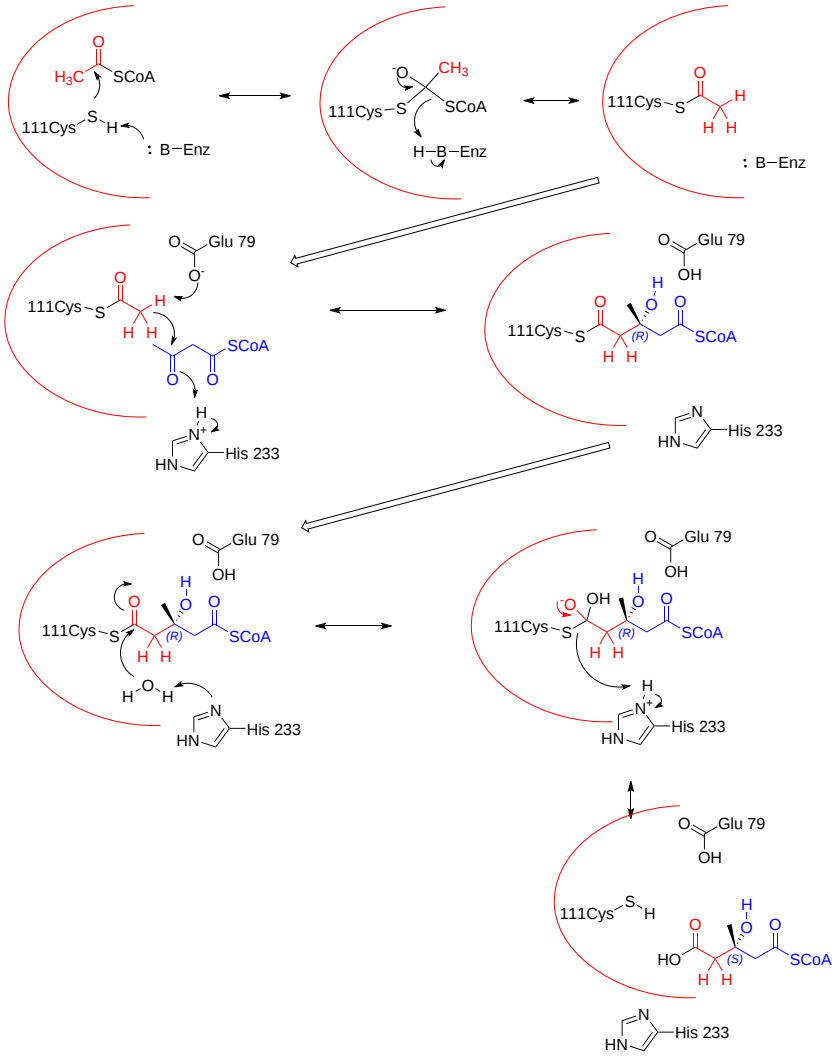
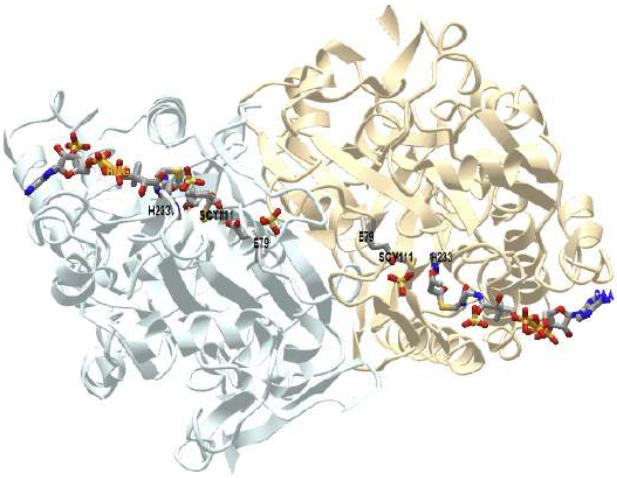


Figure 17.3.5: Reaction mechanism for HMG-CoA synthase

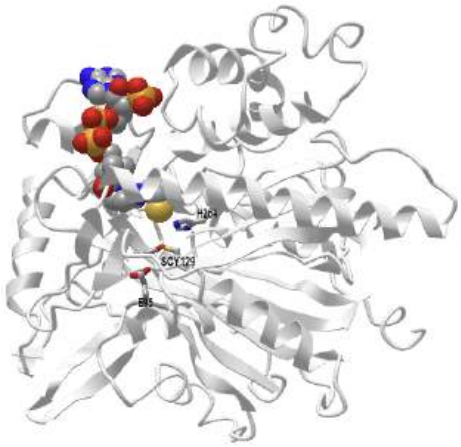
Hence there are three parts of the reaction: acetylation/deacetylation, condensation/cleavage with an enolate intermediate, and C-C formation and hydrolysis/dehydration. Figure 17.3.6 shows an interactive iCn3D model of the Staphylococcus aureus HMG-COA Synthase with bound HMG-CoA and acetoacetyl-CoA (1XPK)



NCBI iCn3D Figure 17.3.6: Staphylococcus aureus HMG-COA Synthase with bound HMG-CoA and acetoacetyl-CoA (1XPK). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...deiFs7JceE5H76

The biologically active unit (homodimer) is shown with each monomer shown in a different color. The A chain (light cyan) has bound HMG-CoA (HMG) while the B chain (light gold) has acetoacetyl-CoA (CAA) bound. The Glu 79, Cys 111, and His 233 in each subunit are shown in CPK sticks and labeled. Note that the Cys 111 is covalently modified in each subunit.

Figure 17.3.7 shows an interactive iCn3D model of the human 3-hydroxy-3-methylglutaryl CoA synthase I with bound CoASH (2P8U)



NCBI iCn3D Figure 17.3.7: Human 3-hydroxy-3-methylglutaryl CoA synthase I (monomer) with bound CoASH (2P8U). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...xxEoUfecTk3eL6

The active site side chains are numbered differently (Glu 95, Cys 129, and His 264) compared to the S. aureus protein. Only the monomer is shown in this model.

Reaction 3: HMG-CoA Lyase

The lyase reaction is a C-C bond cleavage to produce acetyl-CoA and acetoacetate. A gene knockout of this gene in mice is lethal. Figure 17.3.8 shows a plausible mechanism for the reaction.

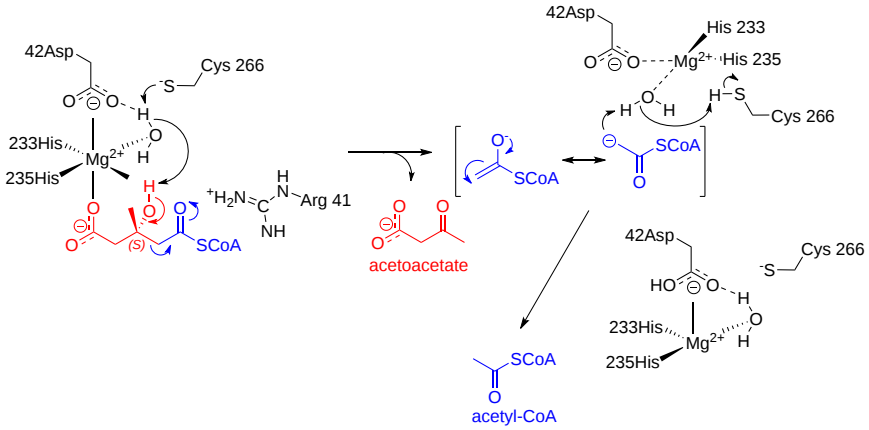
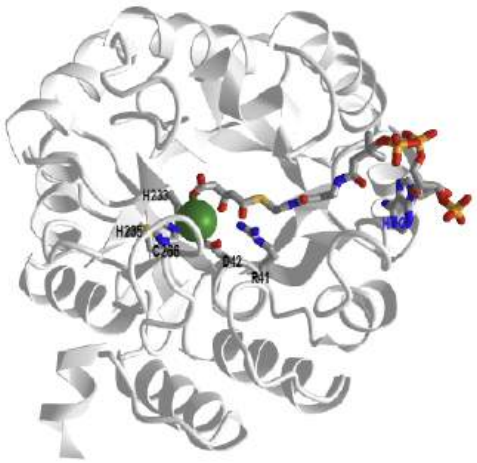


Figure 17.3.8: Mechanism for HMG-CoA lyase of HMG-CoA. (after Fu et al. The Journal of Biological Chemistry, 285, 26341–26349, August 20, 2010)

Figure 17.3.9 shows an interactive iCn3D model of the human HMG-CoA lyase with the competitive inhibitor hydroxyglutaryl-CoA (HG-CoA)(3MP3)



NCBI iCn3D Figure 17.3.9: Human HMG-CoA lyase with inhibitor HG-CoA (3MP3). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...VJBWxXpcKQyij8

Only one subunit of the homodimer is shown. The key residues involved in catalysis and binding (Arg 41, Asp 42, His 233, His 235, and Cys 266) are shown in sticks, CPK colors, and labeled. The competitive inhibitor, 3-hydroxyglutaryl-CoA (without the methyl group) is shown in CPK-colored sticks. The S-stereoisomer of HMG-CoA or inhibitor binds, not the R-isomer. Water may shuttle protons between Asp 42 and the C3 hydroxyl of HMG-CoA. Arg 41 facilitates the proper enol/keto tautomer of the product mutation on reaction product enolization. Mutation of Arg 41 have lhas effect on catalysis.

We won't go into detail on the remaining enzymes since we have encountered variants of NAD^+/NADH dehydrogenase reactions before. The acetoacetate decarboxylase just hastens the already spontaneous decarboxylation of the β-keto ester acetoacetate.

Regulation of Ketogenesis

Ketogenesis can be upregulated by hormones such as glucagon, cortisol, thyroid hormones, and catecholamines which lead to fatty acid mobilization. Insulin is the primary hormonal regulator of this process.

Insulin regulates many key enzymes in the ketogenic pathway, and a state of low insulin triggers the process. A low insulin state leads to:

- Increased free fatty acids (FFAs) arising from decreased inhibition of hormone-sensitive lipase
- Increased uptake of FFAs into the mitochondria arising from decreased activation of acetyl-CoA carboxylase, decreasing malonyl CoA, which disinhibits Carnitine Palmitoyltransferase 1 (CPT1)
- Increased production of ketone bodies arising from increased HMG-CoA activity

17.3.3: KETOLYSIS: UTILIZATION OF KETONE BODIES

Figure 17.3.10 shows the chemical structures in the utilization of ketone bodies.

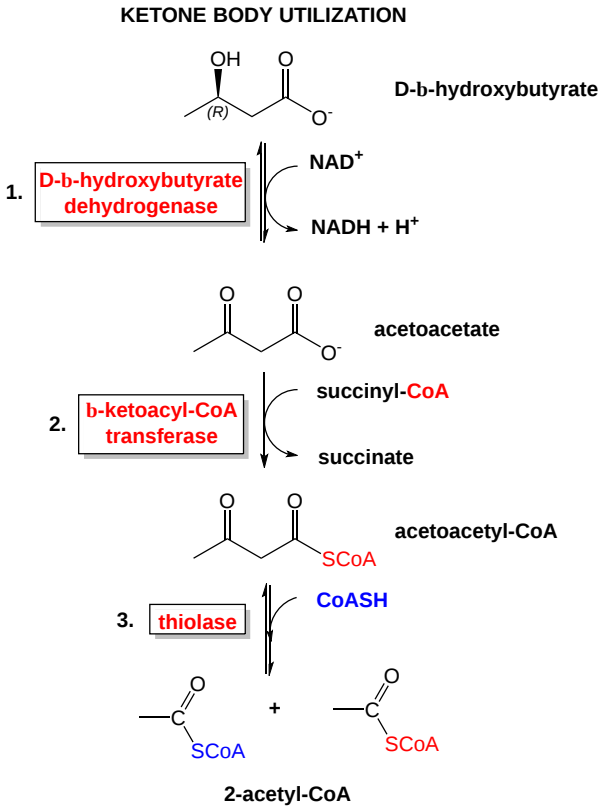


Figure 17.3.10 shows the chemical structures in the utilization of ketone bodies.

Since we have encountered reactions similar to these before, we will not discuss their detailed mechanisms.

Most organs and tissues can use ketone bodies for energy. They are a major source of energy for the brain when glucose is not available. That makes sense given that fatty acids can't cross the blood-brain barrier. The heart prefers fatty acids for energy but can also use ketone bodies. The liver makes but does not use them since the gene for step 2, beta ketoacyl-CoA transferase is not transcribed. D-β-hydroxybutyrate is first converted to the other ketone body, acetoacetate, for the pathway to continue. The net result is the formation of two acetyl-CoAs that can then enter the citric acid cycle for ATP production. Acetone is a dead-end ketone body and is excreted in the urine or eliminated in the breath.

17.3.4: CLINICAL SIGNIFICANCE

An overproduction of ketone bodies through increased ketogenesis can cause ketoacidosis (given the pKas of the ketone bodies). One type is diabetic ketoacidosis (DKA) in Type I and II diabetes which impair glucose use. This leads to liver gluconeogenesis and the resulting depletion of oxaloacetate and pyruvate, impacting the ability of acetyl-CoA to enter the citric acid cycle.

On presentation, patients are usually very dehydrated from being hyperglycemic. High glucose levels in the blood cause higher osmotic pressure in the vessels and decreased water reabsorption in the kidneys. They often have accompanying symptoms like confusion, nausea, vomiting, and abdominal pain. Because of the acidosis, patients often breathe very deeply and rapidly to eliminate carbon dioxide, which can cause and cause respiratory alkalosis. Ketoacidosis also can occur with severe alcoholism and prolonged starvation.

KETOGENIC DIET

The ketogenic diet, characterized by high fat and low carbohydrate consumption, can lead to weight loss. It shifts metabolism to more closely resemble the fasting state. Its long-term effects are not clear but some have suggested the use of ketone-based drugs to mimic the effect of ketone bodies. Fasting can also lead to a ketogenic state and it has been used for the treatment of epilepsy before the drug became available. It is important to differentiate this attempt to induce a ketogenic state from diabetic ketoacidosis.

CHAPTER OVERVIEW

18: NITROGEN - AMINO ACID CATABOLISM

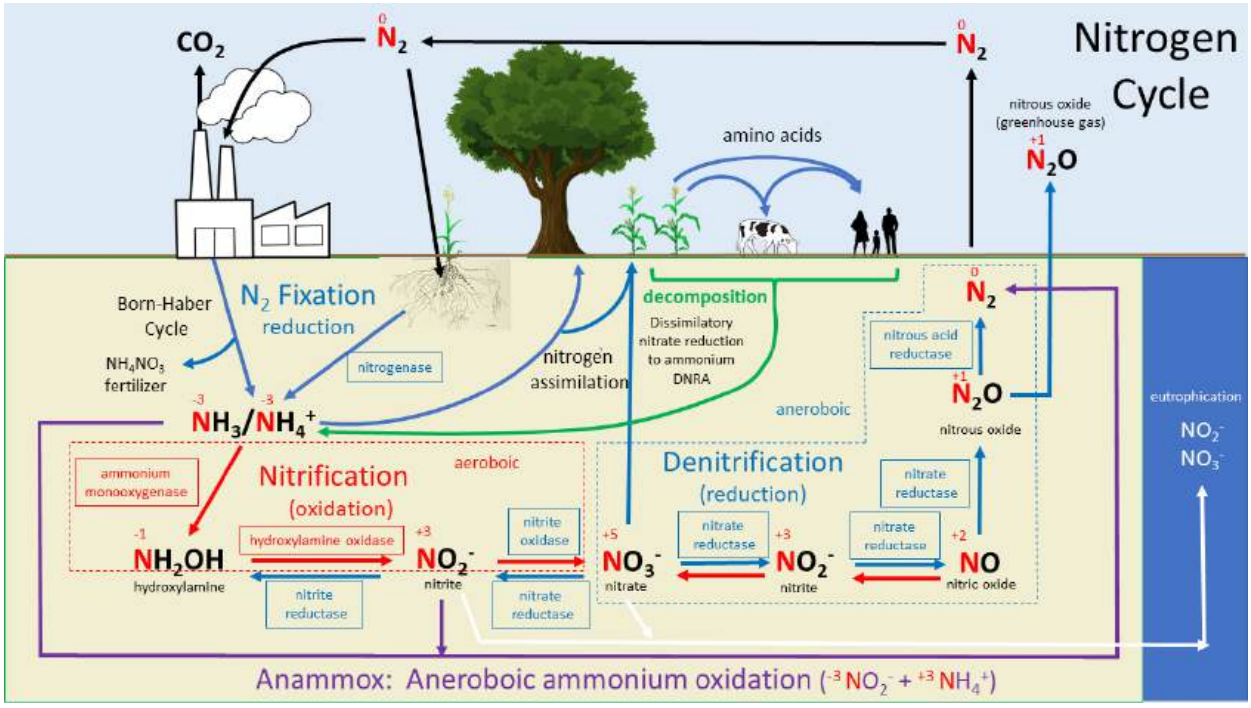
18.1: The Biochemistry of Nitrogen in the Biosphere
18.2: Metabolic Fates of Amino Groups
18.3: Nitrogen Excretion and the Urea Cycle
18.4: An overview of amino acid metabolism and the role of Cofactors
18.5: Pathways of Amino Acid Degradation

18.1: THE BIOCHEMISTRY OF NITROGEN IN THE BIOSPHERE

18.1.1: INTRODUCTION

Organic chemistry is usually described as the chemistry of carbon-containing molecules. But isn't that definition a bit carbon-centric, especially since the prevalence of oxygen-containing molecules is staggering? What about nitrogen? We live in a dinitrogen-rich atmosphere (80%), and all classes of biomolecules (lipids, carbohydrates, nucleic acids, and proteins) contain nitrogen. Dinitrogen is very stable, given its triple bond and its nonpolar nature. We rely on a few organisms to fix N_2 from the atmosphere to form ammonium (NH_4^+), which through nitrification and denitrification can form nitrite (NO_2^-), nitrate (NO_2^-), nitric oxide (NO), and nitrous oxide (N_2O), the latter being a potent greenhouse gas. We'll concentrate on the metabolic fate of amino groups in amino acids and proteins in the next section. Before exploring their fates, look at the figure below which shows an overall view of the biological nitrogen cycle. The study of biochemistry should encompass more than homo sapiens and expand to the ecosystem in which we are such a small but damaging part.

Let's break down the diagram from a biochemical perspective. There are aerobic and anaerobic processes (conducted by bacteria). Nitrogen-containing substances include both inorganic (ammonium, nitrate, nitrite) and organic (amino acids, nucleotides, etc) molecules. The reactions shown are oxidative and reductive (note: the oxidation number of the nitrogen atoms in the molecules is shown in red). Most of the reactions are carried out underground by bacterial and Archaeal microorganisms.



Here are some of the major reactions:

- N_2 fixation (a reduction): N_2 from the air is converted by bacteria to ammonium (NH_4^+) by the enzyme nitrogenase of soil prokaryotes. The energetically disfavored reaction requires lots of ATPs. Ammonium once made can then be taken up by primary producers like plants and incorporated into biomolecules such as amino acids, which animals consume. For those who may still believe that people have marginal effects on our biosphere, consider this. We may soon fix more N_2 (to NH_3) through the industrial Born-Haber reaction (used for fertilizer and explosive productions) that all the N_2 fixed by the biosphere. Much of the nitrogen in use comes from the Born-Haber reaction. The excess NH_4^+ (upwards of 50%) produced industrially and which enters the soil in fertilizers (mostly as NH_4NO_3) has overwhelmed nature's ability to balance the nitrogen cycle and is not taken up by plants. It is metabolized by microorganisms to nitrite and nitrate.
- Nitrification: Ammonium is converted to nitrite by ammonia-oxidizing aerobic microorganisms and further to nitrate by a separate group of nitrite-oxidizing aerobic bacteria. Here are the reactions (Rx 1 and 2) to produce nitrate through a hydroxylamine intermediate, followed by the formation of nitrate (Rx 3).

$$\mathrm{NH_3 + O_2 + 2\,e^- \longrightarrow NH_2OH + H_2O} \qquad \text{(Rx 1)}$$

$$\mathrm{NH_2OH + H_2O \longrightarrow NO_2^- + 5\,H^+ + 4\,e^-} \qquad \text{(Rx 2)}$$

$$\mathrm{NO_2^- + \tfrac{1}{2}O_2 \longrightarrow NO_3^-} \qquad \text{(Rx 3)}$$

These added ions exceed soil capacity and end up runoff water, polluting our rivers and lakes.

- **Denitrification**: This anaerobic reaction pathway reproduces N_2 from nitrate Here is the net reaction:

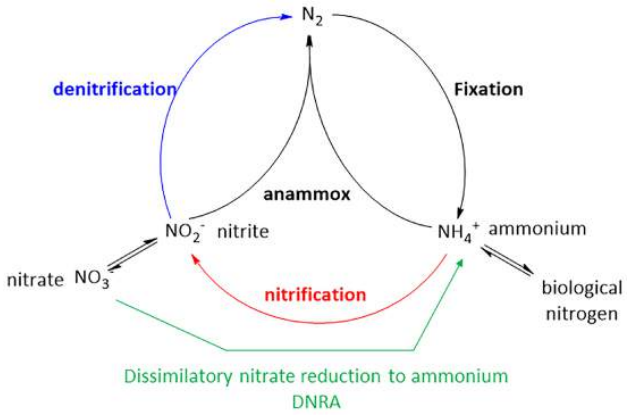
$$2\,NO_3^- + 10\,e^- + 12\,H^+ \longrightarrow 2\,N_2 + 6\,H_2O$$

- **Anammox reaction**: This more recently discovered bacterial anaerobic reaction pathway converts ammonium and nitrate to N_2. Here is the net reaction

$$NO_2^- + NH_3^+ \longrightarrow N_2 + 2\,H_2O$$

- **Ammonification** (not to be confused with mummification) occurs when plants and animals decompose, which returns ammonium to the soil for reuse by plants and microbes.

These reactions are shown in the abbreviated Nitrogen Cycle below.



Nitrogen metabolites are nutrients for plants and perhaps the most important nutrients in the regulation of plant growth (primary productivity) and in regulating life diversity in the biosphere. All living organisms require feedstocks to produce energy and as substrates for biosynthetic reactions. Which ones are used depends on the organism. Plants are primary producers so they use their own synthesized carbohydrates for both energy production and biosynthesis. For carnivores, proteins and their derived amino acids are the source of energy (through oxidation) and serve as biosynthetic precursors. For omnivorous organisms, the source of energy depends on the "fed" state. With abundant food resources, carbohydrates, and lipids are the source of energy. Unlike carbohydrates and lipids, which can be stored as glycogen and triacylglycerols for future use, excess protein, and their associated amino acids can not be stored, so amino acids can be eliminated or used for oxidative energy.

In the fed state, carbohydrates are the main source, while in the unfed state, lipids take a predominant role. Under starving conditions, the organisms' own proteins are broken down and used for oxidative energy production and for any biosynthesis that remains. In diseased states like diabetes, which can be likened to a starving state in the presence of abundant carbohydrates, both lipids and amino acids become the sources of energy.

How are amino acids in animals oxidatively metabolized? Many pathways could be used to do so but it would seem logical that NH_4^+ would be removed and the carbons in the remaining molecule would eventually enter glycolysis or the TCA cycle in the form of ketoacids. NH_4^+ is toxic in high concentrations. Ammonium is not oxidized to nitrite or nitrates in humans as occurs in the soil by microorganisms. It can be recycled back into nucleotides or amino acids, and excess amounts are eliminated from the organism. Both processes must be highly controlled. We will turn out attention to the oxidation of amino acids in the next section.

18.2: METABOLIC FATES OF AMINO GROUPS

From the previous section on the nitrogen cycle, it should be clear that NH_3/NH_4^+ has a central role in metabolism. Nature's fertilizer for plants is ammonium derived from bacterial nitrogenase and human-derived Born-Haber process. Animals get their ammonia mostly from ingested plants (primary producers) and animal protein. In vertebrates, proteins get digested into small proteins and peptides starting in the stomach where pepsin cleaves proteins after aromatic and Leu side chains into smaller fragments. The low pH in the stomach (1-2) facilitates protein unfolding, which allows great access to buried side chains for pepsin cleavage. As the ingested materials move into the small intestine, the protein fragments are further cleaved into small peptides by trypsin, and chymotrypsin and then into individual amino acids by carboxypeptidases (which cleaves C-terminal amino acids) and aminopeptidases (which cleave at the N-terminal end). The individual amino acids are then adsorbed, along with di- and tripeptides into the epithelial cells lining the small intestine. Peptidases cleave short peptides and amino acids are moved into the blood through cell membrane transporters which are coupled to sodium ion intake. Amino acids are taken up by tissues for use in protein synthesis and by the liver of vertebrates for metabolic processing. The gut proteolytic enzymes are secreted into the gut lumen as precursors (proenzymes), where they are activated either autocatalytically or by other proteases into their mature form.

Amino acids are also derived from the degradation of cellular proteins by a supramolecular assembly called the proteasome. Proteins designated for cleavage are covalently modified by ubiquitin, a short ubiquitous protein, which targets them to the proteasome.

18.2.1: DEGRADATION OF AMINO ACIDS TO PRODUCE AMMONIUM

Here is some good advice on giving a seminar: tell what you will tell them, tell them and tell them what you told them. These words imply that repetition is one of the keys to learning. So look at Figure 18.2.1 which shows an overall view of the glycolytic and TCA pathways and see how amino acid catabolism fits into what you already know!

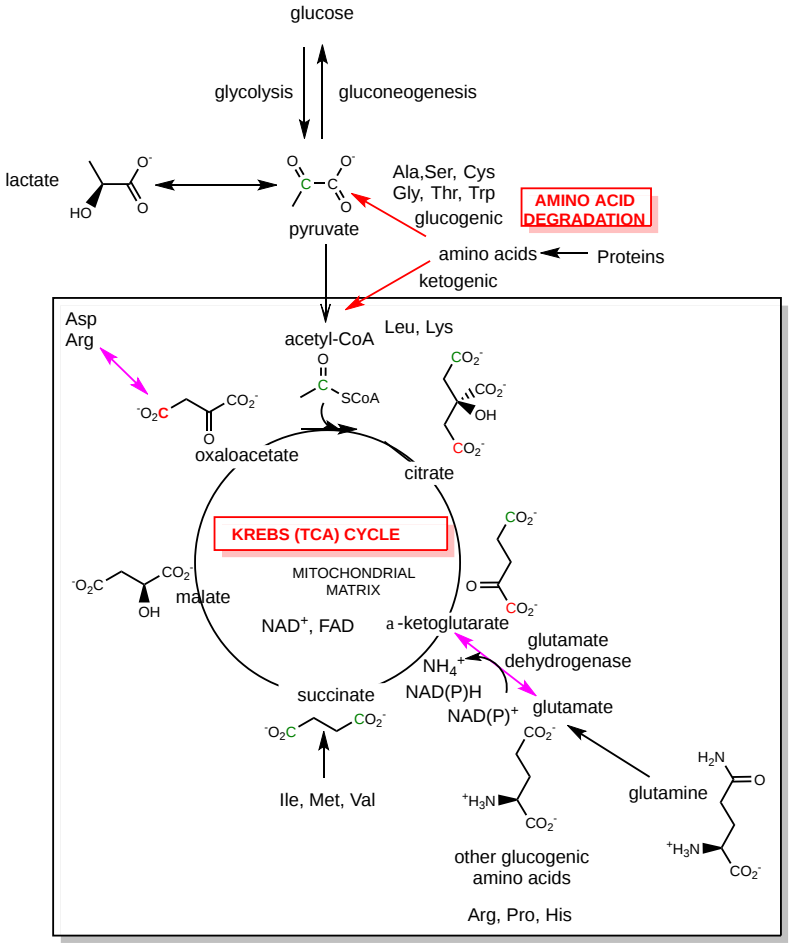


Figure 18.2.1: Amino acids and their involvement with the glycolytic and TCA pathways

A summarizing principle is that amino acids degrade to form intermediates that fit into these two common pathways and can be used for ATP production as well as biosynthesis. Those that form pyruvate are glucogenic amino acids, while those that form acetyl-CoA are ketogenic (form ketones and could form ketone bodies as well). Many form direct intermediates in the TCA cycle.

This chapter will focus on glutamine/glutamic acid, alanine (which forms pyruvate), and aspartic acid. In this section, we will focus on the role of amino acid catabolism in producing ammonium. When new students confront the structures, properties, and reactions of amino acids, they are undoubtedly daunted. Even biochemists would be if amino acid metabolism is outside their active research area. Perhaps some interesting biological properties of **free** amino acids might pique your interest. For example, did you know that ...

- glutamate and glutamine are the most abundant amino acids in red blood cells;
- alanine and glutamine are the most abundant amino acid in the body;
- glutamic acid, glutamine, and alanine are the most abundant amino acids in the cerebrospinal fluid CSF (50-55% of total amino acids);
- glutamine is the most abundant amino acid in blood serum;
- In contrast, leucine is the most abundant amino acid in **proteins**, and free leucine is a prime regulator of protein synthesis through its interaction with the mTOR kinase complex. The other top three amino acids in proteins are serine, lysine, and glutamic acid.

From the above list, it appears that Glu and Gln play special roles in metabolism and signaling. They also play major roles in ammonium metabolism as the pair are major sources of NH_4^+ production in cells. Gln has two Ns (an amine and an amide), so it's not an unexpected source of nitrogen:

- Glutamine can give up its amide nitrogen to form NH_4^+ on conversion to glutamic acid, a reaction catalyzed by **glutaminase**
- Glutamic acid can undergo oxidative deamination to form α-ketoglutarate (a TCA intermediate) and free NH_4^+, a reaction catalyzed by **glutamate dehydrogenase**
- Glutamic acid can also give up its ammonia nitrogen to a ketoacid like pyruvate to form α-ketoglutarate (a TCA intermediate) and another amino acid (alanine if pyruvate was used as the keto acid), a reaction catalyzed by a class of enzymes called **transaminases** (makes sense again), which are also called **aminotransferases**. In this case, free NH_4^+ is not formed but rather is passed to a keto acid to form another amino acid.

In vertebrates, free amino acids are metabolized in the liver. Amino acids that enter the liver transfer their ammonia group to α-ketoglutarate (aka 2-oxoglutarate) to form glutamic acid which enters the mitochondria and can be cleaved by glutamate dehydrogenase to form α-ketoglutarate and free ammonium. The NH_4^+ produced is either recycled or excreted in the form of NH_4^+ in fish, urea, $H_2N(C{=}O)NH_2$ in vertebrates, and uric acid in birds and reptiles. Excess amino acids (which again can't be stored for energy) in other organs pass their ammonia group to glutamic acid to form glutamine, which then heads to the liver for processing. Glutamine then becomes a safe way to transfer 1-2 ammonium equivalents through the blood.

Alanine and aspartic acid also play secondary roles. Ala is one transamination step away, through the removal of its amine group, from pyruvate, a crucial α-ketoacid end product of glycolysis and entry product for the TCA cycle (after pyruvate dehydrogenase). In muscle, excess amino acids pass their ammonia groups to pyruvate to form alanine, another "safe" carrier of ammonium, which heads to the liver for processing. Likewise, aspartic acid is one transamination step way, through the removal of its amino group, to form oxaloacetate, another TCA α-ketoacid.

Figure 18.2.2 shows the key α-keto acids which are also intermediates of the TCA cycle or feed into it (pyruvate) and their respective transamination amino acid products. Their structures immediately show a link between amino acid and carbohydrate metabolism.

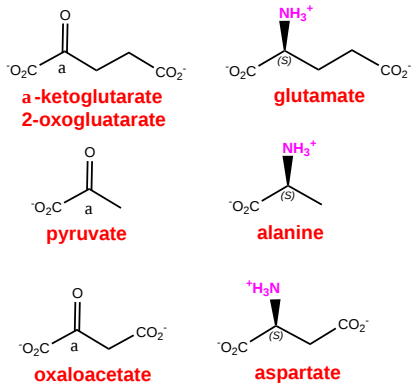


Figure 18.2.2: α-ketoacids and their transamination amino acid products

We will focus most of our attention on glutamine and glutamic acid. Figure 18.2.3 summarizes the important enzymatic steps in the conversion of Gln ↔ Glu ↔ α-KG

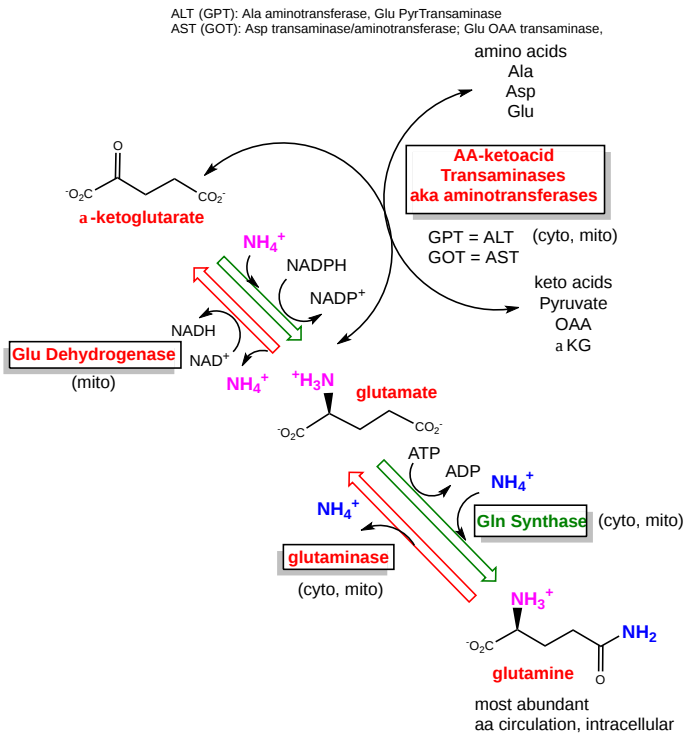


Figure 18.2.3: Important enzymatic steps in the conversion of Glutamine ↔ Glutamic acid ↔ α-ketoglutarate

18.2.2: PYRIDOXAL PHOSPHATE AND TRANSAMINATION REACTIONS

Most free amino acids start their metabolic degradation in the liver with transamination reactions, using PLP as a cofactor. PLP enables other biochemical reactions of amino acids, including racemizations, decarboxylation, and dehydration (of serine). PLP is covalently attached to a Lys through a Schiff base linkage in the enzyme during the reaction cycle so it is not considered a substrate, but rather a cofactor that returns to its original state after the reaction.

Figure 18.2.4 shows the structure of PLP and the imine formed on reaction with an amino acid. The reaction is readily reversible through hydrolysis in the presence of water which would presumably not be present in the active site of PLP-dependent enzymes. A lysine ammonia group in the enzyme forms a covalent adduct with PLP to form an imine. The imine in the figure below is called an **internal aldimine**. Internal implies that the source of the N in the imine link is a Lys internal to the protein. If in your mind you would replace the N in the imine with an O, the functional group with the C=O would be an aldehyde. Hence the imine shown is an aldimine.

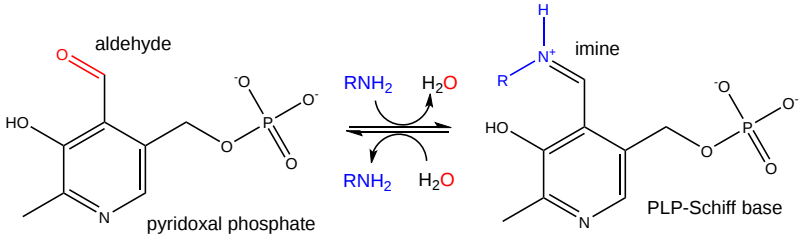


Figure 18.2.4: PLP and its imine-protein conjugate

William Jencks, in his classic text, Catalysis in Chemistry, describes the mechanistic beauty of PLP-dependent enzymes:

> *"It has been said that God created an organism especially adapted to help the biologist find an answer to every question about the physiology of living systems; if this is so, it must be concluded that pyridoxal phosphate was created to provide satisfaction and enlightenment to those enzymologists and chemists who enjoy pushing electrons, for no other coenzyme is involved in such a wide variety of reactions, in both enzyme and model systems, which can be reasonably interpreted in terms of the chemical properties of the coenzyme. Most of these reactions are made possible by a common structural feature. That is, electron withdrawal toward the cationic nitrogen atom of the*

imine and into the electron sink of the pyridoxal ring from the alpha carbon atom of the attached amino acid activates all three of the substituents of this carbon for reactions which require electron withdrawal from this atom."

When PLP-dependent enzymes react with an **external** amino acid acting as a substrate, the amine of the incoming amino acid replaces the amine from the enzyme's active Lys to produce an **external** imine. The figure below shows an external amino acid in imine linkage to PLP, and in the form shown is an aldimine.

Each PLP-dependent reaction involves a protonated and positively charged pyridinium N as an electron sink, which facilitates cleavage of **each bond** around the C_α of the covalently attached amino acid. Bond cleavage would leave a lone pair and negative charge on the α-carbon and no place to stabilize it by resonance. On the formation of a Schiff base with a pKa of around 7.0, the imine nitrogen, especially in its protonated form, is an excellent stabilizer of the negative charge. Continued delocalization of the lone pair and negative charge to the positive pyridinium nitrogen further facilitates the stabilization and hence the reaction. Figure 18.2.5 shows the cleavage of bonds around the C_α of an amino acid-PLP Schiff base.

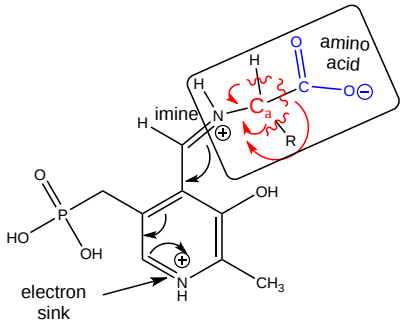


Figure 18.2.5: Cleavage of bonds around the C_α of an amino acid-PLP Schiff base

Figure 18.2.6 shows the mechanism for a transamination reaction using PLP as a cofactor. This is of course the reaction most relevant to this chapter section. The amino acid substrate is first shown in an aldimine linkage to PLP. Also, note the conversion of an aldimine to a ketimine in the next step.

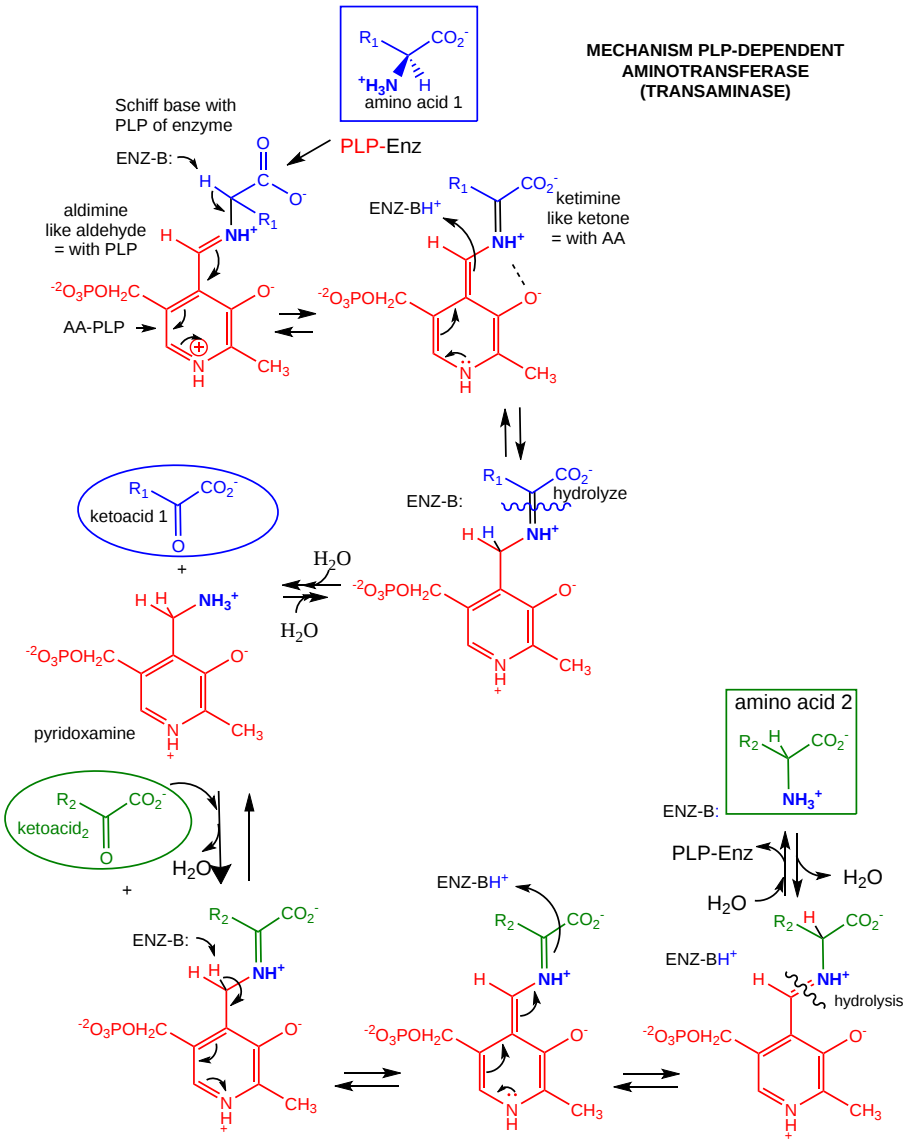
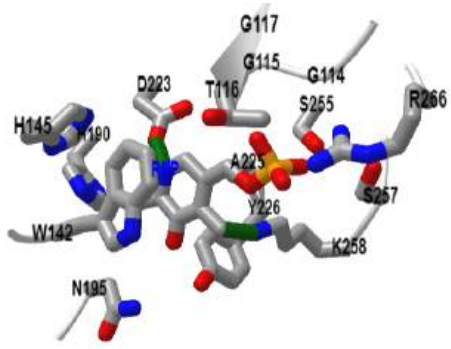


Figure 18.2.6: Mechanism for a transamination reaction using PLP as a cofactor

The model below shows the active site of aspartate transaminase from E. Coli (pdb code 1aam) with PLP in Schiff base linkage with lysine 258. Three amino acids critical for enzyme function (Trp 142, Asp 223, and Lys 258) and all side chains within 5 angstroms from the active site defined by those amino acids, are shown. Lys 258 forms the Schiff base with PLP cofactor and also acts as a general acid/base on interconversion between the aldimine and ketamine forms (see figure above). The ring of Trp 142 forms pi-stacking interactions with the aromatic ring of PLP and helps position it. Asp 223 acts as a general acid/base in the catalytic cycle, facilitating the deprotonation of the amine group of the substrate Asp, which makes it a potent nucleophile in the formation of the external aldimine.

Figure 18.2.7 shows an interactive iCn3D model of the active site of aspartate transaminase from E. Coli with PLP (1aam)



NCBI iCn3D Figure 18.2.7: Active site of aspartate transaminase from E. Coli with PLP (1aam). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...YeQ6kah4EuERT9

The various names and abbreviations for transaminases/aminotransferases (which are reversible) can be confusing. Two key ones whose levels are used to assess liver toxicity in clinical tests are AST (GOT) and ALT (GPT).

- **AS**partate Amino **T**ransferase (**AST)** is the same enzyme as **G**lutamate **O**xalacetate **T**ransaminase (**GOT**), named for the reverse reaction:

$$\text{aspartate} + \alpha\text{−ketoglutarate} \leftrightarrow \text{oxalacetate} + \text{glutamate}$$

- **AL**anine Amino **T**ransferase (**ALT**) is the same enzyme as **G**lutamate **P**yruvate **T**ransaminase (**GPT**), named for the reverse reaction:

$$\text{alanine} + \alpha\text{−ketoglutarate} \leftrightarrow \text{pyruvate} + \text{glutamate}$$

We will consider other reactions of PLP-dependent enzymes when needed.

To summarize, different amino acids coming into the liver can donate their amine groups to α-ketoglutarate to form glutamic acid in a reaction catalyzed by a transaminase/aminotransferase to form glutamic acid.

18.2.3: PRODUCTION OF NH_3 - GLUTAMATE DEHYDROGENASE

Now that we made glutamic acid from various amino acids, we can now break it down to make α-ketoglutarate again and ammonium. The overall reversible reaction catalyzed by the enzyme glutamate dehydrogenase in mitochondria is shown in Figure 18.2.8. Mammalian livers can also use $NADP^+$ as an oxidizing agent.

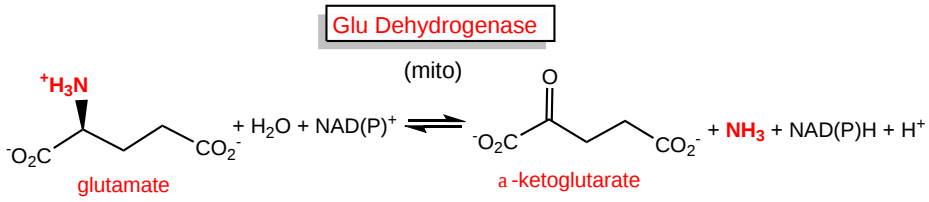


Figure 18.2.8: Overall reversible reaction catalyzed by the enzyme glutamate dehydrogenase in mitochondria

Humans have two isozymes. GluDH1 is expressed at high levels in the liver, brain, pancreas, and kidney, while GluDH2 is found in the retina, testes, and brain. GluDH1 can use both NAD^+ and $NADP^+$ so it can be used in both anabolic and catabolic reactions.

The reaction/product pairs again show this reaction to be a clear link between protein and carbohydrate metabolism. Figure 18.2.9 shows just the first step in the reaction of glutamate dehydrogenase.

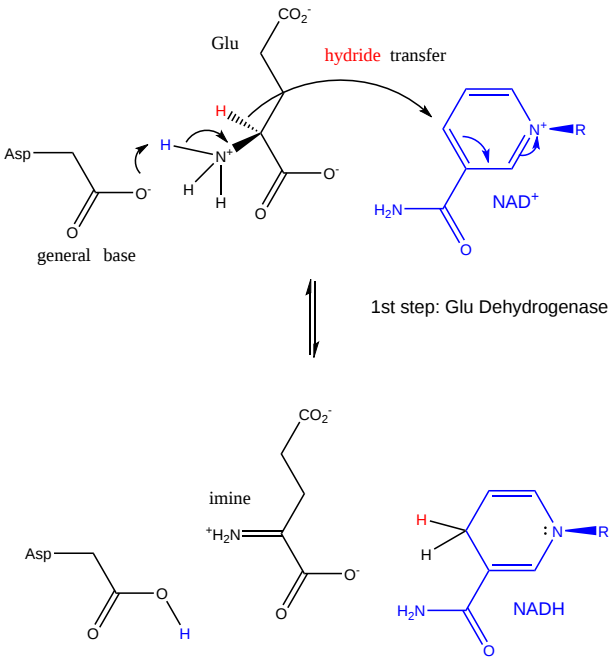


Figure 18.2.9: First step in the reaction of glutamate dehydrogenase

The C-N in Glu becomes C=N in the imine intermediate. You should recognize this as an oxidation step even if the obvious oxidizing agent, $NAD(P)^+$, was not shown. That should also be evident since a hydride (H^-) with a lone pair is removed from glutamic acid and not a H^+ as in an acid/base reaction. Subsequent steps include nucleophilic attack by water (enhanced by a proximal Lys acting as a general base) on the imine C followed by ammonia release on the formation of alpha-ketoglutarate. The overall reaction is an **oxidative deamination**.

NH_3/NH_4^+ are toxic in high concentrations. One possible reason is that the NH_4^+ cation might compete with the transport of other ions across neural membranes, altering transmembrane potentials and hence neural function. Given this toxicity, you expect that glutamate dehydrogenase is highly regulated and it is in mammals. The mammalian enzyme is a hexamer of identical subunits which suggests allosteric regulation of activity (much like tetrameric hemoglobin). ADP and leucine are allosteric activators while GTP, palmitoyl CoA, and ATP act as inhibitors. The GTP/ADP regulation is consistent with the notion that the enzyme is regulated by the need for cellular energy (remember that GTP is formed in the TCA cycle).

Although this reaction is written above in the direction of NH_3 formation, the reaction is reversible, even though the ΔG^0=-6.2 kcal/mol (-26 kJ/mol) and though the human enzymes have a high Km for ammonia (12-62 mM, Brenda Database). As the pH decreases from 8 to 7, the Km for ammonia increases from around 12 to 60 mM. This implies that on glutamate uptake and in acidotic conditions, the reaction runs exclusively in the direction of NH_3 release - as an oxidative deamination reaction.

18.2.4: TRANSPORTING AMMONIUM EQUIVALENTS IN THE BLOOD - GLUTAMINE AND ALANINE

NH_3/NH_4^+ are toxic in high concentration, mechanisms must be in place to transport it in the blood. Ideally, it would be transported in a nontoxic form. Glutamic acid is a key metabolic molecule so glutamine is the preferred molecule for safe "NH_3/NH_4^+" transport from tissue to the liver, while alanine is the choice for muscle tissue. In a way, this less toxic mode of transport is similar in principle to the use of soluble ketone bodies to transport the energy equivalent of the less soluble fatty acids in the blood.

A key enzyme for this process is **glutamine synthetase**, which catalyzes this reaction:

$$NH_4^+ + ATP + Glu \rightarrow Gln + ADP + HPO_4^{2-} + H^+$$

This reaction is important as most ingested glutamine is converted to glutamate on absorption.

NH_3/NH_4^+ made in the intestine and kidney ends up in the circulation where it heads to the liver for the eventual production of urea.

Figure 18.2.10 shows the Glucose-Alanine Cycle in comparison to the Cori cycle we discussed in a previous chapter.

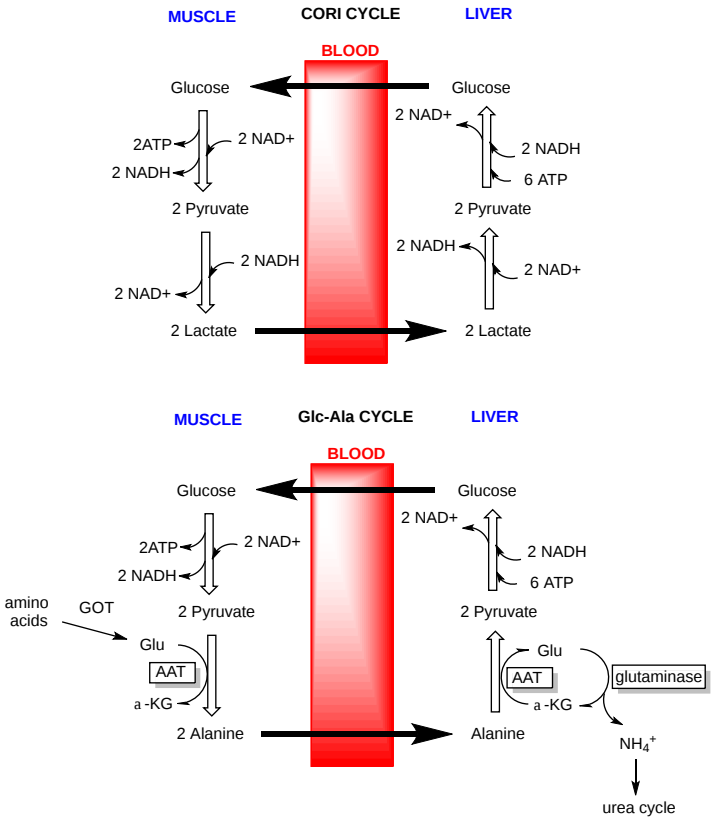


Figure 18.2.10: Glucose-Alanine Cycle in comparison to the Cori cycle

18.2.5: PRODUCTION OF NH_3 - GLUTAMINASE

Most amino acids arriving in the liver for degradation go through two enzymes, transaminases to form glutamic acid followed by glutamate dehydrogenase to form a TCA intermediate (alpha-ketoglutarate) and ammonia. Glutamine is a carrier of excess ammonia in the bloodstream (see below). When it arrives in the liver, it can lose its amide NH_2 as ammonium as it is converted to glutamic acid by the enzyme **glutaminase**, a nonoxidative enzyme that is expressed in the liver, brain, and kidney. Some of the glutamate can lose ammonia by the enzyme glutamate dehydrogenase, but most is reserved for protein synthesis or the creation of anabolic precursors.

As mentioned above, glutamine is the most abundant free amino acid in the body and most cells. It can donate its amide N through reactions catalyzed by **amido**transferase (not to be confused with aminotransferases) leading to the incorporation of nitrogen into many biomolecules.

18.2.6: METABOLIC SUMMARY

- Glutamine is both a carrier of ammonia and a carbon backbone used in metabolism when converted to α-ketoglutarate. It can be metabolized for energy and used in biosynthesis for nucleotides and neurotransmitters. As such its levels are controlled by glutamine synthase and glutaminase. When we get to synthetic pathways, you will find that glutamine is a NH_3/NH_4^+ donor for the synthesis of other amino acids as well as nucleotides, amino sugars, and NAD^+.
- Glutamine that is absorbed in the intestines is mostly converted to glutamic acid in the intestinal epithelial cells and is transported by the blood. The rest heads to the liver for processing. Hence most glutamine must be synthesized by glutamine synthase, a cytosolic enzyme found in most mammalian cells. However, it is most abundant in muscle, liver, and adipose cells which express little glutaminase and from which it can be exported. Glutaminase in the liver and kidney leads to ammonium production
- Glutamine is used as a source of energy and carbon for biosynthesis by tumor cells, which need both energy and intermediates for synthesis by rapidly proliferating cells. Aerobic glycolysis ("Warburg effect") allows that to occur.

18.3: NITROGEN EXCRETION AND THE UREA CYCLE

18.3.1: THE UREA CYCLE

Now we are in a position to see what happens to excess NH_3/NH_4^+ that accumulates in the liver mitochondria. The ammonia that is formed in the liver through oxidative deamination by glutamate dehydrogenase, ends up in urea. Alternatively, it could be added to glutamic acid to form glutamine which can head to the kidney where, after sequential reactions by glutaminase (a deamidation reaction) and glutamate dehydrogenase, it forms ammonia for direct excretion in the urine. If your diet is high in proteins and hence amino acids, since they can't be stored as proteins per se, they are metabolized for energy and the metabolic products can be converted to fat or carbohydrate. That leaves excess nitrogen which ends up mostly as urea for excretion.

Urea is a very water soluble, nontoxic molecule that biochemists use in the lab to chemically denature protein (at concentrations of 3-6 **M**). Clinical blood urea concentrations are expressed as Blood Urea **Nitrogen** (BUN) which range from 7 to 20 mg/dL N, which is equivalent to 2.5 to 7.1 **mM** urea. Serum levels depend predominately on the balance between urea synthesis in the liver and its elimination by the kidney. Normal cellular urea concentrations should be similar.

Urea is produced by a pathway called the urea cycle, as shown in Figure 18.3.1.

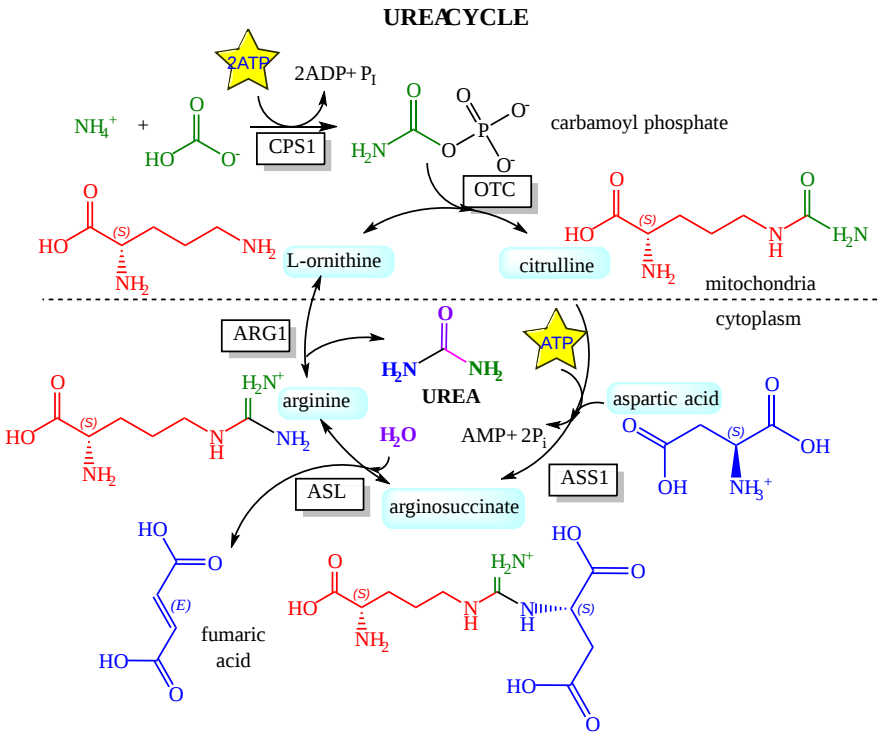


Figure 18.3.1: The Urea Cycle

The enzyme names are:

- CPS1: carbamoyl phosphate synthase I
- OTC: ornithine transcarbamoylase
- ASS1: arginosuccinate synthase 1
- ASL: arginosuccinase lysase (aka arginosuccinase)
- ARG1: arginase

It is color-coded to designate the sources of the C and N atoms. One N comes from ammonia (or indirectly from the amido N of glutamine), while the C atom comes from carbonate. The other comes from the amine of aspartate. Note that 3 ATPs are used to power the cycle. Also note the guanidino group of arginine ($NH(NH_2^+)NH_2$) is the immediate donor of the 2 N atoms in the step that produces urea. Knowing that one amino acid was the immediate donor of the N atoms in urea, you would easily predict that.

Now, let's look at some of the steps, starting with the ones that are powered by ATP, which are also the ones where nitrogen is incorporated into the precursors of urea.

18.3.2: CARBAMOYL PHOSPHATE SYNTHASE I (CPS I)

This enzyme catalyzes the first "committed" path of the pathway and makes the reactant that enters the cycle, carbamoyl phosphate. It requires a specific cofactor, N-acetyl-glutamate (NAG), produced by the acetylation of glutamate by acetyl-CoA using an enzyme NAG synthase which is activated by arginine. The reaction is not part of the cycle but provides the input for it.

A cytosolic version of this enzyme, CPS II, is used to synthesize arginine and pyrimidine nucleotides using glutamine as a donor of NH_4^+which reacts with carboxyphosphate to produce carbamoyl phosphate. It is not regulated by NAG.

CPS I provides a way to form a $(C=O)NH_2$ unit for urea by condensing bicarbonate and ammonia to form a carbamate, which contains a high energy "motif" with respect to its hydrolysis product. (Note: this does not imply that the broken bond is high energy, a misnomer found in many books.) Hence the reaction is powered by 2 ATPs as an energy source. Urea is a molecule that "carries" both NH_3/NH_4^+ and also carbonate.

A reaction mechanism for carbamoyl phosphate synthase is shown in Figure 18.3.2.

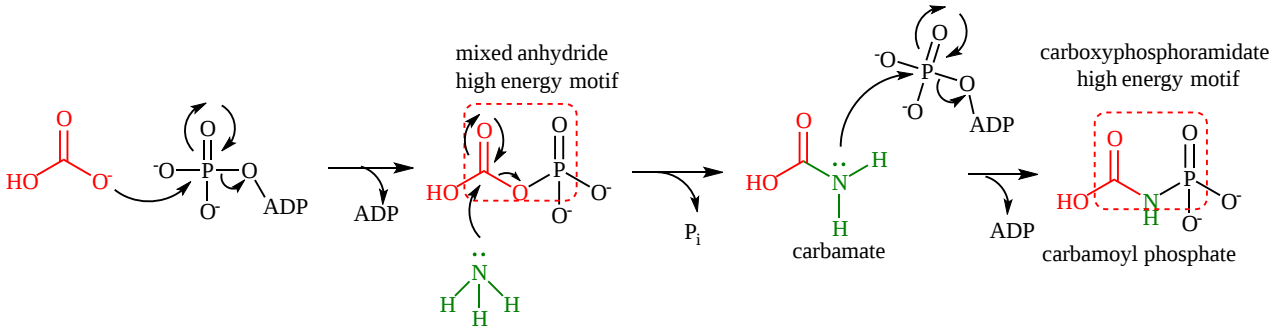
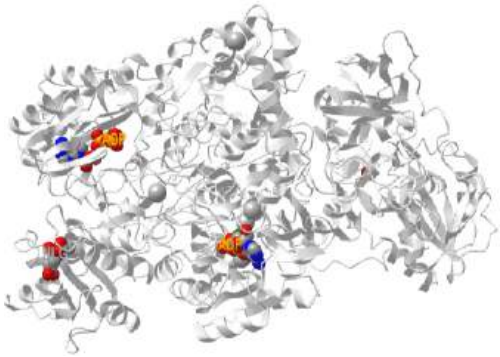


Figure 18.3.2: Reaction mechanism for carbamoyl phosphate synthase

The two high-energy motif molecules (again with respect to their hydrolysis products) are protected from nonspecific hydrolysis as the mixed anhydride pass along a sequestered tunnel to a more distal phosphorylation site in the multimeric enzyme.

Figure 18.3.3 shows an interactive iCn3D model of the human carbamoyl phosphate synthetase I with bound ADP and N-acetyl-glutamate (5DOU).



NCBI iCn3D Figure 18.3.3: Human carbamoyl phosphate synthetase I with bound ADP and N-acetyl-glutamate (5DOU). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...yV8fW3mW9idy6A

The model shows one chain of the multimer. The 2 ADPs and single N-acetylglutamate (NLG) are shown in spacefill and labeled. Additional bound ions are shown (unlabeled).

The protein exists in a monomer-dimer equilibrium mixture. Each monomer has 1 NAG and 2 ADP, which represent the two different binding sites for the two different phosphorylation steps as shown in the reaction above. The two phosphorylation sites are connected by a tunnel allowing passage of the mixed anhydride to the carbamate phosphorylation site. The NAG is an allosteric modifier as it is not bound near the phosphorylation sites. On binding to the apo form of the enzyme, it elicits a large conformation change allowing a competent enzyme conformation with a complete tunnel to predominate.

An image of the tunnel for one monomer of the human carbamoyl phosphate synthetase I is shown in Figure 18.3.4.

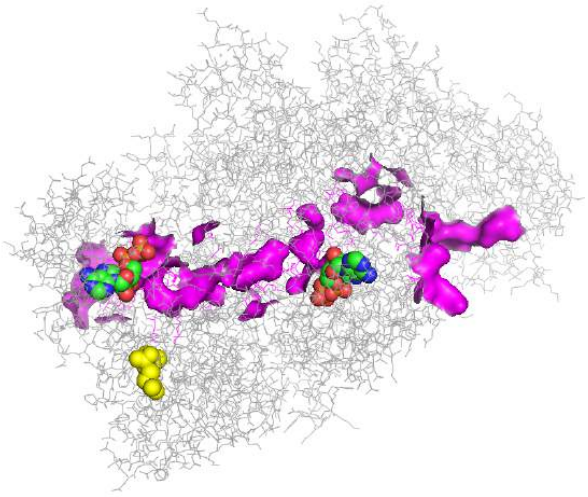


Figure 18.3.4: Tunnel for one monomer of the human carbamoyl phosphate synthetase I enzyme.

Two ADPs are shown below in spacefill with CPK color. The tunnel between the two ADP sites is highlighted in magenta. The yellow molecule is NAG, which again exerts its activation effect by binding to an allosteric site, which effectively opens up the tunnel.

Figure 18.3.5 shows the enzyme monomer in a different orientation with a tunnel made using ChExVis: a tool for molecular channel extraction and visualization. The large red and green spheres show the surface openings.

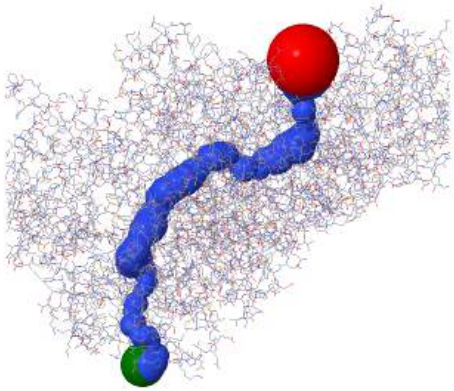


Figure 18.3.5: Second view of the tunnel in human carbamoyl phosphate synthetase I

The top panel in Figure 18.3.6 shows the radius in Angstroms of the tunnel going from the Red sphere to the Green sphere in the image above. The bottom panel shows a measure of the hydrophobicity (red)/hydrophilicity (blue) of the amino acids surrounding the tunnel.

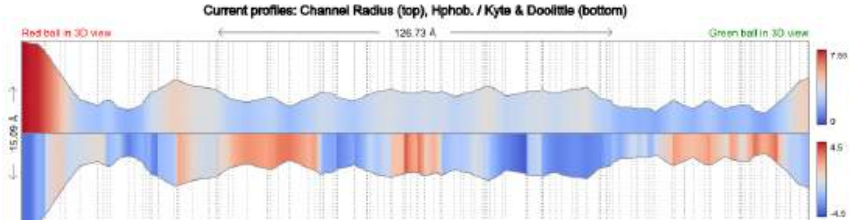


Figure 18.3.6: Tunnel size and polarity profile in human carbamoyl phosphate synthetase I

18.3.3: ASS1: ARGINOSUCCINATE SYNTHASE 1

Another ATP is cleaved to drive the conversion of citrulline to arginosuccinate, as shown in Figure 18.3.7. The product could easily be cleaved in the next step to form the amino acid arginine.

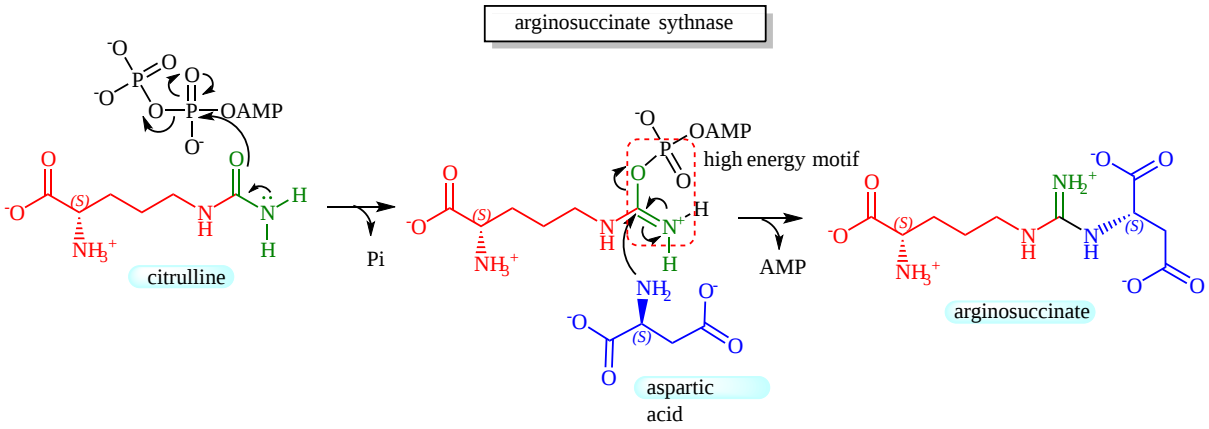


Figure 18.3.7: ATP cleavage drives the conversion of citrulline to arginosuccinate

18.3.4: ASL: ARGINOSUCCINASE LYSASE (AKA ARGINOSUCCINASE)

This enzyme catalyzes a beta-elimination reaction, which may proceed through either an E2 or E1 mechanism. A general but very abbreviated mechanism showing a two-step (E1) elimination proceeding through a carbanion intermediate is shown in Figure 18.3.8: below.

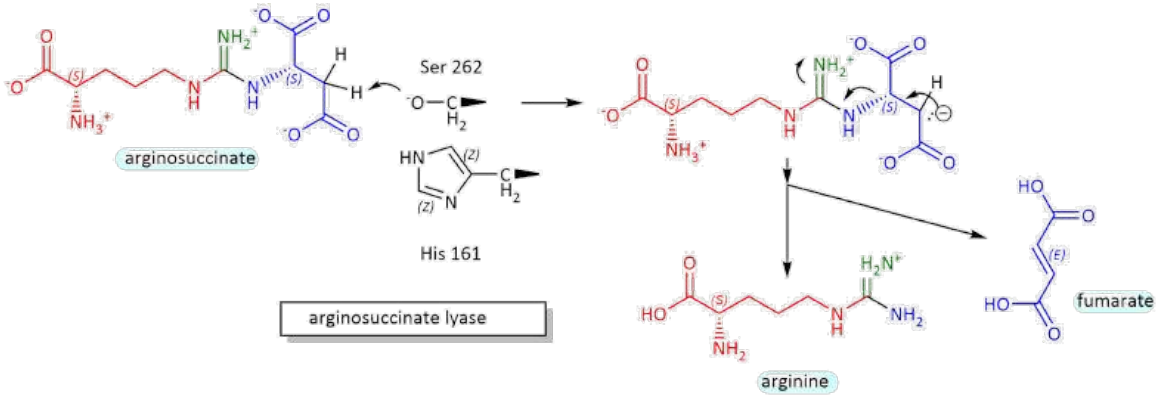


Figure 18.3.8: Mechanism for conversionof arginosuccinate to arginine and fumarate

Figure 18.3.9 shows an interactive iCn3D model of the tetrameric arginosuccinate lyase (ASL) from *Mycobacterium tuberculosis* (6IEN).

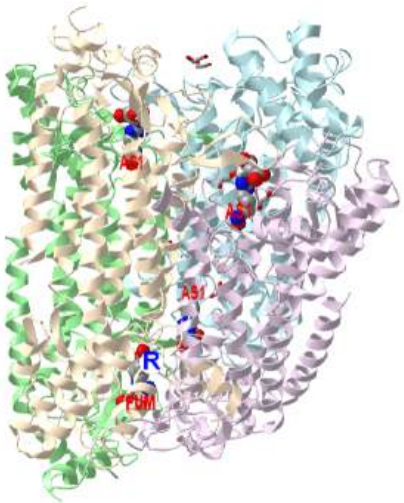


Figure 18.3.9: Tetrameric arginosuccinate lyase (ASL) from *Mycobacterium tuberculosis* (6IEN). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...32EuouGQ6xMVk6

Each monomer of the four monomers is shown in a different color. 3 arginosuccinates (substrate) and 1 fumarate and 1 arginine (products) are shown in spacefill, with CPK colors and labeled. Each monomer has an N-, M-, and a C-domain. The four binding sites consist of

residues from three monomers. The products are in the fourth site. Catalytic residues probably include a serine and histidine act as general acids/bases with a deprotonated lysine making the serine a general base.

18.3.5: ORNITHINE TRANSCARBAMYOLASE

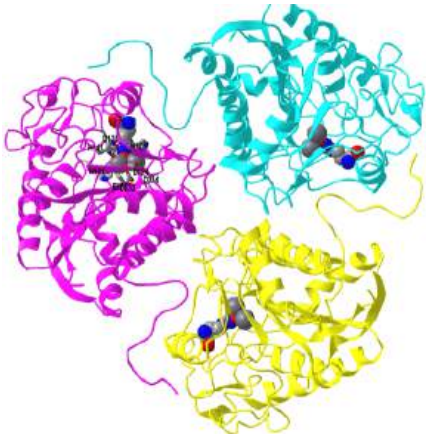
The reaction mechanism here is quite simple. The carbamoyl phosphate is an activated electrophile in which the carbonyl C is attacked by a deprotonated amino of ornithine to produce citrulline. The gene is found on the X chromosome in humans so mutations in males lead to a "hyperammonemic coma" and death. Heterozygous females can be asymptomatic and that generally proves to be fatal. Heterozygous females are either asymptomatic or have problems arising from defects in pyrimidine biosynthesis which a low-protein diet can alleviate with the addition of arginine.

18.3.6: ARGINASE

There are two general variants of this enzyme, Arginase I is a cytosolic enzyme expressed in the liver and is involved in urea synthesis. Arginase II is a mitochondrial enzyme and is involved more generally in arginine metabolism. Arginase I converts arginine to ornithine and urea. The enzyme has 2 Mn ions in the active site. A hydroxide ion bridges the two Mn ions and acts as a nucleophile which attacks the guanidino group of arginine and forms a tetrahedral intermediate. An active site Asp 128 acts as a general acid and protonates the amino leaving group to form ornithine and urea. Water then adds back to the Mn dinuclear cluster and ionizes, donating a proton to solvent through an intermediary His 41 as it reforms the active enzyme.

The model below shows the trimeric human arginase I (3KV2) in complex with the inhibitor N(omega)-hydroxy-nor-L-arginine, in which one of the guanidino Ns is replaced with an OH. The two grey spheres are Mn ions.

Figure 18.3.10 shows an interactive iCn3D model of the trimeric human arginase I in complex with the inhibitor N(omega)-hydroxy-nor-L-arginine (3KV2).



NCBI iCn3D Figure 18.3.10: Trimeric human arginase I in complex with the inhibitor N(omega)-hydroxy-nor-L-arginine (3KV2). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...QaAQpt5t2KKic6

Figure 18.3.11 show an active site water (#480) near the Mn ions.

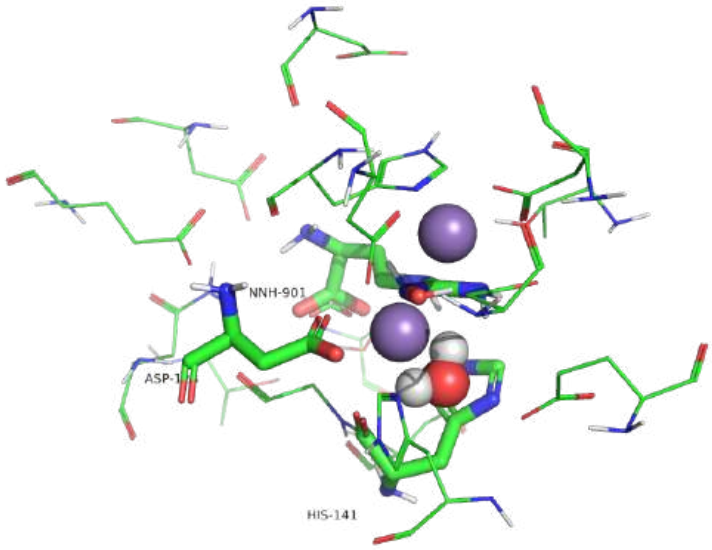


Figure 18.3.11 : Active site of human arginase I with an active site water

Water molecules can act as ligands and bind transition state metal ions through a coordinate covalent bond. The pKa of the bound water shifts to a lower value, making it more likely to deprotonate to form OH^-, which is both a better base and nucleophile. Water 480 in the active site appears to form a coordinate covalent bond to the Mn ion forming an active site hydroxide that attacks the guanidino group of arginine, leading to the formation of urea and ornithine.

18.3.7: REGULATION

If excess amino acids/proteins are consumed, a change in gene expression can increase levels of the enzymes in the urea cycle. In the short term, the activity is regulated by the enzyme that provides access into the cycle, carbamoyl phosphate synthase, whose activity is regulated by N-acetylglutamic acid and which catalyzes the rate-limiting step. As shown above, without the allosteric regulator, NAG, the enzyme is functionally inactive. The levels of NAG are determined by the enzyme NAG synthase, which itself is regulated by free arginine, the immediate precursor of urea in the cycle. Supplemental arginine is given to patients with urea disorder system and its effect probably occurs through the regulation of NAG synthase.

18.3.8: THE LINK BETWEEN THE UREA CYCLE AND THE TCA CYCLE

Two of the major pathways you have explored, the TCA and urea cycles, are not linear but cyclic pathways. What advantages do "circular" offer over linear ones? Perhaps a comparison to economic pathways offers a clue. In a linear economy, raw materials are used to produce a product which is eventually discarded as trash. Unfortunately, linear pathways dominate our world at a great cost to our environment. To reduce the inefficiencies in a linear economy and make it more viable and to decrease environmental damage and associated climate change, the world should move to a circular economy. A circular economy is characterized by what has been referred to as the 3Rs: reduce, reuse, and recycle. This saves resources and energy. These characteristics also apply to circular biochemical pathways as reduced levels of metabolites get reused and recycled.

Given this, it should not come as a surprise that the urea and TCA cycle are linked, especially given that fumaric acid is a product of the urea cycle and a cyclic metabolite of the TCA cycle. In addition, aspartic acid is one transamination step away from oxaloacetate. The interconnections between the urea cycle and the TCA cycle are shown in Figure 18.3.12 along with two additions, the malate aspartate shuttle, and the aspartate arginosuccinate shunt.

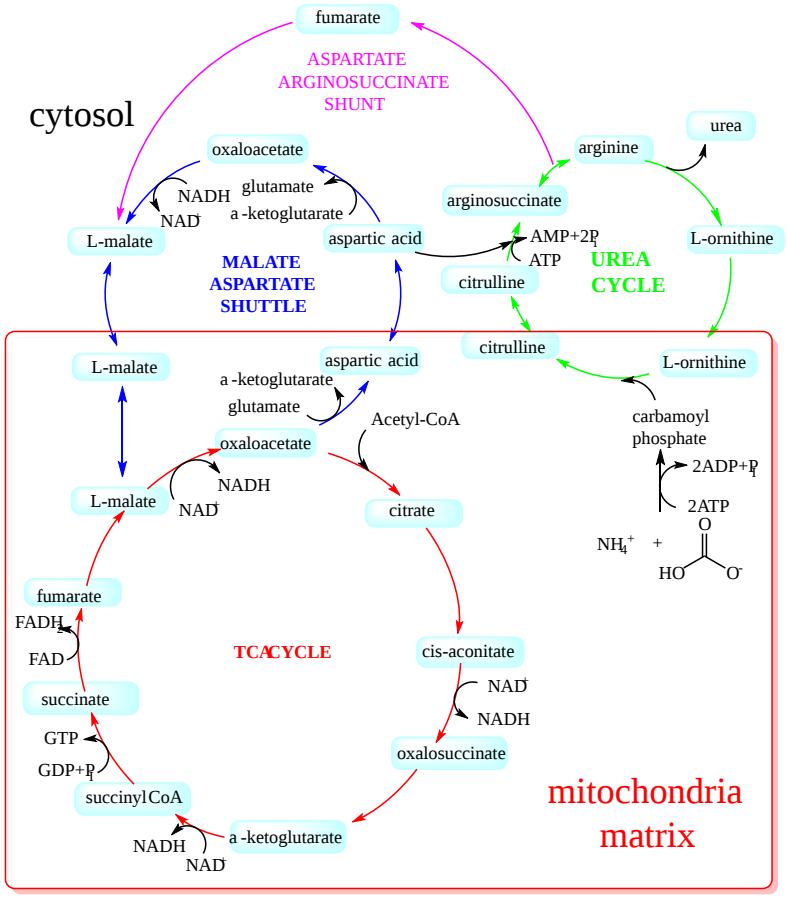


Figure 18.3.12: Urea and TCA cycles linked by the malate aspartate shuttle and the aspartate arginosuccinate shunt.

One of the nitrogens in urea comes from the amine of aspartate, which can be formed through the transamination of oxaloacetate from the TCA cycle to form aspartate. Since the TCA cycle is a cycle, an equivalent of one oxaloacetate must return to the cycle. It does so indirectly from fumarate, a TCA intermediate produced as a by-product of the urea cycle, but the fumarate is produced in the cytoplasm. There, it gets converted to malate, which can get transported into the mitochondria by a transport protein, mitochondrial 2-oxoglutarate/malate transporter, that is part of yet another cycle shown in the figure above, the **aspartate arginosuccinate shunt.** The introduction of fumarate into the shuttle occurs through the **aspartate arginosuccinate shunt** shown in the figure above**.** Once malate enters the mitochondrial matrix, it can be directly converted to oxaloacetate.

Note that these linked cycles require the presence of both mitochondrial and cytoplasmic forms of the several enzymes and balances between two key amino acids across the mitochondrial divide, glutamate, and aspartate. One such key enzyme is **AS**partate Amino **T**ransferase (**AST), aka G**lutamate **O**xaloacetate **T**ransaminase (**GOT**), named for the reverse reaction of aspartate + α−ketoglutarate ↔ oxaloacetate + glutamate. As mentioned in the previous chapter section, this key enzyme can be found in the blood if the liver is damaged, so its levels are routinely used in medical tests for impaired liver function.

Another key function of the malate aspartate shuttle is to bring into the mitochondria reduced "equivalents" of NADH that are produced in the cytoplasm from glycolysis and fatty acid oxidation. There is no membrane transporter for NAD^+ or NADH. Instead, cytoplasmic malate, produced by the cytosolic reduction of oxaloacetate by NADH, can be transported into the matrix which can reform mitochondrial NADH on reconversion to oxaloacetate. The NADH can then be used to power ATP production through mitochondrial electron transport/oxidative phosphorylation pathways.

From an energetic perspective, four phosphoanhydrides bonds in 3 ATP molecules are used to produce one urea molecule. This large energy expenditure is partially compensated for by ATP made from the "fumarate equivalents" that enter the mitochondria through the aspartate arginosuccinate and aspartate arginosuccinate shunts and the passage of NADH through oxidative phosphorylation.

18.4: AN OVERVIEW OF AMINO ACID METABOLISM AND THE ROLE OF COFACTORS

18.4.1: AMINO ACID DEGRADATION

We saw how nitrogen is removed from amino acids to produce urea or NH_4^+ in the previous chapter section. What are the fates of the carbon skeletons that remain? This section is where students might get overwhelmed by the diversity of amino acid degradation pathways, so it helps to realize that carbon skeletons of deaminated amino acids can be used for biosynthesis or energy production and are converted to key glycolytic and TCA intermediates that you have seen many times before. Everything is interconnected which makes the study of metabolism daunting but also fascinating, as organisms try to extract all the energy and molecule-building atoms from a metabolite and minimize waste.

Here are some key features of amino acid catabolism that this chapter section will present.

- some are converted to pyruvate, the end product of glycolysis and the start reactant of gluconeogenesis. Hence, these amino acids are **glucogenic**;
- some are converted to acetoacetate-CoA and or acetyl-CoA. Both of these can be converted to ketone bodies (acetoacetate/β-hydroxybutyrate) so these are considered **ketogenic**. Since the two carbons of the acetyl group of acetyl-CoA are lost as CO_2 in the TCA cycle, and there is no reverse for the pyruvate dehydrogenase reaction (pyr → acetyl-CoA), acetyl-CoA formed by amino acid degradation can not be used to create glucose in net fashion;
- some are metabolized to form TCA intermediates. Since they are added in a **net** fashion to the TCA cycle and don't remove the existing pool of TCA intermediates, they can produce in a **net** fashion directly or indirectly molecules that can be used to produce glucose. These entry reactions to the TCA which replenish or add to TCA intermediates are called anaplerotic (replenishing) reactions. Hence these amino acids are also **glucogenic**.
- some have multiple ways to be degraded and can produce both acetyl-CoA and pyruvate, so they are **both glucogenic and ketogenic.**

Let's get more explicit:

- purely ketogenic: only Leu and Lys (the only amino acids whose name starts with **L** and you have to **L**ove them since there are only 2 amino acids in this category)
- both: 5 including the aromatics - Trp, Tyr, Phe - and Ile/Thr
- purely glucogenic: the rest

General hints as to if an amino acid is glucogenic and/or ketogenic can be derived from their structures.

Fully or Partly Ketogenic Amino Acids: Amino acids except for Arg that have a **continuous chain** of 3 or more carbon atoms in their **side chains** and hence are more "fat-like" are ketogenic or ketogenic/glucogenic. These include Lys, Leu, Ile, and the aromatic amino acids Phe, Trp, and Tyr and exclude Val. (note: Thr, which is both ketogenic and glucogenic, doesn't fit this rule). The fully or partly ketogenic amino acids are shown in Figure 18.4.1.

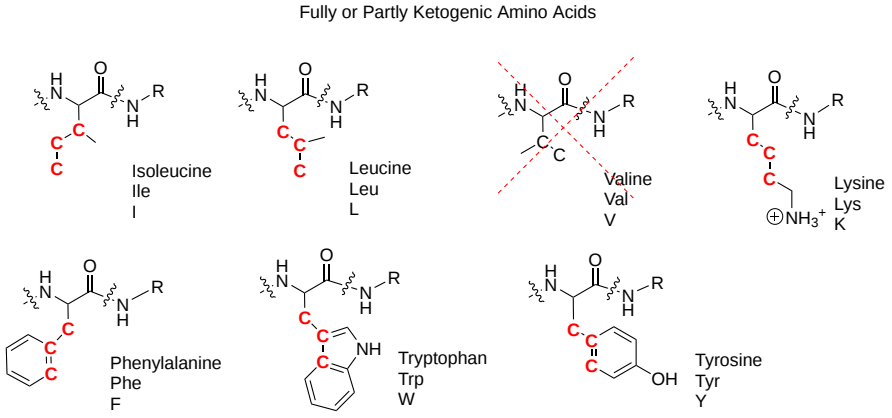


Figure 18.4.1: Fully or Partly Ketogenic Amino Acids

Purely Glucogenic Amino Acids (for Pyr and/or TCA intermediates): Which amino acids produces which intermediate? Except for glycine, it's pretty easy to remember. The number of carbons in the intermediate formed is the same as the number of carbons in the longest chain in the amino acid.

These amino acids form the 5C TCA intermediate alpha-ketoglutarate (2-oxoglutarate) and are shown in Figure 18.4.2.

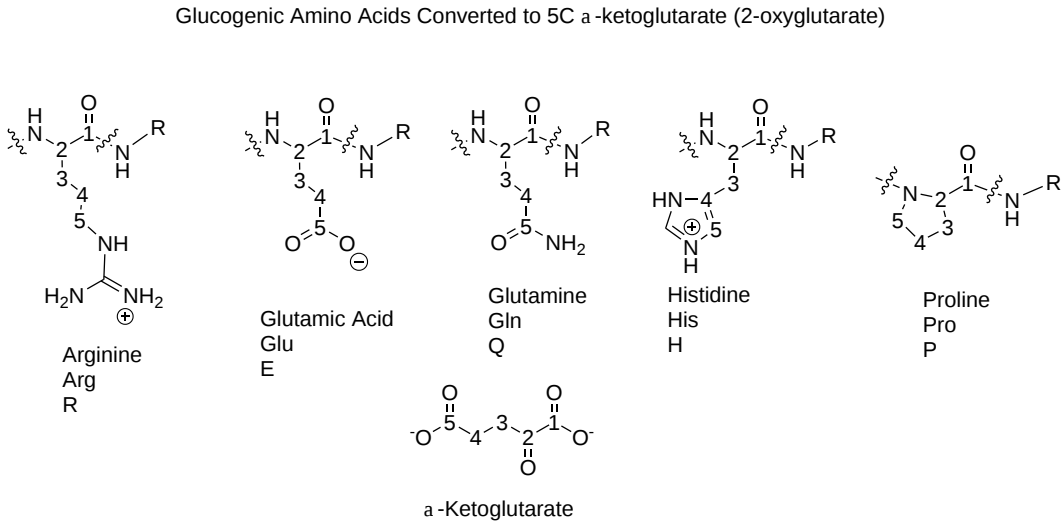


Figure 18.4.2: Glucogenic Amino Acids Converted to 5C α-ketoglutarate (2-oxoglutarate)

These amino acids form either the 4C TCA intermediate succinate or oxaloacetate. Amino acids with more oxidized 4C atoms in the continuous chain produced oxaloacetate (which is more oxidized than succinate), while the least oxidized ones produce succinate acid (which is more reduced than oxaloacetate). These amino acids are shown in Figure 18.4.3.

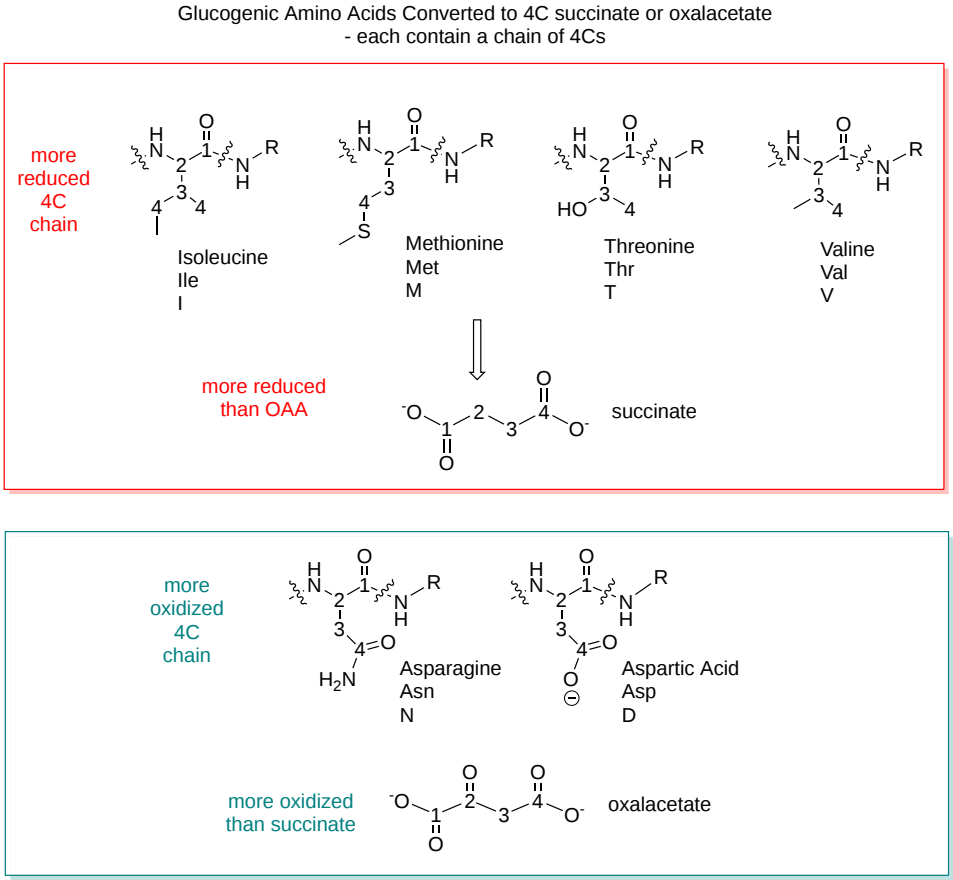


Figure 18.4.3: Glucogenic Amino Acids Converted to 4C oxaloacetate and succinate

The 3C amino acids, Ser, Ala, and Cys, except for glycine, form pyruvate, as shown in Figure 18.4.4.

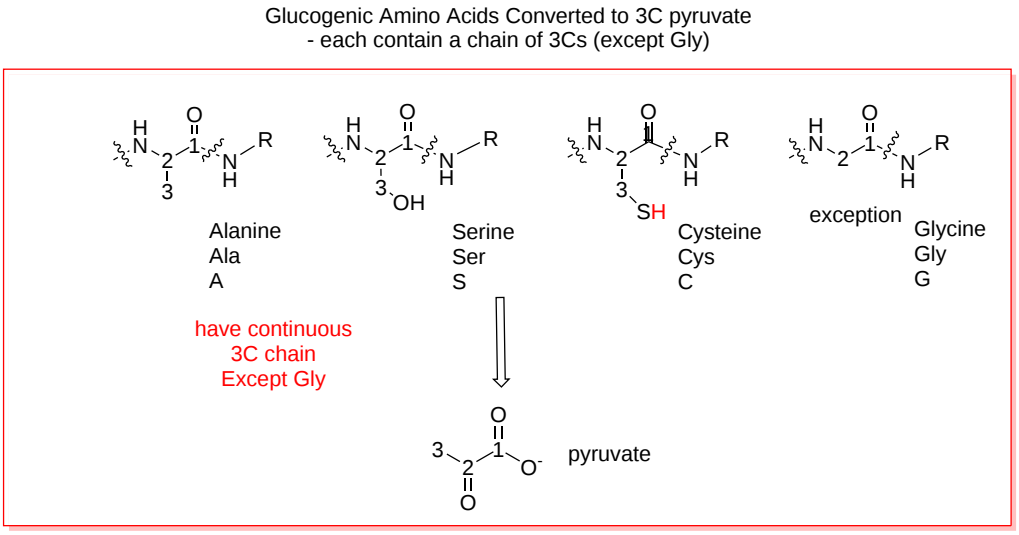


Figure 18.4.4: Glucogenic Amino Acids Converted to 3C pyruvate

A full figure summarizing everything above is shown in Figure 18.4.5!

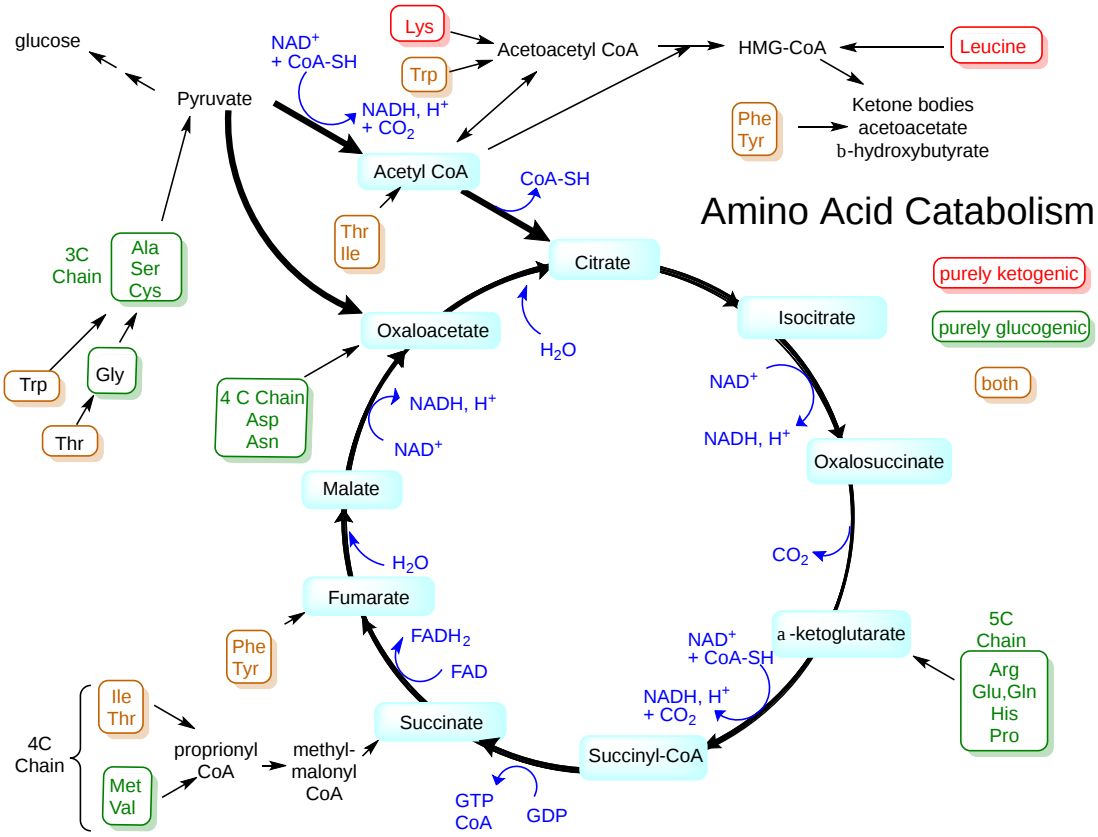


Figure 18.4.5: Summary of amino acid breakdown products

18.4.2: WITH A LITTLE HELP FROM MY FRIENDS: COFACTORS AND AMINO ACID CATABOLISM

The myriad of breaking, making, and rearranging of carbon atoms in amino acid catabolism is daunting. As with all reactions, a pathway for a flow of electrons and stabilization of transition states and intermediates must be in place for the reactions to occur. All methods of catalysis (general acid/base, covalent/nucleophilic catalysis, electrostatic/metal ion catalysis, preferential stabilization of the transition state occur. Yet some reactions need additional "helpers" or "cofactors" to facilitate electron flow and shuttle small motifs (from electrons in

redox reactions to methyl groups in methylases) from one molecule to another. We've seen the important role of pyridoxal phosphate in transaminations/aminotransferases in the previous section. Even the terminology cofactor is a bit confusing since analogous terms are used in different contexts:

- cofactors - nonprotein small molecules or ions (divalent ions, transition metal ions, and Fe/S clusters for example) that must bind to an enzyme, but once bound usually stay put. That is, they don't dissociate during the catalytic cycle. For instance, PLP must initially bind to an enzyme but then becomes covalently attached through a Schiff base through the epsilon amino group of a Lys in the catalytic site. That linkage might swap with an incoming amino acid substrate, for example, but reforms in the catalytic cycle so the enzyme remains functional.
- coenzyme - a vitamin derivative organic cofactor, as compared to inorganic ions, for example. Coenzyme is a poor historical name that persists.
- prosthetic group - this is usually a coenzyme that is either covalently bound (like PLP) or noncovalently bound with such a low K_D (like FAD/$FADH_2$-containing enzymes) that they stay bound
- cosubstrate - these bind as substrate (for example NAD^+), and depart as products (for example NADH).

Here is a table of common nonmetallic cofactors.

Cofactor	Vitamin Derivative	Carrier
biotin	H - biotin	1 C - CO_2 (most oxidized)
tetrahydrofolic acid (FH_4 or THF)	B9 - folic acid	1C - formyl, -(C=O)H, methylene (-CH_2) (more reduced) and methyl -CH_3
cobalamin	B12 - cobalamin	methyl CH_3
thiamine pyrophosphate	B1 - Thiamine	2C group
coenzyme A	B5 - pantothenic acids	acetyl (CH_3-C=O) and acyl (R-C=O)
pyridoxal phosphate	B6 - pyridoxine	amino and carboxyl groups
NAD^+/$NADP^+$	B3 - Niacin	electrons
FAD/FMN	B2 - riboflavin	electrons

Examples of other nonvitamin cofactors include S-adenosylmethionine (SAM or adoMet), a carrier of methyl groups, coenzyme Q, a carrier of electrons, tetrahydrobiopterin, a carrier of oxygen and electrons, and of course heme, a carrier of electrons.

We've already described amino acid deamination reactions using PLP. Now let's consider the biochemistry of two of these cofactors involved in 1 C transfers in amino acid metabolism in more detail, tetrahydrofolate (FH_4) and SAM.

18.4.3: TETRAHYDROFOLATE (FH_4)

Tetrahydrofolate (FH_4) is a key cofactor in the metabolism of amino acids but it is also critical in the biosynthesis of nucleotides. It is also commonly abbreviated as THF, especially in the names of enzymes that use it. To avoid confusion with chemist's use of THF for tetrahydrofuran, it will be referred to mostly here as FH_4 but the enzymes will be named as derivatives of the abbreviation XHF (X = D for dihydro and T for tetrahydro.

FH_4 is a carrier of 1 C units in various oxidation states. It receives 1C units and then transfers them to other species. The structure of FH_4 and derivatives carrying 1 C units with different oxidation numbers (indicating their oxidation state) are shown in Figure 18.4.6.

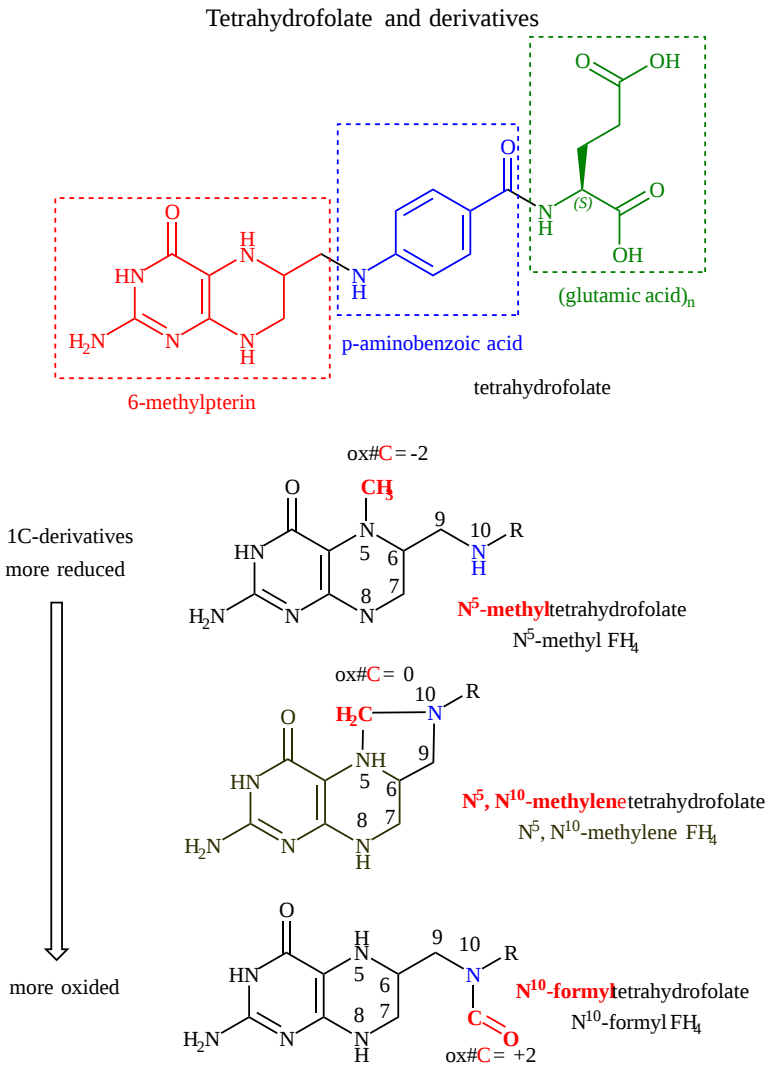
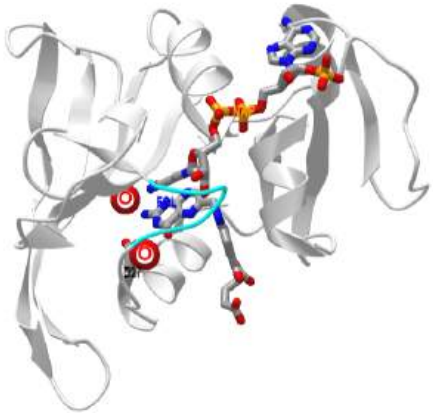


Figure 18.4.6: 1C derivatives in different oxidation states for tetrahydrofolate

FH_4 is made from the vitamin folate (bacteria, fungi, and most plants can synthesize it), which gets converted first to dihydrofolate (FH2), and then tetrahydrofolate (FH_4), using NADPH as a reducing substrate "cofactor", by the enzyme **dihydrofolate reductase**. The model below shows this small enzyme with NADPH ($NADP^+$) and folate bound to DHFR (7dfr). Folate (FOL) is shown in salmon space fill while NADPH is shown in cyan.

Figure 18.4.7 shows an interactive iCn3D model of dihydrofolate reductase with bound NADPH ($NADP^+$) and folate (7dfr).



NCBI iCn3D Figure 18.4.7: Dihydrofolate reductase with bound $NADP^+$ and folate (7dfr). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...ojdpBZuQH8Lm96

Folate (FOL) and $NADP^+$ (NAP) are shown in spacefill and CPK colors. The side chain of Asp 27 is shown in sticks and CPK colors. The oxygen atoms of two water molecules interacting with folate are shown as red spheres. A disordered and hence mobile loop (residues 16-20) that is involved in binding the nicotinamide is shown in cyan.

As FH_4 is a key substrate in many reactions, including nucleotide synthesis (as we shall see later), this enzyme is a key target for chemotherapy to kill cancer cells which require robust nucleotide synthesis for rapid cell proliferation. One such drug, which inhibits the enzyme, methotrexate, is shown in Figure 18.4.8.

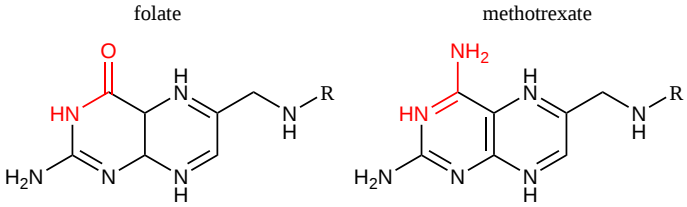


Figure 18.4.8: A structural comparison of the chemotherapeutic drug methotrexate and folate

Before getting into the nitty-gritty of amino acid degradation and tetrahydrofolate chemistry, let's take a look at the mechanism by which dihydrofolate reductase, DHFR, catalyzes the sequential reduction by 2 NADPH molecules of folate to tetrahydrofolate. This is illustrated in Figure 18.4.9.

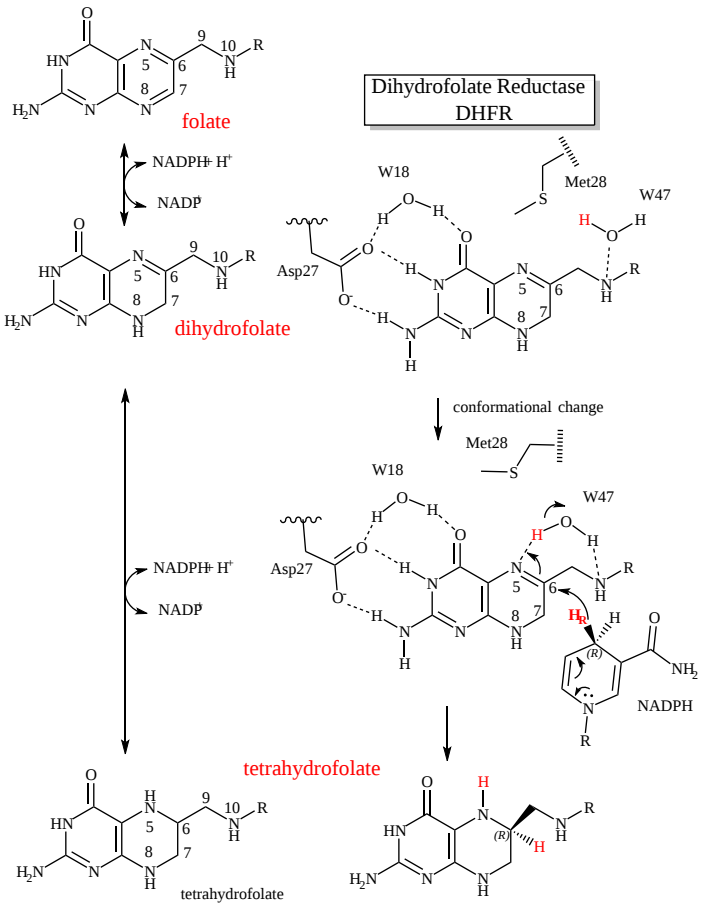


Figure 18.4.9: Conversion of folate to tetrahydrofolate by dihydrofolate reductase

18.4.4: ONE CARBON CHEMISTRY: THE INTERCONVERSION OF FH_4 1C DERIVATIVES

Tetrahydrofolate is a carrier for 1C units in metabolism. As such, is it intimately involved in many anabolic and catabolic reactions. These include thymidine and purine biosynthesis and amino acids metabolism through reactions involving serine, glycine (both with 1C in their side chains), and methionine (with 1C after the sulfur in the side chain). The reactions take place in both the cytoplasm and mitochondria. FH_4 is involved directly or indirectly in epigenetic control of DNA expression as well, as methylation of both DNA and histones is critical to gene expression. The myriad of 1C derivatives of FH_4 make the biochemistry complex, but as the transfer of 1C is a critical job in both the breakdown of the carbon skeleton of amino acids and biosynthesis, we need to explore it. The complexity is simplified by noting that the 1C derivatives have only three different oxidation states (+2, 0, and -2), as noted in Figure 18.4.10.

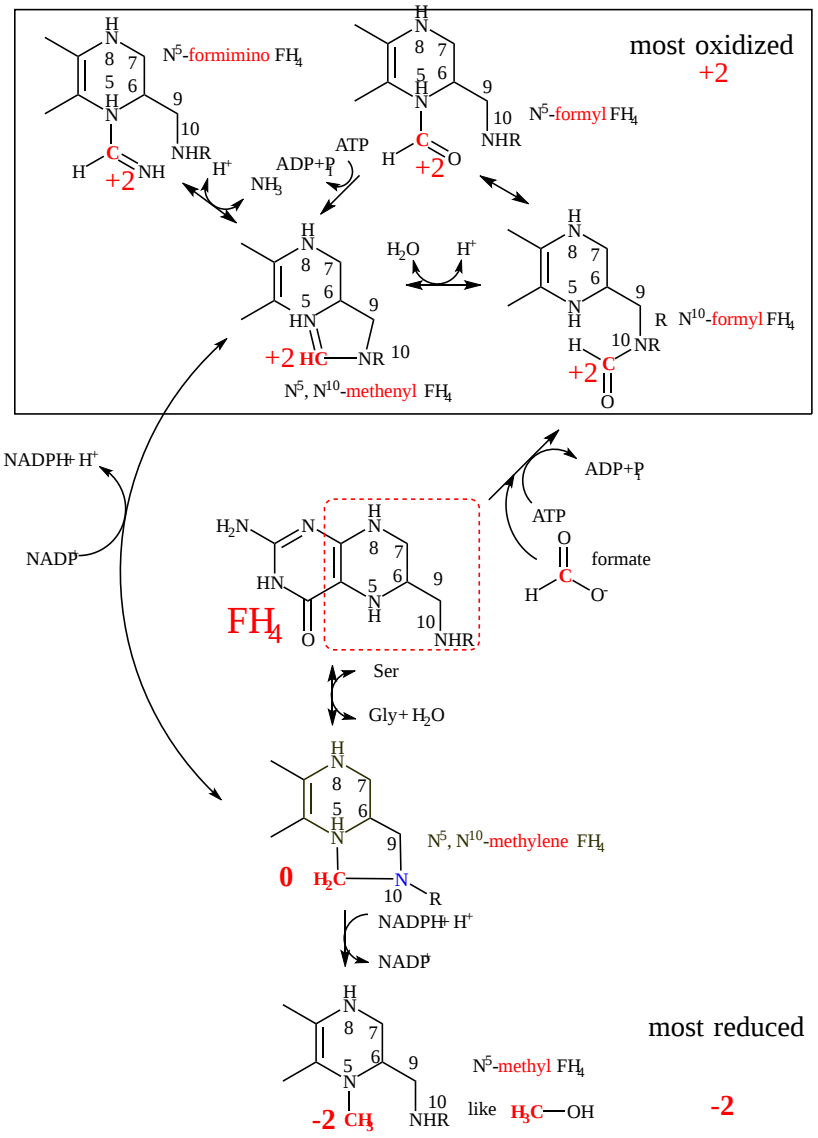


Figure 18.4.10: Interconversion of 1C derivatives of tetrahydrofolate. The oxidation number of the 1C addition is shown in red. (after http://education.med.nyu.edu/mbm/aminoAcids/folate.shtml)

1C units usually enter FH_4 as the N^5,N^{10} methylene unit. This is made in both the cytoplasm and mitochondria since 1C-derivatized FH_4 appears not to cross the mitochondrial membranes. The methylene derivative, with a 1C oxidation # of 0, can be reduced to form methyl (oxidation # -2) or oxidized to methenyl or formyl groups (oxidation # +2).

We will see some of these reactions again in chapters dealing with amino acid and nucleotide biosynthesis. That's not a bad thing - learning occurs best on the repetition of material in different contexts.

18.4.5: SERINE HYDROXYMETHYLTRANSFERASE (SHMT): A COMPLEX REACTION NEEDED TWO COFACTORS - PLP AND FH_4

Let's look in detail at one mechanism that shows how a 1C methylene is added to tetrahydrofolate (FH_4). The mechanism shown is for serine dehydratase, aka serine hydroxymethyltransferase. The enzyme not only uses tetrahydrofolate as a substrate but also PLP, which, as we have seen previously, makes bonds to the alpha-carbon of amino acids labile to cleavage. In this case, the amino acid serine becomes dehydrated through an alpha-elimination reaction. Here is the overall reaction.

$$\text{Serine} + FH_4 \leftrightarrow \text{Glycine} + N^5,N^{10}\text{-}CH_2\text{-}FH_4 + H_2O$$

Figure 18.4.11 shows the dehydration reaction and formation of glycine. using PLP as a cofactor. Figure B shows how the released formaldehyde reacts with FH_4 to form N^5,N^{10}-methylene FH_4. using FH_4 as a cofactor. As this is needed in purine and thymidylate

synthesis, SHMT is a target for malaria treatment as well.

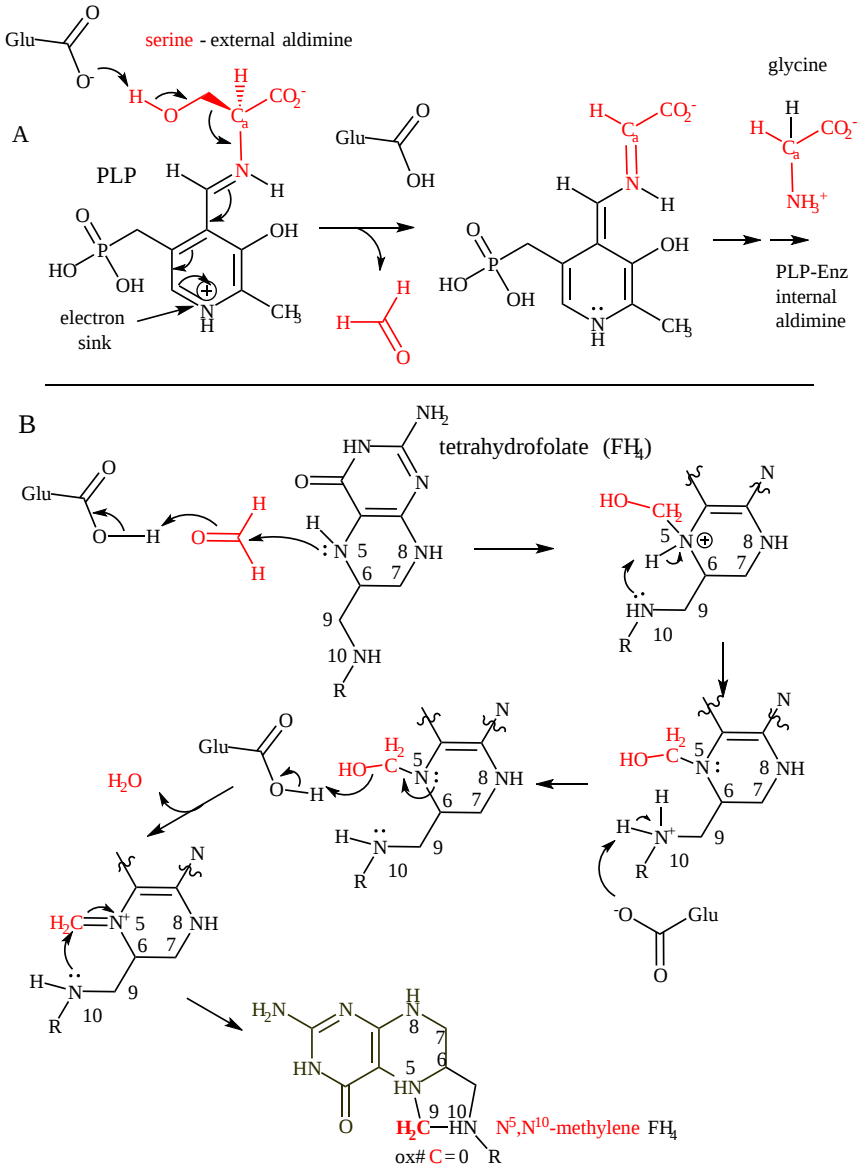


Figure 18.4.11: Serine hydroxymethylltransferase reaction mechanism

Glu 57 plays a key role as a general acid/base throughout the catalytic cycle of the enzyme.

Pathway diagrams showing a myriad of reactants, products, and enzymes can be confusing to students (and to authors as well). It's useful to think about them and see them in different ways. Here is another way to present the conversion of FH_4 and its various 1C intermediates. Figure 18.4.12 concentrates on the **outputs** of 1C FH_4 derivatives (shown in blue eclipses).

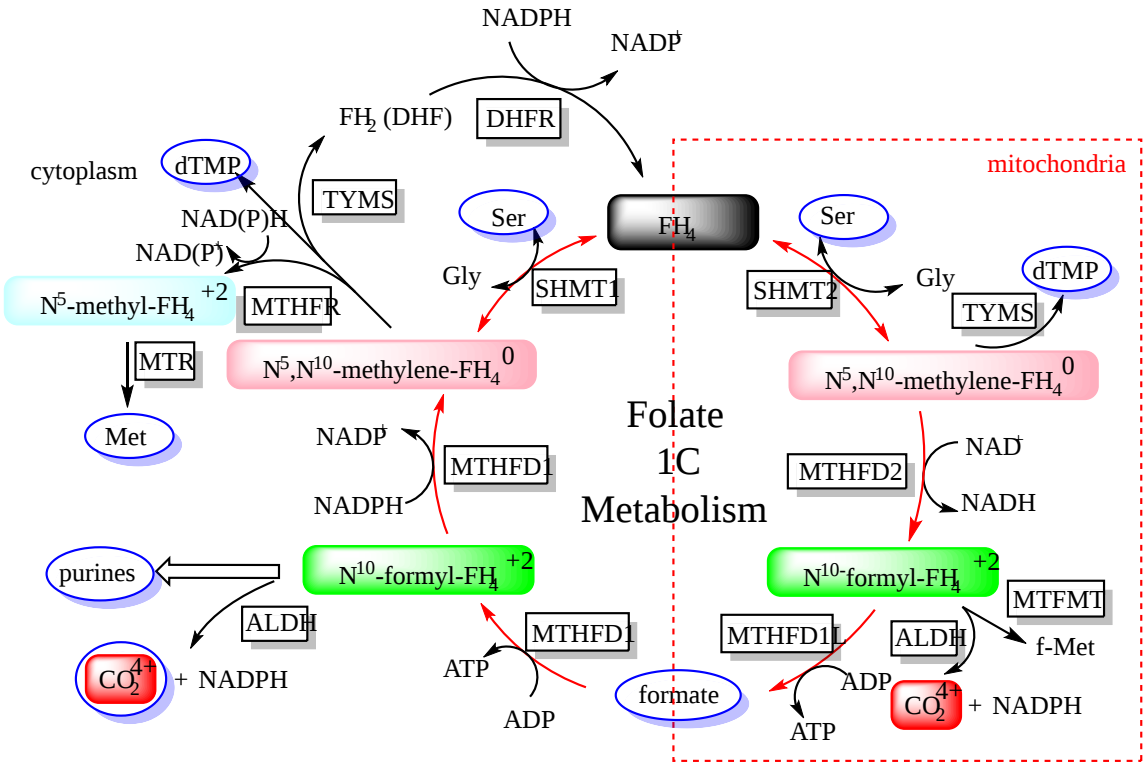


Figure 18.4.12 Folate 1C Metabolism - Cyclic Presentation after Ducker and Rabinowitz, Cell Metab. 2017 January 10; 25(1): 27–42. doi:10.1016/j.cmet.2016.08.009

The enzymes catalyzing these reactions are shown in the table below. This figure applies to both catabolic and anabolic reactions using FH_4 derivatives. Oxidation numbers for the 1C adducts are shown. Any reaction that involves a change in redox state must use $NAD(P)^+$/NAD(P)H as a redox reagent.

- N^5,N^{10}-methylene FH_4 gives thymidine and serine
- N^5-methyl FH_4 gives methionine
- N^{10}-formyl FH_4 gives formate, purines, and CO_2
- N^5-formyl serves more as a passive reservoir of 1C units.

Table 18.4.1: Enzymes catalyzing interconversions of THF derivatives

Abbreviation	Enzyme Name
ALDH	10-formyltetrahydrofolate((aldehyde) dehydrogenase.
DHFR	Dihydrofolate reductase
MTHFD	methylenetetrahydrofolate dehydrogenase
MTHFD1	C-1-tetrahydrofolate synthase, cytoplasmic
MTHFD1L	monofunctional tetrahydrofolate synthase, mitochondrial
MTHFD2/L	methylenetetrahydrofolate dehydrogenase 2/2-like
MTFMT	mitochondrial methionyl-tRNA formyltransferase
MTHFR	methylenetetrahydrofolate reductase
MTR	methionine synthase
TYMS	thymidylate synthetase

A key enzyme in these reactions, methylene-THF reductase (MTHFR), irreversibly removes 1C from the cycle depicted above as it forms N^5-methyl FH_4. As this reaction is a reduction, it requires a reducing agent (NADPH in yeast and animals and NADH in plants). This irreversible removal would deplete the cycle shown so it is allosterically regulated (inhibited) by another methylating agent, S-adenosylmethionine (SAM aka adoMet), as we will see next. Plant versions of the enzyme are reversible so there is no need for regulation by SAM in this feedback loop process.

18.4.6: S-ADENOSYLMETHIONINE (SAM), ALSO KNOWN AS ADOMET

N^5-methyl FH_4 appears to have one function, to methylate a molecule called homocysteine (same as Cys but with an extra -CH_2 in the side chain) to methionine. This is adenosylated (not phosphorylated!) with ATP to produce S-adenosylmethionine (SAM), a more potent methylating agent than N^5-methyl FH_4. SAM is hence part of a cycle involving N^5-methyl-FH_4. The methyl group of N^5-methyl FH_4 reacts with homocysteine to produce methionine, catalyzed by the enzyme methionine synthase, which requires cobalamin (vitamin B12) as a cofactor. The combined Figure 18.4.13 adds the Met and Folate cycles (showing the outputs of the 1C Folate metabolism cycle

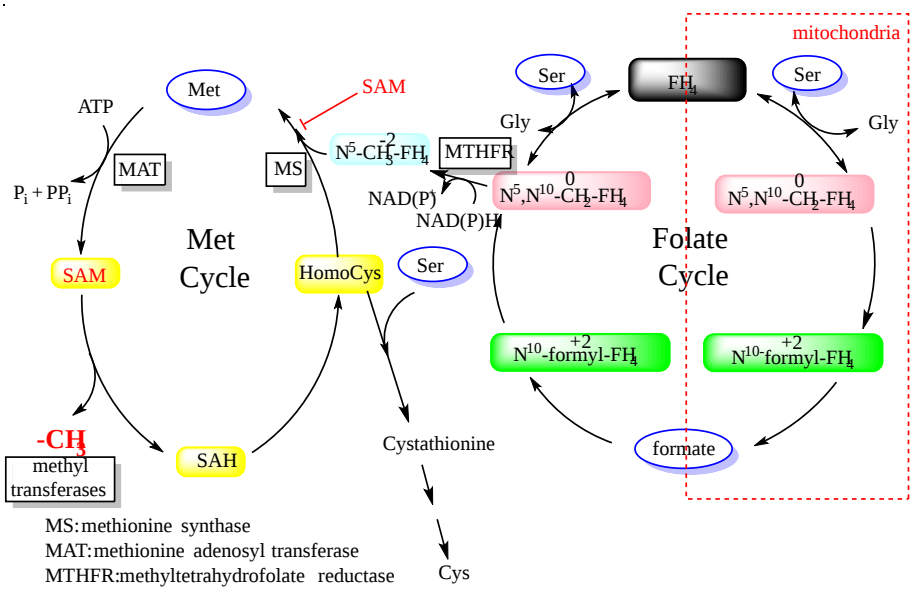


Figure 18.4.13 Folate and Methionine Cycle (after Ducker and Rabinowitz, Cell Metab. 2017 Jan 10; 25(1): 27–42.)

Figure 18.4.14 shows the structures of molecules in the Met Cycle.

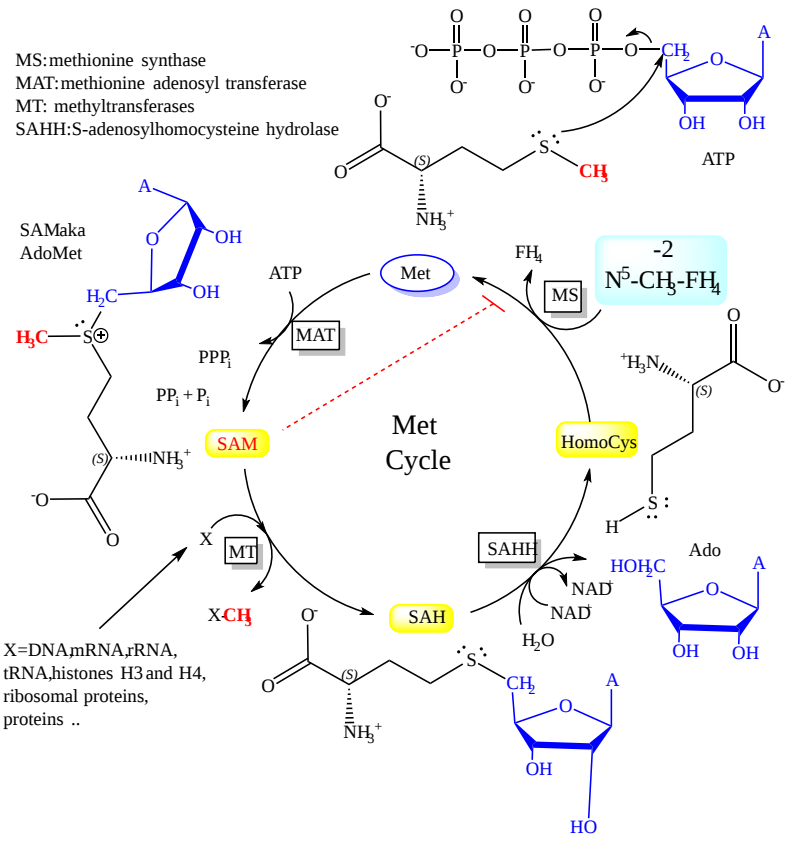
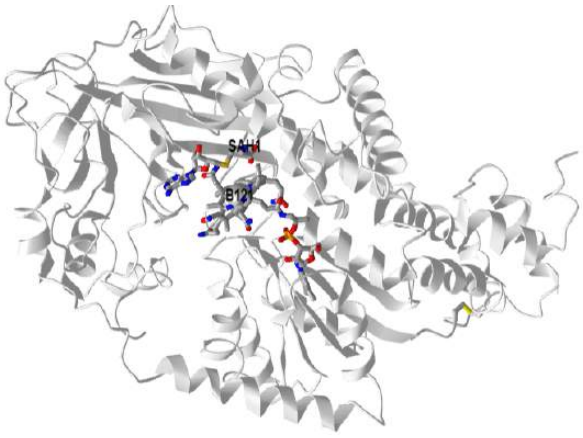


Figure 18.4.14: The Met Cycle and Methylations by SAM

The mechanism of methyl transfers using SAM as the -CH_3 donor involves a SN_2 attack by a nucleophile of the substrate on the CH_3 of SAM, with the electron pair from the C-S bond going to the positively charged sulfonium sulfur, a great "electron sink". An analogous nucleophilic attack on the terminal CH_3 of plain old methionine would not be readily enabled. Hence SAM, with its charged S, has a much high methyl transfer potential than N^5-CH_3-FH_4.

In the actual reaction catalyzed by methionine synthase (MS) in mammals, the methyl CH_3 from N^5-methyl FH_4 is first transferred to cobalamin, a derivative of vitamin B12, to form methylcobalamin, which then transfers it to homocysteine. The structure of the C-terminal half of B_{12}-dependent Methionine Synthase from *E. Coli* with bound adenosylhomocysteine bound (3iva) is shown below and in this link:

Figure 18.4.15 shows an interactive iCn3D model of the C-terminal half of B_{12}-dependent Methionine Synthase from *E. Coli* with bound adenosylhomocysteine bound (3iva)



NCBI iCn3D Figure 18.4.15: C-terminal half of B_{12}-dependent Methionine Synthase from *E. Coli* with bound adenosylhomocysteine bound (3iva). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...E7jTsdvVvx1M1A

In mammals, vitamin B_{12} is a key cofactor in only two enzymes, one being methionine synthase. If vitamin B_{12} is lacking, N^5-methyl FH_4 builds up, and since the enzyme that converts N^5,N^{10}-methylene FH_4 to N^5-methyl FH_4, methylenetetrahydrofolate reductase (MTHFR), is irreversible. This leads to megaloblastic anemia as precursors to red blood cells that can't mature. Folate deficiencies also lead to anemia.

18.4.7: THE SERINE GLYCINE ONE CARBON (SGOC) METABOLIC CYCLE

Figure 12 shows the coupled Folate and Methionine cycles that emphasize the intermediates involved in the 1C -CH_3 transfer reaction, an important part of amino acid metabolism and other anabolic and catabolic reactions as well. Interpreting metabolic figures is complicated. Each is designed to emphasize certain selected features. Another way to present Figure 12 is to emphasize metabolites involved in 1C chemistry in general. Figure 18.4.145 shows what has been called the **Serine Glycine One Carbon (SGOC)** metabolic cycle. It is just a redrawn version of Figure 12 with attention drawn to non-FH4 molecules involved in 1C transfers, namely serine, glycine, and also formate.

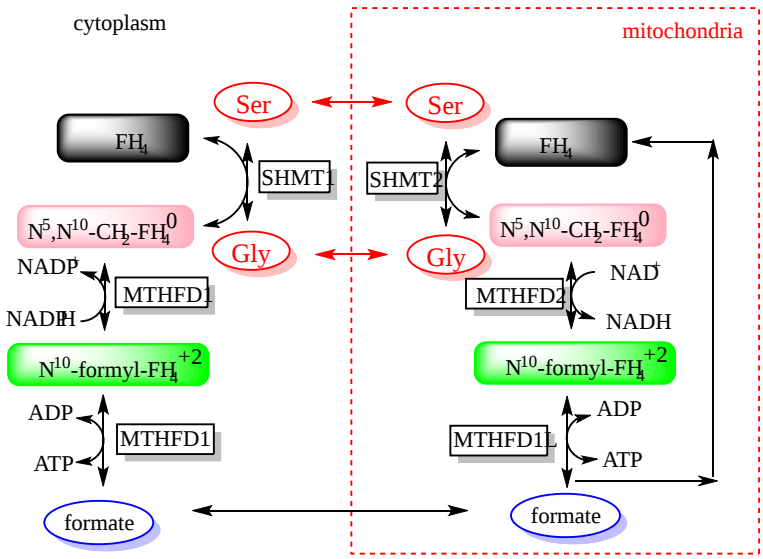


Figure 18.4.15: The Ser-Gly One Carbon (SGOC) Cycle

This cycle and its key substrates, serine, and glycine, integrate many metabolic pathways and controls the conversion of serine and glycine into outputs essential for other pathways. We will see this cycle again in the chapter on the biosynthesis of amino acids. The pathway is especially important in tumor cells, which need precursors for nucleic acid, protein, and lipid synthesis.

18.5: PATHWAYS OF AMINO ACID DEGRADATION

18.5.1: INTRODUCTION

In previous sections, we saw how nitrogen is removed from amino acids to produce urea or NH_4^+, that some amino acids are glucogenic, ketogenic, or both, and the role of tetrahydrofolate derivatives and S-adenosylmethionine in 1C transfer reactions. Now we can focus on how the carbon skeletons of amino acids are processed during degradation.

Here are some key features of amino acid catabolism that were discussed in the previous section.

- some are converted to pyruvate, the end product of glycolysis and the start reactant of gluconeogenesis. Hence, these amino acids are **glucogenic**;
- some are converted to acetoacetate-CoA and or acetyl-CoA. Both of these can be converted to ketone bodies (acetoacetate/β-hydroxybutyrate) so these are considered **ketogenic**. Since the two carbons of the acetyl group of acetyl-CoA are lost as CO_2 in the TCA cycle, and there is no reverse for the pyruvate dehydrogenase reaction (pyr → acetyl-CoA), acetyl-CoA formed by amino acid degradation can not be used to create glucose in net fashion;
- some are metabolized to form TCA intermediates. Since they are added in a **net** fashion to the TCA cycle and don't remove the existing pool of TCA intermediates, they can produce in a **net** fashion directly or indirectly molecules that can be used to produce glucose. These entry reactions to the TCA which replenish or add to TCA intermediates are called **anaplerotic** (replenishing) reactions. Hence these amino acids are also **glucogenic**.
- some have multiple ways to be degraded and can produce both acetyl-CoA and pyruvate, so they are **both glucogenic and ketogenic.**

Let's get more explicit:

- purely ketogenic: only Leu and Lys (the only amino acids whose name starts with **L** and you have to **L**ove them since there are only 2 amino acids in this category)
- both: 5 are, including the aromatics - Trp, Tyr, Phe - along with Ile/Thr
- purely glucogenic: the rest

Figure 18.5.1, also shown in the previous sections, summarizes the fates of the 20 amino acids in their catabolic reactions

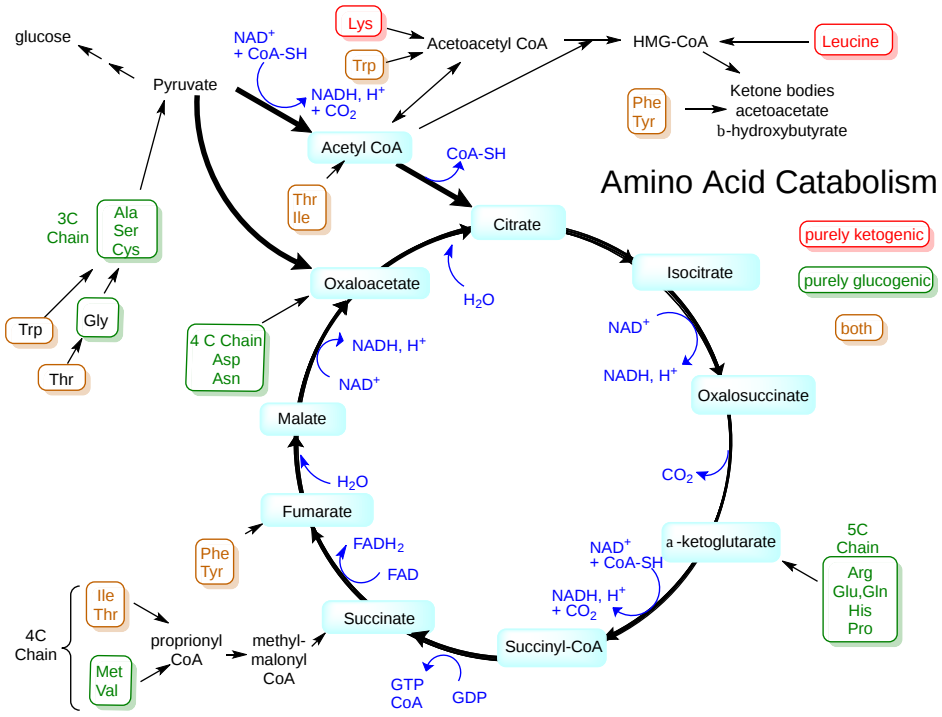


Figure 18.5.1: Fates of the 20 amino acids in their catabolic reactions

Given the myriad of enzymes and pathways involved, we won't delve into the mechanisms for the reactions or the structures of the enzymes, with the exception of one for lysine metabolism.

18.5.2: CONVERSION TO PYRUVATE: ALA, TRP, CYS, SER, GLY, THR

We ended section 18:3 with a discussion of the Ser Gly One Carbon Cycle (SGOC), so some of this will be a bit of a review.

Figure 18.5.2 shows an overview of the conversion of amino acids to pyruvate. More details are provided for each of the steps below.

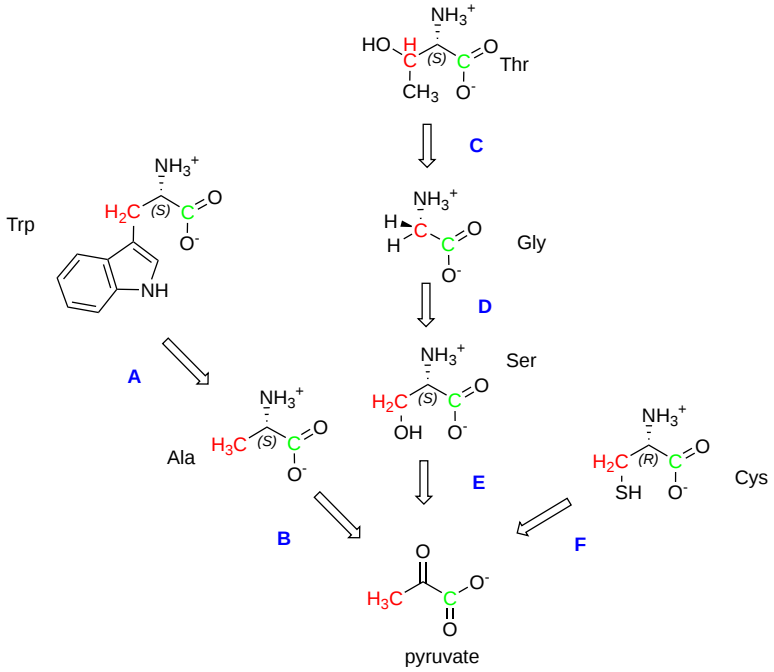


Figure 18.5.2: Overview of conversion of amino acids to pyruvate

The metabolic steps for the chemical transformations shown in A-F are described in more detail below.

18.5.2.1: TRYPTOPHAN TO ALANINE AND ON TO ACETOACETATE

This is a multistep process as shown in Figure 18.5.3.

A. Trp to Ala (and to acetoacetate and NAD+)

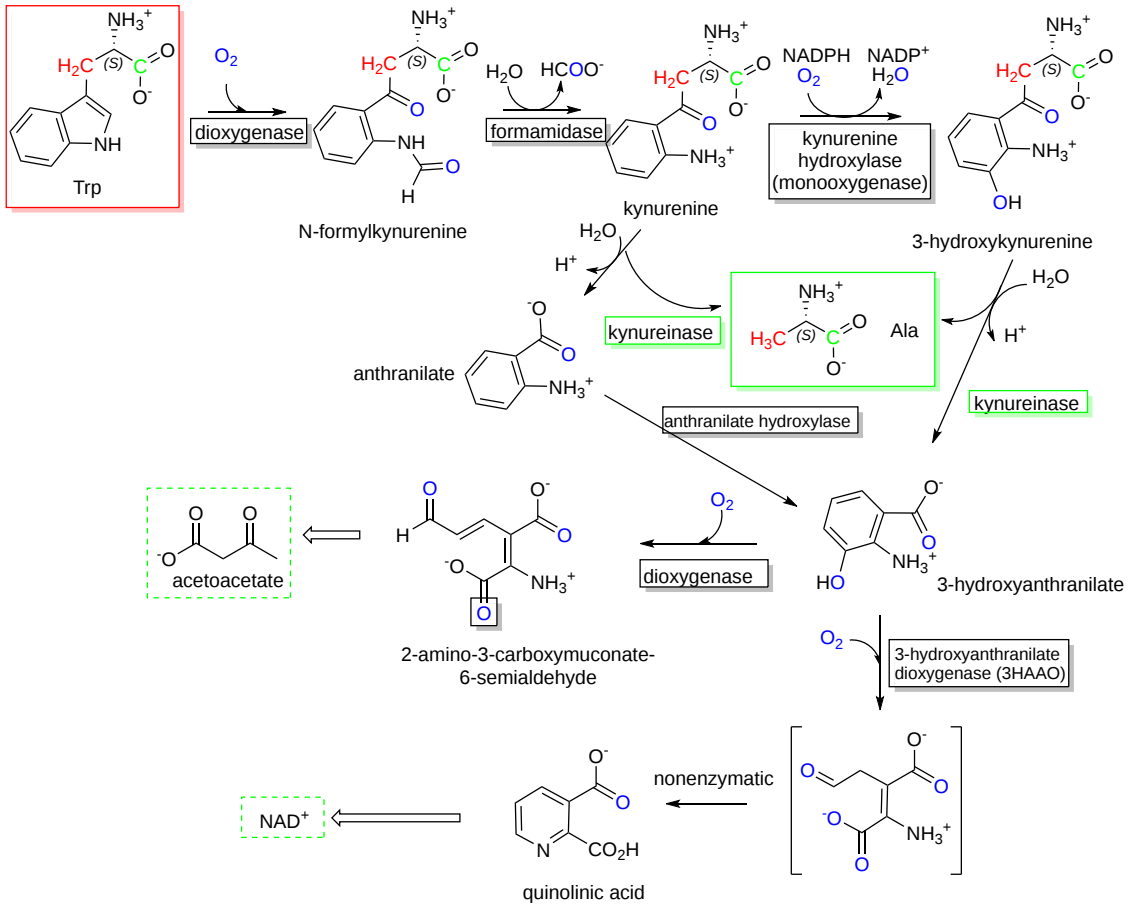


Figure 18.5.3: Conversion of Tryptophan to Alanine and to acetoacetate

The starting material, tryptophan, is highlighted in a red box while the end product of specific interest, Ala, is highlighted in a green box. No reaction occurs in isolation in a cell, but rather as part of a more complex pathway. In the figure below, Ala is presented almost as a side product as the modified aromatic ring found in either anthranilate or 3-hydroxyanthranilate continues on to form either acetoacetate, a ketone body which can breakdown to acetyl-CoA (making tryptophan ketogenic as well as glucogenic) or NAD^+.

18.5.2.2: ALANINE TO PYRUVATE

As described in 18.2, and shown in Figure 18.5.4, the PLP-dependent enzyme **AL**anine Amino **T**ransferase (**ALT**), also known as **G**lutamate **P**yruvate **T**ransaminase (**GPT**), catalyzes this simple transamination reaction:

alanine +α−ketoglutarate ↔ pyruvate + glutamate

B. Alanine to Pyruvate

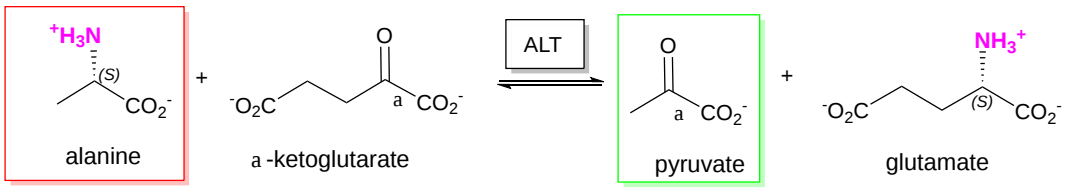


Figure 18.5.4: Alanine to pyruvate

The glutamate produced in this reaction can be oxidatively deaminated to give NH_4^+ and α-ketoglutarate again, giving the net reaction:

18.5.2.3: THREONINE TO GLYCINE

There are several pathways for this conversion.

One involves the conversion of Thr to 2-amino-3-ketobutyrate by **threonine-3-dehydrogenase**.

$$\text{Rx: Thr} + NAD^+ \leftrightarrow \text{2-amino-3-ketobutyrate} + NADH$$

This is followed by the conversion of 2-amino-3-ketobutyrate to glycine by the enzyme **2-amino-3-ketobutyrate coenzyme A ligase.**

$$\text{Rx: 2-amino-3-ketobutyrate} + CoASH \leftrightarrow \text{Gly} + \text{acetyl-CoA}$$

The net of these reactions is

$$\text{Rx: Thr} + NAD^+ + CoASH \leftrightarrow \text{Gly} + \text{acetyl-CoA} + NADH$$

These reactions are illustrated in Figure 18.5.5.

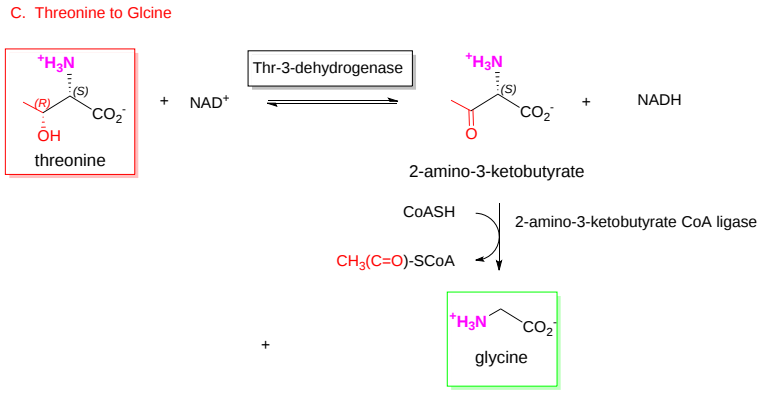


Figure 18.5.5: Threonine to glycine

A second and predominate reaction involves the conversion of Thr to NH_4^+ and α-ketobutyrate by the PLP-dependent enzyme **Ser/Thr dehydratase** (also called threonine ammonia-lyase), an enzyme we have seen in the previous section. Note this reaction does NOT produce glycine but is an intermediate, α-ketobutyrate.

$$\text{Rx: Thr} \leftrightarrow NH_4^+ + \text{α-ketobutyrate}$$

α-ketobutyrate can then be converted to propionyl-CoA.

$$\text{Rx: α-ketobutyrate} + NAD^+ + CoASH \leftrightarrow \text{propionyl-CoA} + NADH + CO_2 + H^+$$

This reaction, catalyzed by the inner mitochondrial membrane branched-chain α-ketoacid dehydrogenase complex (BCKDC or BCKDH complex) is an oxidative decarboxylation reaction. BCKDC is a member of two other enzymes, pyruvate dehydrogenase and alpha-ketoglutarate dehydrogenase, both of which act on short alpha-keto acids to produce key Kreb cycle metabolites.

Propionyl CoA is then converted eventually in several mitochondrial steps to succinyl CoA for entrance into the TCA cycle. Three enzymes are required for this conversion: propionyl CoA carboxylase, methylmalonyl-CoA epimerase, and methylmalonyl-CoA mutase. Propionyl carboxylase, like another alpha-keto acid carboxylase (pyruvate carboxylase), requires **A**TP, **B**iotin, and **C**O_2 (as a substrate) for the carboxylation reaction and hence is often referred to as an **ABC** enzyme.

The three-step conversion pathway of propionyl CoA to succinyl CoA is also used for in the degradation of **V**aline, **O**dd-chain fatty acids (which form multiple 2-carbon acetyl CoA units and 1 3-C propionyl CoA unit), **M**ethionine, and **I**soleucine along with **T**hreonine. This three-step pathway is sometimes referred to as VOMIT pathway.

The third pathway, which we just saw in the previous section, is catalyzed by serine hydroxymethyltransferase (SHMT) (but also called glycine hydroxymethyltransferase or threonine aldolase) and requires the use of both PLP and tetrahydrofolate as cofactors. A 1C methylene is added to tetrahydrofolate (FH_4). PLP makes bonds to the alpha-carbon of amino acids labile to cleavage. In this case, the amino acid threonine becomes dehydrated through an alpha-elimination reaction. However, threonine has an extra CH_3 group which is released as acetaldehyde. Here is the overall reaction.

$$\text{Rx: Thr+ } FH_4 + \leftrightarrow \text{ Glycine} + N^5,N^{10}\text{-}FH_4 + \text{acetaldehyde} + H_2O$$

The enzymes involved in this reaction are a bit unclear in the literature. It appears that SHMT can act on Thr at a lower rate, but that a second enzyme, threonine aldolase, which seems to be afunctional in mammals, acts in other organisms.

18.5.2.4: GLYCINE TO SERINE

As mentioned above, this reversible reaction is catalyzed by serine hydroxymethyltransferase (SHMT) (see the mechanism in section 18.4) and uses tetrahydrofolate and PLP as cofactors. Here is the overall reaction, the reverse of the Gly ↔ Ser we saw in 18.4.

$$\text{Rx: Glycine} + N^5,N^{10}\text{-}CH_2\text{-}FH_4 + H_2O \leftrightarrow \text{Serine} + FH_4$$

Figure 18.5.6 shows the serine dehydratase reaction presented in Chapter 18.4 Figure A below shows the dehydration reaction and formation of glycine. using PLP as a cofactor. Figure B shows how the released formaldehyde reacts with FH_4 to form N^5,N^{10}-methylene FH_4 using FH_4 as a cofactor.

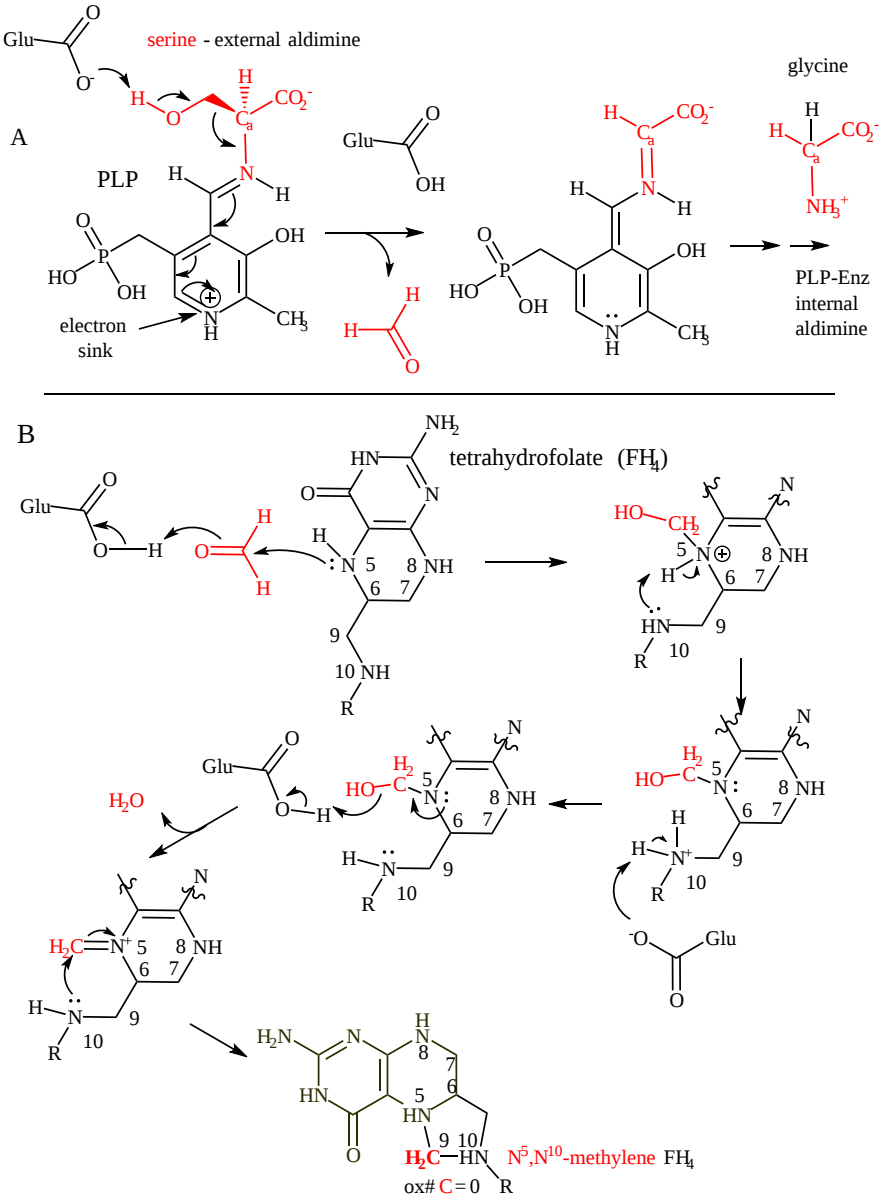


Figure 18.5.6: Reversible reaction of Serine to Glycine

18.5.2.5: SERINE TO PYRUVATE

This reaction is analogous to the Ala → Pyr reaction in Rx B above and is catalyzed by the PLP-dependent enzyme serine/threonine dehydratase/threonine deaminase.

Rx: Serine ↔ Pyr + NH_4^+

The enzyme is found in the cytoplasm and is mainly involved in gluconeogenesis.

18.5.2.6: CYSTEINE TO PYRUVATE

The overall reactions for this conversion are shown in the figure below. The aspartate aminotransferase used in the production of 3 sulfinylpyruvate is cytosolic and not the same as the more abundant version in the mitochondria.

The reaction pathway is shown in Figure 18.5.7.

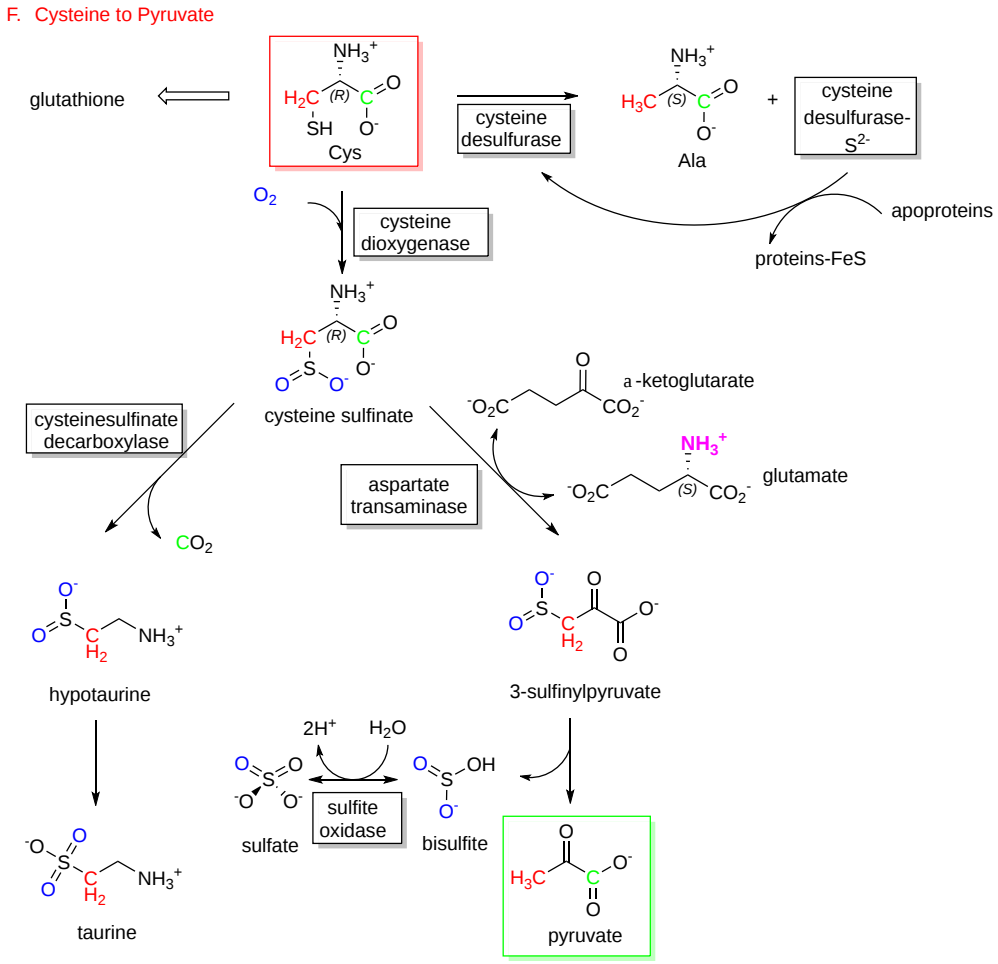


Figure 18.5.7: Cysteine to pyruvate

Other important metabolites are made from cysteine catabolic pathways. One is taurine, which is the most abundant free amino acid in the body and is especially abundant in development and early milk. It is synthesized predominately in the liver. It is unclear if hypotaurine is converted to taurine in a non-enzymatic fashion or by an oxidase/dehydrogenase.

The sulfate produced in the pathways is used to make an interesting derivative of ATP, 3′-phosphoadenosine-5′-phosphosulfate (PAPS), which is used to produce sulfated sugars using in glycolipid and proteoglycan synthesis. This is illustrated in Figure 18.5.8.

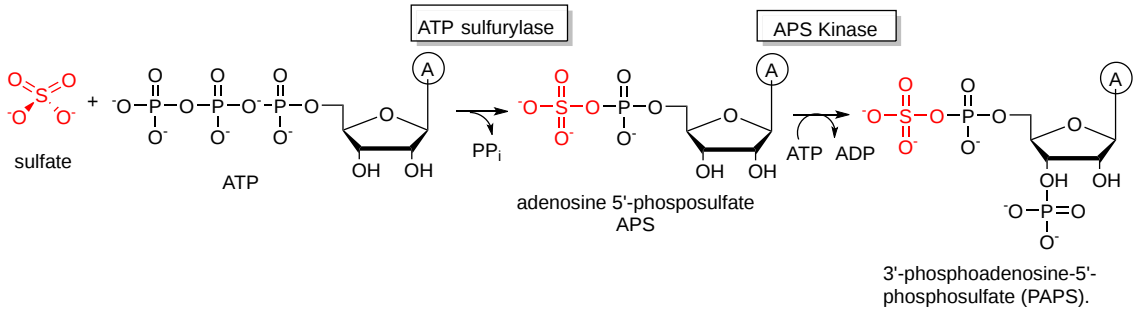


Figure 18.5.8: Sulfate conversion to PAPS

18.5.3: CONVERSION TO ACETYL-COA: TRP, LYS, PHE, TYR, LEU, ILE, THR

An overview of the many reactions in ketogenic amino acid degradation is shown in Figure 18.5.9. The red-boxed amino acids are those that form either acetoacetate (a ketone body) or acetyl-CoA directly (green boxes). Some of the carbons are color-coded red or green to indicate where they end up.

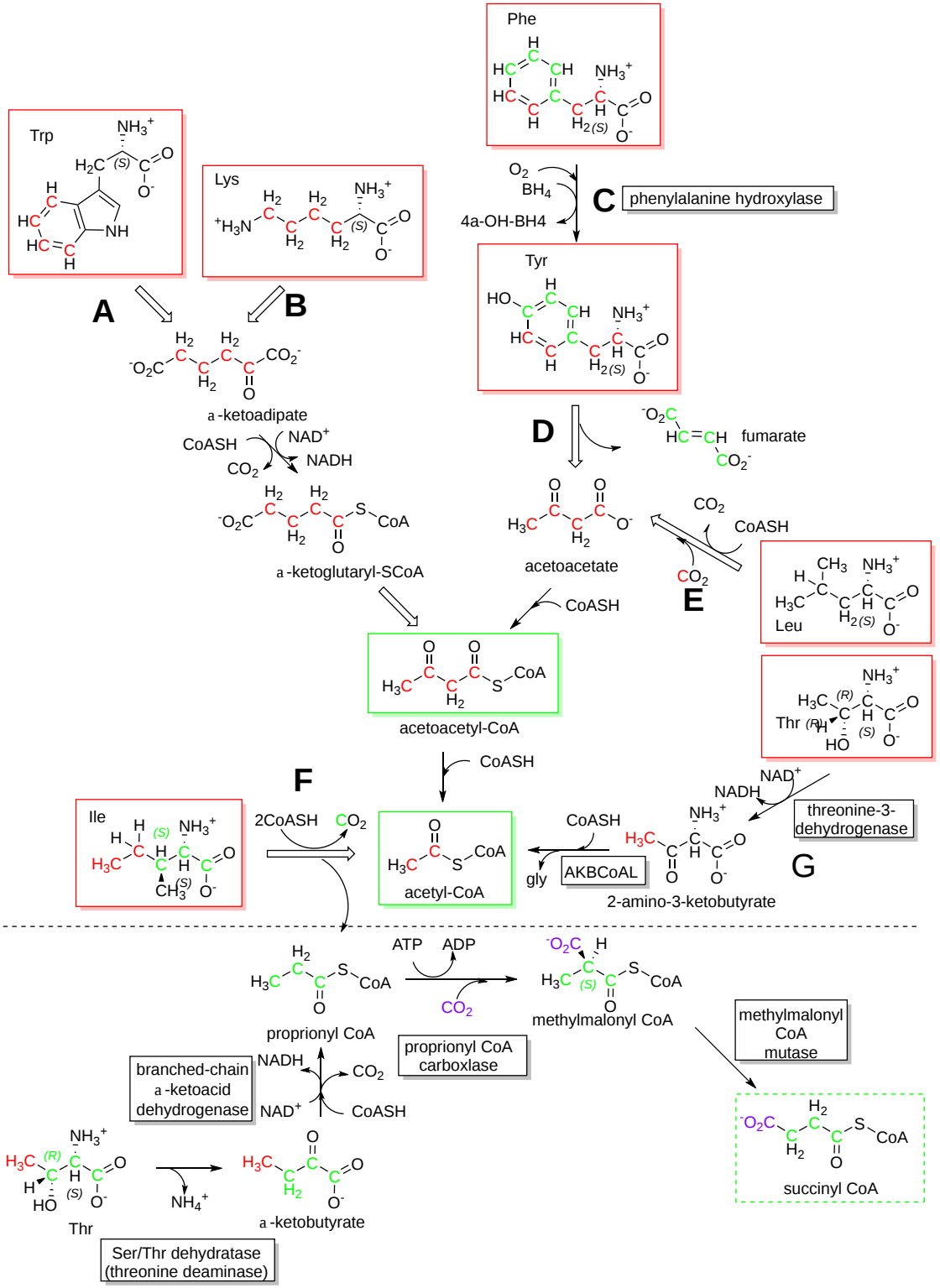


Figure 18.5.9: Ketogenic amino acid pathway

18.5.3.1: TRP TO ACETYL-COA

Fortunately, we have explored the conversion of the non-ring part of tryptophan to alanine and a precursor of acetoacetyl Coa (2-amino-3-carboxymuconate 6-semialdehyde - ACMS) and to NAD^+ (quinolinate). ACMS, through the action of ACMS decarboxylase leads to acetoacetyl CoA and then to acetyl-CoA as shown in Figure 18.5.10. As Trp is a ketogenic amino acid, it seems appropriate to show the steps that lead to acetyl-CoA even at the risk of providing too much detail.

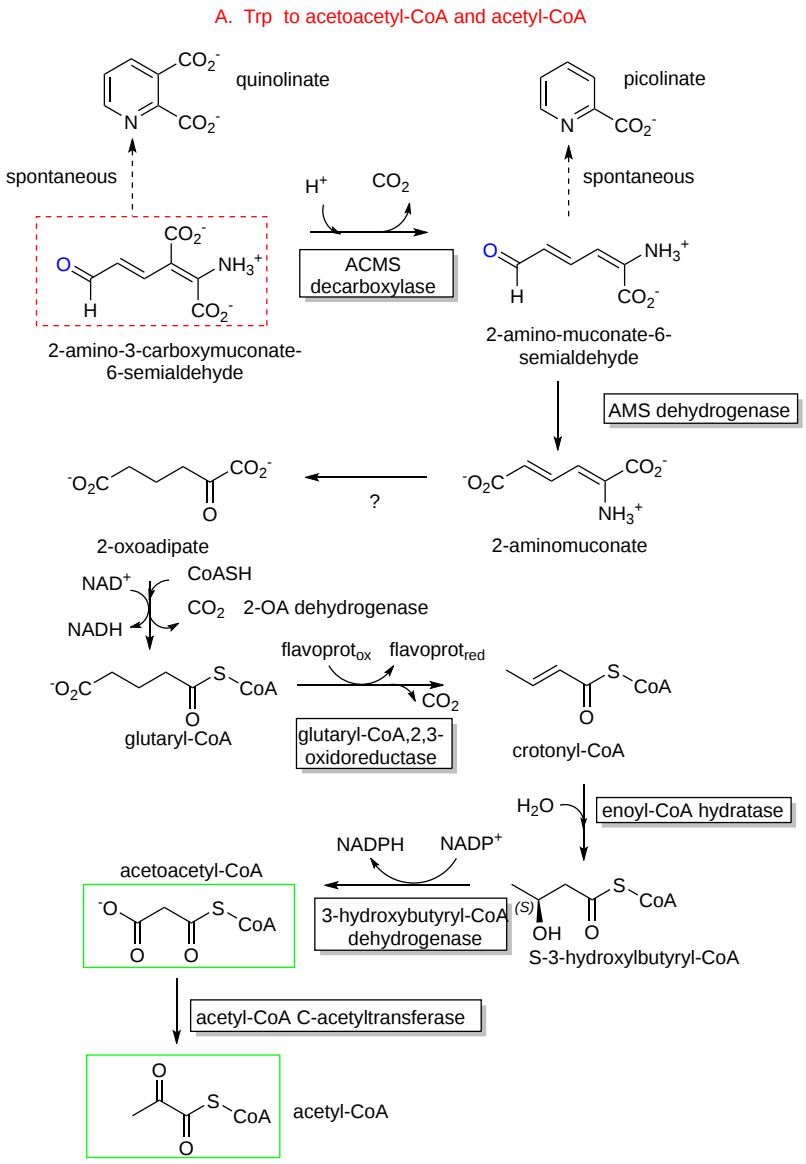


Figure 18.5.10: Part 2 - Tryptophan to acetyl-CoA

18.5.3.2: LYS METABOLISM

In the liver, the main pathway (of several) starts with the formation of saccharopine from the transamination reaction of lysine and α-ketoglutarate, allowing the ε-amino group of lysine to enter the nitrogen metabolic pool. This transamination does not use pyridoxal phosphate (PLP). The first two steps of the reaction are catalyzed by an enzyme, α-aminoadipic semialdehyde synthase, with two activities (condensation/reduction and hydrolysis/oxidation). Lysine is an essential amino acid since the transamination is not reversible. Figure 18.5.11 shows pathways for the conversion of lysine to acetoacetyl-CoA and acetyl-CoA.

B. Lys to acetoacetyl-CoA and acetyl-CoA

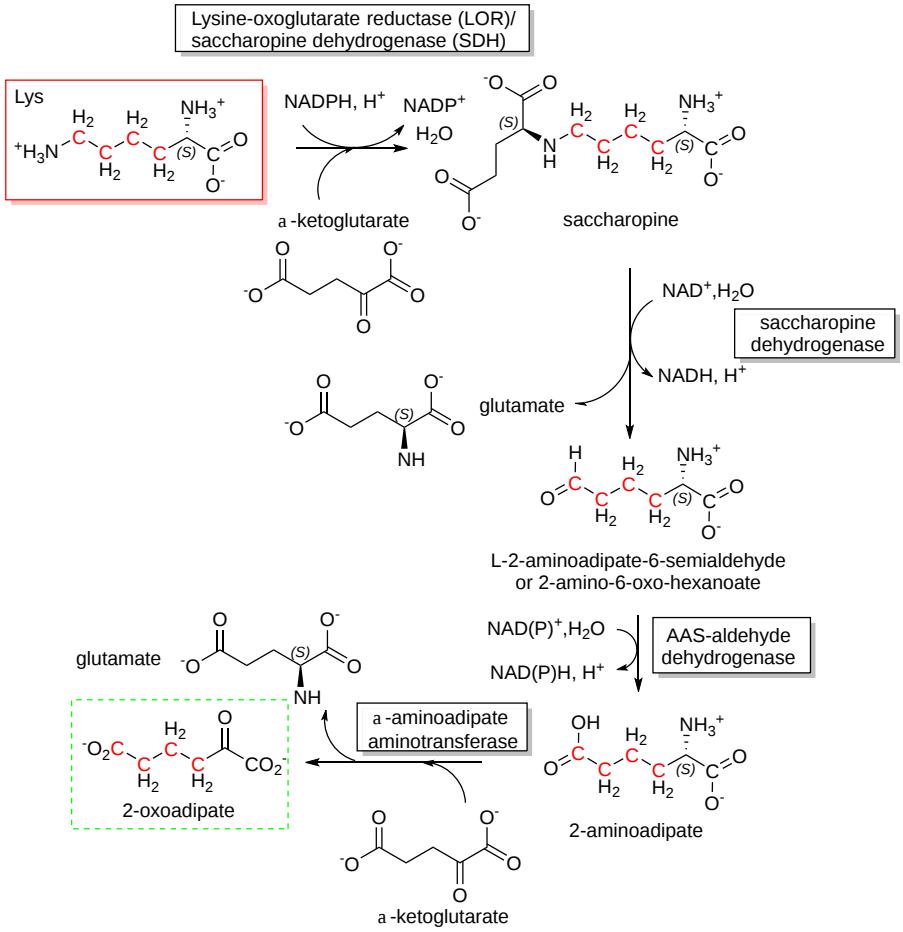
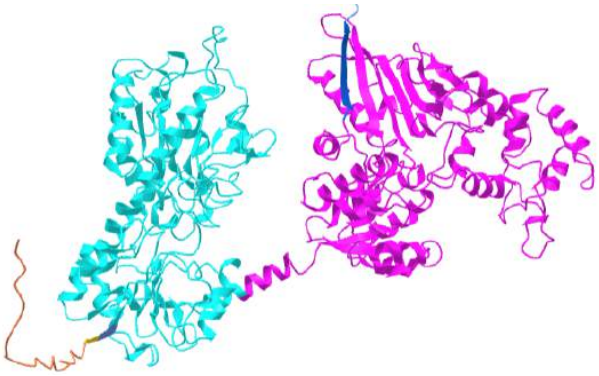


Figure 18.5.11: Pathways for conversion of lysine to acetoacetyl-CoA and acetyl-CoA.

The lysine-oxoglutarate reductase (LOR) and saccharopine dehydrogenase (SDH) are found in one bifunctional enzyme, often called aminoadipic semialdehyde synthase.

Figure 18.5.12 shows an interactive iCn3D model of the AlphaFold predicted structure of aminoadipic semialdehyde synthase (Q9UDR5)



NCBI iCn3D Figure 18.5.12: AlphaFold predicted structure of aminoadipic semialdehyde synthase (Q9UDR5). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...Pi1c7Y5A9DmLa8

The cyan domain represents the lysine-oxoglutarate reductase (LOR) domain. It is connected by a well-predicted alpha helix to the magenta saccharopine dehydrogenase (SDH) domain.

18.5.3.3: PHENYLALANINE CONVERSION TO TYROSINE AND CONTINUES TO ACETOACETATE

We'll follow the conversion of phenylalanine to tyrosine, which continues on to acetoacetate, making Phe and Tyr both ketogenic amino acids, and in subsequent steps that produce fumarate. They can enter the TCA cycle leading to the net production of oxaloacetate, which can be pulled off into gluconeogenesis, making both Phe and Try glucogenic as well.

A new cofactor that facilitates electron flow in the conversion of Phe to Try in the first step, catalyzed by the enzyme tyrosine hydrolase, is required. That cofactor is **tetrahydrobiopterin** (BH_4). The reaction involves the hydroxylation of BH_4 and then its transfer to phenylalanine. Figure 18.5.13 shows a possible mechanism for the conversion of phenylalanine to tyrosine with tetrahydrobiopterin (BH_4).

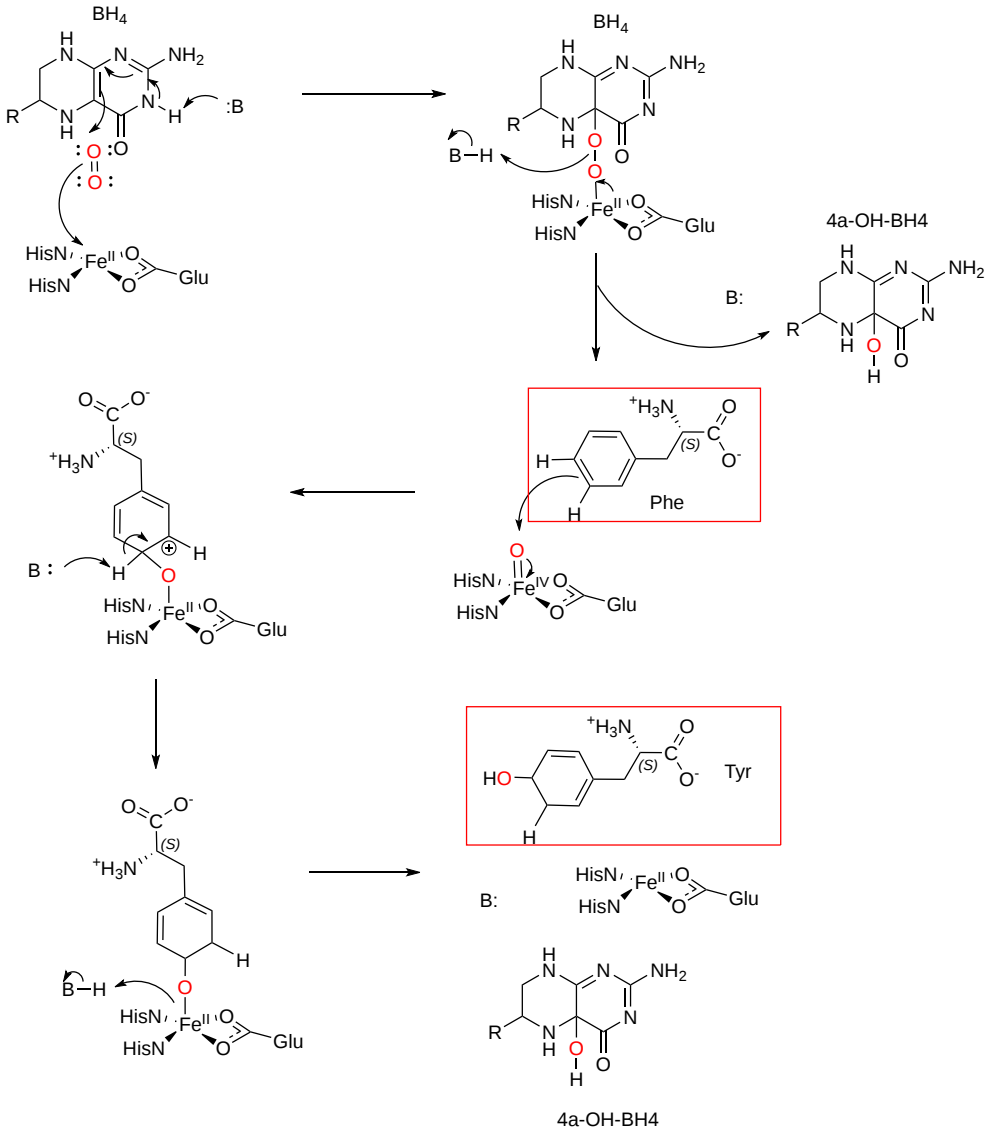


Figure 18.5.13: Conversion of phenylalanine to tyrosine with tetrahydrobiopterin (BH_4)

As in the case with the conversion of dihydrofolate back to tetrahydrofolate (FH_4) by dihydrofolate reductase, the 4a-OH-BH4 is converted to dihydrobiopterin and then to tetrahydrobiopterin by dihydrobiopterin reductase.

Figure 18.5.14 shows the full pathway for the conversion of Phe and Tyr to acetoacetate and fumarate.

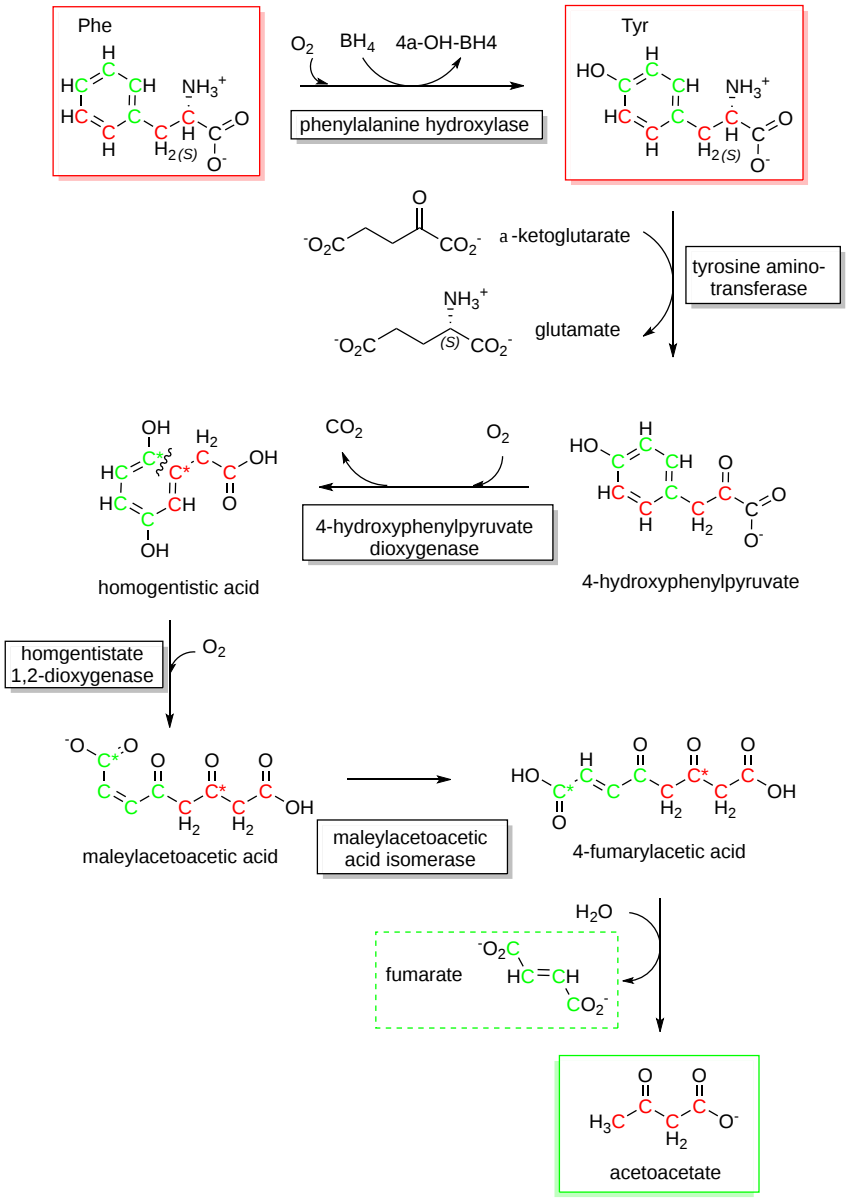


Figure 18.5.14: Conversion of phenylalanine and tyrosine to acetoacetate and fumarate

18.5.3.4: LEU TO ACETOACETATE

The conversion of leucine to acetoacetate is shown in Figure 18.5.15.

E: Leu to Acetoacetate

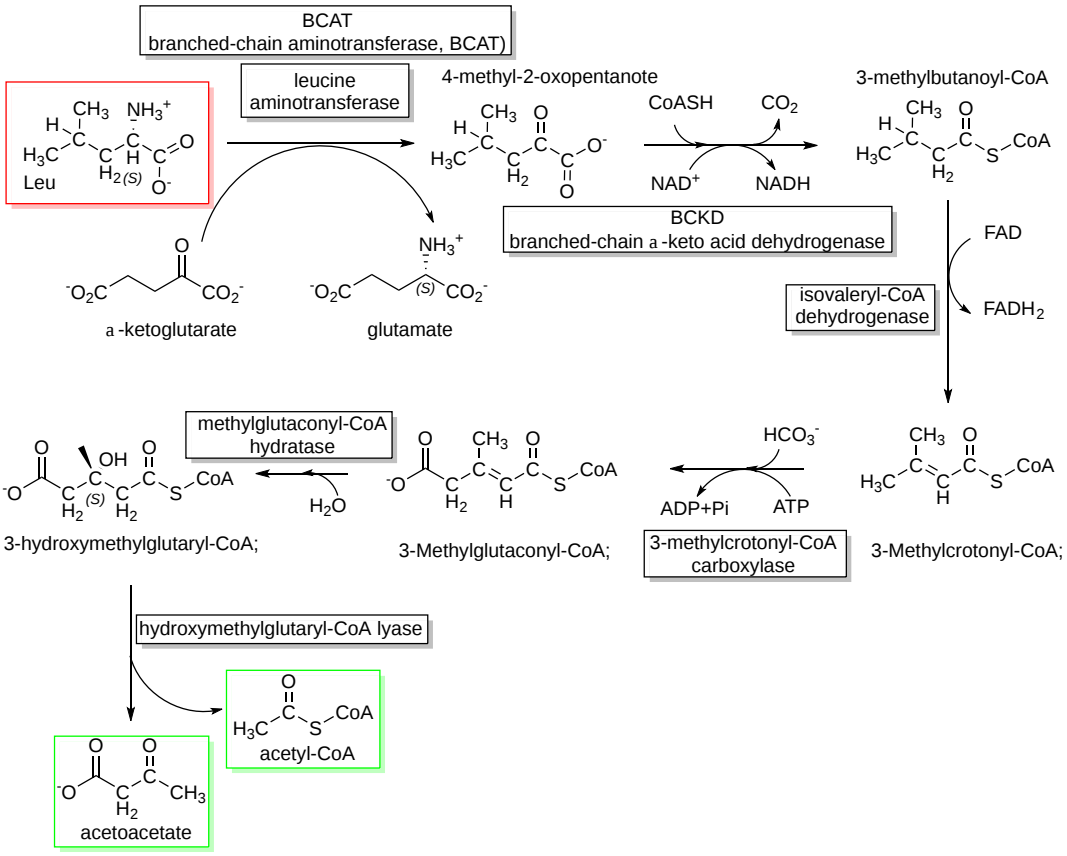


Figure 18.5.15: Conversion of leucine to acetoacetate

The first reaction is a transamination using the PLP-dependent branched-chain aminotransferase (BCAT) with α-ketoglutarate.

18.5.3.5: ISOLEUCINE TO ACETYL-COA

Figure 18.5.16 shows the pathway for the conversion of isoleucine to acetyl-CoA.

E: Ile to Acetyl-CoA

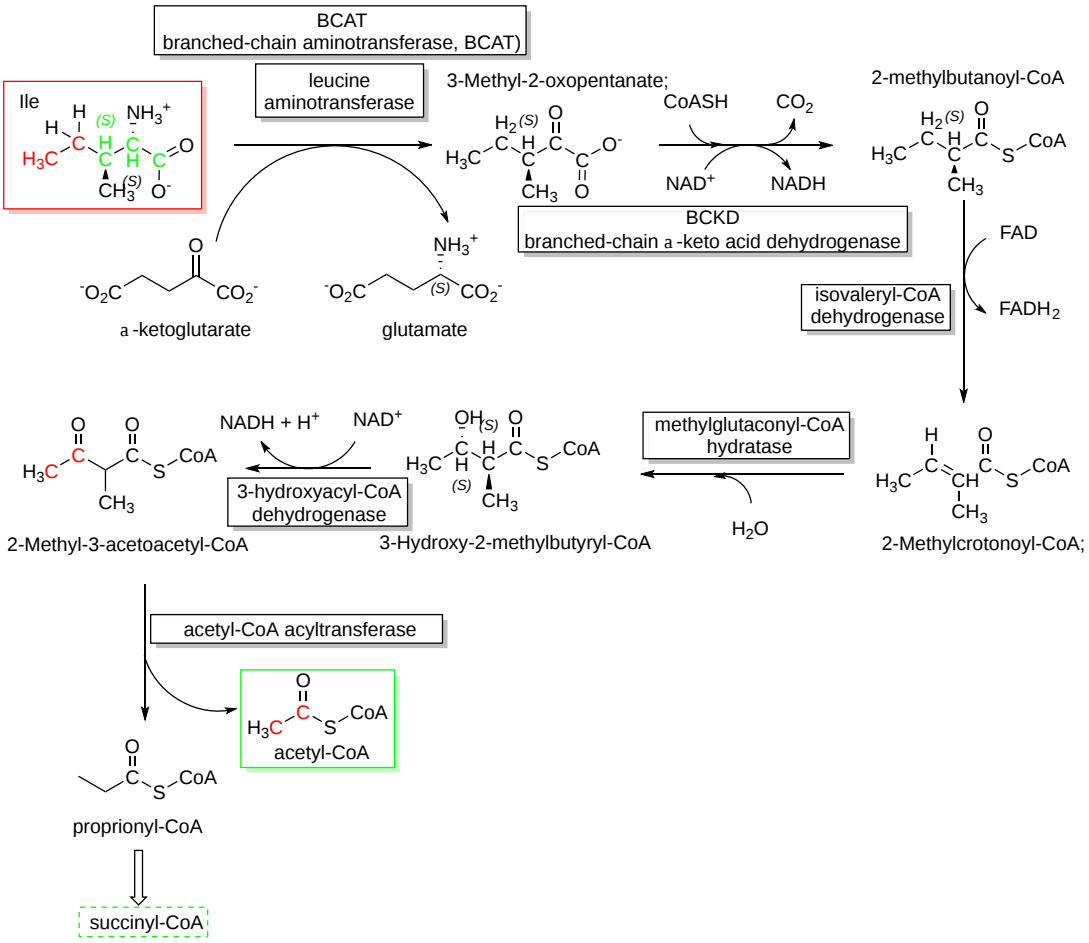


Figure 18.5.16: Pathway for conversion of isoleucine to acetyl-CoA.

18.5.4: CONVERSION TO A-KETOGLUTARATE: PRO, GLU, GLN, ARG,HIS

18.5.4.1: PROLINE AND ARGININE

The conversion of proline (bottom left) to glutamate (top left) is shown in Figure 18.5.17.

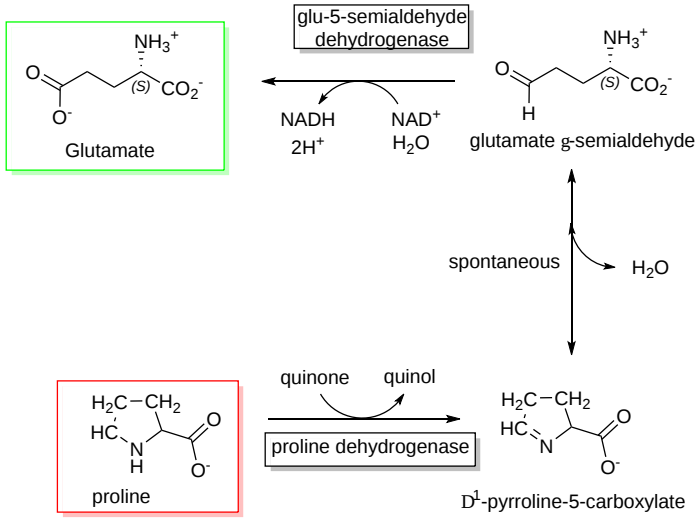


Figure 18.5.17: Conversion of proline (bottom left) to glutamate (top left)

Glutamate can then form α-ketoglutarate so the reaction is glucogenic.

The conversions of arginine (and proline) to α-ketoglutarate are shown in Figure 18.5.18.

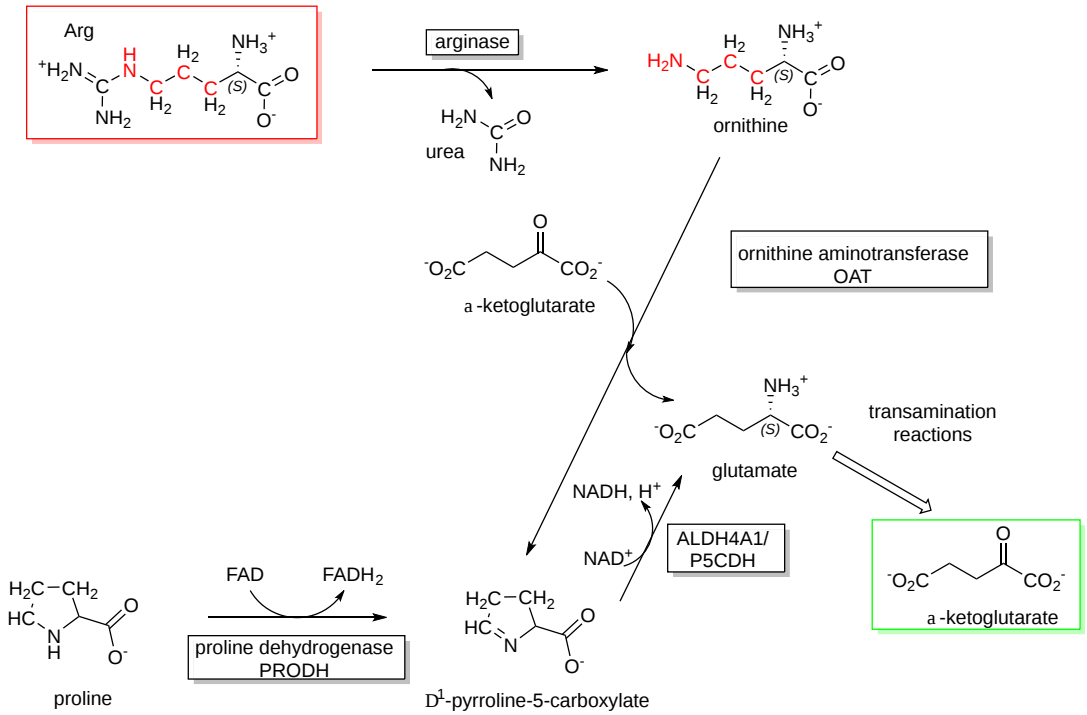


Figure 18.5.18: Conversion of arginine and proline to α-ketoglutarate

18.5.4.2: HISTIDINE

The conversion of histidine to α-ketoglutarate is shown in Figure 18.5.19.

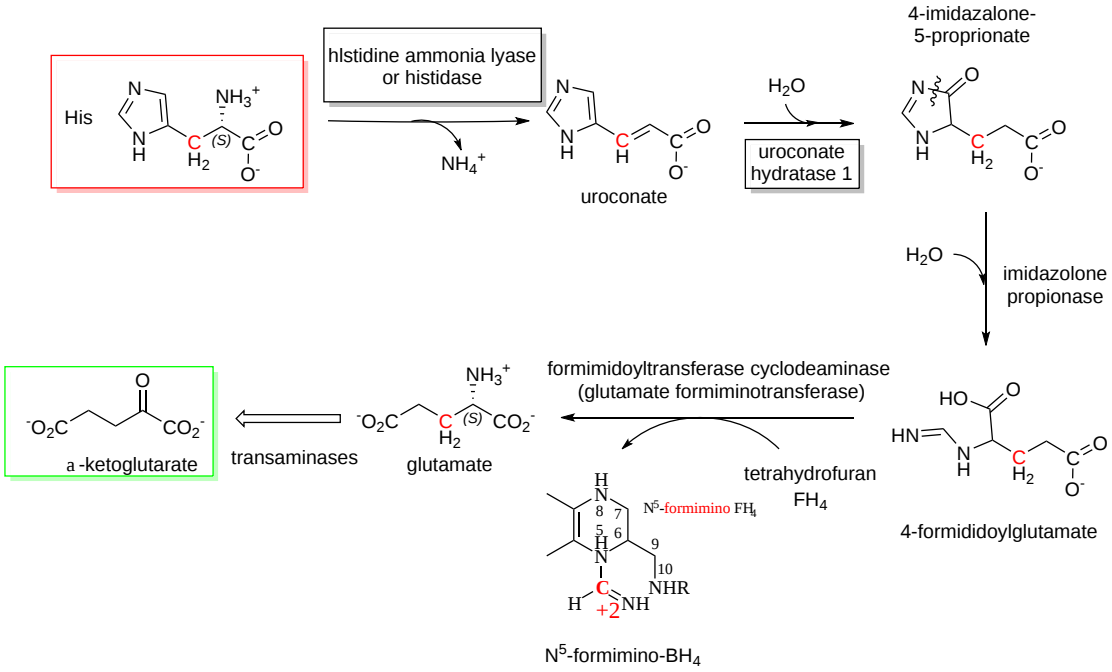


Figure 18.5.19: Conversion of histidine to α-ketoglutarate

As described in the reactions above, can be converted to α-ketoglutarate through transamination reactions. Also, we described in a previous section how glutamine can be deaminated through the action of glutaminase to form glutamine which can likewise form α-ketoglutarate, a gluconeogenic intermediate.

18.5.5: CONVERSION TO SUCCINYL-COA: MET AND THE BRANCHED-CHAIN AMINO ACIDS ILE, THR, VAL

We just saw that two branched-chain amino acids, Leu and Ile, are converted to acetyl-CoA and hence are ketogenic (E and F above). Another branched chain hydrophobic amino acid, Val, and also Leu again, can be converted to succinyl-CoA which can be converted to α-ketoglutarate in the Kreb's cycle in net fashion and hence are glucogenic amino acids. We saw in the introduction to amino acids that produce acetyl-CoA that threonine and isoleucine, two branched-chains amino acids, also form propionyl-CoA which goes on to succinyl CoA. So, let's consider Val, another branched-chain amino acid before we consider Met, both of which have 3 Cs in their side chains.

18.5.5.1: VALINE

The conversion of valine to succinyl-CoA is shown in Figure 18.5.20.

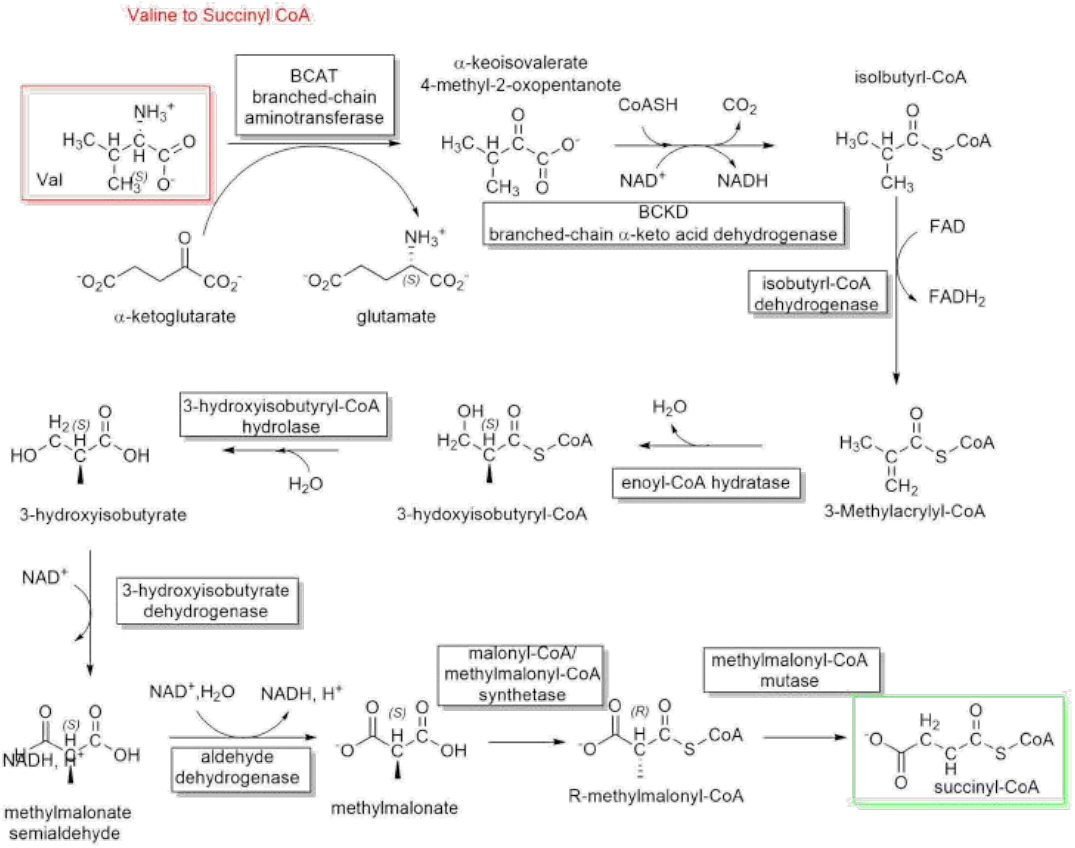


Figure 18.5.20: Conversion of valine to succinyl-CoA

The other two amino acids with branched-chain carbon chains (isoleucine and leucine), use the same enzymes as valine to enter the degradation pathway. They start with branched-chain transaminases (BCATc or BCATm) followed by oxidative decarboxylation reactions catalyzed by branched-chain ketoacid dehydrogenase (BCKD). Three different enzymes are required for the following dehydrogenase reaction. These are short/branched-chain acyl-CoA dehydrogenase (SBCAD) for isoleucine, isovaleryl-CoA dehydrogenase (IVD) for leucine, and isobutyryl-CoA dehydrogenase (IBD) for valine.

18.5.5.2: METHIONINE

Methionine can be metabolized to S-adenosylhomocysteine (SAM) and on to cysteine and α-ketobutyrate, which can also be produced by a transsulfuration reaction, the produces cysteine. That product is metabolized using branched chains dehydrogenases to eventually produce succinyl-CoA, a TCA intermediate. Three enzymes are needed to convert the α-ketobutyrate to succinyl-CoA. Propionyl-CoA carboxylase uses ATP, biotin, and CO_2 while the methylmalonyl-CoA mutase requires vitamin B12. An addition enzyme is an epimerase reaction in which D-methylmalonyl-CoA into L-methylmalonyl-CoA The conversion of propionyl-CoA to succinyl-CoA also occurs for branched-chain amino acids (**V**al, **I**le, **T**hr) as well as **M**et, and in addition **O**dd number fatty acids. This odd assortment of substrates for conversion to succinyl-CoA leads to the name **VOMIT** pathways. These reactions are illustrated in Figure 18.5.21.

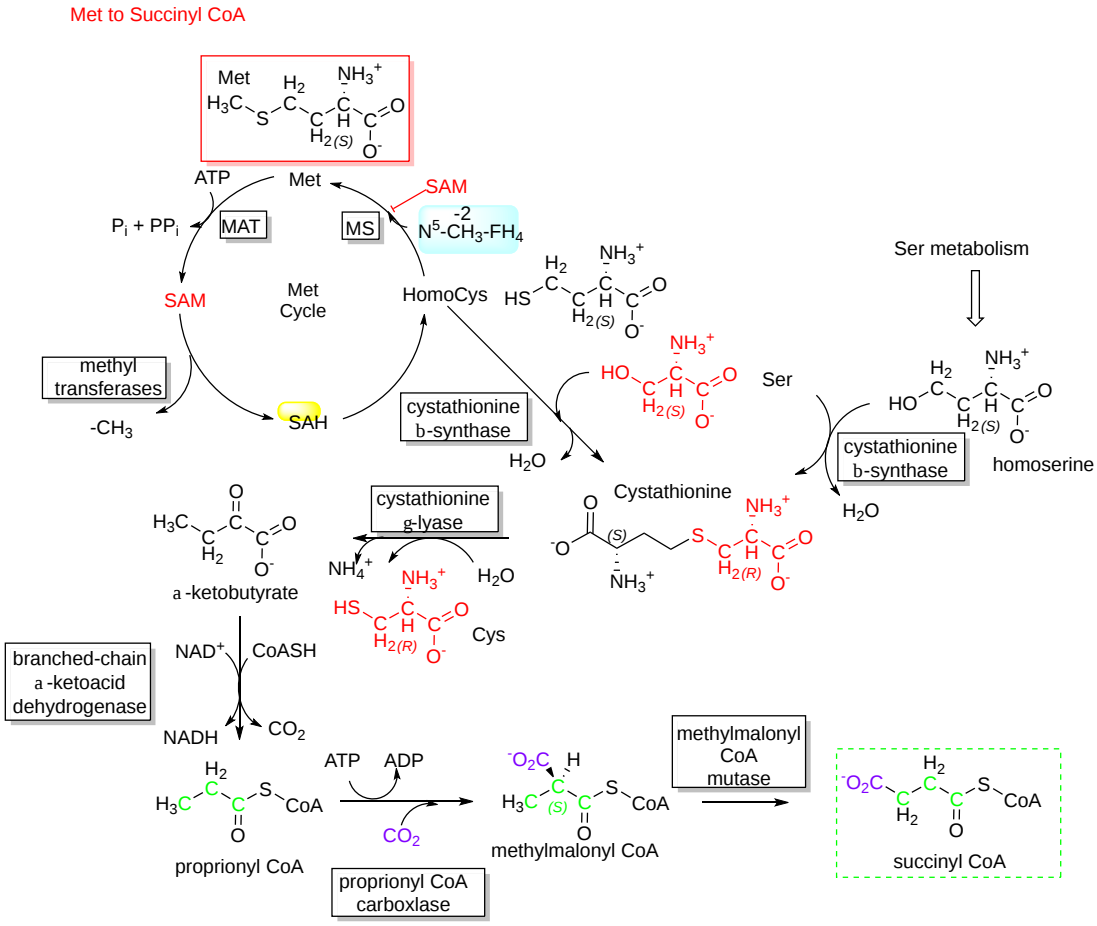


Figure 18.5.21: Methionine conversion to succinyl-CoA

18.5.6: FINALLY, ASPARTATE AND ASPARAGINE

Asparagine is converted to NH_3 and aspartate using the enzyme asparagine. Aspartate is then used in a transamination reaction to form oxaloacetate, a gluconeogenic precursor. Aspartate participates in the urea cycle as a way to eliminate nitrogen. Glutamate also acquires NH_3 through the reaction catalyzed by glutamine synthase.

Both amino acids are substrates for transamination reactions to produce TCA intermediates. Glutamate dehydrogenase can lead to alpha-ketoglutarate.

Glutamate and aspartate are important in collecting and eliminating amino nitrogen via glutamine synthetase and the urea cycle, respectively. The catabolic path of the carbon skeletons involves simple 1-step aminotransferase reactions that directly produce net quantities of a TCA cycle intermediate. The glutamate dehydrogenase reaction operating in the direction of 2-oxoglutarate (α-ketoglutarate) production provides a second avenue leading from glutamate to gluconeogenesis.

CHAPTER OVERVIEW

19: OXIDATIVE PHOSPHORYLATION

19.1: Electron-Transfer Reactions in Mitochondria

19.2: ATP Synthesis

19.3: Regulation of Oxidative Phosphorylation

19.1: ELECTRON-TRANSFER REACTIONS IN MITOCHONDRIA

Search Fundamentals of Biochemistry

19.1.1: AN OVERVIEW OF MITOCHONDRIAL ELECTRON TRANSPORT

The main oxidizing agent used during aerobic metabolism in the citric acid cycle is NAD^+ (although FAD is used in one step). In the process, these oxidizing agents get reduced to form NADH (and $FADH_2$). Unless NAD^+ is regenerated, glycolysis and the citric acid cycle will grind to a halt. This occurs under anaerobic conditions when NADH formed in glycolysis is reoxidized back to NAD^+ by pyruvate which is converted to lactate.

Under aerobic conditions, we are continually breathing one of the best oxidizing agents around, dioxygen. NADH is oxidized back to NAD^+ not directly by dioxygen, but indirectly as electrons flow from NADH through a series of electron carriers to dioxygen, which gets reduced to water. This process is called electron transport. No atoms of oxygen are incorporated into NADH or any intermediary electron carrier. Hence the enzyme involved in the terminal electron transport step, in which electrons pass to dioxygen, is an oxidase. The enzymes of the citric acid cycle and electron transport are localized in mitochondria, which is shown in Figure 19.1.1.

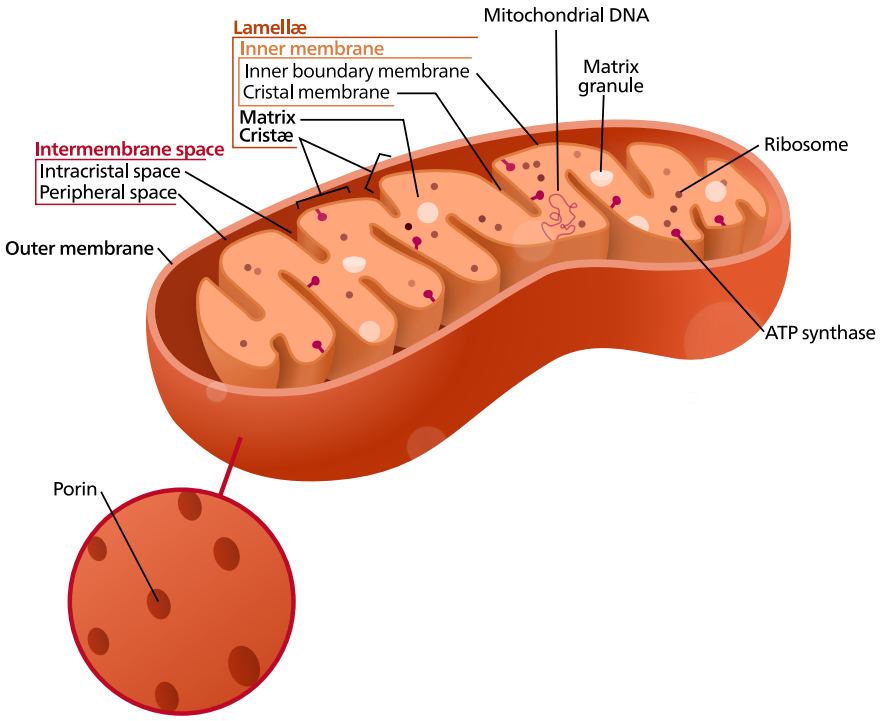


Figure 19.1.1: Mitochondrion. (Public Domain; Kelvinsong via Wikipedia)

By analogy to the coupling mechanism under anaerobic conditions, it would be useful from a biological perspective if this electron transport from NADH to dioxygen, a thermodynamically favorable reaction (ΔG of about -55 kcal/mol, -230 kJ/mol), were coupled to ATP synthesis. It is! For years scientists tried to find a molecule with "high energy" (with respect to its hydrolysis product) phosphorylated intermediate similar to 1,3-bisphosphoglycerate (formed by glyceraldehyde-3-phosphate dehydrogenase in glycolysis), which could drive ATP synthesis in the mitochondria. (**Remember, there is no such thing as a "high energy" bond.**) None could be found. A startling hypothesis was put forward by Peter Mitchell, which was proven correct and for which he was awarded the Nobel Prize in Chemistry in 1978. The immediate source of energy to drive ATP synthesis was shown to come not from a phosphorylated intermediate, but a proton gradient across the mitochondrial inner membrane. All the enzyme complexes in electron transport are localized in the inner membrane of the mitochondria, as opposed to the cytoplasmic enzymes of glycolysis. A pH gradient is formed across the inner membrane in respiring mitochondria. In electron transport, electrons are passed from mobile electron carriers through membrane complexes back to another mobile carrier.

There are four inner mitochondrial membrane complexes involved in the flow of electrons to dioxygen. The first, **Complex I**, passes two electrons from NADH, which engages like NAD^+ in 2 electron redox steps), to a flavin derivative, FMN, covalently attached to Complex I. The oxidation of NADH occurs, as expected, by hydride transfer. The reduced form of FMN then passes electrons in single electron steps (characteristic of FAD-like molecules, which can undergo 1 or 2 electrons transfers) through the complex to the lipophilic electron carrier, **ubiquinone, UQ** as shown in Figure 19.1.2.

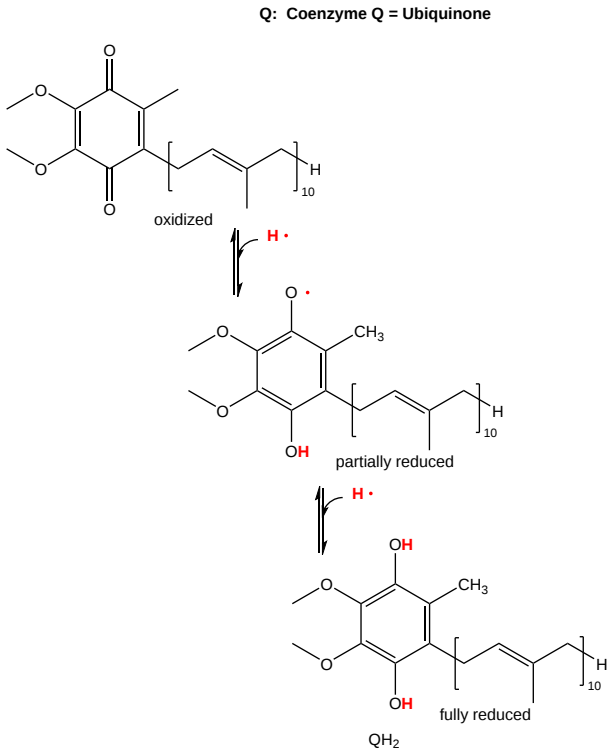


Figure 19.1.2: lipophilic electron carrier, ubiquinone, UQ (Copyright; author via source)

This then passes electrons through **Complex III** to another mobile electron carrier, a small protein, cytochrome C. This protein has a covalently attached heme with a central iron ion that can undergo one-electron redox reactions (Fe^{3+} + $1e^{-}$ ↔ Fe^{2+}). Reduce cytochrome C (Fe^{2+}) passes electrons through **Complex IV**, **cytochrome C oxidase**, to dioxygen to form water. At each step, electrons are passed to better and better oxidizing agents, as reflected in their increasing positive standard reduction potential. Hence the oxidation at each complex is thermodynamically favored. 4 electrons must be added to dioxygen, along with $4H^{+}$ to form water.

$$O_2 + 4\,e^- + 4\,H^+ \rightarrow 2\,H_2O.$$

What happened to Complex II? **Complex II** (also called succinate:quinone oxidoreductase) contains the citric cycle enzyme succinate dehydrogenase, which catalyzes the oxidation of succinate to fumarate by FAD. The bound $FADH_2$ does not dissociate from the enzyme, which is functionally dead for further formation of fumarate until it is oxidized back to FAD by additional components of Complex II. In this process, electrons from $FADH_2$ are passed to the same lipophilic mobile electron carrier from Complex I, ubiquinone. Reduced ubiquinone then can transfer electrons to cytochrome C through Complex III.

For Complex I, III, and IV, the energy released by the oxidative event is used to drive protons from the matrix to the intermembrane space of the mitochondria. The oxidative energy released is NOT used to form a "high-energy" mixed anhydride as we saw in the glyceraldehyde-3-phosphate dehydrogenase reaction from glycolysis. The resulting proton gradient collapses in a thermodynamically favorable process that energetically drives the endergonic synthesis of ATP by the last inner membrane complex, **ATP synthase,** which is sometimes called **Complex V**. This aerobic coupling of electron transfer (from NADH to dioxygen) and ATP synthesis, is shown schematically in Figure 19.1.3. The electron transfer is usually called electron transport, which might suggest to you that it requires energy as in the case of active transport. It does not, however, since it is strongly favored thermodynamically with dioxygen used as the final electron acceptor.

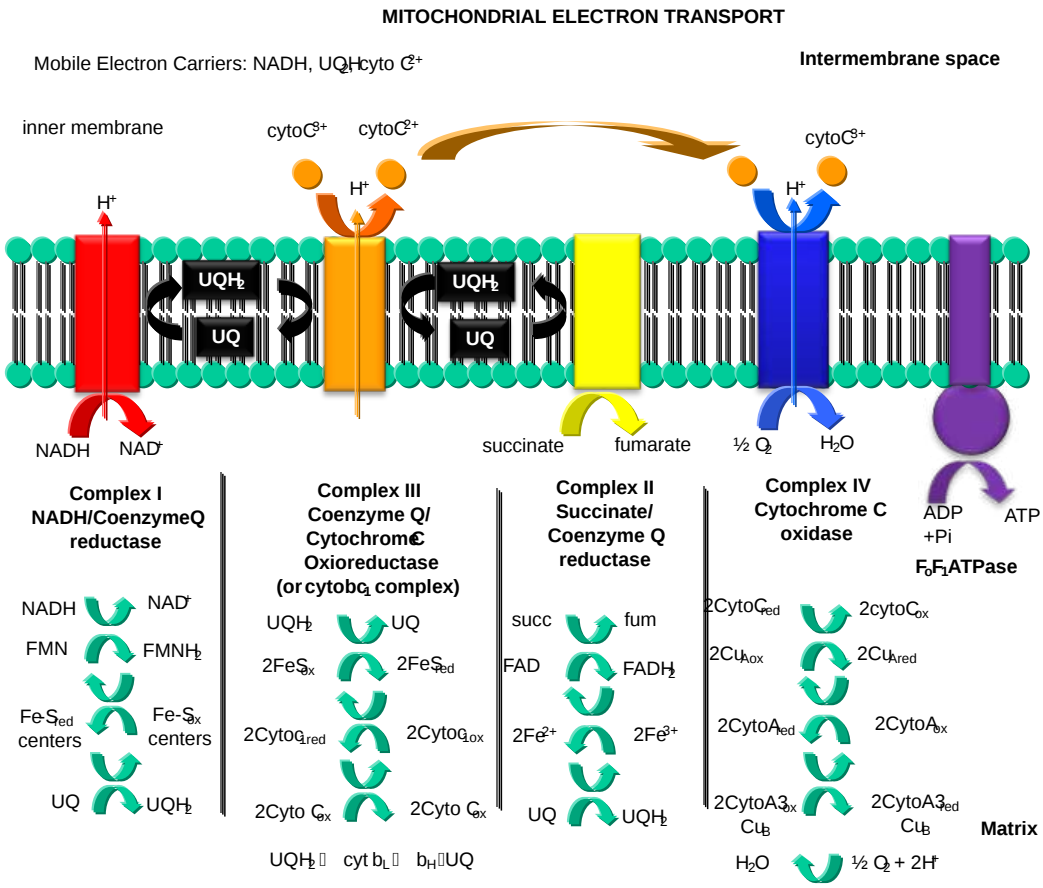


Figure 19.1.3: ELECTRON TRANSPORT AND PROTON GRADIENT FORMATION IN THE MITOCHONDRIA (after Voet and Voet)

You may wish to question the need for such a complicated series of reactions to move the simple electron from one chemical species to another and a simple H^+ from one side of a membrane to another to form a proton gradient that collapses anyway. The movers of these species are aligned in the membrane to optimize the processes and allow the vectorial, nonrandom flow of electrons and protons. This is what evolution has produced. One reason must be to maintain optimal control over the efficiency of the coupled reaction. Another powerful explanation is to reduce dangerous side reactions. In redox chemistry, this involves the formation of highly reactive free radicals that can react quickly and uncontrollably with any species nearby. This is especially true when dioxygen is involved, as we saw in Chapter 13.4 - The Chemistry and Biology of Dioxygen. It takes 4 electrons to fully reduce dioxygen to water. If this process is interrupted, highly reactive oxygen species (ROS) like superoxide and peroxide can form, which can damage membranes, proteins, and nucleic acids.

Before we explore each complex in more detail, let's get a lower resolution view of the entire mitochondrial electron transport and ATP synthesis (often called oxidative phosphorylation or ox-phos) which is shown in cartoon form in Figure 19.1.4.

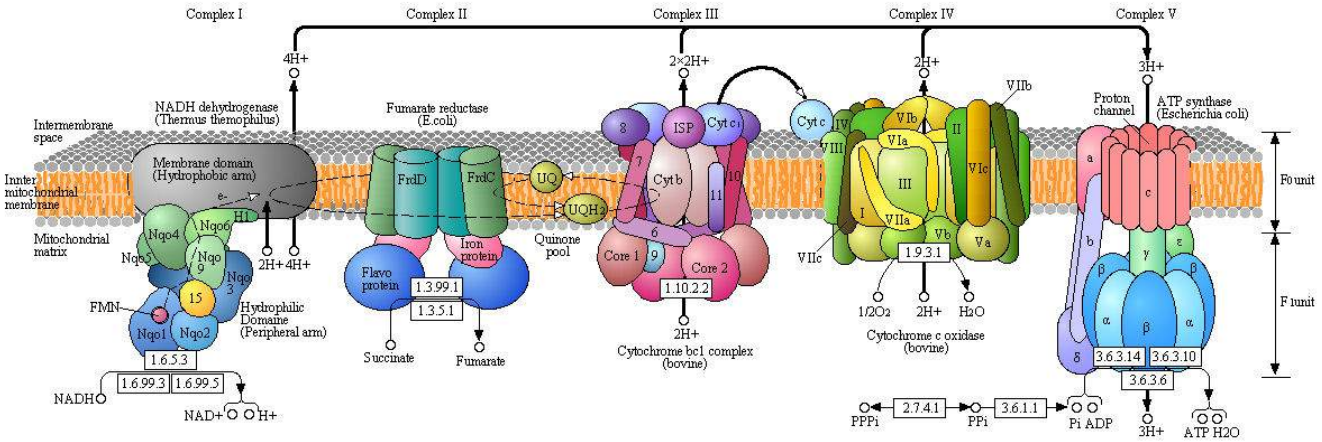


Figure 19.1.4: Overview of coupled electron transport and ATP synthesis (reprinted with permission from Kanehisa Laboratories and the KEGG project: www.kegg.org). Numbers represent Enzyme Commission Numbers. Original KEGG Map with embedded links.

Especially note:

- the subunit complexity of each Complex;
- the location of the subunits (many integral membrane proteins, others more hydrophilic and located in the polar matrix);
- the mobile electron carriers NADH/NAD^+, lipophilic UQH_2/UQ used for both complex I and II, cytochrome C (a protein), and O_2;
- the number of H^+s moved from the matrix to the intermembrane space through Complexes I, III, and IV and back through ATP synthase.

Also note that the prokaryotic structures are displayed, even though they are placed in the mitochondrial membrane! Bacteria do not have organelles like mitochondrial but they have the same complexes inserted into their plasma membranes. For bacteria, protons are ejected into the periplasm space between the plasma membrane and cell wall.

Diagrams like the one shown in Figure 19.1.4 are really useful in simplifying complex biological systems. At the same time, they constrain our ideas about how, in this case, the individual complexes are arranged in the membrane. Much evidence shows that Complex I, III, and IV (those that pump protons) form a very large supercomplex called the **respirasome,** which has some 80 subunits in mammals**.** Clustering the subunits in a supercomplex increases the likelihood that shared mobile electron carriers will not diffuse away, but rather continue passing electrons along the electron transport chain. As mentioned earlier, some reactions generate reactive oxygen species, which need to be kept in the complex to avoid their toxic effects. The stability of the individual complexes is also most likely increased in the larger aggregate.

Using cryoelectron microscopy (cryoEM), the structures of two large super- and mega-complexes have been solved showing the arrangement of all of the respiratory complexes. Two different structures showing the SuperComplex-$I_1III_2IV_1$ ($SCI_1III_2IV_1$) and MegaComplex-$I_2III_2IV_2$ ($MCI_2III_2IV_2$) are shown below in Figure 19.1.5.

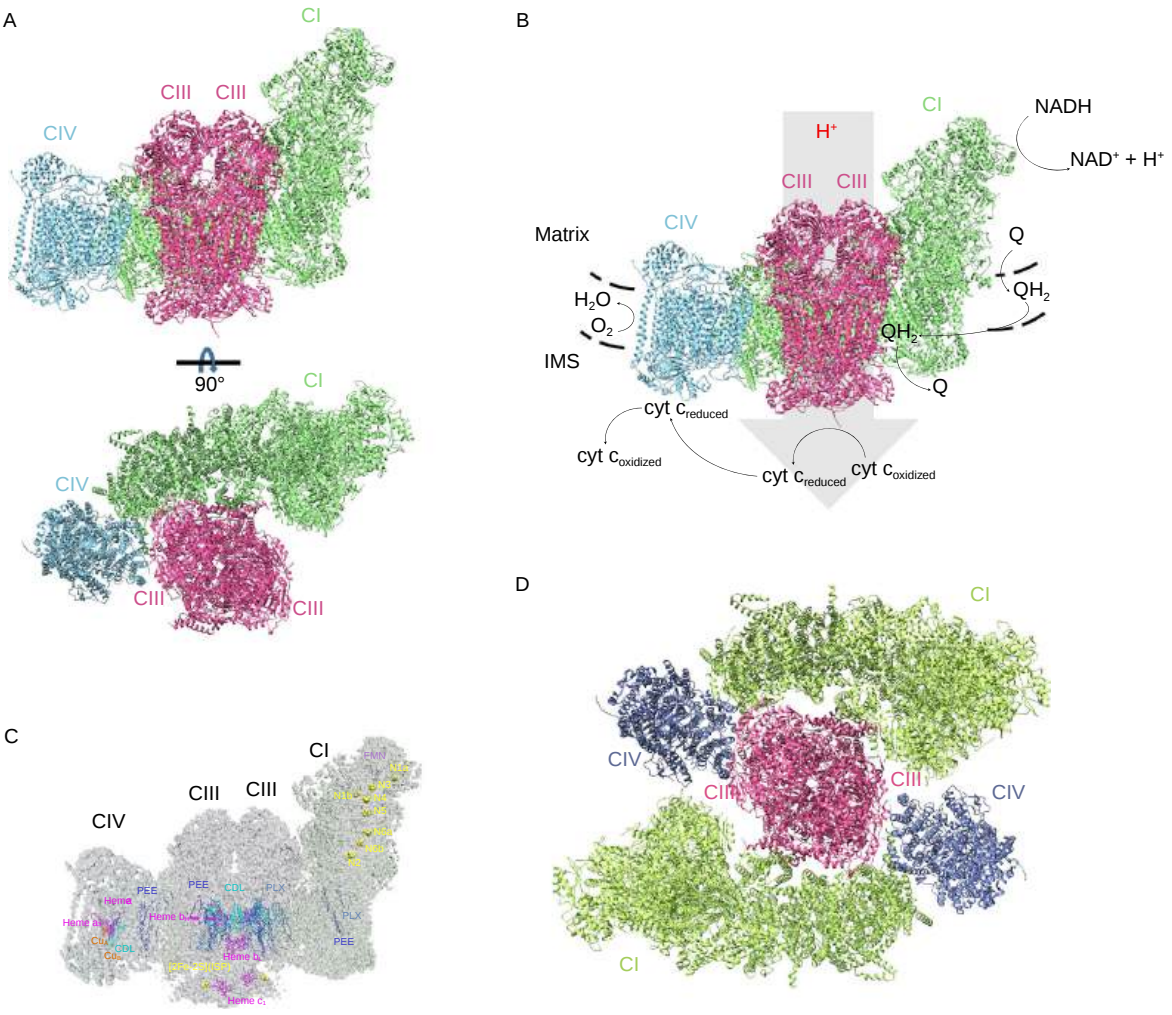


Figure 19.1.5: Structures of Mammalian SCI1III2IV1 and MC I2III2IV2. Meng Wu et al. Protein Cell 2020, 11(5):318–338. https://doi.org/10.1007/s13238-019-00681-x. Creative Commons Attribution 4.0 International License, https://creativecommons.org/licenses/by/4.0/.

Panel (A) shows human SCI1III2IV1 (PDB ID: 5XTH) in side view (along the inner membrane) and a top-down view. Note that CI is adjacent to CIII (located in the center) which is adjacent to CIV. This is the expected orientation given the flow of electrons from NADH to dioxygen. In panel (B) substrates and products are included into the side view, along with the flow of H^+ (from all of the complexes) into the matrix. Panel C shows the added cofactors (FeS clusters, FMN, hemes, etc). Panel (D) shows a very beautiful top-down view of the human megacomplex, which has two copies of each of subunit. Again note the centrality of Complex III.

Figure 19.1.6 shows another model of the megacomplex with Complex II (CII, which does not translocate protons across the membrane) added. It also shows a tetramer of Complex V (ATP synthase)

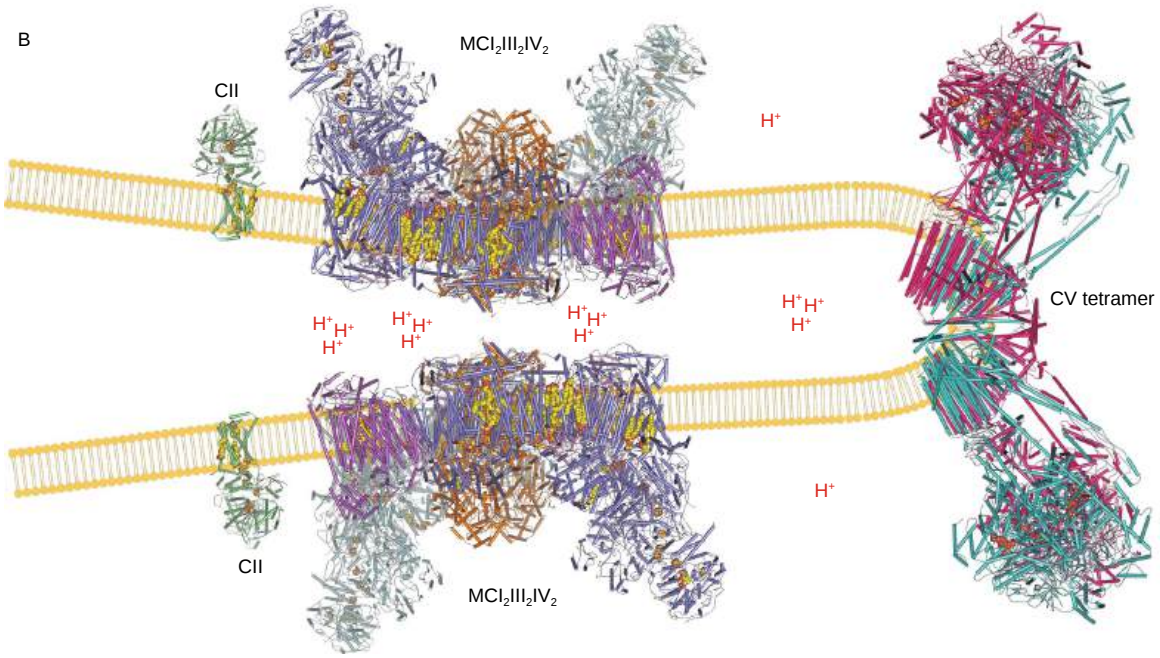


Figure 19.1.6: Arrangement of CII, MCI2III2IV2, and CV tetramer on mitochondrial cristae. (PDB ID of CII: 1ZOY. PDB ID of MCI2III2IV2: 5XTI. PDB ID of SCI1III2IV1: 5XTH. PDB ID of CV: 6J5K. Meng Wu et al. Protein Cell 2020, 11(5):318–338. https://doi.org/10.1007/s13238-019-00681-x. Creative Commons Attribution 4.0 International License, https://creativecommons.org/licenses/by/4.0/.

With this low-resolution background view of the entire mitochondrial oxidative phosphorylation pathway done, now let's explore each pathway in more detail.

19.1.2: COMPLEX I - NADH-QUINONE OXIDOREDUCTASE

Complex I is located in the inner mitochondrial membrane in eukaryotes and in the plasma membrane of bacteria. Bacteria have only 14 subunits while the mammalian Complex I has 45, including the core 14 found in bacteria. Another 31 are term **supernumerary** (in excess of the "normal number") subunits that support the function of the core. The significance of the extra mammalian subunits is still unclear. A cartoon model with the actual crystal structure of the Thermus thermophilus (bacterial) complex is shown in Figure 19.1.7. The hydrophilic or peripheral domain catalyzes electron transfer while the membrane domain (encoded by mitochondrial DNA) is involved in the active transport of protons. It is one of the biggest enzymes around, so it's exciting to try to understand it.

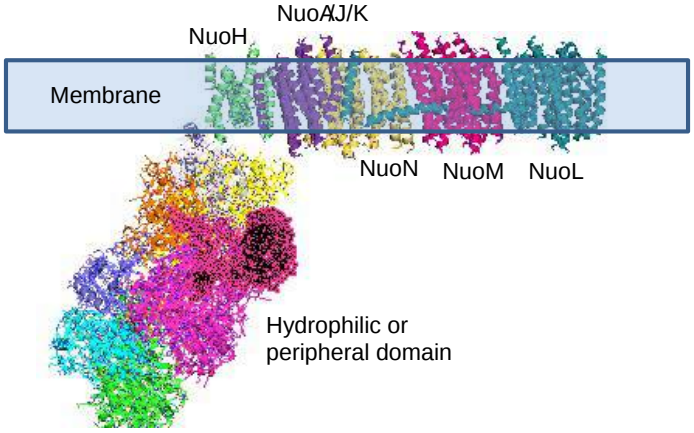


Figure 19.1.7: Detailed view of Complex I from Thermus Thermophilus (Copyright; author via source)

Figure 19.1.8 shows an interactive iCn3D model of the mammalian respiratory complex I (5LDW)

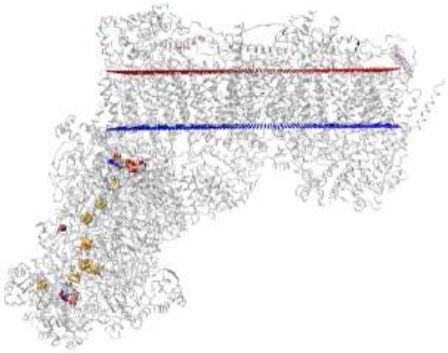


Figure 19.1.8: Mammalian respiratory complex I (5LDW). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...xVjRPLLgaLCi17 (both load slowly given the size of the structures). For the external link, use the mouse scroll to zoom the structure to see the FeS clusters and FMN, all of which are labeled.

The red and blue parallel "dummy" atoms represent the boundary of the mitochondrial inner membrane.

19.1.2.1: ELECTRON TRANSFER IN COMPLEX I

Electron flow occurs from NADH to UQ through a series of one-electron carriers in the hydrophilic or peripheral domain of Complex I. The initial handoff of electrons occurs to a flavin cofactor, FMN, and then through a series of Fe/S clusters, to ubiquinone. These electron acceptors include tetranuclear (Fe_4S_4) and binuclear (Fe_2S_2) iron-sulfur clusters, a FMN flavin mononucleotide, and a Mn (II) ion. These are shown in Figure 19.1.9.

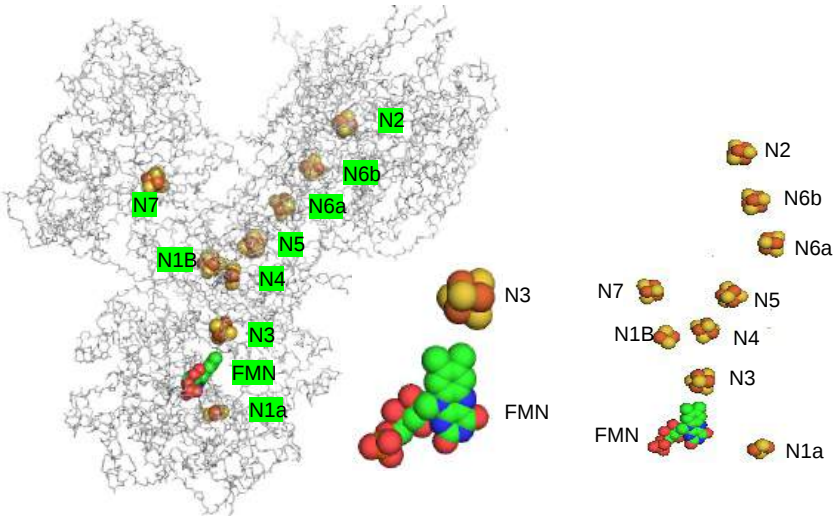


Figure 19.1.9: Electron Flow in Complex I from T. Thermophilus (Copyright; author via source)

The tetranuclear Fe_4S_4 cluster is based on the cubane structure with Fe and S occupying alternating corners of a square in a tetrahedral geometry. Each Fe is also coordinated to thiolate (RS^-) from coordinating cysteines and also sulfide (S^2)-anions. The actual structure is a distorted cube as shown in Figure 19.1.10, along with that of the binuclear Fe_2S_2 cluster, whose bond angles also deviate from those in a perfect tetrahedron.

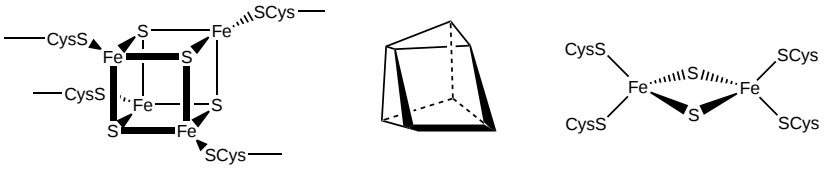


Figure 19.1.10: Fe/S Clusters in Complex I (Copyright; author via source)

Many possible micro-redox states with different standard reduction potentials are possible for tetranuclear Fe_4S_4 clusters, much as polyprotic acids have multiple pKa values. The two relevant to the tetranuclear clusters in Complex I are shown below:

a. $Fe(II)Fe(III)_3/S_2(CysS)_2^{1-} + e^- \leftrightarrow Fe(II)_2Fe(III)_2/S_2(CysS)_2^{2-}$ (less positive standard reduction potential)

b. $Fe(II)_2Fe(III)_2/S_2(CysS)_2^{2-} + e^- \leftrightarrow Fe(II)_3Fe(III)/S_2(CysS)_2^{3-}$ (more positive standard reduction potential)

PK$_A$S AND E^0 VALUES

You have studied Bronsted acids in many courses and know the factors that affect the pK_a of different acids that are **structurally similar**. These include the initial charge on a potential proton donor. The relative strength of a series of structurally similar acids can be predicted by examining the factors that stabilize the negative charge on the resulting conjugate base. The more stable the conjugate base, the more acidic the parent acid. Factors that stabilize a negative conjugate base are electronegativity of the atom holding the negative charge, resonance which might delocalize the charge, inductive/electron release effect of substituents, and the hybridization of the atom holding the charge (sp anions are more stable due to more s character with electrons pulled more closely to the nucleus).

What about a single acid? The pKa of given acid is a constant for a given set of conditions (temperature, solvent polarity, etc) but it can be altered by changing conditions. The diagram below in Figure 19.1.11 shows ways to "tip" the pKa of a given acid (in this case a simple carboxylic acid) to a lower value, making the acid stronger in the new set of conditions.

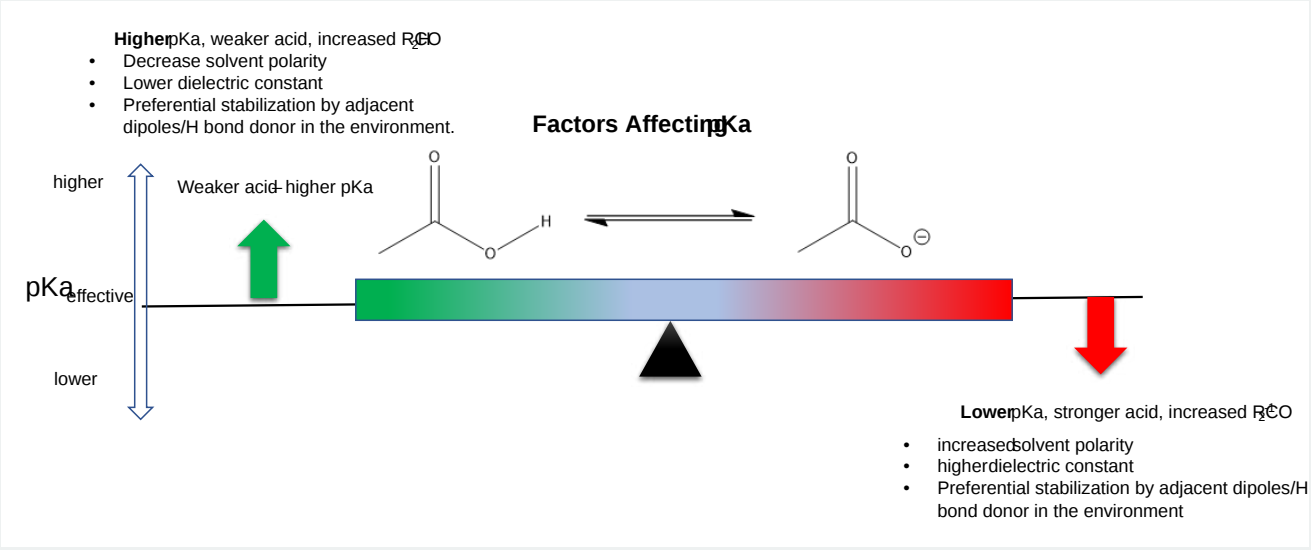


Figure 19.1.11: Factors affecting pKa of a given acid

The same ideas apply to the strength of a potential oxidizing agent (electron acceptor) as described by the E^0 value (higher for stronger oxidizing agents/electron acceptors). Figure 19.1.12 shows the factors that could alter the E^0 value for a binuclear FeS cluster. In this case, changes in coordinating ligands, analogous to changes in substituents for an acid) are included since most students have probably spent more time thinking about pKa values than E^0 values.

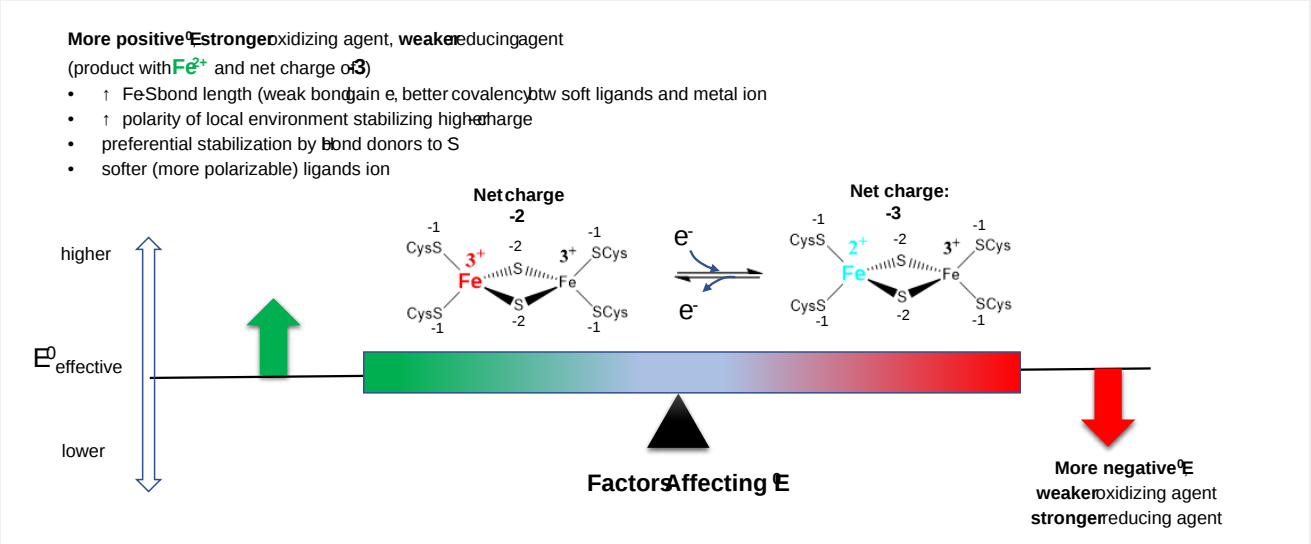


Figure 19.1.12: Factors affecting E^0.

In either case, it can be understood through Le Chatelier's principle. The equilibrium can be shifted to the right (stronger acids with lower pKa or stronger oxidizing agent/electron acceptor with a higher E^0) by preferentially stabilizing products compared to the initial

reactants.

Electrons are passed singly to oxidized UQ in one electron steps to form UQH_2.

Fe-S clusters are synthesized predominately in the mitochondria where they serve as redox cofactors in electron transport as described above. They are ubiquitous in all life forms and serve other roles in addition to redox cofactors per se as they serve structural roles in proteins and are used in redox signaling within the cell as they change oxidation states. Many proteins that interact with DNA (repair enzymes, polymerases, and helicases) contain an FeS cluster.

Evidence suggests that they placed critical roles in the abiotic evolution of life in the absence of oxygen as a terminal electron acceptor in exergonic oxidation reactions. When oxidation became available, they became potentially became toxic to the cell as the Fe^{2+} could participate in reactions (such as the Fenton reaction) leading to the generation of deleterious reactive oxygen species (such as superoxide). To prevent toxicity, when delivered to the cytoplasm and nucleus they are carried and delivered by cytoplasmic iron-sulfur assembly (CIA) proteins.

Figure 19.1.13 shows models of the complex and paths for electron flow.

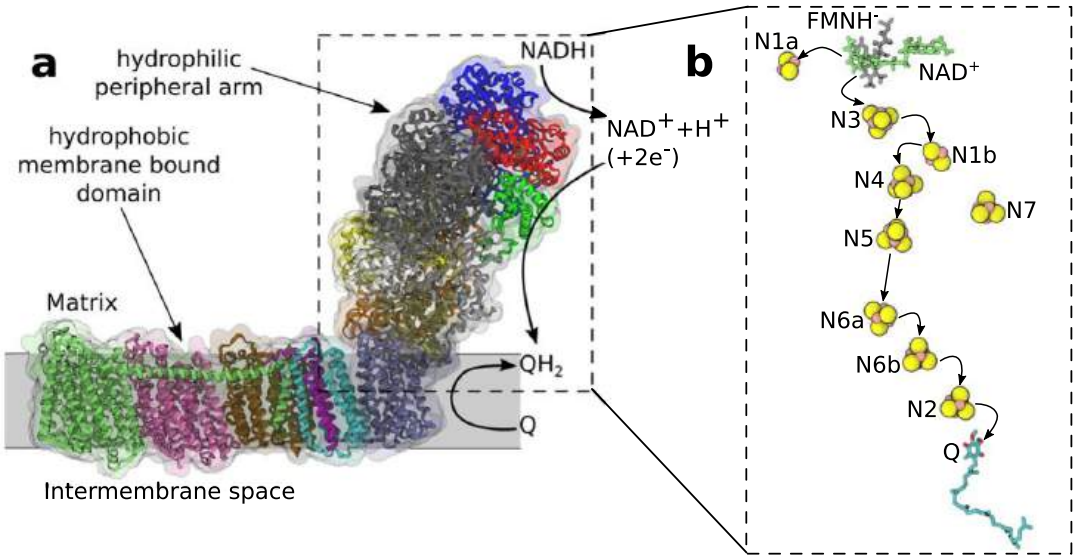


Figure 19.1.13: Complex I and electron transfer. Martin, D.R., Matyushov, D.V. Electron-transfer chain in respiratory complex I. *Sci Rep* **7,** 5495 (2017). https://doi.org/10.1038/s41598-017-05779-y. Creative Commons Attribution 4.0 International License, http://creativecommons.org/licenses/by/4.0/.

Electrons "tunnel" from one cofactor to another in Complex I through quantum mechanical "hops". Since all of the redox cofactors after the favored hand-off of electrons from NADH to FMN are FeS clusters, which would have similar E^0 values (modulated only by local environments), there is no significant thermodynamic barrier to electron flow. The final N2 FeS cluster has a E^0 value about 100 mV higher than the started ones so it most readily accepts an electron.

The separation of the hydrophilic mitochondrial matrix domains, where electron transfer occurs, from the transmembrane domains, where proton transfer occurs, is an elegant design that keeps the key and reactive particles, e^- and H^+, from reacting with each other. The electrons stay on one side of the membrane, as the protons move to the other side (until the return in the last step when Complex V produces ATP).

19.1.2.2: PROTON TRANSFER IN COMPLEX I

Proton transport occurs in the membrane domain. Available evidence suggests that 4 protons move from the cytoplasm to the periplasmic space against a concentration gradient during a catalytic cycle of Complex I in bacteria and other organisms.

Figure 19.1.14 shows multiple views of both the key electron transfer step and H^+ ejection from the matrix into the intermembrane space. The key to proton transports is the discrete membrane domains shown in the bottom left part of Panel (a). There are 4 such domains involved from left to right, each one associated with the transport of $1H^+$. The individual membrane domains have different names and abbreviations depending on the organism. They are from left to right:

- Nqo12, NuoL or ND5
- Nqo13, NuoM, or ND4
- Nquo14, NuoN or ND2
- Nqo8, NuoH, or ND1.

The Nqo8, NuoH, or ND1 domain is proximal and closely linked to the site for the reduction of UQ to UQH_2.

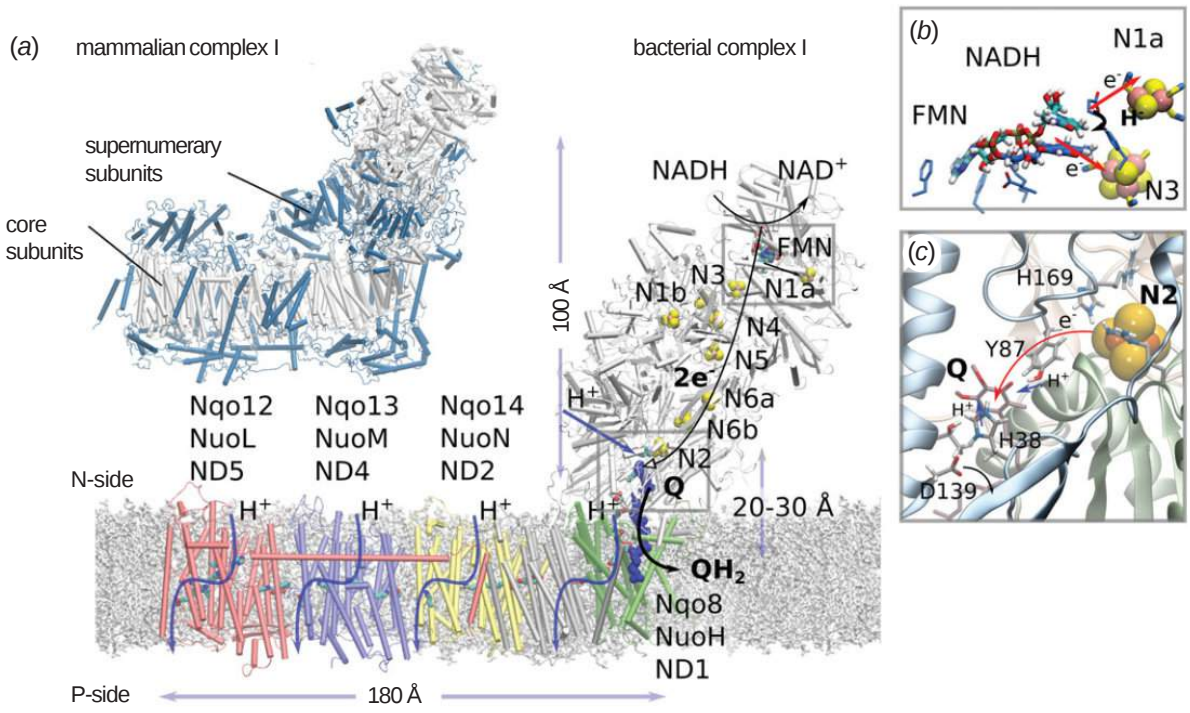


Figure 19.1.14: Structure and function of complex I. Panel (a) shows a large structural model of complex I from Thermus thermophilus (4HEA) showing the flow of electrons in the hydrophilic domain, while the mammalian complex (5LC5), with its supernumerary subunits in blue, is shown at the top left. Panel (b) shows the NADH/FMN site (3IAM) and its close proximity to the N3 tetranuclear cluster. Panel (c) shows the N2/Q reduction site with pi-stacking interactions of the ubiquinone ring with the aromatic His 38 ring. The P (positive side) accumulated H^+s while the N (negative side) becomes depleted in them. Kaila VRI. 2018 Long-range proton-coupled electron transfer in biological energy conversion: towards mechanistic understanding of respiratory complex I. J. R. Soc. Interface 15: 20170916. http://dx.doi.org/10.1098/rsif.2017.0916. Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/,""

The NuoN/Nqo14/ND2, NuoM/Nqo13/ND4 and NuoL/Nqo12/ND5 domains act as and are structurally similar to antiporters. Note also a very long helix that crosses the left three-most antiporter domains. This helix runs parallel to the membrane and probably acts as a lever allowing coupled conformational changes of the three H^+ transporters. We looked in a previous chapter at transporters and showed in both cartoon and structural models that they alternated between inward and outward-open conformations.

Figure 19.1.15 shows more details of the antiporter domains and the unique structural features allowing proton transport. Each antiporter domain has two 5-helix bundles that display pseudo-symmetry with transmembranes helices number TM4-8 and TM9-13. The next to last helix in each bundle (7 and 9) is interpreted, with unbroken parts of the helices number TM7a/b and TM12a/b. Running along the middle of each of the antiporter-like domains of the membrane domain is a series of conserved charged amino acids (aspartate, glutamate, lysine, and histidine) that participate in a hydrogen bonding network. Molecular dynamics simulations show water molecules (panel b), represented by red dots (for the oxygen atoms), transiently occupying a hydrogen-bonded network of waters that serve as a conduit for H^+ flow. An unbound "naked" proton (H^+) would not persist long in any aqueous biological system. Panel C shows the conformational flexibility in broken interrupted helices that presumably gates the proton pore open and closed.

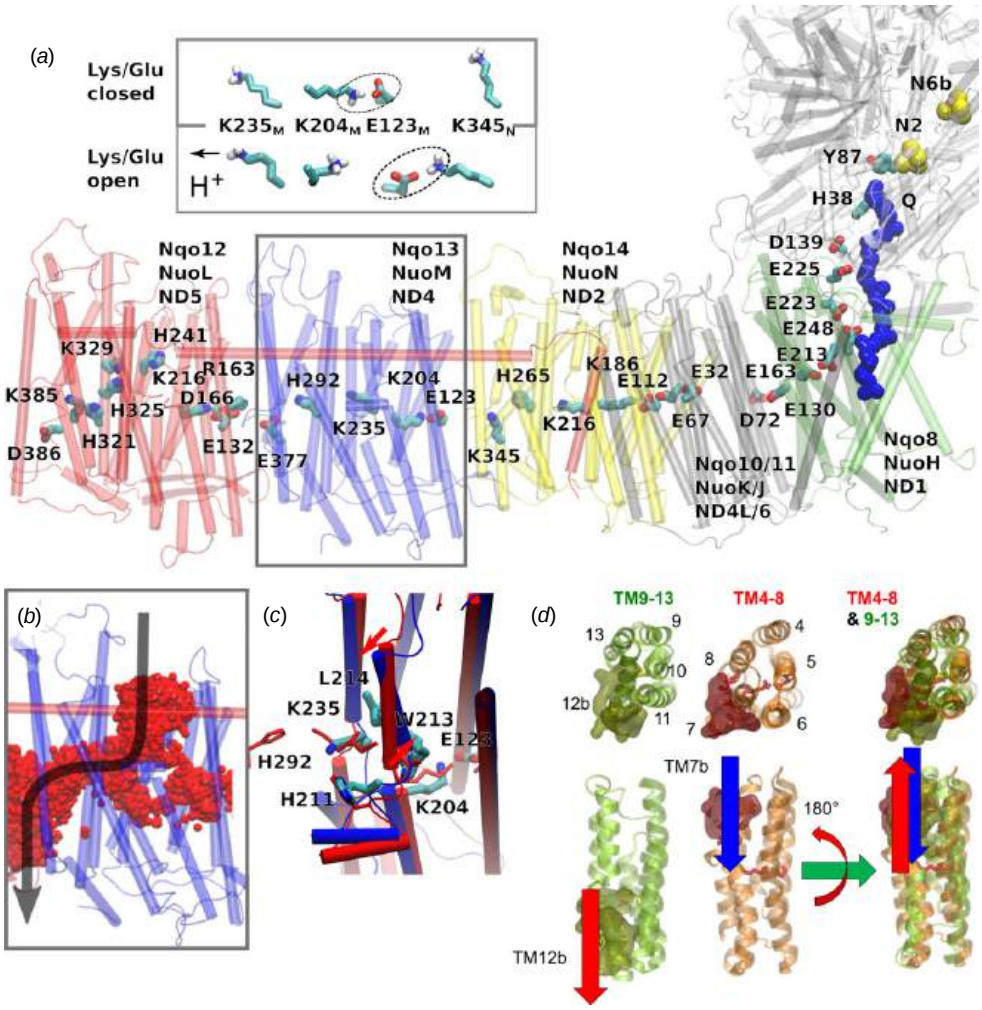


Figure 19.1.15: Structure of the proton-pumping membrane domain in complex I.

(a) The membrane domain of complex I have conserved buried charged/hydrophilic residues. NuoL (red); NuoM (blue); NuoN (yellow); NuoH (green); Ubiquone - Q (blue van der Waals representation). Inset: The conformation of the Lys/Glu ion pair (here Lys-204M/Glu-123M) can modulate the pKa of the middle Lys (Lys-235M).

(b) The NuoM antiporter-like subunit with water molecules (red dots). The arrow shows the flow of protons through a hydrogen bond-linked network of waters that "visit" the structure during μs molecular dynamics simulation.

(c) Snapshot of structures obtained from molecular dynamic simulations of open (in blue) and closed (in red) proton channels from the N-side in Nqo13 (NuoM/ND4), showing conformational changes in the broken helix element.

(d) The structural symmetry of the antiporter-like subunits with an N-side input channel near broken helix TM7b and output channel near broken helix TM12b.

How is the reduction of ubiquinone and proton transport coupled? The mechanism likely involves conformational changes in flexible regions in the interface of the hydrophilic and membrane domains of the complex.

19.1.2.3: INHIBITION OF COMPLEX I

Complex I is inhibited by more than 60 different families of compounds. They include the classic Complex I inhibitor rotenone and many other synthetic insecticides/acaricides. The classes include Class I/A (the prototype of which is Piericidin A), Class II/B (the prototype of which is Rotenone), and Class C (the prototype of which is Capsaicin). They appear to bind at the same site. Figure 19.1.16 shows their structures. From the structure of the 3 prototypes, what are the characteristics of the pharmacophore, the “ideal binding ligand”? Where do they likely bind? How “promiscuous” is the binding site?

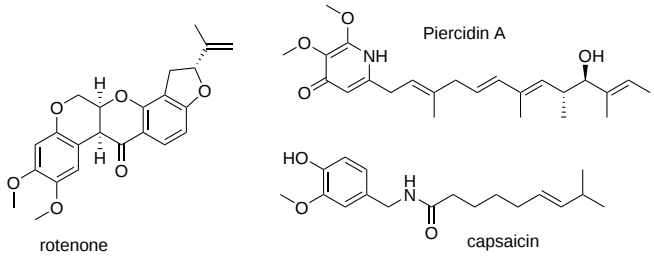


Figure 19.1.16: Inhibitors of Complex I (Copyright; author via source)

Many devastating neurological diseases are associated with defects in Complex I. In addition to major problems with oxidative ATP production, reactive oxygen species (ROS) increase. The major sites for the generation of ROS are Complex I and Complex III. Given the locations of the electron carriers at the periphery and internal within the protein complex, which electron carriers might most readily leak electrons to dioxygen? What ROS is likely to form in the process?

Inhibitors might block access to UQ or conformational changes necessary for the final reduction of the ubiquinone free radical. Class A inhibitors dramatically increase ROS production. The actual site of ROS production in Complex I is a bit controversial. One possible electron donor to dioxygen is FMN since both can engage in 1 e^- transfers. Mutants that lack N2 iron-sulfur cluster showed ROS production.

In submitochondrial preparations, normal Complex I activity occurs (which leads to the formation of a sustained proton gradient). Also reverse electron transport, powered by an artificial proton gradient can occur, which leads to the reduction of NAD^+, as shown in Figure 19.1.17.

Figure: Normal and Reverse Electron Transport Complex I

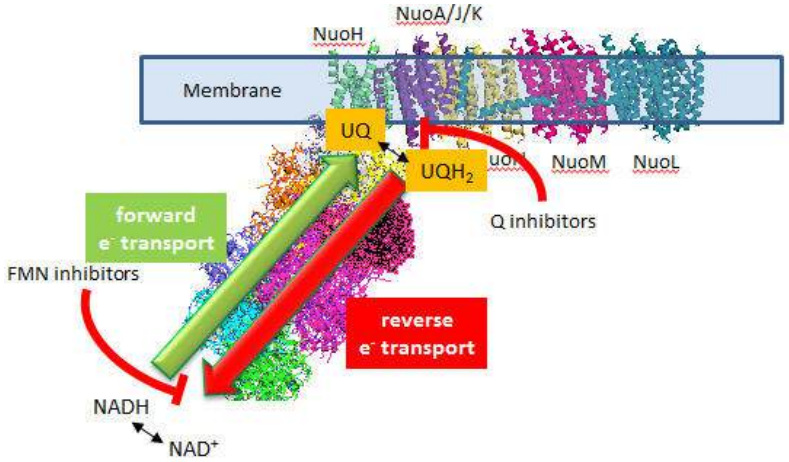


Figure 19.1.17: Overview of forward and reverse electron transport in Complex I.

A summary of the finding on superoxide production by Complex I show that:

- superoxide production is inhibited by flavin site inhibitors but not Q site inhibitors.
- reverse electron transport leads to NAD^+ and O_2 reduction
- Reverse electron transport superoxide production is inhibited by both flavin and Q site inhibitors

Figure 19.1.18 shows an expanded view of electron transport and sites where ROS production are enhanced. Refer to the legend for details.

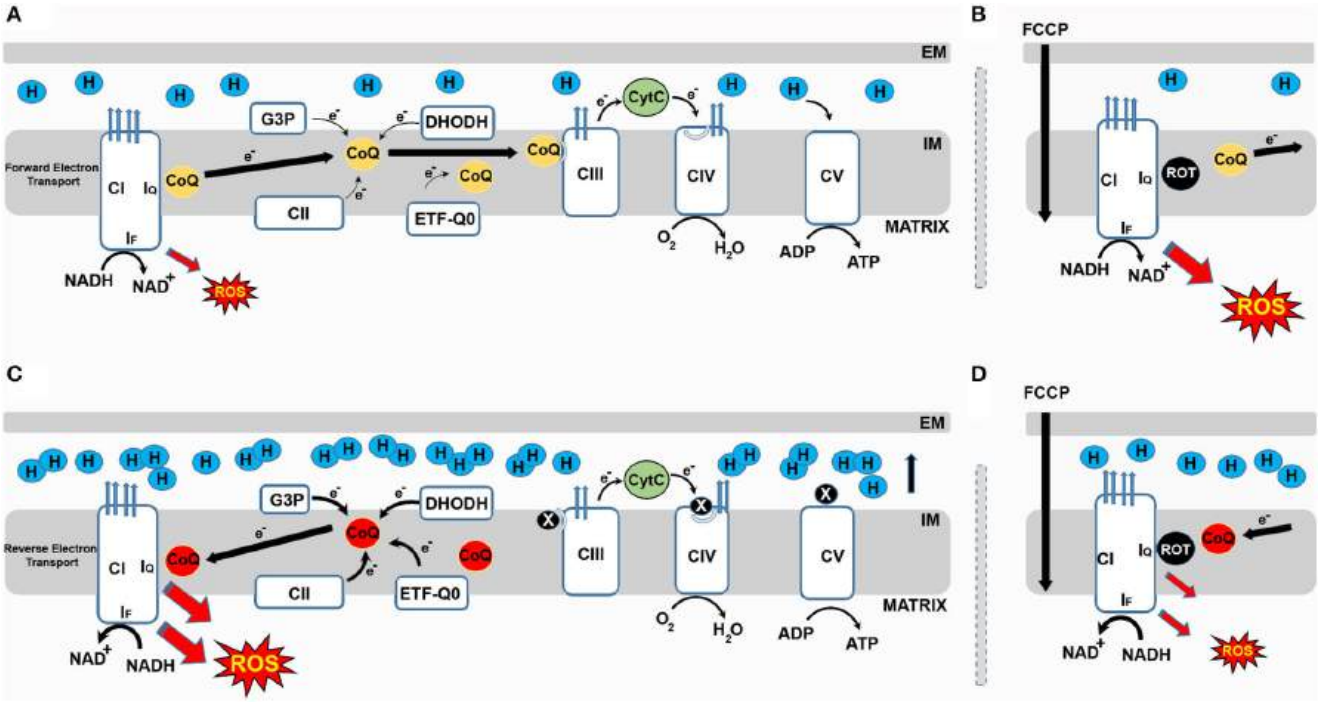


Figure 19.1.18: Complex I (CI) produces ROS in both the forward and reverse directions. Scialò Filippo et al. Frontiers in Physiology (8), 2017. https://www.frontiersin.org/article/...hys.2017.00428. DOI=10.3389/fphys.2017.00428. **Creative Commons Attribution License (CC BY)**.

(A) During **forward electron transfer (FET)**, ubiquinone, also called Coenzyme Q (CoQ) receives electrons from complexes I and II, glycerol-3-phosphate dehydrogenase (G3P), electron-transferring flavoprotein (ETF-QO) or dihydroorotate dehydrogenase (DHODH). During this process electrons mainly leak to produce superoxide from the FMN binding site (I_F) of CI during the oxidation of NADH to NAD^+.

(B) In these conditions, if rotenone blocks the ubiquinone binding site (I_Q) site, electrons cannot be transferred to CoQ, so they leak and generate ROS. FCCP, a molecule that collapses the H^+ gradient by simple diffusion across the membrane, bypassing Complex V - ATP synthase) also increases mitochondrial ROS generation during FET.

(C) When the CoQ pool becomes over-reduced, a high membrane potential favors the reverse transfer of electrons from ubiquinol to CI in a process called **reverse electron transport (RET)**. During RET electrons leak at either I_F or I_Q generating a significant amount of superoxide.

(D) Blocking the I_Q site with rotenone during RET prevents CoQ from transferring electrons back to CI and reduces ROS production. Similarly, FCCP reduces ROS by dissipating membrane potential.

19.1.3: COMPLEX III

Complex III is a complicated, multisubunit protein that is at the heart of the respirasome. The subunits involved in electron transfer are cytochrome b, cytochrome c_1, and the Rieske iron-sulfur protein (ISP). Cytochrome b has two hemes. One is cyto b_{562} which is also called the low potential heme or cyto b_L. The other is cyto b_{566} which is called the high-potential heme or cyto b_H. The cytochrome c_1 subunit has one heme. The Rieske iron-sulfur protein contains a Fe_2S_2 cluster. Figure 19.1.19 shows the relative position of the bound mobile electron carrier, cytochrome C, and the internal ones, the Rieske Fe/S cluster and cytochrome b_L and b_H. Note also the molecule stigmatellin A, which binds to the site where UQ becomes reduced (called the Qo site) and inhibits the complex. This shows that UQ/UQH_2 are in a position to react readily with the Rieske center and cytochrome b_L heme.

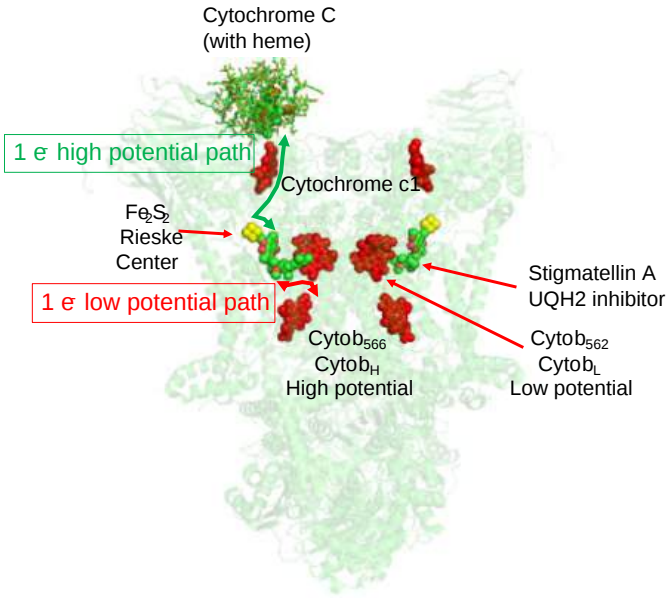
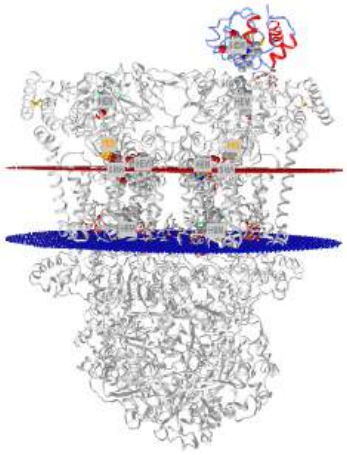


Figure 19.1.19:

Figure 19.1.20 shows an interactive iCn3D model of a eukaryotic respiratory complex III (3CX5)



NCBI iCn3D Figure 19.1.20: eukaryotic respiratory complex III (3CX5). (Copyright; author via source).
Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...5riT4GSSGhCHfA. (loads slowly given the size of the structures).

The Rieske iron-sulfur protein has a Fe_2S_2 iron-sulfur cluster which differs from other such clusters in that each Fe is also coordinated to two His side changes, as shown in Figure 19.1.21. Alterations in H bonds to the histidines and to the sulfurs in the complex can dramatically affect the standard reduction potential of the cluster.

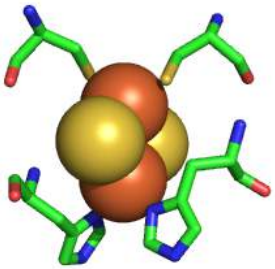


Figure 19.1.21: The Rieske Fe_2S_2 cluster

As with Complexes I and IV, proton and electron transfer are coupled processes. However, in contrast to Complex I, in which protons pass through protein domains that have homology to K^+/H^+ antiporters, and Complex IV, in which they pass through a combination of a water channel and the H-bond network, the protons in Complex III are carried across the intermembrane space by ubiquinone itself. Two reduced ubiquinones (UQH_2) from complex I pass their four matrix-derived protons into the intermembrane space. In the process, four electrons are removed in a multiple-step process called the Q cycle.

The two electrons from each UQH_2 take different paths. One electron moves to a Fe/S Rieske cluster and the other to cytochrome b_L. The electrons moved to the Rieske center then move to cytochrome c_1 and then to the mobile electron carrier cytochrome C which is bound to the complex in the intermembrane space. The electrons moved to cyto b_L are transferred to cytochrome b_H in the complex. Through this latter path, two electrons (from two UQH_2) are then moved to oxidized UQ, and two matrix protons are added to reform one UQH_2. Hence, only one UQH_2 participates in the net reaction shown below.

$$QH_2 + 2 \text{ cyto } c^{3+} + 2H^+_{matrix} \rightarrow Q + 2 \text{ cyto } c^{2+} + 4H^+_{IMS}$$

This net overall reaction, the Q cycle, is illustrated below. This net overall reaction, the Q cycle, is illustrated in Figure 19.1.22.

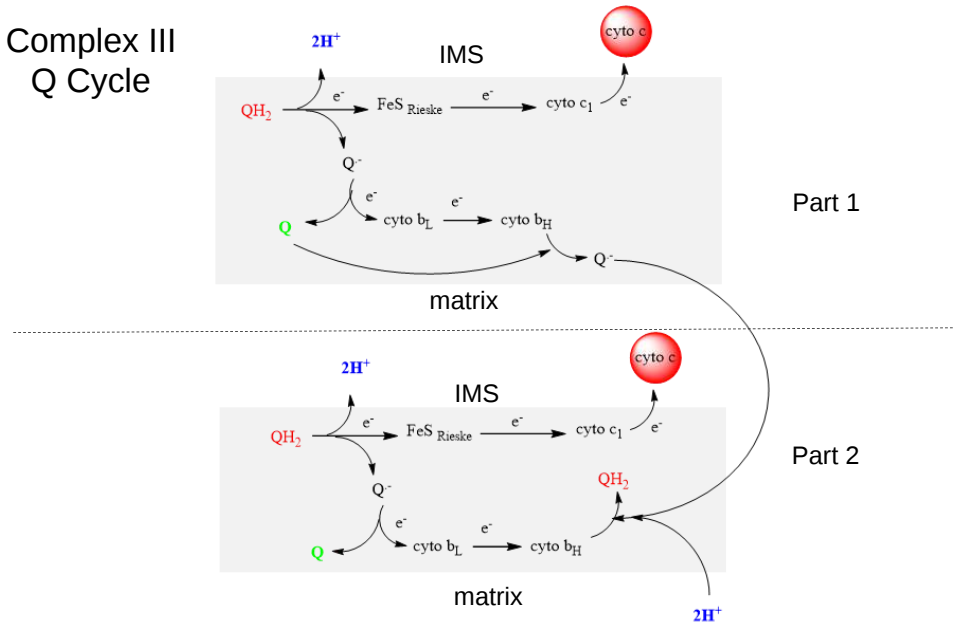


Figure 19.1.22 Q cycle of Complex III

Once again, there are no “proton” channels or hydrogen-bonded networks in the protein for proton transfer across the inner membrane.

Another way to think about the electron transfer process from UQH_2 to cytochrome C is that the 2 electrons from UQH2 take two different paths, one a high potential path to the Rieske center and on to cytochrome C, and another low potential path to the b_L heme and on to the b_H heme and then to UQ to reform UQH2 (see figure above).

Complex III, along with Complex I, can also produce unwanted reactive oxygen species (ROS). Only three of the protein subunits, cytochrome b (with the b_L and b_H hemes), cytochrome c_1, and the Rieske iron-sulfur protein (ISP) are involved in electron transfer, so one of those is most likely involved in ROS production. Experiments and mathematical models support a mechanism that involves a reduction of UQ by the addition of one electron from cytochrome b_L to form $UQ^{\cdot}$ which then passes its electron to dioxygen to form superoxide ($O_2^{-\cdot}$).

As two ubiquinones must bind to the complex, there must be two proximal sites. One is the Qi site where oxidized UQ binds and receives an electron. The other is the Qo site where UQH_2 binds.

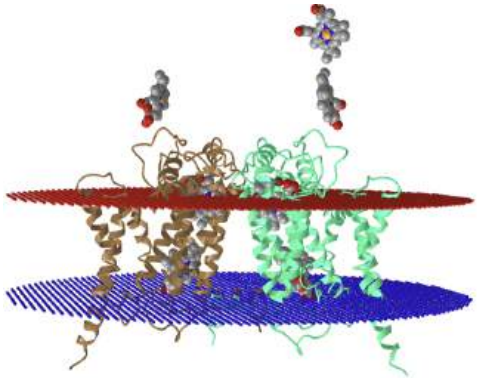
From a kinetic perspective, the first UQH_2 binds and transfers two electrons, one to the Rieske cluster (and on to cytochrome c_1 and then to cytochrome C) and one to cytochrome b_L (and on to heme b_H) and then to an oxidized UQ bound at the Qi site. The $UQ^{\cdot}$ radical is stabilized by the adjacent b_H heme which has a lower affinity for electrons. Now a second UQH_2 binds to the Qo site, and transfers two electrons, again one via the Rieske cluster and the second through cytochrome b_L and b_H to the $UQ^{\cdot}$ radical present at the Qi site to form UQH_2 after two protons are transferred to it from the matrix.

Now a second UQH_2 binds to the Qo site, and transfers two electrons, again one via the Rieske cluster and the second through cytochrome b_L and b_H to the $UQ^{\cdot}$ radical present at the Qi site to form UQH_2 after two protons are transferred to it from the matrix.

Antimycin A, an extremely toxic drug, binds to the UQ Qi site and hence blocks electron transfer from cytochrome b_L to b_H at the Qi site. Heme b_L can then pass its electron to dioxygen to produce superoxide.

Before we leave Complex III, just a quick observation. The complex has two identical cytochrome b subunits, each with a b_L and b_H heme. The standard reduction potential of the b_L and b_H hemes in a subunit differ. This is another example that shows that the standard reduction potential, E^0, depends on the environment of the electron acceptor/donor, just as the pKa of an acid depends on its environment.

Figure 19.1.23 shows an interactive iCn3D model of just the cytochrome b subunits and all the hemes in the eukaryotic respiratory complex III (3CX5) tp give more clarity to the differing environments of the hemes in the two identical cytochrome b subunits.



NCBI iCn3D Figure 19.1.20: Cytochrome b subunits and all the hemes in the eukaryotic respiratory complex III (3CX5). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...BaLQH2KwahNFt9

https://structure.ncbi.nlm.nih.gov/i...BaLQH2KwahNFt9

19.1.4: COMPLEX IV - CYTOCHROME C OXIDASE (CCOX)

This is the final complex in the electron transport chains and is the one that passes 4 electrons and 4 "substrate" protons to dioxygen to produce water. In addition, it moves as many as 4 H^+s from the matrix to the intermembrane space, for a net change of 8 H^+s. Here is a modified version of Equation 1 that represents the transported Hs.

$$O_2 + 4\,e^- + 4\,H^+ + 4\,H^+(in) \rightarrow 2\,H_2O + 4\,H^+(out).$$

The structure of complex IV is shown in Figure 19.1.24. KEGG pathways (with permission).

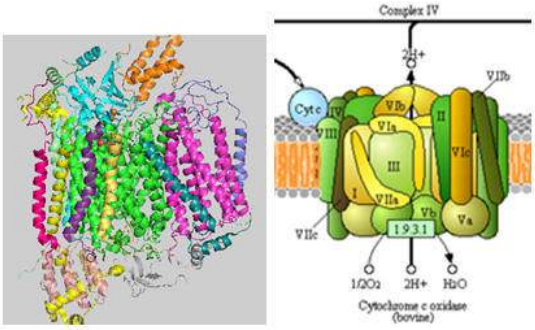


Figure 19.1.24 Cytochrome C Oxidase structure (left) and cartoon model (right) showing subunits (taken KEGG pathways with permission). (Copyright; author via source)

The 4 H^+s that are moved across the membrane are not transported by protein domains that are structural equivalents of antiporters, as in Complex I. Rather elaborate H-bonded water channels with contribution from a space of polar backbone and side chain H-bond donors and acceptors are involved in proton transport. Hence a key question to understand is how electron transfer is coupled to H^+ transfer through this hydrogen-bonded network. Since it's quite complicated we will present multiple different but somewhat redundant figures of low to high resolution to help explain this coupled process.

Figure 19.1.25 shows an interactive iCn3D model of the 14-subunit human cytochrome c oxidase (5z62) with bound cofactors required for electron transfer. Note that it does not show cytochrome C.

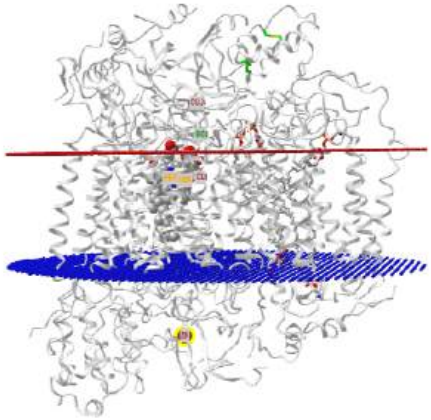


Figure 19.1.25: 14-subunit human cytochrome c oxidase (5z62). (Copyright; author via source).
Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...XoHRyCkpeRN917. (loads slowly given the size of the structure).

Reduced cytochrome C, a mobile electron carrier protein, binds to Complex IV in the intermembrane spaces and singly passes 4 electrons to a series of electron carriers (oxidizing agents) in the protein, eventually to dioxygen as the terminal electron acceptor. Figure 19.1.26 shows the arrangement of the electron acceptor cofactors in Complex IV. Electron transfer presumably occurs through quantum mechanical tunneling.

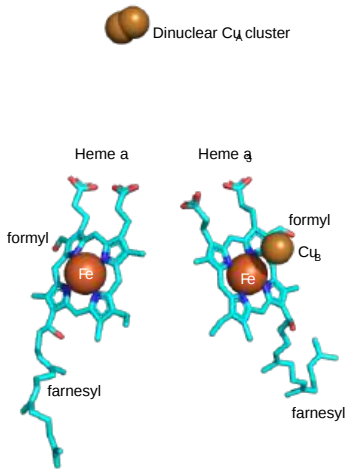


Figure 19.1.26: Electron acceptor cofactors in Complex IV.

Cytochrome C, the initial “substrate” of this complex, delivers electrons from its heme cofactor to a dinuclear copper cluster, Cu A, where the copper ions are collectively coordinated to two histidines and two cysteines, and a methionine. Figure 19.1.27 shows an interactive iCn3D model of the coordinating ligands for Cu A cluster in Complex IV (5z62)

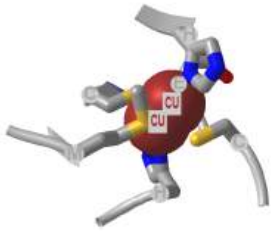
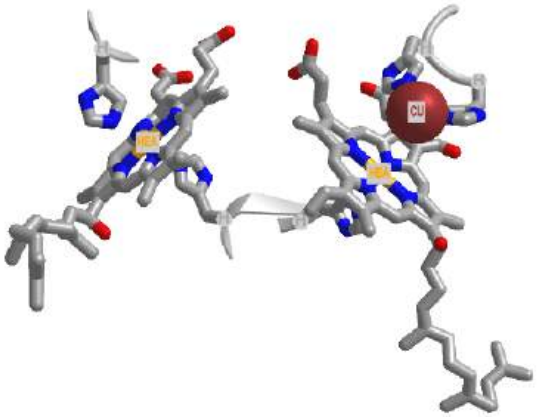


Figure 19.1.27: coordinating ligands for Cu A cluster in Complex IV (5z62). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...7ARphDU4oyDS4A

From there, electrons flow to an adjacent **heme a** (low spin), which has two coordinating histidine ligands. This heme hence does not bind dioxygen. From there, electrons move to another heme, **heme a_3** (high spin), and then finally to dioxygen which is coordinated to the Fe in heme a_3 and to an adjacent Cu B. Heme a_3 has only one coordinating histidine ligand, which allows dioxygen binding to the unligated site.

The heme a_3 Fe:Cu dinuclear cluster is unique among all hemes. Figure 19.1.28 shows an interactive iCn3D model of the coordinating ligands for the hemes and Cu B in Complex IV (5z62)



NCBI iCn3D Figure 19.1.28: coordinating ligands for CuB, heme a and heme a_3 in Complex IV (5z62). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...4EUooXkwh13Fm8

First, let's consider the transfer of electrons from heme a to heme a_3 and then on to dioxygen (we will consider the entry of electrons into the complex later). If dioxygen, a substrate for the reaction, dissociated from the heme a_3 Fe before it was completely reduced, toxic ROS would result. This suggests a reason for the evolution of this key enzyme to have produced the unique heme a_3 Fe:Cu dinuclear cluster

Heme a and a_3 vary from the heme in hemoglobin as they both have a formyl group replacing a methyl and a hydroxyethylfarnesyl group added to a vinyl substituent. Its structure is shown in Figure 19.1.29. What is its overall charge of the heme in its reduced state? In its oxidized state?

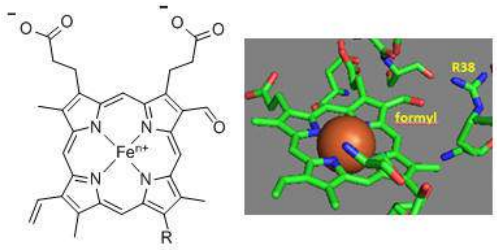


Figure 19.1.29: Heme-Formyl group of Cytochrome C Oxidase

As hemes are prosthetic groups, they must bind to the apoprotein. The hydroxyethylfarnesyl group, with its nonpolar tail, facilitates the binding of the heme to both the Fe^{+2} and Fe^{+3} forms by about 6.3 kcal/mol (26 kJ/mol). The formyl group in heme a increases the E^0 value (i.e. make the heme a better oxidizing agent) by about 179 mV (4.1 kcal/mol, 17 kJ/mol). The effect appears to be mediated by the binding of the heme group to the protein again, only in this case there is a differential effect. The formyl group preferentially stabilizes the binding of the Fe^{+2} heme compared to the Fe^{+3} heme, which in a thermodynamic cycle, would promote the binding of an electron to Fe^{+3}-heme (increasing its E^0 value) to form the more stably bound Fe^{+2} heme.

The key challenge has been to understand the redox coupling to H^+ transport. How is this done? The electrostatic environment of the hemes must be considered. In heme a, the two axial ligands are uncharged imidazole side chains of histidines. The overall charge on Fe^{2+}-heme is zero as the two propionates cancel the charges on Fe^{2+}. However, there is a net +1 charge in the Fe^{3+}-heme. This charge could be delocalized within the conjugated pi electrons of the planar heme, which could contribute to the deprotonation of a side chain near heme a in a process that would couple redox and H^+ movement. The formyl group of heme a is coplanar with the heme ring in both Fe^{3+}-heme and Fe^{2+}-heme, which allows it to participate in the pi-conjugated electron system of the ring. As the formyl group forms an H bond with Arg 38, changes in oxidation might affect Arg 38 as well.

Many other amino acids are involved as well. One is Asp-51 of subunit I which contains heme a. This undergoes a conformational change which moves it to the surface of the inner membrane on reduction. It is near the matrix in the oxidized state. Hence it is likely involved in a proton transport pathway. In fact, crystal structures of oxidized and reduced Complex IV show water channels and small "cavities" which calculations show can hold 1-3 water molecules. Hence groups around the heme, including R38 are assessable to water. Water would be involved in H^+ transport through "handshaking" transfer of protons through a hydrogen-bonded network of waters and selective side chains. Some of the amino acid residues associated with the water channels are shown in Figure 19.1.30: and include R38 (the one hydrogen bonded to the formyl group of the heme), S34, T 424, S461, S382, and H413.

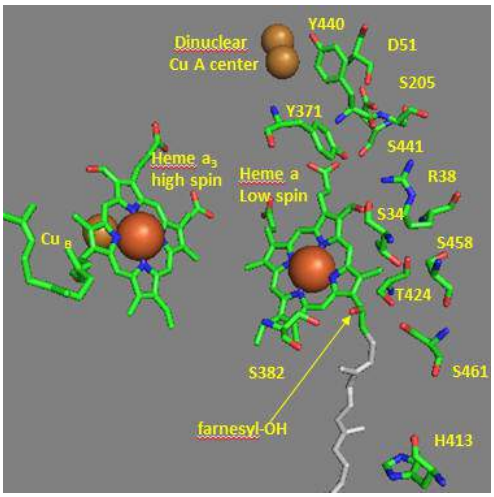


Figure 19.1.30: Possible amino acids involved in H^+ transport through a hydrogen-bonded network in Complex IV

Another consequence of electron transfer to heme a involves the interaction of S382 and the farnesyl OH group, which are close in space and proximal to a water cavity. On reduction of the heme, a conformation change occurs which increases the standard reduction potential of the S382-farnesyl OH group. What effect would this have on the interaction of the two and the S382-L381-Val380 localized conformation? A new water cavity appears to emerge on reduction in this region. How might this impact proton transfer from the matrix?

Now let's consider the entry site of electrons into the complex and how they might influence proton transport. As mentioned above, Asp 51 (D51) appears to play a key role. It is shown in Figure 19.1.29 above and also in Figure 19.1.30 below. In the oxidized state, D51 interacts with two OH side chains (S205 and S441) and amide NH backbone groups but is not exposed to water. On reduction, D51 lies on the surface in an aqueous environment. Near D51 is Y440-S441. The backbone carbonyl group between 440 and 441 forms an "indirect" interaction with R38 (right panel in Figure 19.1.31), which we showed earlier is affected by the redox state of heme a. They are too far apart to form H bonds. Add two water molecules and envision a bridging interaction between the carbonyl O and the side chain R38 via Y371. Likewise in your mind add a water to allow a bridged hydrogen bond interaction between Y371 also forms a H bond and the heme a propionate.

Figure: Role of Amino Acids near D51 in Cytochrome C Oxidase

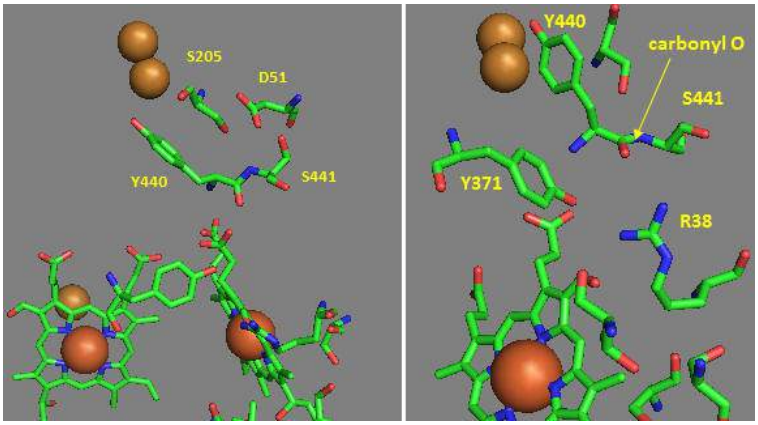


Figure 19.1.31: Local environment of D51, R38 in hemes in Complex IV

A more nuanced understanding of the mechanism and linkage between H^+ and e^- movement derives from high-resolution structures determined by Yano et al (2016). In their model (shown in the figures below based on the oxidized form of the protein, pdb 5b1a), protons from the negative (matrix) N side of the complex enter through a water channel and proceed to the positive (intermembrane side) through a H bond network (as described above and depicted below). These comprise the **H-Pathway** for H^+ transfer across the membrane. Two other H^+ transfer pathways, the **D- and K-pathways,** are used as a source of substrate H^+ for the 4 H^+s added to dioxygen to form $2H_2Os$.

Directional movement is mediated by proton:proton repulsion aided by an increase in + charge on heme a when it transfers an electron to heme a_3. Of course, proton:proton repulsion would move protons in both directions. Reverse flow back through the water channel is prevented by a conformational change on oxygen binding that closes the channel.

Ultimately 4 electrons are transferred from cytochrome Cs (in single electron steps) to the dicopper cluster, CuA, and then sequentially to heme a to heme a_3 (near the copper B ion) to dioxygen to form water. The motion of electrons and protons is coupled electrostatically.

Figure 19.1.32 gives an overview of these movements. The small red dots are the oxygen atoms of internal water molecules (the rest have been removed using Pymol). It should be apparent, given the number and location of the internal water molecules, that many would be involved in the proton translocation pathways.

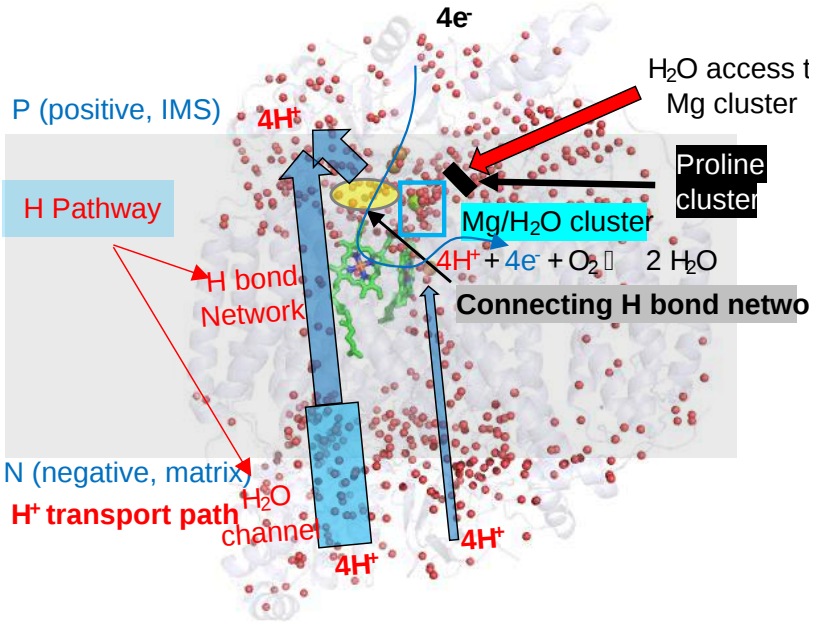


Figure 19.1.32 Cartoon showing movement of H^+s and e^-s in Complex II

What's so interesting about this model is the detailed description of two types of protons, the ones that add to dioxygen and end up in water (substrate protons), and those that are vectorially (directionally) transported to the IMS. In their model, the H^+s that end up being transported move through the water and H bond network through a connecting H bond link region to a Mg^{2+}/water cluster. Since the binding of oxygen leads to structural changes that close off the water channel, all protons to be transported to the IMS must be bound in the cluster before dioxygen binding.

Figure 19.1.33 shows that initially, 4 H^+s move through the H system to the Mg^{2+}/H_2O cluster. Oxygen binding then closes the water channels. This buildup of positive charges would certainly lead to enhanced electrostatic attractions for the next phase of the reaction, the movement of electrons into the heme cofactors. Additionally, the 4 H^+s in the cluster are probably prevented from leaking to the P side through waters that are proximal (see above figure) by a proline cluster, which presumably restricts the dynamical motion of the protein in that region necessary for proton movement. The figure does not show charge changes in the electron carriers.

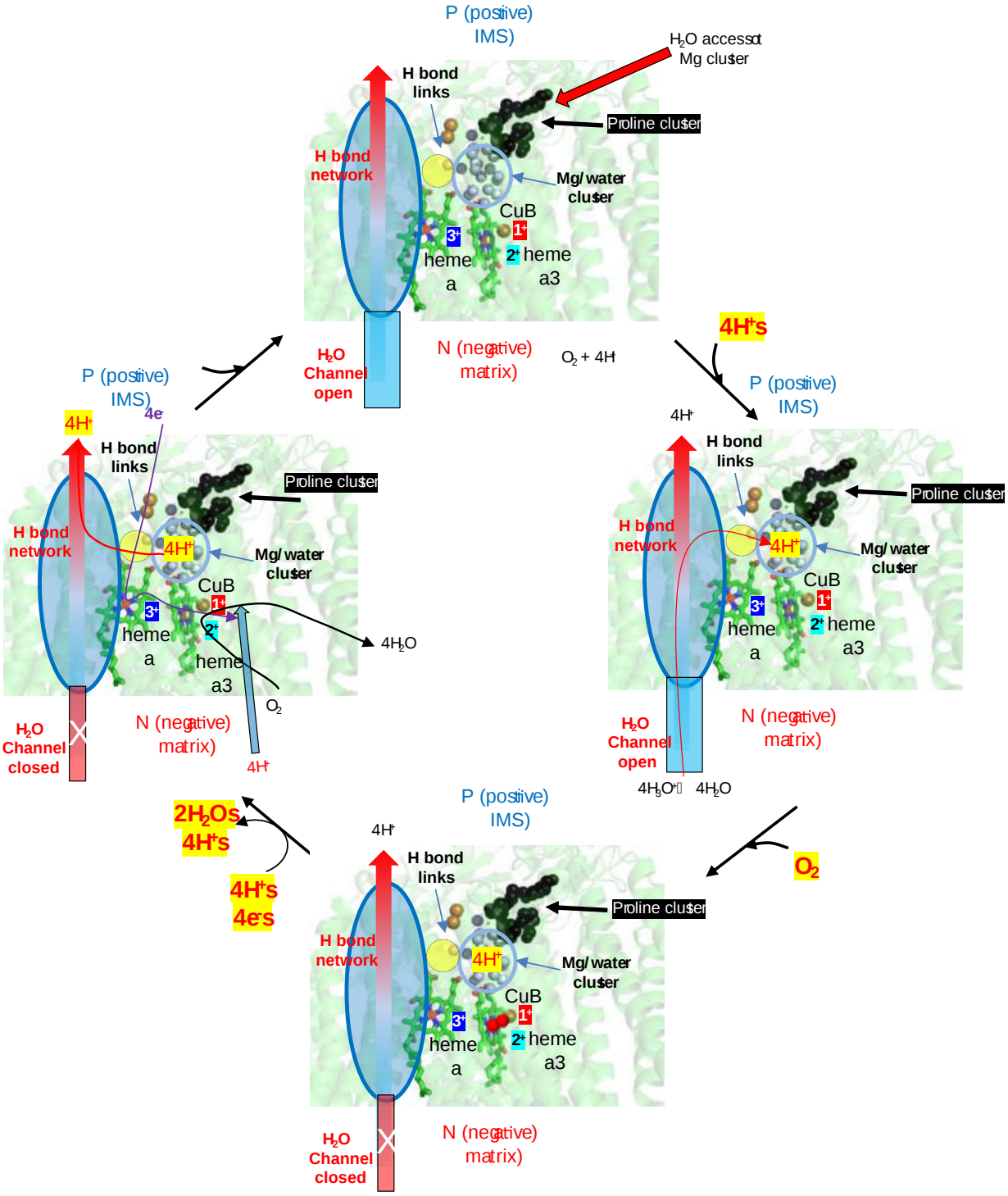


Figure 19.1.33: Proton movement in Complex IV

The figure below breaks downs the mechanism to show the addition of the first electron to the CuA (dicopper cluster), delivered from cytochrome C, and the subsequent transport of one proton from the fully proton-loaded Mg^{2+}/water cluster after dioxygen binding. This figure does show the stepwise redox changes in the electron carriers.

After CuA receives an electron from cytochrome C, it donates it to heme a and not to heme a_3, even though both are close. The extra negative on heme a facilitates proton pumping though the H pathways shown.

Figure 19.1.34s shows a cartoon description of the movement of protons and electrons through Complex IV.

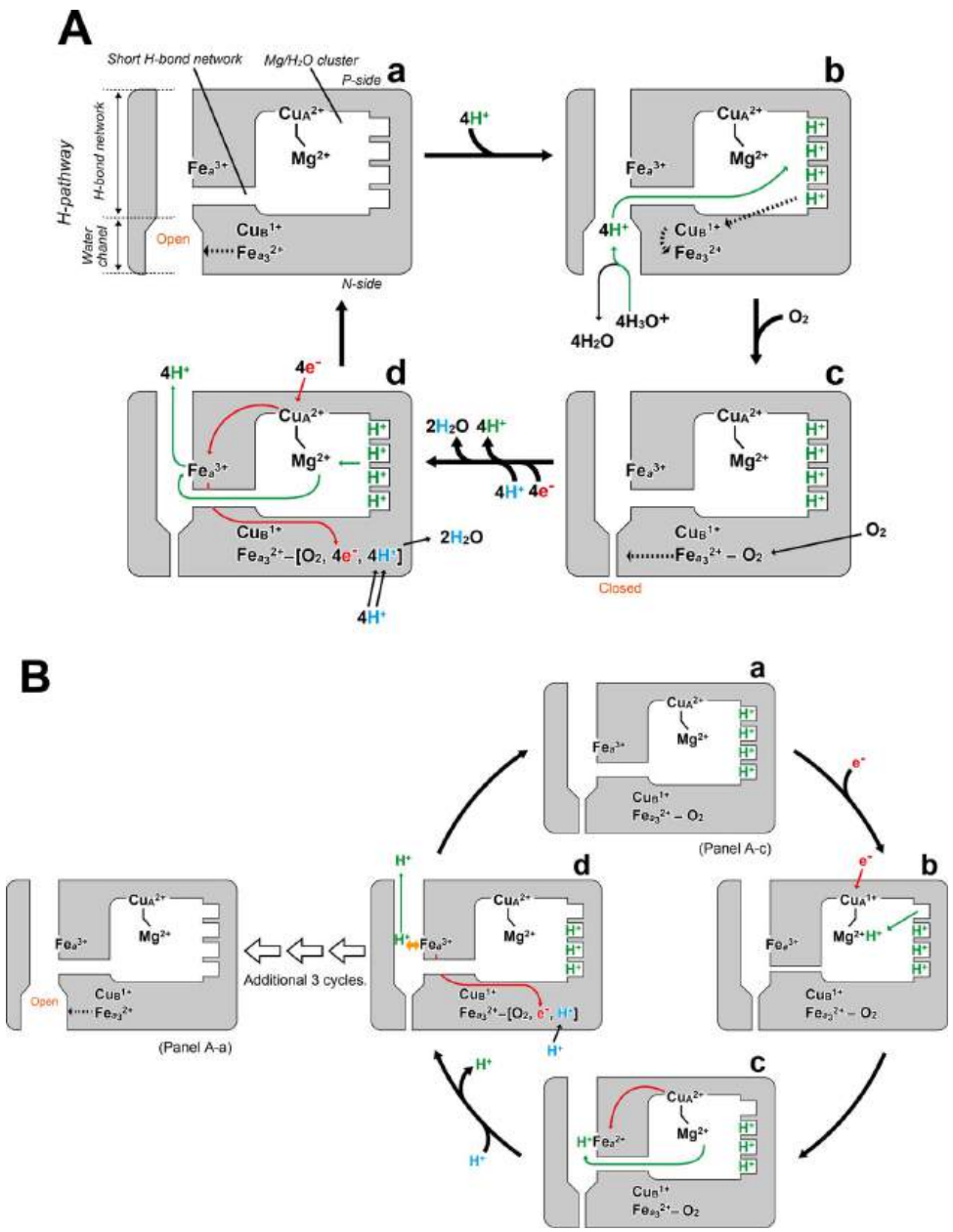


Figure 19.1.34: **A schematic representation of the role of the Mg/H_2O cluster in the proton-pumping mechanism of bovine heart CcO.** Yano N et al. J Biol Chem. 2016 Nov 11;291(46):23882-23894. doi: 10.1074/jbc.M115.711770. Epub 2016 Sep 7. PMID: 27605664; PMCID: PMC5104913. Creative Commons Attribution (CC BY 4.0)

The Mg/H_2O cluster and the H-pathway are connected with the short hydrogen bond network, as *labeled* in *A* (*a*). The location of heme a is shown by Fe_a^{3+} or Fe_a^{2+} attached to the hydrogen bond network of the H-pathway. The Cu_A site is on the water-accessible surface of the Mg/H_2O cluster. The conformation of the Cu_A-Mg^{2+} complex is shown by the *shape* of a *line* connecting the two metal ions. The proton-accepting sites are shown by the *four hollows* on the water-accessible surface of the Mg/H_2O cluster. The O_2 reduction site is shown by Cu_B^{1+} and Fe_{a3}^{2+}. The pumping and chemical (water-forming) protons are *labeled* in *green* and *blue*, respectively. Electron transfers are shown by *red curves*. *A*, overall catalytic cycle of CcO. *B*, a typical single electron transfer from Cu_A to the O_2 reduction site, coupled with the uptake and release of protons. For the sake of simplicity, the uptake of one chemical proton equivalent (the average number) is given in *B*. Various oxidation and ligand binding states in both Cu_B and Fe_{a3} in the intermediate states shown by *A* (*d*) and *B* (*d*) are not included for the sake of simplicity.

Figure 19.1.35 converts the schematic representation into a "structural cartoon" representation showing one cycle of coupled electron and proton transfer into Complex IV.

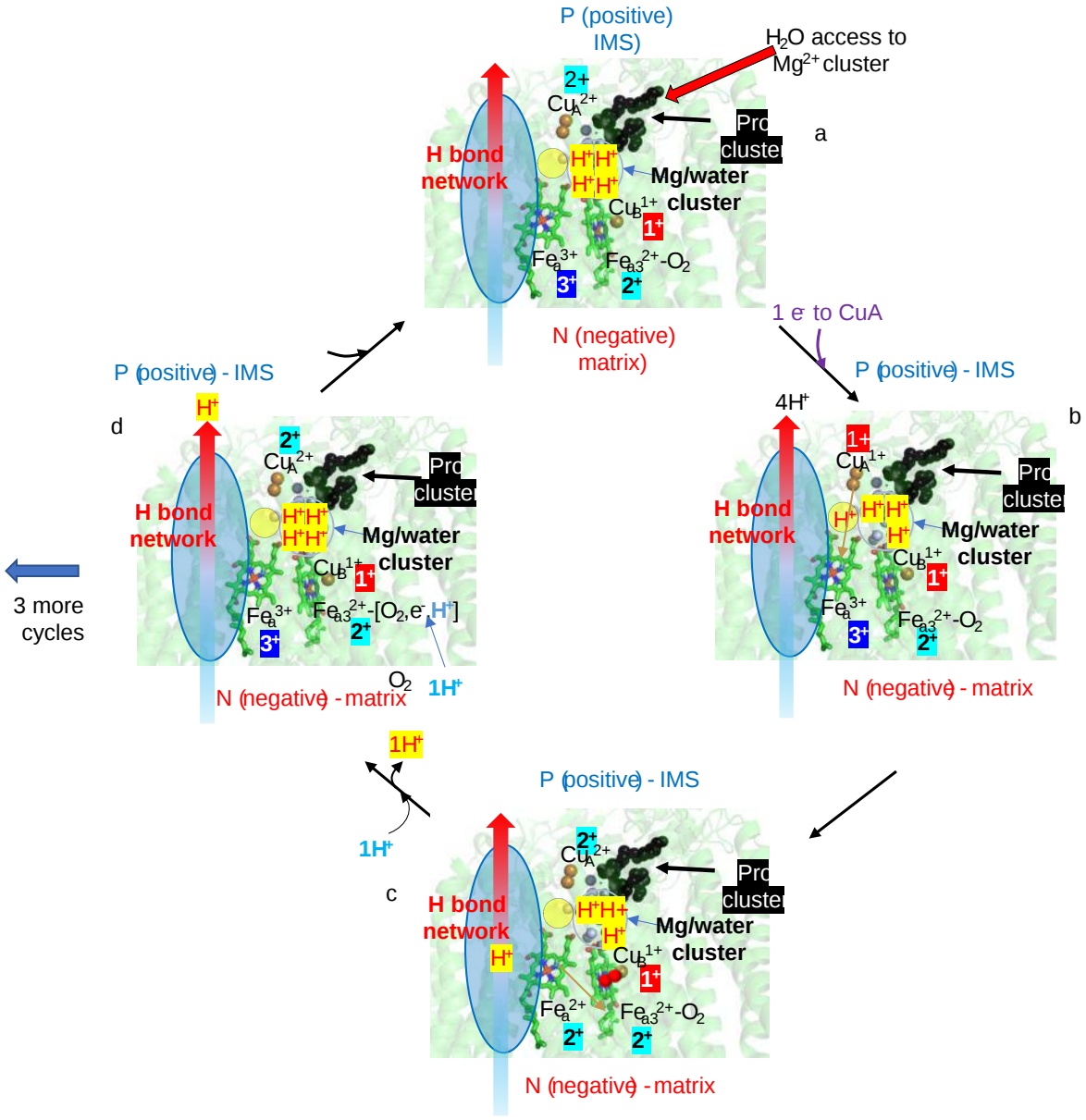


Figure 19.1.35: Structural cartoon representation showing one cycle of coupled electron and proton transfer into Complex IV.

The H- (for transport of $4H^+$s across the membrane) and the D- and K-pathways (sources of substrate $4H^+$s for water synthesis from O_2 are shown in Figure 19.1.36.

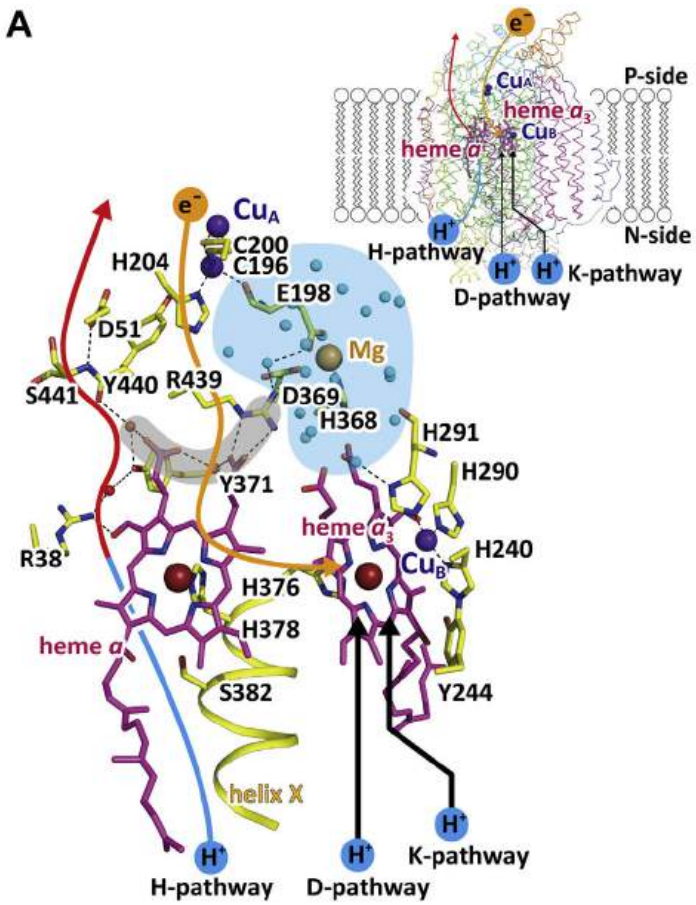


Figure 19.1.36: The H-, D- and K-pathways in bovine heart Complex IV (5B1A). Atsuhiro Shimada et al. J. Biol. Chem.(2021) 297(3) 100967. Creative Commons. https://creativecommons.org/licenses/by/4.0/

Two black arrows show the D- and K- for substrate H^+s that form water, while the blue arrow shows the H-pathway for the hydrogen bonding network of water in the water channel. The Mg/H_2O cluster (blue area) is attached to the hydrogen-bond network of the H-pathway by a short hydrogen-bond network (gray area with water shown as blue sphere). The formyl group and one of the propionate groups of heme a are hydrogen-bonded with Arg38 and a fixed water molecule in the hydrogen-bond network of the H-pathway. Compare this diagram to the ones presented in the figures above it. Again we show multiple representations to give readers many opportunities to conceptualize this complicated complex.

19.1.5: COMPLEX II - SUCCINATE:UBIQUINONE OXIDOREDUCTASE (SQR) - SUCCINATE DEHYDROGENASE

You've seen complex II before with another name, succinate dehydrogenase from step 6 of the citric acid cycle, which catalyzes the following reaction.

$$\text{succinate} + \text{FAD} \leftrightarrow \text{fumarate} + \text{FADH}_2 \quad \Delta G^o = 0 \text{ kcal/mol}$$

The ΔG^o for this reaction is about 0 kcal/mol so it is readily reversible.

However, something vital is left out of this description. The above reaction appears to suggest that the FAD/$FADH_2$ are readily diffusable and bind and unbind after catalysis. That's not true. Instead, FAD/$FADH_2$ are covalently attached to the enzyme so after one cycle of enzyme catalysis, the enzyme is dead. Another substrate/product pair must interact with enzymes to reconvert the covalently attached $FADH_2$ back to FAD so catalysis can continue. The other pair is ubiquinone (UQ/UQH_2). A better description of the reaction equation is shown in Figure 19.1.37. Complex II clearly links the tricarboxylic acid cycle to electron transport!

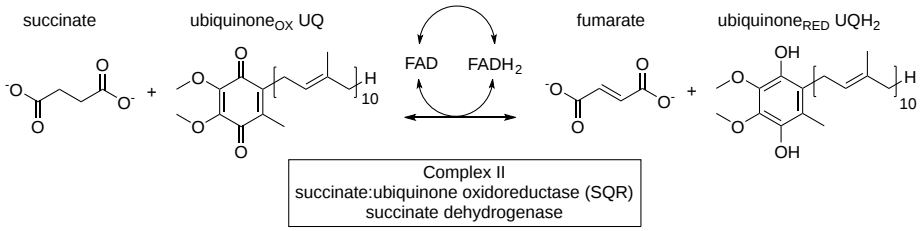
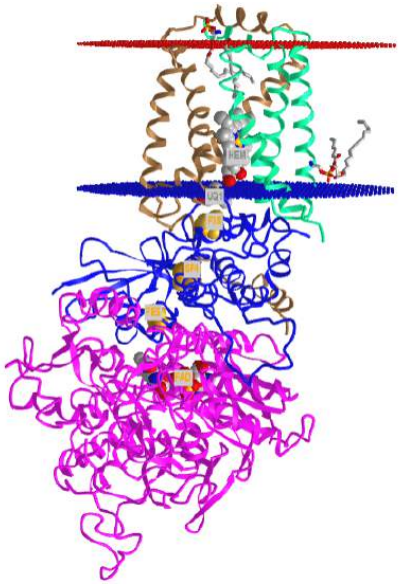


Figure 19.1.37: Complex II - Succinate Dehydrogenase Reaction

As with the other mitochondrial electron transport complexes, Complex II is a multisubunit integral membrane protein. The hydrophilic section protruding into the matrix contains two proteins, a flavin-binding protein (Fp) and an iron/sulfur protein (Ip). The transmembrane domain has two "cytochrome binding proteins", a lighter one (CybL) and a heavier one (CybS). The complex has binding sites for dicarboxylic acids (succinate/fumarate) and two binding sites for ubiquinone, a proximal higher affinity site (Qp) on the matrix side closer to the succinate binding site, and a distal lower affinity site (Qd) near. The ubiquinone sites bind 2-thenoyltrifluoroacetone (TTFA) which acts as an inhibitor of ubiquinone reduction. 3-nitropropionate - $(NO_2)(CH_2)_2CO_2^-$ is structural similar to succinate - $(CO_2^-)(CH_2)_2CO_2^-$ and is an inhibitor of succinate oxidation.

Figure 19.1.38 shows an interactive iCn3D model of the Mitochondrial Respiratory Complex II from porcine (pig) heart (1ZOY). The chains are:

- hydrophilic FAD-binding protein (lavender)
- hydrophilic FeS protein (blue)
- integral membrane large cytochrome binding protein (brown)
- integral membrane small cytochrome binding protein (green)



NCBI iCn3D Figure 19.1.38: Mitochondrial Respiratory Complex II from porcine heart (1ZOY) . (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?KKPWG9dW1ARq3wQf9. (loads slowly given the size of the structure).

Given the role of the enzyme in both the mitochondrial citric acid and electron transfer pathways, you would expect high electron transfer rates between the dicarboxylic acid and the ubiquinone binding sites. Any disruptions in electron transfer (by inhibitors or mutations) could lead to the production of toxic reactive oxygen species (ROS). Binding an inhibitor to the proximal Qp site would lead to a buildup of $FADH_2$ which could transfer electrons in single-electron steps to dioxygen to produce ROS. This would occur if the dicarboxylic site were empty.

As there is no H^+ transfer from the matrix to the intermembrane space for Complex II, we'll only focus on electron transfer from succinate to FAD and on to ubiquinone.

The transfer of electrons from succinate to ubiquinone hence takes place in two steps.

Step 1: Succinate + FAD ↔ Fumarate + $FADH_2$. Figure 19.1.39 shows a plausible mechanism using a histidine in the active site as a general base. This is a simple hydride ($2e^-$) transfer reaction to FAD. Note that no FeS cofactors are required.

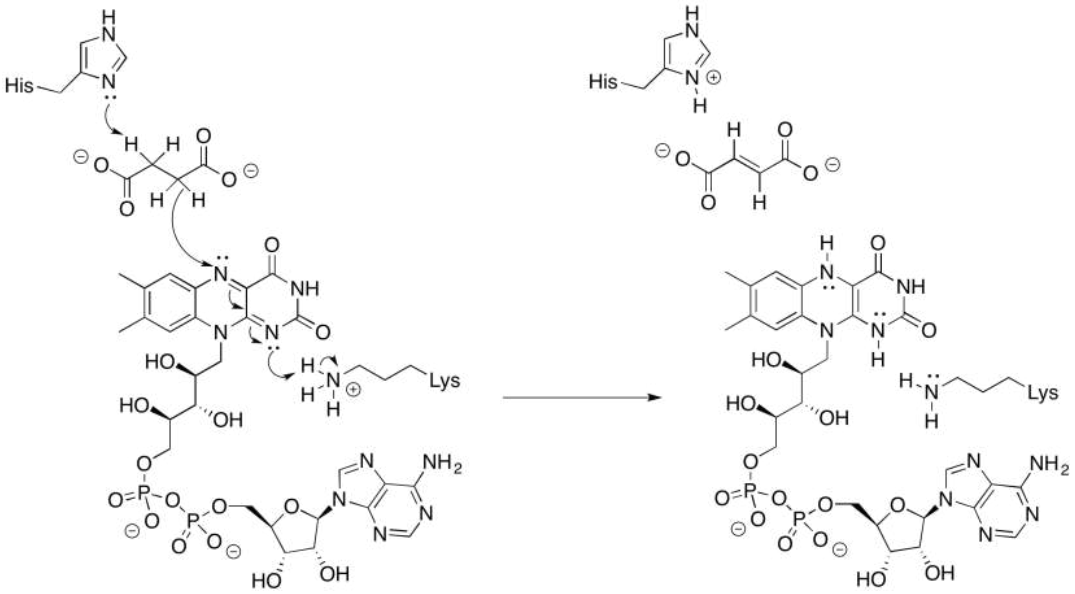


Figure 19.1.39: A mechanism for Succinate oxidation by FAD to Fumarate and $FADH^2$ (after Schaller, LibreText)

Step 2: $FADH_2$ + UQ ↔ FAD + UQH_2. This does not proceed by transfer of 2 e^- from $FADH_2$ directly to UQ. Rather it occurs through a series of the FeS cofactors (Fe_2S_2, Fe_4S_4, and Fe_3S_4 clusters) and heme. First, let's consider the path of electron transfer which will occur in single electron steps through the intermediary FeS clusters to ubiquinone as shown in Figure 19.1.40. The standard reduction potential E^0 is shown for each step. Note that FAD is shown instead of the actual electron donor, $FADH_2$, to correspond to the standard reduction E^0 values shown in tables.

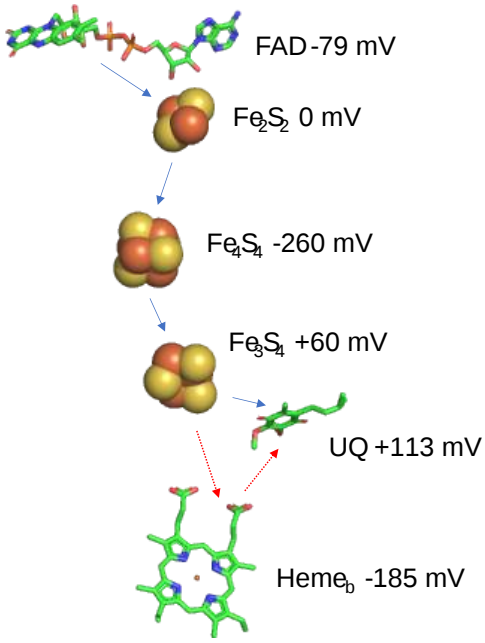


Figure 19.1.40: Pathway for electron transfer in Complex II

Electrons flow from $FADH_2$ (lowest, most negative E^0 value, -79 mV) to UQ (highest, most positive value E^0 value, +113 mV) or from weakest to strongest oxidizing agent. Electron transfer depends on distance as well. It appears that the transfer of electrons to heme $_b$ is not likely, given the presence of the closer UQ with the most positive E^0 value. Hence the function of heme $_b$ is unclear. However, if electrons are transferred to it, it would quickly and favorably send them on to UQ. Note that the FeS clusters have progressively more positive E^0 values except for Fe_4S_4 whose E^0 = -260 mV. Its E^0 is depressed by the presence of charged/polar groups which would make the transfer of a negative electron less likely. The net transfer standard potential is clearly favored thermodynamically. As the E^0 and corresponding G^0 values are state functions, the final ΔG^0 doesn't depend on the path. The high value of E^0 for Fe_4S_4 presents an activation energy barrier for transfer in both directions and probably prevents reverse electron transport in this complex.

The membrane domains have two polar (Asn, Ser) and 5 polar-charged (Asp, Glu, and Lys) side chains.

Now let's look at the terminal step in the transfer of electrons from Fe_3S_4 to ubiquinone. Figure 19.1.41 shows a plausible mechanism. Ubiquinone is shown in black. TTFA, which binds to the ubiquinone site and inhibits electron transfer, is shown in brackets. Tyrosine 91, Serine 41, and Histidine 216 are likely candidates for involvement in catalysis as shown below. Arginine 46 and aspartate 90 might be involved in the protonation steps of UQH_2.

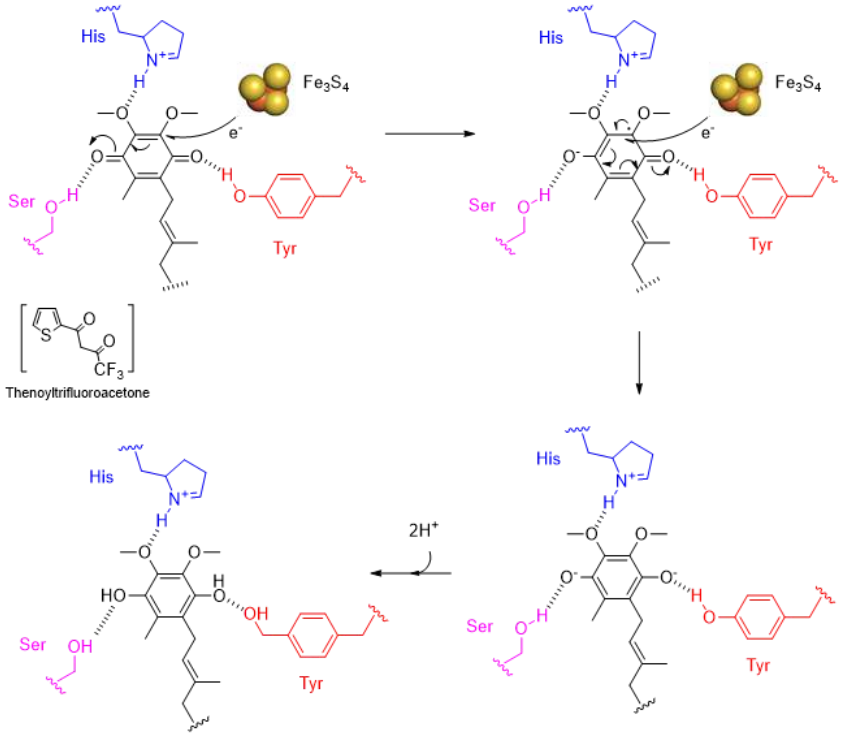


Figure 19.1.41: Terminal step in the transfer of electrons from Fe_3S_4 to ubiquinone

Figure 19.1.42 shows an interactive iCn3D model of the iCn3D TTFA1 binding site (where ubiquinone binds).

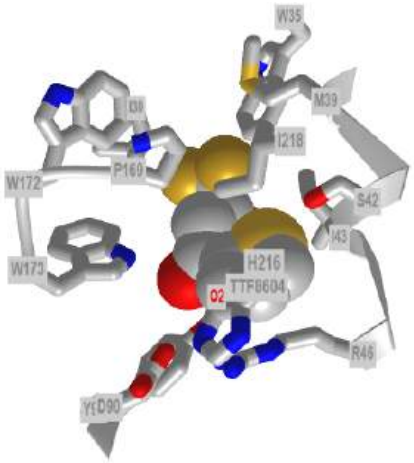


Figure 19.1.42: Mitochondrial Respiratory Complex II bound with 2-thenoyltrifluoroacetone (1zp0). (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?dY3niSfLVa9WiNNe7

19.2: ATP SYNTHESIS

Search Fundamentals of Biochemistry

19.2.1: INTRODUCTION

ATP synthase, also called F_1F_oATPase, is a rotary motor enzyme. This enzyme is found in the inner membrane of mitochondria, the analogous thylakoid membranes of chloroplasts, and the cell membrane of bacteria. The enzyme consists of two parts, the membrane-bound F_o which is a proton translocator, and the F1 part which has catalytic (ATP synthesis or hydrolysis activity. The F_o part can be considered to be a rotary electrical motor powered by proton flow, while the F_1 part acts as a rotary chemical motor powered (in reverse) by ATP hydrolysis. The F_o part is named since it is sensitive to **o**ligomycin (note that it should theoretically be read as F_o and not F_{zero}). The F_1 part is named since it was eluted from a column chromatography column in **F**raction **1**.

The elucidation of its structure and mechanism by many but especially by Paul D. Boyer and John E. Walker, was a major feat of scientific study. They, along with Jen Skou, who studied the mechanism of the analogous Na/K ATPase, were awarded the Nobel prize for their work.

Let's start with the known structure of this complicated membrane-bound rotary enzyme and then work towards an understanding of how it works. The structure of the bovine F_1F_oATPase is shown in Figure 19.2.1 .

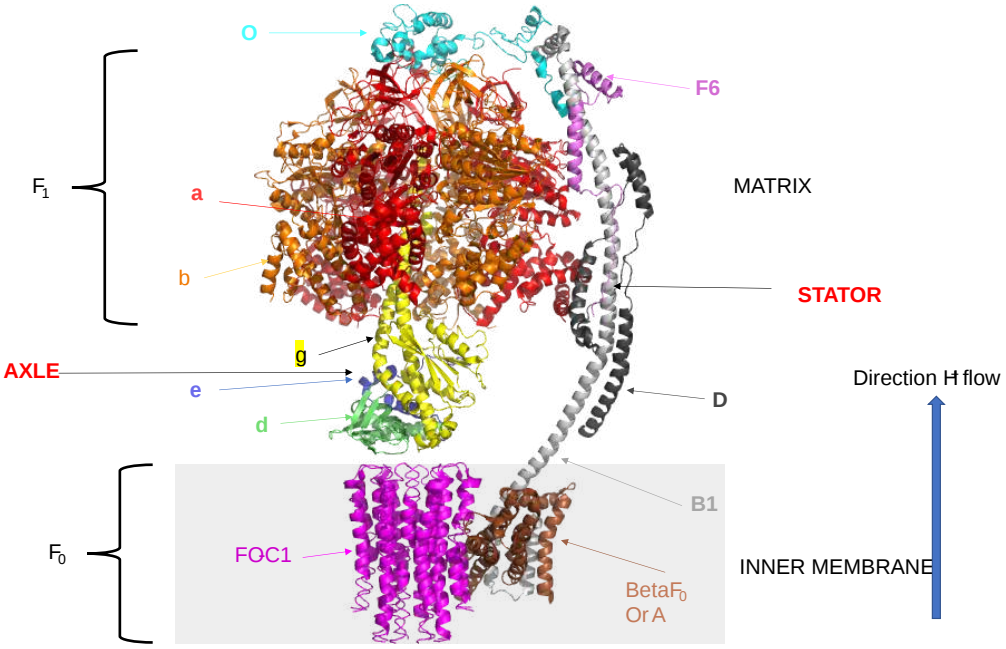


Figure 19.2.1 : Bovine F_1F_oATPase.

In the previous section, we discussed how oxidative electron transport in the mitochondria is accompanied by proton transport against a concentration gradient from the matrix through the inner member to the intermembrane space. The synthesis of ATP from ADP and P_1 is endergonic and requires an energy source. That energy source is provided by the thermodynamically favored collapse of the pH gradient across the mitochondrial inner membrane. Protons flow through the F_o membrane spanning C helices from the intermembrane space (not shown in Figure 1) to the matrix. This powers the synthesis of ATP from ADP and P_i by the F_1 part of the enzyme in the matrix lumen.

Figure 19.2.2 shows an interactive iCn3D model of the entire bovine mitochondrial ATP synthase (5ARA). The subunits are shown in the same color as Figure 1 above.

Figure 19.2.2: Bovine mitochondrial ATP synthase (5ARA). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...3YUBWrX91nnua9

The enzyme is reversible. If protons flow down a concentration gradient through F_o, ATP is synthesized by F_1. Alternatively, ATP hydrolysis by F_1 leads to the transport of protons through F_o and against a concentration gradient. Isolated F_1 can only break down ATP, and not synthesize it.

We called this structure a rotary enzyme, so what rotates? A rotary "axle" protein complex comprised of the γ, ε, and δ subunits (some also describe the γ, ε, and δ subunits as part of F_1) rotates. Its interactions with the multimeric complex of c alpha helices (shown in purple in the figures above) in the main F_o complex causes the "c-ring complex" complex in F_o to rotate as well. The F_1 part does not rotate because of the conformational stability of the β subunit and the connection to the long alpha helices of the D and B1 proteins, which comprise the "stator (the stationary) part of an electric motor", which keeps F_1 stationary.

Hence the enzyme does truly act like an electric motor, whose components are shown in Figure 19.2.3.

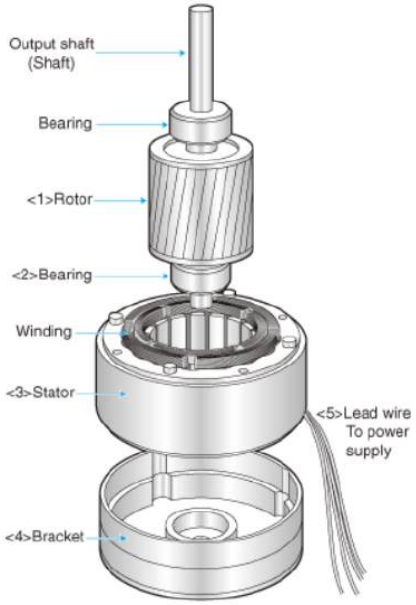


Figure 19.2.3: Electric motor with rotor, shaft, and stator. http://www.cdmmotor.com/en/article-22982-28485.html

Now let's look at the mechanisms and the evidence for them for each part of this remarkable enzyme.

19.2.2: F_1

The F_1 unit (with a quaternary structure of $a_3\beta_3$ forming a hexagonal ringed structure with a central cavity, occupied by a gamma subunit)is about 80 angstroms from the Fo subunit and both are connected to the rod-shaped γ subunit which spans the center of the $a_3\beta_3$ ring. Energy transduction (necessary to capture the negative free energy change associated with the collapse of the proton gradient to drive the positive free energy change for ATP synthesis) occurs between the two subunits.

Figure 19.2.4 shows a top view of the α (red) and β (green) subunits in the $a_3\beta_3$ F_1 complex

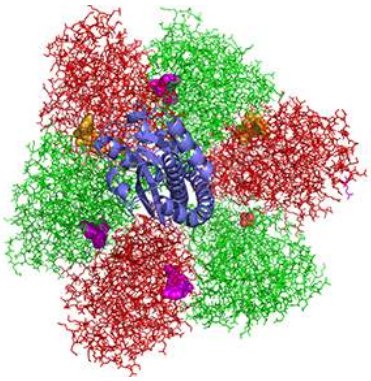


Figure 19.2.4: Top view of the α (red) and β (green) subunits in the $a_3\beta_3$

The γ subunit is shown coming out of the plane of the ring. ATP is shown in gold spheres, ADP in purple, and the smaller SO_4^{2-} (from a crystallizing agent) is shown in CPK spheres. They are sandwiched between the a and β subunits.

Boyer, in the absence of the complete structure of the F_1F_o ATP synthase, was able to deduce from experimental evidence that the $a_3\beta_3$ complex, which can be viewed as three aβ dimers (with catalysis occurring between subunits of individual dimers where ATP and ADP bind), have three different, interconvertible conformation defined as a Loose (L), Open (O) and Tight (T) states, with names describing the strength of substrate binding in each dimer.

- O - open state with very low affinity for substrates and has no catalytic activity;
- L - loose state with low affinity for substrates and also no catalytic activity;
- T - tight state with high affinity for substrates and with catalytic activity.

The three states rotate not physically with respect to some central axis but conformationally, depending on their interaction with the γ subunit which binds perpendicularly in the central junction of the $a_3\beta_3$ ring. Changes in the orientation of the central γ subunit due to its rotation with respect to the $a_3\beta_3$ ring cause the conformation of the O, L, and T states to change in situ with the orientation of the rotating γ subunit

The conversion of the LOT conformations, their binding of substrates (ADP and P_i), the conversion of bound ADP and P_i, and the release of the product (ATP) proposed by Boyer are shown in Figure 19.2.5.

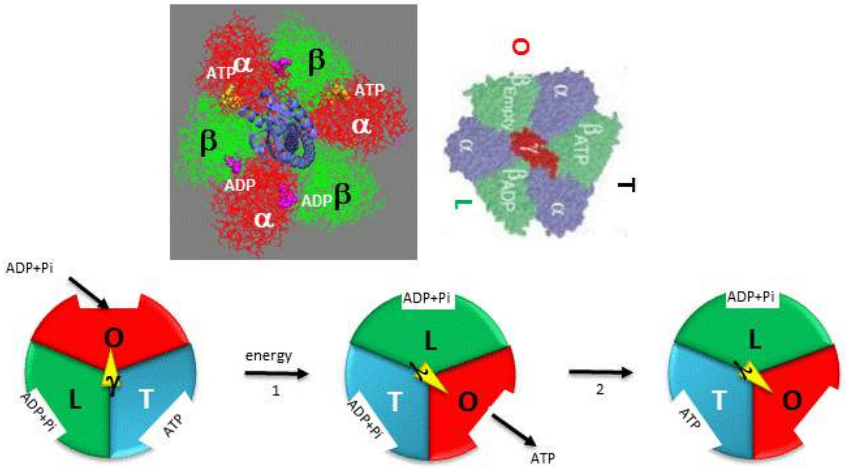


Figure 19.2.5: Boyer's three-state conformational model (L-O-T) for ATP synthesis

The collapse of the proton gradient (i.e. the proton-motive force) causes the γ subunit to rotate like a crankshaft relative to the F1 subunit, forcing the β subunit to change conformation from the T to the O (releasing ATP) and then to the L (binding ADP and Pi) states. The γ subunit does not appear to undergo any significant conformational change on ATP hydrolysis as evidenced by tritium exchange studies of amide protons.

To prove that the γ subunit rotates, you'd have to observe a single molecule. Since the γ subunit was too small to visually discern its rotation, Noji et al covalently attached a fluorescein-labeled actin filament to the γ subunit (near where F_o would bind). The whole F1 molecule was fixed to a glass slip through a His-tag such that the $a_3\beta_3$ ring was effectively immobilized. The γ subunit was free to rotate, which could be

detected by observing the fluorescence under a fluorescent microscope from the attached actin filament. This experiment and the outcomes are described in Figure 19.2.6.

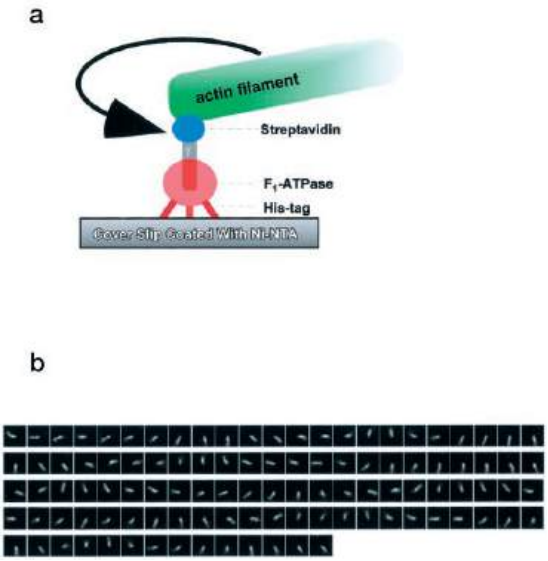


Figure 19.2.6: The direct observation of the γ rotation in the F_1 motor. Noji and Yoshida. JBC (2001) DOI:https://doi.org/10.1074/jbc.R000021200. Creative Commons Attribution (CC BY 4.0)

Panel **A** shows the experimental system for the observation of the γ rotation using an optical microscope. The F_1 motor tagged with 10 His residues at the N terminus of the β subunit was immobilized upside down on a coverslip coated with nickel-nitrilotriacetic acid (*Ni-NTA*). An actin filament (*green*) labeled with fluorescent dyes and biotins was attached to the biotinylated γ subunit (*gray*) through streptavidin (*blue*).

Panel **B** shows the rotary movement of an actin filament observed from the bottom, the membrane side, with an epifluorescent microscope. Length from the axis to tip, 2.6 μm; rotary rate, 0.5 revolutions per s; the time interval between images, 133 ms.

The actin filament rotated only in the presence of ATP. It rotated only counterclockwise, indicating that the motion was not random, but a specific motion of the γ subunit. At extremely low concentrations of ATP, rotation occurred only in 120^o increments, implying one step per molecule of ATP hydrolyzed. (Remember the β subunits are separated by 120^o). As the rotation occurs, there is viscous resistance to the movement of the actin filament. He calculated that for a single 120^o step caused by hydrolysis of a single ATP molecule, the amount of work was 80 piconewton which is about the free energy of hydrolysis of a single ATP molecule.

Later experiments in which a colloidal gold nanoparticle (40 nm diameter, with less frictional resistance to movement) was used instead of an actin filament showed the same result. At low [ATP], the motor rotates in 120^o steps. At high [ATP], the rotation rate becomes continuous and saturates (with Michaelis/Menten kinetics) at 130 revolutions per second.

Other experiments using immobilized ATPase and magnetic tweezers have addressed the timing of substrate binding and product release when the enzyme is run in reverse (ATP hydrolysis). On rotation of the γ subunit, the three binding sites change properties. In hydrolysis, ATP binds to the open site and helps promote the 120-degree rotation. In the next step, ATP is hydrolyzed. In the final step, products dissociate. Pi dissociation occurs last from the third site. Hence each of the 3 beta-binding sites has different roles. One binds the substrate, one performs catalysis and third releases products. Assuming the synthesis pathway is the reverse of the ATPase reaction, the final release of Pi in ATP cleavage predicts that Pi binds first in the synthetic direction. This would preclude the binding of ATP next which is critical since its concentration during synthesis can be 10x higher than that of ADP. As Pi is bound first, only ADP, not ATP can bind next.

The γ subunit rotation plays a "catalytic" role as its rotation induces cyclic conformational changes in the beta subunit of the synthase. Can ATP synthesis occur without the gamma subunit by a mechanism that involves a less proficient, but a concerted set of cyclic changes in beta subunit conformation? It can. Uchihashi et al have used high-speed atomic force microscopy (AFM) to study the $a_3\beta_3$ ring from the F1 subunit without the gamma subunit. They found that upon ATP hydrolysis, the beta subunits underwent conformational changes in the same counterclockwise rotary direction as when the gamma subunit was present. This is illustrated in Figure 19.2.7.

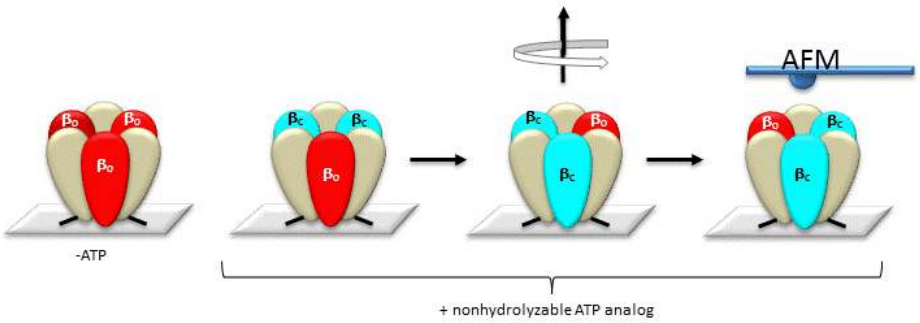


Figure 19.2.7: AFM Study of Conformational Changes in F1 "gammaless" subunit

These experiments conclusively show that the F1 subunit is effectively a rotary motor with the gamma subunit acting as a rotor in the stationary hexagon ring composed of the 3 pairs of alpha/beta subunits which acts as the stator (stationary part of an electric rotary motor).

The actual amino acids involved in the mechanism of ATP synthesis/hydrolysis are still not clearly defined but Glu 190 on the beta subunit acts as a general base. Figure 19.2.8 shows bound ADP and the proximity of Glu 188. Ala 158 is thought to move towards the active site after a conformational change, with the nonpolar methyl side chain displacing an adjacent water molecule which could leave as a product of ATP synthesis.

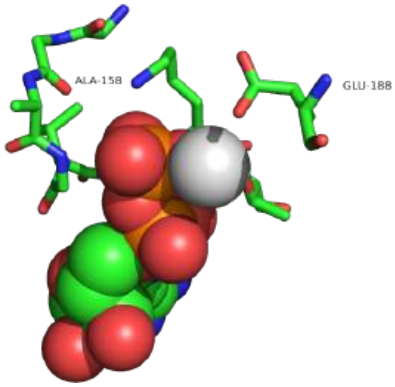


Figure 19.2.8: Active site residues for the synthesis of ATP. ADP is shown in spacefill, and CPK colors. Mg^{2+} is shown as a gray sphere (1E79).

19.2.3: F_O AND PROTON TRANSFER

The mechanism by which the proton gradient drives ATP synthesis involves a complex coupling of the F_o and F_1 subunits. Proton translocation occurs through the interface of the inner mitochondrial (or cell membrane of bacteria) membrane proteins C1 and A (or beta), whose structures from the bovine enzyme are shown again in Figure 19.2.8.

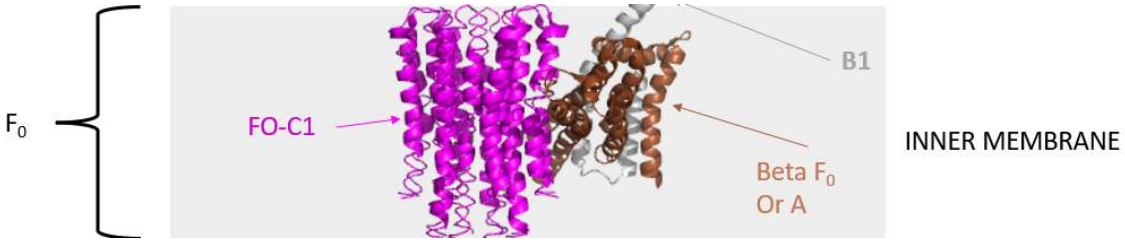


Figure 19.2.8: Inner membrane proteins C1 and A (or beta)

For protons to translocate through the membrane, there must be an aqueous channel on the matrix side and the lumenal side (bounded by the inner and outer mitochondrial membranes). So we have two understand two features:

- how the rotary axle part of the complex interacts with the C1 hexameric ring, causing the ring to rotate in the membrane.
- The nature of the water channels and the pathway for proton translocation.

The number of c monomers in the c-ring can vary between 8-17. For example, the bovine ATP synthase shown in Figures 1 and 2 has 8 c-monomers. Each has a critical proton donor and acceptor located near the a or beta-F_o chain. In the bovine case, it is glutamate 58. (The side chains in the bovine structure presented in Figure 2 do not show since they were not resolved in the cryoEM structure.) In E. Coli, which we will explore further below, it is aspartate 61. In yeast, it is at position 59. They are positioned near the center of the membrane helices. Mutations in these residues cause significant decreases in enzyme activity. An E56Q mutation in some bacilli species prevents proton pumping and ATP synthesis. Another key charged group, arginine, in the a or beta F_o protein is also key.

Figure 19.2.9 shows an interactive iCn3D model of the E. Coli A_1-C_{12} subcomplex of F_1F_O ATP Synthase (1C17).

Figure 19.2.9: E. Coli A_1-C_{12} Subcomplex OF F_1F_O ATP Synthase (1C17). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...sfRJBHEBiC5sy9

The key aspartate 61 (sticks and labeled) in each C ring subunit (red) and the proximity of two of the aspartate 61s to arginine 210 in the A subunit (brown) are evident.

Two classic inhibitors (structures shown below) of ATP synthase interact with the F_o subunit. One, oligomycin A, binds between the a and c subunits and blocks proton transport activity of the F_o subunit. The O protein at the top of the F_1 complex is also called the Oligomycin-sensitivity-conferring protein (OSCP), even though oligomycin does not bind there. The soluble F_1 by itself is not sensitive to oligomycin, but when it's linked to F_o in part through the O or OSCP peripheral stalk protein, it becomes sensitive to oligomycin.

Another inhibitor, dicyclohexylcarbodiimide reacts with a protonated Asp 61 in c subunits of F_o. It does so even at pH 8.0 which indicates that the pKa of the Asp 61 is much higher than usual. This might occur if the Asp is in a very hydrophobic environment. The modification of one Asp 61 in only one c subunit is necessary to stop Fo activity. The protonated carboxyl group donates a proton to a nitrogen atom in DCCD, which then reacts with the deprotonated Asp to form an O-acyl isourea derivative. Figure 19.2.10 shows the structures of oligomycin A and DCCD, inhibitors of proton transport by F_o.

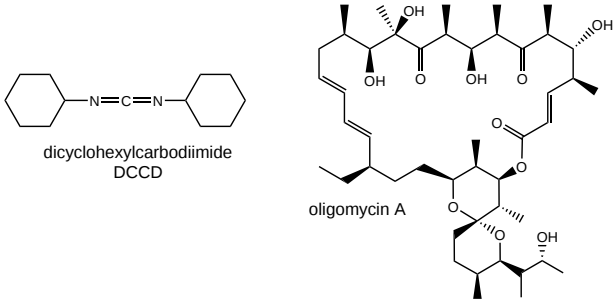
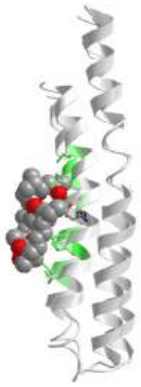


Figure 19.2.10: structures of oligomycin A and DCCD

Figure 19.2.11 shows an interactive iCn3D model of the F_1F_oATPase c10 ring with bound oligomycin (4F4S)



NCBI iCn3D Figure 19.2.11: E. Coli A_1-C_{12} Subcomplex OF F_1F_O ATP Synthase (1C17). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...pJw1pFr8jTPmK8

Only two of the c-subunits are shown interacting with one oligomycin. The actual complex has 10 c-subunits and 8 oligomycins bound. The oligomycin is shown in spacefill with CPK colors. The two c-subunits are shown in gray. The side chains from the two c-subunits interacting with oligomycin are shown as CPK-colored sticks. They are conserved between humans and yeast. Note the single H bond between the critical Glu 59 and oligomycin. The green colors indicate hydrophobic parts of the c-subunits which interact with the mostly nonpolar parts of the antibiotic.

How do these key Asp or Glu side chains participate in proton translocation across the membrane and c-ring rotation and how do they access the matrix and lumenal water channels?

We can start to understand the answer to these question by examining the structure of the c-ring and its interaction with the a-chain from *Bacillus* PS3 ATP synthase, as shown in Figure 19.2.12, which show the key Glu59 residues in each of the c subunits of the 10-mer ring along with Arg 176 and two other glutamates (Glu 223 and Glu 162) in the A ring.

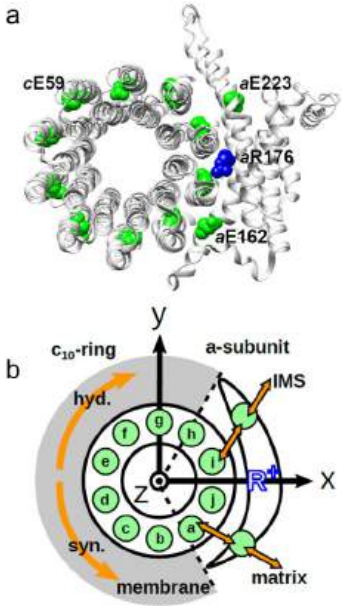


Figure 19.2.12: Schematic picture of the *a*-subunit and *c*-ring of F_o. Mitome et al. (2022). *eLife* **11**:e69096. https://doi.org/10.7554/eLife.69096. Creative Commons Attribution License

Panel (**a**) shows the ac_{10} part of the F_o region with the key amino acids shown as colored spheres (and numbered based on the yeast protein).

Panel (**b**) shows a cartoon of a simulation model developed by Mitome et al (ibid). The Glu 59s are shown as green circles and labeled a-j. The gray represents the membrane.

Protons move between the protonated Glu 59 in c subunits that are proximal to the deprotonated Glu 223 and Glu 162 in a hydrogen bond "handshake". The Glu 59s in the c chains in the 10-mer ring are likely to be protonated given their position in the center of the membrane, whereas the proximal Glu 223 and 162 are more likely to be initially deprotonated given their proximity to aqueous channels, where they

"hand off" their protons to complete the translocation from the intermembrane luminal space to the matrix. Now of course, once a c-ring Glu 59 is deprotonated, it has a negative charge.

Simulations suggest that there are 2-3 deprotonated Glu 59 with a negative charge proximal to the a-subunit. One is likely to take up a proton in a process leading to ATP synthesis. The orange arrows show the direction of proton flow. If the c-ring rotates in the clockwise direction, ATP is hydrolyzed while counter-clockwise rotation leads to ATP synthesis.

Now let's look at the water channel forming at the interface between the c-ring and a-subunit and an additional protein ASA6 from the unicellular green alga *Polytomella sp*. The interactions and nature of the water channels are shown in Figure 19.2.13.

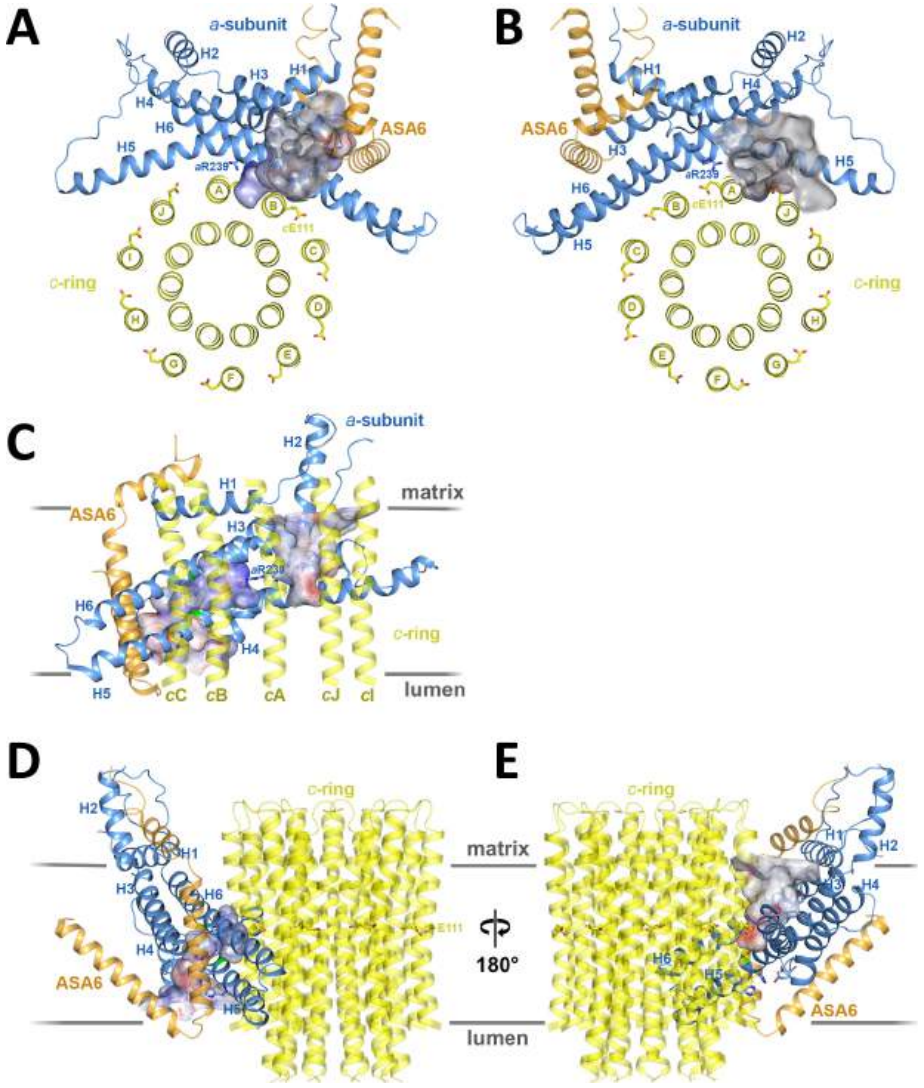


Figure 19.2.13: Two aqueous channels in F_o. Niklas Klusch, Bonnie J Murphy, Deryck J Mills, Özkan Yildiz, Werner Kühlbrandt (2017) Structural basis of proton translocation and force generation in mitochondrial ATP synthase eLife 6:e33274. https://doi.org/10.7554/eLife.33274. Creative Commons Attribution License

Subunit-a is shown in blue, the c_{10}-ring in yellow, and the associated ASA 6 protein in brick. The channels are shown as potential surfaces (red, negative; blue, positive; grey, neutral). (**A**) and (**B**) display a 5 Å slice of the c_{10}-ring at the level of the protonated *c*Glu111.

Panel(**A**) shows a 5 Å slice of the c_{10}-ring at the level of the protonated c-chain Glu111, with the lumenal channel seen from the crista lumen.

Panel (**B**) shows a 5 Å slice of the c_{10}-ring at the level of the protonated c-chain Glu111, with the matrix channel seen from the matrix.

Panel (**C**) shows a side view of both channels seen from the *c*-ring, with outer *c*-ring helices in transparent yellow. Lumenal channel, left; matrix channel, right. The strictly conserved a-chain Arg239 in helix 5 (H5) separates the lumenal and matrix channels.

Panel (**D**) shows that the lumenal channel passes through the H5/H6 hairpin at the small sidechains *a*Ala246, *a*Gly247 (H5), and *a*Ala292 (H6) (green).

Panel (**E**) shows H4, the N-terminal half of H5, and the connecting H4/H5 loop at the matrix channel.

Figure 19.2.14 a probable proton pathway through the F_o subcomple

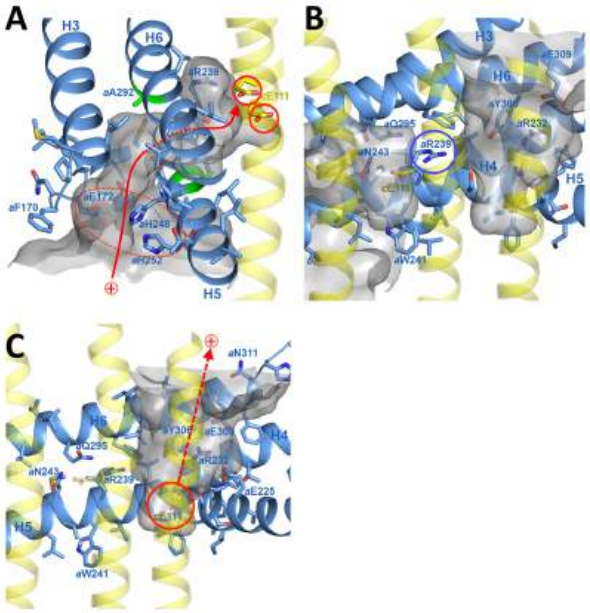


Figure 19.2.14: Proton pathway through the F_o subcomplex. Klusch et al, ibid.

The a-subunit is shown in blue, the adjacent *c*-ring helices in transparent yellow, and the aqueous channels in translucent grey.

Panel (**A**) shows the movement of protons (red arrow) in the lumenal channel. They likely move between Glu172, His248, His252, and Glu288 in the a-chain (dashed red ellipse) through the H5/H6 helix hairpin at the small sidechains of a-chain Ala246 and Gly247 (H5) and Ala292 (H6) (green) to c-chain Glu111 in the rotor ring *c*-subunits (red circles).

Panel (**B**) shows that a-chain Arg239 (blue circle) is located halfway between the lumenal channel on the left and the matrix channel on the right, forming a seal to prevent proton leakage. *c*-ring helices (transparent yellow) with *c*Glu111 are seen in the foreground.

Panel (**C**) shows the movement of protons (dashed red arrow) in the matrix channel as they pass directly from the deprotonated c-chain Glu111 to the pH 8 matrix.

A final summary view showing c-ring rotation powered by the collapse of the pH between the lumenal (intermembrane space) channel (pink) and matrix channel (light blue) is shown in Figure 19.2.15.

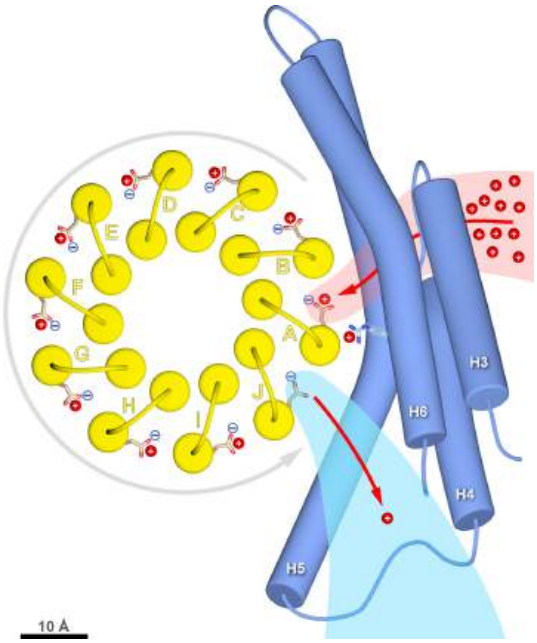


Figure 19.2.15: c-ring rotation is powered by the potential gradient between the lumenal channel (pink) and matrix channel (light blue). Klusch et al, ibid

The *c*-ring (yellow) and the membrane-intrinsic four-helix bundle of subunit *a* (blue) are drawn to scale as seen from the matrix. Protons (red) pass from the crista lumen below the projection plane through the lumenal channel between H5 and H6 to protonate Glu111 of *c*-subunit *A*, while subunit *J* is deprotonated by the higher pH of the matrix channel. The positively charged a-chain Arg239 is likely to interact with the deprotonated c-chain Glu111 during its short passage to the lumenal channel. The lumenal and matrix channels approach one another to within 5–7 Å. A protonmotive force of 200 mV between the closely spaced channels creates a local electrostatic field in the range of 40 million to 100 million V/m, depending on the protein dielectric. The field exerts a force on the deprotonated c-chain Glu111 that results in a net counter-clockwise rotation of the *c*-ring (grey arrow). Scale bar, 10 Å.

Click the images below to see (in a new window) a series of stunning videos showing proton translocation across the channel! (Niklas Klusch, Bonnie J Murphy, Deryck J Mills, Özkan Yildiz, Werner Kühlbrandt (2017) Structural basis of proton translocation and force generation in mitochondrial ATP synthase eLife 6:e33274. https://doi.org/10.7554/eLife.33274. Creative Commons Attribution License)

The three-dimensional arrangement of subunit *a* (blue), *c*-ring (yellow) with the lumenal channel in pink and matrix channel in light blue.
Arrangement of channel-lining sidechains for the lumenal channel. Sidechains in stick representation are colored as subunit *a*, blue; *c*-ring, yellow; ASA 6, brick; lipids, grey. Channels are shown as potential surfaces (red, negative; blue, positive; grey, neutral).
Arrangement of channel-lining sidechains for the matrix channel. Sidechains in stick representation are colored as subunit *a*, blue; *c*-ring, yellow; ASA 6, brick. Channels are shown as potential surfaces (red, negative; blue, positive; grey, neutral).

In summary, F_1F_oATP synthase is a rotary enzyme that ultimately couples the collapse of a proton gradient (a chemical potential gradient that contributes to the transmembrane electrical potential) to a chemical (phosphorylation) step. The rotor, which is in contact with both the F_o proton pore and the F_1 synthase, moves with respect to both subunits, which couples them. Both the axle and c-ring of F_o rotate. The interaction of the delta protein with the F_1 $a_3\beta_3$ hexamer does not cause $a_3\beta_3$ to physically rotate as a whole, as does the c-ring. Rather it causes a sequential rotation in the conformation of the three aβ dimers from the Loose to Open to Tight (LOT) conformations. The F_o pore can hence be considered an electrical motor and the F_1 synthase a chemical motor. Carrying the analogy of a motor even further, the F_o electrical motor turns the F_1 chemical motor into a generator, not of electricity but of ATP. The figure and link below, taken from the Protein Data Bank, go into more depth about this nanomotor.

The concerted movement of the c-ring and axle protein with respect to the not physically rotating $a_3\beta_3$ F_1 part can be seen in Figure 19.2.16, which is derived from cryoEM structure of different rotated states of bovine ATP synthase.

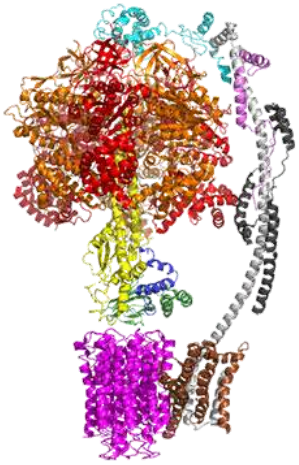


Figure 19.2.16: Structure and conformational states of the bovine mitochondrial ATP synthase (5ARA, 5ARE, 5ARH, 5ARI, 5FIJ, 5FIK, 5FIL)

A phenomena video is available showing the detailed steps in ATP synthesis by ATP synthesis Click on Figure 19.2.17 to get the link to the video.

Figure 19.2.17: ATP synthase in action. Muzzey and Lue, HHMI Institute. For educational and non-commercial use only.

This amazing enzyme can be purified and reconstituted in a lipid vesicle along with bacteriorhodopsin. If these proteins are oriented in the correct fashion, light can cause proton translocation into the lumen of the vesicle creating a transmembrane potential. The proton gradient can collapse as protons move from the lumen through the F_o part of F_1F_o ATP synthase, causing ATP synthesis from ADP and Pi in the outside solution. This is shown in Figure 19.2.18.

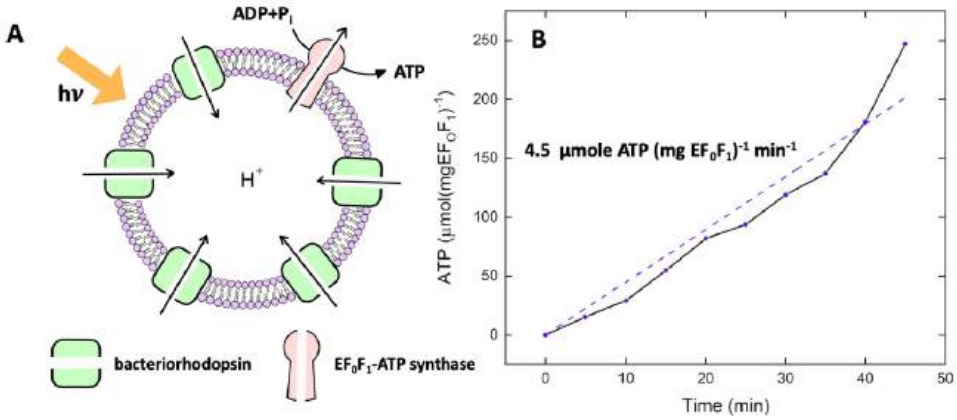


Figure 19.2.18:. Light-driven ATP production. Ahmad et al. *ACS Synthetic Biology* 2021. https://pubs.acs.org/doi/10.1021/acssynbio.1c00071. https://creativecommons.org/licenses/by/4.0/

Panel(A) shows a cartoon of the light-driven ATP synthesis in lipid vesicles. Panel (B) shows the increase in [ATP] over time.

19.2.4: PROTON GRADIENT COLLAPSE AND ATP SYNTHESIS - THERMODYNAMICS

Mathematical analyses show that the free energy change on proton gradient collapse can easily power the endergonic synthesis of ATP. Consider a typical pH gradient (-1.4 pH units) across the inner membrane of respiring mitochondria (with the outside having a lower pH than the inside making the inside more depleted in protons). There is a chemical potential difference in protons across the membrane. However, another factor determines the thermodynamic driving force for proton translocation across the membrane. A transmembrane potential exists across the inner membrane of the mitochondria, as it does across most membranes. The source of the membrane potential will be discussed in the signal transduction chapter (Chapter 28). The inside is more negative than the outside, giving the membrane a transmembrane electrical potential. of about -0.14 V. Clearly, protons would be attracted to the other side of the membrane (into the matrix) by this potential difference, which then augments the chemical potential difference as well. A simple mathematical derivation shows that indeed, a proton gradient can supply enough energy for ATP synthesis, especially when coupled to a transmembrane electrical potential. We reshow a small image of the complex and its activities in Figure 19.2.19 to help you through this section

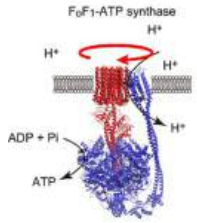


Figure 19.2.19: ATP synthase and its activities. Noji, H., Ueno, H. & Kobayashi, R. Correlation between the numbers of rotation steps in the ATPase and proton-conducting domains of F- and V-ATPases. *Biophys Rev* **12,** 303–307 (2020). https://doi.org/10.1007/s12551-020-00668-7. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Chemical Potential (use G instead of μ)

Since most students use free energy and not chemical potential (which is just free energy /mole at constant T and P, we will derive these equations using free energy G.

Let's consider the free energy of the protons on the outside (low pH, higher concentration) and in the inside (matrix, higher pH, low concentration):

$$\mathrm{H^+}_{\mathrm{out}} \longrightarrow \mathrm{H^+}_{\mathrm{in}}\ \mathrm{G} = \mu = \text{ chemical potential (kcal/mol)} \tag{19.2.1}$$

Now express this as the change in free energy, ΔG

$$\Delta G_{\mathrm{chem}} = G_{\mathrm{in}} - G_{\mathrm{out}} = G^0 + RT \ln \left[H^+\right]_{\mathrm{in}} - \left(G^0 + RT \ln \left[H^+\right]_{\mathrm{out}}\right) \tag{19.2.2}$$

This can be rewritten as:

$$\Delta G_{\mathrm{ohem}} = \mathrm{RT} \ln \frac{[H^+]_{\mathrm{in}}}{[H^+]_{\mathrm{out}}} = 2.303RT \left[-\log \left[H^+\right]_{\mathrm{out}} - \left(-\log \left[H^+\right]_{\mathrm{in}}\right)\right] \tag{19.2.3}$$

Hence

$$\Delta \mathrm{G}_{\mathrm{chem}} = 2.303 \mathrm{RT}\left(\mathrm{pH}_{\mathrm{out}} - \mathrm{pH}_{\mathrm{in}}\right) = 2.303 \mathrm{RT} \Delta \mathrm{pH} \tag{19.2.4}$$

In respiring mitochondria, ΔpH = -1.4, so

$$\Delta \mathrm{G}_{\mathrm{chem}} = 2.303 \mathrm{RT} \Delta \mathrm{pH} = 2.303 (1.99 \mathrm{cal/mol \cdot K})(298\ \mathrm{K})(-1.4) = -1.91 \mathrm{kcal/mol H^+} \tag{19.2.5}$$

Electrical Potential

Consider the relationship between ΔG and the transmembrane electrical potential

$$\Delta G_{\mathrm{elect}} = +ZF\Delta\psi \tag{19.2.6}$$

where Δψ is the transmembrane potential, Z=+1 (charge on a proton) and F is the Faraday constant

This equation is similar to

$$\Delta G \text{ elect } = -nF\Delta E \tag{19.2.7}$$

where ΔE is the cell potential or EMF. If ΔE is + and Δψ is -, then ΔG is -.

In respiring mitochondria, Δψ = -0.140 V (remember that 1 V = 1 J/C). Hence

$$\Delta G_{\text{elect}} = +ZF\Delta\Psi = 1(96.485\text{C/mol})(-0.140\text{ J/C}) \times (1\text{cal}/4.18\text{ J}) = -3.23\text{kcal/molH}^{+} \tag{19.2.8}$$

Electrochemical potential

Now let's add the equations to get the full driving force, the electrochemical potential:

$$\Delta G_{\text{tot}} = 2.303\text{RT}\Delta\text{pH} + \text{ZF}\Delta\psi = -5.15\text{kcal/molH}^{+} \tag{19.2.9}$$

This gives

$$\frac{\Delta G_{\text{tot}}}{F} = \frac{2.303RT\Delta pH + +zF\Delta\psi}{F} \tag{19.2.10}$$

Hence

$$\Delta p = \Delta\psi + \frac{2.303RT\Delta pH}{F} = \Delta\psi + 0.059\Delta pH = 0.224\text{ V} \tag{19.2.11}$$

where Δp is the protonmotive force.

19.3: REGULATION OF OXIDATIVE PHOSPHORYLATION

19.3.1: INTRODUCTION

The main function of oxidative phosphorylation (oxphos) is to produce, under aerobic conditions, lots of ATP. It should make sense that to a first approximation, the regulation of oxphos depends primarily on the energy state of the cell, which is reflected by the ADP/ATP ratio. In contrast to glycolysis and the citric acid cycle, which produce NADH, the primary and first electron carrier in oxphos, the regulation does not depend on the binding of allosteric effectors to key enzymes in the pathway. Likewise, hormonal regulation is not involved. Perhaps this occurs since the oxphos pathway is downstream from the glycolytic and citric acid pathways, where allosteric and hormonal regulation robustly affects their outputs.

If you consider the primary reactants in electron transport to be NADH, $FADH_2$, O_2, ADP, and P_i and if you assume that O_2 levels are constant, ADP appears to be the main molecule that determines the rate of the oxphos reactions. This is in accordance with what we have previously described: AMP, ADP, and ATP levels regulate many of the key enzymes in the glycolytic and citric acid pathways, as reviewed in Figure 19.3.1.

Coordinate regulation of ATP production:

Regulation by:

- [ATP]
- [ADP]- the "PO_4" acceptor
- Mass action ratio, [ATP]/[ADP][P

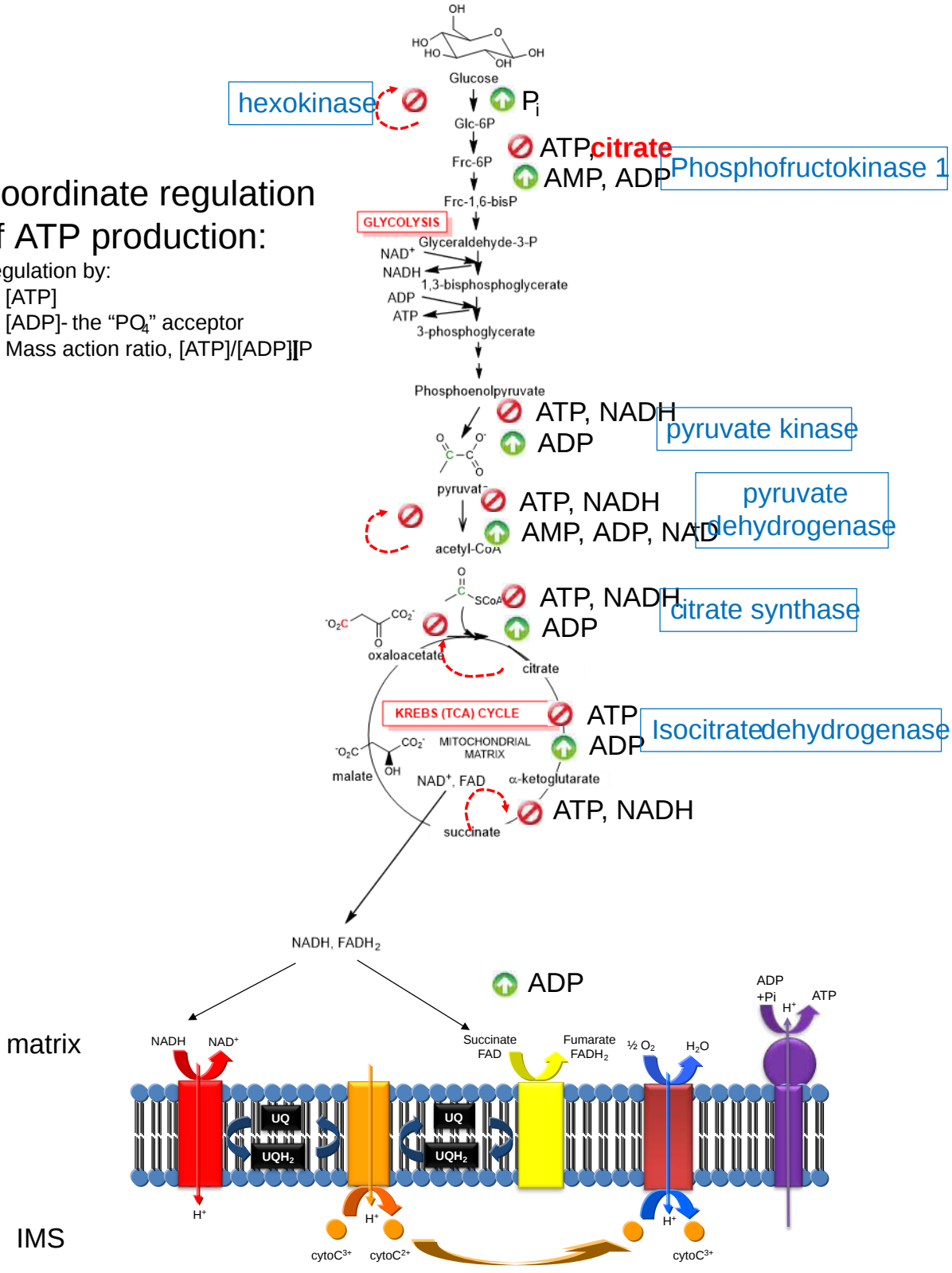


Figure 19.3.1: Regulation of glycolysis, the citric acid cycle, and oxphos by ATP, ADP, and AMP.

Overall there are several levels of oxphos regulation, which occurs at the level of the kinetics of the electron transport chain complexes, the efficiency of electron and proton transfer, the structure of the mitochondria, and mitochondria formation and their degradation.

19.3.2: COUPLING AND UNCOUPLING OF ELECTRON AND PROTON TRANSPORT

What is key is that the regulation of oxidation (electron transport) is coupled to the regulation of ATP synthesis (phosphorylation). These processes can also be uncoupled if electrons "leak" away from their transfer from NADH to O_2, which leads to the production of harmful reactive oxygen species (ROS) such as superoxide. Proton can also "leak" away from the proton gradient if they bypass the ATP synthase complex and pass through the inner membrane with the help of specific carriers and protein uncouplers, as shown in Figure 19.3.2.

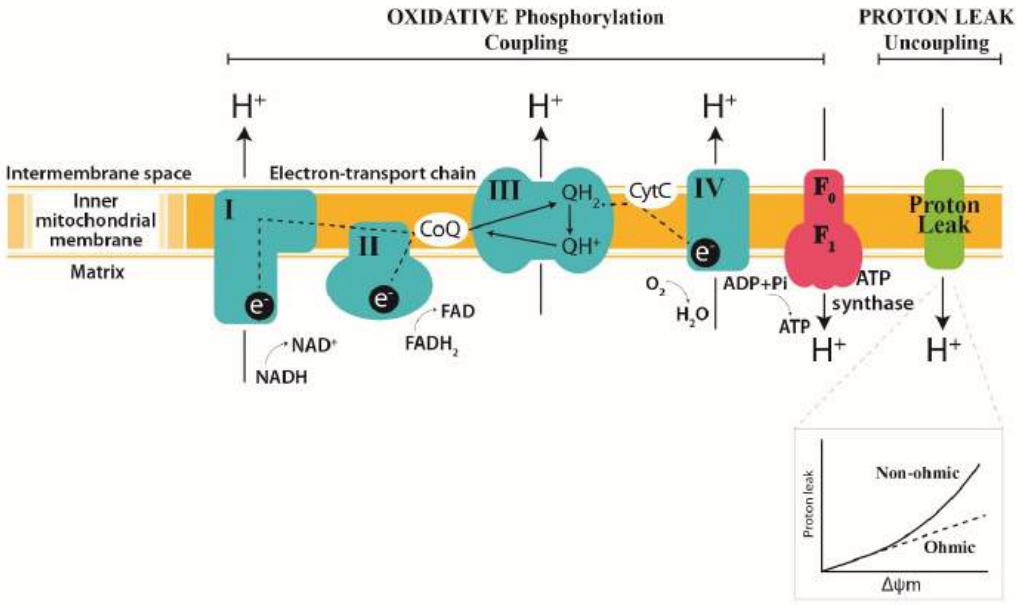


Figure 19.3.2: Coupling and uncoupling of oxidative phosphorylation. Eyenga, P.; Rey, B.; Eyenga, L.; Sheu, S.-S. Regulation of Oxidative Phosphorylation of Liver Mitochondria in Sepsis. *Cells* **2022**, *11*, 1598. https://doi.org/10.3390/cells11101598. **Creative Commons Attribution License**

The far right light green proton and the insert show the proton leak's kinetics. The dotted line indicates the electron pathways.

There are two types of proton leaks:

- a basal leak through mitochondrial anion carriers in a non-regulated and minor process.
- an inducible leak through specific proteins, adenine nucleotide translocase (ANT), and uncoupling proteins (UCPs), which leads to loss of energy as heat. They can be activated by fatty acids as well as ROS. A particular UCP, UCP1, is found most abundantly in brown fat in mammals.

The basal leak occurs in the presence of inhibitors of ANT (carboxyatractylate) and UCP1 (GDP). It can be significant (about 25% of the resting metabolic rate of liver cells and up to 50% in respiring skeletal muscle cells). It is less understood than inducible leaks and does appear to be attributed to ANT in some fashion.

When oligomycin, which inhibits the collapse of the proton gradient through ATP synthase, is added to mitochondria, the proton gradient (protonmotive force, Δp) decreases. The rate of leakage increases exponentially with increasing Δp, so it represents **non-Ohmic leakage**. Ohmic leakage would give a linear dependence on Δp. This can be seen from a rearrangement of Ohm's Law (V=IR) to give I = V/R = conductance(V), where V is voltage, I is the current (analogous to the leak), R is resistance and 1/R is conductance. The exponentially increasing non-Ohmic leakage (usually measured by O_2 uptake in the presence of oligomycin) shown in the inserted graph in Figure 2, is not associated with non-specific membrane damage.

A particular uncoupling membrane protein, UCP1, is especially prevalent in hibernators, small mammals, and human infants, which allows them to stay warm. The UCP family is activated by fatty acids, which may act as required cofactors for the protein or as antagonists of the inhibition of UCPs by nucleotide diphosphates.

Two methods are used to determine the efficiency of coupling of oxidation and phosphorylation.

- One is to determine the ADP/O ratio as both are substrates for oxphos, where O represents the amount of O_2 consumed in the conversion of a defined amount of ADP to ATP when ADP is at high levels (which might not be physiologically relevant). This would give the maximum phosphorylation rate (defined as state 3).

- A second is to measure ATP (or ADP)/O ratio directly through methods that give spectrophotometric signals such as the production of NADH by a coupled hexokinase/glucose-6-phosphate dehydrogenase assay in which ATP is used in the following reactions to make NADH, as shown in Figure 19.3.3.

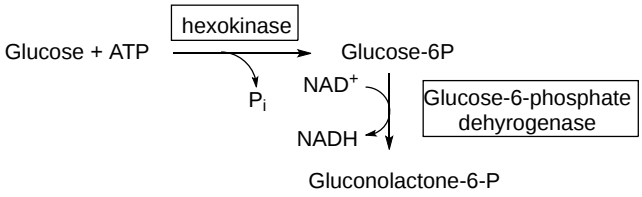


Figure 19.3.3: Generation of NADH on the combined reaction of hexokinase and glucose-6-phosphate dehydrogenase.

In this coupled assay, one ATP used produces one NADH, which absorbs at 340 nm. Also, the ATP used as a substrate in the coupled reaction has left the mitochondria for it to have access to hexokinase, so the internal ratios of ADP/ATP in the respiring mitochondria are unchanged during the time course of the reaction. This method also allows the measurement to be made under nonsaturating ADP concentrations.

Uncoupling of electron transport and ATP synthesis can occur through a combination of mechanisms, including proton loss at ANT and UCPs, at the complexes in electron transport that transport electrons and protons, and through ATP synthase itself.

Other mechanisms would include substrate depletion or effects on the **mitochondrial permeability transition pore (mPTP)**, which is an inner membrane transmembrane protein, which is gated open by increases in matrix Ca^{2+}, adenine nucleotide decreases and increases in Pi. It is also a voltage-gated channel. When open the mitochondrial membrane depolarizes, ATP is depleted and cell death can be triggered. We'll discuss this complex later.

Figure 19.3.4 shows the chemistry used to analyze the efficiency of oxphos and their graphical results.

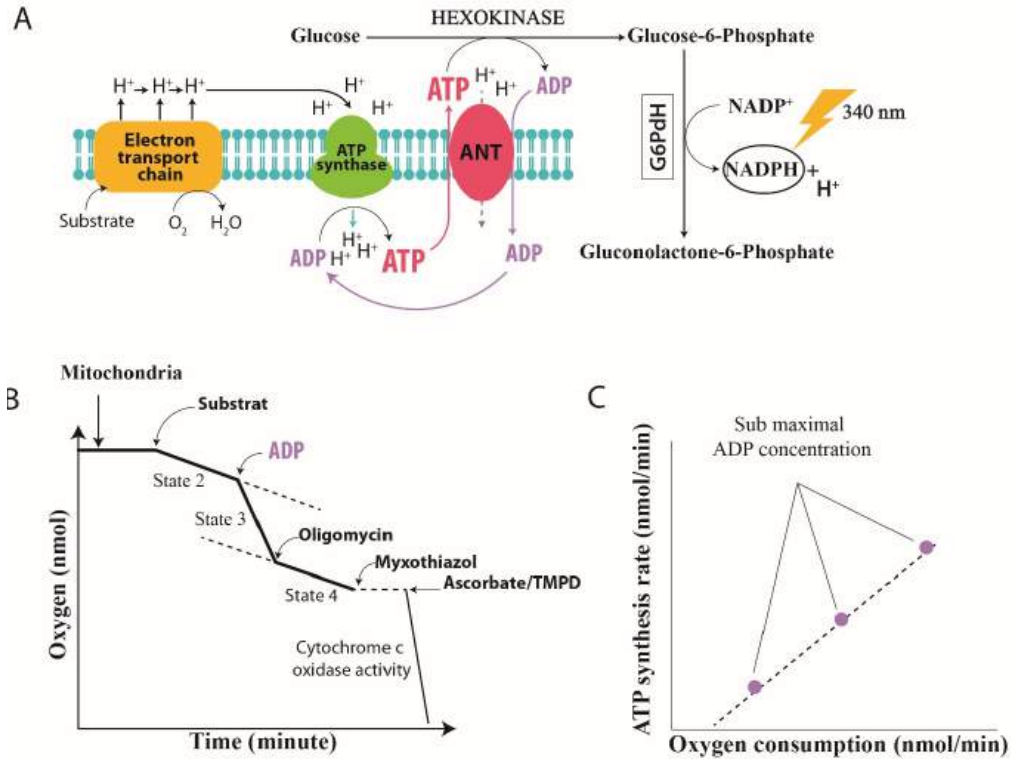


Figure 19.3.4: Measurement of coupling efficiency in respiring mitochondria. Eyenga et al. Ibid.

Panel (**A**) shows the chemistry involved in the determination of the ATP/O ratio to measure coupling efficiency. Note that this step allows the regeneration of ADP and efficiency can be measured at any ADP concentration.

Panel (**B**) shows changes in O_2 with time. The respiration rate is given by the slope of the curve. Until substrate is added, O_2 levels are constant. On the addition of substrate, O_2 is consumed in electron transport so levels O_2 levels fall linearly. When ADP is added, the rate of O_2 consumption increases which is reflected in a steady linear decline in O_2 levels. This state which is at fixed ADP levels (as illustrated in Panel (A)), is called state 3 (maximum ATP synthesis at a given ADP level). When oligomycin, an inhibitor of proton flow across ATP

synthase which inhibits ATP synthesis (and ADP use) is added, state 4 ensues with shows decreased O_2 consumption when ATP synthesis is stopped. Note the parallel lines for stages 2 and 4. The Cytochrome c oxidase activity at the end of the curve is the maximal respiration with ascorbate (as an oxidizing agent) and N,N,N′N′-tetra methyl-p-phenylenediamine (TMPD) as a substrate.

Panel (**C**) shows a graph of the rate of ATP synthesis vs the rate of O_2 consumption. The slope gives the rate of ATP generation to that of O_2 consumption (ie. the ATP/O ratio). Note that the graph is linear:O is linear over a range of fixed, non-saturating ADP concentrations.

19.3.3: CHANGES IN PROTON GRADIENT FROM "SLIPPAGE"

Slippage in the electron transport chains: Redox Slipping

Proton translocation from the matrix to the inner membrane space is an integral part of the function and activity of the electron transport chains that establish the pH gradient and hence the Δp. These components are Complexes I, III, and IV. Complex IV (cytochrome C oxidase - CCO) appears to be especially important. Protons are removed by CCO from the matrix on the reduction of O_2 to H_2O. This requires exactly 4 electrons and 4 protons so the stoichiometry is fixed. Also, CCO moves protons into the inner membrane space in the formation of the pH gradient. This is where the electron-to-proton ratio may vary depending on redox and coupling conditions. Alterations in CCO or even decreases in its concentration would impact the movement of protons as well. Instead of calling these alterations in nominal proton translocation leakages, they are called slippage since an ordinary function of CCO is proton translocation. Slippages through CCO would decrease proton transport and decrease ATP/O. Reactive nitrogen species (NO and peroxynitrite) can also decrease ATP/O.

Slippage at ATP synthase

Proton leakage at ATP synthase would also result in a decreased ATP/O ratio. This occurs especially if nucleotides are low or if the holoenzyme is compromised by low subunit availability. ATP synthase forms dimers and other multimers to form supercomplexes, and this affects (or is determined by) the formation of invaginations in the inner membrane which form cristae.

Given that the multimeric c-ring of F_0 effectively moves protons across the inner membrane, it is a likely source of slippage. It most likely occurs through a modified version of the F_0 ATP synthase unit called the **ATP synthase c-subunit leak channel (ACLC).** This is different from the mitochondrial permeability transition pore (mPTP) mentioned above. Alternatively, it may be that the ACLC might be part of the leak channel of the mitochondrial permeability transition (MPT) pore discussed above. Studies show that the c-ring can form large conductance channels that are voltage-gated without regulatory subunits. In addition, added F_1 subunit inhibits it. In neurons stimulated by the excitatory neurotransmitter glutamate, F_1 dissociates from the F_1F_0 complex, allowing the c-ring to slip protons through the membrane. Finally knockdown of the c-subunit eliminates high conductance. Hence the c-ring is part of a type of mPT that is regulated by cyclophilin D (CypD), which is a mitochondrial peptidyl-prolyl cis-trans isomerase, that regulates the mitochondrial permeability transition pore (PTP).

In vitro, c-rings incorporated into vesicles form a large voltage-gated channel that is inactivated by the addition of F_1. Also in neurons stimulated with the excitatory neurotransmitter glutamate, which leads to increases in cytoplasmic Ca^{2+}, binding and ensuring conformational changes in the complex leads to dissociation of F_1 from the complex, freeing F_0 for proton "slippage" and transport uncoupled from ATP synthesis.

A Proposed “bent-pull-twist” model of ACLC gating in physiological and pathological conditions is shown in Figure 19.3.5.

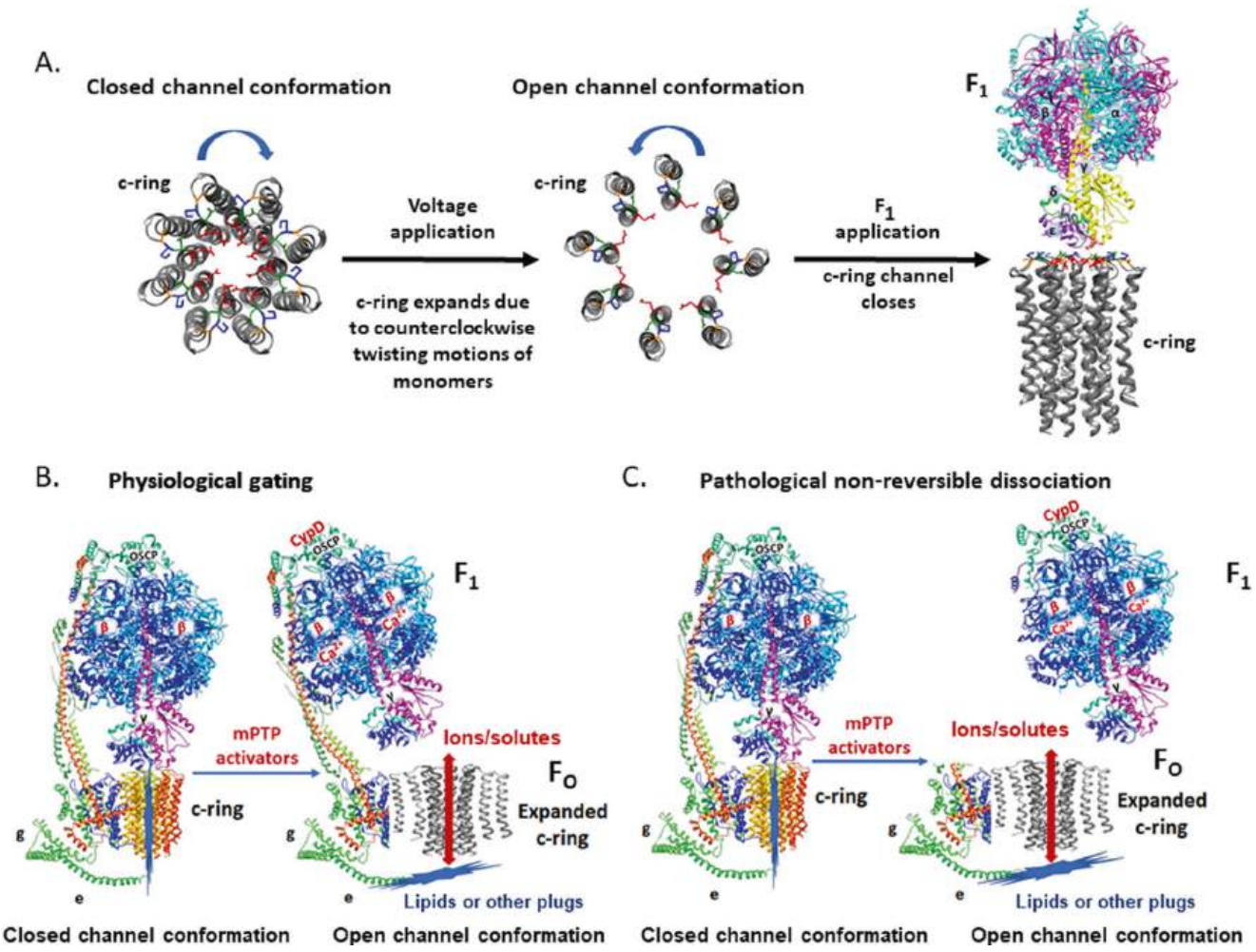


Figure 19.3.5: B**ent-pull-twist model of ACLC gating in physiological and pathological conditions.** Mnatsakanyan, N., Park, HA., Wu, J. *et al.* Mitochondrial ATP synthase c-subunit leak channel triggers cell death upon loss of its F_1 subcomplex. *Cell Death Differ* (2022). https://doi.org/10.1038/s41418-022-00972-7. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Panel A shows c-ring diameter changes during the electrophysiology recordings, upon the application of voltage and F_1. F_1 added during c-ring recordings probably inactivates the channel due to the specific interactions between F_1 and c-ring that induce twisting motions in c-subunit monomers in a clockwise direction, to stabilize the ring, reduce the pore diameter and close the channel.

Panel B shows reversible brief openings of ACLC in physiological conditions.

Panel C shows that non-reversible dissociation of F_1 from F_O occurs during long-lasting openings of the c-ring channel in severe pathology. For simplicity, only ATP synthase monomer is shown. ATP synthase subunits are drawn as ribbon representations (modified PDB ID code: 6J5I) [14]. In **B** and **C**, red arrows indicate the path of ion flow through the channel. Closed and open conformations of the channel are noted.

19.3.4: A CLOSER LOOK AT UNCOUPLER PROTEIN 1 (UCP1)

UCP1 is an inner membrane protein in the mitochondria that bypasses the movement of protons through the c-ring of F_1. UCP1 is activated by fatty acids and inhibited by purine nucleotides. The mechanism of proton translocation is unclear. It could arise from a "handshake" of protons through a series of linked hydrogen bond donors and acceptors in the membrane protons. Evidence suggests that what moves are bound and protonated fatty acids which flip across the inner membrane where the carboxyl group deprotonates in the high pH matrix side of the membrane.

The role of UCP1 in this process could be to flip the deprotonated fatty acid back across the membrane. This model is called the protonophoretic model in which the fatty acid acts as a "protonophore" much like valinomycin, which transports K^+ across the membrane and is called an ionophore. UCP1 doesn't interact with the H^+ directly but just flips the deprotonated fatty acid back to continue the cycle. This model would give a stoichiometry of one flipped negatively charged fatty acid per proton transferred.

In the second "H^+-shuttling" model, UCP1 is considered a fatty acid/H^+ symporter (carries both species in the same direction). In this model, a fatty acid is strongly bound to UCP1 with the head group inside a cavity in the protein. Both models required UCP1 to carry deprotonated fatty acids back across the membrane.

The NMR structure of a similar protein, UCP2, is known. It can bind fatty acids with the acyl tail near the matrix side. However, UCP2 does not seem to be involved in thermogenesis and doesn't uncouple oxphos. It binds and catalyzes the exchange of malate, oxaloacetate, and aspartate for phosphate and H and hence has other metabolic roles. UCP 1 and 2 are about 60% homologous but have similar alpha-helical structures based on modeling.

Perturbation of NMR signals from assigned side chains was used to show the fatty acids bind to UCP1 and it's closer to the matrix than the inner membrane space. Two key lysines (K56 on helix 1 and K269 at the matrix side of helix 6) probably bind with the carboxylate end of the fatty acid. UCP1also has a hydrophobic grove (T31, F32, L34, D35, L278, G279, S280, W281, and V283) to accommodate the fatty acid. In addition, both UCP1 and UCP2 transport/flip alkyl sulfonates, analogs of fatty acids. This is consistent and required with both models.

Figure 19.3.6 shows an interactive iCn3D model of the predicted AlphaFold model of human UCP1 (P25874). The subunits are shown in the same color as Figure 1 above.

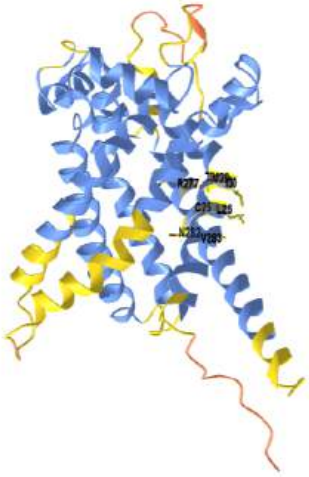


Figure 19.3.6: Predicted AlphaFold model of human UCP1 (P25874). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...Qf8W1MgVH6u546

The blue colors represent assigned structures with the highest predicted confidence, and the red the lowest. The amino acid side chains that are implicated in binding to fatty acids based on perturbation of NMR signals on fatty acid binding (http://dx.doi.org/10.1016/j.str.2017.07.005) are shown in sticks, with CPK colors and labeled. The membrane would be positioned perpendicular to the average vertical orientation of the helices in the above orientation.

19.3.5: ADDITIONAL CONTROLS

From the perspective of moving electrons into the electron transport pathway, the key deliverers of electrons are NADH derived from integrated multiple pathways at Complex I (as illustrated in Figure 1) and $FADH_2$ from the citric acid cycle at Complex II. Hence if intermediates in the citric acid cycle are low, the flux through electron transport is compromised. Flux through Complex I seems to determine the rate-limiting steps for oxygen use.

In the previous chapter section, we calculated the total chemical potential driving force across the membrane. A pH difference of 1.4 across the membrane provides a -1.9 kcal/mol (-8 kJ/mol) H^+ chemical potential compared to the -3.23 kcal/mol (13.5 kJ/mol) from the transmembrane potential, $\Delta\psi_m$. Both contribute to the thermodynamically favorable movement of H^+ across the inner membrane in respiring mitochondria, but $\Delta\psi_m$ is a large contributor. Hence $\Delta\psi_m$ and ATP synthesis are also coupled with maximal synthesis occurring between -100 to -150 mV. At values around -200 mV, give rise to increased permeability for protons (which decreases Δp) but also the production of superoxide and peroxides. Hence $\Delta\psi_m$ is also regulated.

ADP levels control ATP synthesis as well, with higher concentrations leading to increased ATP synthesis and a decrease in Δp. Decreased ATP synthesis from increased Δp affects coupling, slippage in cytochrome C oxidase, and heat generation.

Reduction of O_2 to water by stepwise addition of electrons involves the generation of superoxide, O_2^-. As shown in an earlier section, superoxide formed by Complex I would move to the matrix, while those formed by Complex III to both sides of the membrane. Hydrogen peroxide can also form from monoamine oxidase and α-ketoglutarate dehydrogenase, which produces increased peroxides with increasing

NADH concentrations. Respiratory state 4 (when proton transport is blocked by oligomycin) leads to higher levels of ROS. Hence ROS levels are also determined by not only the respiratory state but also Δp.

Other control mechanisms are potentially derived from other protein subunits that bind to the complex within a lipid bilayer. These include A6L, DAPIT (diabetes-associated protein in insulin-sensitive tissues), and the proteolipid 6.8PL. Another is the ATPase inhibitory factor 1 (IF1), which turns out to be a significant regulator. When the $\Delta\psi_m$ falls too low, IF1 inhibits ATP depletion which leads to ATP hydrolysis by the complex to move protons from the matrix to the inner membrane space. The interaction of IF1 with ATP synthase is shown in Figure 19.3.7.

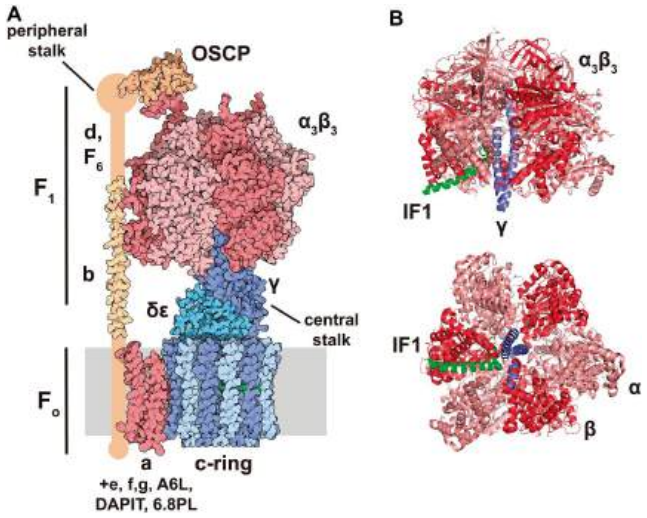


Figure 19.3.7: Structure of the bovine H^+-ATP synthase and binding site of IF1. Esparza-Moltó, P.B., Nuevo-Tapioles, C. & Cuezva, J.M. Regulation of the H^+-ATP synthase by IF1: a role in mitohormesis. *Cell. Mol. Life Sci.* **74,** 2151–2166 (2017). https://doi.org/10.1007/s00018-017-2462-8. Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/)

Panel **(a)** shows the soluble F_1-ATPase domain composed by the 3α3β subassembly (*salmon*/*red*) and γ (*dark blue*), δ and ε (*light blue*) subunits, while the membrane-embedded F_o domain is formed by subunit a (*red*) and a ring of 8c subunits (*light*/*dark blue*). Both domains are linked together by a central stalk (γ, δ, ε subunits of the F_1 domain) and a peripheral stalk (b, d F6, A6L, OSCP subunits, in *orange*). The 3D structure of the peripheral stalk is not fully resolved. Except for a and A6L subunits, the remainder is encoded in the nucleus. (model comprised of multiple PDB structures)

Panel **(b)** shows the lateral and basal view of the bovine F_1 domain (α subunit is shown in salmon, β in *red*, and γ in *blue*) complexed with a fragment of IF1 (*green*). IF1 binds to the αβ interface through residues 1–37 and also contacts the γ subunit. (PDB: 1OHH)

IF1 has an intrinsically disordered N-terminal domain that binds to ATP synthase between the α and β subunits of the F_1, where it adopts an alpha-helical conformation. It can dimerize through self-interactions with the C-terminal domain. Interestingly, bound IF1 inhibits both ATP synthesis and hydrolysis as well as H^+ translocation through the complex. These activities are illustrated in Figure 19.3.8.

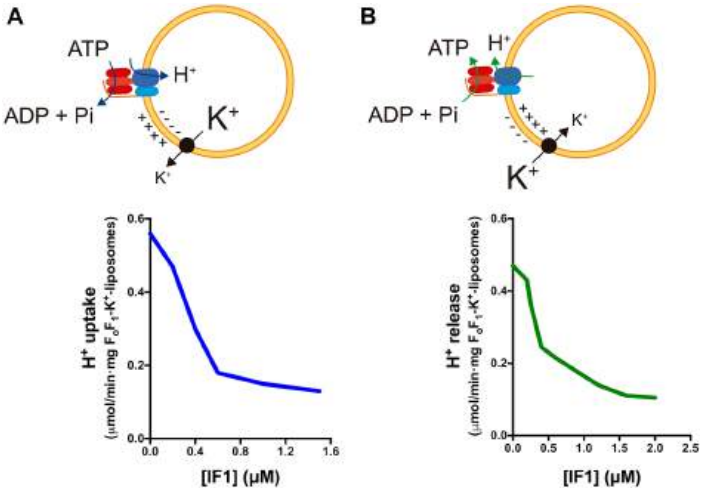


Figure 19.3.8: IF1 inhibits H^+ translocation through the H^+-ATP synthase in both synthetic and hydrolytic modes. Esparza-Moltó, ibid.

Panel **(a)** shows H^+ uptake is induced by valinomycin-mediated K^+ release from F_oF_1-K^+ liposomes with the H^+-ATP synthase functioning in the hydrolytic mode.

Panel **(b)** shows H^+ release is induced by valinomycin-mediated K^+ uptake in F_oF_1-K^+ liposomes with the H^+-ATP synthase functioning in the synthetic mode. The rates of H^+ uptake **(a)** and H^+ release **(b)** are reduced when the liposomes are incubated with increasing concentrations of isolated IF1. *Black circle* valinomycin.

Five key histidines in IF1 play a role in the inhibition of the ATPase activity and the oligomerization of IF1. When the pH gradient collapses, and the matrix becomes more acidic (as seen in hypoxia and ischemia), IF1 form dimers that inhibit ATP hydrolysis. At higher pHs, IF1 forms a dimer of IF1 dimers, or effectively a tetramer, which prevents its binding to ATP synthase and hence prevents the inhibition of ATP synthesis. An IF1-H49K mutant is an active inhibitor even at higher pHs. These properties are illustrated in Figure 19.3.9.

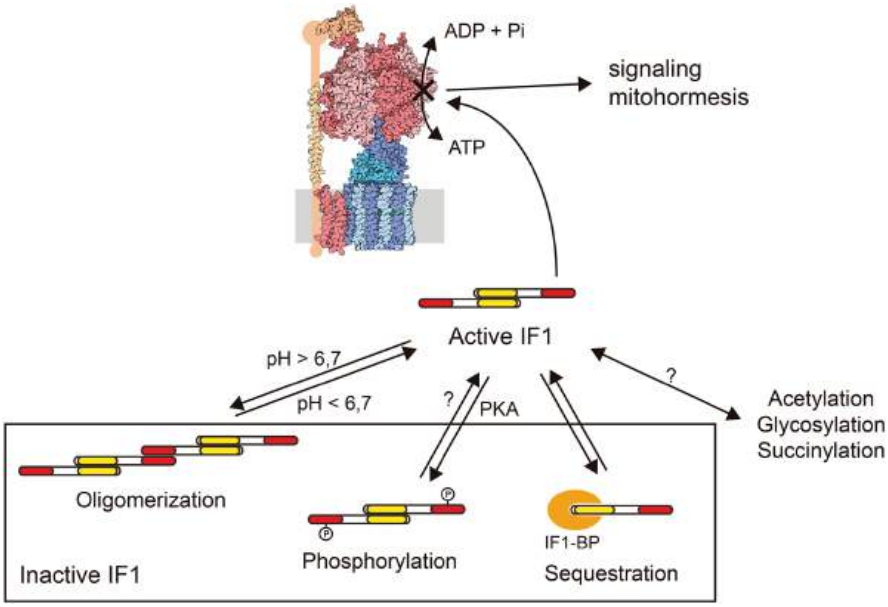


Figure 19.3.9: Post-translational regulation of IF1 activity. Esparza-Moltó, ibid.

IF1 inhibition of ATP hydrolysis is not only relieved by pH-dependent oligomerization as illustrated above, but also by phosphorylation by a mitochondrial Protein kinase A-like molecule, or by its interaction with yet another protein, BP. A conserved serine 39 is important in its inhibitory action. When dephosphorylated, IF1 binds to ATP synthase and inhibits both ATP synthesis and hydrolysis. When phosphorylated it can't bind and hence can't inhibit ATP synthase.

19.3.6: MITOCHONDRIA - NUCLEAR SIGNALING

We would be remiss to not include the role of the mitochondrial genome and its relationship to the nuclear genome in the control of mitochondrial function, including oxidative phosphorylation. Mitochondrial arose from the internalization of a distant bacteria into a eukaryotic cell. They developed an endosymbiotic relationship in which some of the genes from the original bacterial species remain in mitochondria today. The mitochondrial genome is shown in Figure 19.3.10.

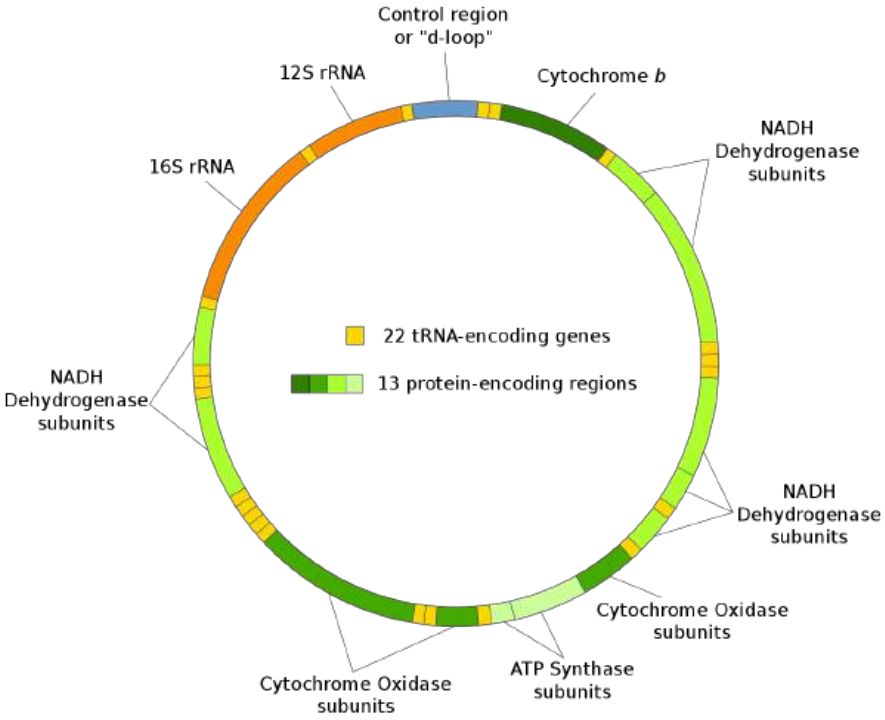
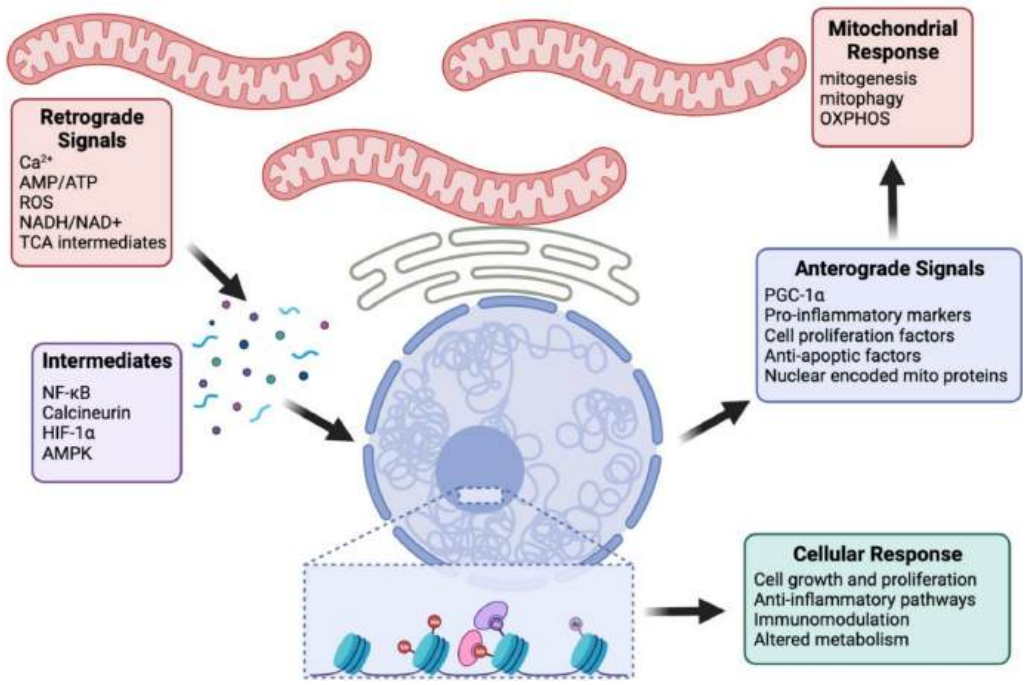


Figure 19.3.10: Mitochondrial genome. https://commons.wikimedia.org/wiki/C...ial_DNA_en.svg

Mitochondrial DNA is a double-stranded molecule of 16.5 kilobases. It encodes some proteins, as well as its own tRNA and ribosomal RNAs. There are many copies (100s to 1000s) of the circular genome in each cell.

Some (actually 13), but not all of the genes for proteins required for oxidative phosphorylation are found in the mitochondrial genome. The rest reside in the nuclear genome, and after transcription and translation, the resulting protein must be imported into mitochondria. In fact, over 95% of mitochondrial proteins are encoded by nuclear genes. Mitochondria move in the cells guided by the internal cytoskeleton to sites where energy (ATP) is needed. They also participate in buffering intracellular Ca^{2+} ions in the cell, much like the ER with which mitochondria form intimate membrane contacts. The ER also makes contact with the outer nuclear membrane, so, likely, the mitochondria can closely associate with the nucleus as well.

Hence there must be regulation of the respective genomes to produce the correct type and amounts of proteins required for oxidative phosphorylation. To accomplish this, there is a bidirectional regulated signaling between the two organelles. There is direct signaling between the nucleus and mitochondria and retrograde signaling in the other direction. Signals and processes affected in these signaling pathways are shown in Figure 19.3.10.



jdjkfj

Figure 19.3.11: Signaling from the mitochondria to the nucleus. Brittni R. Walker and Carlos T. Moraes. *Biomolecules* **2022**, *12*(3), 427; https://doi.org/10.3390/biom12030427. Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/)

Signals originating from the mitochondria, including calcium, reactive oxygen species (ROS), and AMP/ATP, often stimulate pathways, leading to transcriptional changes in the nucleus. Nuclear responses can involve upregulating cell proliferation and anti-apoptotic factors, as well as proteins involved in mitogenesis, such as PGC-1α and nuclear-encoded mitochondrial proteins. The levels of other molecules, such as TCA intermediates Acetyl CoA and α-ketoglutarate, can influence the epigenome by modifying methylation and acetylation, and consequently, the cell physiology.

We've encountered several key enzymes involved in the regulation of metabolism, including:

- AMP-activated protein kinase (AMPK) senses the energy state (AMP/ATP ratio) of the cell. When ATP is low (AMP is high), it activates pathways to produce ATP and inhibits pathways that require ATP. AMPK has been found at or near the mitochondrial membrane and may also be found in it. AMPK also promotes mitochondrial biogenesis (mitogenesis)
- mTOR (mammalian target of rapamycin) which is a serine/threonine kinase that regulates protein synthesis and other anabolic pathways based on the energy status of the cell. processes in response to growth factors, energy status, and oxygen levels. It also simulated mitogenesis.

CHAPTER OVERVIEW

20: PHOTOSYNTHESIS AND CARBOHYDRATE SYNTHESIS IN PLANTS

20.1: Light Absorption in Photosynthesis - An Overview
20.2: The Kok Cycle and Oxygen Evolving Complex of Photosystem II
20.3: Plant Electron Transport and ATP Synthesis
20.4: CO_2 uptake - Calvin Cycle and C3 organisms
20.5: CO2 uptake - C4 and CAM Pathways
20.6: Biosynthesis of Starch, Sucrose and Cellulose

20.1: LIGHT ABSORPTION IN PHOTOSYNTHESIS - AN OVERVIEW

20.1.1: INTRODUCTION

We have seen how we can transduce the chemical potential energy stored in carbohydrates, into the chemical potential energy of ATP. This occurs namely through coupling the energy released during the thermodynamically favored oxidation of carbon molecules through intermediaries (high energy mixed anhydride in glycolysis or a proton gradient in aerobic metabolism) to the thermodynamically uphill synthesis of ATP. There is a situation that occurs when we wish to reverse the entire process:

$$CO_2 + H_2O \rightarrow \text{carbohydrate} + O_2$$

This process is, of course, photosynthesis, which occurs in plants and certain photosynthetic bacteria and algae. Given that this process must by nature be an uphill thermodynamic battle, let us consider the major requirements that must be in place for it to occur:

- A **strong oxidizing agent** must be formed which can oxidize water to dioxygen. We know that redox reactions occur in the direction of a stronger to a weaker oxidizing agent (just as acid/base reactions are thermodynamically favored in the direction of strong to weak acid). Somehow we must generate a stronger oxidizing agent than product dioxygen, which often has the most positive standard reduction potential in common tables.
- Plants must have high concentrations of a **reducing agent** for the reductive biosynthesis of glucose from CO_2. The reducing agent used for most biosynthetic reactions in nature is NADPH, which differs from NADH only by the addition of a phosphate to the ribose ring. This phosphate differentiates the pool of nucleotides in the cells used for reductive biosynthesis ($NADPH/NADP^+$) from those used for oxidative catabolism ($NADH/NAD^+$)
- Finally, plants need an abundant source of **ATP** which will be required for reductive biosynthesis.

The *light reaction* of photosynthesis produces these three molecules, O_2, NADPH, and ATP. We will explore the light reaction in the next few sections of this chapter. The **dark reaction**, which as the name implies can occur in the dark, involves the actual fixation of carbon dioxide into carbohydrates using the ATP and NADPH produced in the light reaction.

The energy to power the light reactions comes directly from sunlight. Clue two is that plants have an organelle that animal cells don't - the chloroplast. Its structure is in many ways similar to mitochondria in that it has an outer membrane, an intermembrane space, and an inner membrane. In addition, it has a series of stacked, interconnected compartments called thylakoids bounded by a thylakoid membrane surrounding a lumen. A schematic is shown in Figure 20.1.1 below.

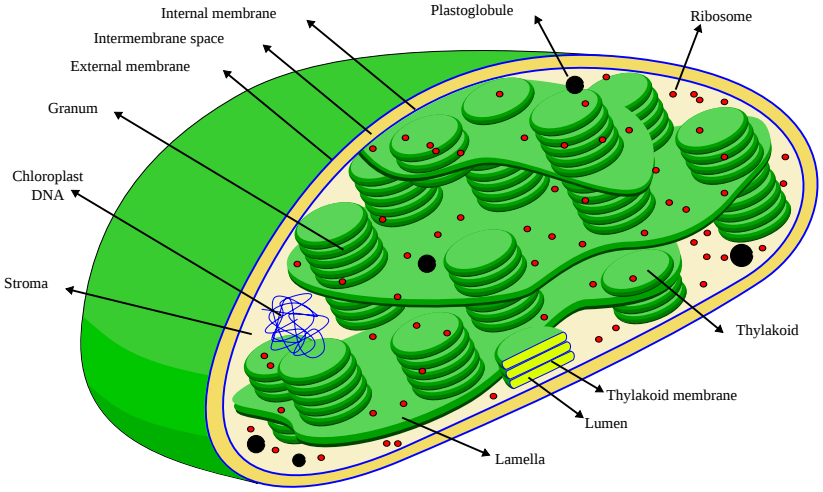


Figure 20.1.1: The chloroplast. chloroplast: https://commons.wikimedia.org/wiki/F...roplast-en.svg. Creative Commons Attribution-Share Alike 4.0

Stacks of thylakoids are called grana. Each thylakoid represents one granum. The thylakoids are where the **light reactions** occur. In contrast, the stroma contains the enzymes for the **dark reactions** of photosynthesis.

The enzyme complexes involved in the light reaction are aligned in the thylakoid membrane just as the membrane complexs involved in mitochondrial electron transport/oxidative phosphorylation were aligned in the mitochondrial inner membrane. That process used Complex I, Complex III, Complex II (succinate dehydrogenase), and Complex IV to transfer electrons from NADH and $FADH_2$ to increasingly potent oxidizing agents (ubiquinone, cytochrome C) ending with dioxygen. The energy released during that thermodynamically favored process was captured in the formation of a proton gradient, which collapsed through the F_0F_1ATPase to drive ATP synthesis.

For the light reaction, three complexes, the **Light Harvesting ComplexII -Photosystem II** (also called the **LHC-PSII supercomplex**), **cytochrome b_6f,** and the **light-harvesting complex Photosystem I** (called the LHC-PSII supercomplex) are used to carry out the light reactions of photosynthesis. These include the photooxidation of water to produce O_2 and the transfer of the lost electrons through a series of mobile electron carriers (plastoquinone and plastocyanin) to a terminal acceptor, $NADP^+$ to form the reducing agent NADPH needed for carbohydrate synthesis in the dark reaction. The energy released during the electron transfer process is likewise captured in the form of a proton gradient as protons are moved from the chloroplast stroma to the lumen of the thylakoid. The collapse of the resulting proton gradient powers ATP synthesis. These reactions are shown in Figure 20.1.2.

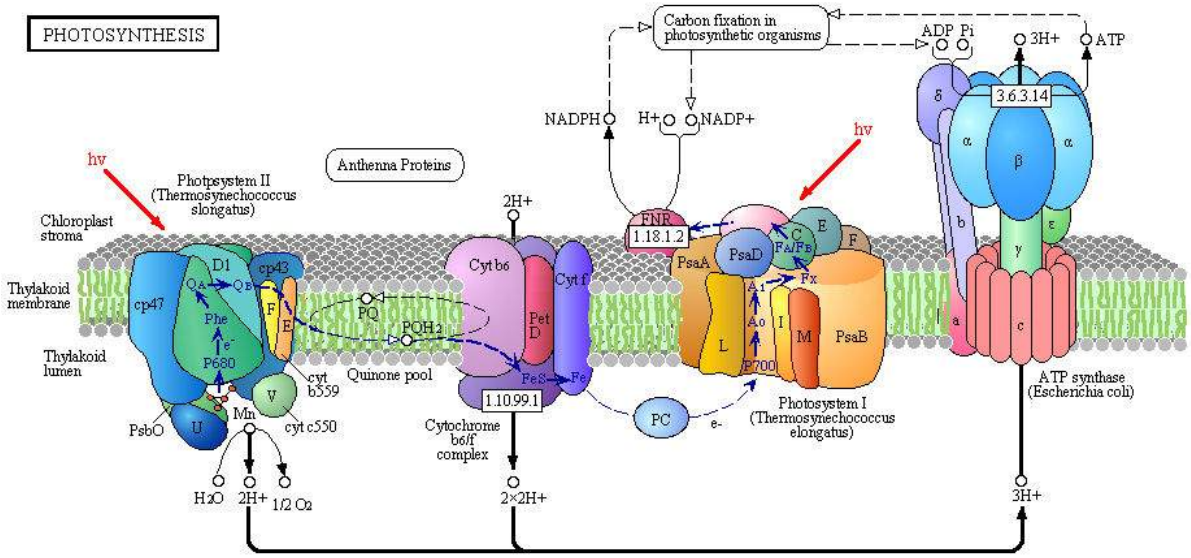


Figure 20.1.2: Light reaction of photosynthesis. The boxed numbers represent Enzyme Commission Number. Original KEGG Map with embedded links. (reprinted with permission from Kanehisa Laboratories and the KEGG project: **www.kegg.org**)

In this section, we will explore the absorption of light by the light-harvesting complex (LHCII) of the LHC-PSII supercomplex. Before we get into too much detail, let's start with a simplified review of light absorption in the LHCII.

20.1.2: ABSORPTION OF LIGHT

Plants have many pigments (chlorophylls, phycoerthryins, carotenoids, etc.) whose absorption spectra overlap that of the solar spectra. The main pigment, chlorophyll, has a protoporphyrin IX ring (same as in heme groups) with Mg^{2+} at its center instead of Fe^{2+}. When the chlorophyll absorbs light, the excited electrons must eventually relax to their ground state. It can do this by either radiative or nonradiative processes. In radiative decay, a photon of lower energy is emitted (after some energy has already been lost by vibrational transitions) in a process of either fluorescence or phosphorescence. In nonradiative decay, the energy of an excited electron can be transferred to another similar molecule (in this case other chlorophyll molecules) in a process that excites the electron in the second molecule to the same excited state. It is as if a photon is released by the first excited molecule, which then is absorbed by an electron in a second molecule to excite it to the same excited state. However, no photon is involved in the energy transfer. In this fashion, energy is transferred from one chlorophyll to another. This type of energy transfer is called resonance energy transfer or **exciton** transfer, as shown in Figure 20.1.3.

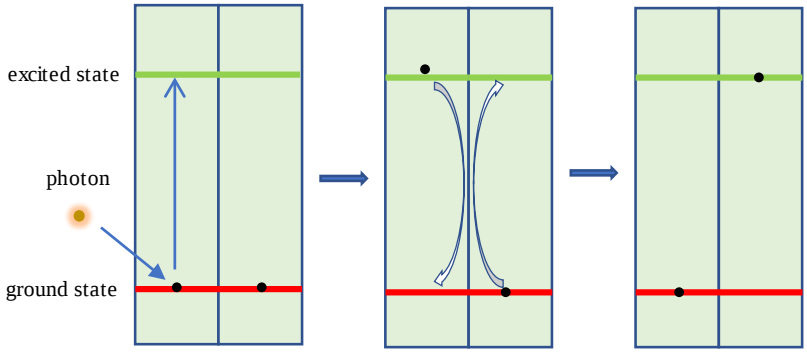


Figure 20.1.3: Resonance Energy (Exciton) Transfer

Because of its unique environment, one type of chlorophyll has slightly different characteristics. The energy level of the first excited state in the chlorophyll reaction center is lower than in the rest of the chlorophyll molecules, in much the same way that pKa values of amino acid side chains differ with the local and solvent environment, and the standard reduction potential of FAD molecules that are tightly bound to enzymes differ due to the different environment of bound FAD/$FADH_2$.

Instead of a radiationless transfer of energy to this special chlorophyll, an actual electron from the excited state chlorophyll is transferred, which by definition is a redox reaction. The electron donor (the excited chlorophyll) loses an electron (an oxidation reaction) as the recipient molecule gains one (a reduction). This charge (electron) transfer reaction produces charge separation in the formation of a positively charged chlorophyll (now an oxidizing agent) and a negatively charged chlorophyll (now a reducing agent). The chlorophylls directly involved in this final process are collectively called the **reaction center.** This process is shown in Figure 20.1.4 below.

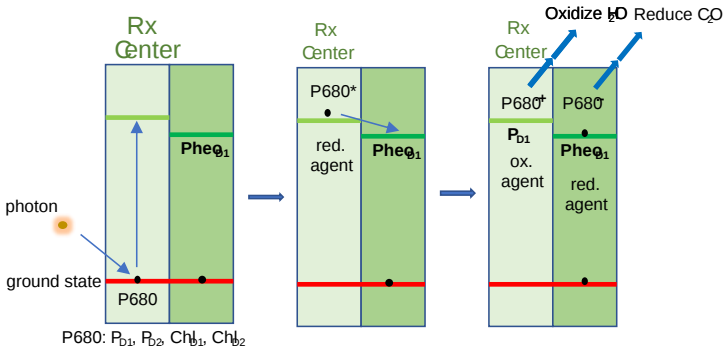


Figure 20.1.4: Reaction Center

The reaction center chlorophylls absorb light at 680 nm so sometimes these chlorophylls are labeled P680. There are 4 unique chlorophylls (P_{D1}, P_{D2}, Chl_{D1}, and Chl_D) that are the main players in the reaction center. Both sets of labels are shown in the figure above.

Figure 20.1.5 shows a cartoon of the absorption of photons and subsequent handoff of energy to "antennae" chlorophylls leading to photoexcitation of the reaction center and subsequent transfer of an electron to the special chlorophyll which has a lower first excited state energy. This molecule is named pheophytin A. It is identical to chlorophyll A but lacks the central Mg^{2+} ion.

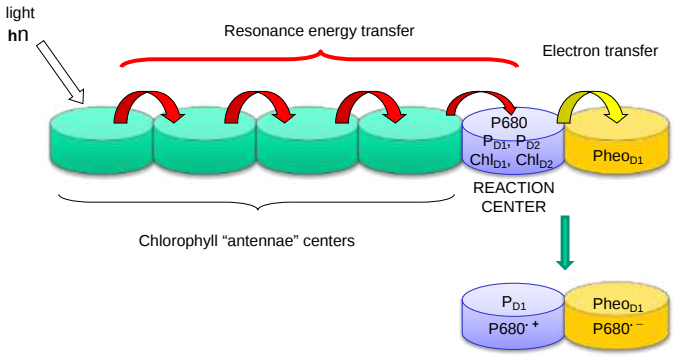


Figure 20.1.5: Photoexcitation and electron transfer of chlorophylls

Photosystems I and II contain many chlorophyll molecules that act as antennas that transfer energy to the reaction centers. The "antenna" proteins involved in photon adsorption and energy transfer in Photosystems I and II are shown in Figure 20.1.6 below. Note the D1, D2, cp47, and cp43 protein subunits in the figure below are also shown in the figure above. The D1 and D2 subunits contain the reaction center chlorophylls.

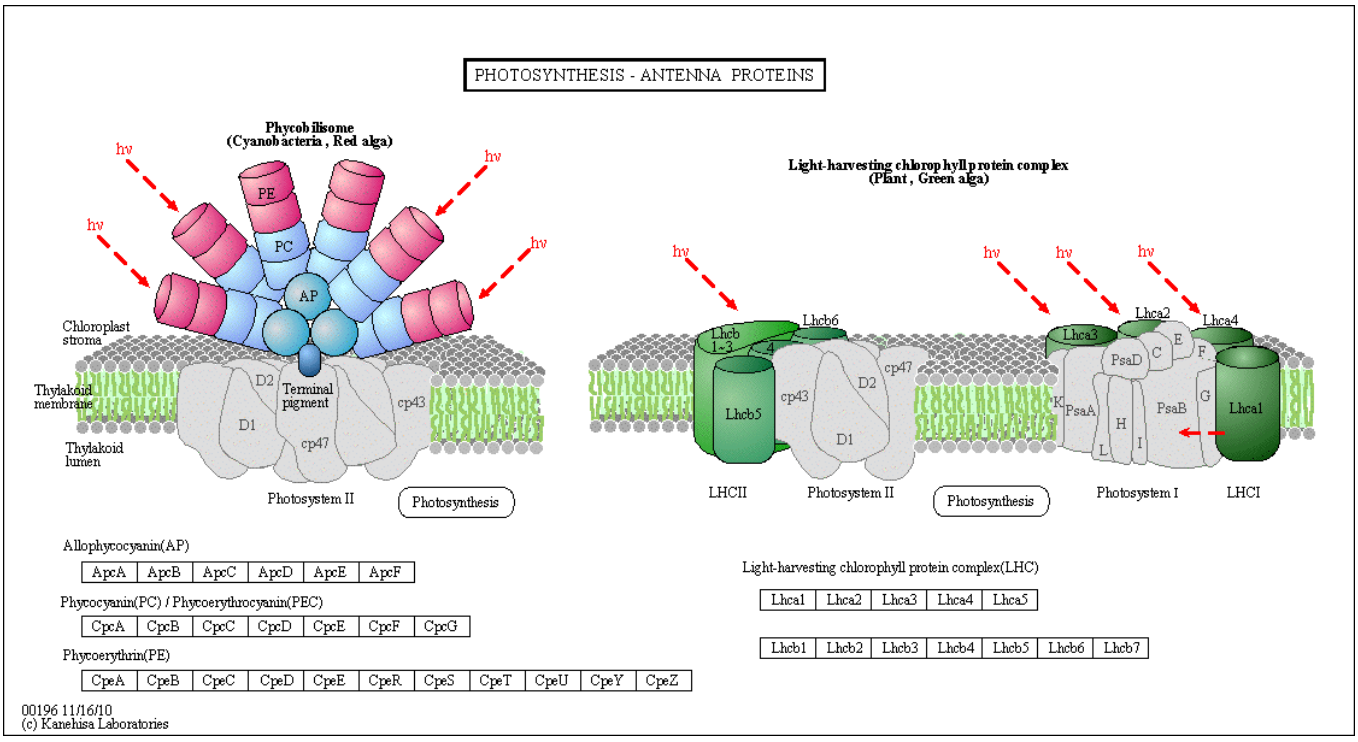


Figure 20.1.6: Antennae Proteins (reprinted with permission from Kanehisa Laboratories and the KEGG project: www.kegg.org)

The different chlorophylls have absorption spectra that overlap reasonably well with the solar spectrum, as illustrated in Figure 20.1.7.

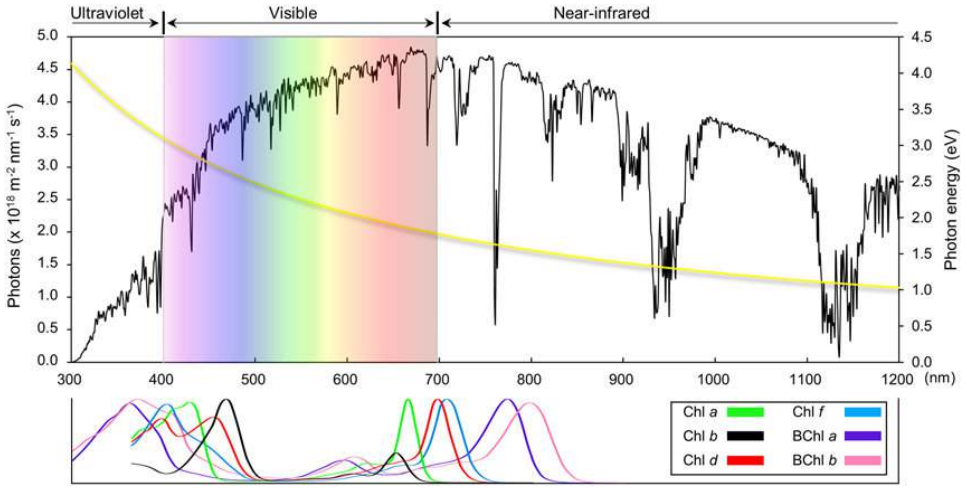


Figure 20.1.7: Solar spectrum and absorption profiles of chlorophyll and bacteriochlorophyll pigments. Cardona T, Shao S, Nixon PJ. Enhancing photosynthesis in plants: the light reactions. *Essays Biochem*. 2018;62(1):85-94. Published 2018 Apr 13. doi:10.1042/EBC20170015Creative Commons Attribution License 4.0 (CC BY).

The black lines show the solar flux spectrum (photons per meter square per wavelength) from 300 to 1200 nm. The yellow line is the photon energy (E=hc/λ) at each wavelength. The absorbance spectra of the chlorophyll species are normalized to the same maximal value.

If you use the energy required to make glucose from CO_2 and H_2O as a standard, about 95% of the incoming solar energy is wasted since not all of the incident light photons have the right wavelength. Waste also occurs due to reflectance and heat generation.

20.1.3: THE LIGHT HARVESTING COMPLEX (LHCII) - PHOTOSYSTEM II (PS II) SUPERCOMPLEX

Now let's look in more detail at the chloroplast thylakoid membrane complex that interacts with light and results in the oxidation of water to form O_2. This first structure is called the **Light Harvesting Complex II (LHCII) - Photosystem II (PS II) Supercomplex**. It is a super complex (a pun) to understand. The supercomplex has a **PSII core complex** interacting with a variable number of light-harvesting complex IIs (**LHCIIs)** complexes. Sometimes the entire super complex is more simply called Photosystem II. The supercomplex consists of

- a PSII core **C** with a Mn_4CaO_5 cluster called the oxygen-evolving complex (OEC or OEX) that oxidizes water to O_2;

- peripheral antennae complexes **M** and **S**, also called light-harvesting complexes II (LCHIIs)

The core part of the Photosystem II supercomplex contains the key to the existence of aerobic organisms as it produces the dioxygen in our atmosphere. Its chemistry is phenomenal. Many are seeking to modify it to produce, in addition to O_2, a green energy source, H_2. We will explore the mechanisms for the oxidation of water to O_2 by a PSII-bound inorganic Mn_4CaO_5 cluster, the oxygen-evolving complex (OEC) in the next section. Let's first look at PSII in light of the discussion above. Where are all the chlorophylls? The mystical "reaction center"?

The PSII Core Complex (within the supercomplex) has over 20 subunits. These include:

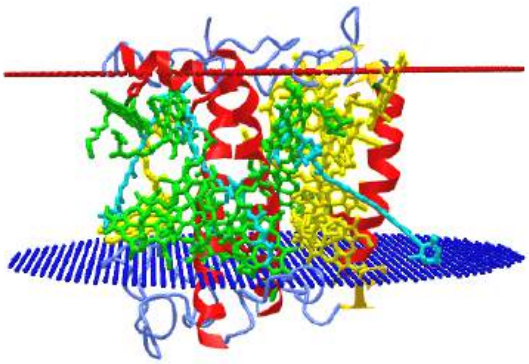
- reaction center subunits D1 and D2
- inner antennae subunits CP43, CP47
- many other small protein subunits
- peripheral subunits that project into the lumen that interact with the OEC

The peripheral antennae complexes (light-harvesting complexes - LHCIIs M and S - consist of proteins encoded by 6 different Lhcb genes (1-6).

- Lhcb1-3 monomers form trimeric (homo- or hetero) LHCIIs (Lhcb1-3). An example is $(Lhcb3)_3$. That trimer is often called the M-LHCII. There is also an S-form trimer as well as L- and N-forms. S represents a strong association with the reaction center, M for moderate and L for loose.
- Lhcb4-6 are called the minor components and consist of the monomeric proteins named **CP29, CP26, and CP24**, respectively.

20.1.4: LIGHT-HARVESTING COMPLEX PROTEINS

Each Lhc apoprotein consists of 3 alpha-helices and binds 8 chlorophylls a, 6 chlorophyll b, and 4 carotenoid molecules. Figure 20.1.8 shows an interactive iCn3D model of the monomeric pea chlorophyll a-b binding protein AB80 (LHCII type I CAB-AB80) (2BHW)



NCBI iCn3D Figure 20.1.8: Monomeric pea chlorophyll a-b binding protein AB80 (LHCII type I CAB-AB80) (2BHW). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...QtEEWeQ5Y9n2i8

The color scheme is as follows:

- protein - secondary structure colors
- Chlorophyll a - Green
- Chlorophyll b - yellow
- carotenoids - cyan

Now let's look at a trimeric form of the same protein. of LHCII, the chlorophyll-binding protein, from PSII. It's also called LHCII type I CAB-AB80. We showed the role of the LHCII trimer in the figures above. The subunits function to absorb light for the light reaction and through resonance transfer energy to the reaction centers in PSII and in addition PSI. Figure 20.1.9 shows an interactive iCn3D model of the trimeric Light Harvesting Complex (LHC_II) from pea photosystem II (2bhw).

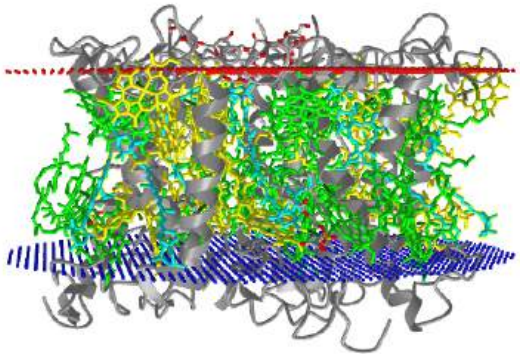


Figure 20.1.9: Light Harvesting Complex (LHC_II) from pea photosystem II (2bhw). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...BK6WvzrxABPDh6

The protein is shown in grey. To simplify the model, chlorophyll A molecules are shown in magenta, and chlorophyll B isomeric variants are in green. Note a large number of chlorophylls (42) in the LHCII complex, which differentiates it from most protein complexes which might have just a few ligands bound.

Photosynthetic bacteria and plants have an abundance of molecules that interact with visible light. The main one found in PS II is chlorophyll a whose structure is shown in Figure \(\PageIndex{10}\) below, along with the variant, chlorophyll b. Remember the pheophytin A is identical to chlorophyll A but lacks the central Mg^{2+} ion.

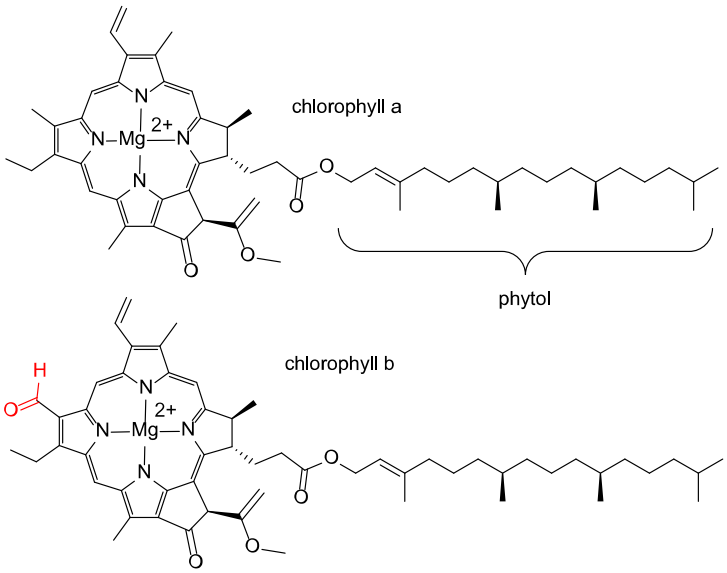


Figure 20.1.10: Chlorophyll a and b

20.1.5: PHOTOSYSTEM II (PS II) SUPERCOMPLEX STRUCTURES

The most common version of the supercomplex is a dimer. A simplified cartoon of the monomeric version of the **PSII-LHCII supercomplex** is shown in Figure 20.1.11. The PSII core is shown in the blue rectangle. The CP29, CP26, and CP24 in the peripheral antennae complexes (light-harvesting complexes - LCHIIs) are shown outside of the blue-outlined rectangle. They surround the PSII core.

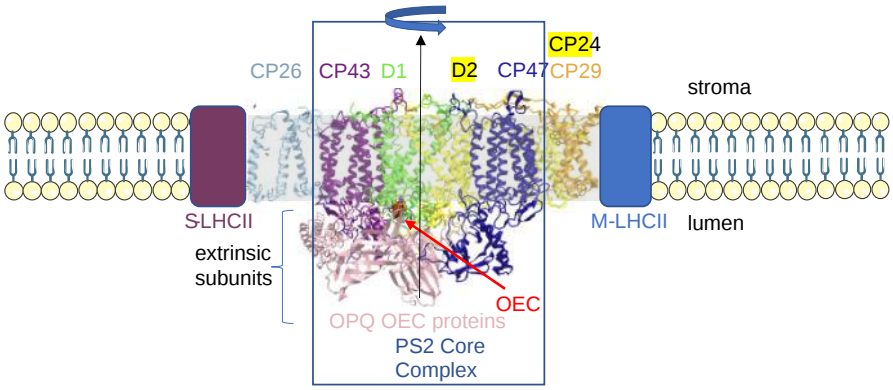


Figure 20.1.11: Cartoon of the monomeric PSII-LHCII supercomplex structure of the PSII-LHCII supercomplex from Pisumsativum

- The PSII core **C** (outlined in the blue box) contains two 2 reaction center subunits D1 and D2, as well as two inner antennae subunits CP43 and CP47. The core also contains extrinsic subunits (pink) surrounding the OEC (red spacefill)
- A **S**trongly bound LHCIIs protein, **S -LHCII**, left (maroon rectangle) interacting with CP26. The S-LHCii complex is a trimer of individual Lhcb monomers.
- A **M**oderately bound LHCIIs protein complex, **M-LHCII,** right (lighter blue rectangle) interacting with the CP24/CP29 dimer. The M-LHCII complex is a trimer of 3 lcbh-3 subunits $(Lcbh3)_3$.

There can be a variable number of LHCIIs in the complex. The predominant form of the supercomplex is the $\mathbf{C_2S_2M_2}$ complex. There are also many other proteins in the supercomplex that are not shown in Figure 11. A cartoon showing the top-down view of the actual supercomplex dimer, $C_2S_2M_2$, is shown in Figure 20.1.12.

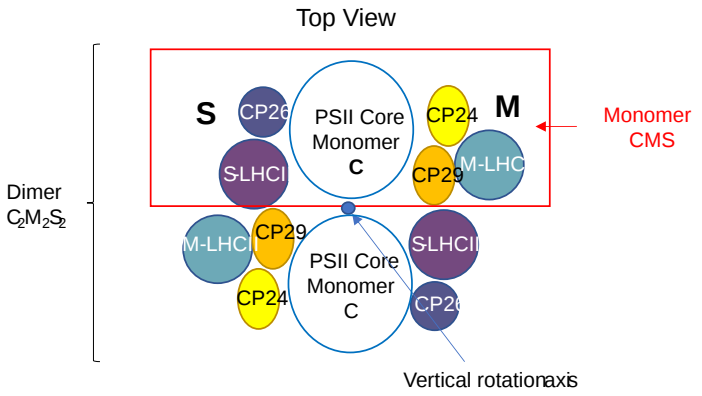
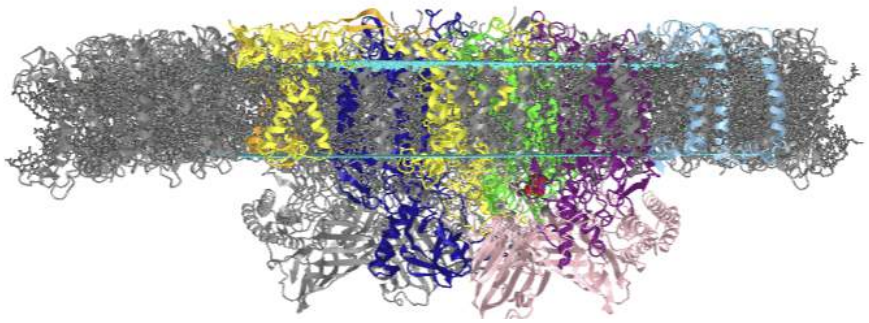


Figure 20.1.12: Top town view of the **PSII-LHCII supercomplex dimeric complex.**

You could imagine forming the dimer by spinning the monomer CSM 180^0 and reproducing the monomer there.

Figure 20.1.13 shows an interactive iCn3D model of the full C2S2M2-type PSII-LHCII supercomplex from Pisum sativum (5XNL). (long load time)



NCBI iCn3D Figure 20.1.13: C2S2M2-type PSII-LHCII supercomplex from Pisum sativum (**5XNL)**. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...F4WFfHoevxaVV8 (long load time)

Only a single CSM monomer of the C2S2M2 dimeric supercomplex is colored. The colors match those in Figures 8 and 9. The molecules shown in gray sticks that outline a bilayer include chlorophyll Bs (98 of them), chlorophyll As (216), and carotene or its derivatives (88) as well. The different proteins in the supper complex are colored as shown below.

- CP47 Rx Center (inner antenna protein) Dark Blue
- CP43 Rx Center (inner antenna protein) - Purple
- D1: green
- D2: yellow
- OEC: spacefill Red
- O, P, Q - OEC peripheral: Light pink
- CP24 (peripheral antenna protein): yellow
- CP29 (peripheral antenna protein): orange
- CP26: Light blue
- LHCIIs (S and M) - not shown
- membrane bilayers - cyan spheres

(A quick load version of the C2S2M2-type PSII-LHCII supercomplex with different coloring than above)

20.1.6: PSII SUPERCOMPLEX REGULATION

Plants have evolved a great ability to absorb light over the entire visible range of the spectra. Can they absorb too much energy? The answer is yes, so plants have developed many ways to protect themselves. IF too much light is absorbed, the pH gradient developed across the thylakoid membranes becomes greater. This is sensed by a protein, PsbS, and through subsequent conformational changes transmitted through the light-harvesting antennae, the excess light energy is dissipated as thermal energy. Mutants lacking PsbS showed decreased seed yield, a sign that it became less adaptable under conditions of stress (such as exposure to rapidly fluctuating light levels). Molecules called xanthophylls and other carotenoids such as zeaxanthin are also important in excess energy dissipation. These molecules appear to cause excited state chlorophyll (a singlet-like excited state dioxygen) to become deexcited. Without the xanthophylls, the excited state chlorophyll could deexcite by transfer of energy to ground state triplet dioxygen, promoting it to the singlet, reactive state, which through electron acquisition, could also be converted to superoxide. These reactive oxygen species (ROS) can lead to oxidative damage to proteins, lipids, and nucleic acids, alteration in gene transcription, and even programmed cell death. Carotenoids can also act as ROS scavengers. Hence both heat dissipation and inhibition of the formation of ROS (by such molecules as vitamin E) are both mechanisms of defense against excessive solar energy

Given that both plants and animals must be protected from ROS, antioxidant molecules made by plants may prove to protect humans from diseases such as cancer, cardiovascular disease, and general inflammatory diseases, all of which have been shown to involve oxidative damage to biological molecules. Humans, who can't synthesize the variety and amounts of antioxidants that are found in plants, are healthier when they consume large amounts of plant products. These phytomolecules also have other properties, including regulation of gene transcription which can also have a significant effect on disease propensity.

Plants have to respond to different qualities and quantities of light. When light is low, they seek to maximize light capture. Too much light could damage a plant so molecular adaptions are made to prevent it. Part of the regulation occurs in the stoichiometry of the supercomplex by altering the antenna protein (the LHCs) composition. The most abundant component of the complex is $C_2S_2M_2$ and C_2S_2M. C_2S and C_2S_2 increase as an adaption to increasing levels of light.

Light levels can also promote grana membrane association mediated by interactions of a PSII-LHCII supercomplex (PSII-LHCIIsc) on one thylakoid membrane interacting with a PSII-LHCIIsc on an adjacent thylakoid membrane to form a large PSII-LHCIIsc dimer. The structure and properties of paired PSII-LHCIIsc are illustrated in Figure 20.1.14.

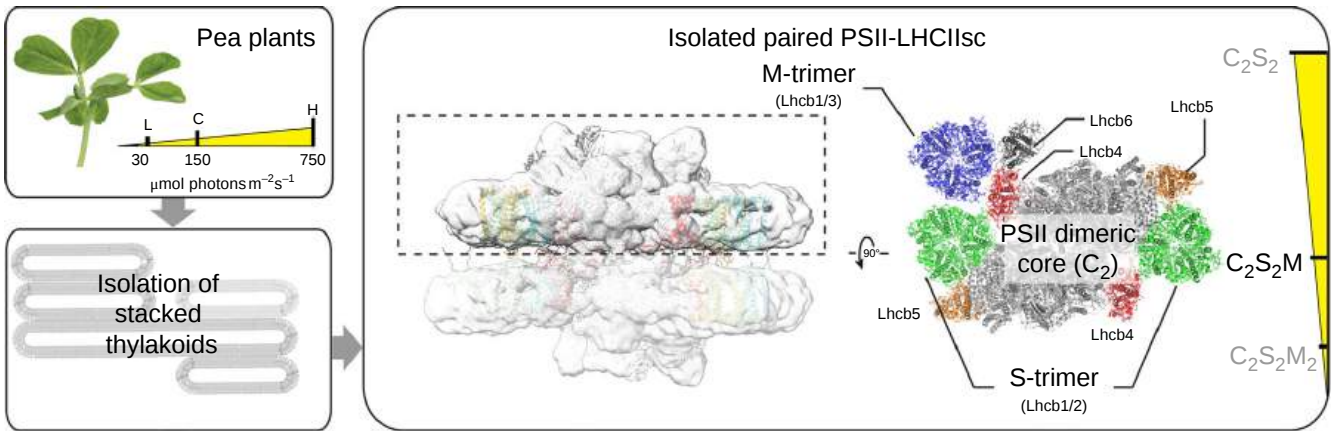


Figure 20.1.14: Light-driven modulation of paired PSII–LHCIIsc. Heterogeneous mixtures of PSII–LHCIIsc from pea plants grown at different light intensities (L, low; C, moderate used as control; H, high) were isolated and analyzed. Albanese et al. Nature Communications | (2020) 11:1361 | https://doi.org/10.1038/s41467-020-15184-1. Creative Commons Attribution 4.0 International License. http://creativecommons.org/ licenses/by/4.0/.

The figure also shows the variation in the PSII–LHCIIsc with light intensity. At low light $C_2M_2S_2$ prevails while at high light intensity C_2S_2 is most abundant. The system moves to regulate activity by increasing the abundance of LHCIIs in low light and decreasing them in high light intensities!

Figure 1 shows the stacking of individual granum in the chloroplasts. The stacking is maintained at various light levels and is mediated by loops of the LHCII trimers that are exposed in the stroma and Lhcb4 subunits on adjacent membranes. The stromal surfaces are flat and tightly stacked in grana. Stacking is a dynamic process and depends on cross-membrane interactions between and reorganization of the PSII -LHCIIsc.

The PSII–LHCIIsc is regulated through light-dependent post-translational phosphorylation, particularly on LHCII, and acetylation. which further regulates energy distribution. In plants, the core proteins CP43, D1 and D2, and PsbH are phosphorylated by the kinase Stn8. The peripheral antenna LHCII proteins are phosphorylated by Stn7. Phosphorylation of PSII is not seen in cyanobacteria and red algae. Higher light levels promote higher levels of core protein phosphorylation. Stn7 appears to be inhibited at high light levels. Figure 20.1.15 shows putative Stn8 phosphorylation sites in the PSII -LHCIIsc

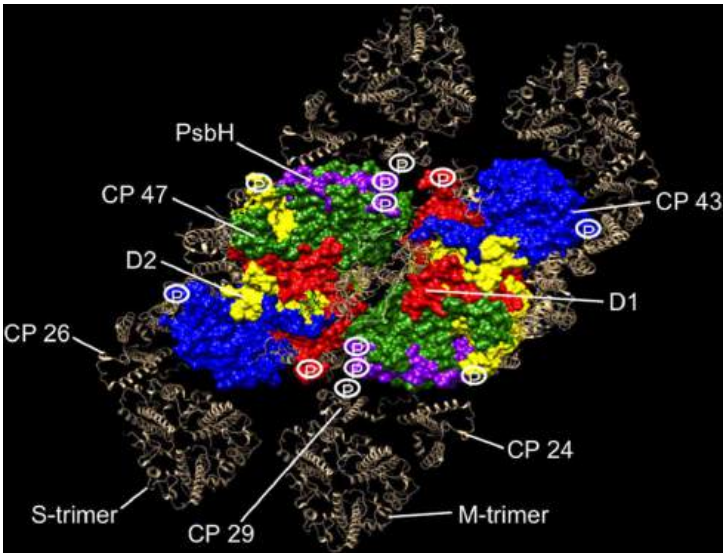


Figure 20.1.15: **Phosphorylation map of the C2S2M2 supercomplex as catalyzed by the Stn8 kinase**. Puthiyaveetil Sujith and Kirchhoff Helmut, Frontiers in Plant Science, 4 ,2013. https://www.frontiersin.org/article/...pls.2013.00459. DOI=10.3389/fpls.2013.00459. **Creative Commons Attribution License (CC BY)**

The structure is a model pieced together from multiple pdb files. The approximate positions of phosphorylation sites of D1, D2, CP43, PsbH, and CP29 are shown.

Most free N-terminal loops in LHCII can be phosphorylated. Both phosphorylation and lysine acetylation on Lhcb2 N-terminal loops help regulate the redistribution of LHCII from PSII (in grana stacks) to PSI (in single-layered thylakoid regions). Grana stacking by N-terminal loop association between facing PSII–LHCIIsc and their N-terminal acetylation appear to strengthen grana stacking.

How does PSII respond when not the intensity but the wavelength of light is changed? For this, we have to briefly discuss photosystem I (**PSI**), which we will explore in more detail in Chapter 20.3. Both PSII and **PSI** work together to transduce light into chemical energy so you would expect that their activities are regulated in a linked fashion. They have different absorbance spectra characteristics as well, with PSI absorbing more in the red region. If the effective absorbance (normalized for concentrations and LHCIIs) were the same, you would predict that the effective absorbance ratio over a broad wavelength range for the two photosystems, **PSI**/(**PSI**+PSII) would be 0.5. This is approximately the case over the entire spectral wavelength except for between 670-730 nm, where the ratio is close to 1, showing that PSI absorbance and hence activity is tilted toward the red end of the absorbance spectra. Changes in light characteristics (i.e. wavelength) would then affect each photosystem differently and can cause imbalances in their activities and their states, which should lead to a restorative balance. When exposed to far-red light, the systems move to state I. In this state, the major mobile antenna proteins (LHCIIs) move to PSII to restore a "photoabsorption" balance. When exposed to light depleted in the high end of visual spectra, the system moves to state II, in which mobile LHCIIs move to PSI. What an interesting reciprocal way to balance activity, even if it is hard to conceptualize regulation involving the movement of membrane proteins. Of course, such movement is seen often in the clustering of ligand-bound membrane receptors.

As mentioned above, the location and hence movement of the LHCII is regulated by phosphorylation by LHCII kinase. We'll explore that more in Chapter 20.3 when we discuss PSI in more detail.

20.1.7: A CLOSER VIEW OF THE REACTION CENTER AND ITS LOCAL ENVIRONMENT

Let's look at the structure of a simpler PSII complex from Thermostichus vulcanus (3WU2). It has 70 chlorophyll a molecules, 4 special chlorophylls, 4 pheophytin As (PHOs), 20 beta-carotenes, 4 plastoquinol-9s (PL9s), 4 hemes, and 2 caroten-3-ols. Figure 20.1.16 shows the arrangement of chlorophyll molecules in the CP47 and CP44 (not CP43 as shown in several of the above diagrams) antennae subunits of PSII from T. vulcanus (3WU2).

.

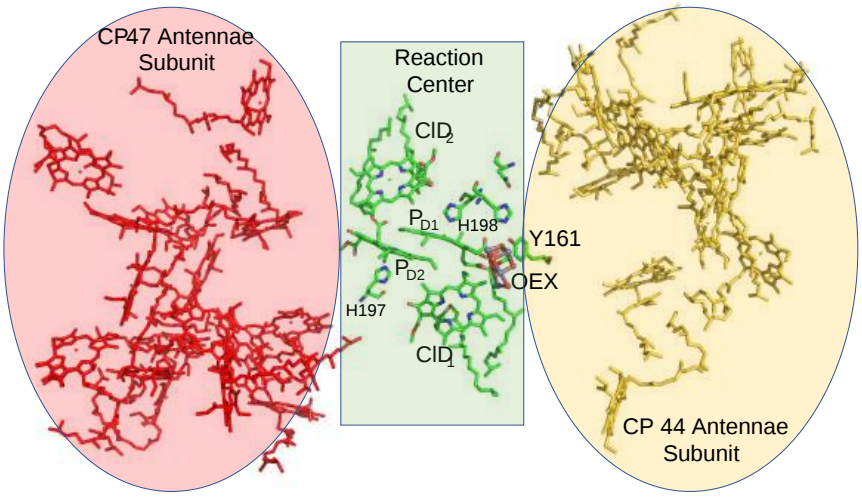


Figure 20.1.16: Reaction Center with special chlorophylls sandwiched between antennae subunits

- CP44 is the **Photosystem II CP44 reaction center protein, psbC gene,** Photosystem II 44 kDa reaction center protein;
- CP47 is the **Photosystem II CP47 reaction center protein, psbB gene;** Photosystem II CP47 chlorophyll apoprotein

Sandwiched in between them is the reaction center containing the 4 special chlorophylls, the inorganic metal cluster called the oxygen-evolving complex (OEC or OEX), and a key amino acid near the OEC, Tyr 161 (pdb 3ARC).

Two special chlorophylls, $ChlD_1$ and $ChlD_2$, accompanied by partner chlorophylls, P_{D1} and P_{D2}, that are coplanar to each, are found near the OEC and are the key chlorophylls in the reaction center that turn OEC into a powerful enough oxidant to oxidize H_2O

A special reaction center chlorophyll/pheophytin absorbs a photon of light at 680 nm so the species that absorbs the photon is given the label P680. On absorption, it forms the excited state, P680*. This transfers the excited state electron to pheophytin which forms a pheophytin radical anion, while the electron donor P680* becomes $P680^{\cdot+}$, a radical cation. The radical cation, an unstable species, can oxidize another molecule to regain stability, and in a series of addition linked oxidation stems, water is oxidized (O in water has an oxidation state of 2-) to O_2 (in which each oxygen has an oxidation state of 0). This happens through the oxygen-evolving complex, which we will explore in the next section.

Which of the chlorophyll molecules absorbs the light (ie. which is P680)? The answer is probably all four chlorophylls in the reaction center, P_{D1}, P_{D2}, Chl_{D1}, and Chl_{D2} through a delocalization of the excited electron, which would be described by a wave function for the combined chlorophylls. On donation of the electron from the P680*, the positive charge, which can be also described as a "hole" (similar to

transistors) delocalizes as well. Quantum mechanical calculations show a coupling between P_{D1} and P_{D2}, such that 80% of the positive charge and radical character is situated on P_{D1} and 20% is on P_{D2}.)

Figure 20.1.17 shows the four chlorophylls in the reaction center of T. Vulcanus. P_{D1} appears closest to the OC.

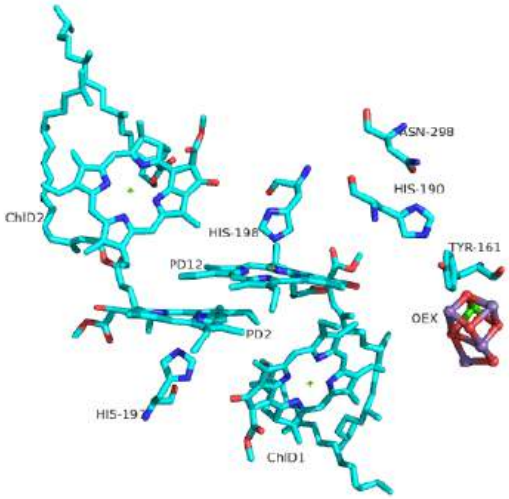
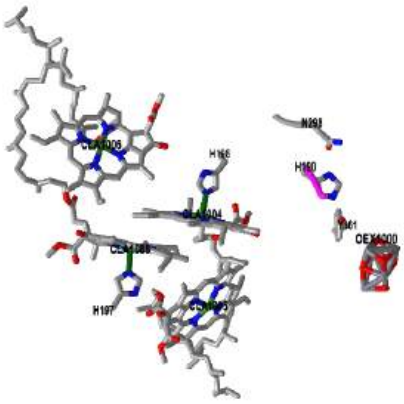


Figure 20.1.17: Closeup of the reaction center chlorophylls and OEC

Figure 20.1.18 shows an interactive iCn3D model of the key chlorophylls and OEC of photosystem II from Thermostichus vulcanus (3WU2)



NCBI iCn3D Figure 20.1.18: Key chlorophylls and OEC of photosystem II from Thermostichus vulcanus (3WU2). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...qYxWN5sKcquGK7

Now in any enzymatically catalyzed reaction mechanism, the enzyme must return to the beginning state and a path for electron flow must be apparent. How does the radical cation $P680^{\cdot+}$ return to the ground state? As we will see in the next section, it "grabs" an electron from the nearby Y161 (also called **YZ**), which then forms a radical cation, $Y^{\cdot+}$. This likewise returns to the ground state by grabbing an electron from the OEC which ultimately grabs one from water. This process repeats four times to remove the four electrons from two waters needed to form dioxygen.

Figure 20.1.19 shows the first step in process of reforming P680 and the formation of the Y161 radical cation.

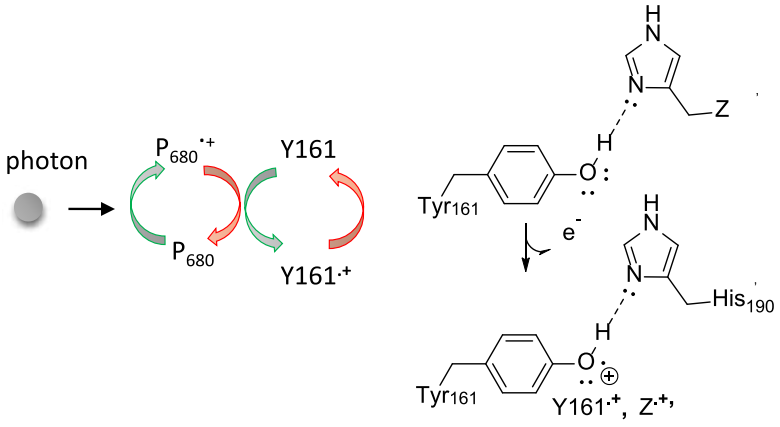


Figure 20.1.19: First step in process of reforming P680 and formation of the Y161 radical cation.

A different process occurs to remove electrons from the radical anion, $P680^{-}$. This is passed on to a series of other electron acceptors/carriers as part of the Z scheme for the light reaction of photosynthesis. The electron is ultimately passed on to $NADP^{+}$ to form NADPH for the reductive biosynthesis of carbohydrates.

Light is also absorbed in photosystem I, which does not form O_2. Rather it receives electrons from a mobile electron carrier and passes them on $NADP^{+}$ to form NADPH, a reducing agent needed for the reductive biosynthesis of carbohydrates. It also helps produce a proton gradient which helps drive ATP synthesis.

With this background, we can now explore in greater detail the key reactions that enable the evolution of aerobic organisms.

20.2: THE KOK CYCLE AND OXYGEN EVOLVING COMPLEX OF PHOTOSYSTEM II

20.2.1: INTRODUCTION

We have just seen how photoexcitation of the non-reaction center chlorophyll turns that molecule into a good reducing agent, which transfers its electron to the nearest excited state level of the reaction center chlorophyll. If you count both steps together, the non-reaction center chlorophyll gets "photooxidized", in the process producing the "strong" oxidizing agent which is the positively charged chlorophyll derivative. The extra electron passed onto the second molecule will eventually be passed on to $NADP^+$ to produce NADPH.

These reactions occur in the presence of light and hence are called the **light reactions**. The light reactions of photosynthesis in green plants are shown in Figure 20.2.1, along with the standard reduction potentials of the participants, the Z scheme.

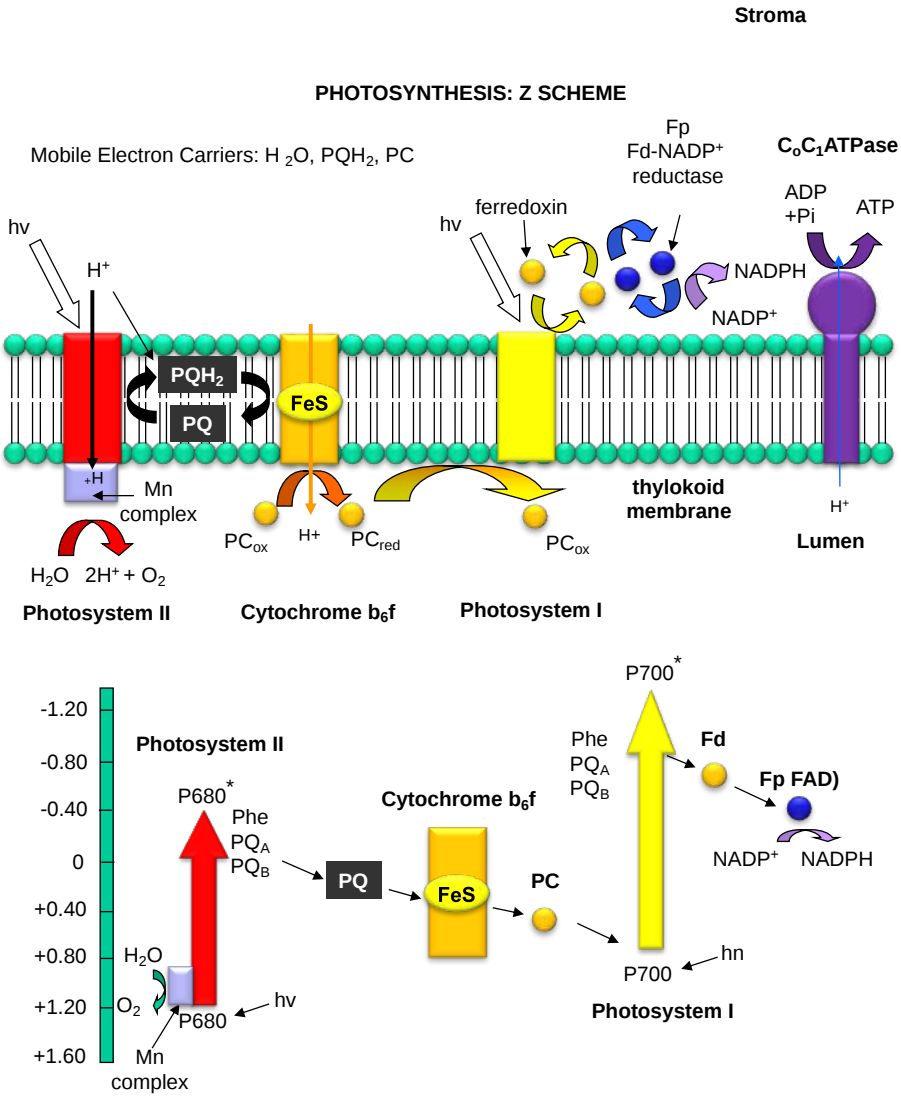


Figure 20.2.1: Light reaction of photosynthesis and associated standard reduction potentials

The combined processes of PSII and PSI resemble a "Z" scheme (rotate in your mind the standard reduction potential figure 90 degrees clockwise). In an organization reminiscent of electron transport in mitochondria, water is oxidized by **photosystem II (PSII)**. Electrons from water are moved through PSII to a mobile, hydrophobic molecule, plastoquinone (PQ) to form its

reduced form, PQH_2. Another photosystem, **photosystem I (PS1)**, is next in the electron transport pathway. It takes electrons from another reduced mobile carrier of electrons, plastocyanin (PC_{red}) to ferredoxin, which becomes a strong reducing agent. Ferredoxin is a protein with an Fe-S cluster (Fe-S-Fe-S in a 4-membered ring, with 2 additional cysteine residues coordinating each Fe). It ultimately passes its electrons along to $NADP^+$ to form NADPH. Note the complexes that produce a transmembrane proton gradient. In contrast to mitochondria, the lumen (as compared to the mitochondrial matrix) becomes more acidic than the stroma. Protons then can move down a concentration gradient through the C_0C_1ATPase to produce ATP required for the reductive biosynthesis of glucose.

Figure 20.2.2 a more detailed view of the molecular players in the light reaction.

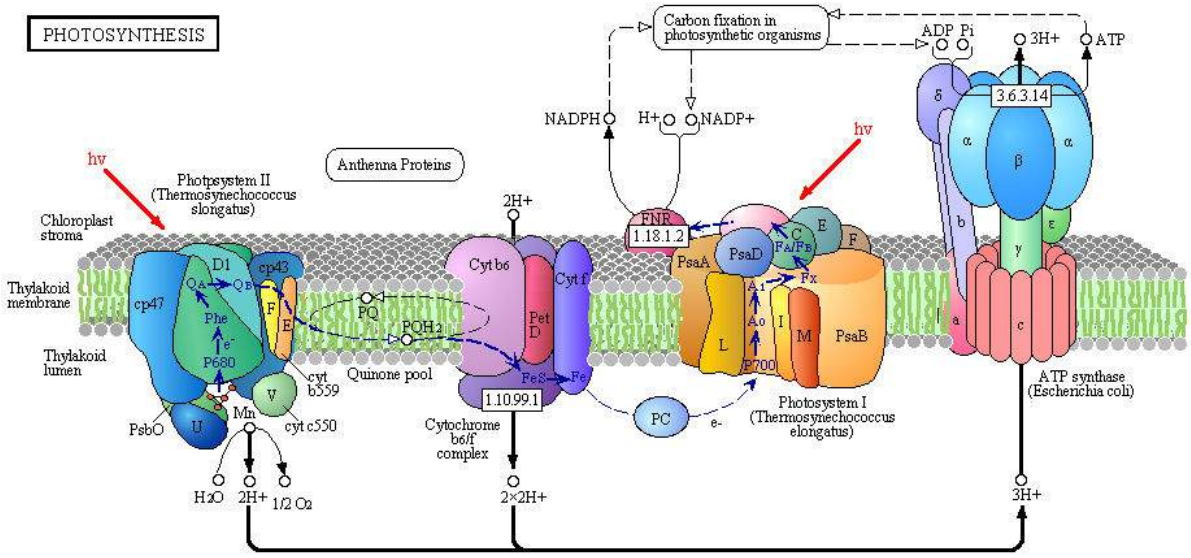


Figure 20.2.2: Detailed View of Light Reaction of Photosynthesis (reprinted with permission from Kanehisa Laboratories and the KEGG project. The boxed numbers represent Enzyme Commission Number. Original KEGG Map with embedded links.

20.2.2: PHOTOSYSTEM II

PSII has a complicated structure with many polypeptide chains, lots of chlorophylls, and Mn, Ca, and Fe ions. A Mn cluster, called the **oxygen-evolving complex, OEC** (also called the OEX) is directly involved in the oxidation of water. Two key homologous 32 KD protein subunits, D1 and D2, in PSII are transmembrane proteins and are at the heart of the PSII complex. It has been said of PSII that "Of all the biochemical inventions in the history of life, the machinery to oxidize water — photosystem II — using sunlight is surely one of the grandest." (Sessions, A. et al, Current Biology 19 (2009)

The net reaction carried out by PSII is the oxidation of water and reduction of plastoquinone.

$$2PQ + 2H_2O \rightarrow 2PQH_2 + O_2\,(g)$$

The oxidation number of oxygen in water is -2 and 0 in O_2, so this is a loss of electrons or oxidation of the water. Note that water is not converted to $2H_2 + O_2$, as in the electrolysis of water. Rather the Hs are removed from the water as protons in the lumen of the chloroplast, since the part of PSII that oxides water is near the lumenal end of the transmembrane complex. Protons are required to protonate the reduced (anionic) form of plastoquinone to form PQH_2, an activity of PSII found closer to the stroma, derive from the stroma. That being said, researchers actively trying to develop a photosynthetic scheme or mimic that does produce H_2 for use as a clean and essentially boundless fuel source to replace climate-warming fossil fuels.

A quick look at standard reduction potentials (SRP) shows that the passing of electrons from water (dioxygen SRP = +0.816 V) to plastoquinone (approx SRP of 0.11) is **not** thermodynamically favored. The process is driven thermodynamically by the energy of the absorbed photons.

The crystal structure of PSII from a photosynthetic cyanobacterium consists of 17 polypeptide subunits with metal and pigment cofactors and over 45,000 atoms. Of particular interest is the **P680 chlorophyll reaction center**, which consists of four monomeric chlorophylls adjacent to a key Tyr 161 side chain. When H_2O gets oxidized to form dioxygen, 4 electrons must be removed by photoactivated P680. In PSII, this process occurs in 4, single electron steps, with the electrons first being transferred to the oxygen-evolving complex. The electrons passed through the Mn complex are delivered to P680 by a photoactive **Tyr 161 (Tyr Z or YZ)** free radical.

Five discrete intermediates of the OEC, S_0-S_4, are suggested from the experimental data and are consistent with the Kok cycle, which we will discuss below. These were postulated from experiments in which spinach chloroplasts were illuminated with short light pulses. A pattern of dioxygen release was noted that repeated after 4 flashes. Ultimately, light absorption by P680 forms excited state P680*, which donates an electron to **pheophytin**, which passes them to quinones. Hence P680 gets photooxidized as it forms the cationic $P680^+$. This then removes an electron from Tyr 161 (YZ) producing the tyrosine radical cation, **Tyr 161$^{\cdot+}$**. Given its positive charge, its reactive nature as a free radical, and its proximity to Mn ions in the OEC, it pulls an electron from a Mn ion in the OEC. This process repeats itself 4 times for the oxidation of two H_2Os, which injects 4 electrons back into the OEC to return to the basal state.

The mechanism is very complicated and still not fully understood. It is perhaps easiest to think about the mechanism involving a series of sequential electron and proton transfer and their accompanying change in charge and redox states. Most biochemistry students have a limited understanding of transition state complexes and chemistry but even the experts struggle with the mechanism.

In summary, for PSII in plants:

1. a pair of chlorophylls (P680) in the D subunits absorb light (maximum absorbance around 680 nm) and reach an excited state
2. electron transfer from P680 to a nearby chlorophyll with a lower energy level for the excited state electron occurs, which produces an anionic chlorophyll. This chlorophyll has 2 H^+ ions in the chlorophyll instead of Mg^{2+} (again note the charge balance). After electron transfer, P680 now becomes the cation $P680^+$.
3. This "anionic" chlorophyll transfers an electron to oxidized plastoquinone.
4. The $P680^+$, a strong oxidizing agent, removes one electron from an adjacent Tyr 161 to reform P680 and the radical cation **Tyr 161$^{\cdot+}$**. Its proximity to the OEC complex leads to it removing an electron from the OEC, making it a more potent oxidizing agent
5. This process repeats a total of 4 times to fully oxidize two water molecules to produce 1 O_2 with the 4 electrons removed from 2 glasses of water added back to the metal centers of the OEC.

This suggests that there are 5 states of the OEC, an initial state, which we will call S_0, and four other states (S_1, S_2, S_3, and S_4). S_1 forms after the removal of one electron from the OEC by the adjacent radical cation **Tyr 161$^{\cdot+}$** (formed after absorption of one photon). S_2, S_3, and S_4 are sequentially formed after the removal of one electron by a newly regenerated **Tyr 161$^{\cdot+}$** after another round of photoexcitation. S_4 then returns to its original state, S_0. This series of reactions is called the Kok cycle, which is shown in Figure 20.2.3.

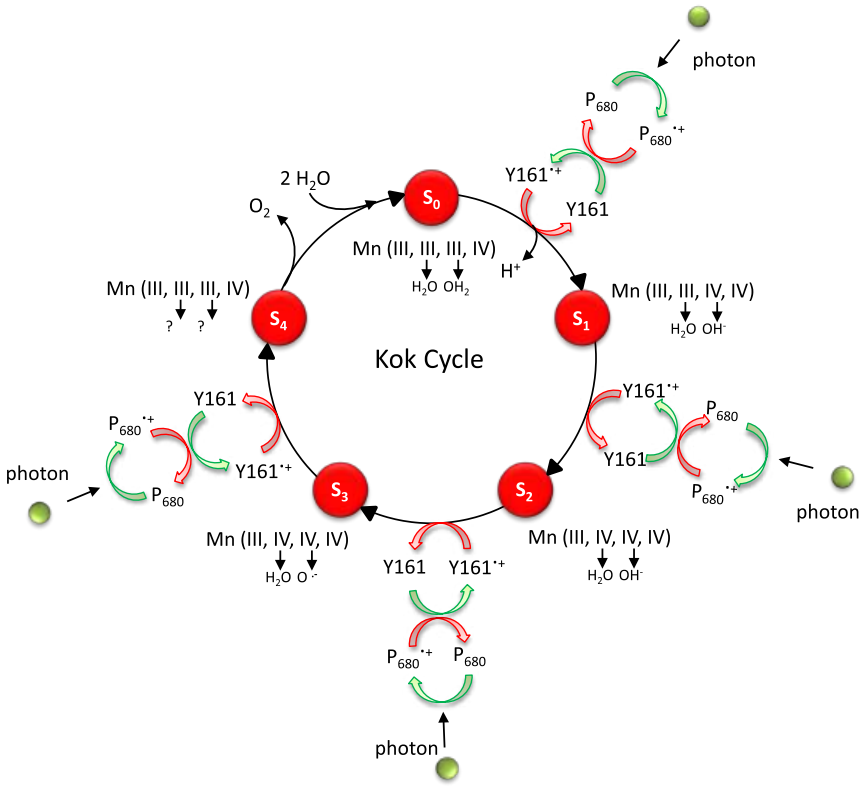


Figure 20.2.3: The Kok cycle

There is no structural information given in the above figure. What is shown instead are possible and consistent oxidation numbers of the four Mn^{n+} ions in the OEC that are consistent with charge balance and the changes in the oxidation number (-2) of the oxygen atom in water as it progresses to O_2 with an oxidation number of 0. The Mn ion states in the Kok diagram denote different discrete oxidation states where n is the number of oxidative “equivalents” stored in the OEC during cycle progression. Think of the OEC as the key catalyst, which will interact with substrate H_2O molecules. We start with the S_0 state and must return to it in the full cycle.

Remember that when O_2 acts as an oxidizing agent in combustion reactions, it forms $2H_2O$. That requires the addition of four electrons. If done sequentially, the oxygen intermediates include superoxide, peroxide, and oxide, the latter of which when protonated is water. Hence two waters and four cycles are required to remove the four electrons required to produce dioxygen. Intermediate but transient oxygen states are also presumably important in this mechanism.

A similar mechanism is found in PSI, except plastocyanin, not dioxygen is oxidized, with electrons moved to ferredoxin. This is likewise a difficult process since the reduction potential for oxidized plastocyanin (the form that can act as a reducing agent) is +0.37 while for ferredoxin it is -0.75. This transfer of electrons is an uphill thermodynamic battle since the more positive the standard reduction potential, the better the oxidizing agent and the more likely the agent becomes reduced. What drives this uphill flow of electrons. Of course, it is the energy input from photon. We won't go into any more detail about PSI since it is very similar to PSII but of course, does not have the OEC.

20.2.3: THE OXYGEN EVOLVING COMPLEX - OEC

Even though this is not a bioinorganic textbook, we must move past the "simple" Kok cycle diagram and look at the actual structure of the minicatalyst, the OEC, and the protein and water (substrate) environment around it to understand the mechanism. The mechanism of the OEC is still not fully understood. It's experimentally difficult to unravel given its complexity as the intermediates are very labile and the x-ray-induced transient alterations in the structure of OEC complicate matters more. Paradoxically it is quite simple in overall terms. Here is the essential reaction:

$$2H_2O + 4 \text{ photons} \rightarrow 4\ H^+(\text{lumen}) + 4\ e^- + O_2.$$

The crystal structure of PS2 from T. vulcanus has significantly improved our understanding of the OEC and electron flow on water oxidation. We will concentrate on developing an understanding of the amazing Photosystem II from Thermosynechococcus vulcanus, a cyanobacterium (19 subunits with 35 chlorophylls, two pheophytins, 11 beta carotenes, 2 plastoquinones, 2 heme irons, 1 non-heme iron, 4 Mn ions, 3-4 Ca ions, 3 Cl ions, 1 carbonate ion, and around 2800 water molecules).

Nature has appeared to evolve a single gene for the central protein in PSII that binds the OEC. The cluster, Mn_4CaO_5, appears identical in all photosynthetic organisms and is shown below. Researchers were surprised to find that the Ca ion was an integral part of the basic geometric “framework” of the OEC instead of a Mn which was found to be “dangling” from the basic geometric framework. A detailed structure of the OEC from T. vulcanus is shown in Figure 20.2.4.

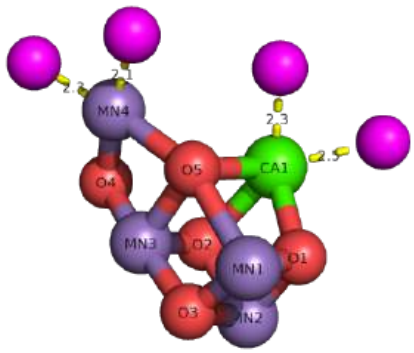
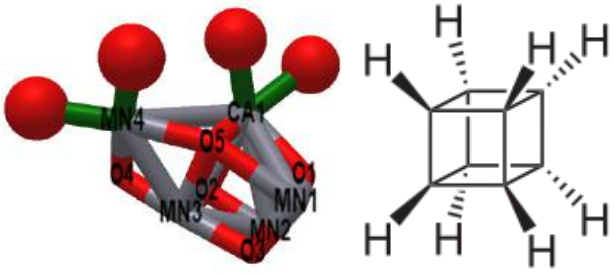


Figure 20.2.4: Structure of the OEC

Note the basic structure is a distorted cube the metal ions at every other corner separated by oxides. Again, it was a surprise that not all of the 4 Mn^{2+} ions were in the cubic structure. Note one "dangling" Mn^{2+} with the other last metal site in the distorted cube occupied by Ca^{2+}. Four oxygens (presumably from waters) are shown interacting with MN4 and CA1.

It is very difficult to visualize this structure correctly from a 2D figure. Figure 20.2.5 shows an interactive iCn3D model of the OEC with bound water of photosystem II from Thermostichus vulcanus (3WU2) which should help in visualizing this structure. (very long load time!)



NCBI iCn3D Figure 20.2.5: OEX with bound water of photosystem II from Thermostichus vulcanus (3WU2). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...F8vwgkDXV6C1S6. (very long load time)

The shape outlined by O5-CA-O1-MN1 and MN3-O2-MN2-O3 is similar to cubane (shown above right)

Now let's zoom out and view some of the amino acid side chains that interact with the OEC from T. vulcanus. These are shown in Figure 20.2.6.

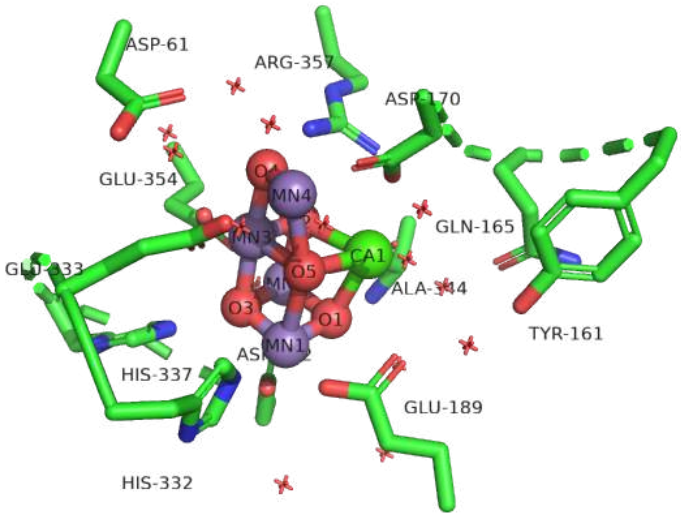
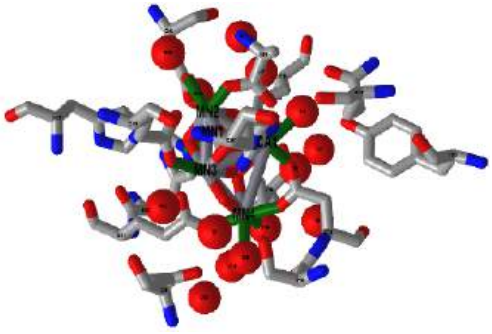


Figure 20.2.6: OEC and surrounding amino acids from T. vulcanus

The coordination number for all the Mn ions (including those interacting with water) is identical. Note the proximity of Tyr 161 which is involved in electron removal from the OEC after it becomes the radical cation **Tyr 161**$^{\cdot+}$ in the primary photooxidation event.

Figure 20.2.7 shows an interactive iCn3D model of the OEX with surrounding amino acids and water in photosystem II from Thermostichus vulcanus (3WU2). (long load time) presented to once again help you better understand the 2D structure shown above.



NCBI iCn3D Figure 20.2.7: OEX with surrounding amino acids and water in photosystem II from Thermostichus vulcanus (3WU2). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...bP6ttbaWUmQAg7

The OEC can be thought of as a distorted cubane with a MN3-O4-MN4-O5 back. Bonds to O5 are longer than the other bonds which suggest they are weaker than the other metal-oxygen bonds. This could suggest that it may not be an oxo (O^{2-}) ligand but another variant

such as OH^- (lower charge), which may imply involvement in the splitting of dioxygen in the reaction mechanism. From a mechanistic perspective, an **O-O bond must form between two waters**. Both sets of bound "waters" (purple spheres in Figure 4 and red spheres in Figure 5 shown without Hs attached if they are present) are close to O5.

It is important to remember that electrons removed from the metal ions in the OEC by **Tyr 161$^{\cdot+}$** must be restored to the OEC to allow the catalytic cycle to continue. These electrons come from the waters that get oxidized. Such a reversible loss and gain of electrons most readily occur from the transition state Mn ions, which you all remember from introductory chemistry have multiple oxidation states. To bring back introductory chemistry again, we present the standard reduction potentials of different Mn ions in Table 20.2.1 below.

Reduction reaction	Standard Reduction Potential
Mn^{2+} (aq) + 2 e^- → Mn (s)	-1.185
MnO_4^- (aq) + 2 H_2O (l) + 3 e^- → MnO_2 (s) + 4 OH^-	+0.595
MnO_2(s) + 4H^+ + e^- → Mn^{3+} + 2H_2O	+0.95
MnO_2(s) + 4H^+ + 2e- → Mn^{2+} + 2H_2O	+1.23
MnO_4- (aq) + 8 H^+ (aq) + 5 e^- → Mn^{2+} (aq) + 4 H_2O (l)	1.507
MnO_4^- (aq) + 4 H^+ (aq) + 3 e^- → MnO_2 (s) + 2 H_2O (l)	1.679
$HMnO_4^-$ + 3H^+ + 2^{e-} → MnO_2(s) + 2H_2O	+2.09
O_2(g) + 4H^+ + 4e^- → 2H_2O +1.229	+1.229

Table 20.2.1: Standard reduction potentials for Mn ions compared to O_2.

You should be able to determine the oxidation number of the Mn ion in each compound. Based on standard reduction potentials, which oxidation states might be sufficient for the oxidation of H_2O in the OEC?

How does this translate into structural/chemical changes in the OEC? Figure 20.2.8 provides a recent mechanism consistent with each of the Kok states (S_0-S_4).

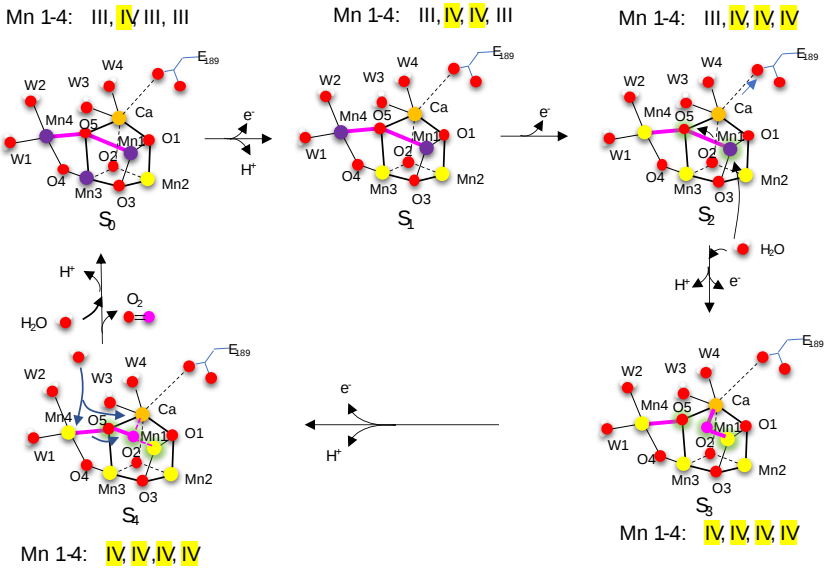


Figure 20.2.8: Mechanism for water oxidation and O_2 formation through the OEC consistent with the Kok cycle (adapter from Britt & Marchiori, 10.1126/science.aaz4522)

The proposed change in redox state for each Mn ion is illustrated with Mn (III) ions in purple and Mn (IV) ions in yellow. Note the change in the oxidation state of the 4 Mn ions from (III, IV, III, and III) in S_0 to (IV, IV, IV, and IV) for all of them in S_3 and S_4. A flip in the side chain of E180 in S_2 allows the binding of Mn1 through an oxy link to Ca. There are many possible different forms of oxygens in the structure including waters, oxides (bridging oxos and possibly terminal oxides), and hydroxides, and the exact form at some sites are still a bit uncertain. Note also the elegance of having a Mn_4 cluster to catalyze the 4 electron oxidation of 2 water through the loss of 4 electrons. Also, the redox state change in the Mn ions is different than the one shown in the Kok diagram in Figure 3. In addition, the final O2-producing step going from S_4 → S_0 is still uncertain. The above mechanism is based on x-ray structures of intermediates and quantum calculations. In it, S_4 has an Mn(IV)O· that bonds with the bridging O5 to form O_2.

20.2.4: WATERS

As water is a reactant in PSII, there must be water channels leading to the OEC that provide a way for water to enter and for protons to be removed and directed to the lumen to develop a proton gradient. Another rendering of the Kok cycle, the position of the OEC in PSII on the luminal side of the membrane, and the presence of water channels (Cl1, O4, O1) and the Yz network, which connect Tyr 161 (Yz) to the lumen, are shown in Figure 20.2.9.

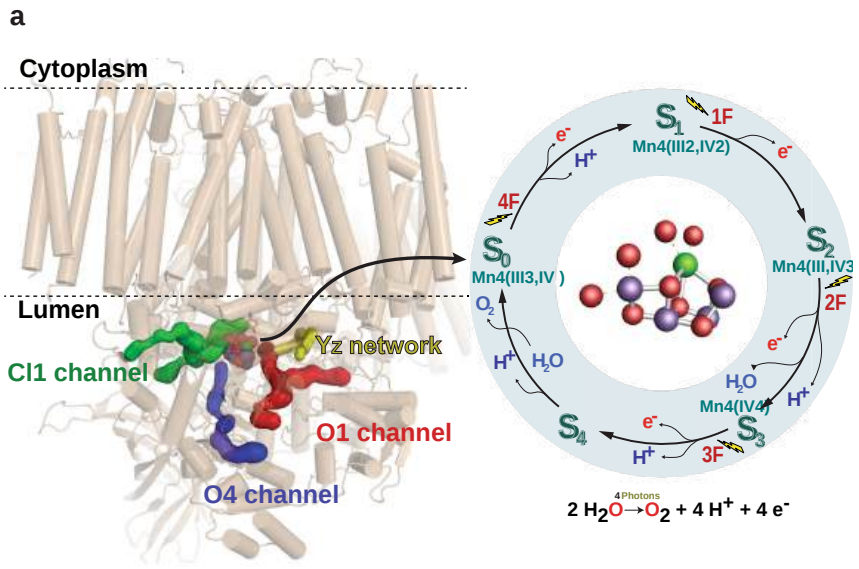


Figure 20.2.9: An overview of Photosystem II and the main water channels and networks from the OEC to the lumenal side. Hussein, R., Ibrahim, M., Bhowmick, A., *et al*. Structural dynamics in the water and proton channels of photosystem II during the S_2 to S_3 transition. *Nat Commun* **12,** 6531 (2021). https://doi.org/10.1038/s41467-021-26781-z. Creative Commons Attribution 4.0 International License. http://creativecommons.org/ licenses/by/4.0/.

The left part of the figure illustrates the structure of PS II showing the membrane-embedded helices and the extrinsic subunits in beige. The OEC and the water channels, in addition to the Yz network, are shown in color. The Kok cycle of the water oxidation reaction that is triggered by the absorption of photons is shown on the right and highlighted with a blue circle. F represents a photon.

A more detailed representation of water channels and the Yz networks is shown in Figure 20.2.10.

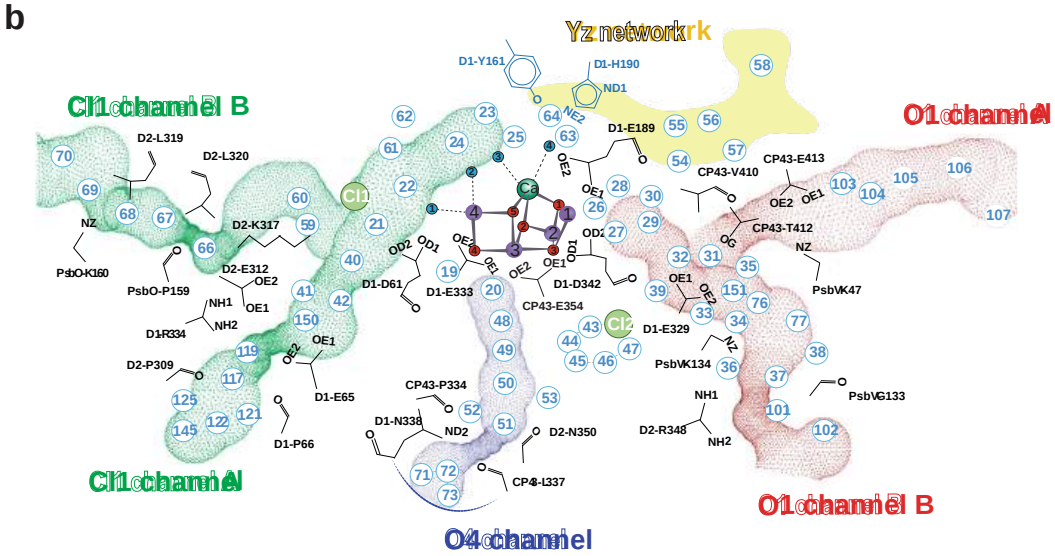


Figure 20.2.10: A detailed view of the water channels showing the waters within each channel (O1 Channel red dotted, O4 Channel blue dotted, and Cl1 Channel green dotted). The region highlighted with solid yellow represents the Yz network. Residues involved in forming bottlenecks in the channels are shown in black. Water molecules found in the crystal structures are labeled with blue numbers with the channel. Hussein, R., Ibrahim, M., Bhowmick, A., *et al.*, ibid.

These structures show that there is no direct water pathway from across the OEC and that all channels restrict water movement to some degree. The O4 and Cl1 channels are narrower than the O1 channel so water in those is less mobile. In the O4 channels, waters 50-53 are

near charged groups and are close to a major bottleneck (residues D1-N338, D2-N350, and CP43-P334, -L334).

Based on the x-ray structures and molecular dynamic simulations, it appears that the O1 channels allow access of water to the OEC. The Cl1 channel A, which is more rigid, may be involved in H^+ transfer during $S_2 \rightarrow S_3$.

20.2.5: THE LAST STEP: ELECTRON TRANSFER TO PLASTOQUINONE

We are almost ready for the next section in which we will present the flow of electrons away from PSII through mobile electron carriers, leading to the synthesis of NADPH for the reductive biosynthesis of carbohydrates. Before we leave PSII, let's look at what happens to the radical anion $P680^-$, also known as $PheA^-$ (**pheophytin A**) or $PheA_{D1}$, the reaction center chlorophyll without a central Mg^{2+} ion) which received an electron from photooxidation of the reaction center P680, as summarized again in Figure 20.2.11.

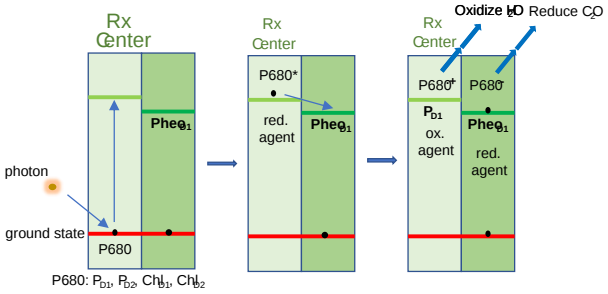
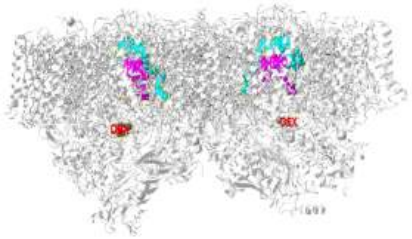


Figure 20.2.11: The reaction center electron transfer reactions in PSII.

Pheophytin A$^-$ passes its electron to plastoquinone A (in PSII). which passes it on to the lipophilic mobile electron carrier in the thylakoid membrane plastoquinone. It is similar to the mobile electron carrier in mitochondrial electron transport, ubiquinone. Ultimately these are passed to $NADP^+$ to form NADPH for reductive biosynthesis.

Figure 20.2.12 shows an interactive iCn3D model highlighting just the OEX, pheophytin A (PHO), and plastoquinone A (PL9) in photosystem II from Thermostichus vulcanus (3WU2). (long load time).



NCBI iCn3D Figure 20.2.12: the OEX, pheophytin A (PHO), and plastoquinone A (PL9) in photosystem II from Thermostichus vulcanus (3WU2). (long load time). (Copyright; author via source). Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...H7Mycbx1PsFbt6

The OEX (OEC) is shown in spacefill, CPK colors, PHO is shown in spacefill magenta, and PL9 in spacefill cyan. Note the proximity of PHO and PL9 for easy electron transfer to plastoquinone A.

Given the number and proximity of high-energy reactive species in the reaction center, it should not be surprising that side reactions can occur. These would decrease the efficiency of light energy transduction and also could damage molecular components of PSII (and similarly PSI)

Figure 20.2.13s shows a standard reduction potential diagram for the **P**680 - **Phe**ophytin A - Plasto**Q**uinone **A** triad (abbreviated P Phe Q) in PSII.

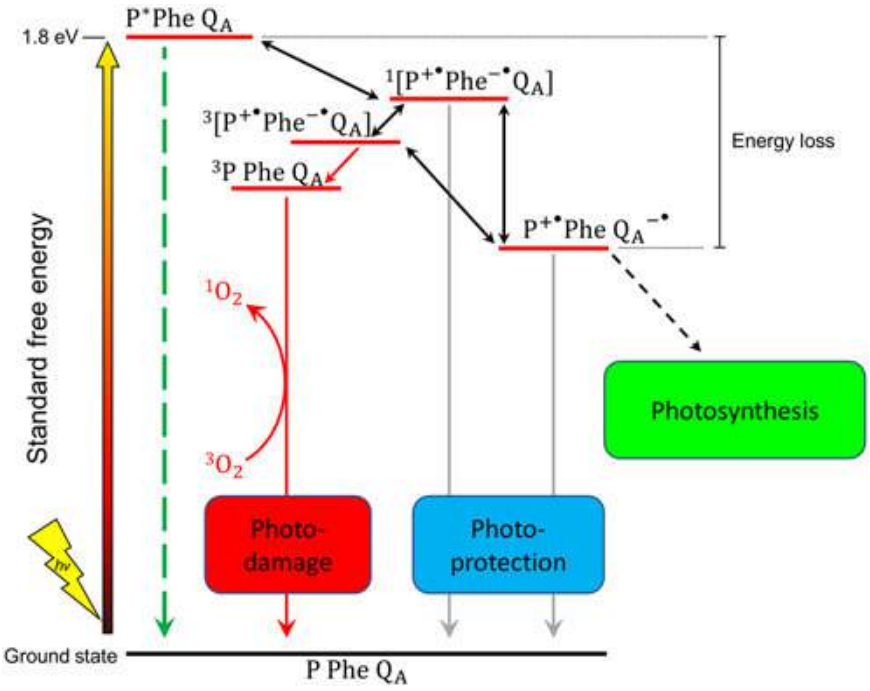


Figure 20.2.13: Charge recombination pathways in PSII (The energy gaps are not drawn to scale). Cardona T, Shao S, Nixon PJ. Enhancing photosynthesis in plants: the light reactions. *Essays Biochem*. 2018;62(1):85-94. Published 2018 Apr 13. doi:10.1042/EBC20170015Creative Commons Attribution License 4.0 (CC BY).

Safe routes for charge recombination between P^+ and Q_A^- are indicated in blue, the damaging route producing $^1O_2$ in red, and the radiative pathway in green. Stabilization of the $P^+PheoQ_A^-$ state helps prevent reverse electron flow to form $P^+Pheo^-Q_A$ and subsequent charge recombination to form $PPheoQ_A$. For clarity, the details of the additional electron transfer steps, including the oxidation of water and the reduction of plastoquinone to plastoquinol by PSII, collectively termed photosynthesis, are omitted. Abbreviations: P, primary electron donor of PSII; Pheo, pheophytin electron acceptor; Q_A, primary plastoquinone electron acceptor; $^1O_2$, singlet oxygen; $^3O_2$, triplet oxygen; ^3P, triplet excited state of P.

20.3: PLANT ELECTRON TRANSPORT AND ATP SYNTHESIS

Search Fundamentals of Biochemistry

20.3.1: INTRODUCTION

In the previous sections, we studied light absorption by chlorophylls, the transfer of energy to the reaction center of photosystem II, the oxidation of H_2O by the oxygen-evolving complex (OEC), and the transfer of electrons from these events to the lipophilic carrier of electrons, plastoquinone. Now we are ready to see how the process continues as electrons are passed on from reduced plastoquinone to the cytochrome b_6f complex, through photosystem I (which has no OEC) and on to the terminal electron acceptor $NADP^+$, which forms NADPH. This is used for reductive biosynthesis of glucose after fixation of atmospheric CO_2 by ribulose bisphosphate carboxylase (RuBisCo). As we saw in mitochondrial electron transport, this passage of electrons is accompanied by the movement of protons from the lumen to the stroma with the ultimate collapse back into the lumen through a rotatory ATP synthase to form the ATP required for reductive biosynthesis. Figure 20.3.1 reviews again the light reactions of photosynthesis.

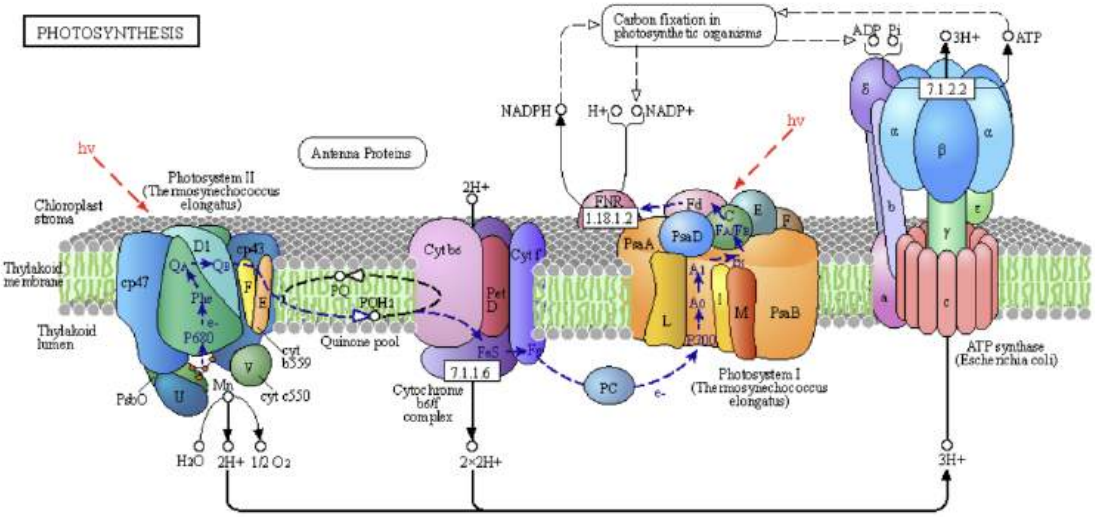
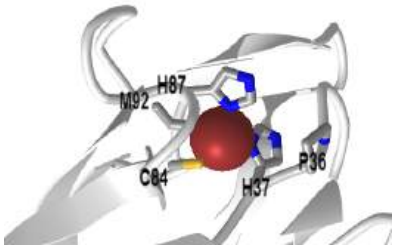


Figure 20.3.1: Light reactions of photosynthesis

20.3.2: CYTOCHROME B_6F

This complex moves electrons from the mobile lipophilic electron carrier reduced plastoquinol (PQH_2), an isoprenoid quinone, to the mobile Cu-containing protein **plastocyanin**, which plays an analogous role to the mobile protein carrier in mitochondrial electron transport, cytochrome C. It catalyzes the rate-limiting step in electron transport in the light reactions. Figure 20.3.2 shows an interactive iCn3D model of the spinach plastocyanin (1AG6)



NCBI iCn3D Figure 20.3.2: Spinach plastocyanin (1AG6). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...DwSDDrYMUb9Xn7

The complex that mediates electron transfer between reduced plastoquinone and the protein plastocyanin, cytochrome b_6f, is centrally positioned between the two photosystems. In addition, it moves 2 H^+s from the stroma into the lumen. These are joined by 2 H^+s from the oxidation of water by the OEC (on the luminal side of PSII) to create a transmembrane proton gradient, which will power ATP synthesis.

Electrons transfer with cytochrome b_6f takes place through the quinol (Q) cycle in a similar fashion to Complex III in mitochondrial electron transport so we won't go into much detail here. Figure 20.3.3 shows a summary diagram with electron and proton flow.

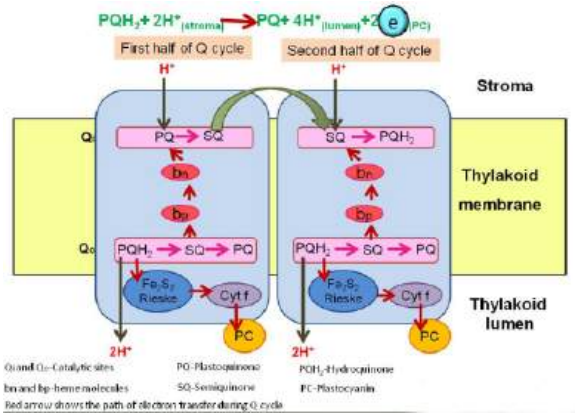


Figure 20.3.3: Q cycle in the cytochrome b6f complex of photosynthesis. https://twitter.com/BiologyNowadays/...232768/photo/1

Figure 20.3.4 shows another version of the Q cycle as an alternative representation.

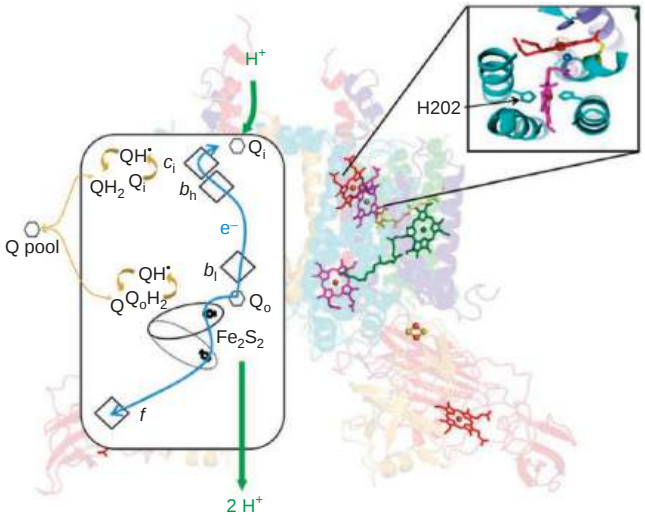


Figure 20.3.4: Alternative version of the photosynthetic cytochrome b_6f Q cycle. Malnoë et al. *Nat Commun* **2,** 301 (2011). https://doi.org/10.1038/ncomms1299. Creative Commons Attribution-NonCommercialShare Alike 3.0 Unported License. http://creativecommons.org/licenses/by-nc-sa/3.0/

The left box shows the *b*$_6$*f* complex transfers two protons (green arrows) per electron transferred (blue arrows) along high (Fe_2S_2 cluster, cytochrome $_f$) and low potential chains (b_l, bh, c_i hemes) as well as Quinol (QH_2) oxidation at Q_o site, Quinone (Q) reduction at Q_i site. The right structure depicts haems *b* (purple), c_i and f (red), Fe_2S_2 cluster (yellow and orange ball-and-stick model), cytochrome b_6 (cyan), subunit IV (blue), Rieske subunit (yellow), cytochrome *f* (red), PetG, L, M and N subunits (green). Magnification of Q_i site comprising b_h and c_i haems.

Both cytochrome bc1 and cytb_6f are dimeric complexes, with 2 Fe_2S_2 clusters, two cytochrome bs, and a cytochrome c. The cytb_6f complex also has 9-cis β-carotene and additional c1 heme. The electrons move from PQH_2 through the complex in a similar fashion as in the bc1 complex.

PQH_2 is oxidized at the Q_p site with a bifurcation of electrons:

- one electron moves through the high potential Fe_2S_2 center and cyt f pathway (bottom left in pathway diagram of Figure 3), with the electron moving to the soluble peripheral protein plastocyanin and one to photosystem I.
- the other moves through the low potential b_l, bh, c_i hemes pathway (top left in pathway diagram of Figure 3), with the electron moving to a plastoquinone at the Qn site near the stroma. A second round of oxidation of PQH_2 at the Q_p site eventually donates another electron to a plastoquinone$^{-\cdot}$ which regenerates PQH_2 after addition of 2 H^+ from the stroma.

Having two successive oxidations of PQH_2 leads to twice the number of protons moving into the lumen. In the process 2 H^+ move into the lumen. Cytochrome b_6f also is a redox sensor of the status of the plastoquinol/plastoquinone pool.

Figure 20.3.5 shows an interactive iCn3D model of the spinach cytochrome b6f complex (6RQF)

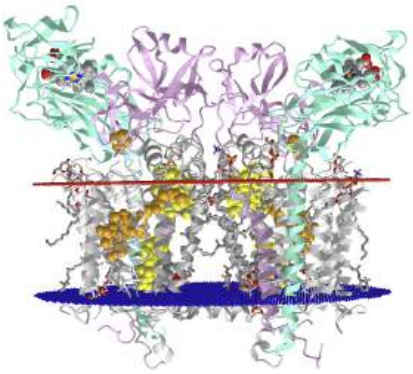


Figure 20.3.5: Spinach cytochrome b6f complex (6RQF). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...tgUhjekB4Tndw6

The region away from the red leaflet represents the lumen side of the complex. The FeS cluster and heme C are in the luminal domains. The coloring scheme is as follows:

- cytochrome bcf subunits 4, 5, 6, 7, and 8: light gray
- cytochrome b6: dark gray
- cytochrome b6 FeS subunit: plum
- cytochrome F: cyan
- chlorophylls: orange spacefill, labeled
- FeS clusters: CPK spacefill, labeled
- HEC - Heme C: CPK spacefill, labeled
- heme - porphoryin IX containing Fe: yellow spacefill

Cytochrome b_6f is the rate-limiting step for electron flow but what it's role in regulating the photosynthetic pathway (light reaction plus the dark reaction of carbon metabolism? Data suggest that the complex regulates electron transport in low light conditions but effects a switch to carbon metabolism under saturating light. Johnson and Berry have analyzed electron flow with carbon metabolism in a fashion analogous to transistors in a circuit board, as illustrated in Figure 20.3.6.

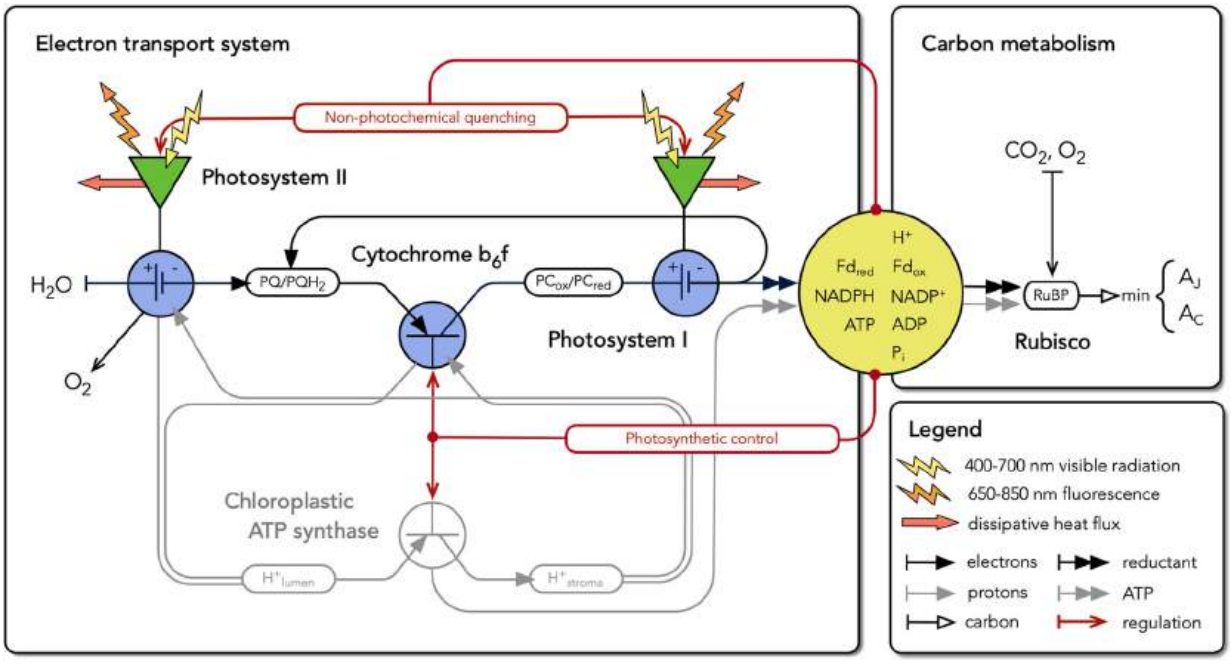


Figure 20.3.6: Electron transport system as an electrical circuit (and associated text below). Johnson and Berry. Photosynthesis Research (2021) 148:101–136. Creative Commons Attribution 4.0 International License. http:// creativecommons.org/ licenses/by/4.0/.

They define a transistor as a regulated circuit element that uses variable conductance to control current flow. The linear flow of electrons from water to reductant is viewed as a light-driven current that is under the control of many regulatory feedbacks stemming from carbon metabolism. In limiting light, Cyt b_6f presents maximal conductance to flow, and feedback from carbon metabolism adjusts the excitation of

PS I and PS II in such a way as to balance the relative rates of linear and cyclic electron flow to the NADPH, Fd, and ATP requirements of the sinks. When the light becomes saturating, feedback from carbon metabolism also decreases the apparent conductance of Cyt b_6f, controlling the linear flow of electrons through the plastoquinone pool and the associated flow of protons into the thylakoid lumen. In this way, the regulation of Cyt b_6f simultaneously permits efficient photosynthesis and protects the system from photodamage.

This model is organized around the idea that the distribution of excitation between PS II and PS I is regulated in such a way as to minimize losses of absorbed light and maximize potential electron transport through Cyt b_6f. The expression for the potential electron transport rate has the form of a Michaelis-Menten expression for a single substrate (i.e., light), but describes the kinetic behavior of the entire electron transport chain (i.e., including both photochemical and biochemical reactions). It predicts that electron transport has a hyperbolic dependence on irradiance, with the maximum efficiency realized at the limit where absorbed irradiance goes to zero and the maximum speed realized at the limit where absorbed irradiance is infinite.

The trade-off between the speed and efficiency of potential electron transport is driven by the need for the supplies of reduced plastoquinone and oxidized plastocyanin to be balanced to sustain Cyt b_6f turnover at the maximum catalytic rate. This causes progressive closure of the PS II and PS I reaction centers, with PS II accumulating in a reduced state and PS I in an oxidized state. As the excitation pressure on PS II and PS I increases, the closure of the reaction centers causes the photochemical yields of PS II and PS I as well as the absorbed quantum yield to decrease as the potential electron flow through Cyt b_6f and the potential photosynthetic rate increase"

The key prediction of the expression for the potential electron transport rate is that the maximum activity of Cyt b_6f defines the upper limit for the theoretical maximum speed of electron transport. The expressions describing feedback control over Cyt b_6f activity are based on the idea that Cyt b_6f functions like a transistor, i.e., a component of an electrical circuit that uses variable conductance to control current.

The fact that Cyt b_6f can modulate its conductance to linear electron flow within milliseconds of a perturbation in light suggests that photosynthetic control is the first line of defense against overexcitation, protecting the acceptor side of PS I from being flooded with highly reduced intermediates.

In response to a sustained increase in light, the induction of photosynthetic control is followed by the induction of PQ_N. As NPQ alleviates the electron overpressure in the PQ pool, photosynthetic control progressively relaxes. The two forms of regulation gradually settle to a steady state at the new light intensity. This interaction seems to allow electron transport to proceed at the Cyt b_6f-limited rate under low light intensities, and then smoothly switch to the Rubisco-limited rate once the light intensity is high enough to become saturating. It also seems to allow photosynthesis to operate safely and efficiently in a wide range of biochemical milieus, from those characteristic of natural variation in photosynthetic capacity (with balanced electron transport and carbon metabolism) to those characteristic of genetic manipulations (with imbalances in electron transport and carbon metabolism).

In this framework, the excitation balance of PS II and PS I and the maximum activities of Cyt b_6f and Rubisco emerge as key limits on system dynamics. For example, the trade-off between the speed and efficiency of electron transport is shown to be controlled by the excitation balance of PS II and PS I and the maximum activity of Cyt b_6f. The development of PQ_N is shown to be controlled by the excitation balance of PS II and PS I and the demand for linear electron flow (LEF) through the light reactions and circular electron flow (CEF) around PSI. The onset of photosynthetic control is shown to be dependent on the maximum activities of Cyt b_6f and Rubisco.

20.3.3: PLANT PHOTOSYSTEM I-LHCI SUPER-COMPLEX

If the goal of the photosystem complexes is to transduce light energy delivered by photons into electrons that can be used for reductive biosynthesis of glucose from atmospheric carbon capture of CO_2, then PSI is quite amazing. It has a "quantum efficiency" close to 1 which implies that one absorbed photon produces 1 electron that can be used to reduce $NADP^+$. This happens since the transfer of photon energy to other molecules in PSI is so quick compared to nonradiative decay processes for the excited state chlorophylls.

The same process for excitation and electron (charge) transfer that we saw in PSII occurs in PSI, with the chlorophyll in the reaction center involved in charge transfers. The light-absorbing molecules of PSI enable light at the far red of the spectra to be absorbed. The complex has 16 proteins, 155 chlorophyll, and 35 carotene derivatives. As with PSII, there is a core complex that is similar to cyanobacterial PSI. The supercomplex has in addition light-harvesting complex I proteins around one side of the complex (looking down on it) that have four LHCI proteins (Lhca 1-4) which allow for more absorption of life. They have 57 chlorophylls and 13 carotene derivatives.

The apoproteins that bind the chlorophylls, including the LHCI proteins, have similar but slightly different topologies, allowing for tuning of the absorbance spectra of the bound chromophores. Seven of the chlorophylls when bound have local environments that allow absorption of far red light. These are not found in PSII. the far red light would be more abundant in the low parts of the plant canopies as photons of lower wavelength would be more filtered out by upper leaves in the canopy.

Because of the complexity of PSI, we will offer several different iCn3D models to illustrate different features of the **photosystem I-LHCI super-complex**. As with PSII, PSI has a core with a surrounding light-harvesting complex (LHCI). Figure 20.3.7 shows an interactive iCn3D model of the Plant (pea) photosystem I-LHCI super-complex (4XK8)

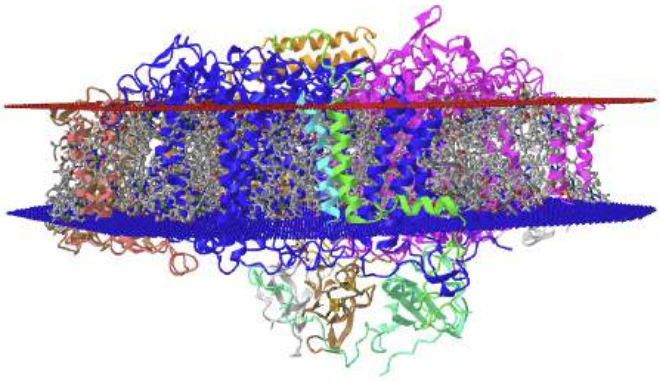


Figure 20.3.7: Plant photosystem I-LHCI super-complex (4XK8). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...uWBzFrVBUDHHQ6 (long load time)

Again the red leaflet represents the lumen side of the complex. Gray represents the chlorophyll and other lipids molecules in the complex. Here is the same complex but just the nonprotein components

Figure 20.3.8 shows an interactive iCn3D model of the lipid and FeS components of Plant photosystem I-LHCI super-complex (4XK8)

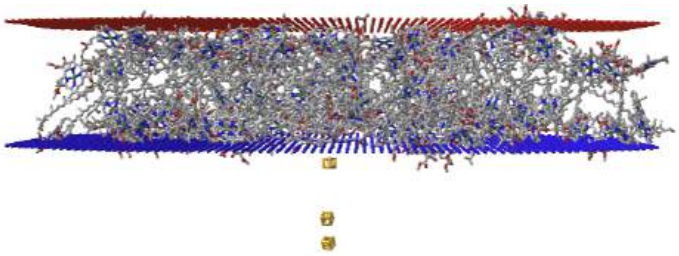


Figure 20.3.8: Lipid and FeS components of plant photosystem I-LHCI super-complex (4XK8). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?VtoCnxLRHQ8RPdESA (long load time)

Again the red leaflet represents the lumen side of the complex. Gray represents the chlorophyll and other lipids molecules in the complex. Rotate the image to see how the chlorophyll and other lipids encircle the membrane protein components.

Figure 20.3.9 shows an interactive iCn3D model of the ApoA1 protein with surrounding chlorophyll and other components from plant photosystem I-LHCI super-complex (4XK8)

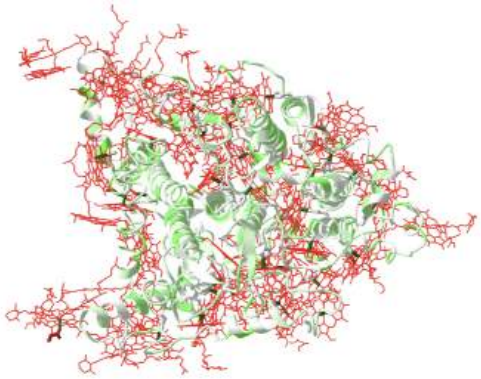
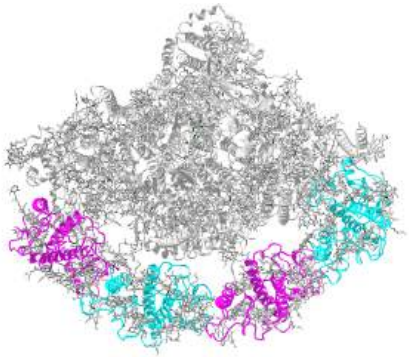


Figure 20.3.9: ApoA1 protein with surrounding chlorophyll and other components from plant photosystem I-LHCI super-complex (4XK8). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...L6tsJ3oMDZ2YQA (long load time)

ApoA1 is shown in cartoon and colored according to hydrophobicity. The chlorophylls, carotenes, and other components within 5 Å are shown in red. Rotate the image to see how the chlorophyll and other lipids encircle the membrane protein components.

Figure 20.3.10 shows an interactive iCn3D model of the Plant (pea) photosystem I-LHCI super-complex highlighting LHCIs (4XK8)



NCBI iCn3D Figure 20.3.10: Plant (pea) photosystem I-LHCI super-complex highlighting LHCIs (4XK8). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...9vHynkRTLxAct7 (long load time)

The LHCI subunits (Lhca 1-4) are shown in alternating magenta and cyan cartoon structure surrounding one side of the core complex. The view in the model above is a top-down view.

We discussed in Chapter 20.1 that PSI absorbance is tilted toward the red end of the spectra, with the effective absorbance ratio over a broad wavelength range for the two photosystems, **PSI**/(**PSI**+PSII) deviating from around 0.5 at the red/far red end of the spectrum (670-730 nm), where the ratio is close to 1. When exposed to far-red light, the systems move to state I. In this state, the major mobile antenna proteins (LHCIIs) move to PSII to restore a "photoabsorption" balance. When exposed to light depleted in the high end of visual spectra, the system moves to state II, in which mobile LHCIIs move to PSI.

The location of LHCs is regulated by the phosphorylation of LHCs controlled by the levels of plastoquinone, which makes great biological sense. When the concentration of plastoquinones in the reduced state, PQH_2 (plastoquinol), is high, it would be optimal to increase the activity of PSI to relieve the high concentration of the substrate for cyto b_6f and shift the system to higher PSI activity and continue electron flow. This regulation is mediated by the phosphorylation of LHCs by LHCII kinase (a Ser/Thr kinase), which is activated by high reduced PQH_2 concentrations. Phosphorylation of LHCs leads to their movement from PSII to PSI. When the oxidized form for plastoquinone is high, LHCII kinase is inactivated by dephosphorylation, causing the mobile LHC to move back to PSII to increase output (PQH_2) from PSII. These events occur in low light. In high-light conditions, when the system is functioning at a high level, the LHCII kinase is inhibited by stromal thioredoxin. Specifically, the phosphorylation status of Lhcb1 and Lhcb2 in LHCII homo- or heterotrimers determine the movement between PSII and PSI. The chloroplastic serine/threonine-protein kinase (STT7 also known as STN7) is another LHCII kinase.

20.3.4: CHLOROPLAST ATP SYNTHASE - CF_1F_OATPASE

Finally, let's take a quick look at the ATP synthase in chloroplasts. It is similar in structure and function to mitochondrial FoF1ATPase so we won't spend much time on the mechanism. Like its mitochondrial counterpart, it is a rotary enzyme that transduces the free energy of a proton gradient collapse into chemical energy in the form of ATP. One difference is the rotary enzyme should be regulated by light levels with its activity decreased at night. This is accomplished in higher plants by conversion between a reduced and oxidized state, an effective redox switch, in one subunit (γ) of the ATP synthase. The deactivation is important at night since if the rotary enzyme runs in reverse, ATP hydrolysis would ensue.

Figure 20.3.11 shows an interactive iCn3D model of the reduced R1 state of chloroplast ATP synthase (R1, CF1FO) (6VON). It should look familiar to you given its similarity to mitochondrial ATP synthase.

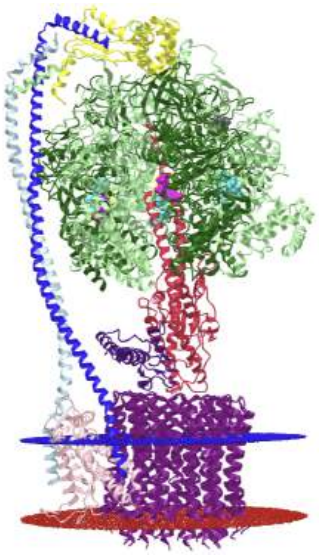


Figure 20.3.11: Reduced R1 state of chloroplast ATP synthase (R1, CF1FO) (6VON) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?hbnDCm1Kgqgu2k1NA (long load)

- alpha - light green
- beta - dark green
- gamma - yellow
- gamma - crimson
- epsilon - indigo
- b - blue
- b' - light blue
- a - light pink
- c ring - purple
- ATP - spacefill, cyan
- TENTOXIN (TTX) - space fill gray
- ADP - spacefill magenta

In the oxidized state, there is a disulfide in the γ subunit which inhibits torsion by stabilizing two β hairpins. This constraint is relieved on reduction and rotation is enhanced. In the reduced structure, tentoxin, an uncompetitive inhibit is present which allowed the structure to be determined.

below Yang, JH., Williams, D., Kandiah, E. *et al*. Structural basis of redox modulation on chloroplast ATP synthase. *Commun Biol* **3,** 482 (2020). https://doi.org/10.1038/s42003-020-01221-8. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

At night, the enzyme is in the low-activity form. When light becomes available in the day, the pH gradient collapse leads to the release of tightly bound ATP from the oxidized state of the enzyme. PSI can engage in photon-induced electron transfer to ferredoxin. This binds to NADP reductase and leads to NADPH formation. These processes are illustrated in Figure 20.3.12.

Figure 20.3.12: Fd bound to PSI is ferrodoxin and FNP is ferrodoxin NADP reductase.

Increased reduced ferredoxin leads to the formation of reduced thioredoxin, a small redox protein involved in redox signaling and protection of cells from oxidative stress. A similar process activates the enzyme sedoheptulose-1,7-bisphosphatase (SBPase) found in the dark reaction Calvin cycle of photosynthesis. Its activation is illustrated in Figure 20.3.13.

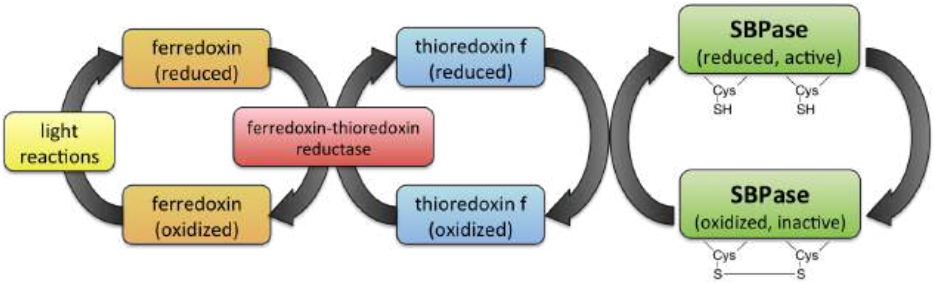


Figure 20.3.13: SBPase regulation by ferredoxin-thioredoxin system. https://commons.wikimedia.org/wiki/F...xin_system.png. Creative Commons Attribution-Share Alike 3.0 Unported l

Reduced thioredoxin also reduces and activates chloroplastic ATP synthase.

Figure 20.3.14 the activities and structures of the ATP synthase (CF_1F_O)

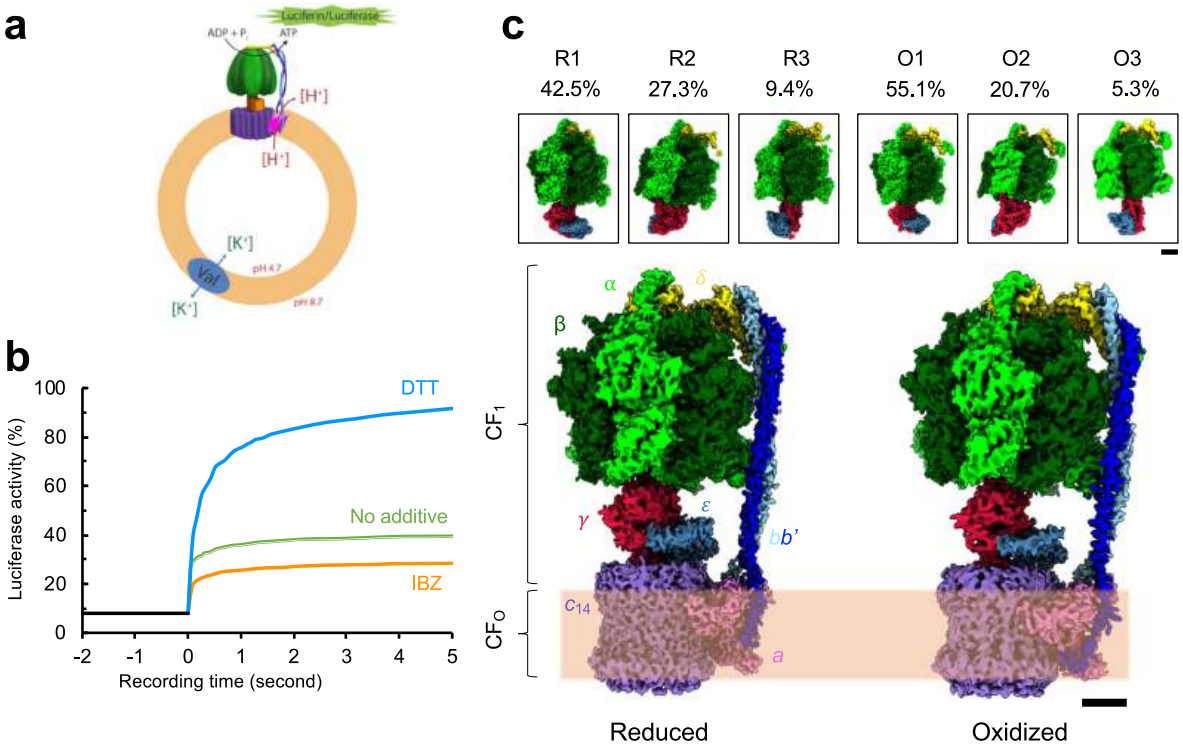


Figure 20.3.14 the activities and structures of the ATP synthase (CF_1F_O). Yang, JH., Williams, D., Kandiah, E. *et al.* Structural basis of redox modulation on chloroplast ATP synthase. *Commun Biol* **3,** 482 (2020). https://doi.org/10.1038/s42003-020-01221-8. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Panel a: For the experiments, purified CF_1F_O was reconstituted into a liposome (orange) mixed with lipids of phosphatidylcholine and phosphatidic acid. The generated pH gradient across the membrane drove the reconstituted CF_1F_O to synthesize ATP molecules, which were detected using a luciferin/luciferase assay (green). Val indicates valinomycin.

Pane b: The activity of the enzyme is shown in the reduced state (in the presence of dithiothreitol - DTT, blue curve), oxidized state (in the presence of iodosobenzoate - IBZ, orange curve, and with no additive (green curve). Note that the activity is reduced by about 80% in the oxidized state.

Panel c: The images are from Cryo-EM density maps of the oxidized and reduced forms of the CF_1F_O. The color codes are as follows: α (light green), β (dark green), δ (yellow), *bb'* (blue and light blue), γ (crimson), ε (indigo), *a* (light pink), and *c* ring (purple). R indicates a reduced state, and O indicates an oxidized state. The three-dimensional (3D) reconstructions are categorized into three different rotary states (states 1, 2, and 3). The upper insets are the density maps of the F_1 domains.

Figure 20.3.15 shows the details of the conformation changes in the γ subunit between the oxidized and reduced forms.

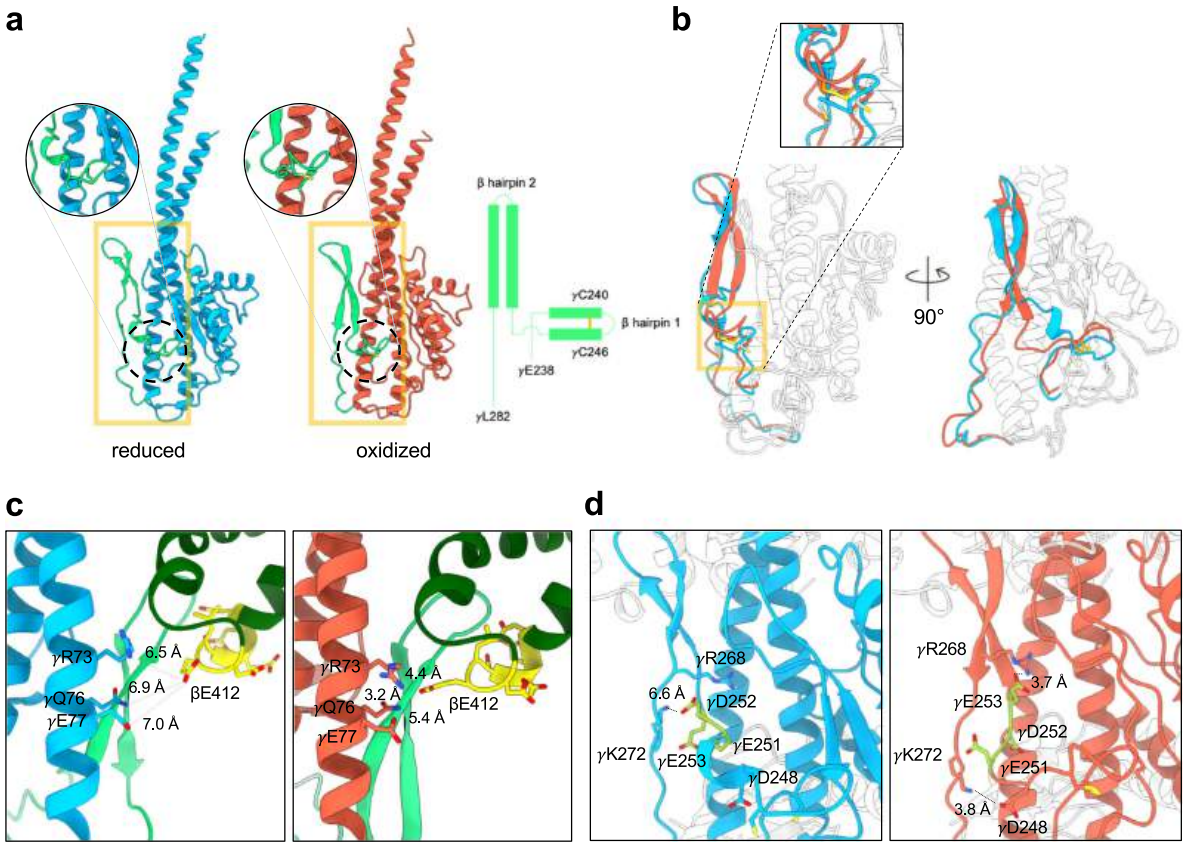


Figure 20.3.15: Conformation changes in the γ subunit between the oxidized and reduced forms

Panel a: Structures of the reduced (light blue) and oxidized (orange) γ subunits. Two β hairpin structures (from γGlu238 to γLeu282) are shown in light green, and the two cysteines of the redox switch are shown in yellow in circular enlarged views. The diagram on right shows the topology of the two β hairpin structures.

Panel b Superposition of the reduced and oxidized γ subunits (RMSD 1.016 Å). The two β hairpins are shown in light blue and orange for the reduced and oxidized forms, respectively. Other regions are shown in white.

Panel c Interaction networks of the β hairpin 2 and βDELSEED motif. The left and right panels are the reduced (γ subunit in light blue) and oxidized (γ subunit in orange) forms. Light green represents the β hairpin 2, dark green for the β subunit, and yellow for the βDELSEED motif. The distances connecting the residues of the γ coiled-coil (γArg73, γGln76, and γGlu77) with the βGlu412 are labeled.

Panel d Interaction of the EDE motif with the γ subunit. The EDE motif (yellow) does not interact with any part of the reduced γ subunit but forms an extensive interaction network with its neighborhood when the γ subunit is oxidized.

Figure 20.3.16 shows a cartoon model illustrating the structures of the overall oxidized and reduced states of ATP synthase

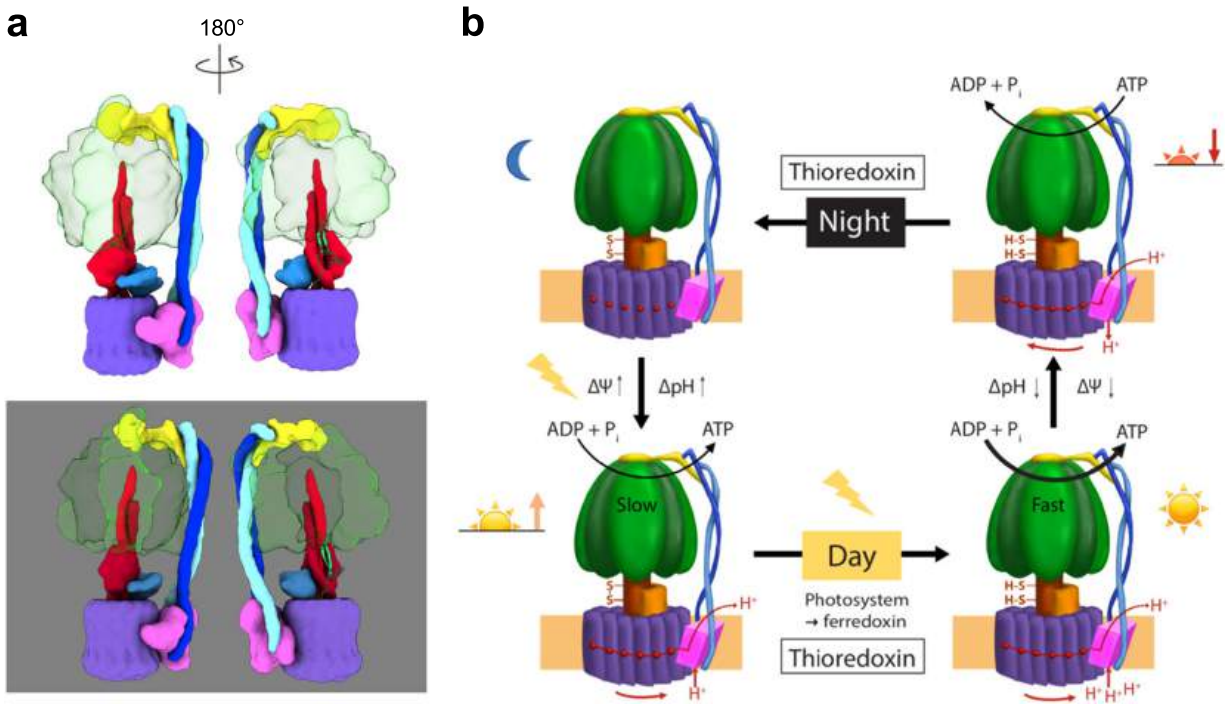


Figure 20.3.16: cartoon model illustrating the structures of the overall oxidized and reduced states of ATP synthase

Panel a: This shows a cartoon schematic of the redox modulation. The upper and lower models are the reduced and oxidized states, respectively. Color codes are the same as in Fig. 1c and the β hairpin structures of the γ subunit are shown in light green. The two redox states are aligned in the same view.

Pane b: At night, no energy input from light is available for the photosynthetic electron transport chain, and thus, no electrochemical potential (ΔΨ) and proton gradient (ΔpH) are generated. The oxidized γ subunit prevents CF_1F_O from hydrolyzing ATP. During the day, light induces charge separation to generate an electrochemical potential across the membrane. Although the CF_1F_O begins to synthesize ATP molecules, the γ subunit is still oxidized while ΔΨ is small. The rate of ATP synthesis is not at its maximum. At sunrise, thioredoxin subsequently reduces the γ subunit, fully activating CF_1F_O. The molecular motor, consisting of the γ-ε central shaft and the c_{14}-ring, is free to rotate at full speed to maximize its ATP synthesis activity. Three ATP molecules per rotation of the c_{14} ring are produced. At sunset, the membrane becomes de-energized, leading to small ΔΨ and ΔpH, and the ATP hydrolysis starts to take place. To prevent ATP loss from excess ATP hydrolysis, the γ subunit is then oxidized again. This process of light regulation and redox modulation on the CF_1F_O will cycle daily.

20.4: CO_2 UPTAKE - CALVIN CYCLE AND C3 ORGANISMS

The source for the organization and some of the text derives from: Sindayigaya and Longhini. https://www.peoi.org/Courses/Courses...chem/biochem18 CC - https://creativecommons.org/licenses...sa/3.0/deed.en

20.4.1: INTRODUCTION

We focused on the light reactions of photosynthesis. Now let's turn our attention to the dark reactions which fix CO_2 from the air and reduce it with NADPH produced, along with O_2, in the light reactions, to produce carbohydrates. The dark reactions don't just occur in the dark. The term is simply used to differentiate them from the light-driven reactions using PSII and PSI. What is so interesting about plants is they produce fuel from CO_2 using photons as a source of energy (they are autotrophs) and also consume the fuels they make, using both anaerobic and aerobic respiration pathways. Their biosynthetic reactions take place mostly in the chloroplast, a type of plastid, which are subcellular organelles with specific functions such as photosynthesis or metabolite synthesis and storage. Plants also can not move to acquire fuel and nutrient molecules. They are subject to a large range of growing conditions (differential light qualities and quantities, temperatures, and rainfall levels). Also, plant cells have cell walls in addition to a cell membrane. A simple cartoon showing the major motifs of photosynthesis is shown in Figure 20.4.1.

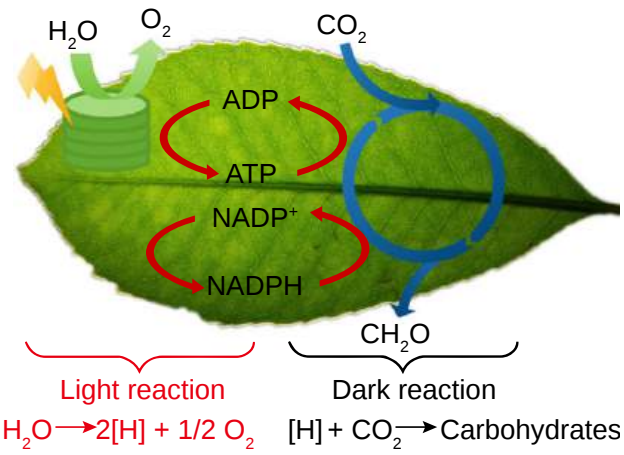


Figure 20.4.1: **a** Schematic depiction of the light and dark reactions in natural photosynthesis. Li, Y., Hui, D., Sun, Y. *et al.* Boosting thermo-photocatalytic CO_2 conversion activity by using photosynthesis-inspired electron-proton-transfer mediators. *Nat Commun* **12,** 123 (2021). https://doi.org/10.1038/s41467-020-20444-1 Creative Commons Attribution 4.0 International License.

In this section, we will discuss how CO_2 from the atmosphere is "fixed" or "captured" in the formation of the simplest sugars (3 carbon molecules like glyceraldehyde-3-phosphate) in a process called the **C3 or Calvin Cycle**, which is also called the Calvin–Benson–Bassham (CBB) cycle, or the reductive pentose phosphate cycle (RPP cycle). Plants that use the C3 cycle are logically called C3 plants There are two other major types of carbon capture pathways, the C4 and CAM pathways, which we discuss in the next section. All use a key enzyme, **ribulose 1,5-bisphosphate carboxylase (RuBisCo)**, to covalently fix CO_2 into small carbohydrates, 3-phosphoglycerate. **RuBisCo** is the most abundant protein in the biosphere. Recent estimates suggest that there are about 0.7 gigatons (Gt = 10^{12} tons) of it, with over 90% in the leaves (about 3% of their weight) of terrestrial plants. It captures about 120 Gt of atmospheric CO_2 each year. This enzyme has a second competing enzymatic activity. It is also an **oxygenase**, which adds to its complexity. That activity captures about 100 Gt of atmospheric O_2 each year. In this chapter section, we will give an overview of the C3 pathway and given the importance of RuBisCo, we will focus on it predominately.

Along RuBisCo, plants have pathways to take the fixed CO_2 to 3C sugars and then a unique pentose pathway which runs in a reductive fashion to ultimately produce the sugar-containing molecules in plants we are most familiar with, sucrose and the glucose polymer starch.

20.4.2: PLASTIDS

There are several types of these organelles. Photosynthesis occurs in chloroplast which has its own genome, like the mitochondria. Another common type is the amyloplasts, which lack pigmented molecules (i.e. they are colorless) and have no inner membrane. In plants, they are filled with starch. Chloroplasts and amyloplasts can interconvert. Chloroplasts are abundant in green leaves while amyloplasts are predominately found in locations like potato tubers where starch is stored. Light can drive the interconversion of plastids as shown in Figure 20.4.2.

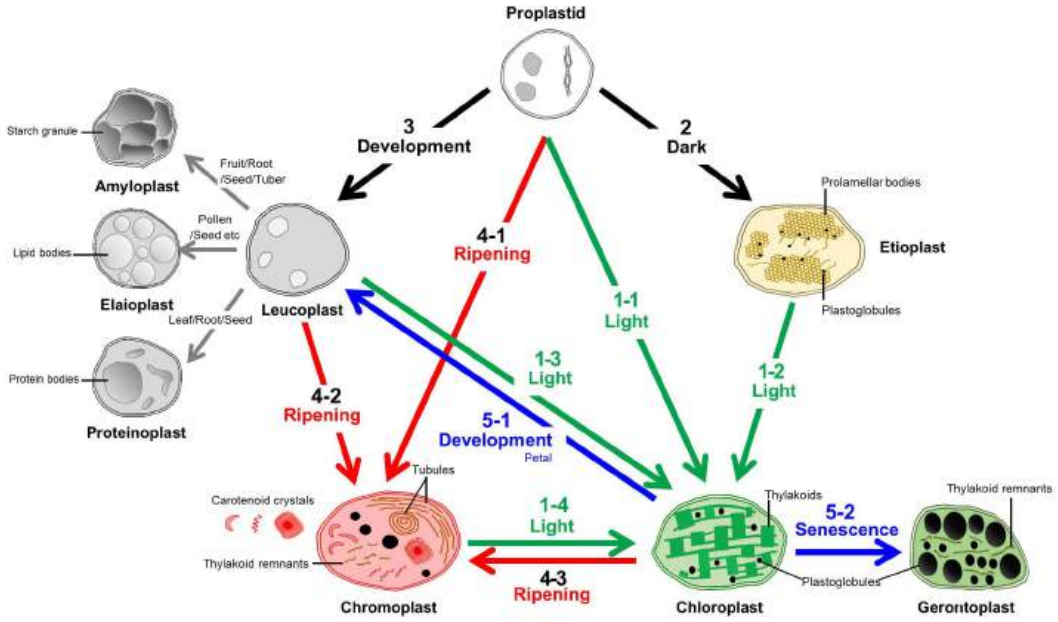


Figure 20.4.2: Transition pathways among various plastids. Choi et al. Frontiers in Plant Science. 12 (2021). https://www.frontiersin.org/article/...ls.2021.692024. DOI=10.3389/fpls.2021.692024 . **Creative Commons Attribution License (CC BY)**

The characteristics and plastid interconversion pathways of the plastids are shown by arrows. The transition to a chloroplast is called “Greening” and is identified with the number “1”. This is mainly triggered by light signals from proplastids, etioplasts, leucoplasts, and chromoplasts. Etioplasts can develop from proplastids in dark conditions and this is identified by the number “2”. The number “3” indicates leucoplast development that is triggered by diverse development processes to generate starch, lipid, and protein-enriched sub-types called amyloplasts, elaioplasts, and proteinoplasts, respectively. Mainly during the ripening stage, diverse types of carotenoid crystals were generated within the plastids called chromoplasts from the proplastids, leucoplasts, and chloroplasts and this is identified with the number “4”. Together with etioplast and leucoplast development (2,3), chromoplast development (4) was identified as a “Non-greening” plastid transition. The loss of green color from the chloroplasts is called “De-greening” and is identified with the number “5”, and these chloroplasts are then transited into leucoplast or gerontoplast by developmental regulation or during senescence, respectively.

20.4.3: CO_2 CAPTURE AND THE C3 CYCLE

There are in the synthesis of the simplest carbohydrates (3 carbon polyhydroxy- aldehydes and ketones:

1. Carbon capture or fixation phase. We prefer the term carbon capture as this term is now used to describe how the world is seeking new ways (other than planting billions of trees) to "capture" excess CO_2 emitted through the use of fossil fuels. In a reaction catalyzed by RuBisCo, atmospheric CO_2 ultimately react with a 5-carbon acceptor molecule, ribulose 1,5-bisphosphate (Ru1,5-BP, 6 carbons in total), to form two molecules of 3-phosphoglycerate (3PG). (2, 3C molecules).
2. Reduction phase: 3-phosphoglycerate is reduced to glyceraldehyde-3-phosphate (G3P). Three CO_2s are captured on reaction with 3 Ru1,5-BP to form 6 glyceraldehyde-3-phosphates (G3P). These can readily interconvert to the keto form, dihydroxyacetone phosphate (DHAP).
3. Regeneration phase: Five of the six G3Ps (15 Cs) react to form 3 three molecules of ribulose 1,5-bisphosphate (15 C2) to allow the catalytic C3 cycle to continue. The other G3P moves into the stroma, in the form of DHAP where it can be used in gluconeogenesis (reductive biosynthesis) of glucose. This can be converted to polymer starch and also the disaccharide sucrose (which we will discuss in a future session).

An overview of the Calvin or C3 cycle is shown below in Figure 20.4.3.

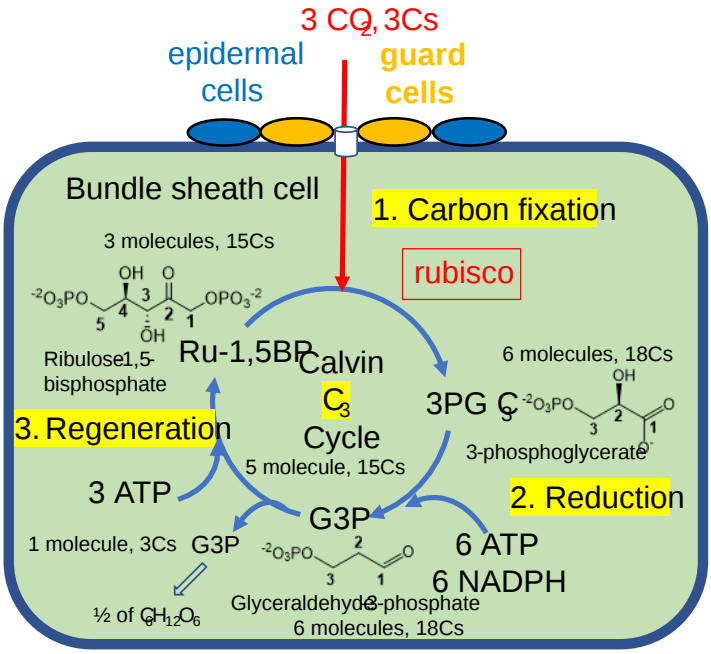


Figure 20.4.3: C3 or Calvin Cycle. The three stages of the cycle in C3 plants and organisms are shown. Three CO2 are captured for the net synthesis of one molecule of glyceraldehyde 3-phosphate.

The stoichiometry can be confusing until you count the actual number of carbon atoms and realize that the cycle has to run 3 times to enable 3 carbon atoms from 3 CO_2 molecules to produce one net glyceraldehyde-3-phosphate (G3P). The G3P leaves the C3 cycle at the low left for glucose synthesis. The conversion of the 5 G3Ps that reform Ru1,5-BP requires ATP as shown below:

$$5\,\text{glyceraldehyde}-3\,\text{P}+3\,\text{ATP}\rightarrow 3\,\text{ribulose}-1\,,5-2\,\text{P}+3\,\text{ADP}+2\,\text{Pi}$$

with Pi indicating inorganic phosphate. Hence the net equation for 3 turns of the cycle, sufficient to produce 1, G3P is:

$$3\,\text{CO}_2+6\,\text{NADPH}+6\,\text{H}^+ +9\,\text{ATP}+5\,\text{H}_2\text{O}\rightarrow \text{glyceraldehyde}-3-\text{phosphate}(\text{G}_3\text{P})+6\,\text{NADP}^+ +9\,\text{ADP}+8\,\text{Pi}$$

Even though glucose is not a product of the Calvin cycle, some texts use the following equation to show the stoichiometry to run the C3 cycle enough times (6) to fix 6 CO_2 molecules, enough to make 1 glucose from a simple carbon atom counting perspective.

$$6\,\text{CO}_2+12\,\text{NADPH}+12\,\text{H}^+ +18\,\text{ATP}+10\,\text{H}_2\text{O}\rightarrow 2\,\text{glyceraldehyde}-3-\text{phosphate}(\text{G}_3\text{P})+12\,\text{NADP}^+ +18\,\text{ADP}+16\,\text{Pi}$$

Remember that NADPH and ATP are produced in the light reactions in about the same ratio as they are used in the C3 cycle (2NADPH/3ATPs). The net 8 P_is made as products will react with 8 ADP to regenerate 8 ATP in the light reaction. The 9th Pi is incorporated in a triose-phosphate in the light reaction, so one Pi must be imported from the cytoplasm by an inner membrane triose-phosphate/phosphate translocator, which we will discuss below. In the dark, when ATP and NADPH are not produced, CO_2 capture also is inhibited.

A more detailed diagram showing the detailed reactions to regenerate ribulose 1,5-bisphosphate (Ru1,5BP) is shown in Figure 20.4.4.

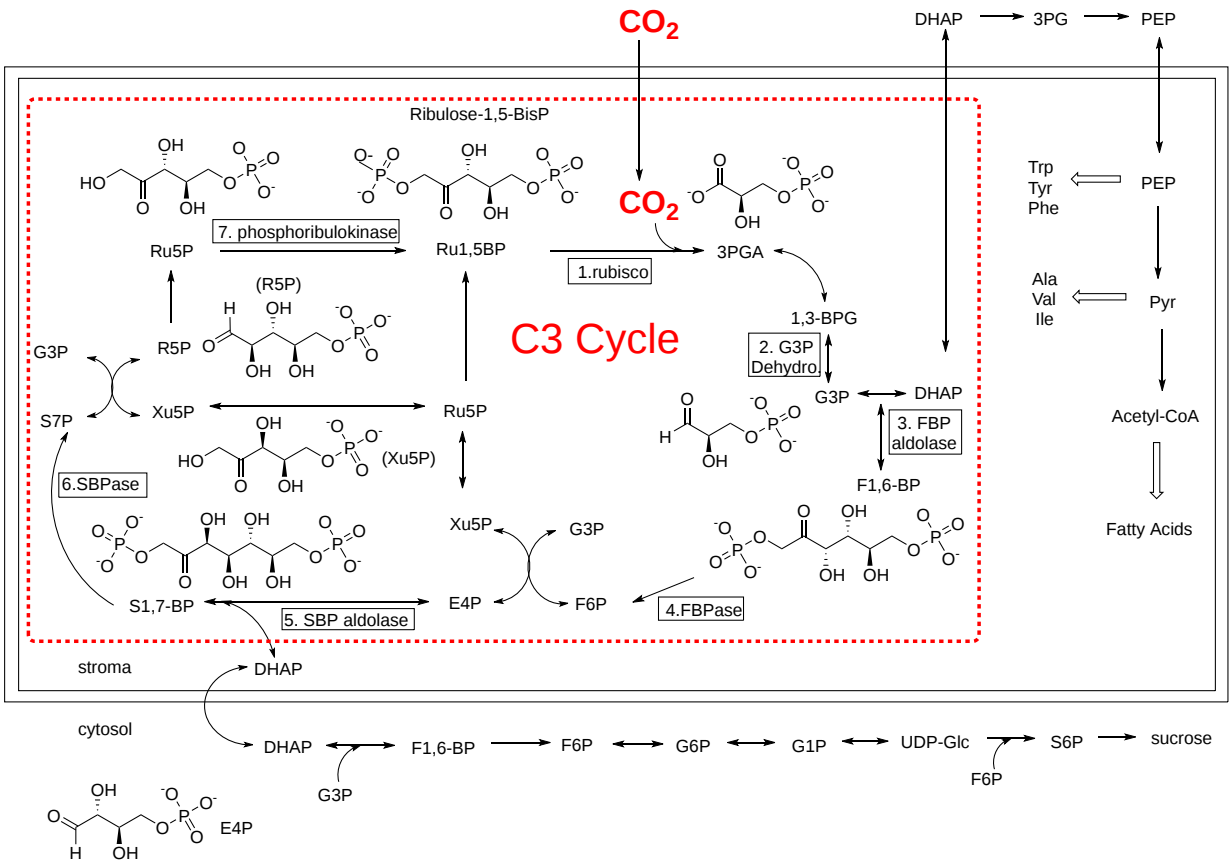


Figure 20.4.4: Detailed reaction diagram for the Calvin (C3) cycle (after Hashunuma et al. Journal of Experimental Botany, 61 (2010). doi:10.1093/jxb/erp374. Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/bync/2.5)

Abbreviated reactions for the synthesis of sucrose, glucogenic amino acids, and fatty acids are also shown.

You should note the reactions for the conversion of the 6C sugar molecule fructose-6-P (F6P) (glycolytic and gluconeogenic intermediate) to the 5C molecule Ru5P and Ru1-5BP, are analogous to the reactions of the nonoxidative part of the pentose phosphate pathway (PPP) pathway which generates 5C sugars for the synthesis of nucleotides, nucleic acids, and some amino acids. Hence we won't discuss them further.

20.4.4: CARBON CAPTURE OF CO2 INTO 3-PHOSPHOGLYCERATE - RUBISCO

This key enzyme requires a Mg^{2+} ion and proceeds through a carbamoylated lysine side chain. The Mg^{2+} ion orients key side chains. The resulting 6C molecule cleaves into two 2 molecules of 3PG.

The RuBisCo family of enzymes can adopt different quaternary structures. A homodimer of two large subunits is the minimum catalytic structure. Often there are in addition two small subunits. The most common form is a 16mer which is found in cyanobacteria, red and brown algae, and all higher plants.

A possible mechanism of RuBisCo from *Synechococcus elongatus*, a unicellular cyanobacterium, is shown in Figure 20.4.5.

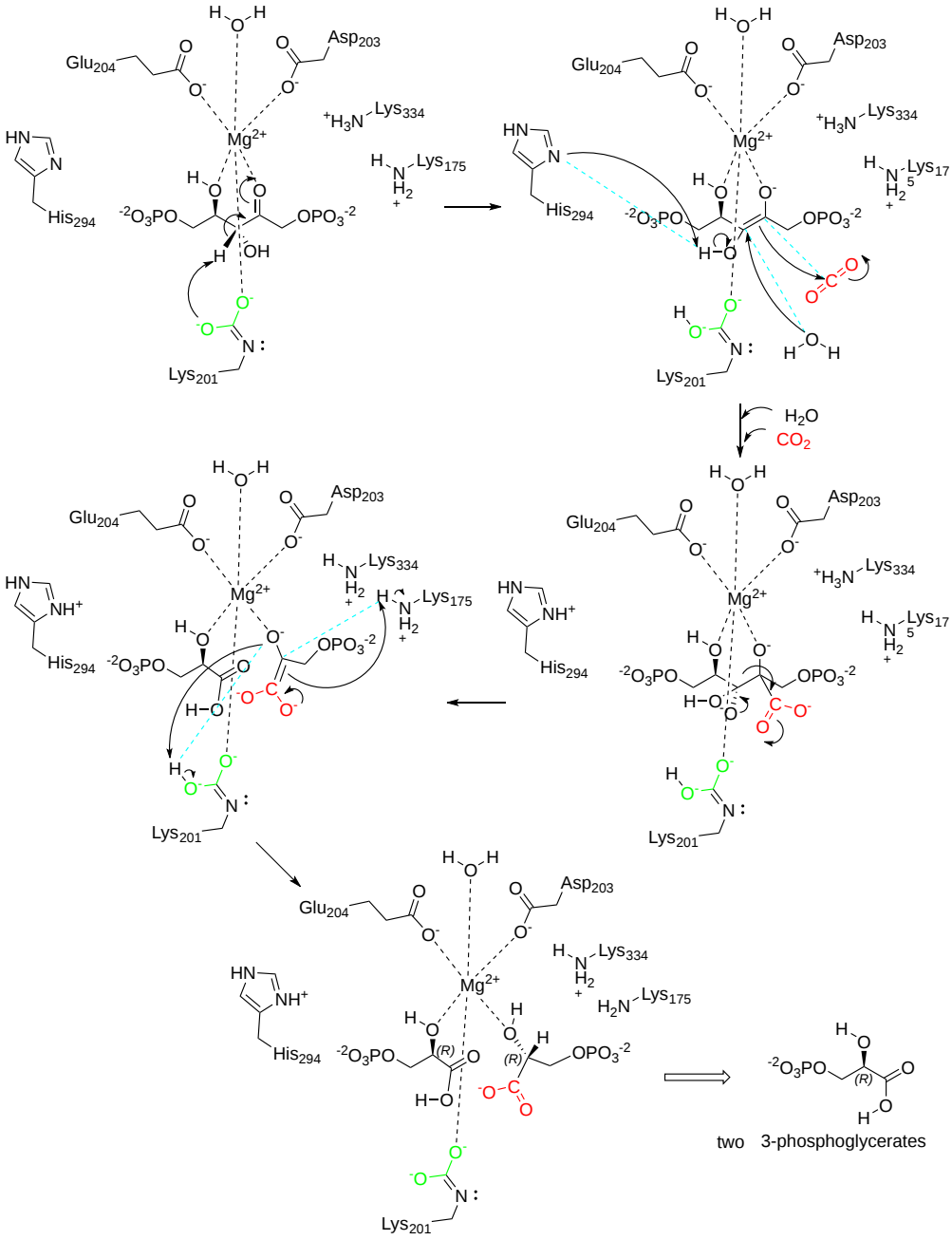
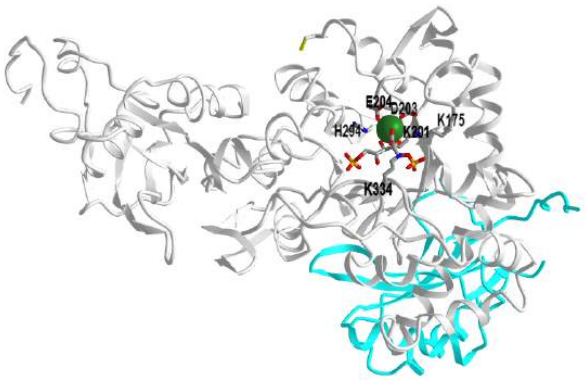


Figure 20.4.5: Mechanism of ribulose-bisphosphate carboxylase (type I) (after https://www.ebi.ac.uk/thornton-srv/m-csa/entry/907/. Creative Commons Attribution 4.0 International (CC BY 4.0) License)

The carbamoylated lysine side chain is shown in green. Ribulose 1,5- bisphosphate is converted to an enediolate which engages in a nucleophilic attack on the CO_2 to form a 6C sugar. Hydroxylation at C-3 of this sugar is followed by aldol cleavage. Ultimately two 3PGs are produced, one of which contains the carbon atom from CO_2 (red).

Figure 20.4.6 shows an interactive iCn3D model of a single heavy and light chain of ribulose 1,5-bisphosphate carboxylase/oxygenase (RuBisCo) from Synechococcus PCC6301 (1RBL). (long load time)



NCBI iCn3D Figure 20.4.6: Ribulose 1,5-bisphosphate carboxylase/oxygenase (RuBisCo) from Synechococcus PCC6301 (1RBL). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...BkwkrBmYVwRJ6A

The light chain is shown in cyan and key residues in the heavy chain are shown in CPK-colored sticks and labeled. Bound to the heavy chain is a substrate analog/inhibitor, 2-carboxyarabinitol-1,5-diphosphate. It is produced in plants and in the dark, it inhibits the enzyme. With increasing lights, its concentration decreases.

Recent Updates (08/14/23) - Rubisco Reacts with both CO_2 and O_2

Rubisco is a slow enzyme with a k_{cat} of around 2-10 CO_2/sec. In addition, it can bind another substrate, O_2, and engage in a competing reaction of photorespiration (oxidation) of ribulose 1,5-bisphosphate to form one molecule of 3-phosphoglycerate (3PG) and one molecule of 2-phosphoglycolate (2PG), as shown (incompletely) in Figure 20.4.7 below.

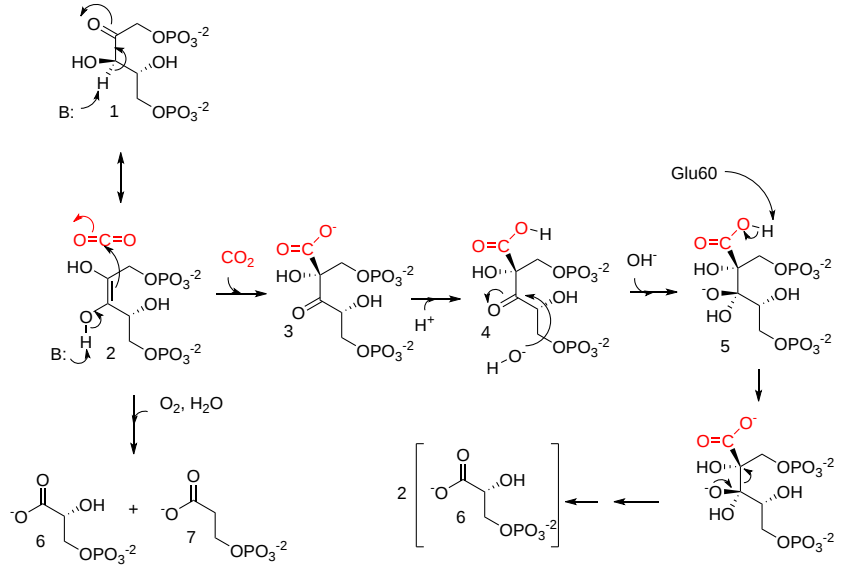


Figure 20.4.7: RuBP conversion by Rubisco through the carboxylase (a) and the oxygenase (b) reactions. Tommasi, I.C. The Mechanism of Rubisco Catalyzed Carboxylation Reaction: Chemical Aspects Involving Acid-Base Chemistry and Functioning of the Molecular Machine. *Catalysts* **2021**, *11*, 813. https://doi.org/10.3390/catal11070813. CC BY) license (https://creativecommons.org/licenses/by/4.0/

Following RuBP (1) enolization, the 2,3-enol(ate) intermediate (2) may react with CO2(a) or O2(b) co-substrates. The carboxylase reaction produces the 2-carboxy-3-keto-arabinitol 1,5-bisphosphate intermediate (3) undergoing protonation to the 2-carboxylic acid before hydration. The C2-C3-scission reaction in C3-gemdiolate (5) is described as occurring in a concerted mechanism upon P1 protonation

producing two molecules of 3-phospho-D-glycerate (3PGA, 6). The oxygenase reaction produces 3-phospho-D-glycerate (3PGA,6) and 2-phosphoglycolate (2PG,7)

How does the enzyme differentiate the two nonpolar substrates, CO_2 and O_2? Given the symmetric arrangement of δ^+ and δ^- charges in CO_2, it has a net 0 dipole, as does O_2, so it would not align/orient in a field generated by two poles (+ and -). However, CO_2, but not O_2, would align in a field generated by four charged poles so it has a **quadrupole moment**, as shown in Figure 20.4.8 below.

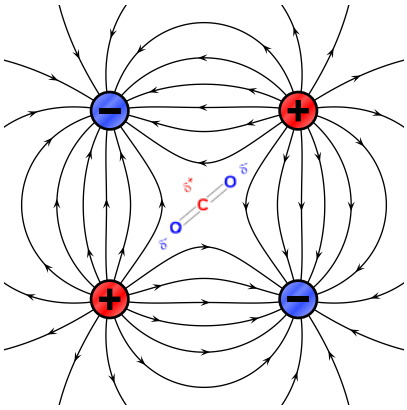


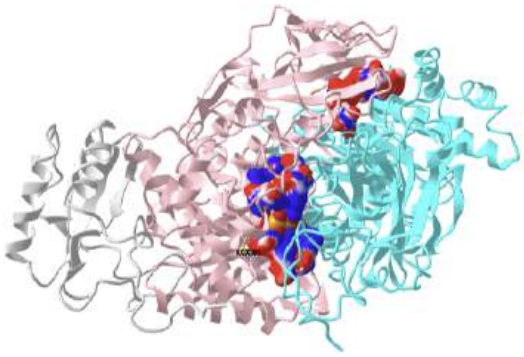
Figure 20.4.8: CO_2 aligning in a quadrupole field. (field lines from https://commons.wikimedia.org/wiki/F...quadrupole.svg)

The dipole unit is the debye and the quadrupole unit can be expressed in debye·Angstrom. Table 20.4.1 below shows some values for dipole and quadrupole moments for simple gases. CO_2 has the highest quadrupole moment of all these simple gases.

Molecule	Dipole moment (D)	Quadrupole moment (D Å)
CO_2	0.000	4.30
CH_4	0.000	0.02
H_2	0.000	0.66
O_2	0.000	0.39
CO	0.112	2.5
N_2	0.000	1.52

Table 20.4.1: Dipole and Quadrupole moments for some simple gases. Castro-Muñoz, R., Fíla, V., 2018. Progress on Incorporating Zeolites in Matrimid®5218 Mixed Matrix Membranes towards Gas Separation. Membranes 8, 30.. https://doi.org/10.3390/membranes8020030

The active site has a high electrostatic field gradient in the dimeric form of the enzyme. Figure 20.4.9 shows an interactive iCn3D model showing the electrostatic potential surface in the active site between two heavy chains of spinach ribulose 1,5-bisphosphate carboxylase/oxygenase (RuBisCo) (8RUC). It shows that is complex.



NCBI iCn3D Figure 20.4.9: Electrostatic potential surface in the active site between two heavy chains of spinach ribulose 1,5-bisphosphate carboxylase/oxygenase (RuBisCo) (8RUC). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...2JG3x4csVnEq67

Two heavy chains are shown in light pink and cyan, and one light chain is shown in gray. The active site regions are shown as an electrostatic surface potential with blue positive and red negative. The model has an activated substrate analog, 2-carboxyarabinitol bisphosphate, shown spacefill (hard to see given the electrostatic surface potential). The residue labeled 201KCX is the carbamate of Lys201.

Presumably, the electrostatic field in the active site facilitates through subtle interaction of the quadrupole CO_2 compared to O_2.

The mechanisms that differentiate the binding of CO_2 and O_2 must also overcome the high intracellular concentration of O_2 (around 250 μM) compared to CO_2 (7–8 μM in C3 plants and 80 μM in C4 plants). The enhanced affinity for CO_2 (about 30-fold) compared to O_2 helps overcomes these concentration barriers. No classic "binding pocket" exists for CO_2 and O_2 so diffusion and binding is likely guided by the electrostatic potential gradients along the diffusion surface. Figure 20.4.10 below the electrostatic potential molecular surface of O_2 and CO_2 (top), calculated from electron density measurement using quantum mechanics, and the electrostatic surface of their binding pockets

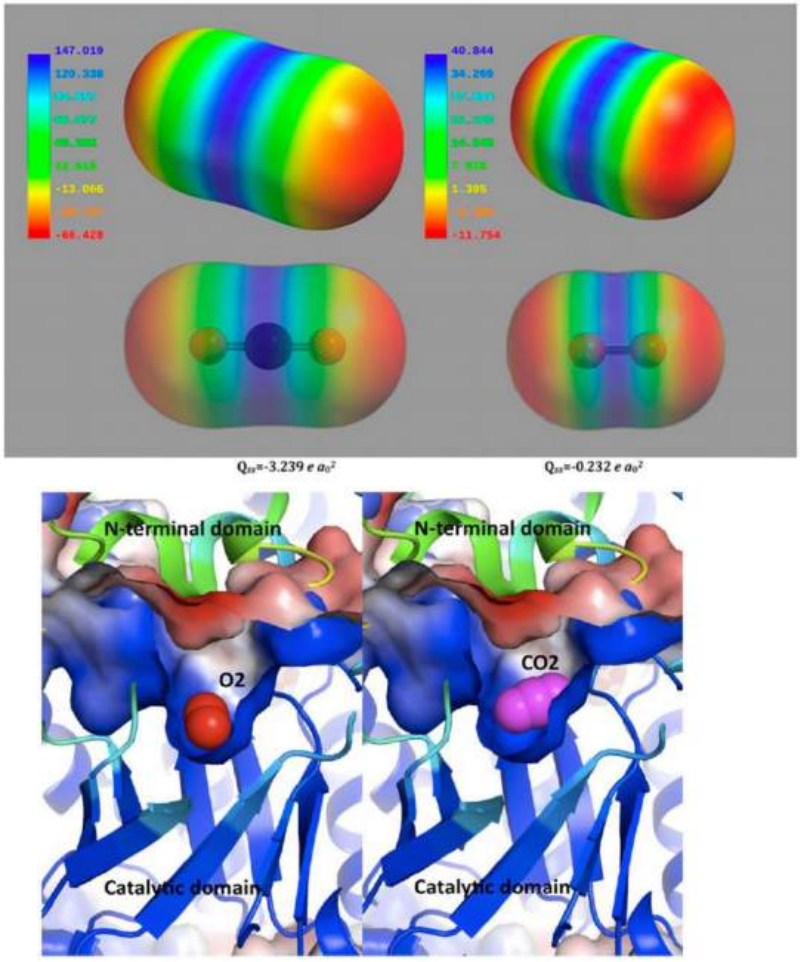


Figure 20.4.10: Electrostatic potential molecular surface of O_2 and CO_2 (top) and the electrostatic surface of their binding pockets. Tommasi, I.C. et al., ibid.

(**top**) Computed electrostatic potential molecular surfaces of CO_2 (left) and O_2 (right). The color scheme follows commonly accepted conventions: blue, positive; red, negative. The value of the Qzz component of the quadrupole moment, as calculated by Stec, is −3.239 e a_0^2 for CO_2 and −0.232 e a_0^2 for O_2. Note that the ratio of these values is about 14, about equal to the earlier quadrupole moments discussed above with units of debeye·Å)

(**bottom**) a ribbon representation of the catalytic domain with bound gaseous ligands and surfaces colored by the electrostatic potential. O_2 (in red) and CO_2 (in purple) lie in a positively charged cavity (blue) of the TIM barrel. (Figure from ref. [15], used by permission of PNAS (copyright © 2012)).

Both CO_2 and O_2 are situated in a tiny "cavity" that is blue (positive potential, C-terminal domain) and just above it red (negative potential, N-terminal domain). The quadrupole moment of CO_2 is 10-15x that of O_2 which helps explain its higher affinity in the localized electrostatic potential gradients around the gas molecules.

Molecular dynamic (MD) simulations show an interaction preference for CO_2. There are many subunit-subunit interfaces and all appear in MD to preferentially interact with CO_2, which probably moves from the solvent through large:small subunit interface to the active site. The CO_2 does not localize long at any residues but seems to occupy areas instead. Since CO_2 has no dipole, it locates more closely to small hydrophobic side chain (Ala, Val, Leu, Ile) and the main chain. CO_2 has more interaction in every active site as well as the large:large subunit interface (whose electrostatic potential in the active site is shown above) and in the large:small subunit interface.

In efforts to quantitate the preference of CO_2 over O_2 using MD, investigators found that over many different species of rubisco, the relative distribution of CO_2 and O_2 to the small and large was on average 1.8 for CO_2 and 1.4 for O_2 with the number of oxygen bound to either subunit lower. This is true even though CO_2 has a lower solvation energy than O_2 in water, so additional energy must be spent to differentially desolvate CO_2. The hydrophobic interactions likely promote the movement of CO_2 to the active site where electrostatic-based potentials likely favor CO_2 binding. These results suggest that the small subunit acts like a "miniresevoir" for O_2 which then diffuse to the large subunit region of the active site. From a simple thermodynamic perspective, CO_2 would be favored to move along the surface and through spaces in the protein guided by transient interactions that be water. The active site in rubisco is not in a deep pocket but rather in shallow groves near the surface. Although the enzyme is slow (2-10 CO_2/s), it's not much slower than the average enzyme. The median turnover number k_{cat} (under saturating conditions) of enzymes is about 10 s^{-1} with most following between 1-100. Its concentration is very high in chloroplasts, which helps increase the fixing of CO_2.

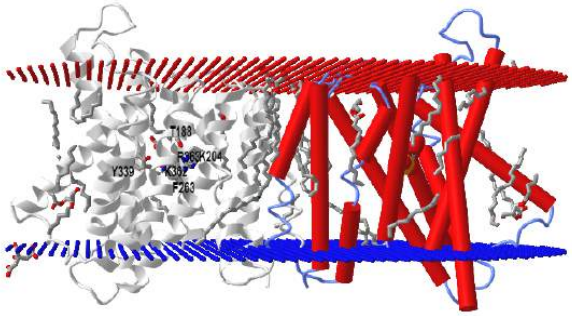
20.4.5: THE NEXT STEP: 3-PHOSPHOGLYCERATE TO GLYCERALDEHYDE 3-PHOSPHATE AND DIHYDROXYACETONE PHOSPHATE

The 3PG produced by RuBisCo is converted to the triose glyceraldehyde-3-P (G3P) (which can readily isomerize to dihydroxyacetone phosphate) using typical glycolytic enzymes run in reverse except that NADPH is used as a reductant instead of NADH. In addition, the stromal and cytosolic enzymes derive from different genes. The remaining G3P not used to resynthesize ribulose 1,5BP can be used for the synthesis of starch, sucrose, etc, as illustrated in Figure 4 above.

20.4.6: EXCHANGE OF TRIOSES AND PHOSPHATE ACROSS THE INNER MEMBRANE

The inner chloroplast membrane has a **triose-phosphate/phosphate translocator (TPT)**, an antiporter that brings into the stroma Pi in exchange for a triose phosphate, either dihydroxyacetone phosphate or 3-phosphoglycerate. The importance of this was discussed above. The exported triose can be used for the synthesis of sucrose, which can be transported around the plant as a source of carbon. Trioses within the chloroplast can also be converted to glucose and onto glycogen as the organelle becomes an amyloplast. If the translocator is inhibited, Pi would decrease in the chloroplast, which would decrease ATP and also starch synthesis.

The structure of TPT has been determined with the bound ligands, 3-phosphoglycerate and inorganic phosphate, in an occluded conformation from *Galdieria sulphuraria*, an extremophilic unicellular species of red algae. Figure 20.4.11 shows an interactive iCn3D model of a triose-phosphate/phosphate translocator from the red algae (5Y78).



NCBI iCn3D Figure 20.4.11: Triose-phosphate/phosphate translocator from red algae (5Y78). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...mVNeFkXibxXr39

The model shows two monomers, one with red cylindrical alpha helices and spacefill 3-phosphoglycerate. The other subunit is shown in gray with the 3PG in colored sticks and conserved residues (T188, K204, F263, Y339, K362, R363) that make interacts with both Pi and 3PG. There would presumably be an outward- and inward-open conformation that is triggered on ligand binding.

20.4.7: ACTIVITY REGULATION BY LIGHT

Given their role in photosynthesis, you would expect even the dark reaction enzymes would be regulated by light. Indeed, four C3 cycle enzymes are. They are ribulose 5-phosphate kinase, fructose 1,6-bisphosphatase, sedoheptulose 1,7-bisphosphatase, and glyceraldehyde 3-phosphate dehydrogenase. The regulation is affected by photon-induced disulfide bond formation between two cysteine side chains in the enzymes. When oxidized (disulfide bond form), the enzymes are inactive. Under light conditions, PSII, cyto b_6f, and PSI work in electron transport to move electrons from H_2O to ferredoxin and onto a small soluble protein thioredoxin which has a disulfide. The enzyme catalyzing this last step is **ferredoxin-thioredoxin reductase.** On reduction, the disulfide in thioredoxin is cleaved, and the now free sulfhydryls in thioredoxin are used to cleave the disulfide in the 4 enzymes mention above, in a similar fashion to how β-mercaptoethanol in excess can cleave disulfides in proteins. This leads to conformational changes in the four enzymes which activate them. In the absence of light, the process reverses, and the enzymes are inhibited. For fuel at night, plants mobilize starch for energy.

A simple mechanism to show how thioredoxin catalyzes disulfide bond reduction in target proteins is shown in Figure 20.4.12.

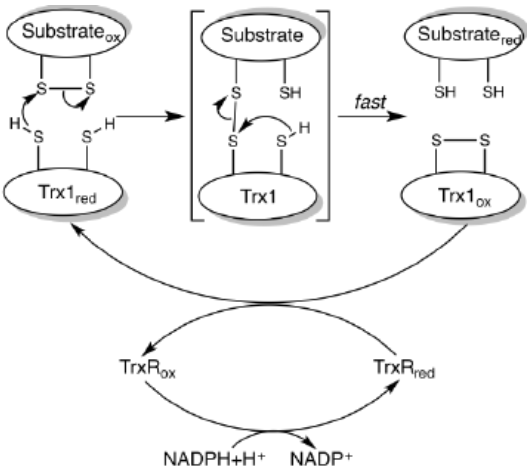
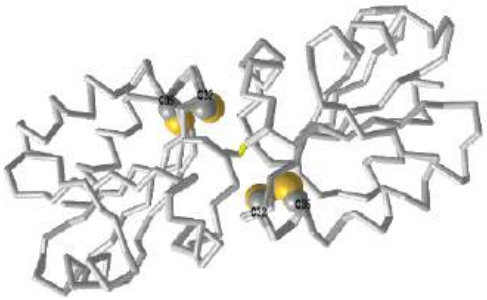


Figure 20.4.12 Mechanism of thioredoxin reduction of target proteins. https://en.Wikipedia.org/wiki/Thiore...le:FigMech.png

The first enzyme in the oxidative branch of the pentose pathway, glucose 6-phosphate dehydrogenase, uses $NADP^+$ as an oxidizing agent, producing NADPH. In the light, there is lots of NADPH produced from the light reactions of photosynthesis so it makes biological sense that under these conditions, glucose 6-phosphate dehydrogenase activity is inhibited. It is so, also by the cleavage of a critical disulfide bond, but in this case, it results in enzyme inactivation.

We saw the role of thioredoxin in the previous chapter section when we discussed the regulation of photosynthesis as well as the ATP synthase of the chloroplast.

Figure 20.4.13 shows an interactive iCn3D model showing a comparison of the structures of oxidized (1ERU) and reduced (1ERT) human thioredoxin.



NCBI iCn3D Figure 20.4.13: Comparison of the structures of oxidized (1ERU) and reduced (1ERT) human thioredoxin. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...XPVsErV2JWxdA6.

The two subunits of thioredoxin, linked by a disulfide are shown in gray. Press the "a" key to toggle between the oxidized form, with Cys32-Cys35 disulfide shown as a yellow stick, and the reduced form with the reduced and separated Cys 32 and Cys 35 shown in colored spheres. Not that the hydrogen covalently attached to the free cysteine side chain does not show in a crystal PDB structure.

20.4.8: SUMMARY

Photosynthesis in vascular plants takes place in chloroplasts. In the CO2-assimilating reactions (the Calvin cycle), ATP and NADPH are used to reduce CO2 to triose phosphates. These reactions occur in three stages: the fixation reaction itself, catalyzed by Rubisco; reduction of the resulting 3-phosphoglycerate to glyceraldehyde 3-phosphate; and regeneration of ribulose 1,5-bisphosphate from triose phosphates. Rubisco condenses CO2 with ribulose 1,5-bisphosphate, forming an unstable hexose bisphosphate that splits into two molecules of 3-phosphoglycerate. Rubisco is activated by covalent modification (carbamoylation of Lys201) catalyzed by Rubisco activase and is inhibited by a natural transition-state analog, whose concentration rises in the dark and falls during daylight. Stromal isozymes of the glycolytic enzymes catalyze the reduction of 3-phosphoglycerate to glyceraldehyde 3-phosphate; each molecule reduced requires one ATP and one NADPH. The cost of fixing three CO2 into one triose phosphate is nine ATP and six NADPH, which are provided by the light-dependent reactions of photosynthesis. An antiporter in the inner chloroplast membrane exchanges Pi in the cytosol for 3-phosphoglycerate or dihydroxyacetone phosphate produced by CO2 assimilation in the stroma. Oxidation of dihydroxyacetone phosphate in the cytosol generates ATP and NADH, thus moving ATP and reducing equivalents from the chloroplast to the cytosol. Four enzymes of the Calvin cycle are activated indirectly by light and are inactive in the dark so that hexose synthesis does not compete with glycolysis—which is required to provide energy in the dark.

20.4.9: PHOTORESPIRATION - RUBISCO/OXYGENASE AND THE GLYCOLATE CYCLE

As autotrophs, plants make their fuels. They use that fuel to make ATP to power endergonic reactions like protein synthesis, cell division, etc. As eukaryotic cells, they have mitochondria and can use both aerobic and anaerobic respiration to produce ATP. In the dark, when photons are not present, they carry out mitochondrial aerobic respiration as they break down carbohydrates to CO_2 and water, the reverse of photosynthesis.

They also use O_2 in another process that is driven by light. The same enzyme that captures carbon, RuBisCo, has oxygenase activity. RuBisCo uses O_2 in a process called **photorespiration**, which produces CO_2 in a competing reaction. Same enzyme, different substrates! The final products of the reaction with CO_2 using RuBisCo are two 3C molecules, 3-phosphoglycerate (3PG). Using O_2 as a substrate produces 1 molecule of the 3C 3PG and 1 molecule of a 2C analog, **2-phosphoglycolate** (not 2-phosphoglycerate). 2-phosphoglycolate is also named **carboxymethylphosphate**. About one out of every four turnovers of the enzyme produced this metabolic dead product. Given this non-trivial side reaction, the enzyme should be called **ribulose 1,5-bisphosphate carboxylase/oxygenase**.

You may ask why would such a critical enzyme evolved to a form which is quite inefficient. One explanation is that the enzyme "finished" its evolution before the great oxygenation event when dioxygen rose to the levels we see now (20%). Before that, little oxygen was available to compete with the trace gas CO_2, which is around 0.04% of the atmosphere. (Even though CO_2 is considered a trace gas, its present concentration is around 420 parts per million (ppm), levels which are warming our planet and which have not been seen for 3 million years, when Arctic forests and camels were present).

Compare the K_M values (9 μM for CO_2 and 350 μM for O_2 or 39x higher for O_2) for the enzyme and the equilibrium concentrations of the gases in aqueous solution (11 μM for CO_2 and 250 μM for O_2 or 23x higher for O_2). The higher K_M for O_2 is nearly offset by O^2's greater solubility so modern conditions lead to a significant waste of the CO_2 capture efficiency of RuBisCo/Oxy. At higher temperatures in a warming world, the equilibrium ratio of solution concentrations of O_2/CO_2 increases as does the affinity (based crudely on K_M values) of CO_2. Both of these exacerbate the wasteful oxygenase activity effect. Finally, as CO_2 is captured by the enzyme, the ratio of the local concentrations of O_2/CO_2 also goes up. All of these make the efficiency of RuBisCo worse.

Figure 20.4.14 shows a mechanism for the reaction of both CO_2 and O_2 with RuBisCo/Oxygenase.

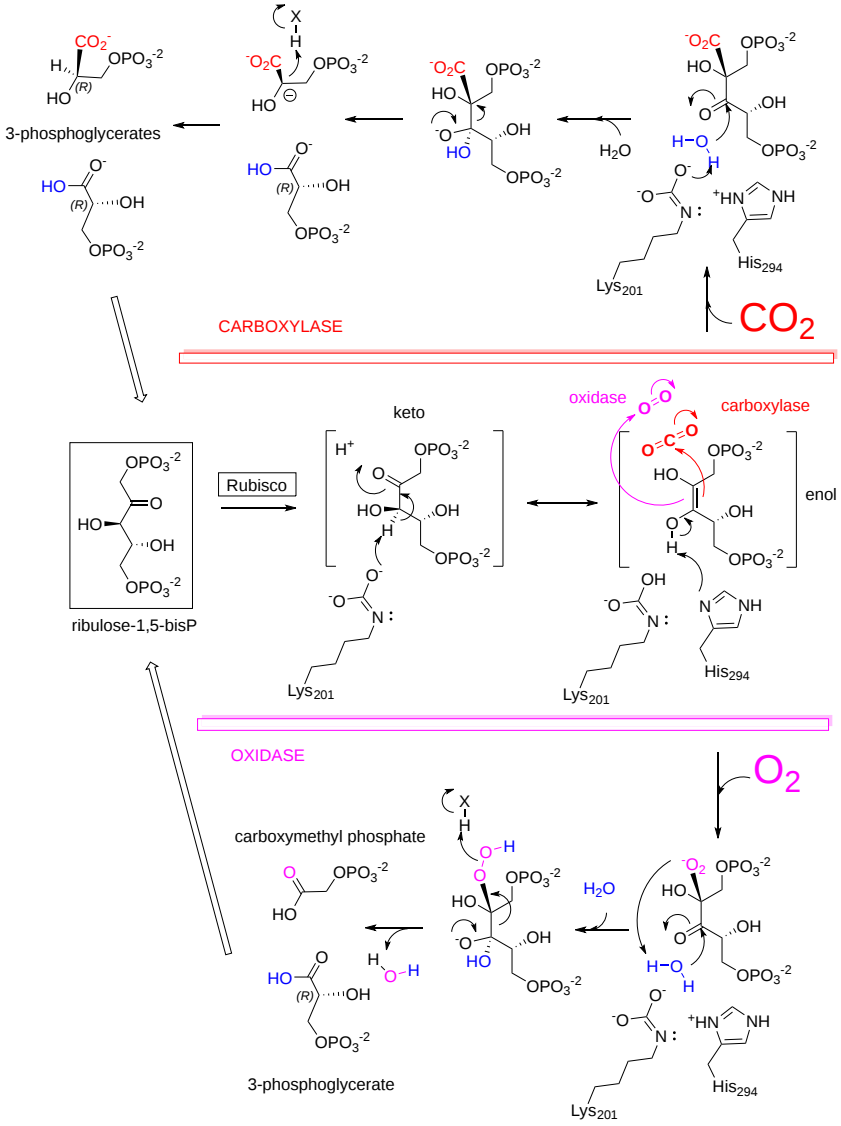


Figure 20.4.14: Mechanism for the reaction of both CO_2 and O_2 with RuBisCo/Oxygenase. (after Kannappan et al. J. Phys. Chem. B 2019, 123, 2833−2843)

Note that in contrast to most oxygenases, no cofactor is required for the RuBisCo/Oxygenase.

20.4.10: THE GLYCOLATE PATHWAY

The 2-phosphoglycolate (carboxylmethyl phosphate) "waste" product of the oxygenase activity of RuBisCo/Oxygenase is recycled through a complex pathway that is called "photorespiration". It occurs in three different organelles, the chloroplast, the peroxisome, and the mitochondria. Part of the generalized pathway is shown in Figure 20.4.15.

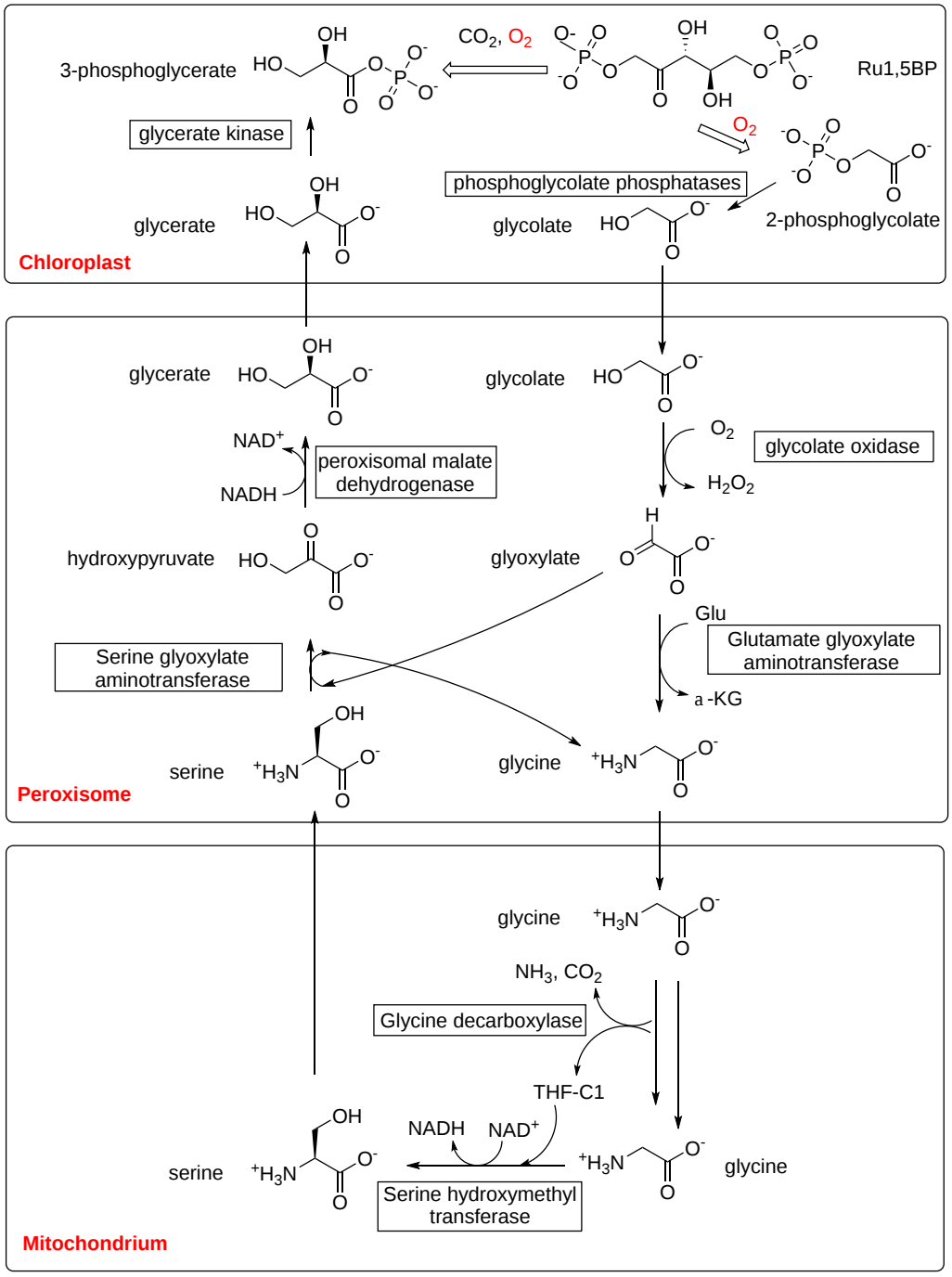


Figure 20.4.15: The glycolate photorespiration salvage pathway (adapted from Hu et al. *Plant Cell*. 2012;24.doi:10.1105/tpc.112.096586)

Multiplying the stoichiometry represented in the figure by 2 shows that 2 molecules of 2-phosphoglycolate produce 2 molecules of glycine. These get converted to two molecules of serine. We will see the mechanisms for some of these reaction in the chapter on amino acid metabolism. The net reaction is:

$$2\,\text{Glycine} + \text{NAD}^{++}\text{H}_2\text{O} \rightarrow \text{serine} + \text{CO}_2 + \text{NH}_3 + \text{NADH}$$

The serine is eventually converted to 3-phosphoglycerate, which can be used again in the C3 cycle. Note that CO_2 is produced in the glycolate pathways that started with the use of O_2 as a substrate by RuBisCo/Oxygenase. Hence the whole system uses O_2 and produces one CO_2 so the combined reactions are usually called **photorespiration**. It's not an ideal term since it is wasteful, compared to mitochondrial respiration. Some prefer to call the combined pathway of Rubisco oxygenase and the glycolate pathways the **C2 cycle**.

20.5: CO2 UPTAKE - C4 AND CAM PATHWAYS

The source for the organization and some of the text derives from Sindayigaya and Longhini. https://www.peoi.org/Courses/Courses...chem/biochem18 CC - https://creativecommons.org/licenses...sa/3.0/deed.en

20.5.1: THE C4 PATHWAY

Photorespiration, caused by the oxygenase activity of RuBisCo/Oxygenase and the ensuring glycolate salvage (C2) pathway, significantly diminishes the efficiency of photosynthesis in C3 plants. Rest assured, some plants have found a way around the problem by adding a few steps before RuBisCo. The altered pathway is called the C4 pathway and plants that use it are called C4 plants. The initial steps involve a temporary capture of CO_2 in the form of HCO_3^-, into a 4C, not 3C sugar. Plants that use the C4 pathway include maize, sorghum, sugar cane, and many tropic plants. An overview of the pathways emphasizing the steps that precede RuBisCo is shown in Figure 20.5.1.

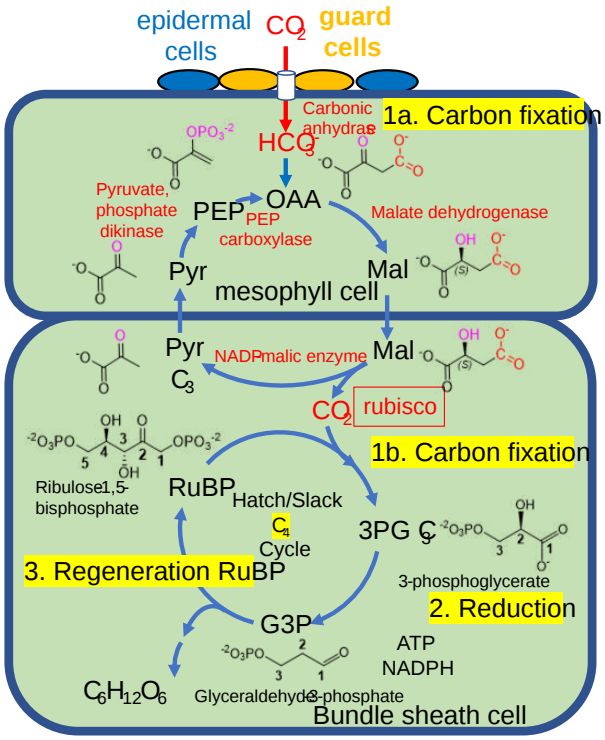


Figure 20.5.1: Carbon assimilation in C4 plant

Experimental evidence shows that radiolabeled $^{14}CO_2$ is captured into the 4C molecule, oxaloacetate (OAA), a citric acid cycle and gluconeogenic intermediate, through the enzyme phosphoenolpyruvate carboxylase, which uses HCO_3^- as a substrate. The reaction takes place in mesophyll cells. OAA can be reduced by NADPH using malate dehydrogenase, or converted to aspartate through a transamination (not shown in the figure). Malate moves in the bundle sheath cell where it is decarboxylated by the NADP-malic enzyme to pyruvate. Pyruvate can move back into the mesophyll cell and be converted to phosphoenolpyruvate (PEP) and then back to OAA by the enzymes **pyruvate phosphate dikinase** and **PEP carboxylase**. CO_2 from the decarboxylation of malate is delivered as a substrate to RuBisCo.

Pyruvate dikinase is used in bacteria, protozoa, C4 plants, and another type, Crassulacean, discussed below. Its non-plant function is to produce ATP, in a fashion similar to pyruvate kinase. The net reaction is:

ATP + phosphate + pyruvate = AMP + PP_i + H^+ + phosphoenolpyruvate

Figure 20.5.2 shows a simplified mechanism for the reaction

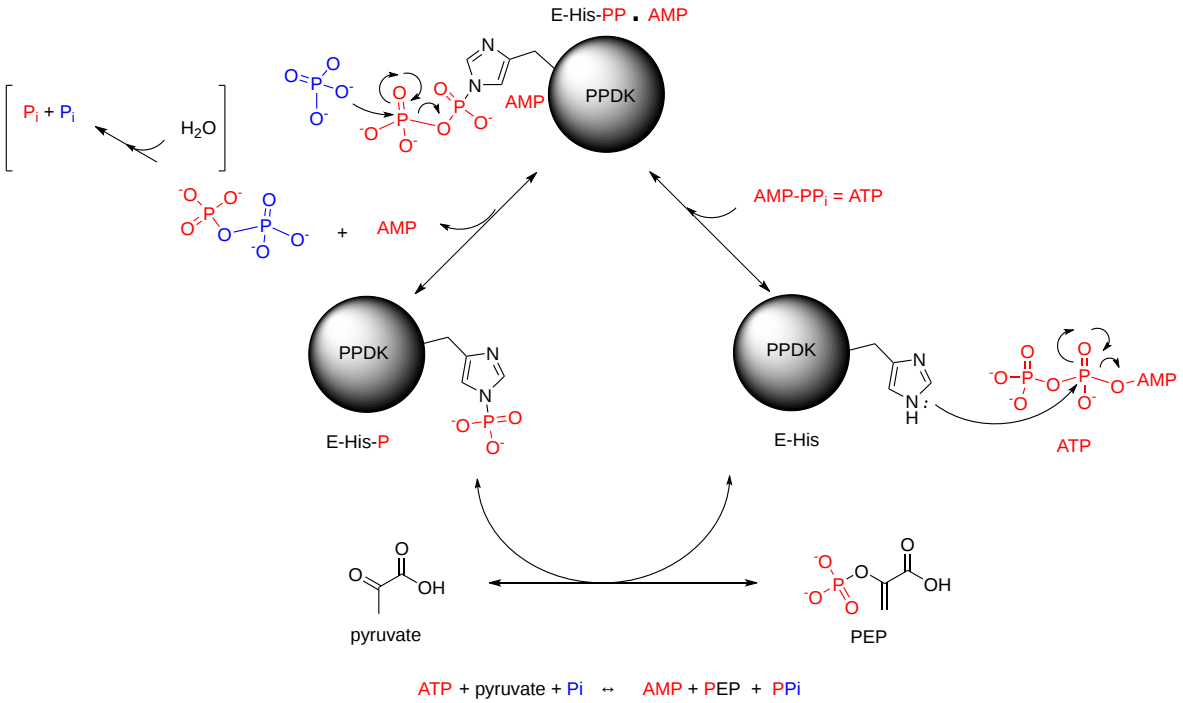


Figure 20.5.2: Mechanism of pyruvate phosphate dikinase

The net reaction shows that in C4 plants, two molecules are phosphorylated by ATP. One is pyruvate and the other is inorganic phosphate (Pi). Hence the name dikinase. The reverse reaction of ATP synthesis occurs in bacteria and protozoan. There are two phosphorylated intermediates, an Enz-His-P, and an Enz-His-PP_i. These served as "activated" phosphate carriers in the phosphotransfer reactions to pyruvate and to P_i, respectively. If PP_i is not hydrolyzed to $2P_i$, as illustrated in the top left of Figure 2 above, the reaction is fully reversible.

In C4 plants there are three reactions

1. Pyr + E-His-P ↔ PEP + E-His
2. E-His + ATP ↔ E-His-PP ·AMP (· indicates a noncovalent interaction)
3. E-His-PP ·AMP + P_i ↔ E-His-P + AMP + PP_i

When PP_i is hydrolyzed, the net input of ATP to phosphorylated pyruvate is two ATP equivalents.

In the next step, PEP is carboxylated in a carbon capture reaction by PEP carboxylase, which as mentioned above used HCO_3^- as a substrate, not CO_2 per se. PEP carboxylase also does not have a competing oxidase activity. The product is malate, which releases locally high "saturating" concentrations of CO_2 in the bundle sheath cells, which significantly suppresses the oxygenase activity of RuBisCo.

Pyruvate phosphate dikinase undergoes a very large change in a conformational change in domain organization during the catalytic cycle. The two kinase activities are located at different sites in the enzyme. The phosphorylation of Pi occurs in the N-terminal domain, while the phosphorylation of pyruvate is in the C-terminal domain. The center domain that links to the N- and C-terminal domains by associated "tethers" is the site of the catalytic histidine involved in phosphotransfer. A swiveling of domains occurs to allow sequential phosphotransfers.

The Pfam domain structure for the protein is shown in Figure 20.5.3.

Figure 20.5.3: The green represents the AMP/ATP nucleotide-binding domain (NBD), the blue the PEP/pyruvate binding domain (PBD), and the red the central domain which catalyzes the phosphotransfer from ATP to pyruvate and from PEP to AMP through phosphorylated histidine intermediates. The NBD has three subdomains.

Figure 20.5.4 shows an interactive iCn3D model of an AlphaFold predicted model of chloroplast pyruvate, phosphate dikinase from *Flaveria brownii (Brown's yellowtops)* (Q39734)

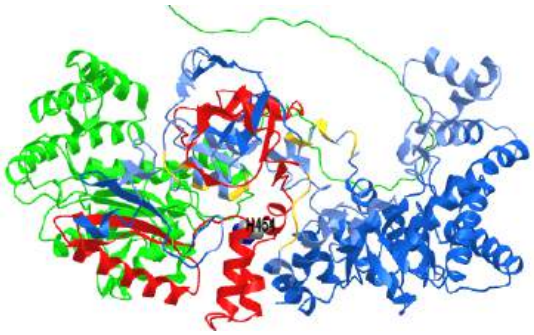


Figure 20.5.4: AlphaFold predicted model of chloroplast pyruvate, phosphate dikinase from *Flaveria brownii (Brown's yellowtops) (*Q39734). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...VB6Aywuc2bLD38

The green represents the N-terminal NBD, the red is the central domain with the catalytic His (side chain shown as CPK-colored spheres and labeled), and the blue the PEP/Pyr PBD).

Figure 20.5.5 illustrates the conformation changes that occur in the central domain (yellow in this figure) and the N-terminal nucleotide-binding domain (NBD, green) in various ligand-bound states.

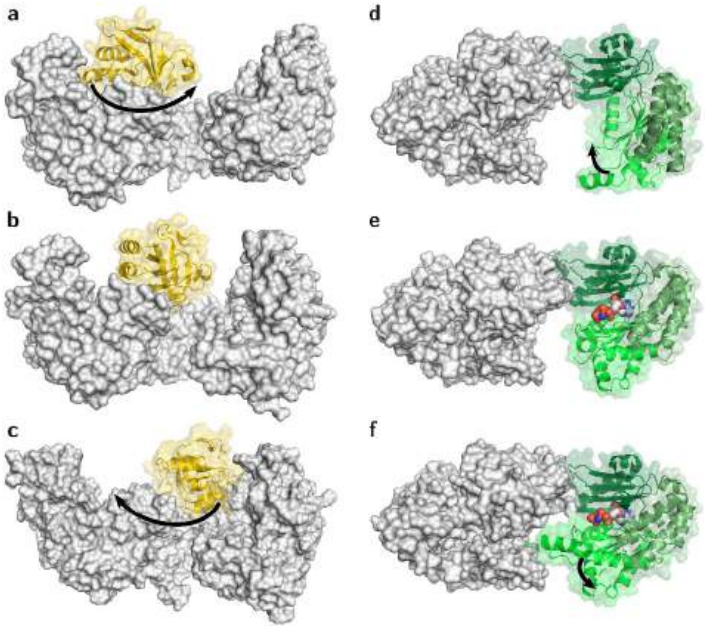


Figure 20.5.5: **Movement of the PPDK domains**

Panel a–c show movement of the central domain (yellow). In (**a**) it swivels to face the PBD domain (PDB 5JVL/C) and in (**c**) it faces the NBD domain (*Tb*PPDK (PDB 2X0S)). In (**b**) it is in an intermediate position.

Panels d-f show the movement of the NBD (the three subdomains are depicted by three different greens). Pane (**d**) is the state without bound nucleotide (PDB 5JVJ/A), while panel (**f**) shows the fully closed, nucleotide-bound state (PDB 5JVL/A). Panel (**e**) shows a semi-closed, nucleotide-bound state (PDB 5JVL/C). Minges, A., Ciupka, D., Winkler, C. *et al.* Structural intermediates and directionality of the swiveling motion of Pyruvate Phosphate Dikinase. *Sci Rep* **7,** 45389 (2017). https://doi.org/10.1038/srep45389. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/

Figure 20.5.6 shows interactions of bound substrates in different conformation states of PPDK.

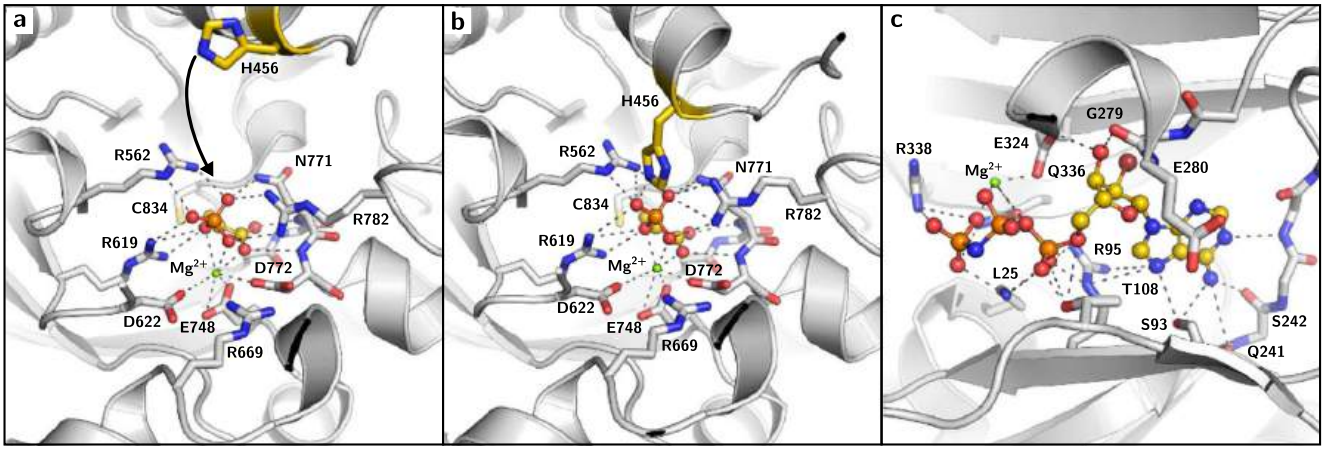


Figure 20.5.6: **Substrate binding sites of *Ft*PPDK.**

Panel (**a**) shows the semi-closed state of the PEP binding site (PDB 5JVL/A) with the catalytic H456 (yellow) pointing away from PEP.

Panel (**b**) shows the closed state of the PEP binding site (PDB 5JVL/C) showing interactions between PEP and surrounding residues, including the catalytic H456 (yellow).

Panel (**c**) shows the closed state of the nucleotide-binding site of 5JVL/D occupied with 2'-Br-dAppNHp, a nonhydrolyzable ATP analog. Interacting residues are highlighted. Minges, A et al. ibid.

After CO_2 is delivered from malate in the bundle sheath cell using RuBisCo, the rest of the reactions are the same as in the C3 pathways.

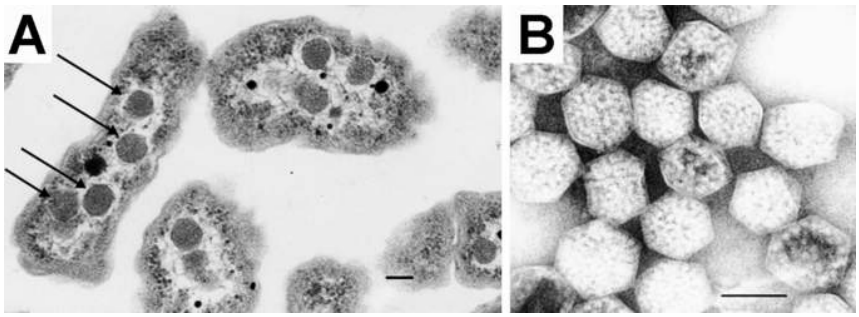
Once CO_2 is fixed into 3-phosphoglycerate in the bundle-sheath cells, the other reactions of the Calvin cycle take place exactly as described earlier. Overall the C4 pathways require more ATP. A molecule of PEP is required for each CO_2 fixed in the C4 pathway which takes the equivalent of two ATPs. So for each CO_2 fixed in the C4 pathway, it takes 5 ATPs compared to 3 ATPs in the C3 pathway. As mentioned above the affinity (estimated from the K_M) of CO_2 for RuBisCo decrease with increasing temperature, which decreases the energetic efficiency of carbon capture. At higher temperatures (28-30 C), the extra energy cost for the C4 pathway balances out with the extra energy cost for the C3 pathway at higher temperatures.

20.5.2: CARBONIC ANHYDRASES

We first encountered carbonic anhydrase when we discussed its mechanism in Chapter x.xx. We'll now discuss its function and activity in the C4 pathway in some detail. Given that we are facing a climate crisis due to the increasing levels of CO_2 in the atmosphere arising from the burning of fossil fuels, removing CO_2 from the atmosphere, a process called for climate purposes carbon sequestration, and understanding the role of carbonic anhydrase in CO_2 sequestration becomes even more important.

Plants have three different genes for carbonic anhydrase (α, β, and γ). Each can be differentially spliced so there many different isoforms of this protein in plants. They are most abundant in the chloroplast, cytoplasm, and in mitochondria and they have many additional roles outside of fixing CO_2 in C4 (and also CAM) plants. In autotrophic (make their own "food") bacteria (such as cyanobacteria, also known as blue-green algae), there are no internal organelles. However, there are **carboxysomes,** which are protein-bounded vesicles (much like a bacteriophage head), which contain in their internal compartment not nucleic acid but RuBisCo and carbonic anhydrase. The carbonic anhydrase converts HCO_3^- to CO_2 for reaction with RuBisCo. The carboxysome hence concentrates the CO2-producing and fixing enzymes for photosynthesis.

Figure 20.5.7:



F

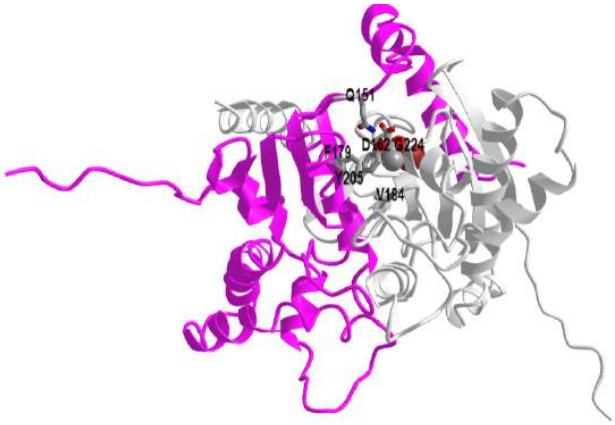
Figure 20.5.7: **Electron Micrograph of a carboxysome**: (A) A thin-section electron micrograph of H. neapolitanus cells with carboxysomes inside. In one of the cells shown, arrows highlight the visible carboxysomes. (B) A negatively stained image of intact carboxysomes isolated from H. neapolitanus. The features visualized arise from the distribution of stains around proteins forming the shell as well as around the RuBisCO molecules that fill the carboxysome interior. Scale bars indicate 100 nm. BioLibreText

In C3 plants, CO_2 (aq), that is dissolved CO_2, is the actual substrate for RuBisCo so available HCO_3^- is converted to CO_2 by carbonic anhydrase. In C4 and CAM plants, CO_2 (aq) is first converted to bicarbonate by carbonic anhydrase. HCO_3^- (aq) is then used as the actual substrate for the "carbon fixation" step. Hence carbonic anhydrase has roles in C3, C4, and CAM plants.

All of the carbonic anhydrases have a Zn^{+2} at the active site. The **alpha** form, the most prominent in plants, was discovered in erythrocytes and is typically active as a monomer. It has one large 10-strand beta sheet surrounded by 7 alpha helices. The Zn ion is tetrahedrally coordinated by 3 histidine side chains and water. **Gamma** carbonic anhydrase is a trimer with three active sites at the interface between pairwise monomers with the Zn-coordinating histidine side chains from two different subunits.

Beta carbonic anhydrase in plants is typically an octamer of identical subunits. The Zn ion is coordinated by two cysteines, one histidine, and water. The monomer has 4 beta strands in a beta-sheet surrounded by alpha helices. An additional beta-strand is involved in monomer interactions. As the active site is in the interface of two subunits, the functional biological unit is the dimer, but a tetramer and even an octamer are typically formed. The substrate binding groups have a one-to-one correspondence with the functional groups in the alpha-carbonic anhydrase active site, with the corresponding residues being closely superimposable by a mirror plane. Therefore, despite differing folds, alpha- and beta-carbonic anhydrases have converged upon a very similar active site design and are likely to share a common mechanism.

Figure 20.5.8 shows an interactive iCn3D model of beta-carbonic anhydrase from Pisum sativum (pea) with bound acetate (1EKJ).



NCBI iCn3D Figure 20.5.8: Beta-carbonic anhydrase from Pisum sativum (pea) with bound acetate (1EKJ). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...VZTEJPy1hPxK27

For the sake of simplicity, only two subunits (A - gray and B - magenta) of the octamer are shown. Acetate, a proxy for HCO_3^{-2}, is shown in spacefill binding between the two subunits in CPK-colored spheres. The side chains involved in Zn^{2+} binding, Cys 160, His 220, and Cys 223 (unlabeled) are on the A chain. The amino acids that bind acetate are distributed between the A chain (Asp 162, Gly 224, Val 184) and the B chain (Gln 151, Phe 169, and Tyr 205) and are labeled. The same groups are involved in substrate binding in alpha carbonic anhydrase but in a mirror image orientation (but with the normal L-amino acids).

Beta carbonic anhydrase has high levels of expression in leaves and is found in chloroplasts, mitochondria, and the cytoplasm. Many plants have it in the cytoplasm and chloroplasts.

- β-carbonic anhydrase (βCA) in C3 plants: Most of βCA in leaves is in chloroplasts in mesophyll cells and may comprise up to 2% of leave protein. Yet studies have shown that you can delete the gene for it with minimal effect on the maximal rate of photosynthesis.

However, plant development was affected so by interference the enzyme is probably most needed to produce HCO_3^- for biosynthesis.

- β-carbonic anhydrase (βCA) in C4 plants: Most of the βCA is found in the cytoplasm of mesophyll cells. There is catalyzes the first reaction of the C4 pathway, CO_2 (aq) to HCO_3^- (aq). Mitochondrial (βCA) and γCA probably function to fix CO_2 arising from oxidative respiration.

20.5.3: CRASSULACEAN ACID METABOLISM (CAM) PATHWAY

Plants that encounter the chronic stress of low water availability have evolved yet another pathway to adapt to low water conditions. Stomata in C3 plants are open during the day to allow carbon capture from CO_2, but they can close when water is limited. This obviously will inhibit plant growth. In the CAM pathway, the stomata stay open at night, allowing carbon capture at a time when water loss through the stomata would be lower. The incoming CO_2 is fixed through carbonic anhydrase and then a series of several enzymes to form malic acid, which is transported for storage and use during the light in vacuoles. This CAM pathway is described in Figure 20.5.9.

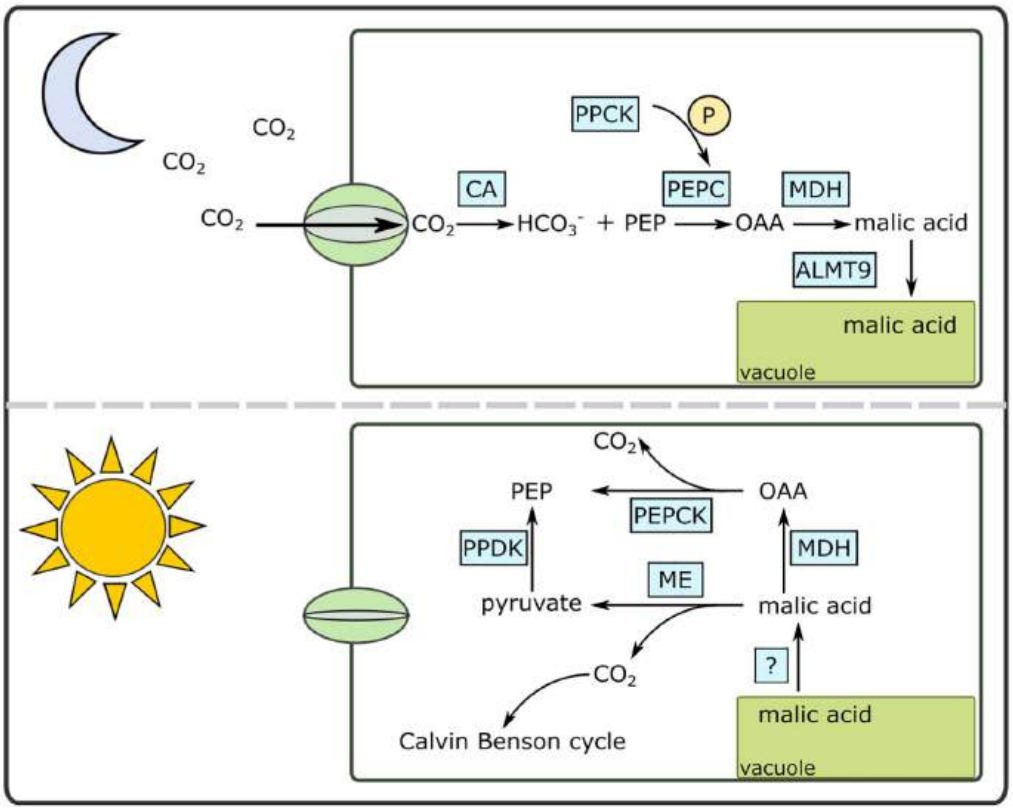


Figure 20.5.9: A simplified diagram of the Crassulacean acid metabolism (CAM) pathway under day and night conditions. Heyduk Karolina, et al. Frontiers in Plant Science.2019. https://www.frontiersin.org/article/...pls.2018.02000. DOI=10.3389/fpls.2018.02000. Creative Commons Attribution License (CC BY).

The proteins and intermediates in the CAM pathways are ALMT9, aluminum-activated malate transporter; CA, carbonic anhydrase; MDH, malate dehydrogenase; OAA, oxaloacetate; ME, malic enzyme (NAD or NADP); P, phosphate; PEPC, phosphoenolpyruvate carboxylase; PEPCK, PEP carboxykinase; PPCK, PEPC kinase; PPDK, pyruvate, phosphate dikinase.

During the day, malic acid moves back into the cytoplasm where it is decarboxylated by malic enzyme and releases locally high CO_2 concentrations for use by RuBisCo in the C3 cycle. In plants that use CAM, a series of other changes occur, including leaf structure and additional regulatory processes that coordinate metabolic genes.

Figure 20.5.10s shows more details of the CAM cycle.

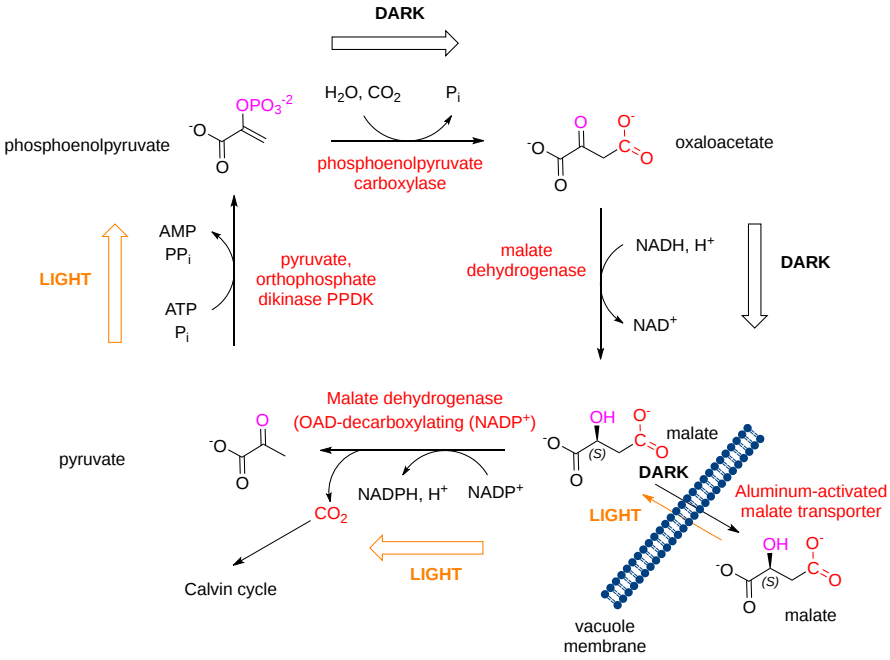


Figure 20.5.10: Details of the CAM cycle

20.5.4: THERMODYNAMIC COMPARISON BETWEEN THE C3/C4 PATHWAYS AND THE OTTO CYCLE

It is interesting to compare the thermodynamic efficiencies of plants and internal combustion engines that are governed by the thermodynamic Otto cycle. In the cycle, chemical energy in the form of gasoline and O_2 are converted to thermal energy which is converted into mechanical energy. Both photosynthesis (the conversion of the energy of photons into chemical energy) have limited efficiencies.

- In internal combustion engines, power is limited in part by air uptake and by different efficiencies when running at non-constant speeds (like stop-and-go).
- In the C3 pathway, efficiency is limited by the oxygenase activity of RuBisCo/Oxygenase (given the much higher concentration of atmospheric O_2 compared to CO_2) and the ensuring photorespiration pathway.
- Water availability also plays a role as the net reaction of photosynthesis and glucose production, in simplified form, is $6CO_2 + 6H_2O \rightarrow 6C(H_2O_6)$ so in high heat and low humidity the process efficiency decreases

Figure 20.5.11 shows a comparison of three stages, the storage component, the basic cycle, and a concentration mechanism, in photosynthesis and the internal combustion engine (ICE).

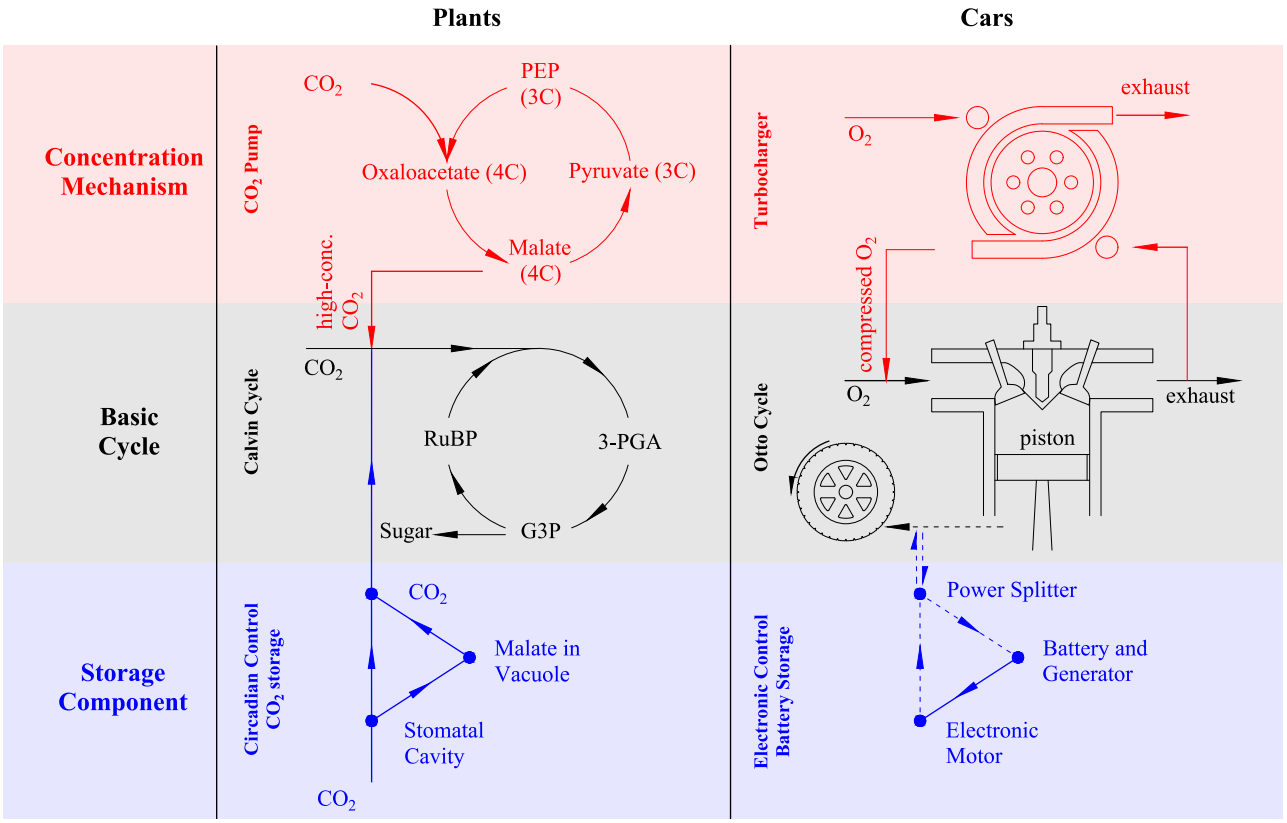


Figure 20.5.11: A comparison of plant photosynthesis and car engine functioning illustrates how the core processes interact with the additional components. Hartzell S, Bartlett M, Yin J, Porporato A (2018) Similarities in the evolution of plants and cars. PLoS ONE 13(6): e0198044. https://doi.org/10.1371/journal.pone.0198044. Creative Commons Attribution License,

A concentrating mechanism in C4 plants and turbocharged cars provides concentrated CO_2 and oxygen, respectively, to the core cycle (upper row). A storage mechanism in CAM plants allows carbon dioxide to be stored as malic acid at night and then passed to the Calvin cycle during the day, while a storage mechanism in hybrid electric vehicles (HEVs) allows energy to be stored in the battery during braking and then passed to the motor to power the drivetrain in parallel with the engine (bottom row).

20.6: BIOSYNTHESIS OF STARCH, SUCROSE AND CELLULOSE

Now that we have seen how carbon is captured and fixed into 3C trioses, which can be converted to fructose and glucose and their derivative, we can now explore the synthesis of the key plant carbohydrates we all know, sucrose, starch, and cellulose.

The source for the organization and some of the text derives from Sindayigaya and Longhini. https://www.peoi.org/Courses/Courses...chem/biochem18. Creative Commons - https://creativecommons.org/licenses...sa/3.0/deed.en

20.6.1: SUCROSE SYNTHESIS

Sucrose is a disaccharide of glucose and fructose with an acetal link between the anomeric carbons to form the nonreducing sugar O-α-D-glucopyranosyl-(1→2)-β-D-fructofuranoside. Its structure is shown in Figure 20.6.1.

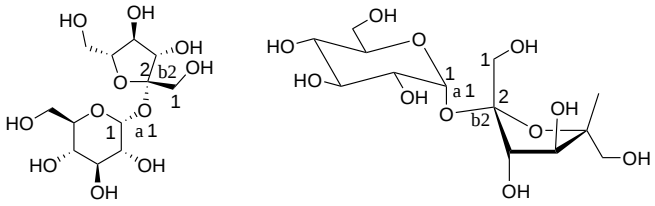


Figure 20.6.1: Sucrose, a disaccharide of glucose and fructose

Sucrose can be considered a transport form of carbon, much like ketone bodies are transport forms of fatty acids. As noted above, the link between the sugars is between the anomeric C-1 link of glucose and the anomeric C-2 link of fructose. As such, it is not cleaved by typical carbohydrate-cleaving enzymes like amylases. Also, it doesn't react with proteins like other sugars with free cyclic hemiacetals that can open and form reactive aldehydes. For example, the cyclic monosaccharide glucose, with its anomeric carbon in a readily reversible hemiacetal link, can form covalent bonds to amine groups in proteins such as hemoglobin, which forms the glycosylated version of HbA1c, as shown in Figure 20.6.2.

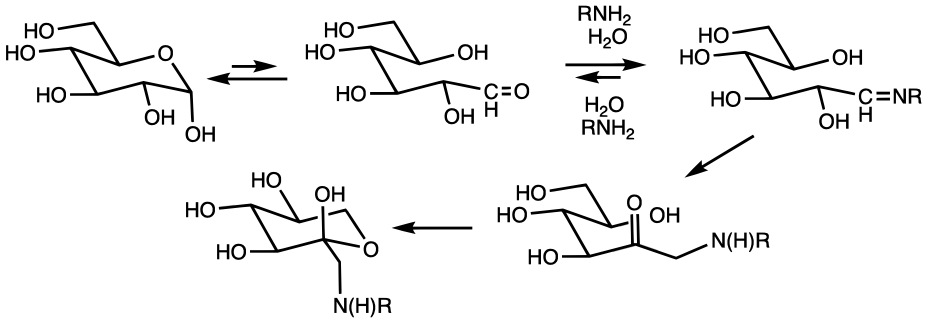


Figure 20.6.2: Reactions of cyclic hemiacetal sugars with amines. https://en.wikipedia.org/wiki/Glycated_hemoglobin. Creative Commons Attribution-ShareAlike License 3.0

The reaction proceeds through the formation of a Schiff base followed by a rearrangement. HbA1c is a marker of diabetes. These features of sucrose may explain its choice as a key form of synthesized carbohydrate in plants.

Sucrose is synthesized in the cytosol, as shown in Figure 20.6.3.

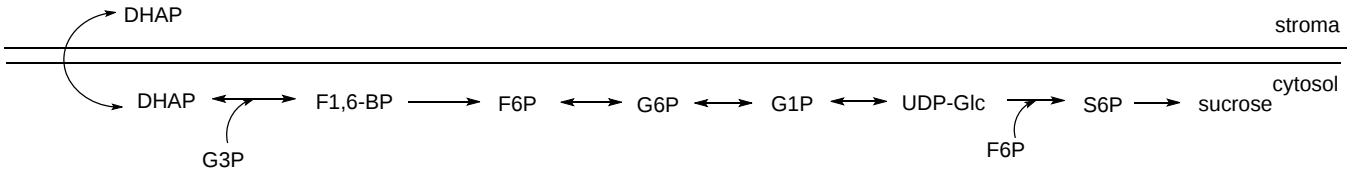


Figure 20.6.3: Synthesis of sucrose in plants.

We have seen the enzymes that catalyze these reactions through G1P before in the glycolytic and gluconeogenesis pathways. Glucose-1-phosphate is then converted to UDP-glucose, which reacts with fructose-6-phosphate to form sucrose 6-phosphate and UDP, catalyzed by the enzyme **sucrose 6-phosphate synthase.** The last reaction is the hydrolysis of sucrose 6-phosphate by sucrose 6-phosphatase, which allows the export and transport of sucrose.

In the reaction scheme above, G1P is converted to UDP-Glc through the following reaction:

$$G1P + H^+ + UTP \leftrightarrow UDP\text{-}Glc + P_i$$

This reaction is just mildly favored thermodynamically.

Before we study the structure of the enzyme, we will first discuss the regulation of the sucrose synthesis pathway and discuss the generic mechanism of glycosyltransferases.

20.6.2: REGULATION

Carbon capture in the light reaction of photosynthesis leads to sucrose (for transport) and starch synthesis. Which product(s) result depends on key regulatory steps. Remember of the 6 trioses formed in the Calvin cycle, 5 are returned into the cycle for the synthesis of ribulose 1,5-bisphosphate, and only 1 is used for the synthesis of sucrose and starch. If too much is removed, the cycle slows. If not enough is used for starch and sucrose synthesis, P_i, which is moved into the stroma from the cytoplasm by an important translocator (see section 20.3) would run low.

The key cytosolic regulatory step is catalyzed by **fructose 1,6-bisphosphatase** (FBPase-1) and also a unique plant enzyme, **PPi-dependent phosphofructokinase** (PP-PFK-1)**,** that catalyzes the reverse reaction of F6P → F1,6BP, in regulatory steps that are similar to ones found in glycolysis and gluconeogenesis.

- **fructose 1,6-bisphosphatase** (FBPase-1) is inhibited by the allosteric modulator **fructose 2,6-bisphosphate** (F2,6BP)
- **PPi-dependent phosphofructokinase (**PP-PFK-1) is activated by **fructose 2,6-bisphosphate**

The coordinate regulation of these two enzymes is shown in Figure 20.6.4.

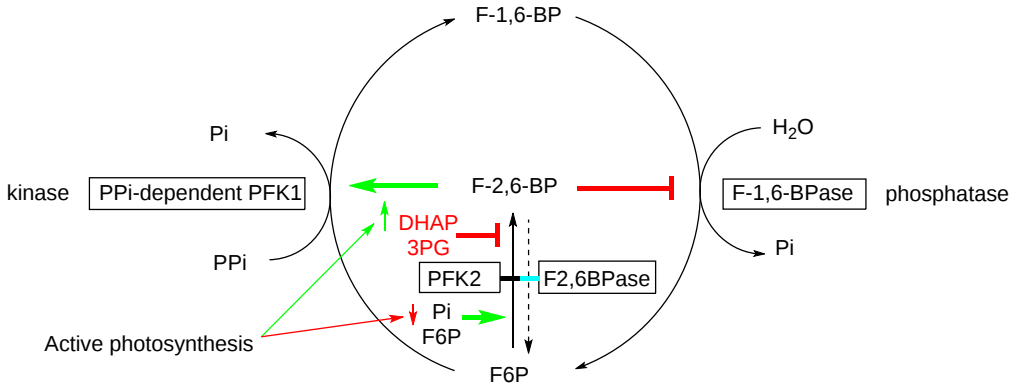


Figure 20.6.4: Regulation of sucrose synthesis at the formation of fructose-6-phosphate

The substrate for the PPi-dependent phosphofructokinase (PP-PFK-1) is PPi, which serves as a phosphodonor. In the plant cytosol, there is no pyrophosphatase to catalyze the cleavage of PP_i into P_i. Also note that Fru-2,6-P_2 itself is synthesized and degraded by the bifunctional enzyme phosphofructokinase 2/fructose-2,6-bisphosphatase, which we studied before in the regulation of the same step in glycolysis/gluconeogenesis.

The levels of F2,6-BP depend on the rate of photosynthesis:

- When photosynthesis is high (in light conditions), [DHAP] and [3PG] increase, which inhibits PFK2, which decreases F26BP, which causes a differential increase in F1,6BPase activity over PP-PFK-1), which increases F6P for sucrose synthesis as well as Pi for the continuation of the light reactions. This allows sucrose synthesis when excess DHAP and 3PG occur in the light reactions, which makes great biological sense.
- When photosynthesis is low (in dark conditions), the same regulations lead to an increase in F2,6BP, which leads to the preferential activation of the glycolytic enzyme PPi-dependent phosphofructokinase-1 (PP-PFK1) and inhibition of the gluconeogenic enzyme fructose 1,6- bisphosphatase (FBPase-1)

We will see later, the main regulatory step in starch synthesis is **ADP-glucose pyrophosphorylase.** In contrast to the inhibition by 3PG of PFK2, 3PG (which increases in active photosynthesis) activates ADP-glucose pyrophosphorylase while Pi inhibits it. Pi increases when the synthesis of ATP from ADP and Pi (by ATP synthase in the light reaction) slows (such as in darker conditions). If sucrose synthesis slows and sufficient 3PG persists, the activation of ADP-glucose pyrophosphorylase stimulates starch synthesis.

20.6.3: GLYCOSYLTRANSFERASES (GTS)

Glycosyltransferases are very important enzymes as they are involved in the synthesis of most of the biomass on the planet. They catalyze the transfer of a monosaccharide from a donor that has been activated by the attachment of a nucleotide in the form of a nucleotide sugar (NDP-sugar) or dolichol-(pyro)phosphate sugar to acceptors. These include other sugars, lipids, and even proteins, which get glycosylated

on alcoholic side chains (Ser, Thr) or amides (Asn). There are over 500,000 different GTs known and deposited in the Carbohydrate-Active enZYmes Database (CAZy database2). Based on sequence and structure there are over 114 different families. Although they depart significantly in primary sequence, only 3 major folds are predominant (GT-A, -B, and -C)

a glycosyl transferase reaction is required for the transfer of glucose from a donor like UDP-glucose to an acceptor like fructose to form sucrose as shown in the reaction below.

NDP-Glc (donor) + F6P (acceptor) → Sucrose-6-P + NDP

Perhaps now is a good time to study their generic reaction mechanisms before we move on to starch synthesis.

About 65% of glycosyltransferase reactions use nucleotide sugars as donors. These enzymes are called Leloir transferases. They are nucleotide-dependent. The activated NDP-sugar donor binds first to the enzyme, followed by the acceptor, to form a ternary complex. A conformational change allows catalysis to occur, which leads to the sequential release of products. The enzyme hence follows a sequential ordered bi-bi mechanism.

The reaction could proceed with either retention or inversion of the anomeric carbon of the donor NDPsugar. This is illustrated for the reaction of a C1 α-NDP donor monosaccharide with a monosaccharide acceptor to produce the α(1,4) link with retention of configuration or the β(1,4) link with inversion as shown in Figure 20.6.5.

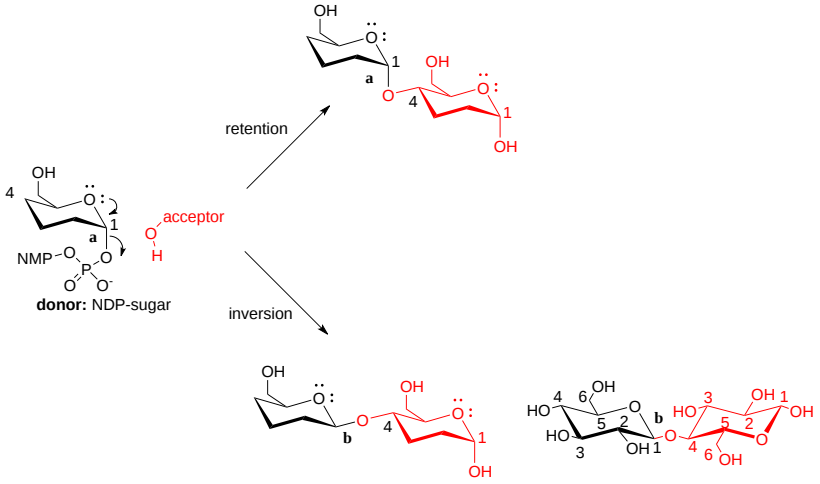


Figure 20.6.5: Reaction of a donor NDP-monosaccharide and an acceptor monosaccharide with retention or inversion of configuration at the anomeric carbon of the donor

The same stereochemical outcomes can occur in the hydrolysis of acetal bonds by glycosyl hydrolases.

Reactions that proceed with inversion react in an S_N2 reaction, similar to the nucleophilic attack on alkyl halides. For the glycosyl transferase that proceeds with inversion, the attacking nucleophile on the acceptor is made more nucleophilic by general base catalysis by a deprotonated glutamic or aspartic acid.

The glycosyl transferase that proceeds with the retention of configuration is less understood. Several alternative mechanisms have been proposed for both inverting and retaining glycosyltransferase in general, as illustrated in Figure 20.6.6.

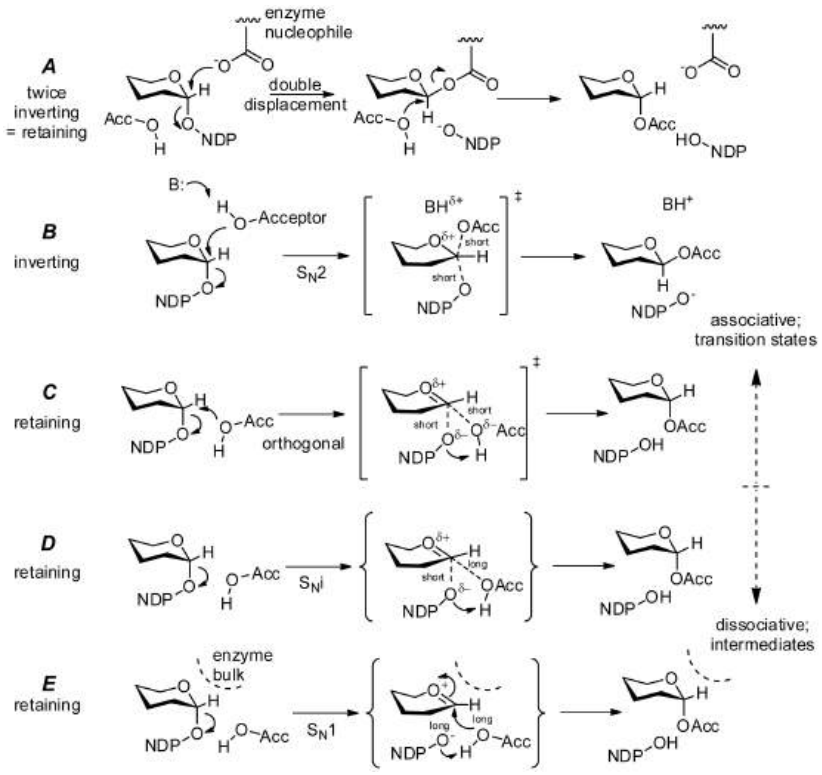


Figure 20.6.6: .Proposed glycosyltransferase mechanisms. Schuman et al. PLOS One (2013). doi:10.1371/journal.pone.0071077.g001. Creative Commons Attribution License

They include the following possible mechanisms:

Panel (A): A double displacement mechanism utilizing two inversions with net retention of stereochemistry involving a covalent glycosyl-enzyme intermediate. The individual steps are inverting via (B) an S_N2 process.

Panel (B): Inverting Leloir glycosyltransferases promote a backside nucleophilic attack on C1 by the acceptor from an inline (usually equatorial) position, resulting in inversion of the anomeric bond stereochemistry.

Panel (C): An orthogonal mechanism consisting of nucleophilic attack on C1 by the acceptor concurrent with leaving group loss from a position approximately at right angles to the C1-leaving group axis.

Panel (D): An S_Ni mechanism involving an intermediate with oxocarbenium character followed by rapid internal nucleophilic attach by the acceptor nucleophile; or

Panel (E): An S_N1 mechanism involving a discrete oxocarbenium intermediate. All mechanisms require proton transfers of the hydroxyl hydrogen of the acceptor to an enzymatic base or the departing leaving group

You will remember from your studies of chemistry that S_N1 reactions are dissociative and form a positively charged carbocation intermediate. In a S_Ni, reaction, the intermediate has cation character but the intermediate is not fully charged. In the case of glycosyltransferase, the intermediates would be oxycarbeniums. (Carbenium ion can be considered carbocations with 3 bonds to the carbon). S_N1 reaction will occur only if the formation of the ion is activated and they are stabilized. A protic solvent is typically required for stabilization if the reaction occurs in solution. In the anhydrous active site of the enzyme, an appropriate arrangement of backbone and side chain negative or partially negative atoms is required to provide stabilization for S_N1 and S_Ni mechanisms.

The double replacement reactions (Panel A) require a side chain nucleophile and likely candidates in retaining glycosyltransferase are not positioned for such a task. The evidence seems to support the orthogonal mechanism. It appears that the binding of the donor is similar in retaining transferase such that it is in a 90^0 position of nucleophilic attack by the acceptor, which leads o a trigonal bipyramidal transition state with the nucleophile axial the leaving group equatorial (orthogonal).

The donor NDP-monosaccharides typically are Mn^2-containing proteins with the inverting and retaining transferase having different coordination geometries for Mn^{2+} binding, as illustrated in Figure 20.6.7.

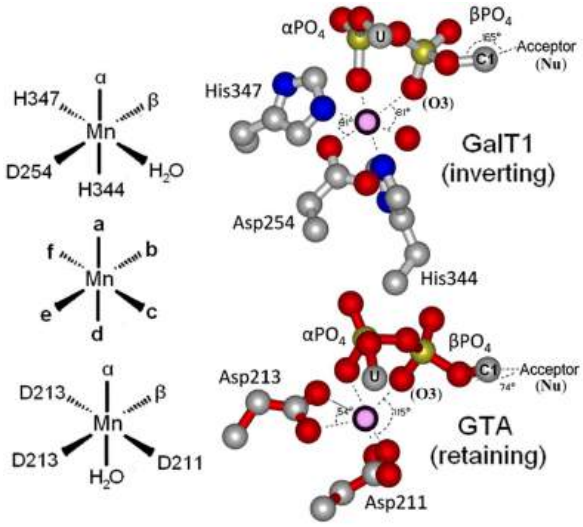


Figure 20.6.7: Coordination geometries for Mn^{2+} binding in glycosyltransferase. Shuman, ibid.

Inverting enzymes such as GalT1 (top) achieve nearly perfect octahedral geometry about the coordinated metal ion (displayed angles of 81° and 91° compared to ideal octahedral 90° bond angles) with subsequent "inline" (approaching 180°) placement of the acceptor nucleophile for classic inverting S_N2 backside attack. Retaining enzymes such as GTA (bottom), however, use an arrangement of acidic residues, often with acute bidentate Asp coordination, which severely skews metal geometry (displayed angles of 54° and 115°) and allows sufficient room between phosphate oxygens for orthogonal attack from the acceptor. **U**is uridine, **C1**is donor galactose **C1**.

Figure 20.6.8 shows a possible mechanism for the transfer of a monosaccharide from the donor ADP-sugar through an oxycarbenium intermediate to an acceptor.

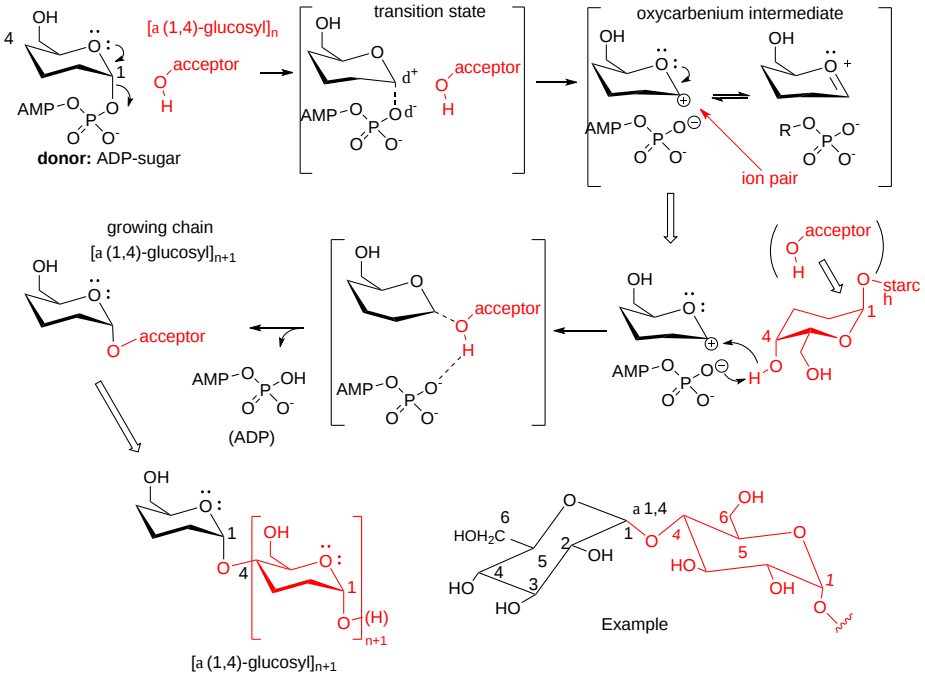


Figure 20.6.8: Possible mechanism for the transfer of a monosaccharide from the donor ADP-sugar through an oxycarbenium intermediate to an acceptor (example - a growing starch chain). Schuman B, Evans SV, Fyles TM (2013) Geometric Attributes of Retaining Glycosyltransferase Enzymes Favor an Orthogonal Mechanism. PLoS ONE 8(8): e71077. doi:10.1371/journal.pone.0071077. Creative Commons Attribution License

As mentioned above, there are three major folds for glycosyltransferases, GT-A, GT-B, and GT-C. Different representations of the structure of the GT-A fold core predicted through analysis by neural networks and deep learning are shown in Figure 20.6.9.

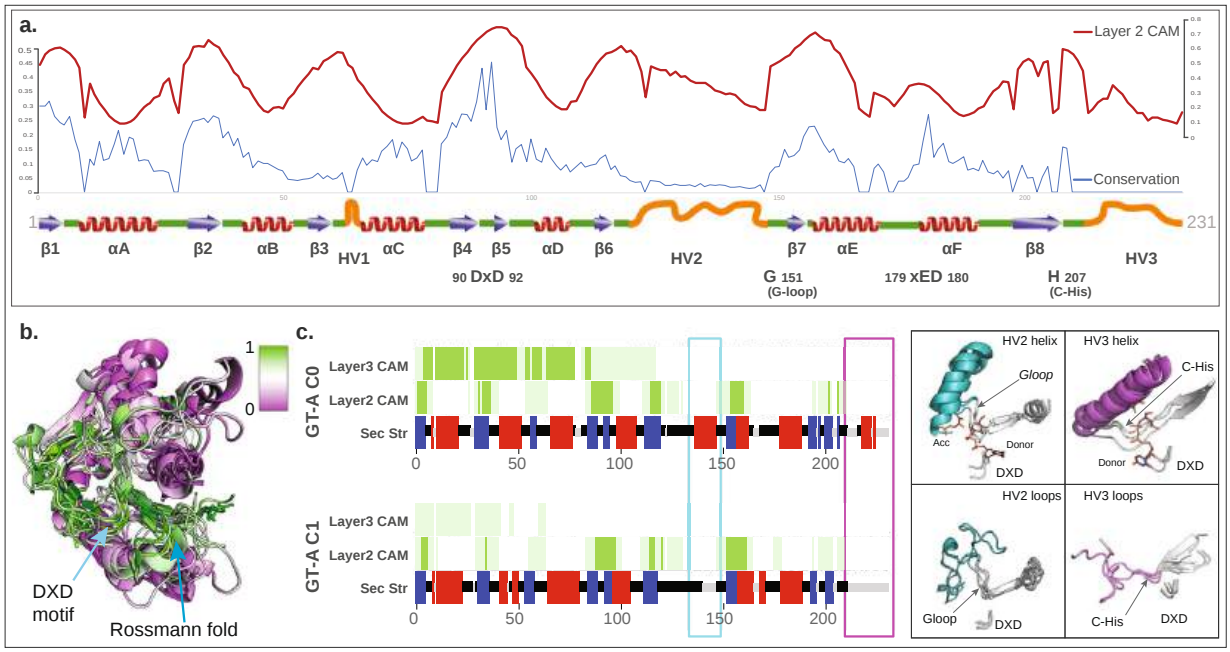


Figure 20.6.9: Fold core of GT-As. Taujale, R., Zhou, Z., Yeung, W. *et al.* Mapping the glycosyltransferase fold landscape using interpretable deep learning. *Nat Commun* **12,** 5656 (2021). https://doi.org/10.1038/s41467-021-25975-9. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

This computational modeling of structure uses simple secondary structure representations generated from primary sequences to predict folds irrespective of sequence. GT folds are predicted with high accuracy by learning secondary structure features free of primary sequence alignments.

Panel (a) shows a linear map of the conserved secondary core structures (below) and graphs (top). The blue line represents a conservation score and the red a CAM score. CAM values correspond to residue positions that distinguish them the most from other class labels (folds and families).

Panel (b) shows a structural alignment of the core with the CAM values mapped onto it. The conserved regions are shown to have a high CAM value indicated by a high intensity of green and a low CAM value of purple.

Panel (c) shows on the left the consensus secondary structure for the aligned positions in the two GT-A fold clusters (blue: beta-sheets; red: helices; green: loops). Average CAM values from using different "layers" of analysis are shown for each aligned position (higher intensity of green corresponds to a higher CAM value). Cyan and magenta boxes highlight the secondary structure differences between the two clusters near the hypervariable HV2 and HV3 region respectively. The conserved DXD motif, G-loop, and C-His are indicated for reference. Donor and acceptor substrates for GT-A0 are shown as sticks

This figure is shown to give readers a sense of the complexity of the analysis required to understand and differentiate structure/function features for these structurally similar but complex enzymes.

20.6.4: STRUCTURE AND ENZYMATIC ACTIVITY OF SUCROSE SYNTHASE (SUSY)

These enzymes are usually homotetramers with a monomeric molecular weight of around 90,000. The monomers typically have an N-terminal domain that directs the targeting of the enzyme to a specific location and a C-terminal GT-B domain. It can be regulated by phosphorylation at a serine near 12 in the N-terminal domain, which presumably regulates its cellular location, and other near 170 that affects its degradation. Two glutamates in the C-terminal GT-B domain (E678 and E686) and phenylalanine (680) are essential for the catalytic activity. Sucrose synthase is reversible as is the synthesis of sucrose-6-P from F6P and UDP-glucose can be reversed in the presence of UDP. The enzyme can also use ADP-glucose as a donor.

The structure of the *Thermosynechococcus elongatus* sucrose phosphate synthase with bound UDP and sucrose-6-Phosphate has been solved and along with other studies a reaction mechanism proposed as shown in Figure 20.6.10.

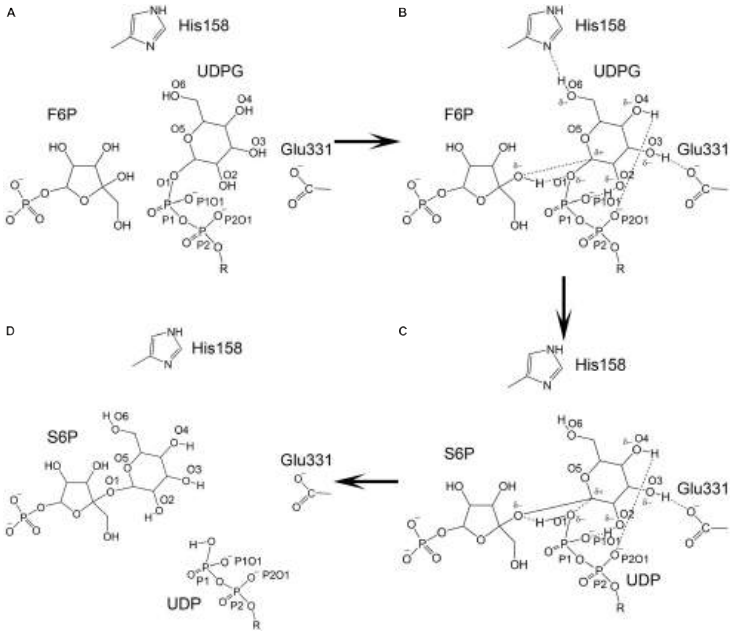


Figure 20.6.10: Catalytic model of TeSPS. Yuying et al. Co-crystal Structure of Thermosynechococcus elongatus Sucrose Phosphate Synthase With UDP and Sucrose-6-Phosphate Provides Insight Into Its Mechanism of Action Involving an Oxocarbenium Ion and the Glycosidic Bond. Frontiers in Microbiology, 11, 2020. https://www.frontiersin.org/article/...icb.2020.01050. **Creative Commons Attribution License (CC BY)**.

(A) The state before the reaction is shown. **(B)**The glucose residue of UDPG forms hydrogen bonds between/among the phosphate groups, His158, Glu331, and F6P. Due to the formation of these hydrogen bonds, the pyranose ring of the glucose becomes negatively charged to promote C1 to form an oxocarbenium ion. **(C)**The relatively weak hydrogen bond formed by His158 and O6 is broken, which causes the pyranose ring to lose some negative charge character and forces the C1 oxocarbenium ion to form a covalent bond with the F6P oxygen atom. **(D)**UDP and S6P are released from the catalytic center.

A 2D view of the active site residues is shown in Figure 20.6.11.

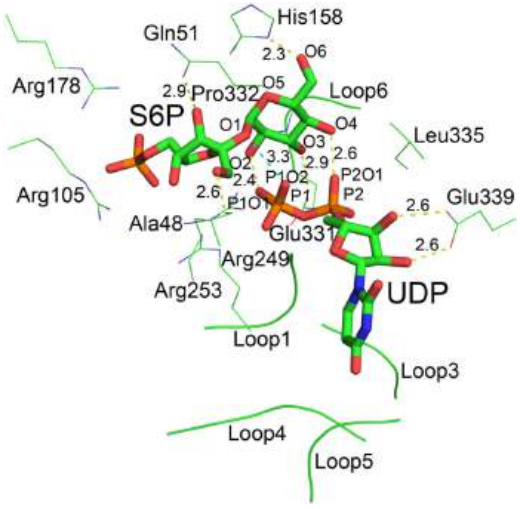


Figure 20.6.11. The catalytic center of TeSPS. Yuying et al. ibid.

The catalytic center of TeSPS. Loops 1, 3, 4, and 5 forms a cave that binds to the uracil moiety of UDP. Glu339 stabilizes the ribose ring via the formation of two hydrogen bonds. Leu335 forces two phosphate groups in UDP to reorient. Several basic amino acids, including Arg105, Arg178, Arg249, and Arg253, interact with the phosphate groups of UDP and S6P via ionic bonds. Pro332 at the turn of loop 6 interacts with the pyranose ring via CH/π bonds. All hydroxyl groups (O2, O3, O4, and O6) of the glucose moiety of S6P form hydrogen bonds with phosphate groups or the side chains of various amino acids. “O2” forms a hydrogen bond with P1O1 of the P1 phosphate group of UDP. “O3” forms a hydrogen bond with the carboxyl group of Glu331. “O4” forms a hydrogen bond with P2O1 of the P2 phosphate

group of UDP. “O6” forms a hydrogen bond with the imidazole side chain of His158. The distances between groups are indicated in the figure

Figure 20.6.12 shows an interactive iCn3D model of the Thermosynechococcus elongatus Sucrose Phosphate Synthase With UDP and Sucrose-6-Phosphate (6KIH)

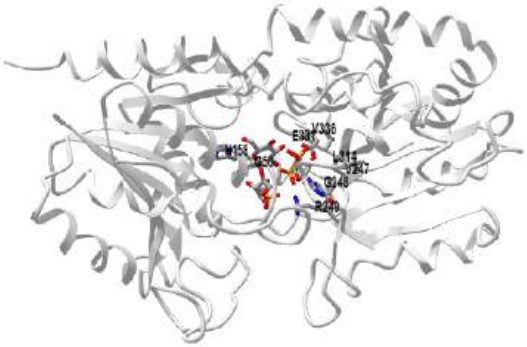


Figure 20.6.12: Thermosynechococcus elongatus Sucrose Phosphate Synthase With UDP and Sucrose-6-Phosphate (6KIH) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...1kCnQ9pxcJjdy7

20.6.5: METABOLISM OF SUCROSE

Sucrose is a primary synthesis product of photosynthesis and is transported to other plant "sink" tissues where it is used for both energy and biosynthetic precursors. Suc transporters can move it from the phloem to the apoplast. It can enter sink cells through Suc transporters or be hydrolyzed by cell-wall invertase (cwINV) to yield glucose (Glc) and fructose (Fru), which enter by hexose transporters. Suc can also pass directly from the phloem to sink cells through plasmodesmata (physical connections between cells). Inside sink cells, Suc can be metabolized or transported to the vacuole, where it can be stored as Suc, transformed into fructans by fructosyltransferases (FTs), or hydrolyzed by vacuolar invertase (vINV) and stored as hexoses. To be metabolized, Suc must be hydrolyzed by either cytosolic invertase (INV) into glucose and fructose, or by the reversible reaction of sucrose synthase (SuSy) using UDP instead of water to yield fructose and UDP-G. These processes are illustrated in Figure 20.6.13.

20.6.5.1:

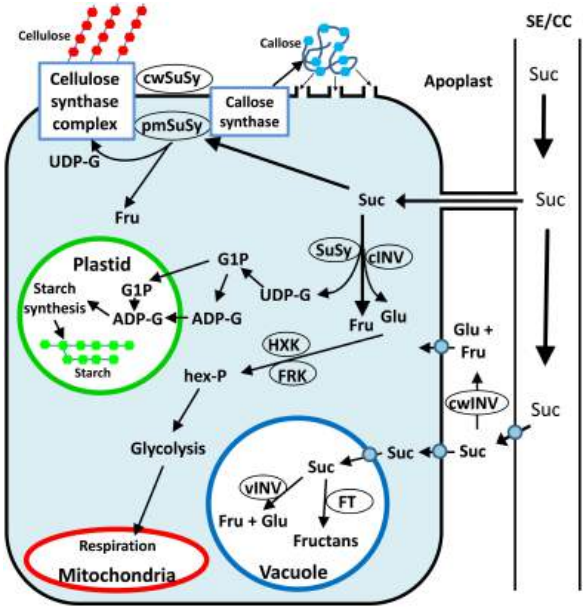


Figure 20.6.13: Simplified schematic presentation of sugar metabolism in sink tissue cells toward cellulose, callose, and starch synthesis. Stein and Granot. Frontiers in Plant Science (10). 2019. https://www.frontiersin.org/article/...pls.2019.00095. DOI=10.3389/fpls.2019.00095 https://www.frontiersin.org/articles...019.00095/full

The hexoses (glucose and fructose) can be phosphorylated to hexose phosphates (hex-P), directed to starch synthesis in the plastid or glycolysis, and then respiration in the mitochondria or directed to other metabolic pathways. Plasma membrane-associated SuSy (pmSuSy)

and cwSUS can generate UDP-G that is used in the synthesis of cellulose for cell walls and callose for plugging plasmodesmata. Callose is a polysaccharide in the form of β-1,3-glucan with some β-1,6-branches in the cell walls of a wide variety of higher plants

20.6.6: BIOSYNTHESIS OF STARCH

During active photosynthesis in bright light, a plant leaf produces more carbohydrates (as triose phosphates) than it needs for generating energy or synthesizing precursors. The excess is converted to sucrose and transported to other parts of the plant, to be used as fuel or stored. In most plants, starch is the main storage form, but in a few plants, such as sugar beet and sugarcane, sucrose is the primary storage form. The synthesis of sucrose and starch occurs in different cellular compartments (cytosol and plastids, respectively), and these processes are coordinated by a variety of regulatory mechanisms that respond to changes in light level and photosynthetic rate.

20.6.6.0.1: STARCH SYNTHESIS

Starch, like glycogen, is a homopolymer of D-glucose in (α1,4) linkage with (α1,6) branches. Glycogen is found in Archaea, Bacteria, and Eukaryotes. In contrast, starch is found in photosynthetic algae, land plants, and in some cyanobacterial species.

Starch is synthesized by starch synthase in chloroplasts for temporary storage and in amyloplasts in seeds, roots, and underground stems (tubers) for long-term storage. As with the synthesis of glycogen, the glycosyltransferase catalyzes the addition of an activated ADP-glucose to the acceptor, the elongating starch polymer.

$$\text{ADP-α-D-glucose-1-phosphate (donor)} + [(1\rightarrow4)\text{-α-D-glucosyl}]_{(n)}\ \text{(acceptor)} \leftrightarrow [(1\rightarrow4)\text{-α-D-glucosyl}]_{(n+1)} + \text{ADP} + H^+$$

The ADP-glucose donor is formed in the following reaction:

$$\text{Glucose 1-phosphate} + \text{ATP} \leftrightarrow \text{ADP-glucose} + \text{PPi}$$

In plastids, there is, in contrast to the cytosol, a pyrophosphatase which makes the reaction irreversible. Hence the next overall reaction is

$$\text{Starch}_n + \text{glucose 1-phosphate} + \text{ATP} \rightarrow \text{starch}_{n+1} + \text{ADP} + 2\text{Pi}$$

Taking into account the hydrolysis by inorganic pyrophosphatase of the PPi produced during ADP-glucose synthesis, the overall reaction for starch formation from glucose 1-phosphate is:

$$\text{Starch}_n + \text{glucose 1-phosphate} + \text{ATP} \rightarrow \text{starch}_{n+1} + \text{ADP} + 2\text{Pi};\ \Delta G'^{o} = -50\ \text{kJ/mol} = -12\ \text{kcal/mole}\ (-50\ \text{kJ/mol})$$

$$\Delta G'^{o} = -50\ \text{kJ/mol} = -12\ \text{kcal/mol}$$

In glycogen synthesis, the donor is UDP-glucose, and it is added to the reducing end (C1) of the growing glycogen polymer (with the C1 OH acting as a nucleophile) so the polymer extends from that end. Kinetic models suggest that starch synthases, which use ADP-glucose, may use two different active sites that appear to alternately add glucose to the nonreducing C4 end (with the C4 OH acting as a nucleophile), with the reducing end of the linear α(1,4) polymer being alternately covalently attached to one site, then the other, with the attachment activating that end for reaction with C4-OH of the polymer at the other site.

Figure 20.6.14 shows a very simplified structure of the starch synthase using two different active sites as the reaction proceeds through the first steps.

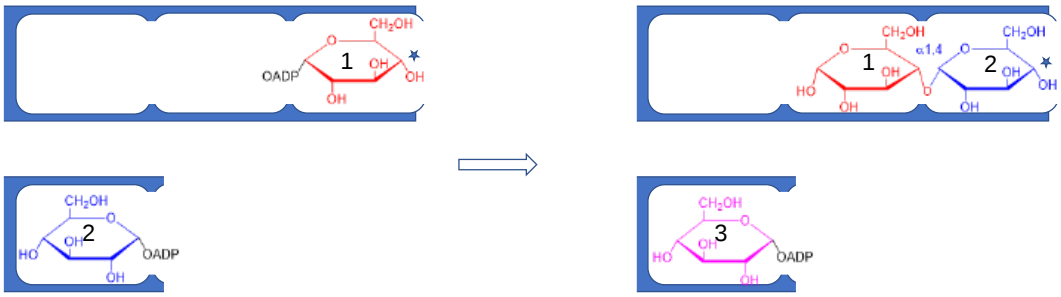


Figure 20.6.14 shows a simplified structure of the starch synthase using two active sites

Note that additional glucose units are added to the nonreducing C4 end denoted by a star. α(1,6) branches are added by branching enzymes as in the case of glycogen. Bacteria starch is made in a fashion similar to glycogen, but they use ADP-glucose as do plant starch synthase. Structural models show that the protein has one active site so the kinetic models suggesting the use of two active sites may refer to a movement of chains between different monomers in oligomeric forms of the protein.

Both the donor (ADP-Glc) and product (starch_{n+1}) have their participating electrophile (ADP) and now substituted nucleophile (C4-O-R) in the α-anomeric form so they are retaining glycosyltransferases. The enzymes could act in two different ways:

- processive mechanism: the acceptor ($starch_n$) stays bound to the enzyme after each addition of the next glucose from the donor ADP-glc, and one chain extends quickly
- distributive mechanism: the enzyme dissociates from the product ($starch_{n+1}$) after the addition of the glucose from ADP-glc, and must rebind to catalyze the next addition, so many new chains start and the growth of each chain is slow

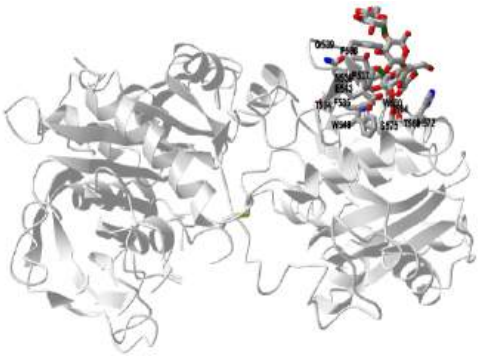
Kinetic evidence suggests that some starch synthases are processive and others are distributive. Along with starch synthase, three other enzymes are involved, ADP-glucose pyrophosphorylase (AGPase), starch branching enzyme (SBE), and starch debranching enzyme (DBE)

Starch synthesis is regulated by gene transcription, phosphorylation, and redox conditions. A key regulatory enzyme is an ADP-glucose pyrophosphorylase (AGPase). This enzyme catalyzes the formation of ADP-glucose (the donor) and PPi from glucose 1-phosphate and ATP. Dithiothreitol, a reducing agent, increases starch synthesis by inactivating AGPase. We will see below that a key disulfide bond is some starch synthases that must be reduced (cleaved) to open an active site cleft between the N-terminal and C-terminal lobes of the catalytic domain. In vivo, thioredoxins are probably involved in redox regulation.

20.6.7: STRUCTURES OF STARCH SYNTHETASES (SS)

The are four classes of soluble starch synthetases (SSI-SSIV) and one starch granule-bound one (GBSS). All have two catalytic domains as noted in the mechanism above except SSII which has 3 CHO binding domains. GBSS appears to form amylose and long chains of amylopectins (amylose with around 5% α(1,6) branching. Loss of function mutants of GBSS have much-reduced amylose concentrations. SSI-SSIII produces most of the amylopectin. SSI is most active with stands (outer ones in branched structures) that have around 7-9 glucoses. In leaves, it generates chains of up to about 10 residues in amylopectin. The SSs can be chloroplastic or amyloplastic.

Figure 20.6.15 shows an interactive iCn3D model of barley starch synthase I in complex with maltooligosaccharide (4HLN).

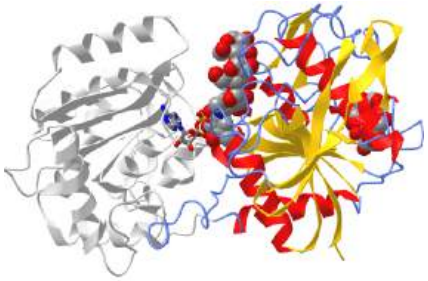


NCBI iCn3D Figure 20.6.15: Barley starch synthase I in complex with maltooligosaccharide (4HLN). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...LGwFUeJAMskBcA

The protein is shown in gray in an open conformation a glucose 5-mer bound on the outside. The protein has a GT-B fold with distinct N- and C-terminal Rossmann-like domains and a central linker. Side chains within 5 Å of oligosaccharide are shown in sticks and labeled. Note especially the disulfide bridge between cysteines 126 and 506 in the central part of the protein that prevents the formation of the active site. This clearly shows the importance of redox signaling to activate the enzyme.

The maltose is not bound in the active site but at a surface secondary binding site (SBS). The role of this site is a bit unclear but may be involved in carbohydrate:carbohydrate interactions. Specifically, they may assist in recruiting starch chains for further elongation. SBSs are found in many but not all starch synthases. Note that in this structure there is only one occluded active site and not two as suggested in Figure 14 above.

Figure 20.6.16 shows an interactive iCn3D model of the catalytic domain of starch synthase IV from Arabidopsis thaliana bound to ADP and acarbose (6GNE).



NCBI iCn3D Figure 20.6.16: Catalytic domain of starch synthase IV from Arabidopsis thaliana bound to ADP and acarbose (6GNE) https://structure.ncbi.nlm.nih.gov/i...MioUaKM73M3Wd9

This structure is just the catalytic domain (representing about half of the total protein sequence). The N-terminal domain is colored by the secondary structure. Acarbose (spacefill, CPK colors) again occupies both the donor and acceptor sites in the active site (central regions). The structure has a secondary binding site (SBS) occupied by the disaccharide maltose. Again there are not two active sites, but the migration of starch chains between monomers in oligomeric forms could support the model shown in Figure 14.

The structure of beta-acarbose, an inhibitor, is shown in Figure 20.6.17.

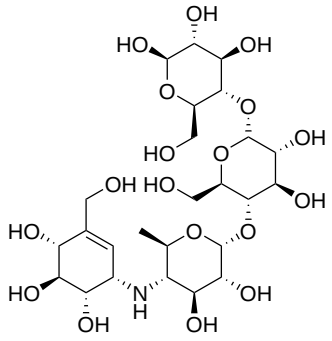
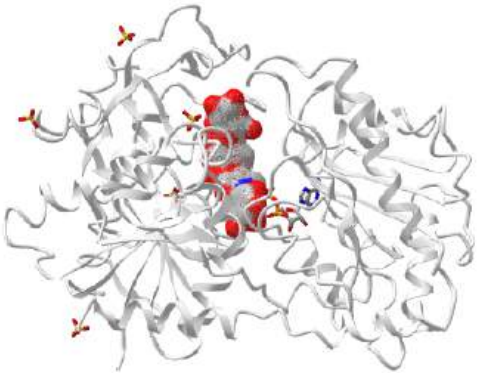


Figure 20.6.17: structure of beta-acarbose

Figure 20.6.x18 below shows an interactive iCn3D model of the Granule Bound Starch Synthase from Cyanobacterium sp. CLg1 bound to acarbose and ADP (6GNF)



NCBI iCn3D Figure 20.6.18: Granule Bound Starch Synthase from Cyanobacterium sp. CLg1 bound to beta-acarbose and ADP (6GNF). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?F1tJBDAh6QRhxpH88

ADP, occupying part of the donor site that would usually bind NTP-Glc, is shown in CPK-colored sticks. The rest of the donor site (where Glc of NTP-Glc would bind) and the acceptor site (growing starch chain) is occupied by beta-acarbose, It is shown in the CPK-colored wire-frame surface. Histidine 181, which probably stabilizes an " oxocarbenium-like anomeric carbon in the transition state", is shown in ball and sticks with CPK colors.

20.6.8: CELLULOSE SYNTHESIS

We can't leave plant carbohydrate metabolism without considering the synthesis of cellulose, which is catalyzed in plants by members of the superfamily cellulose synthase (CesA) and cellulose synthase-like (CsI) enzymes, which are part of the glycosyltransferase GT2 family and have similar structures. Cellulose and hemicellulose are, of course, chief components of the 1^0 and 2^0 cell walls. Members of the CesA family have a conserved motif (DDDQxxRW) as well as a zinc-finger domain. Different members catalyze the synthesis of the 1^0 and 2^0 cell walls. Members of the Csl family are involved in additional cell wall glycans including (1,4)-β-D-mannan (CsIA) and xyloglucan cytoskeleton (CslC). UDP-glucose is the donor in the creation of the β(1,4) acetal linkages between glucose monomers.

Plant growth must respond to environmental triggers through a balance of cell expansion and cell division and a key regulator of these processes is the flexibility of the cell walls which can maintain turgor pressure by expansion. Cells that are nonexpanding (for example those that line the xylem vessels and in woody tissue) have secondary cell walls beneath their primary walls.

We have previously discussed the structure of the primary and secondary cell walls in Chapter 7.3. In brief, the primary cell walls contain cellulose, hemicellulose, and pectins. Cellulose, the main component that provides strength, is synthesized by CesA which forms a very

large rosette-shaped complex (CSC). These complexes seem to move intracellular microtubules which guide that synthase complex through the interactions of microtubule-associated cellulose synthase compartments (MASCs), whose numbers increase during stress. Likewise, there are uncoupling proteins that inhibit microtubule movement from the CSC. These protein complexes stay aligned during cell growth. The cell wall hence is a key player in signal transduction that allows growth and cell division.

The structure of the rosette-shaped complex (CSC) has been determined by cyroEM and is shown in Figure 20.6.19.

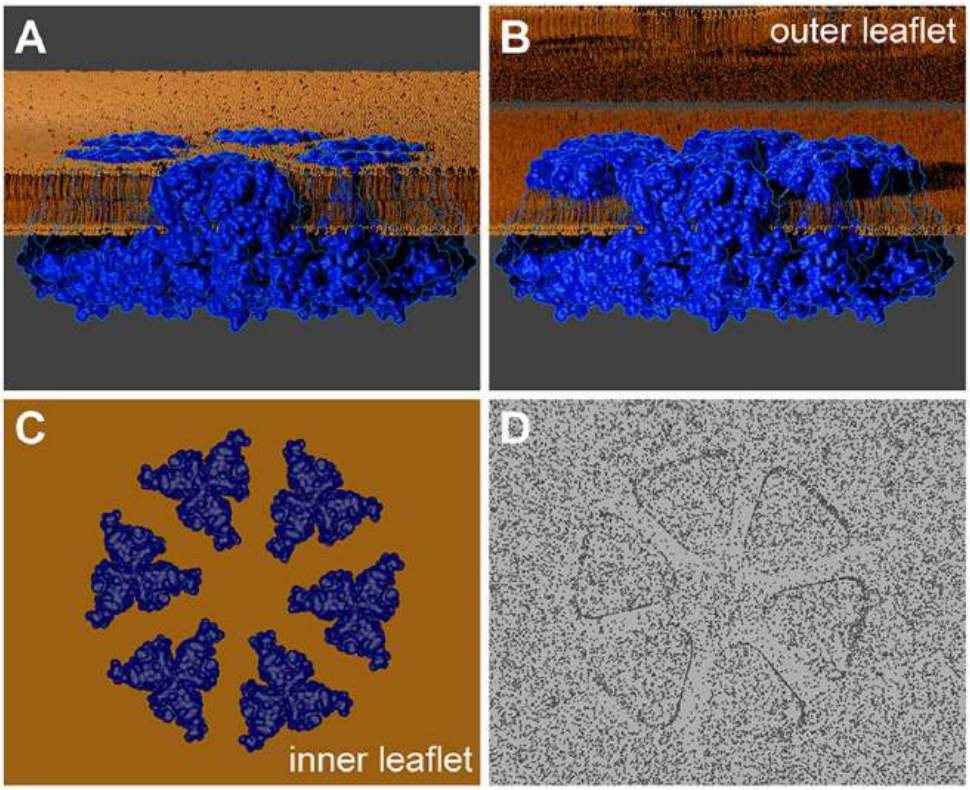


Figure 20.6.19: Structural cartoons of the CESA CSC complex. Nixon, B., Mansouri, K., Singh, A., *et al.* Comparative Structural and Computational Analysis Supports Eighteen Cellulose Synthases in the Plant Cellulose Synthesis Complex. *Sci Rep* **6,** 28696 (2016). https://doi.org/10.1038/srep28696. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/

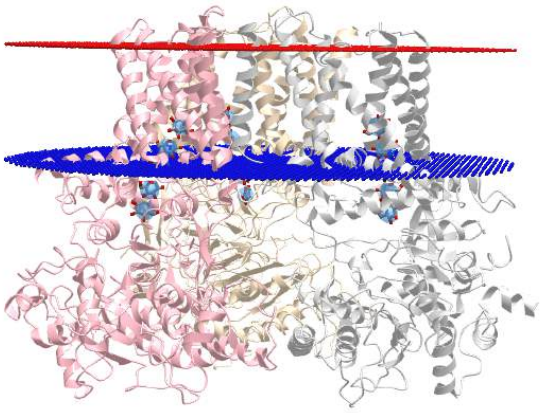
Panel (A) shows the complex as a series of trimers of CESAs, with each monomer spanning the membrane with 7 alpha helices. The catalytic domain is in the cytoplasm.

Panel (B) shows the CSC complex with the top leaflet removed

Panel (C) shows a top-down view showing 6 sets of trimers of CESA.

Panel (D) shows a metal replica viewed in the TEM after the removal of the biological material.

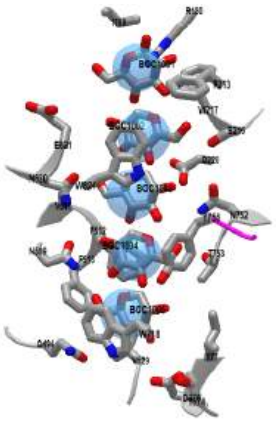
Each trimer synthesizes a cellulose strand. There are 18 trimers in the complex allowing the concomitant synthesis of cellulose strands that can easily self-associate through hydrogen bonding to form near the extracellular surface cellulose protofibrils.Figure 20.6.20 shows an interactive iCn3D model of the catalytically active homotrimeric poplar cellulose synthase (6WLB)



NCBI iCn3D Figure 20.6.20: Catalytically active homotrimeric poplar cellulose synthase (6WLB). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...zCm6fYSwAFA897

Each of the monomers within the trimers is given a different color. A 5 residue β(1,4) glycan is shown in cartoon form emerging into the middle of the membrane complex.

Figure 20.6.21 shows an interactive iCn3D model of the homotrimeric poplar cellulose synthase isoform glycan binding site (6WLB).



NCBI iCn3D Figure 20.6.21: Homotrimeric poplar cellulose synthase isoform glycan binding site (6WLB. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...HqYhTQfuw857PA

The actual transmembrane channels start just above the active site. Key amino acid side chains (Trp 718, Phe 513, Val 529, and Gln 494) help form the portal opening. The actual channel is lined with both aromatic and hydrophilic residues, which supply sufficient but not too strong noncovalent interactions that allow sequential movement of the continually-synthesized cellulose as it ratchets forward toward the extracellular side of the membrane. The aromatic residues interact through pi stacking with the glucose residues and through interactions with equatorial OH groups on the β-glucose polymer. Remember that cellulose is especially stable from a steric perspective since all its OH groups and the acetal linkage are equatorial.

CHAPTER OVERVIEW

21: LIPID BIOSYNTHESIS

21.1: Biosynthesis of Fatty Acids and Eicosanoids
21.2: Biosynthesis of Triacylglycerols
21.3: Biosynthesis of Membrane Glycerolipids
21.4: Biosynthesis of Membrane Sphingolipids
21.5: Biosynthesis of Cholesterol and Steroids
21.6: Biosynthesis of Isoprenoids

21.1: BIOSYNTHESIS OF FATTY ACIDS AND EICOSANOIDS

21.1.1: INTRODUCTION

We present the full fatty acid synthase pathway based on the structure of the yeast fatty acid synthase (FAS) complex, whose full structure is known. Then we will explore each step in more detail. The mammalian FAS complex is a bit different and we will summarize those differences next. We will end with a summary of the bacterial pathway.

Let's look at the net reaction first and then go back to the structure and mechanism from that. Equation 1 below shows the net reaction.

$$8\ \text{Acetyl-CoA}\ (2\text{C}) + 14\text{NADPH} + 14\text{H}^+ + 7\text{ATP} \rightarrow 1\ \text{Palmitate}\ (16:0) + 8\text{CoA} + 6\text{H}_2\text{O} + 14\text{NADP}^+ + 7\text{ADP} + 7\text{Pi} \qquad (21.1.1)$$

Eight 2C-acetyl-CoAs condense to make the 16:0 fatty acid palmitate. That acetyl-CoA is used to synthesize and elongate fatty acids suggests an immediate explanation for the fact that most fatty acids have an even number of carbon atoms. The net reaction is a reductive biosynthesis making one long molecule from multiple short ones. Hence there must be an energy source to drive the reaction (7ATPs) and a reducing agent (typically NADPH). This reaction takes place in the cytosol and uses the nicotinamide-based redox pair (NADPH/$NADP^+$). These factors differentiate fatty acid synthesis from beta-oxidation, which occurs in the mitochondria and which uses the redox pair NAD^+/NADH.

From a chemical perspective, the acetyl-CoAs, high-energy molecules with respect to their hydrolysis products, must be activated using ATP in some way to drive C-C bond formations. That reaction is catalyzed by the enzyme **acetyl-CoA carboxylase**, which uses the ATP already mentioned and the bound cofactor biotin that adds CO_2 to acetyl-CoA to form malonyl-CoA. Neither malonyl-CoA nor CO_2 is shown in the above equation since they are present on both sides of the net reaction and cancel out from the final balanced equation. C-C bond formation occurs on the addition of the growing acyl group with the 3C malonyl with the actual addition reaction driven by the release of CO_2. The malonyl-CoA produced by acetyl-CoA carboxylase feeds now into the fatty acid synthesis cycle.

Only one other enzyme complex is required for the entire reaction, the **fatty acid synthase (FAS) complex**. Don't let that fool you into thinking the mechanism is simple though! There are two types of fatty acid synthase complexes, I and II. Type II FAS is found in plants, most bacteria, and in mitochondria. (Its presence in the mitochondria might seem odd since fatty acid oxidation occurs there. We will discuss that in a bit.) There are multiple separate enzymes to catalyze the synthesis of fatty acids in the Type II FAS systems. We won't discuss those. Instead, we will focus on the **Type I FAS complex** found in some bacteria, fungi like yeast, and higher eukaryotes. The Type I FAS is one very large complex. In mammals, it consists of two α chains with a total molecular weight of 540,000. In yeast, it is an $\alpha_6\beta_6$ heterododecamer with a molecular weight of 2.6 million. In either complex, the alpha and beta chains have multiple different enzyme catalytic domains.

Large enzymatic complexes typically allow bound substrates to proceed to products without dissociation, with the bound intermediates moving through "channels" to the next active site. In the case of Type I FAS, the growing intermediate, which eventually reaches the end product size 16:0 is tethered to an **acyl-carrier protein (ACP)**. The tethered intermediate can "swing" from one active site to another to allow the iterative stepwise chemical reactions to complete before the addition of another 2C acetyl CA to the growing chain. The actual extra 2Cs come from the 3C malonyl-CoA (formed by the extracyclic (outside of the pathway) enzyme acetyl-CoA carboxylase) and immediate release of CO_2 which helps drive the reaction forward.

Now we can introduce the fatty acid synthase cycle and feeder acetyl-CoA carboxylase reaction in its entirely, as shown in Figure 21.1.1 for yeast FAS1.

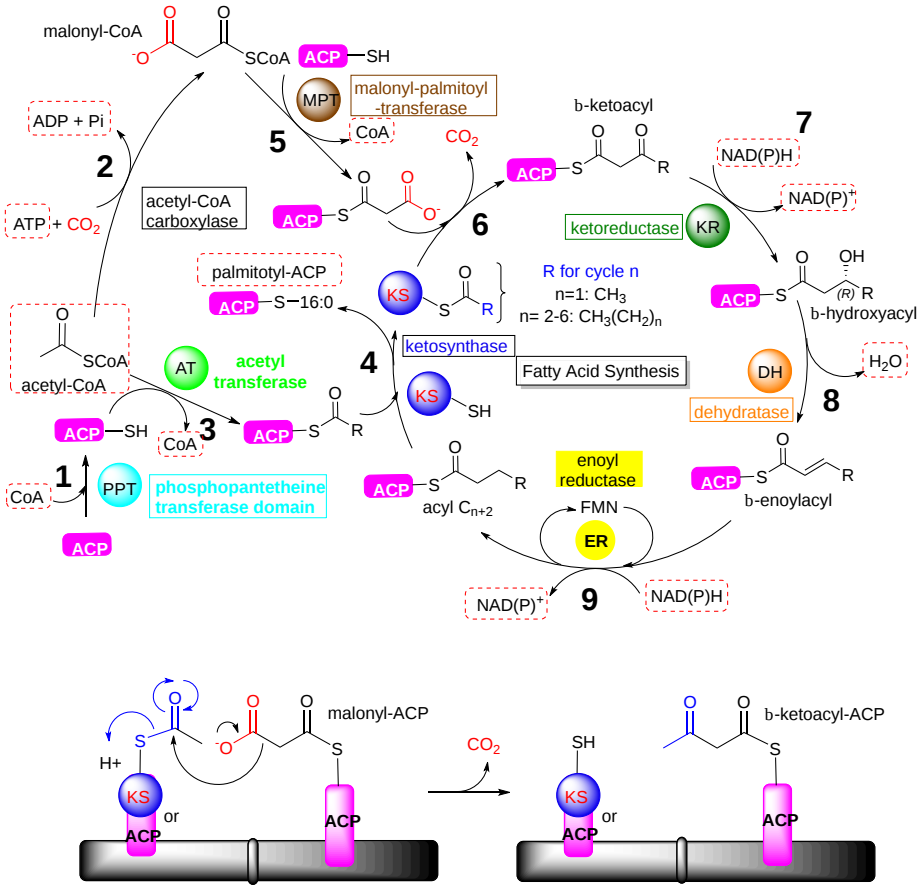


Figure 21.1.1: The Yeast Fatty Acid synthesis cycle with the feeder acetyl-CoA carboxylase reaction. The bottom part of the figure shows a mechanism for the condensation between acetyl-CoA attached to ACP or KS and malonyl-CoA, driven by the release of CO_2. (adapted from Singh et al. Cell, 180 (2020), https://doi.org/10.1016/j.cell.2020.02.034 and Fischer et al. *Sci Rep* **10,** 895 (2020). https://doi.org/10.1038/s41598-020-57418-8. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/

This is a difficult figure to understand but keep in mind that it scales down the complexity of FAS1 enormously. Let's deconstruct the figure to make it more understandable.

The dotted red boxes show the substrates and products that match chemical equation 1. The substrate acetyl-CoA comes into the cycle in two places shown in the extracyclic section to the upper left. It comes in as acetyl-CoA from **reaction 3** after attachment to the acyl carrier protein (part of FAS1). It also comes in "effectively" after it has been carboxylated by acetyl-CoA carboxylase in **reaction 2** to malonyl-CoA, which enters the cycle as malonyl-ACP after **reaction 5**. The other substrates/products include CoAs the NAD(P)H/$NADP^+$ couples and H_2O shown around the right-hand side of the circular fatty acid synthesis cycle (2 -6 o'clock positions) as well as the final product, 16:0-ACP (9:30 clock position), which forms 16:0-SCoA.

The individual enzymatic domains in the $\alpha_6\beta_6$ hetero-dodecamer yeast FAS1 are shown in colored spheres as shown in Figure 21.1.2.

Domain	Chain	Name
ACP	α FAS2	acyl carrier protein
PPT	α FAS2	Phosphopantetheine transferase
AT	β FAS1	acetyl transferase
MPT	β FAS1	malonyl-palmitoyl transferase
KR	α FAS2	keto reductase
DH	β FAS1	dehydratase
ER	β FAS1	Enoyl reductase
KS	α FAS2	ketosynthase

Figure 21.1.2: Enzymatic domains in the $\alpha_6\beta_6$ heterododecamer yeast FAS1

Figure 21.1.3 summarizes pictorially the repetitive Claisen condensation reactions of acetyl-ACP and malonyl-ACP in the first step, and the growing acyl-ACP chain and malonyl-CoA in the next two cycles.

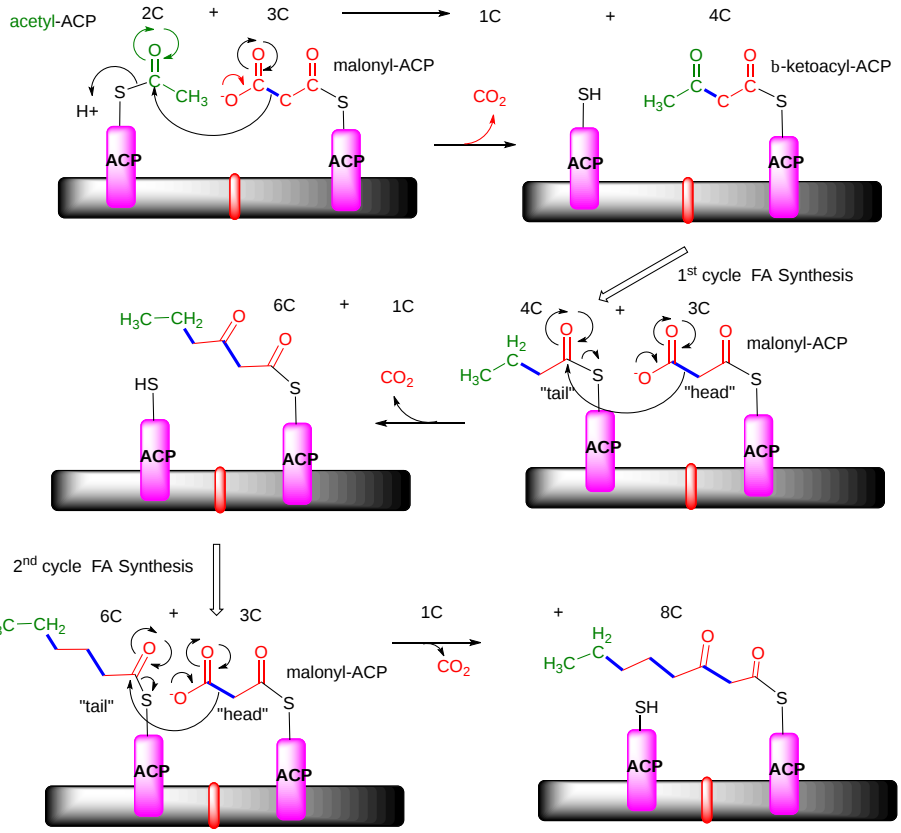


Figure 21.1.3: Pictorial summary of the first three condensation reactions in fatty acid synthesis. Each acyl group is shown connected to ACP for simplicity.

In each step, the condensation (C-C bond formation) is driven by the decarboxylation of malonyl-CoA. (Of course, earlier expenditure of ATP in step 1 is required). Electrons in malonyl-CoA that form a bond with acetyl-ACP in the first condensation step and acyl-ACP in subsequent condensations are shown as **bold blue lines**. Note that the condensation is from the head of the malonyl-ACP to the tail of the elongating acyl chain. The ACP or equivalent chains are in the interior of the complex which allows movement of the acyl groups attached to pantetheine chains to interactively reach nearby catalytic domains for each step in the cycle.

21.1.2: MECHANISM OF INDIVIDUAL REACTIONS

Now we can explore some of the reactions in more detail. Then we will look at the structure of yeast FAS and its domain organization.

21.1.2.1: REACTION 1: PHOSPHOPANTETHEINE TRANSFERASE

Holo-(acyl carrier protein) synthase (AcpS) from *Bacillus subtilis* is a member of the phosphopantetheinyl transferase superfamily. AcpS post-translationally modifies ACP to its holo form to activate it. AcpS catalyzes the transfer of the 4'-phosphopantetheinyl (P-pant) moiety of coenzyme A to a serine residue on the ACP. This gives the activated ACP enzyme and adenosine 3'5'-bisphosphate as products. This process is important as ACP enzymes play important roles in several biosynthetic pathways, such as the synthesis of fatty acids, and vitamins, AcpS is essential in the initiation of the biosynthesis of fatty acids, polyketide antibiotics, and non-ribosomal peptide

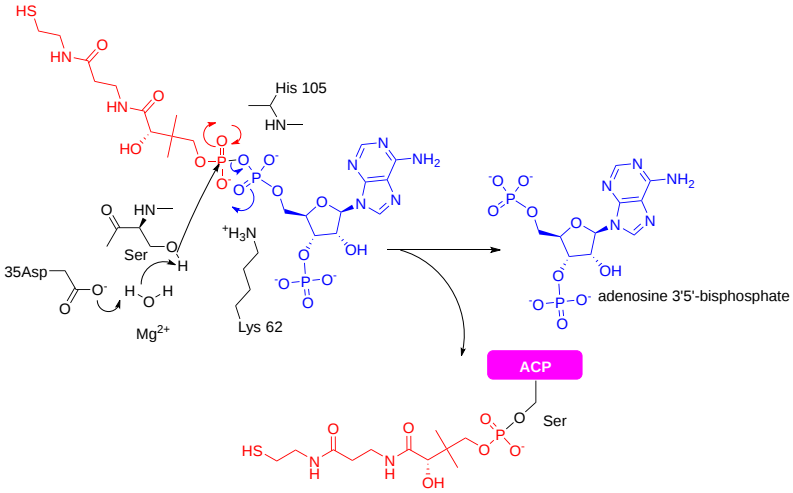


Figure 21.1.4: https://www.ebi.ac.uk/thornton-srv/m-csa/entry/152/

21.1.2.2: REACTION 2: ACETYL-COA CARBOXYLASE (ACC)

This enzyme catalyzes the carboxylation of acetyl-CoA to malonyl-CoA. which then enters the fatty acid synthesis cycle. This is a key enzyme as it is the rate-limiting step, and is regulated. The reaction has two steps:

- carboxylation of biotin on a biotin carboxyl carrier domain of the enzyme at the expense of ATP hydrolysis. The carboxylation, using bicarbonate as a substrate, leads to the formation of a molecule with high energy (with respect to its hydrolysis product), the carboxylated biotin intermediate. (**Remember, there is no such thing as a "high energy" bond.**)
- transfer of the carboxyl group to acetyl-CoA to form malonyl-CoA which requires the formation of a C-C covalent bond.

Biotin, an essential nutrient and cofactor, "carries" activated carboxyl groups for transfer. It is a carboxyl (not acetyl) donor. It is linked to a lysine side chain in the protein through an amide link.

The enzyme is downregulated by 16:0-SCoA, the end product of the pathway for fatty synthesis by FAS, and by phosphorylation by kinase activation through the cAMP pathways from glucagon binding. You don't want to synthesize fatty acids when your energy state is low (signaled by increases in glucagon). Its activity is upregulated by the binding of citrate to an allosteric site. This makes sense since high citrate, a citric acid cycle intermediate, signifies abundant energy reserves are available so fatty acids would be synthesized for future needs. The dephosphorylated form of the enzyme would signify the need to increase fatty acid synthesis.

Figure 21.1.5 shows a few steps in the carboxylation of biotin and the subsequent transfer of the carboxy group to acetyl-CoA, forming malonyl-CoA.

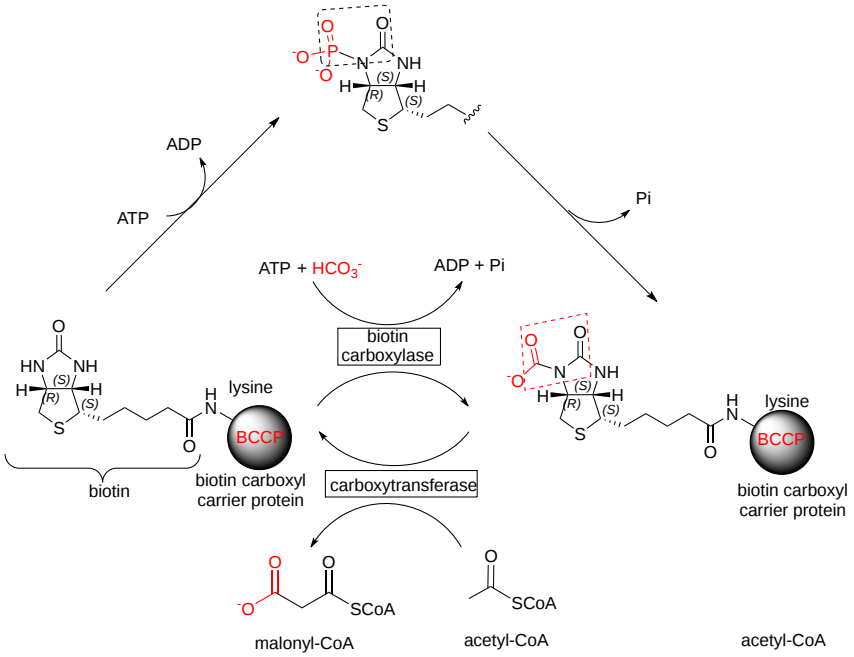


Figure 21.1.5: Simplified reactions of acetyl-CoA carboxylase

BCCP represents the biotin carboxyl carrier domain of acetyl-CoA carboxylase.

Biotin carboxylases are found in many enzymes and pathways, not just fatty acid synthesis. Scientists are trying to devise new pathways and enhanced carboxylases to pull CO_2 from the air into biosynthetic reactions producing fuels, which could be burned. This would release CO_2 back into the atmosphere, in a process that would theoretically but perhaps not practically carbon neutral with respect to greenhouse gas emission.

The abbreviated mechanism shown above shows two of the three activities of acetyl-CoA carboxylase. They are **biotin carboxylase (BC)** and **carboxyltransferase (CT)** activity. The third is the **biotin carboxyl carrier protein (BCCP)** which links biotin to it through a lysine side chain. Different enzymes have different acceptors (acetyl-CoA, pyruvate, etc) of the activated carboxy group so the structure of those centers varies significantly. The BC and BCCP components are structurally similar for different carboxylases.

A mechanism for the E. Coli biotin carboxylase component of the reaction is shown in Figure 21.1.6.

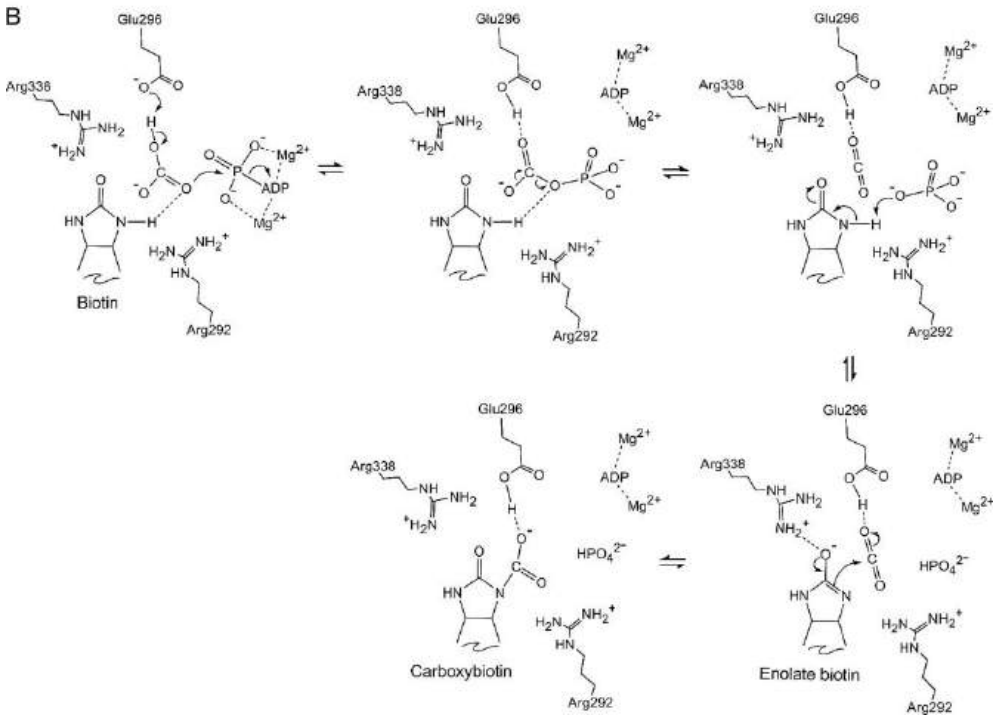


Figure 21.1.6 Reaction mechanism for E. Coli biotin carboxylase. Chou et al. J Biol Chem. 2009 Apr 24; 284(17): 11690–11697. doi:10.1074/jbc.M805783200. Creative Commons Attribution (CC BY 4.0)

Glu 296 acts as a general acid/base while Arg 338 stabilizes negative charge in intermediates and transition states.

Structure of acetyl-CoA carboxylase (ACC)

There are two forms of ACC, soluble cytosolic ACC1 and mitochondrial membrane-associated ACC2 (the latter which regulates beta-oxidation of fatty acids. The overall domain structure of yeast acetyl-CoA carboxylase is in Figure 21.1.7.

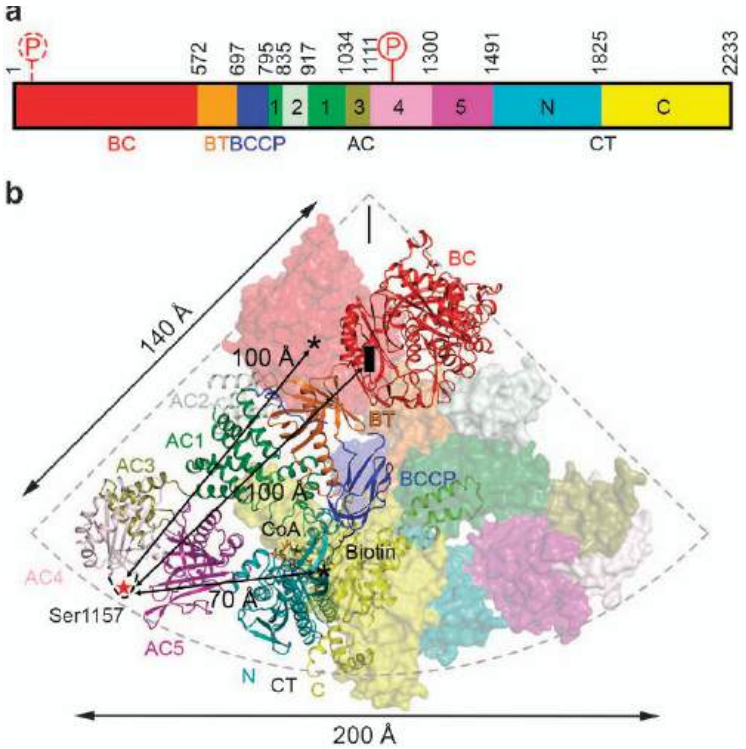


Figure 21.1.7: Overall structure of yeast acetyl-CoA carboxylase (ACC) (ScACC). Wei et al. Cell Discovery (2016) 2, 16044; doi:10.1038/celldisc.2016.44. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Panel (a) shows the domain organization of ScACC. The domains are labeled and given different colors. The five domains of ACC Central (AC1–AC5) are labeled 1–5. The phosphorylation site in the central region is indicated. The phosphorylation site before the biotin carboxylase (BC) domain core is indicated with the dashed lines, as it is absent in ScACC.

Panel (b) shows the structure of the ScACC holoenzyme dimer [7]. One protomer is shown as ribbons, while the other as a surface. The domains in the monomers are colored according to panel (a) and labeled. Ser1157 (red star) is located in a loop missing in the structure (dashed lines), and its distances to the BC and carboxyltransferase (CT) active sites (black asterisks) and the BC dimer interface (black rectangle) in the holoenzyme are indicated.

A chain domain structure 5CSKA_Yeast Acety-lCoA Carboxylase is shown in Figure 21.1.8.

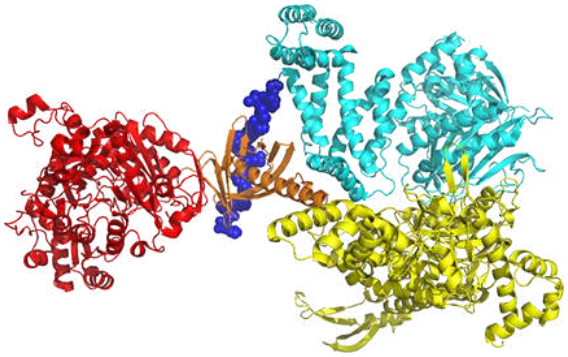


Figure 21.1.8: General domain structure of yeast acetyl-CoA carboxylase A monomer

The colors used to show the domains are matched to Figure 21.1.x above and are as follows:

- Red - BC, biotin carboxylase (has enzymatic activity)
- Gold - BT, an interaction domain
- Blue - BCCP (biotin carboxyl carrier protein (disordered blue spheres)
- Cyan - AC (central bridging noncatalytic domain comprising AC1-AC5).
- Yellow - CT, carboxyl transferase (has enzymatic activity), comprising the N and C subdomains

A similar domain structure is found in human ACC1. The central bridging domain in human ACC1 is called the Central Domain (CD) with four parts, CD_N, CD_L, $CD_{C1,}$ and CD_{C2}.

Regulation of ACC

ACC1 activity is regulated by product inhibition, allosteric effector, and phosphorylation by different kinases. The enzyme is inhibited by its product, malonyl-CoA, and also the end product of fatty acid synthesis, 16:0-CoA.

The protein is inactive as a monomer and is active as a dimer. Anything that can perturb that equilibrium can affect ACC activity. Phosphorylation generally inhibits the protein by promoting the dissociation of the active dimer. The active dimer, in the presence of the allosteric activator, citrate can aggregate to form fibrils, which are even more active than the dimer.

The protein is post-translationally modified by phosphorylation at several sites but a few stand out. AMP-activated protein kinase (AMPK) phosphorylates human ACC at AMP-activated protein kinase at Ser 80 (in the BC domain). In yeast, the kinase SNF1 (equivalent to human AMPK) does not phosphorylate Ser 80 but it phosphorylates a key Ser 1157 in the AC subdomain AC4 in yeast.

- In humans, phosphorylation at Ser 80 in the BC domain prevents dimerization of the protein, through the BC domain and hence inhibits activity;
- In yeast, phosphorylation st Ser 1157 in AC subdomain AC4 in yeast produces a conformational change in the BC domain region that appears to lead to dimer dissociation.

Human ACC is phosphorylated at a multitude of sites with Ser 80 and 1201 (which is phosphorylated both by AMPK and PKA) being the most important in regulation. In addition, it is phosphorylated at Ser 1263 by a cyclin-dependent protein kinase (CDK). When phosphorylated, the pSer 1263 facilitates the binding of BRAC1, the tumor suppressor protein which when mutated can dramatically increase breast cancer. The binding of BRAC1 leads to an inactive fibril form of ACC. Figure 21.1.9 summarizes the regulation of human CoA carboxylase.

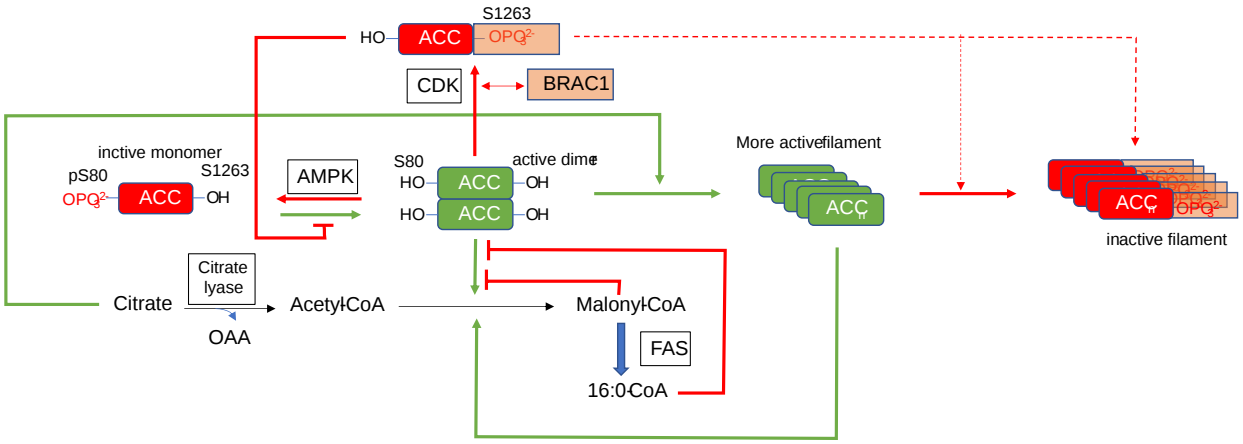


Figure 21.1.9: Regulation of human acetyl-CoA carboxylase. Red indicates inactive states or inactivating reactions, while green indicates active states and activating reactions.

21.1.2.3: 5. MALONYL-PALMITOYL TRANSFERASE (MPT)/MALONYL-COA-ACYL CARRIER PROTEIN TRANSACYLASE (MAT)

These protein activities are found in the yeast (MPT) and human (MAT) domains. The mechanism is shown in Figure 21.1.10.

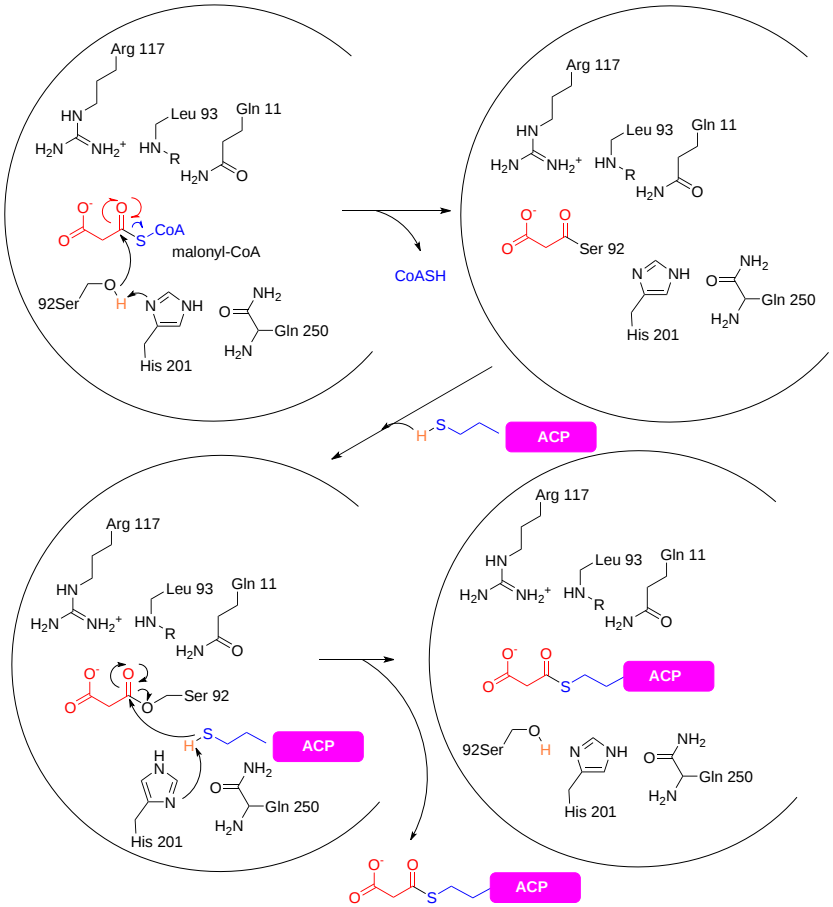


Figure 21.1.10: Reaction mechanism for Malonyl-Palmitoyl Transferase (MPT)/Malonyl-CoA-acyl carrier protein transacylase (MAT). (after https://www.ebi.ac.uk/thornton-srv/m-csa/entry/291/, Creative Commons Attribution 4.0 International (CC BY 4.0) License.

Two key catalytic residues, Ser 92 and His 201, are involved in an acylation/deacylation of the catalytic serine.

The mechanism is shown in Figure 21.1.11.

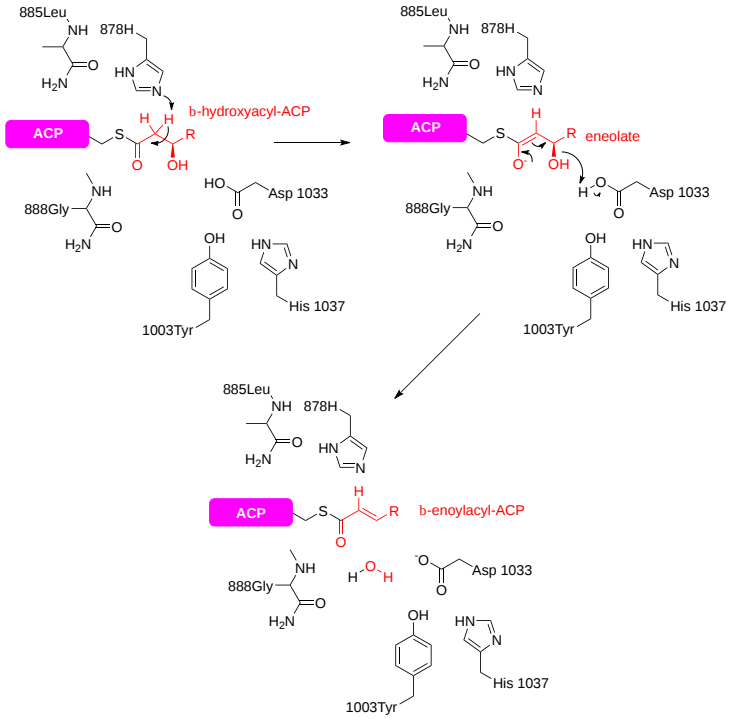
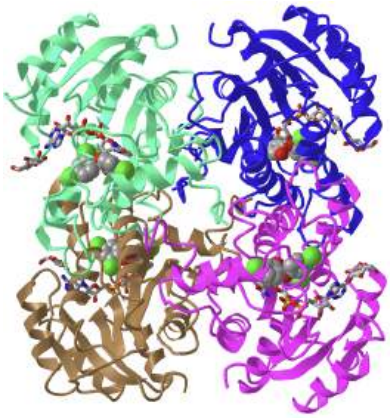


Figure 21.1.11: Reaction mechanism for β-ketodehydratase (after https://www.ebi.ac.uk/thornton-srv/m-csa/entry/972/, Creative Commons Attribution 4.0 International (CC BY 4.0) License.

21.1.2.4: REACTION 9. ENOYL REDUCTASE

This is the last reaction in the fatty acid synthase cycle, which of course repeats until a 16:0-SCoA is made. The yeast enzyme is different in that it has a tightly-bound FMN which is involved in electron transfer (reduction) along with NADPH. The mechanism is likely ping-pong with $NADP^+$ released before the enoyl acyl-CoA binds.

Figure 21.1.12 shows an interactive iCn3D model of the enoyl-acyl carrier protein reductase (ER) in complex with NAD^+ and triclosan (1QSG)



NCBI iCn3D Figure 21.1.12: enoyl-acyl carrier protein reductase (ER) in complex with NAD^+ and triclosan (1QSG) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...Z2ZCQaTYabQWa9

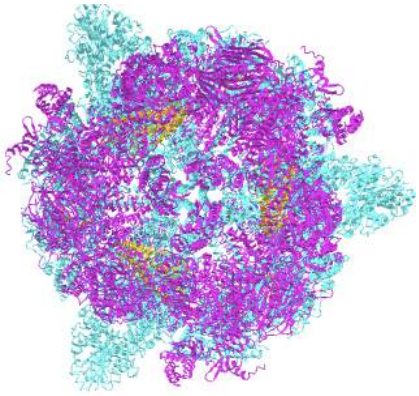
Triclosan is used in many commercial products, even soaps, as it has broad-spectrum antimicrobial activity. In a world with increasing antibiotic resistance, it is unwise to use triclosan in soaps since the evidence shows that the surfactant properties of soaps are sufficient to remove bacteria from the skin. Triclosan is an active site inhibitor of bacterial enoyl-reductases but in humans, it appears to be an allosteric inhibitor, since it is bound at a protein-protein interface and not the active site.

21.1.3: STRUCTURE OF YEAST FATTY ACID SYNTHASE I (FAS1)

Now you have enough background to explore the actual structure of fatty acid synthase. Animal FAS1 is an α_2 homodimer, some bacterial FASs are α_6 hexamers, while fungal FASs are $\alpha_6\beta_6$ dodecamers. In both cases, the monomers are multi-domain proteins with each domain

having a different catalytic activity. There is another type of fatty acid synthase II (FAS2), that consists of separate enzymes each with their own catalytic function. Type II FAS around found in plants, most bacteria, and in mitochondria.

We'll explore the "more interesting" Type I and start with yeast (Saccharomyces cerevisiae) FAS (6QL5), whose structure has been solved with a regulatory γ subunit that 2.8 angstrom 6QL5. Figure 21.1.13 shows an interactive iCn3D model of the fatty acid synthase complex with bound gamma subunit from Saccharomyces cerevisiae (6QL5) very long load time.



NCBI iCn3D Figure 21.1.13: Fatty acid synthase complex with bound gamma subunit from Saccharomyces cerevisiae 6QL5. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...YnMNSaTbTXYGe7 (very long load time).

The alpha subunits are shown in cyan and the beta subunits are shown in magenta. The regulatory γ subunit is shown in orange.

The 6 alpha subunits are arranged in an equatorial wheel or disk. The beta subunits are arranged in trimer top and bottom "domes" that cover the alpha subunit disk. Inside are 3 spaces or chambers where the reactions occur. 5 openings allow outside substrate access and six openings allow internal metabolite access. The FAS is more porous than in large structures with encapsulated volumes like the proteasome or viral capsids. It is more similar to pyruvate dehydrogenase. The ACP protein domain is found inside the structure and is tethered and "swingable" to allow the transfer internal transfer of acyl intermediates to the different enzymatic functional domains of the complex.

Figure 21.1.14 shows animated images of yeast fatty acid synthase complex, which should give you a sense of the arrangement of the equatorial alpha-helical wheel, the two beta subunit domes, and the interior and interior volumes available for reaction chemistry and intermediate movement. They switch between the full structure, followed by separate views of the alpha (cyan), beta (magenta), and regulatory (orange) chains.

fatty acid synthase complex - side view	fatty acid synthase complex - top view

Figure 21.1.14: Alpha chains (cyan), beta chains (magenta), regulatory (orange)

The α and β monomers are multi-domain proteins with different domains having different catalytic properties. Just color coding the monomers in just one color doesn't give insight into the amazing catalytic properties of each chain. Figure 21.1.15 shows colored-coded domain and protein structures for yeast $\alpha_6\beta_6$ dodecamer. The color-coding for the domains for each subunit shown at the bottom of panel A is also used for the domain structures in the actual protein structures.

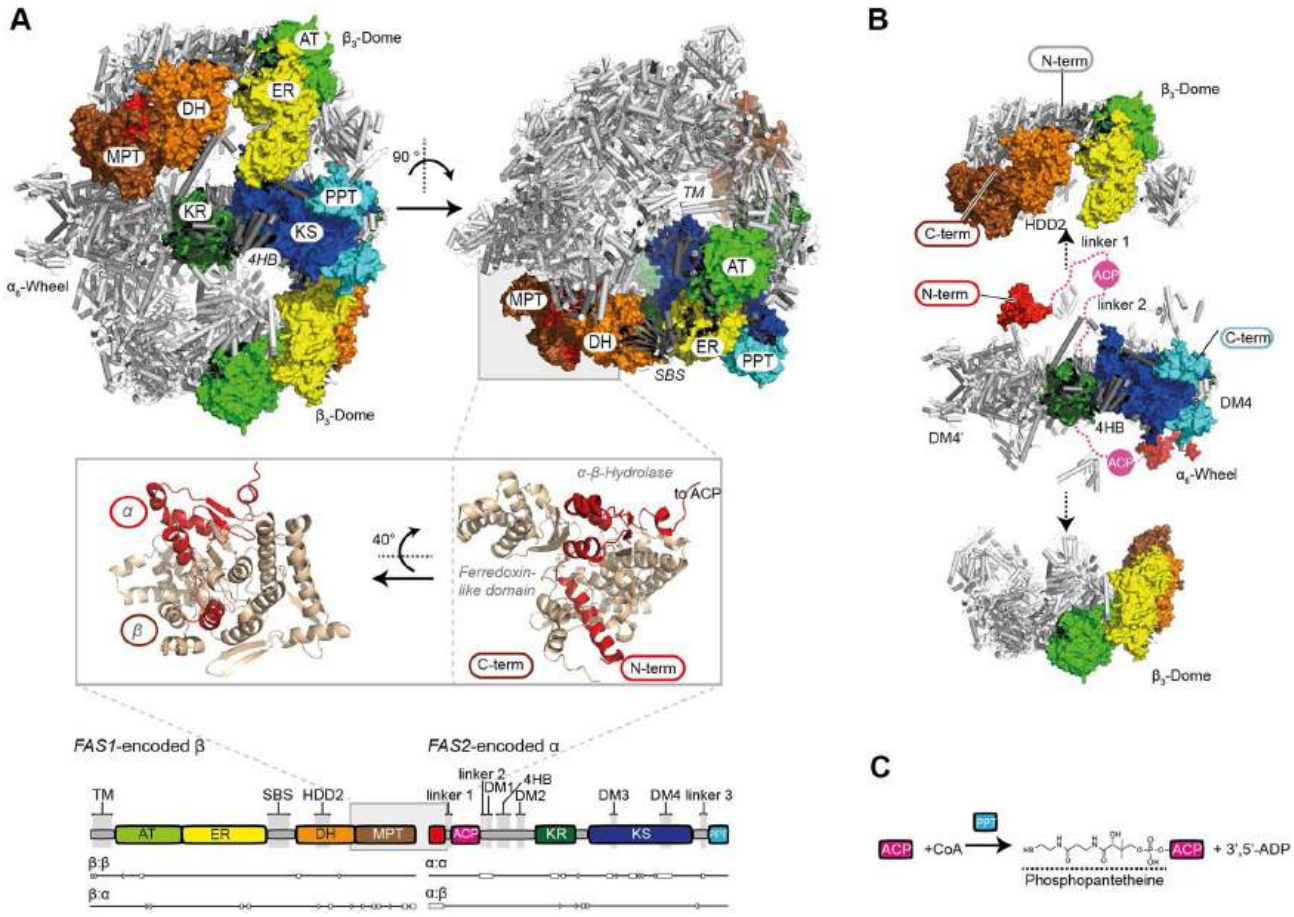


Figure 21.1.15: Structure of yeast FAS.

Panel (**A**) shows the structure of *S. cerevisiae* FAS (3hmj). The side (left) and top view (right) with two β-subunits and two α-subunits are shown. Note that ACP is in the FAS interior, but is not shown. The MPT fold is comprised of both subunits and shown in cartoon representation (β part in brown and its α part in red). Nomenclature: acetyltransferase (AT), enoyl reductase (ER), dehydratase (DH), malonyl-palmitoyl-transferase (MPT), an acyl carrier protein (ACP), ketoacyl reductase (KR), ketoacyl synthase (KS) and phosphopantetheine transferase domain (PPT). Insertion elements are highlighted in grey; trimerization module (TM), 6-stranded β-sheet (SBS), hotdog-domain 2 (HDD2), dimerization module 1–4 (DM1-4), 4-helical bundle (4HB)). Please note that DM2 is not visible in this structure.

Panel(**B**) shows three yeast FAS barrels as a central D3-symmetric α hexamer (α_6-wheel) and two C3-symmetric β trimers (β_3-domes). β_3-domes have been shifted for clarity (see arrows). ACP domains are shown for two α-subunits and are modeled by spheres in magenta. ACP linkers are indicated by dashed lines.

Panel (**C**) shows the post-translational modification of ACP. For phosphopantetheinylation, ACP and PPT have to physically interact. Fischer, M., Joppe, M., Mulinacci, B. *et al.* Analysis of the co-translational assembly of the fungal fatty acid synthase (FAS). *Sci Rep* **10,** 895 (2020). https://doi.org/10.1038/s41598-020-57418-8. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/

In Figure 21.1.16 we connect the domain structures of each monomer to the actual catalytic cycle of yeast FAS.

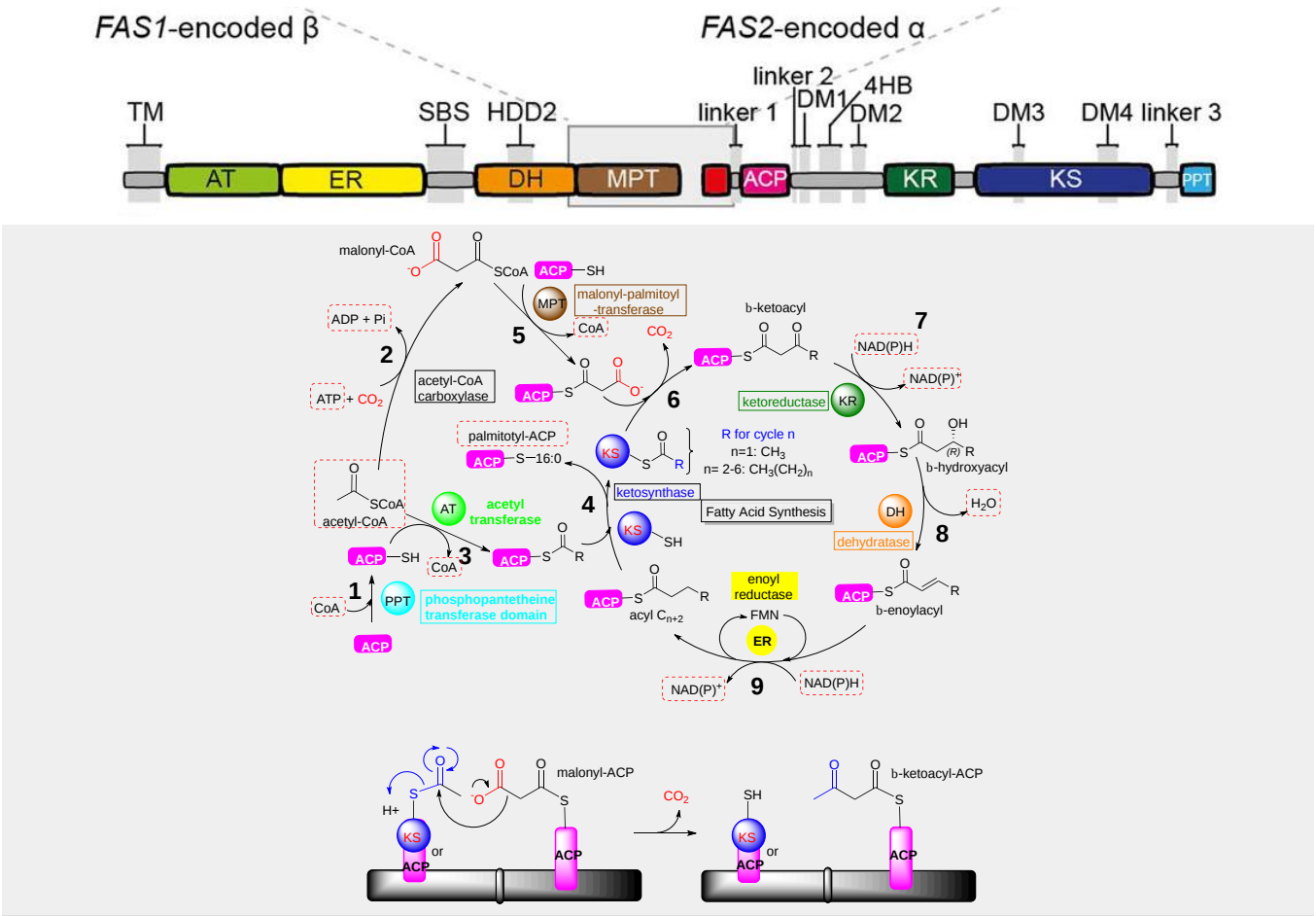
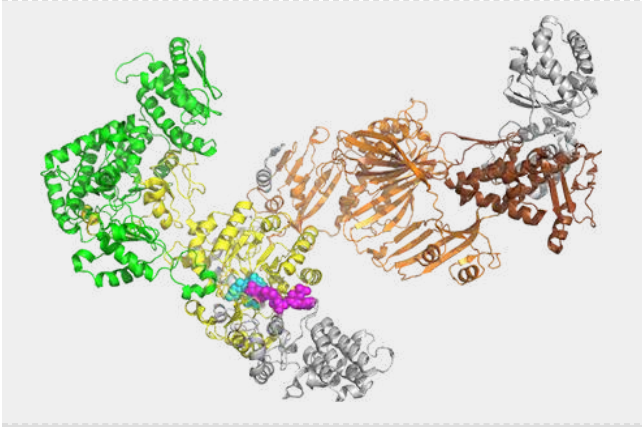


Figure 21.1.16: Domain structure and the catalytic cycle of yeast fatty acid synthase (adapted from Singh et al. Cell, 180 (2020), https://doi.org/10.1016/j.cell.2020.02.034 and Fischer et al. *Sci Rep* **10,** 895 (2020). https://doi.org/10.1038/s41598-020-57418-8. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/

Figure 21.1.17 shows the actual structure of the isolated α and β monomers. The color coding of the domains is the same as used above.

Fatty acid synthase 1 (FAS1) beta subunit (FAS1) (6U5U G chain)



Fatty acid synthase 2 (FAS2) alpha subunit (2UV8A - P19097 AlphaFold)

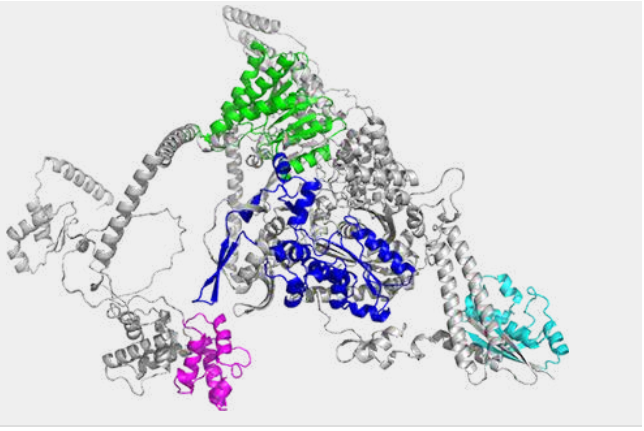


Figure 21.1.17 shows the approximate domain structures and locations for the yeast **fatty acid synthase 1 (FAS1) beta** subunit (6U5U G chain) and fatty acid synthase 2 (FAS2) alpha subunit**.**

The color coding for the domains in each subunit is as follows: FAS1 beta subunit: AT, ER, DH, MPT, flavin mononucleotide (FMN) spacefill cyan NADP nicotinamide-adenine dinucleotide phosphate (NAP), spacefill magenta; FAS2 alpha subunit: ACP, KR, KS, PPT

Now, let's zoom in to see the actual interactions of bound 4'-phosphopantetheine and flavin mononucleotide (FMN). Phosphopantetheine is bound between the alpha (cyan) and beta (magenta) subunits as shown in Figure 21.1.18.

4'-phosphopantetheine between alpha (cyan) and beta (magenta) subunits

flavin mononucleotide (FMN) within beta (FAS1) subunit

Figure 21.1.18:

The γ subunit crosses the entire inner cavity of FAS interfering with the activities of the reductases.

A central difference in the FAS1 found in humans and other animals occurs in the first reaction in which the acetyl group of acetyl-CoA is transferred to the pantetheine sulfhydryl of the acyl carrier protein (ACP) domain catalyzed by the malonyl-/acetyltransferase (MAT) domain. It is then transferred to the active site cysteine of the β-ketoacyl synthase (KS) domain. This allows the malonyl group to be transferred to the now free ACP domain for a second transfer reaction. A Claisen condensation reaction driven by the decarboxylation of the malonyl-ACP occurs to form the β-ketoacyl intermediate bond to the KS domain. An alternative view of the mammalian fatty acid cycle is shown in Figure 21.1.19.

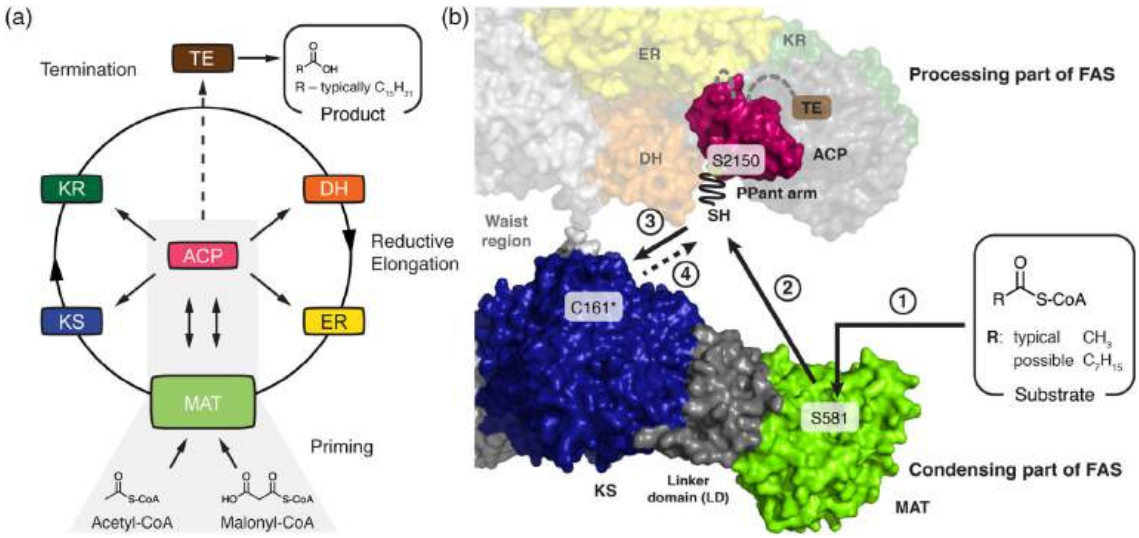


Figure 21.1.19: Mammalian fatty acid synthase cycle. Rittner et al. Protein Sci. 2020 Feb; 29(2): 589–605. doi:10.1002/pro.3797. Creative Commons Attribution License. http://creativecommons.org/licenses/by/4.0/

Panel (a) shows the priming of animal fatty acid synthesis. Panel (b) shows how in the first step, the substrate is selected by the MAT domain and transferred to the ACP domain (Step 2) from where it is passed on to the KS domain (Step 3). Important active site residues are highlighted and C161 is marked with an asterisk. The porcine FAS structure (2png) is shown. Domains of one protomer of FAS homodimer are colored. ACP, acyl carrier protein; DH, dehydratase; ER, enoyl reductase; FAS, fatty acid synthase; KR, ketoreductase; KS, ketosynthase; MAT, malonyl-/acetyltransferase; PPant arm, expand fully in the figure; TE, thioesterase

Figure 21.1.20 shows the yeast FAS cycle (as a proxy for the mammalian cycle) placed in context with the mitochondrial matrix pathways that feed acetyl-Coa and NADPH into the fatty acid synthesis pathway.

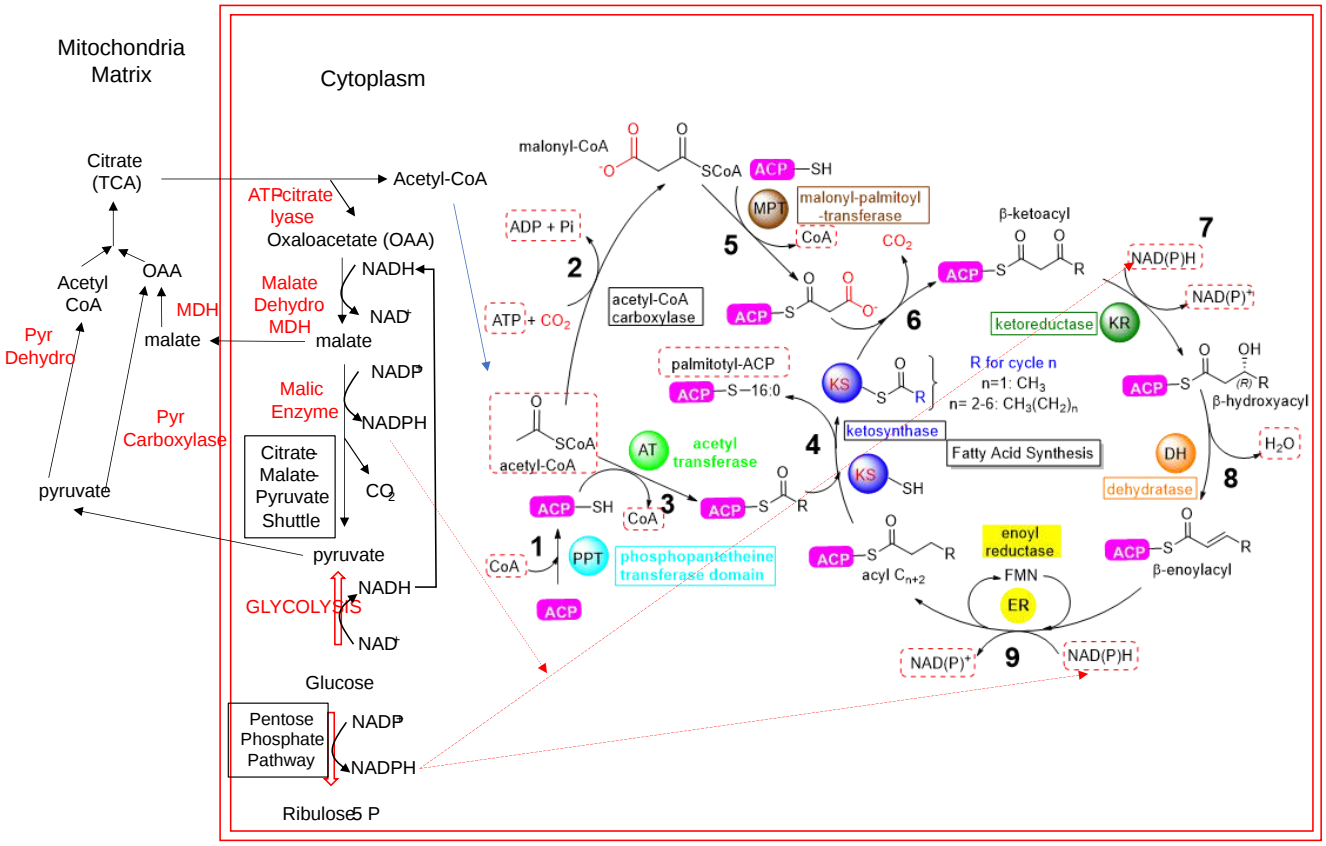


Figure 21.1.20: yeast FAS cycle (as a proxy for the mammalian cycle) placed in context with the mitochondrial matrix pathways that acetyl-Coa and NADPH into fatty acid synthesis

21.1.4: FATTY ACID ELONGATION AND DESATURATION

21.1.4.1: OVERVIEW

Elongation of fatty acids occurs in the cytoplasm of mammals from 16:0 made through fatty acid synthase. As with the synthesis of 16:0 by FAS, the elongation consists of two carbon addition driven by the decarboxylation of a malonyl-CoA substrate. Elongation is carried out by a family of **ELOngation of Very Long-chain fatty acid enzymes (ELOVLs)**. Desaturation is catalyzed by **Stearoyl-CoA Desaturases (SCD1 and SCD2**) and **Fatty Acid Desaturases (FADS1 and FADS2)**.

Figure 21.1.21 shows an overview of de novo fatty acid synthesis with elongation and desaturation in mammals as well as elongation and desaturation of dietary fatty acids.

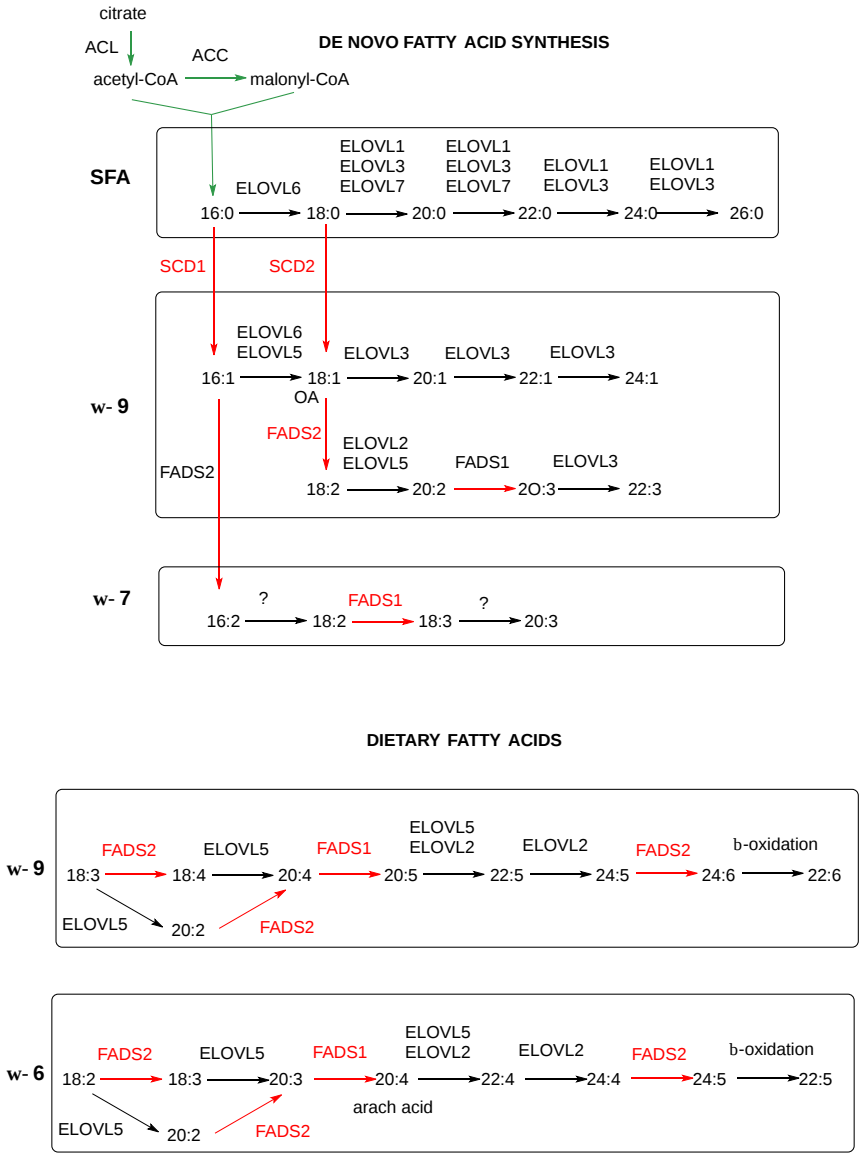


Figure 21.1.21: Scheme of FA synthesis in mammals. After *Corrales et al. Int. J. Mol. Sci.* **2021**, *22*(15), 8159; https://doi.org/10.3390/ijms22158159. Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/)

FAs are elongated (ELOVL1–7) and/or desaturated (SCD, FADS) to generate complex FAs. Long-chain saturated FAs (LCSFA) and unsaturated FAs of ω9 and ω7 can be synthesized from palmitic acid (PA, C16:0) produced by the *de novo* FA synthesis. Long-chain unsaturated FAs of the ω6 and ω3 series can only be synthesized from essential diet-derived FAs (OA, oleic acid; NA, nervonic acid; ALA, α-linolenic acid; DHA, docosahexaenoic acids; LA, linoleic acid; ARA, arachidonic acid). SFA are saturated fatty acids.

21.1.4.2: ELONGATION

Mammals use endoplasmic reticulum very long-chain fatty acid enzymes (ELOVLs). There appear to be seven in mammals with different substrate specificities as shown in the figure above. The nomenclature for fatty acids is a bit strange. Those with 11-20 carbons are called long-chain FAs (LCFAs) and those with more than 20 carbons are very long-chain FAs (VLCFAs). Those with more than 26 carbons are called ultra long-chain FAs (ULCFAs), which are found in the skin, retina, meibomian gland, testis, and brain. The same four steps (condensation, reduction, dehydration, and reduction) used in FAS are used in elongation. The enzymes are abbreviated a bit differently (3-ketoacyl-CoA reductase or KAR, 3-hydroxyacyl-CoA dehydratase or HACD, *trans*-2-enoyl-CoA reductase or TER and LCFA). The enzyme is separate and not part of a multifunction complex like FAS. The elongation cycle is shown in Figure 21.1.22.

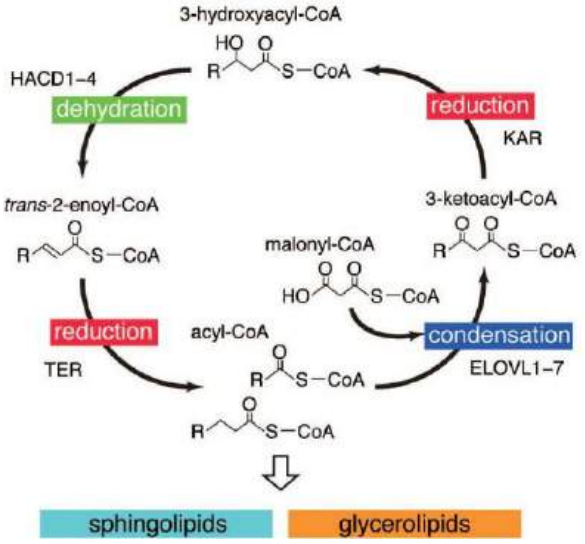


Figure 21.1.22 Mammalian FA elongation cycle. In each cycle, acyl-CoA incorporates two carbon units from malonyl-CoA. Sassa et al. Biomol Ther (Seoul). 2014. http://dx.doi.org/10.4062/biomolther.2014.017. Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/)

Figure 21.1.23 shows fatty acid composition in sphingomyelins which typically have long fatty acids.

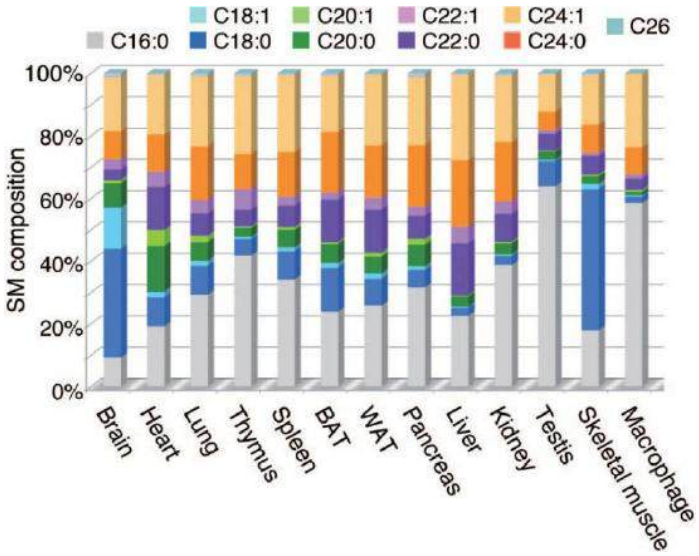


Figure 21.1.23 FA compositions of SM. FA compositions of SM in indicated mouse tissues determined by liquid chromatography-mass spectrometry analysis are illustrated. SM: sphingomyelin; BAT: brown adipose tissue; WAT: white adipose tissue

21.1.4.3: MECHANISM OF ELOVL7

ELOVL condensing enzymes have a long hydrophobic tunnel and an active site nucleophilic histidine. More traditional nucleophiles like serine or cysteine were eliminated as potential candidates for the active site nucleophile since none were in proximity. The active site with a key HxxHH motif lies deep in the membrane. Nucleophilic catalysis leads to an N-acyl intermediate and is consistent with the observed ping-pong kinetics. An alternative reaction for a "bisubstrate" reaction is the formation of a ternary complex with both substrates. The narrow size of the binding pockets precludes this mechanism, so one product, after it is formed, must dissociate before the next substrate can bind. ELOVL7 preferentially catalyzes the elongation of C18-CoA presumably arising from the size of the acyl-binding product.

A plausible mechanism for human long-chain fatty acid enzyme 7 (ELOVL7) elongation is shown in Figure 21.1.24 and Figure 21.1.25. Figure 21.1.24 shows the beginning acylation step of the reaction.

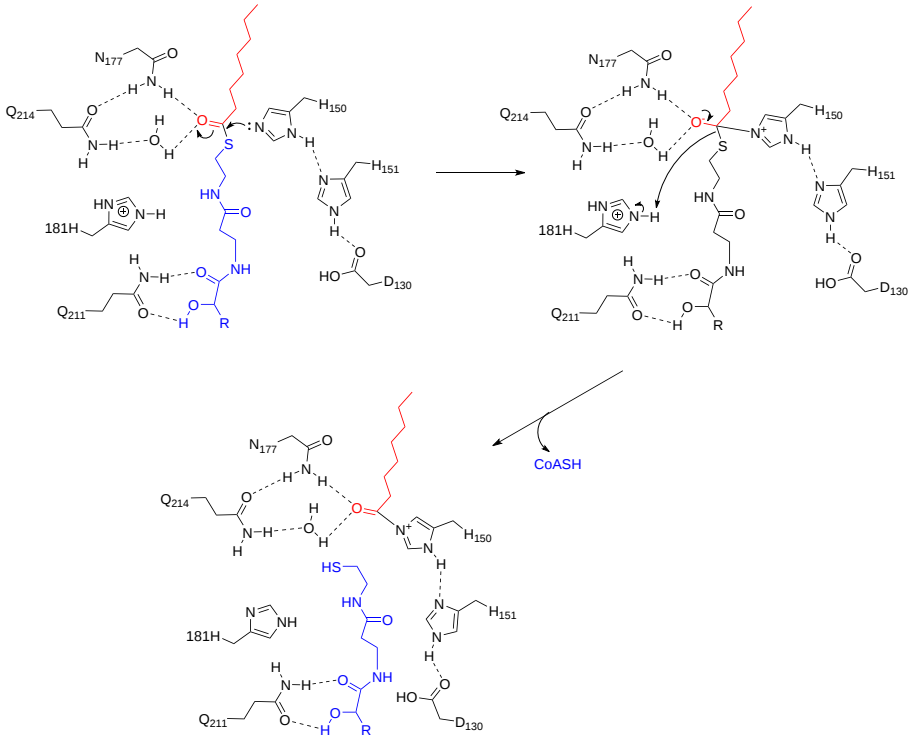


Figure 21.1.24: Acylation step in the human long-chain fatty acid enzyme 7 (ELOVL7) elongation reaction

Figure 21.1.25 shows the condensation (with the elimination of CO_2), followed by the reduction and dehydration of human ELOVL7.

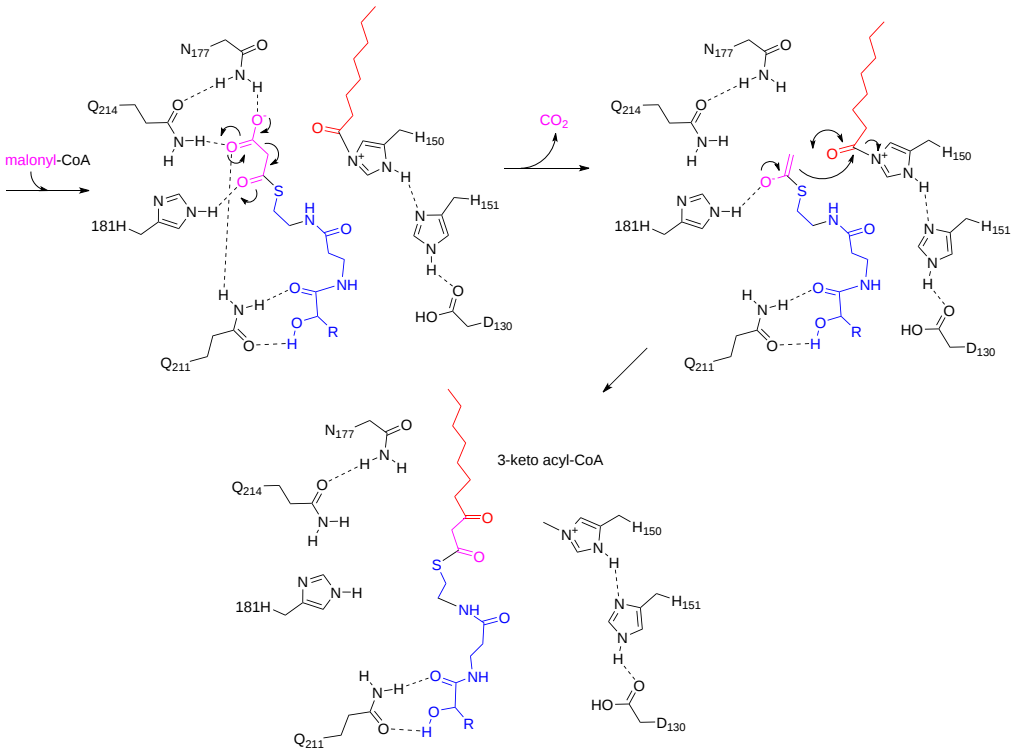
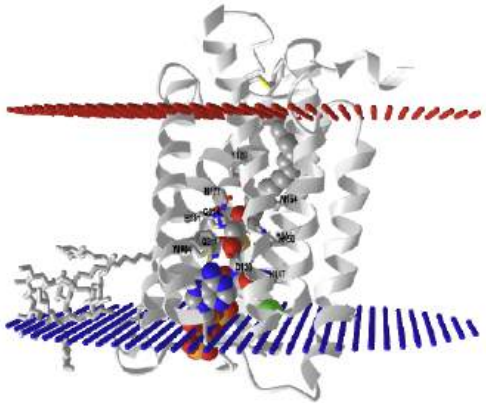


Figure 21.1.25: Part 2 - Plausible mechanism for mammalian fatty acid elongase 7 (ELOVL7) - condensation (with the elimination of CO_2), followed by reduction and dehydration. (adapted from Nie et al. *Nat Struct Mol Biol* **28,** 512–520 (2021). https://doi.org/10.1038/s41594-021-00605-6)

Figure 21.1.26 shows an interactive iCn3D model of the human ELOVL fatty acid elongase 7 (ELOVL7) with bound 3-keto eicosanoyl CoA (6Y7F).



NCBI iCn3D Figure 21.1.26: Human ELOVL fatty acid elongase 7 (ELOVL7) with bound 3-keto eicosanoyl CoA (6Y7F). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...eajMeiRgfHd9s9

The bound 3-keto eicosanoyl CoA is shown in spacefill, CPK colors. Key amino acids involved in catalysis are shown in sticks, CPK colors, and labeled.

21.1.4.4: DESATURATION

The introduction of double bonds requires a different set of enzymes, acyl-coenzyme A (CoA) desaturases. Mammals have desaturases that can produce double bonds at Δ^9, Δ^6, and Δ^5. There are two types, stearoyl-CoA desaturases (SCDs), which introduce a double bond at C9 of saturated fatty acid, and fatty acid desaturases (FADS), which work on unsaturated fatty acids. These are shown in Figure 21.1.20. Plants, but not mammals, have Δ^{12} and Δ^{15}-desaturases, so they, but not mammals, can synthesize ω6 and ω3 fatty acids. These must be supplied by the diet as essential fatty acids and are precursors for the synthesis of longer fatty acids like arachidonic acid (20:4$^{\Delta 5,8,11,14}$), necessary for prostaglandin synthesis (see below) and docosahexaenoic acids (DHA). A plausible mechanism for the stearoyl-CoA desaturase from castor seeds is shown below in Figures 27-30. The mechanisms and explanations are adapted from https://www.ebi.ac.uk/thornton-srv/m-csa/entry/136/. Creative Commons Attribution 4.0 International (CC BY 4.0) License

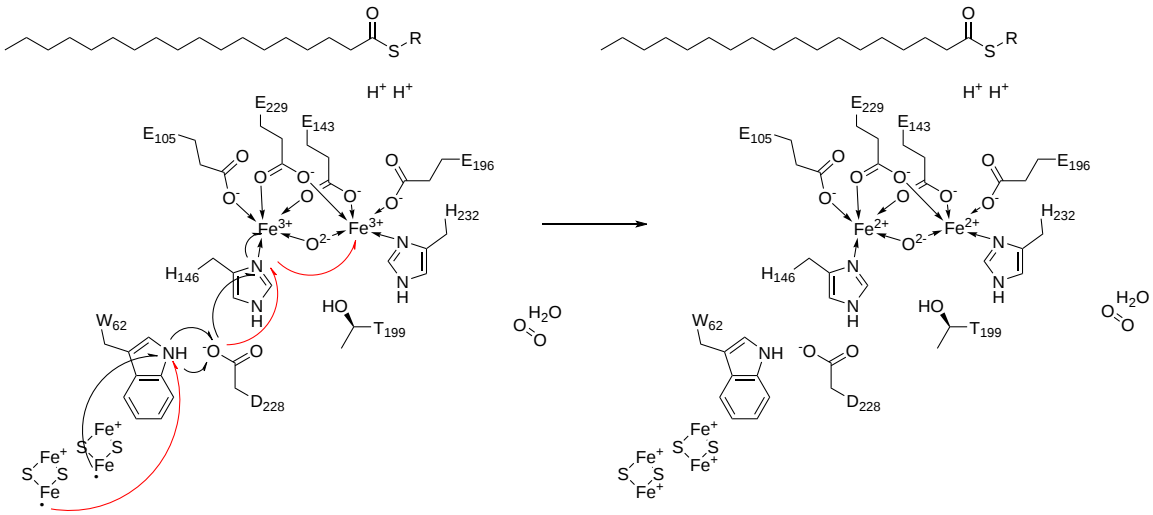


Figure 21.1.27: Part 1 of the stearoyl-CoA desaturase

At the start of the reaction, ferredoxin, a small iron-sulfur protein donates a single electron, through Trp62, Asp228, and His146, to one of the Fe(III) centers in the desaturase. Ferredoxin then donates a second single electron, through Trp62, Asp228, and His146, to the second Fe(III) center. Both of these are shown in Figure 21.1.27

Part 2 of the reaction is shown in Figure 21.1.28.

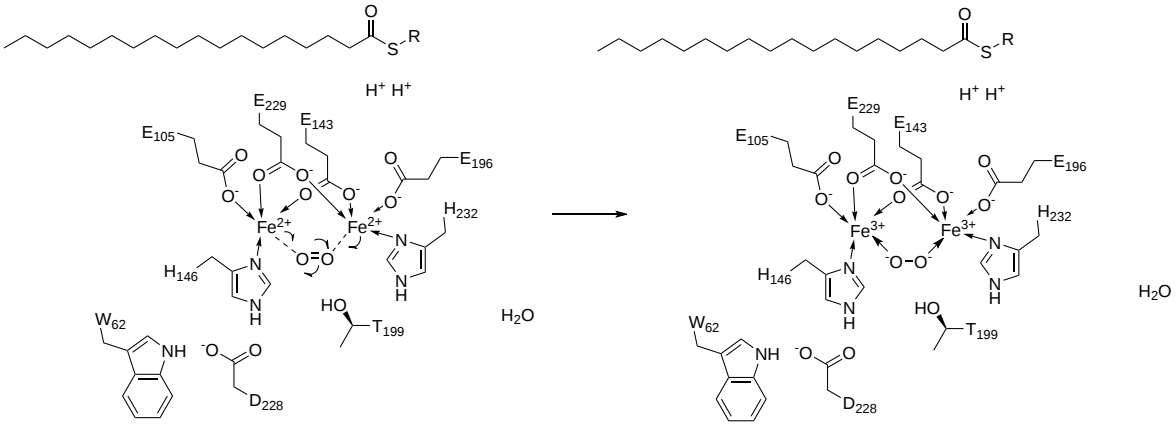


Figure 21.1.28: Part 2 of the stearoyl-CoA desaturase

The reduced charges on the Fe ions lead to loss of the oxide bridging ion (O^{2-}) and its replacement with dioxygen. Water coordinates to one of the Fe(II)s, which causes it to donate a single electron to the dioxide molecule. This starts the first of two homolytic additions of the dioxygen molecule to both Fe(II) centers and the second Fe(II) center also donates a single electron to the dioxygen bridge.

Part 3 of the stearoyl-CoA desaturaseFigure 21.1.29 is shown below.

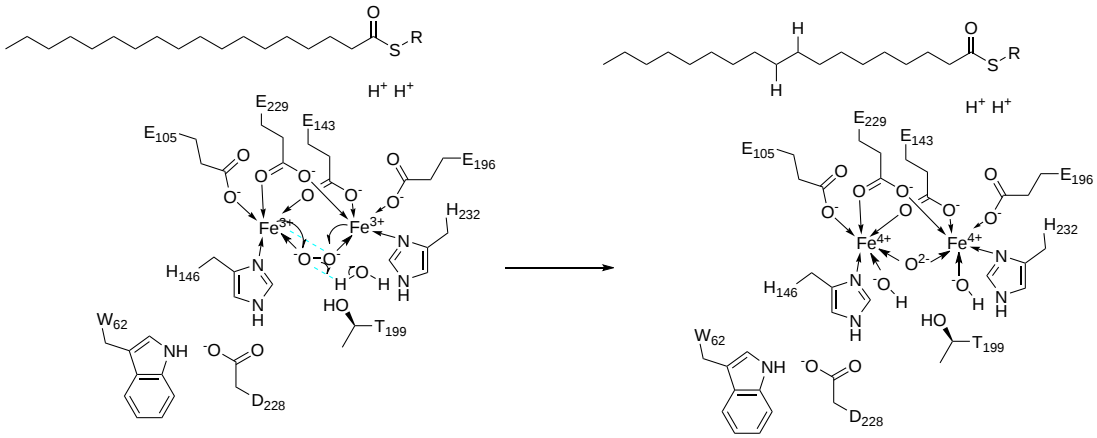


Figure 21.1.29: Part 3 of the stearoyl-CoA desaturase

Part 4 of the stearoyl-CoA desaturase is shown in Figure 21.1.30.

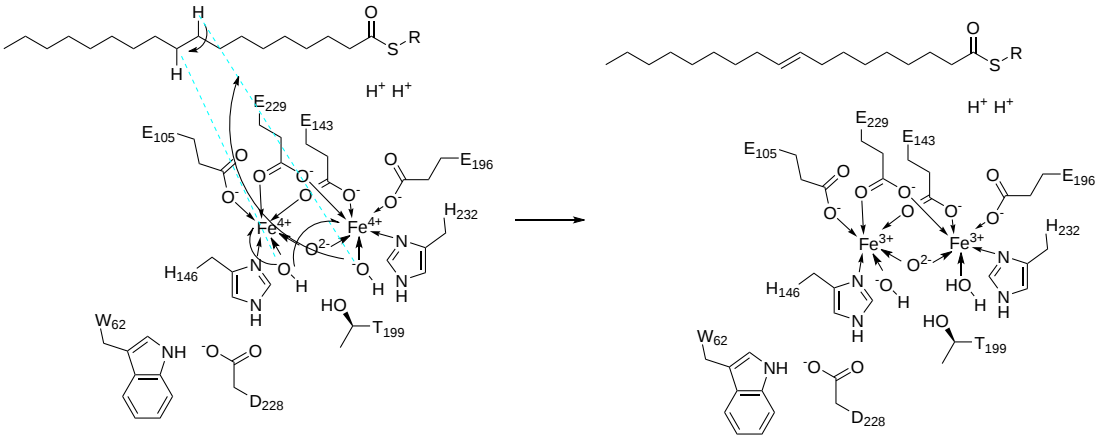


Figure 21.1.30: Part 4 of the stearoyl-CoA desaturase

In this reaction the hydroxide on the first Fe(IV) center deprotonates the stearoyl-[acyl-carrier protein] substrate, initiating the elimination of a hydride ion, which attacks the second Fe(IV) bound hydroxide. The excess electrons are donated singly to both Fe(IV) centers, regenerating the enzyme.

Figure 21.1.31 shows an interactive iCn3D model of the Stearoyl-acyl-carrier protein desaturase from castor seeds (1afr)

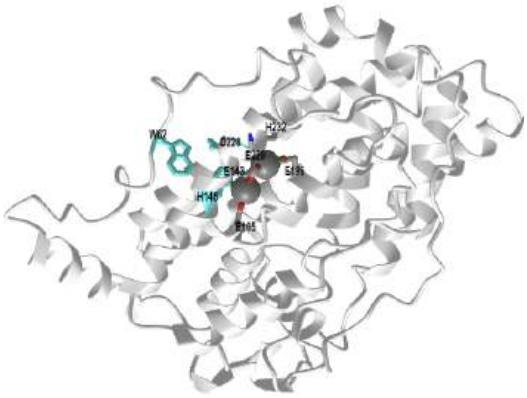


Figure 21.1.31: Stearoyl-acyl-carrier protein desaturase from castor seeds (1afr) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...UonrwyWfcg9wU8

An electron transport chain carries an electron from ferredoxin to one of the iron centers via several residues. A second electron is then carried to the second iron center. Several redox reactions take place involving oxygen, water, and iron centers, which result in the deprotonation of the substrate and the formation of a double bond.

There are two possible paths for the electron transport chain - the one described here has been chosen because of its analogy to what has been suggested for *E. coli* ribonucleotide reductase protein R2.

21.1.5: MITOCHONDRIAL FA SYNTHESIS

Mitochondria also engage in limited fatty acid synthesis. They use separate discrete enzymes (at least 6) as do most bacteria. They don't employ a multienzyme complex as described above for yeast and mammals. Nevertheless, there are commonalities. Synthesis proceeds by the addition of two carbon units to a growing acyl-ACP chain using malonyl-CoA and the release of CO_2 to drive the condensation process (in addition of course to ATP cleavage). The cytoplasmic fatty acid synthesis produces 16:0 (palmitate). In contrast, mitochondrial FAS produces two major products, 8:0 (octanoate). This is then converted to the 8C derivative lipoic acid. Their structures are shown in Figure 21.1.32.

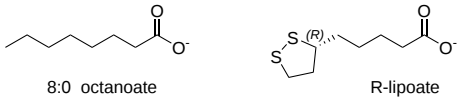


Figure 21.1.32 Structure of octanoate and R-lipoate

We have seen lipoic acid before as a cofactor in pyruvate dehydrogenase and α-ketoglutarate dehydrogenase, two enzymes that catalyze the more difficult decarboxylation of α-ketoacids (which don't have a built-in electron sink at a β carbon adjacent to the departing carboxyl group. It is also used in branched-chain amino acid dehydrogenases. The absence of lipoic acid is lethal in mice.

Figure 21.1.33 shows the fatty acid synthesis pathway in mitochondria.

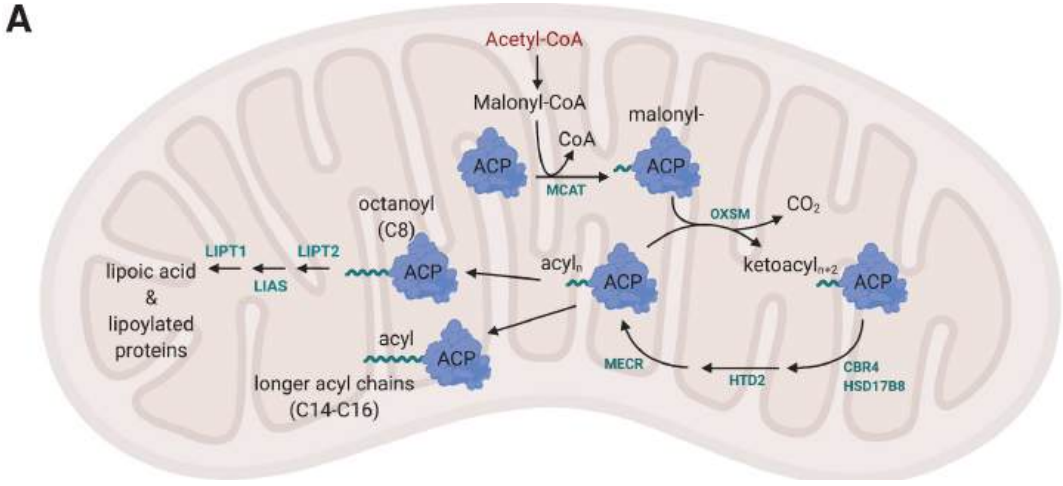


Figure 21.1.33: mitochondrial fatty acid synthesis pathway and downstream lipoic acid synthesis. Nowinski et al. eLife 2020;9:e58041. DOI: https://doi.org/10.7554/eLife.58041. Creative Commons Attribution License

Some of the enzymes in the pathway include the following: **Mcat** is malonyl-CoA ACP transacylase, and **Ox**sm is beta-ketoacyl synthase that catalyzes the Claisen condensation of malonyl-ACP with the growing acyl chain. **CBR4** is 3-oxoacyl-[acyl-carrier-protein] reductase. **HSD17B8** is (3R)-3-hydroxyacyl-CoA dehydrogenase. **HTD2** is hydroxyacyl-thioester dehydratase type 2). **Mecr** is the terminal enoyl-[acyl-carrier-protein] reductase. **LIPT2** is lipoyltransferase 2. **LIAS** is lipoyl synthase. **LIPT1** (Lipoyltransferase 1) catalyzes the terminal step in lipoic acid and lipoylated protein synthesis.

It now appears that longer fatty acyl chains (up to 14C) are synthesized in mitochondria. Mutations in key enzymes above have been made that don't affect lipoic acid synthesis but impair mitochondrial function.

21.1.6: EICOSINOID AND PROSTAGLANDIN SYNTHESIS

As shown in Fig 20, arachidonic acid, $20{:}4^{\Delta 5,8,11,14}$, an ω-6 fatty acid (5,8,11,14- eicosatetraenoic acid), is synthesized from linoleic acid, $18{:}2^{\Delta 9,12}$, another ω-6 polyunsaturated fatty acid, through a series of elongation and desaturation steps. Arachidonic acid is released from membrane phospholipids on activation of phospholipase A2 and converted to a class of molecules called prostaglandins, which have powerful hormone-like effects.

The prostaglandins are part of a group of icosanoids derived from cleavage of membrane ω-3 ($20{:}5^{\Delta 5,8,11,14,17}$, eicosapentaenoic acid) and ω-6 C_{20} fatty acids, which are metabolized to form leukotrienes (LTs), prostaglandins (PGs), prostacyclins (PCs), and thromboxanes (TXAs). These have differing and sometimes opposing physiologically effects in animals. Prostaglandins have been found in almost every tissue in humans and other animals

Ω-3 AND Ω-3 AND FATTY ACIDS

The most common polyunsaturated fats (PUFAs) in our diet are the ω-3 and ω-6 classes. Most abundant in the ω-6 class in plant food is linoleic acid (18:2 ω-6, or $18{:}2^{\Delta 9,12}$), while linolenic acid (18:3 ω-3 or $18{:}3^{\Delta 9,12,15}$) is the most abundant in the n-3ω class. These fatty acids are essential in that they are biological precursors for other PUFAs. Specifically,

- linoleic acid (18:2 ω-6, or $18{:}2^{\Delta 9,12}$) is a biosynthetic precursor of arachidonic acid (20:4 ω-6 or $20{:}4^{\Delta 5,8,11,14}$)
- linolenic acid (18:3 ω-3, or $18{:}3^{\Delta 9,12,15}$) is a biosynthetic precursor of eicosapentaenoic acid (EPA, 20:5 ω-3 or $20{:}5^{\Delta 5,8,11,14,17}$) and to a much smaller extent, docosahexaenoic acid (DHA, 22:6 ω-3 or $22{:}6^{\Delta 4,7,10,13,16,19}$).

These essential precursor fatty acids are substrates for intracellular enzymes such as elongases, desaturases, and beta-oxidation type enzymes in the endoplasmic reticulum and another organelle, the peroxisome (involved in the oxidative metabolism of straight chain and branched fatty acids, peroxide metabolism, and cholesterol/bile salt synthesis). Animals fed diets high in plant 18:2(n-6) fats accumulate 20:4(n-6) fatty acids in their tissues while those fed diets high in plant 18:3(n-3) accumulate 22:6(n-3). Animals fed diets high in fish oils accumulate 20:5 (EPA) and 22:6 (DHA) at the expense of 20:4(n-6).

Figure 21.1.34

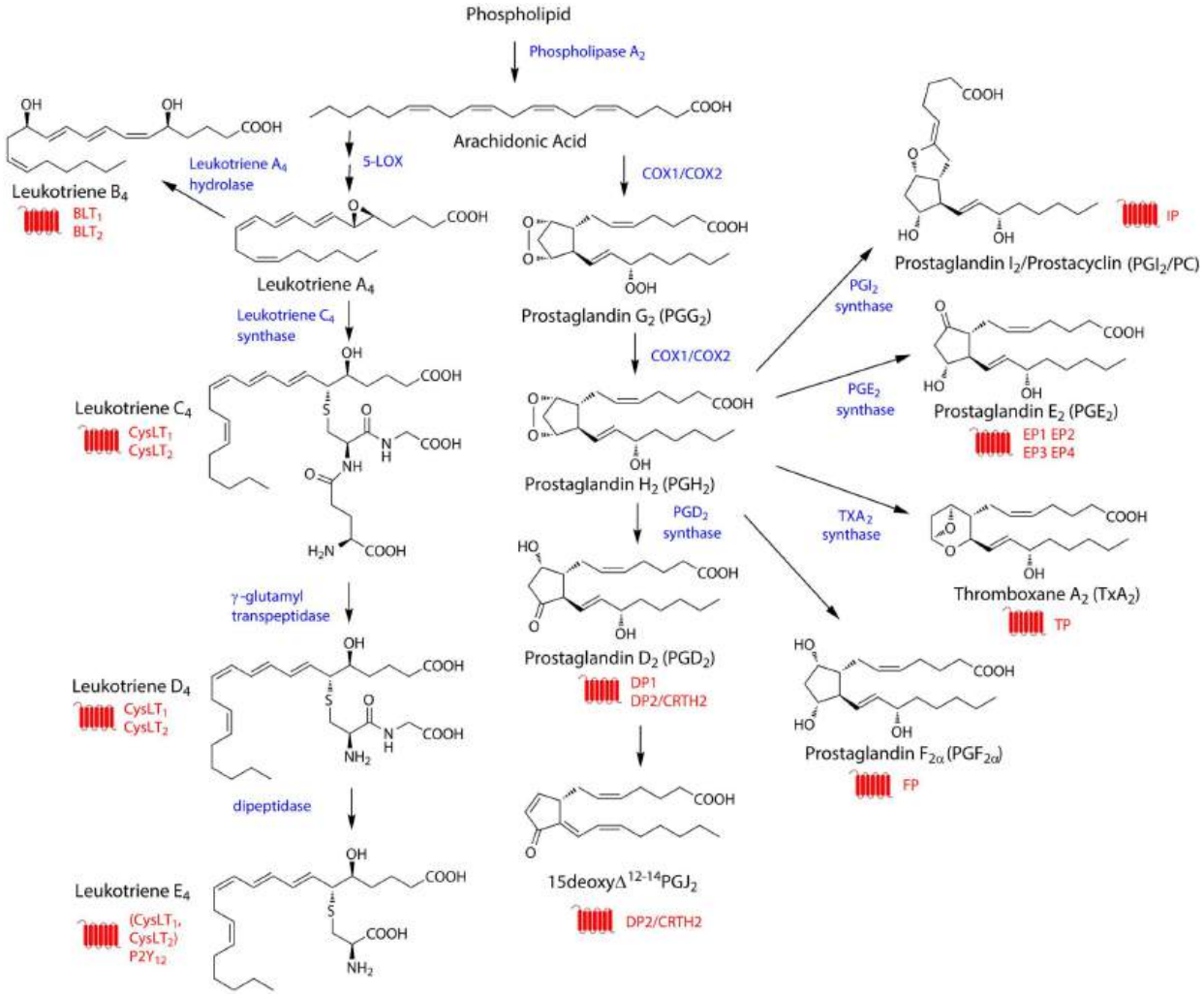


Figure 21.1.34Lone, Anna & Taskén, Kjetil. (2013). Proinflammatory and Immunoregulatory Roles of Eicosanoids in T Cells. Frontiers in immunology. . https://doi.org/10.3389/fimmu.2013.00130. https://creativecommons.org/licenses/by/3.0/

We will focus our attention on the conversion of arachidonic acid (20:4 ω-6 or $20{:}4^{\Delta 5,8,11,14}$) to prostaglandin G_2 (PGG_2) and H_2 (PGH_2) by the enzyme cyclooxygenase I and II. These enzymes are also called by their more formal names, prostaglandin endoperoxide H synthases (PGHSs) I and II. COX-1 prefers arachidonic acid as a substrate, while COX-2 has a broader substrate specificity and can use eicosapentaenoic acid (EPA, 20:5 ω-3 or $20{:}5^{\Delta 5,8,11,14,17}$) and even neutral substrate derivatives of arachidonic acid, including 2-arachidonoylglycerol and anandamide, both ligands for cannabinoid receptors.

COX 1/2 use two dioxygens and 2 electrons to make PGH_2. Two different, separate but connected catalytic sites are used to catalyze the two steps (shown in Figure 21.1.34)

1. Arachidonic acid ($20{:}4^{\Delta 5,8,11,14}$) is converted to PGG_2. This step is catalyzed by a **bis-oxygenase** domain. The word oxygenase is used since both atoms of dioxygen are added to the reactant. Bis indicates that two different molecules of dioxygen are added. Since one of the dioxygens added forms an endoperoxide cyclic bridge in PGG_2, this enzymatic domain, and the entire enzyme is usually called **cyclooxygenase** (COX). The second dioxygen is added as a peroxide (both oxygen atoms of dioxygen added);
2. Prostaglandin G_2 (PGG_2) is then converted to prostaglandin H_2 (PGH_2) in which the noncyclic peroxide form in the first step is converted to hydroxide. This is also a **reduction** step since the oxidation numbers of the oxygen atoms in the peroxide, -1, change to -2 (a gain of electrons) in the hydroxide in PGH_2. This reaction is catalyzed by the peroxidase (POX) activity of the enzyme. A cofactor heme in the POX site facilitates catalysis.

The COX reaction proceeds through radical intermediates. but it requires the activation of peroxide by the heme in the POX site. We'll show two different mechanisms for both reactions (COX and POX) catalyzed by the enzyme prostaglandin endoperoxide H synthase, which we will also call COX.

The first mechanism shown below in Figure 21.1.35 emphasizes the free radical nature of the reactions. It shows just one amino acid (Tyr 385) involved in the reaction. Note that the first step in the reaction shows the added dioxygen as an excited state "singlet" O_2, but it is

parenthetically noted that it most likely involves O_2 binding to and being activated by the heme iron in POX site.

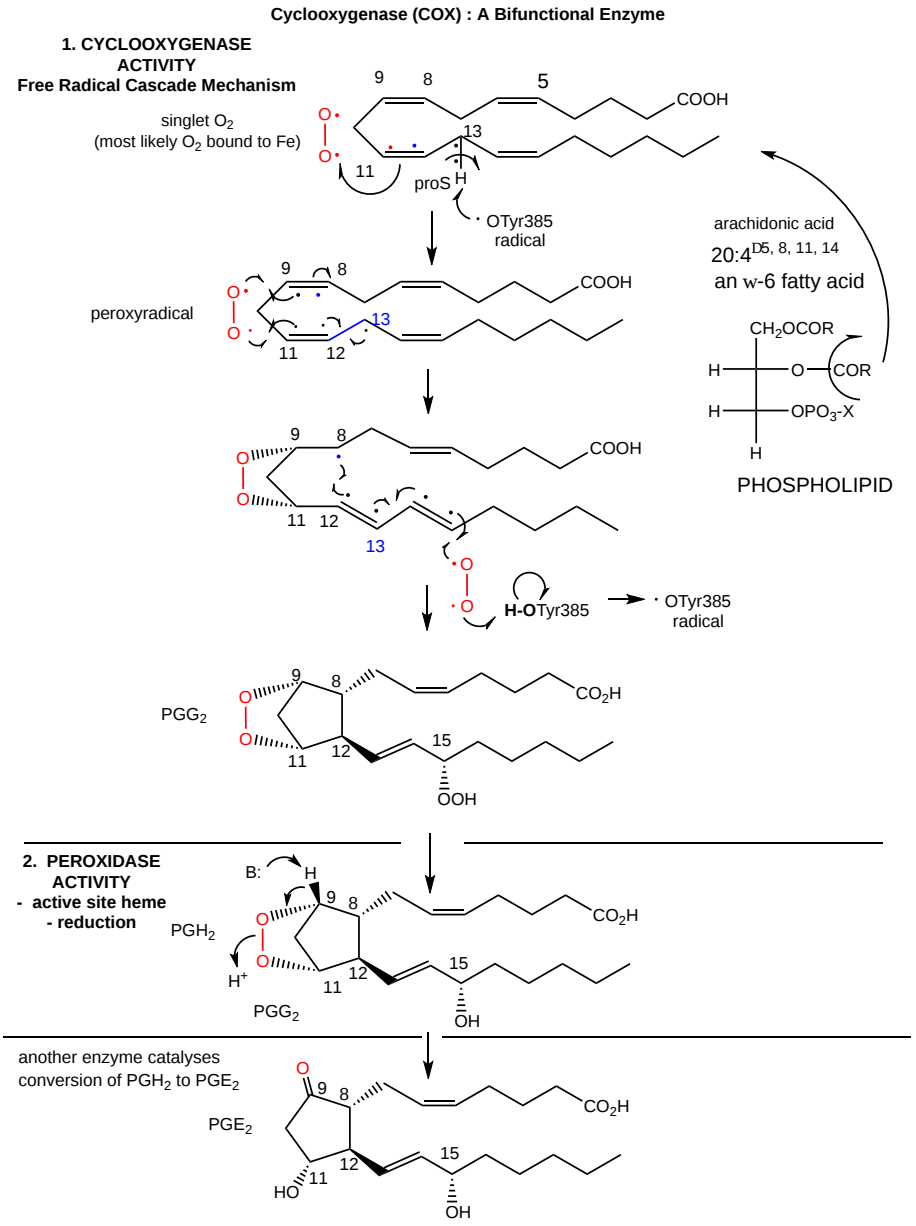


Figure 21.1.35: Simplified mechanism of prostaglandin endoperoxide H synthase (COX) emphasizing the radical nature of the mechanism.

A more detailed reaction mechanism for the synthesis of PGG_2 by prostaglandin endoperoxide H synthase (COX), emphasizing the role of active site amino acid and the heme in the POX site, is shown in Figures 36-38 below. (The mechanism and its explanation are adapted from https://www.ebi.ac.uk/thornton-srv/m-csa/entry/37/. Creative Commons Attribution 4.0 International (CC BY 4.0) License)

The 1st step involves a peroxide (denoted H-O-O*) but it might also be a ROO peroxide or NO. The generation of a Tyr-385· free radical in the COX site. is critical for the reaction. This is preceded by the heme radical cation in the POX site.

Figure \(\PageIndex{36}\) below shows the first step in which His 207 deprotonates the alkyl peroxide, which then coordinates to the heme Fe(III) in the POX site.

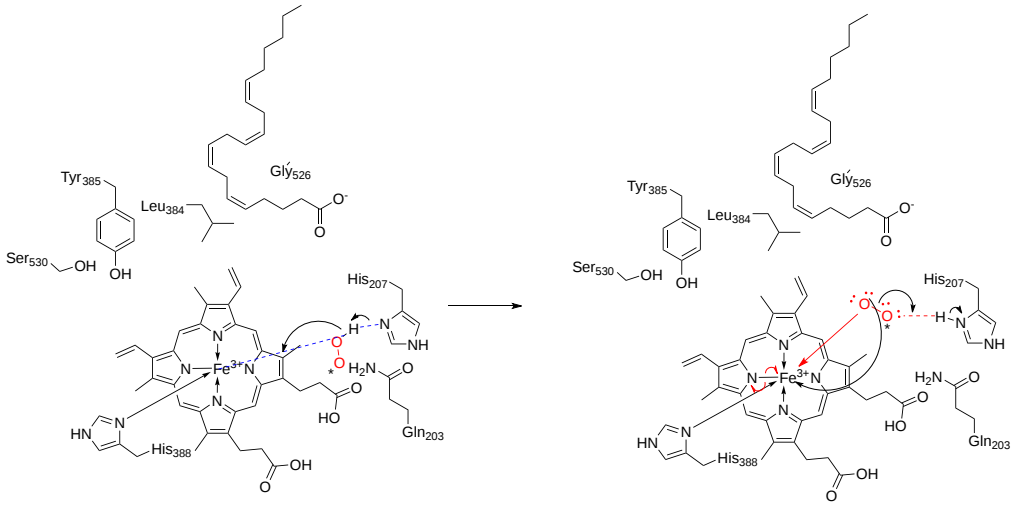


Figure 21.1.36: First step in the prostaglandin endoperoxide H synthase (COX) reaction.

In the next step, the peroxy bond donates two electrons to the heme Fe(III), one of which moves into the heme ring, forming Fe(IV). The heme ring then removes an electron from Tyr 385. This then removes a H. from the substrate, arachidonic acid. These steps are illustrated in Figure 21.1.37.

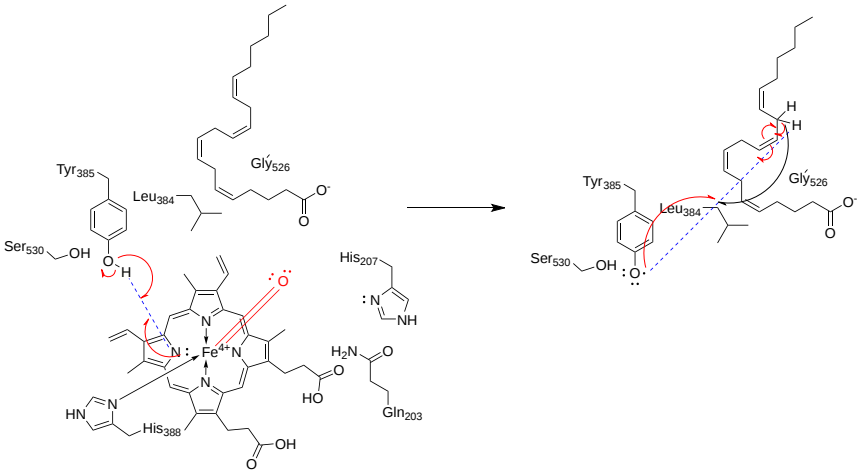


Figure 21.1.37: Next steps in the prostaglandin endoperoxide H synthase (COX) reaction.

The heme site is not shown in the next steps. Next, the reactive dioxygen, itself a ground state diradical, form a bound to the carbon-free radical in arachidonic acid. The rest of the reaction ensues as shown in Figure 21.1.38.

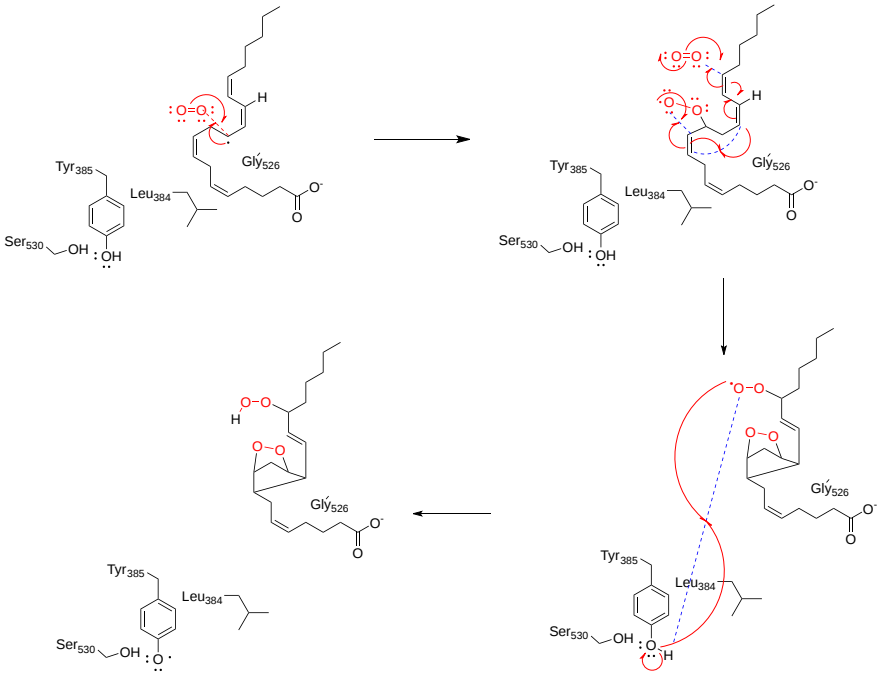
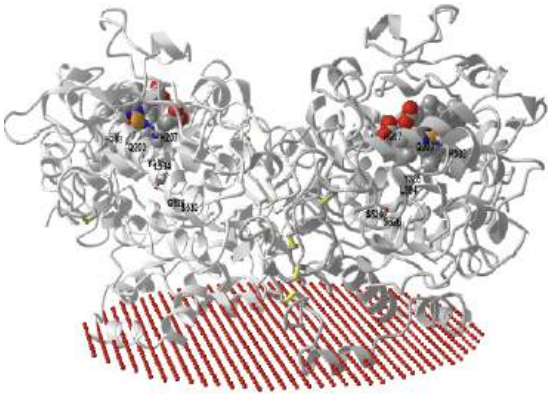


Figure 21.1.38 Figure 21.1.37: Final steps in the formation of PGG_2 by prostaglandin endoperoxide H synthase (COX) reaction.

Figure 21.1.39 shows an interactive iCn3D model of the mouse cyclooxygenase 2 (5COX)



NCBI iCn3D Figure 21.1.39: mouse cyclooxygenase 2 (5COX) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...wt9ECkgsyu4Vs8

Here is a model of human cyclooxygenase 1 structure (6Y3C): https://structure.ncbi.nlm.nih.gov/i...joK7oV63vLSi1A

Figure 21.1.39 shows an interactive iCn3D model of the mouse COX-2 with bound arachidonic acid (3KRK)

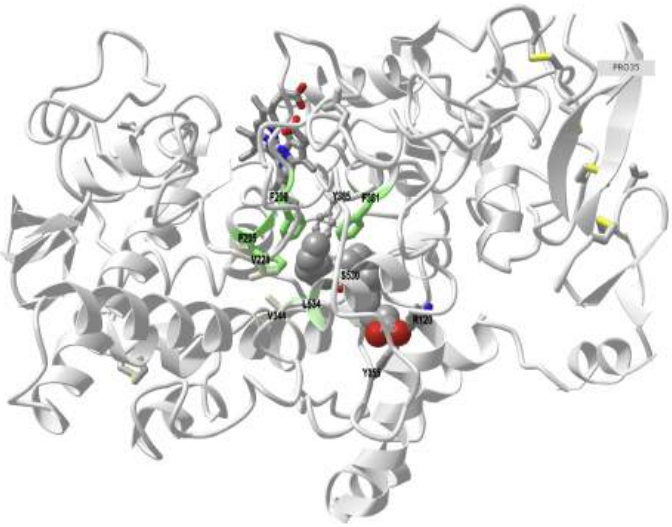


Figure 21.1.39: Mouse COX-2 with bound arachidonic acid (3KRK) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...yAtJnTkTmuNreA

The following key binding and catalytic residues are shown and labeled:

- Arg-120 and Tyr-355 are close to the carboxylate of AA
- Phe205, Phe209, Val228, Val344, Phe381, and Leu534 form a hydrophobic groove for the ω-end of AA.
- Ser 530, which is above this, gets acetylated by aspirin
- Tyr 385, near C13 in arachidonic acid, forms the free radical which removes a single electron from C13

Prostaglandins, which were first isolated from prostate glands, serve as powerful, but labile local hormones which are mediators of pain, inflammation, and immune and clotting activity. The cyclooxygenase activity is inhibited by aspirin, which probably accounts for most of its anti-inflammatory and analgesic properties. Aspirin, acetylsalicylic acid, acetylates a reactive Ser 530 in the active site. Another nonsteroidal anti-inflammatory drug (NSAID) with similar properties is Ibuprofen (Advil). Acetaminophen (Tylenol) is also considered a member of this drug class, even though it doesn't have anti-inflammatory properties. The question has arisen as to why. It now turns out that there are three different types of COX, I, II, and III. COX III is expressed in the brain and might be involved in pain pathways. Acetaminophen appears to work on this COX, as shown in Table 21.1.1 below (Bazan et al.).

Cyclooxygenase Activities

COX	Expression	Function	Inhibitors
COX 1	constitutively	organ pain, platelet function, stomach protection	NSAIDs including aspirin
COX 2	induced by growth factors, neurotransmitters, inflammatory cytokines, oxidative stress, and injury. Constitutively in the brain, kidney	Inducible COX2: inflammation, pain, fever Constitutive COX2: synaptic plasticity	NSAIDs, COX 2 inhibitors including celecoxib (Celebrex) which has few GI problems associated with its use
COX 3	constitutively, high in brain, heart	pain pathways, not inflammation pathways	acetaminophen (no GI problems, great fever reducer), some NSAIDs

21.1.7: FISH N-3 FATTY ACIDS AND HEALTH

We mentioned the importance of arachidonic acid in signal transduction in the lipid chapter. In addition, the importance of n-3 fatty acids to health was discussed as well. As mentioned above, arachidonic acid is cleaved from the C2 or sn-2 position of membrane phospholipids and modified by cyclooxygenase or lipoxygenase to form prostaglandins and leukotrienes, both potent local biological mediators. Linoleic acid and 22:6n-3 (DHA or docosahexaenoic acid) are also found in membrane phospholipids at the sn-2 position. What is the mechanism for the health-protective effects of n-3 fatty acids like DHA?

In human tissue, DHA, 22:6n-3 or $22{:}6^{\Delta 5,8,11,14,17,20}$ is the most abundant n-3 polyunsaturated fatty acids (PUFAs). Since it is synthesized from linolenic acid (as is EPA), a deficiency of linolenic acid in the diet will lead to lowered levels of 22:6n-3 in tissues, with ensuing health effects. Since these lipids are involved in membrane structure, signal transduction, and hormone synthesis, diverse effects of dietary n-3 PUFA deficiency will be observed. 50% of all fatty acids in the sn1 and sn2 position of membrane phospholipids of rod outer segments (in the retina) are 22:6(n-3). Cognitive dysfunctions (loss of memory, etc.) have been linked to decreased levels of 22:6(n-3) in the brain. This fatty acid binds to retinoid X receptors which then activate (through linked binding reactions) nuclear receptors, leading to alterations in gene transcription in the CNS.

In other tissues, 22:6(n-3) rarely exceeds 10% of membrane fatty acids, but this percentage can be increased in cells with increases in a precursor, 20:5(n-3). DHA might affect lipid rafts in the membrane, which would affect the movement of important membrane protein receptors (and associated proteins) in the membrane, altering cell response to environmental stimuli. DHA and EPA affect arachidonic acid conversion to prostaglandins and leukotrienes. EPA binds less tightly to cyclooxygenase I and is a poor substrate for the enzyme, both effects which inhibit the formation of prostaglandins and signaling processes mediated by them. This explains why n-3 fatty acids have anti-inflammatory effects.

In addition, n-3 fatty acids have noticeable effects on gene transcription, which remain as long as these fatty acids are present in high levels in the diet These and other fatty acids bind to fatty acid-activated transcription factors called PPARs (peroxisome proliferator receptors - alpha, beta and gamma 1 and 2). These receptors regulate, through alterations in gene expression, proteins involved in lipid metabolism. Other fatty acid-dependent transcription factors are known as well. PPARs bind 20:5(n-3) with a micromolar Kd and change the conformation of the protein to a form than can bind other proteins, ultimately altering gene expression.

Table 21.1.2: Some biological effects of n-3 polyunsaturated fatty acids EPA (20:5) and DHA (22:6). Adapted from Jump D. The Biochemistry of n-3 Polyunsaturated fatty acids. J. Biol. Chem. 277, pg 8755 (200)

Organ(s)	Effect	Mechanism acts through
central nervous system	improve cognitive function	membrane composition; retinoic X receptor alpha
retina	improve acuity	membrane composition
immune	immunosuppressive; anti-inflammatory	membrane composition; rafts
cardiovascular	anti-arrhythmia; anti-clotting	membrane composition; rafts; eicosanoids
serum lipids	lowers triglycerides (a risk factor for cardiovascular. disease)	peroxisome proliferator receptor alpha and gamma
liver	decrease lipid synthesis; increase fatty acid oxid. decrease VLDL synthesis	sterol reg. element bind. protein; PPAR alpha PPAR alpha

21.1.8: RECOGNITION OF UNSATURATED FATTY ACIDS BY MEMBRANE G-PROTEIN-COUPLED RECEPTORS

The positive health effects of the n-3 (ω-3) fatty acids appear to involve those fatty acids acting as hormones that bind to a special membrane receptor called G protein-coupled receptor 120 (GPCR120). We will discuss GPCRs in greater detail in Chapter 28.2. When a hormone bounds to the membrane GPCR (a transmembrane protein with 7 membrane-spanning helices), the intracellular domain of the GPC alters how it interacts with a variety of binding partners inside the cell. The GPCRs are given that name since most interact with heterotrimeric G proteins inside the cell. These G proteins have a β, γ, and a variety of different α subunits (α_q, α_i, α_s). Activation of different α subunits leads to different downstream effects as part of an elaborate signaling system in the cell. When the health-promoting polyunsaturated n-3 (ω-3) fatty bind to GPCR120, here are some of the beneficial health effects on metabolism, as outlined in Table 21.1.3 below.

Protein involved in GPCR120 signaling	Downstream effect
$G\alpha_q$	Increases Ca^{2+} levels in the cell, which act as a second signal (messenger) Increases GLUT4 (glucose transporter) translocation to the cell membrane; Increases secretion of glucagon-like peptide 1 (GLP1)
$G\alpha_i$	Increases insulin secretion Inhibits Ghrelin (hunger hormone) secretion
$G\alpha_s$	Controls fat synthesis
non G protein GPCR kinase (GRK)/β-arrestin 2	Inhibits the NLRP3 inflammasome and is ant-iinflammatory

The data in the table above is summarized in Figure 21.1.40 below.

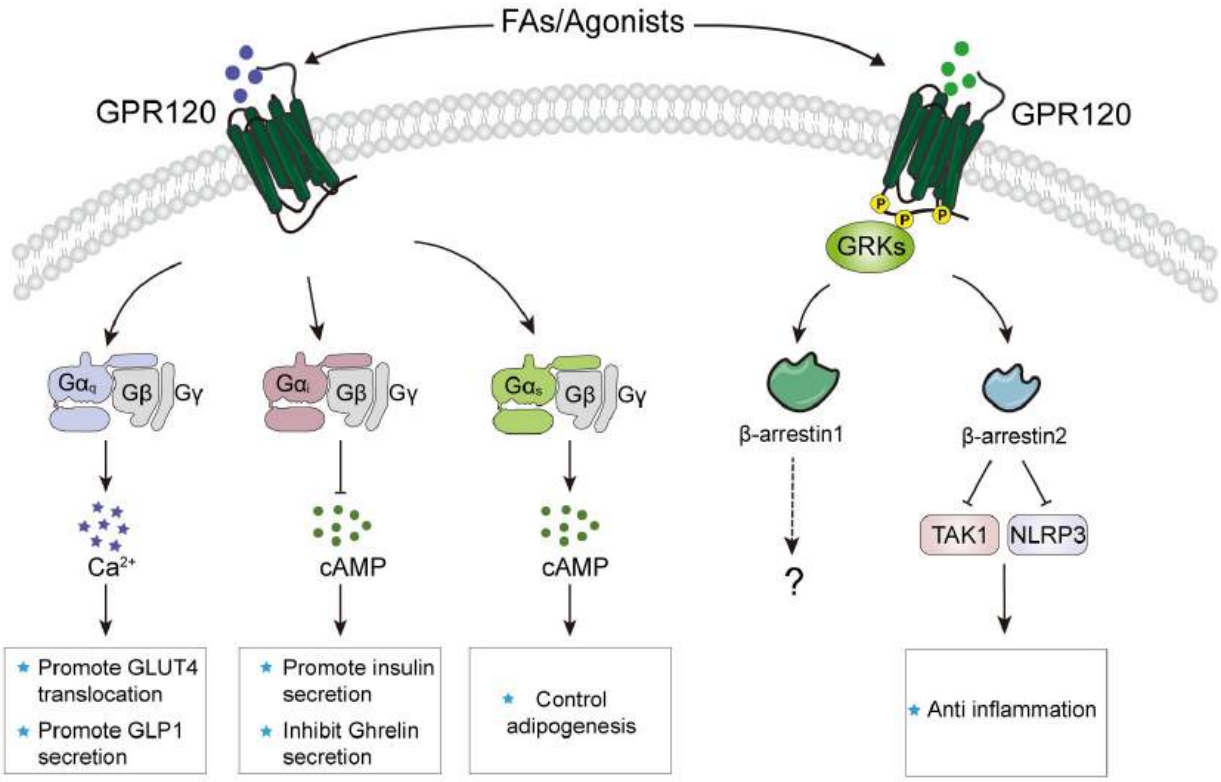
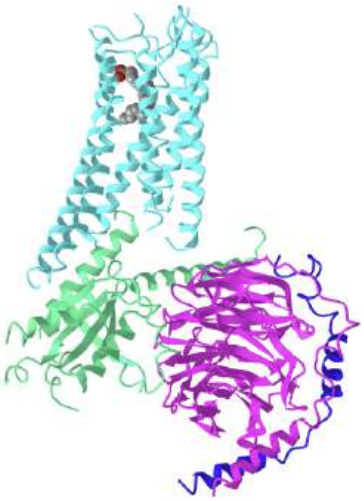


Figure 21.1.40: Schematic overview of G protein–and arrestin–mediated GPR120 signaling and related functions. Chunyou Mao et al. *Science* **380**, eadd6220 (2023). https://www.science.org/doi/10.1126/science.add6220. With permission from the AAAS.

GPCR120 and differentiate fatty acids based on the number and position of the double bond in the acyl chain of the ligand. Even though saturated fatty acids can bind to GPCR 120, only the n-3 (ω-3) fatty acids elicit the health effects.

CryoEM structures show the fatty acids buried in an L-shaped conformation in the GPCR. The protein engages in noncovalent interactions, specifically π:π interactions, with the double bonds in acyl chain. These interactions, which depend on the number of double bonds in the n-3 (ω-3) fatty acid differentially biases the GPCR towards the promotion of the specific effects for each different fatty acid.

Figure 21.1.41 shows an interactive iCn3D model of the eicosapentaenoic acid bound GPR120-Gi complex (8ID9).

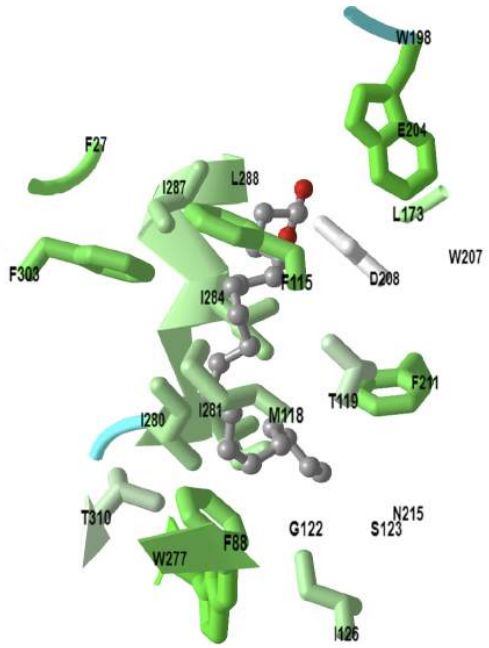


NCBI iCn3D Figure 21.1.41: Eicosapentaenoic acid bound GPR120-Gi complex (8ID9). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...vgMGJRTr3emNj7

The color coding is as follows:

- GPCR120 - cyan (the membrane would be perpendicular across the 7 membrane helices) with the bound EPA
- Gβ - magenta
- Gγ - blue
- EPA -spacefill

Figure 21.1.42 shows an interactive iCn3D model showing the nonpolar environment of eicosapentaenoic acid bound to GPR120 (8ID9).



NCBI iCn3D Figure 21.1.42: Nonpolar environment of eicosapentaenoic acid bound to GPR120 (8ID9). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...SszD1K3B9ieND7

Dark green shows the hydrophobic amino acids

21.2: BIOSYNTHESIS OF TRIACYLGLYCEROLS

by William (Bill) W. Christie and Henry Jakubowski.

This section is an abbreviated and modified version of material from the Lipid Web, an introduction to the chemistry and biochemistry of individual lipid classes, written by William Christie.

21.2.1: INTRODUCTION

All eukaryotic organisms and even a few prokaryotes can synthesize triacylglycerols, and in animals, many cell types and organs have this ability, but the liver, intestines, and adipose tissue are most active with most of the body stores in the last of these. Within all cell types, even those of the brain, triacylglycerols are stored as cytoplasmic '**lipid droplets**' enclosed by a monolayer of phospholipids and hydrophobic proteins such as the perilipins in adipose tissue or oleosins in seeds. These lipid droplets are now treated as distinctive organelles, with their own characteristic metabolic pathways and associated enzymes - no longer boring blobs of fat. However, they are not unique to animals and plants, as Mycobacteria and yeasts have similar lipid inclusions.

The lipid serves as a store of fatty acids for energy, which can be released rapidly on demand, and as a reserve of fatty acids for structural purposes or as precursors for eicosanoids. In addition, lipid droplets serve as a protective agency in cells to sequester any excess of biologically active and potentially harmful lipids such as free fatty acids, oxylipins, diacylglycerols, cholesterol (as cholesterol esters), retinol esters, and coenzyme A esters.

While triacylglycerols are essential for normal physiology, an excessive accumulation in human adipose tissue and other organs results in obesity and other health problems, including insulin resistance, steatohepatitis, and cardiomyopathy. Accordingly, there is considerable pharmaceutical interest in drugs that affect triacylglycerol biosynthesis and metabolism.

21.2.2: BIOSYNTHESIS OF TRIACYLGLYCEROLS

Three main pathways for triacylglycerol biosynthesis include the *sn*-glycerol-3-phosphate and dihydroxyacetone phosphate pathways, which predominate in liver and adipose tissue, and a monoacylglycerol pathway in the intestines. In maturing plant seeds and some animal tissues, a fourth pathway has been recognized in which a diacylglycerol transferase is involved.

The most important route to triacylglycerols is the *sn*-glycerol-3-phosphate or Kennedy pathway, first described by Eugene Kennedy and colleagues in the 1950s, from which more than 90% of liver triacylglycerols are produced. Figure 21.2.1 shows the pathway from glycerol-3-phosphate to triacylglycerols (the Kennedy pathway) and some preceding reactions that generate glycerol-3-phosphate.

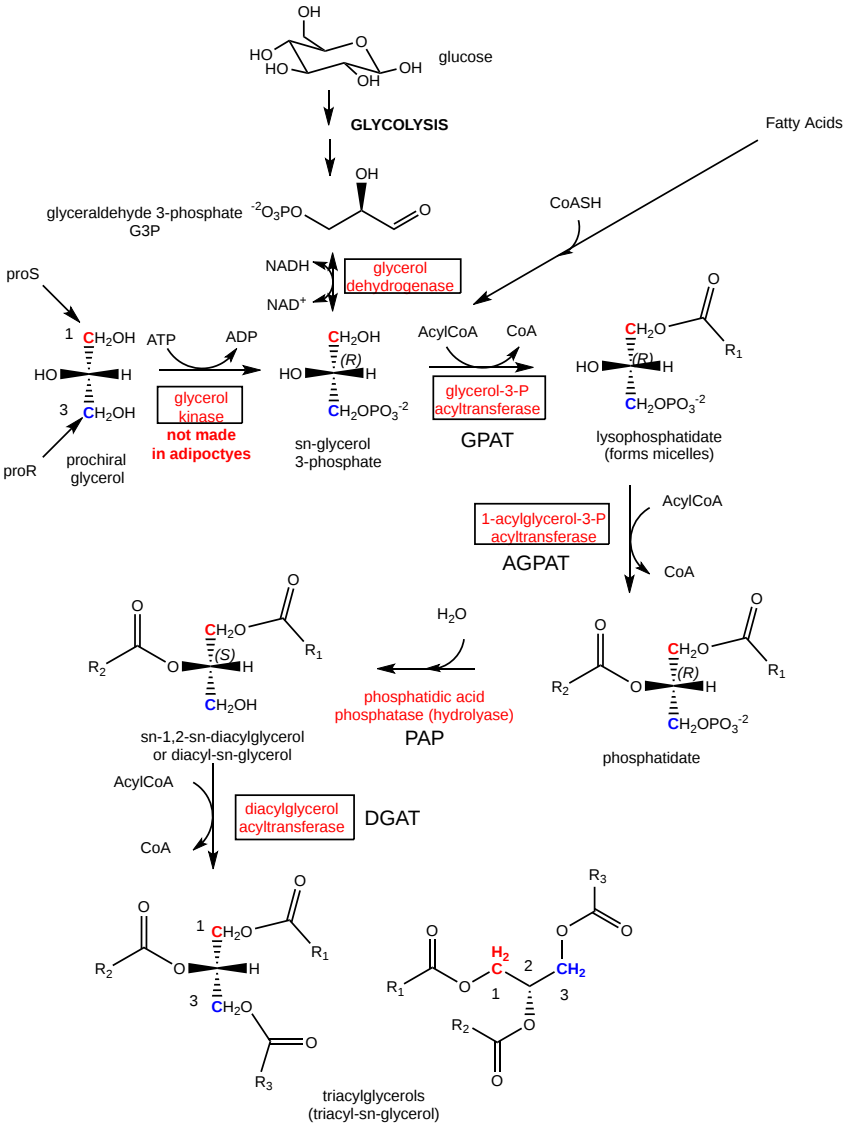


Figure 21.2.1: Synthesis of triacylglycerols from glycerol-3-phosphate (Kennedy pathway) and some preceding reactions that generate glycerol-3-phosphate.

In this pathway, the main source of the glycerol backbone has long been believed to be *sn*-glycerol-3-phosphate produced by the catabolism of glucose (glycolysis) or to a lesser extent by the action of the enzyme glycerol kinase on free glycerol. However, there is increasing evidence that a significant proportion of glycerol is produced *de novo* by a process known as glyceroneogenesis via pyruvate. Indeed, this may be the main source in adipose tissue.

Subsequent reactions occur primarily in or at the endoplasmic reticulum. First, the precursor *sn*-glycerol-3-phosphate is esterified by a fatty acid coenzyme A ester in a reaction catalyzed by a glycerol-3-phosphate acyltransferase (GPAT) at position *sn*-1 to form lysophosphatidic acid, and this is in turn acylated by an acylglycerophosphate acyltransferase (AGPAT) in position *sn*-2 to form a key intermediate in the biosynthesis of all glycerolipids - **phosphatidic acid.** Numerous isoforms of these enzymes are known; they are expressed with specific tissue and membrane distributions, and they are regulated in different ways.

Let's look at some of the enzymes involved in some key reactions in the synthesis of triacylglycerols.

GPAT Enzymes

Glycerol-3-phosphate acyltransferase (GPAT) catalyzes the first step in the pathway in most tissue. From an enzymatic perspective, it can be considered rate-limiting for the pathways since its specific activity is slow. In mammals, there are four isoforms (GPAT1 in the outer membrane of the mitochondria, GPAT2 also in the mitochondria, and GPAT3 and 4, both in the ER. The activity of GPAT1 is activated by insulin, presumably through phosphorylation) and inhibited by AMP-activated protein kinase (AMPK).

Fatty acyl-CoA in the outer membrane could be used for TAG synthesis through the Kennedy pathway in times of energy abundance or imported into the matrix through the carnitine-acyl-CoA cycle in times of energy need. Overexpression of GPAT1 in liver cells increases

TAG synthesis and decreases β-oxidation. GPAT1 deficient mice in contrast have higher levels of β-oxidation.

Figure 21.2.2 shows an interactive iCn3D model of the AlphaFold predicted structure of human mitochondrial glycerol-3-phosphate acyltransferase 1 - GPAT1 (Q9HCL2).

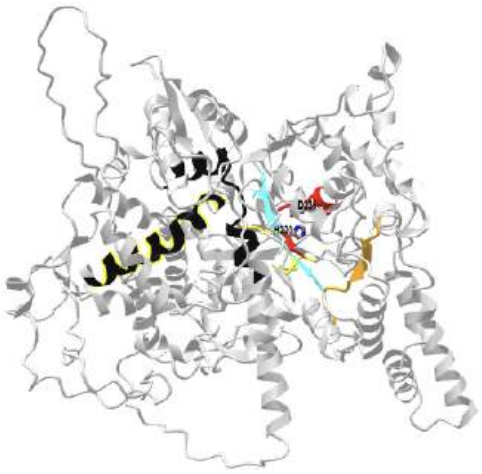


Figure 21.2.2: AlphaFold predicted structure ofhuman mitochondrial glycerol-3-phosphate acyltransferase 1 - GPAT1 (Q9HCL2). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...kuvBTPo5mNK8p9

Four conserved motifs involved in catalysis (red and cyan) and binding (orange and yellow) are shown. Significant sections are still disordered in the model. Two highly conserved amino acids, His 230 and Asp 235 in the red catalytic motif are shown in CPK-colored sticks and labeled. The black helices are predicted outer mitochondrial membrane transmembrane helices. Both the N- and C-terminal amino acids project into the cytoplasm.

Phosphatidic acid phosphohydrolases (PAPs or 'phosphatidate phosphatases' or 'lipid phosphate phosphatases').

PAPs are also important as they produce ***sn*-1,2-diacylglycerols** as essential intermediates in the biosynthesis not only of triacylglycerols but also of phosphatidylcholine and phosphatidylethanolamine (and of monogalactosyldiacylglycerols in plants). This is a key branch-point in lipid biosynthesis as it may dictate the flow of lipids for storage or membrane biogenesis.

Much of this phosphatase activity leading to triacylglycerol biosynthesis in animals reside in three related cytoplasmic proteins, termed **lipins**, i.e., lipin-1, lipin-2 and lipin-3, which have tissue-specific roles in glycerolipid synthesis. Unusually, these were characterized and named before the nature of their enzymatic activities was determined. Each of the lipins appears to have distinctive expression and functions, but lipin-1 (PAP1) in three isoforms (designated 1α, 1β, and 1γ) accounts for most of the PAP activity in adipose tissue and skeletal muscle in humans. Lipin 2 is the most abundant lipin in the liver but is also expressed substantially in the small intestine, macrophages, and some regions of the brain, while lipin 3 activity overlaps with that of lipin 1 and lipin 2 and is found in the gastrointestinal tract and liver. Lipins are cytosolic enzymes but associate transiently with membranes to access their substrate, i.e., they are translocated to the endoplasmic reticulum in response to elevated levels of fatty acids within cells, although they do not have trans-membrane domains. Lipin-1 activity requires Mg^{2+} ions and is inhibited by *N*-ethylmaleimide, whereas the membrane-bound activity responsible for synthesizing diacylglycerols as a phospholipid intermediate is independent of Mg^{2+} concentration and is not sensitive to the inhibitor.

Perhaps surprisingly, lipin-1 has a dual role in that it operates in collaboration with known nuclear receptors as a transcriptional coactivator to modulate lipid metabolism (lipin 1α) while lipin 1β is associated with the induction of lipogenic genes such as fatty acid synthase, stearoyl-CoA desaturase, and DGAT. They can have profound effects on signaling in a variety of cell types. Abnormalities in lipin-1 expression are known to be involved in some human disease states that may lead to metabolic syndrome and inflammatory disorders. Lipin 2 is a similar phosphatidate phosphohydrolase, which is regulated dynamically by fasting and obesity (in mice).

PAP depend on divalent Mg^{2+} ions and are sensitive to N-ethylmaleimide. Lipin has a N-Lip (N-terminal) and C-Lip (C-terminal) domain. The crystal structure of lipin with a 250 amino acid regions between the two domains removed has been solved. The domain structures of humans, the yeast Saccharomyces cerevisiae (SC) and its truncated form (Tt-Pah2), and the structure of the Tt-Pah2 truncated fragment of the enzyme from Tetrahymena thermophila, are shown in Figure 21.2.3.

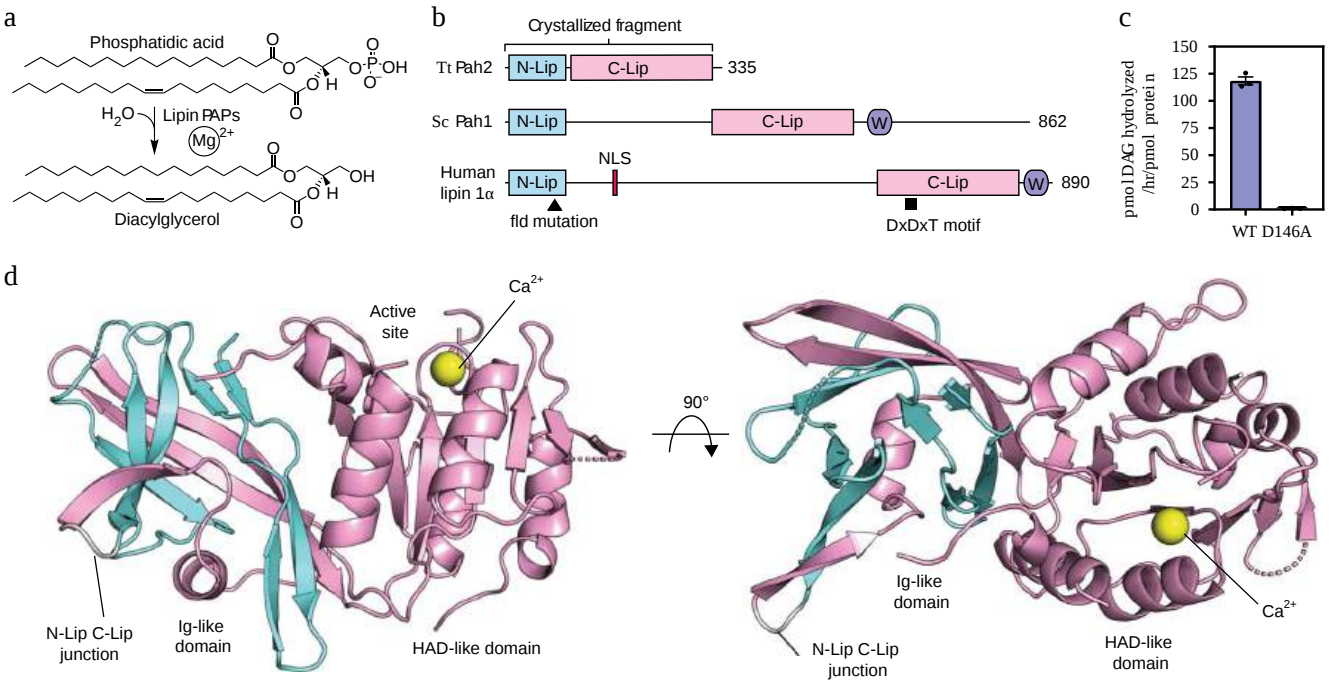


Figure 21.2.3: The domain structures of human, yeast Saccharomyces cerevisiae (SC) and a truncated form (Tt-Pah2) of the Tetrahymena thermophila enzyme. Khayyo, V.I., Hoffmann, R.M., Wang, H. *et al.* Crystal structure of a lipin/Pah phosphatidic acid phosphatase. *Nat Commun* **11,** 1309 (2020). https://doi.org/10.1038/s41467-020-15124-z. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Panel (a) shows the PAP-catalyzed dephosphorylation reaction.

Panel (b) shows the domain architecture of PAPs drawn to scale. The positions of the nuclear localization signal (NLS), conserved Trp-motif (purple W), catalytic DxDxT motif, and fatty liver dystrophy (fld^{2J}) mutation are indicated.

Panel (c) shows the the wild-type (WT) *Tt* Pah2 is catalytically active and the D146A mutant of the DxDxT motif eliminates activity.

Panel (d) shows the overall structure of *Tt* Pah2, which contains an immunoglobulin-like (Ig-like) domain and a HAD-like catalytic domain. The Ig-like domain is formed by the N-Lip (cyan) and C-Lip (pink) regions connected by a short linker (gray loop) that would be replaced by the extended 500-residue linker in human lipins and a 250-residue linker in *Sc* Pah1. A calcium ion (Ca^{2+}, yellow sphere) is bound in the active site of the HAD-like catalytic domain. The top view (right) of the enzyme shows the N-Lip co-folding with the C-Lip to form the Ig-like domain.

Figure 21.2.4s shows a closeup of the active site of *Tt* Pah2 and a mechanism for the hydrolytic removal of the phosphate on DAG.

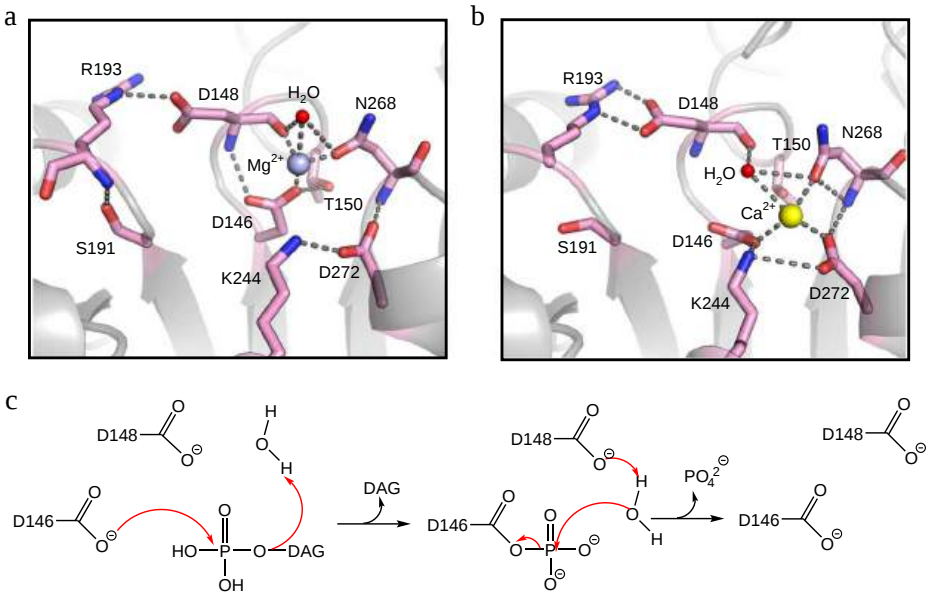
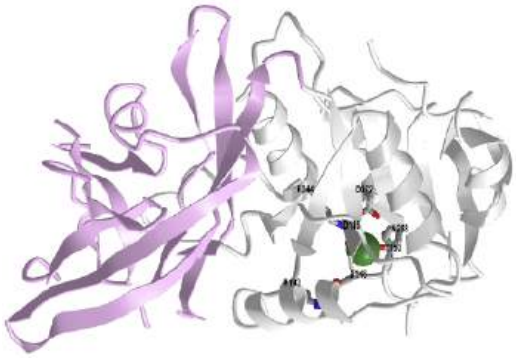


Figure 21.2.4: **a** Detailed view of the *Tt* Pah2 active site with magnesium bound (light blue sphere). Polar contacts are indicated by gray dashes. **b** Detailed view of the structural rearrangement of key residue interactions in the active site with calcium bound (yellow sphere). **c** Proposed schematic of lipin/Pah PAPs catalytic mechanism. Catalysis proceeds through a trigonal bipyramidal intermediate. Khayyo, V.I. et al, ibid.

Figure 21.2.5 shows an interactive iCn3D model of the Tetrahymena Thermophila lipin phosphatidic acid phosphatase with magnesium (6TZZ).



NCBI iCn3D Figure 21.2.5: Tetrahymena Thermophila lipin phosphatidic acid phosphatase with magnesium (6TZZ). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...GTusCCDP5eRrt7

The N-terminal domain is shown in lavender and the C-terminal domain is in gray. Key amino acids in the active site are shown in CPK-colored sticks and labeled.

Diacylglycerol acyltransferases (DGAT)

In the final step in this pathway, the 1,2-diacyl-*sn*-glycerol intermediate is acylated by **diacylglycerol acyltransferases (DGAT)**, which can utilize a wide range of fatty acyl-CoA esters to form the triacyl-*sn*-glycerol. There are two DGAT enzymes, which are structurally and functionally distinct. In animals, DGAT1 is located mainly in the endoplasmic reticulum and is expressed in skeletal muscle, skin, and intestine, with lower levels of expression in the liver and adipose tissue. It is believed to have dual topology contributing to triacylglycerol synthesis on both sides of the membrane of the endoplasmic reticulum but esterifying only pre-formed fatty acids of exogenous origin. Perhaps surprisingly, DGAT1 is the only one present in the epithelial cells that synthesize milk fat in the mammary gland. Also, DGAT1 can utilize a wider range of substrates, including monoacylglycerols, long-chain alcohols (for wax synthesis), and retinol, and it is reported to have an important role in protecting the endoplasmic reticulum from the lipotoxic effects of high-fat diets. Orthologs of this enzyme are present in most eukaryotes, other than yeasts, and they are especially important in plants.

Figure 21.2.6 shows an interactive iCn3D model of Human Diacylglycerol Acyltransferase 1 in complex with oleoyl-CoA (6VP0)

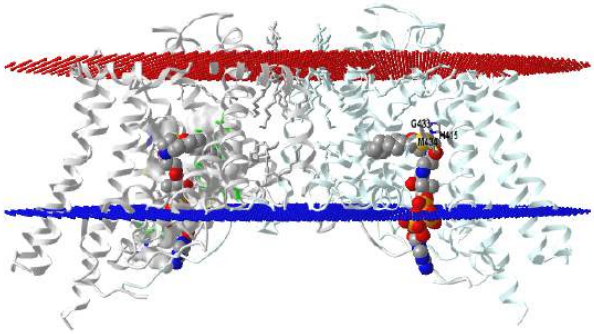


Figure 21.2.6: Human Diacylglycerol Acyltransferase 1 in complex with oleoyl-CoA (6VP0) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...adyWg9fy2hh8r8

The red layer is the extracellular leaflet and the blue is the intracellular leaflet. The protein is a dimer with one subunit shown in gray and one in light. A substrate, oleoyl-CoA, is shown in spacefill with CPK colors in both monomers. The acyl chain with a clear kink induced by the double bond in the oleoyl acyl group is clearly in the hydrophobic interior of the bilayer. The gray monomer shows the atoms within 4 Å from the oleoly-CoA is shown as surface color-coded by hydrophobicity with the green most hydrophobic. The active site His 415 and Met 434 which hydrogen bonds to it are shown in the light cyan subunit in sticks, CPK colors, and labeled. Three other conserved polar side chains, Asn378, Gln437, and Gln465, which are in the active site region, are also shown.

Each monomer in the homodimer has 9 transmembrane helices with 8 forming a unique fold called the MBOAT which encloses a chamber with catalytic side chains. The substrates, fatty acyl-CoA, and DAG have different entrances to the chamber with the acyl-CoA occupying a cytosolic tunnel. The cytosolic face interacts with the CoA and the acyl chain extends into the ER membrane. The channel appears bent which probably allows the acylation of DAGs and long-chain alcohols, but not cholesterol, a rigid, planar molecule that is acylated to form an ester by a different enzyme, acyl CoA:cholesterol acyltransferases. Hence it appears that DAG, the acyl-acceptor, binds in a hydrophobic tunnel from within the bilayer and the acyl-CoA binds in another tunnel, allowing the transfer of the acyl group.

DGAT2 is the main form of the enzyme in hepatocytes and adipocytes (lipid droplets), although it is expressed much more widely in tissues. It is associated with distinct regions of the endoplasmic reticulum, at the surface of lipid droplets, and in mitochondria, and it esterifies fatty acids of both endogenous and exogenous origin. DGAT2 is believed to have a targeting domain that enables it to tether between the endoplasmic reticulum and lipid droplet thereby channeling triacylglycerols from the synthesis site in the endoplasmic reticulum to the nascent lipid droplet, where they accumulate and lead to the expansion of the latter (see **below**). Both enzymes are important modulators of energy metabolism, although DGAT2 appears to be especially important in controlling the homeostasis of triacylglycerols *in vivo*. As the glycerol-3-phosphate acyltransferase (GPAT) has the lowest specific activity of these enzymes, this step may be the rate-limiting one. However, DGATs are the dedicated triacylglycerol-forming enzymes, and they are seen as the best target for pharmaceutical intervention in obesity and attendant ailments; clinical studies of DGAT1 inhibitors are at an early stage.

In a second pathway for triacylglycerol biosynthesis, dihydroxyacetone-phosphate in peroxisomes or endoplasmic reticulum can be acylated with fatty acid CoA esters by a specific acyltransferase to form 1-acyl dihydroxyacetone-phosphate, which is reduced by dihydroxyacetone-phosphate oxidoreductase to lysophosphatidic acid; this can then enter the pathway above to triacylglycerols. The precursor dihydroxyacetone-phosphate is important also as part of the biosynthetic route to **plasmalogens**, and neutral plasmalogens can be significant components of cytoplasmic droplets in many mammalian cells types but not in adipose tissue.

Figure 21.2.7: Biosynthesis of triacylglycerols from dihydroxyacetone-phosphate

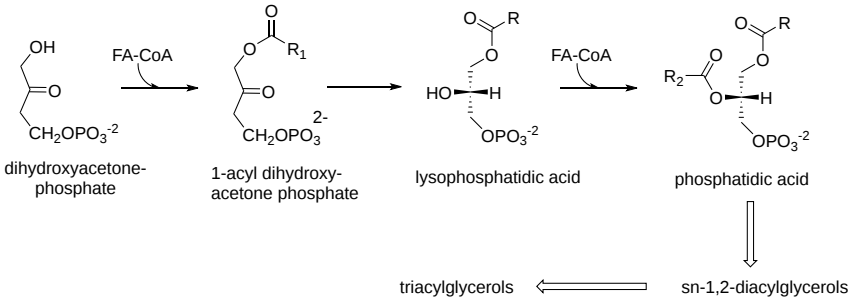


Figure 21.2.7: Biosynthesis of triacylglycerols from dihydroxyacetone-phosphate

In prokaryotes, the glycerol-3-phosphate pathway of triacylglycerol biosynthesis only occurs, but in yeast, both glycerol-3-phosphate and dihydroxyacetone-phosphate can be the primary precursors and synthesis takes place in cytoplasmic lipid droplets and the endoplasmic reticulum. In plants, the glycerol-3-phosphate pathway is most important.

In the enterocytes of the intestines after a meal, up to 75% of the triacylglycerols are formed via a monoacylglycerol pathway. In this, 2-monoacyl-*sn*-glycerols and free fatty acids released from dietary triacylglycerols by the action of pancreatic lipase within the intestines (see below) are taken up by the enterocytes. There, the monoacylglycerols are first acylated by an acyl-coenzyme A:monoacylglycerol acyltransferase with the formation of *sn*-1,2-diacylglycerols mainly as the first intermediate in the process, though some *sn*-2,3-diacylglycerols (~10%) are produced. In addition, 1-monoacylglycerols can be synthesized by acylation of glycerol for further acylation. There are three isoforms of the monoacylglycerol acyltransferase in humans of which MGAT2 is most active in the intestines, but also in the liver where an appreciable proportion of the triacylglycerols are formed by the monoacylglycerol pathway, while MGAT1 functions in adipose tissue; the role of MGAT3 is not clear. Finally, the acyl-coenzyme A:diacylglycerol acyltransferase (DGAT1) reacts with the *sn*-1,2-diacylglycerols (not the *sn*-2,3 form) to produce triacylglycerols (DGAT1 can also acylate monoacylglycerols). Figure 21.2.8 shows the biosynthesis of triacylglycerols from monoacylglycerols.

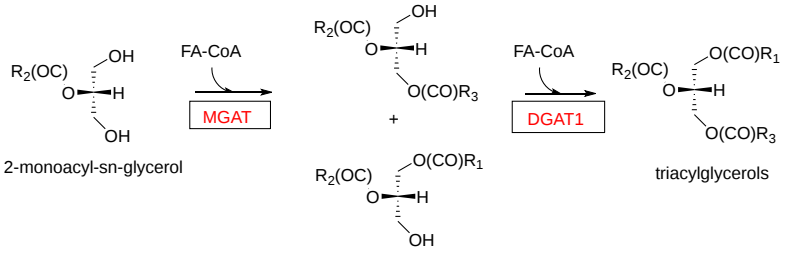


Figure 21.2.8: Biosynthesis of triacylglycerols from monoacylglycerols

Other pathways: In a fourth biosynthetic pathway, which is less well known, triacylglycerols are synthesized by a transacylation reaction between two racemic diacylglycerols that are independent of acyl-CoA. The reaction was first detected in the endoplasmic reticulum of intestinal micro villus cells and is catalyzed by a diacylglycerol transacylase. Both diacylglycerol enantiomers participate in the reaction with equal facility to transfer a fatty acyl group with the formation of triacylglycerols and a 2-monoacyl-*sn*-glycerol. A similar reaction has been observed in seed oils. Figure 21.2.9: Biosynthesis of triacylglycerols from diacylglycerol transacylases

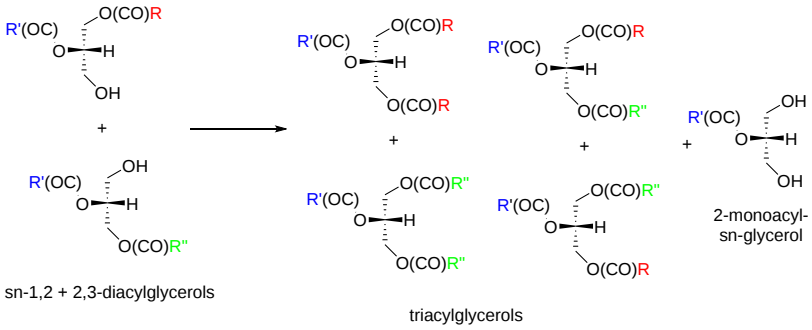


Figure 21.2.9: Biosynthesis of triacylglycerols from diacylglycerol transacylases

It has been suggested that this enzyme may function in remodeling triacylglycerols post-synthesis, especially in oil seeds, and it may be involved in similar processes in the liver and adipose tissue, where extensive hydrolysis/re-esterification is known to occur. There is evidence for selectivity in the biosynthesis of different molecular species in a variety of tissues and organisms, which may be a consequence of the varying biosynthetic pathways. Also in adipose tissue, fatty acids synthesized *de novo* are utilized in different ways from those from external sources in that they enter positions *sn*-1 and 2 predominantly, while a high proportion of the oleic acid synthesized in the tissue by desaturation of exogenous stearic acid is esterified to position *sn*-3.

Among other potential routes to the various intermediates, lysophosphatidic acid and phosphatidic acid can be synthesized in mitochondria, but they must then be transported to the endoplasmic reticulum before they enter the pathway for triacylglycerol production. 1,2-Diacyl-*sn*-glycerols are also produced by the action of phospholipase C on phospholipids and can be utilized for triacylglycerol biosynthesis.

In the glycerol-3-phosphate and other pathways, the starting material is of defined stereochemistry and each of the enzymes catalyzing the various steps in the process is distinctive and can have preferences for particular fatty acids (as their coenzyme A esters) and for particular fatty acid combinations in the partially acylated intermediates. It should not be surprising, therefore, that natural triacylglycerols exist in enantiomeric forms with each position of the *sn*-glycerol moiety esterified by different fatty acids.

21.2.3: TRIACYLGLYCEROL METABOLISM IN THE INTESTINES, LIVER, AND MAMMARY GLAND

Fat comprises up to 40% of the energy intake in the human diet in Western countries, and a high proportion of this is in the form of triacylglycerols. The process of fat digestion is begun in the stomach with acid-stable gastric or lingual lipases, the extent of which depends on species but may be important for efficient emulsification. However, this is insignificant in quantitative terms in comparison to the

reaction with pancreatic lipase, which occurs in the duodenum. Entry of triacylglycerol degradation products into the duodenum stimulates the synthesis of the hormone cholecystokinin and causes the gall bladder to release **bile acids**, which are strong detergents and act to emulsify the hydrophobic triacylglycerols so increasing the available surface area. In turn, cholecystokinin stimulates the release of the hydrolytic enzyme pancreatic lipase together with a co-lipase, which is essential for the activity of the enzyme. Pancreatic lipase, co-lipase, bile salts, and calcium ions act together in a complex at the surface of the emulsified fat droplets to hydrolyze the triacylglycerols. The process is regiospecific and results in the release of the fatty acids from the 1 and 3 positions and the formation of 2-monoacyl-*sn*-glycerols as shown in Figure 21.2.10.

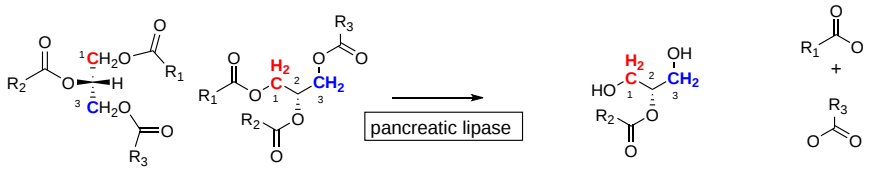


Figure 21.2.10: Hydrolysis of triacylglycerols in the intestines

Spontaneous isomerization of the latter to 1(3)-monoacyl-*sn*-glycerols occurs to some extent, and these can be degraded completely by the enzyme to glycerol and free fatty acids. Other lipases hydrolyze the phospholipids and other complex lipids in foods at the same time.

This process is somewhat different in neonates and young infants, in whom pancreatic lipase is less active but is effectively replaced by lipases in breast milk and by an acid gastric lipase (pH optimum 4-6).

There is evidence that the regiospecific structure of dietary triacylglycerols affects the uptake of particular fatty acids and may influence further lipid metabolism in humans. Incorporation of palmitic acid into the position *sn*-2 of milk fat may be of benefit to the human infant (as a source of energy for growth and development), although it increases the atherogenic potential for adults. In addition, 2-monoacylglycerols and 2-oleoylglycerol especially have a signaling function in the intestines by activating a specific G-protein coupled receptor GPR119, sometimes termed the 'fat sensor'. When stimulated, this causes a reduction in food intake and body weight gain in rats and regulates glucose-stimulated insulin secretion. The free fatty acids released have a similar effect, though by a very different mechanism, via the receptor GPR40. Overall, it has become evident that triacylglycerol metabolism in the intestine has regulatory effects on the secretion of gut hormones and subsequently on systemic lipid metabolism and energy balance.

The free fatty acids and 2-monoacyl-*sn*-glycerols are rapidly taken up by the intestinal cells, from the distal duodenum to the jejunum, via specific carrier molecules but also by passive diffusion. A specific fatty acid binding protein prevents a potentially toxic build-up of unesterified fatty acids and targets them for triacylglycerol biosynthesis. The long-chain fatty acids are converted to the CoA esters and esterified into triacylglycerols by the monoacylglycerol pathway as described above. In contrast, short and medium-chain fatty acids (C_{12} and below) are absorbed in unesterified form and pass directly into the portal blood stream, where they are transported to the liver to be oxidized.

Subsequently, the triacylglycerols are incorporated into **lipoprotein complexes** termed **chylomicrons** in the enterocytes in the small intestines. In brief, these consist of a core of triacylglycerols together with some cholesterol esters that is stabilized and rendered compatible with an aqueous environment by a surface film consisting of phospholipids, free cholesterol, and one molecule of a truncated form of apoprotein B (called apo B48). A cartoon structure of a chylomicron is shown in Figure 21.2.11.

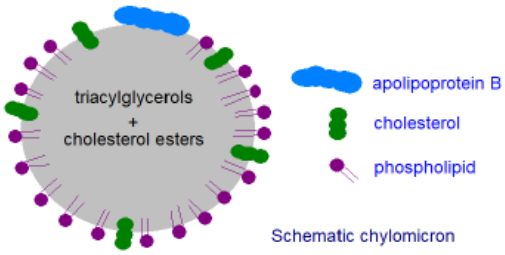


Figure 21.2.11: Cartoon structure of a chylomicron

These particles are secreted into the lymph and thence into the plasma for transport to the peripheral tissues for storage or structural purposes. Adipose tissue in particular exports appreciable amounts of the enzyme **lipoprotein lipase**, which binds to the luminal membrane of endothelial cells facing into the blood, where it rapidly hydrolyses the passing triacylglycerols at the cell surface releasing free fatty acids, most of which are absorbed into the adjacent adipocytes and re-utilized for triacylglycerol synthesis within the cell.

The chylomicrons remnants eventually reach the liver, where the remaining lipids are hydrolyzed at the external membranes by a hepatic lipase and absorbed. The fatty acids within the liver can be utilized for a variety of purposes, from oxidation to the synthesis of structural lipids, but a proportion is re-converted into triacylglycerols, and some of this is stored as lipid droplets within the cytoplasm of the cells (see

next section). In addition, phosphatidylcholine from the high-density lipoproteins is taken up by the liver, and a high proportion of this is eventually converted to triacylglycerols. In a healthy liver, the levels of triacylglycerols are low (<5% of the total lipids), because the rates of acquisition of fatty acid from plasma and synthesis *de novo* within the liver are balanced by rates of oxidation and secretion into plasma. On the other hand, excessive accumulation of storage triacylglycerols is associated with fatty liver, insulin resistance, and type 2 diabetes.

Most of the newly synthesized triacylglycerols are exported into the plasma in the form of very-low-density lipoproteins (VLDL), consisting again of a triacylglycerol and cholesterol ester core, surrounded by phospholipids and free cholesterol, together with one molecule of full-length apoprotein B (100 kDa), apoprotein C and sometimes apoprotein E. These particles in turn are transported to the peripheral tissues, where they are hydrolyzed and the free acids absorbed. Eventually, the remnants are returned to the liver.

In the **mammary gland**, triacylglycerols are synthesized in the endoplasmic reticulum and large lipid droplets are produced with a monolayer of phospholipids derived from this membrane. These are transported to the plasma membrane and bud off into the milk with an envelope comprised of the phospholipid membrane to form milk fat globules as food for the newborn. The process is thus very different from that involved in the secretion of triacylglycerol-rich lipoproteins from other organs.

21.2.4: TRIACYLGLYCEROL SYNTHESIS AND CATABOLISM (LIPOLYSIS) IN ADIPOCYTES AND LIPID DROPLETS

Adipose tissue and the adipocytes are characterized by accumulations of triacylglycerols, which act as the main energy store for animals, although they also cushion and insulate the body. Large fat depots occur around internal organs such as the liver, and also subcutaneously, and each of these may react differently to metabolic constraints. Thus, triacylglycerols stored when there is a surplus of nutrients are mobilized for energy production during starvation. Adipose tissue also functions as a reserve of bioactive lipids, such as eicosanoids and lipid-soluble vitamins, and when required provides structural components, including fatty acids, cholesterol, and retinol, for membrane synthesis and repair. By buffering against fatty acid accumulation that might exceed their capacity, non-adipose cells defend themselves in this way against lipotoxicity while providing a rapid source of energy and essential metabolites by sensing and responding rapidly to changes in systemic energy balance. Brown and beige fat have special properties and are discussed below, while bone marrow adipocytes (70% of the available space) have distinctive functions also.

Similarly, within most other animal cells, including most cell types in the brain, a proportion of the fatty acids taken up from the circulation is converted to triacylglycerols as described above and incorporated into cytoplasmic lipid droplets (also termed 'fat globules', 'oil bodies', 'lipid particles' or 'adiposomes'). In adipocytes, the lipid droplets can range from up to 200 μm in diameter, while other cell types contain smaller lipid droplets of the order of 50 nm in diameter. The triacylglycerol droplets together with cholesterol esters and other neutral lipids are surrounded by a protective monolayer that includes phospholipids, cholesterol, and hydrophobic proteins. The phospholipid component of the monolayer consists mainly of phosphatidylcholine and phosphatidylethanolamine derived from cytosolic leaflets of the endoplasmic reticulum and plasma membrane. Among the proteins are many that function directly in lipid metabolism, and they include acyltransferases, lipases, perilipins, caveolins, and the Adipose Differentiation Related Protein (ADRP or adipophilin). Acting in concert with other cellular organelles, they function in many different metabolic processes facilitating coordination and communication between different organelles and acting as vital hubs of cellular metabolism.

Cytosolic lipid droplets with similar metabolic activities are found in most eukaryotic cells, including those of the fruit fly *Drosophila melanogaster*, and many aspects of triacylglycerol processing and regulation parallel those in humans. They are also present in the cytoplasm of some prokaryotes and in the plastids and other organelles of plants (see below).

Lipid droplet assembly

This process takes place in sub-domains of the endoplasmic reticulum, where at least one isoform of each of the enzymes of triacylglycerol biosynthesis, from acyl-CoA synthetases through to glycerol-3-phosphate acyltransferases, is located probably in a protein assembly or 'interactome'. As triacylglycerols accumulate, they reach a critical level when a spontaneous condensation or nucleation by phase separation occurs, leading to the formation of an oil blister within the hydrophobic bilayer region so attracting perilipins and other proteins that allow lipid droplets to grow further in patches of the membrane as lens-like swellings between the two membrane leaflets.

The protein **seipin** stabilizes the nascent droplets with minimal disruption to the membrane and enables them to mature. Seipin monomers assemble into a decameric cage-like structure and sit in the ER to provide a space that permits triacylglycerol molecules to interact with each other, rather than with phospholipid acyl chains, a process that is probably aided by trans-membrane protein segments. This enables phase separation of the triacylglycerols, lens formation, and growth to a point where the seipin oligomer opens toward the cytoplasm so the lens can form a budding lipid droplet.

Figure 21.2.12 presents a cartoon that shows the role of ER membrane seipins and other proteins in the assembly of lipid droplets.

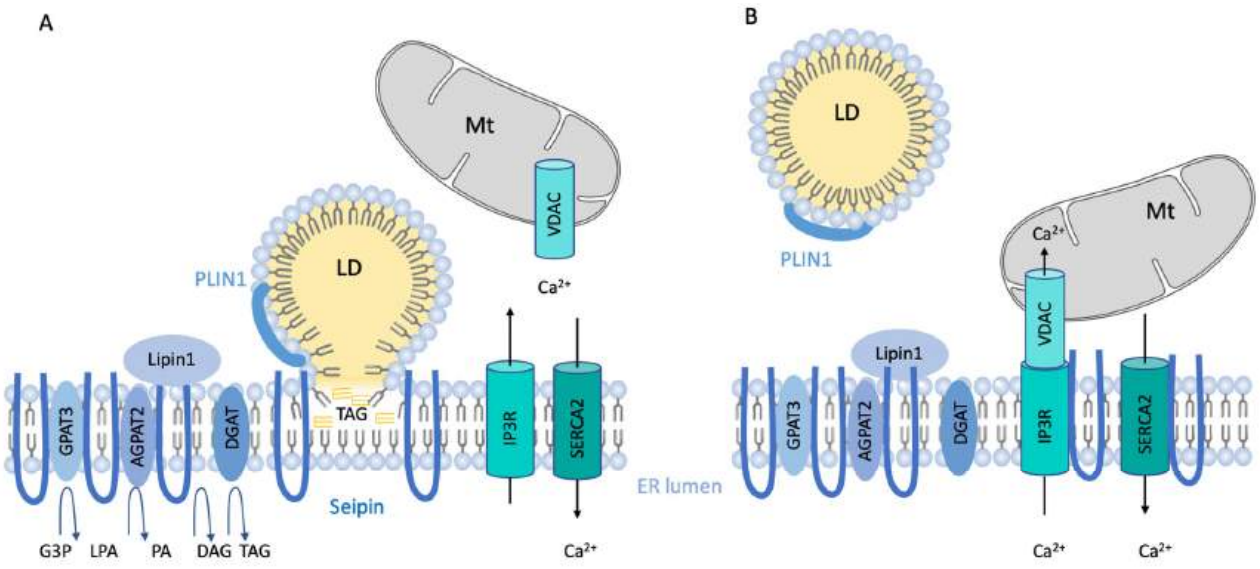
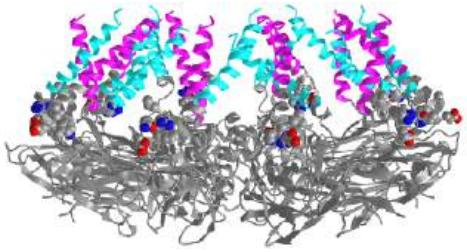


Figure 21.2.12 Seipin function in mature adipocytes. Magré J, Prieur X. Seipin Deficiency as a Model of Severe Adipocyte Dysfunction: Lessons from Rodent Models and Teaching for Human Disease. Int J Mol Sci. 2022 Jan 11;23(2):740. doi: 10.3390/ijms23020740. Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/

In panel (A), seipin is enriched at the endoplasmic reticulum (ER)/lipid droplet (LD) contact sites and is crucial in triglycerides (TAG) flow from the ER to the LD. Seipin interacts with Perilipin1 (PLIN1) and with several TAG synthetic enzymes such as glycerol-3-phosphate acyltransferase (GPAT3), 1-acyl-sn-glycerol-3-phosphate acyltransferase beta (AGPAT2), and LIPIN. No interaction with diacylglycerol acyltransferases (DGAT) has been formally reported. Magré J et al, Int J Mol Sci. 2022 Jan 11;23(2):740. doi: 10.3390/ijms23020740. Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/

In panel (B), representing the fasting state, seipin is enriched at the ER/mitochondria (Mt) contact sites, also named mitochondria-associated membranes (MAM) and participates in ER/mitochondria calcium (Ca^{2+}) flux and mitochondrial activity. Seipin is near MAM Ca^{2+} regulators IP3(inositol 1,4,5-trisphosphate) receptor (IP3R), voltage-dependent anion channel (VDAC), and sarco-/ER Ca^{2+} ATPase 2 (SERCA2) as well as glycerol-3-phosphate (G3P), lysophosphatidic acid (LPA), phosphatidic acid (PA), and diacylglycerol (DAG).

Figure 21.2.13 shows an interactive iCn3D model of the S. cerevisiae seipin Homo 10-mer flexible cage at lipid droplet formation sites (7RSL).



NCBI iCn3D Figure 21.2.13: S. cerevisiae seipin homo 10-mer flexible cage at lipid droplet formation sites (7RSL). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...BjYLiQXVJjBpk8

The N-terminal transmembrane segment (TMS) on each monomer is shown in cyan while the C-terminal TMS is shown in magenta. These form the cage sides and top and anchor the complex in the ER membrane. The side chains of a conserved switch sequence (F_{232}xxGLR) on each monomer are shown in spacefill, CPK colors. The gray ribbons represent the part of the complex that is in the lumen of the ER.

Adjacent monomers in the homo 10-mer cage, even though they are identical in sequences, adopt different orientations for two helices that enable a binary switch in the conformation of the complex. Two adjacent subunits, the A chain in blue and the B chain in orange, are shown, along with their overlap, in Figure 21.2.14.

Figure 21.2.14: Conformational flexibility in adjacent monomers in the seipin homo 10-mer.

A highly conserved "switch" region is shown in green with its highly conserved motif F_{232}xxGLR shown as sticks with CPK colors. Note the large conformational change in the switch region which occurs with a large conformational change of the two helices. This suggests that there are two overall conformations of the complex, a closed form allowing accumulation of TAGs and an open form which allows the budding and release of lipid droplets. A possible model for the formation and budding of lipid droplets is illustrated in Figure 21.2.15. As in the above figure, the A (open) chain is shown in blue while the B (closed) chain is shown in orange.

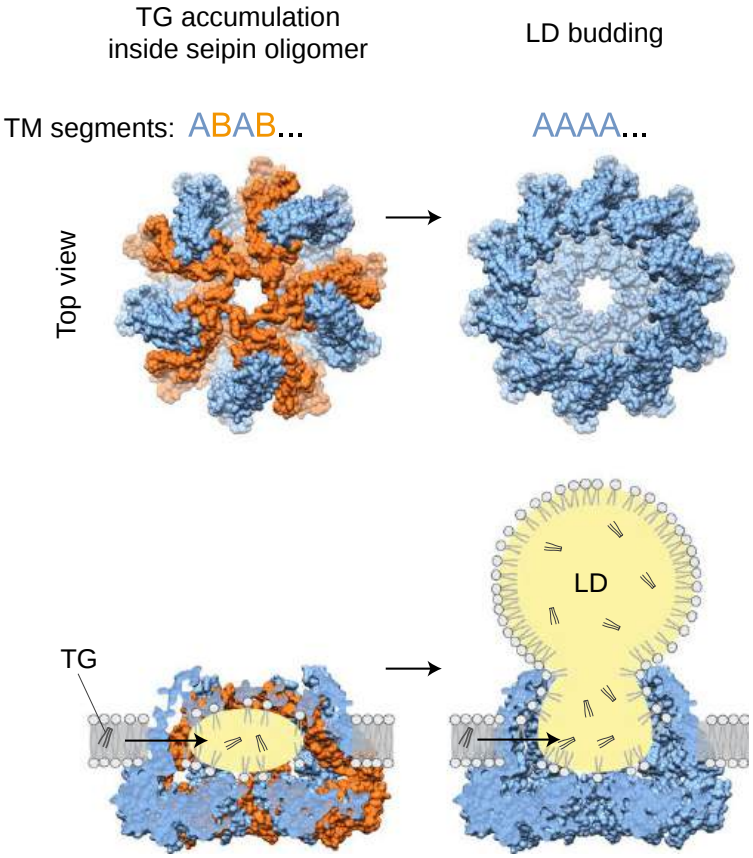


Figure 21.2.15: Accumulation and phase separation of triacylglycerides (TG) and budding of lipid droplets (LD) from the seipin complex. Arlt et al. https://doi.org/10.1038/s41594-021-00718-y. Nature Structural & Molecular Biology, 194–202 (2022). Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0l. The left-hand side shows alternating monomer conformations (A in blue and B in orange, based on the solved structure), which forms an overall "closed" form of the complex, allowing the accumulation of TGs. The right-hand side shows a possible open form, with all monomers in the A or blue form, which allows lipid droplet formation and release.

Some of the surface proteins on lipid droplets can extend long helical hairpins of hydrophobic peptides deep into the lipid core. For example, perilipins constitute a family of at least five phosphorylated proteins that bind to droplets in animals and share a common region, the so-called ‘PAT’ domain, named for the three original members of the family that include perilipin and ADRP. Proteins related evolutionarily to these are found in more primitive organisms, including insects, slime molds, and fungi, but not in the nematode *Caenorhabditis elegans*. In mammals, perilipin A (or 'PLIN1' or more accurately the splice variant 'PLIN1a') is a well-established regulator

of lipolysis in adipocytes, and it is believed to be involved in the formation of the large lipid droplets in white adipose tissue. The perilipins PLIN1 and PLIN2 have functions in triacylglycerol metabolism in tissues other than adipocytes, and PLIN2 is the main perilipin in hepatocytes; PLIN5 operates in tissues that oxidize fatty acids such as the heart. Other surface proteins of lipid droplets are enzymes intimately involved in triacylglycerol metabolism, although there is a suggestion that cytoplasmic droplets may act as a storage organelle for hydrophobic proteins whose function is elsewhere in the cell.

Lipolysis

When fatty acids are required by other tissues for energy or other purposes, they are released from the triacylglycerols by the sequential actions of three cytosolic enzymes at neutral pH, i.e., adipose triacylglycerol lipase (ATGL), hormone-sensitive lipase (HSL) and monoacylglycerol lipase, which cycle between the cytoplasmic surfaces of the endoplasmic reticulum and the surface layer of lipid droplets. Simplistically, ATGL hydrolyses triacylglycerols to diacylglycerols, which are hydrolyzed by HSL to monoacylglycerols before these are hydrolyzed by the monoacylglycerol lipase to complete the process. Lipolysis proceeds in a highly ordered manner with stimulation through cell-surface receptors via neurotransmitters, hormones, and autocrine/paracrine factors that activate various intracellular signaling pathways and increase kinase activity. A protein perilipin (PLIN1) has been described as "the gatekeeper of the adipocyte lipid storehouse" that regulates lipolysis by acting as a barrier to lipolysis in non-stimulated cells. However, on β-adrenergic stimulation during fasting, it is phosphorylated by the cAMP-protein kinase, which changes its shape and reduces its hydrophobicity, and in the process activates lipolysis. An isoform, perilipin A, is the main regulatory factor in white adipose tissue. However, many other proteins interact with the three enzymes to modulate their activity, location, and stability.

The **adipose triacylglycerol lipase**, which initiates the process, was discovered surprisingly recently. It is structurally related to the plant acyl-hydrolases in that it has a patatin-like domain in the NH_2-terminal region (patatin is a non-specific acyl-hydrolase in potatoes). Specific transport mechanisms guide ATGL from the endoplasmic reticulum membrane to lipid droplets, where it is located on the surface both in the basal and activated states. This lipase is specific for triacylglycerols containing long-chain fatty acids, preferentially cleaving ester bonds in the *sn*-1 or *sn*-2 position (but not *sn*-3), and it yields diacylglycerols and free fatty acids as the main products, with low activity only towards diacylglycerols, and none to monoacylglycerols and cholesterol esters. However, it also has transacylase and phospholipase activities, and it hydrolyses retinol esters in hepatic stellate cells. Adipose triacylglycerol lipase can be activated at the same time as hormone-sensitive lipase and is now believed to be rate-limiting for the first step in triacylglycerol hydrolysis. Figure 21.2.16 shows the sequential hydrolysis of TAGs in adipocytes and lipid droplets.

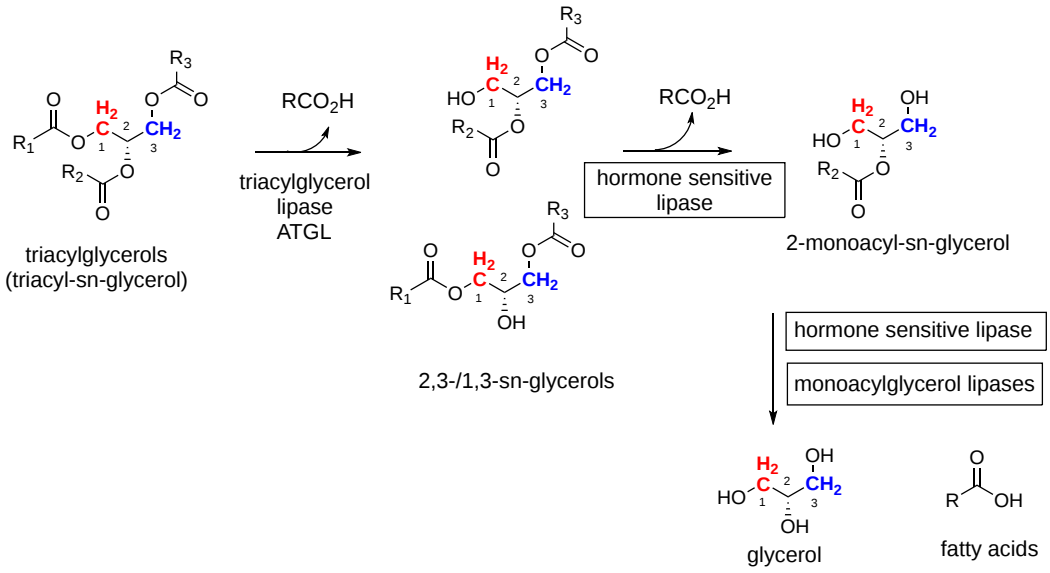


Figure 21.2.16: Sequential hydrolysis of TAGs in adipocytes and lipid droplets

Regulation of the enzymatic activity is a complex process, and for example, a lipid droplet protein designated Gene identification-58 (CGI-58 or ABHD5), is known to be an important activating factor and is required for hydrolysis of fatty acids from position *sn*-1. In the resting state, this protein binds to perilipin (PLIN1), but on hormonal stimulation, the latter is phosphorylated leading to dissociation and interaction of CGI-58 with phosphorylated ATGL to commence the first step in triacylglycerol hydrolysis. Mutations in adipose triacylglycerol lipase or CGI-58 are believed to be responsible for a syndrome in humans known as ‘neutral lipid storage disease’. A second protein (G0S2) inhibits the enzyme.

Hormone-sensitive lipase in various isoforms is a structurally unique member of the large Ser-lipase/esterase family of enzymes in animals. It is regulated by the action of the hormones insulin and noradrenaline by a mechanism that ultimately involves phosphorylation of the enzyme by cAMP-protein kinase (as with perilipin), thereby increasing its activity and causing it to translocate from the cytosol to the

lipid droplet to initiate the second step in hydrolysis. Its activity is regulated further by a variety of proteins that include PLINs and fatty acid binding proteins (FABP). Hormone-sensitive lipase has a broad substrate specificity compared to other neutral lipases, and in addition to its activity towards triacylglycerols, it will rapidly hydrolyze diacylglycerols, monoacylglycerols, retinol esters, and cholesterol esters. Diacylglycerols are hydrolyzed ten times as rapidly as triacylglycerols. Within the triacylglycerol molecule, hormone-sensitive lipase preferentially hydrolyses ester bonds in the *sn*-1 and *sn*-3 positions, leaving free acids and 2-monoacylglycerols as the main end products.

The **monoacylglycerol lipase** is believed to be the rate-limiting enzyme in lipolysis, i.e., the final step in triacylglycerol catabolism releasing free glycerol and fatty acids, and it is found in the cytoplasm, the plasma membrane, and in lipid droplets. It is specific for monoacylglycerols, but with no positional specificity, and has no activity against di- or triacylglycerols. As it is the enzyme mainly responsible for the deactivation of the endocannabinoid **2-arachidonylglycerol** and is highly active in malignant cancers, it is attracting pharmaceutical interest. A further enzyme, α/β hydrolase containing-6 (ABHD6), is located on the inner leaflet of the plasma membrane and preferentially hydrolyses fatty acids at the *sn*-1 position over the *sn*-2 position of monoacylglycerols, and it also hydrolyses lysophospholipids and bis(monoacylglycerol)-phosphate. This has been associated with the development of insulin resistance and the progression of cancer. Additional lipolytic enzymes, including carboxyesterases, are believed to operate against triacylglycerols in cytoplasmic lipid droplets in the liver.

Unesterified fatty acids released by the combined action of these three lipases are exported into the plasma for transport to other tissues in the form of albumin complexes, while the glycerol released is transported to the liver for metabolism by either glycolysis or gluconeogenesis.

Eventually, the whole organelle can disappear, including the proteins, when they undergo a process of autophagy (**lipophagy**), i.e., the delivery of the organelles to lytic compartments for degradation. This can occur through direct lysosomal invagination or more often by a multistep process involving the formation of double-membrane vesicles termed 'autophagosomes' around droplets with subsequent lysosomal fusion and degradation of the triacylglycerols by the lysosomal acid lipase. This process is important for the regulation of cellular lipid levels in various tissues and disease conditions, especially during starvation. It is relevant to tumorigenesis and cancer metastasis, and to neurodegenerative and neuroinflammatory diseases. While lipophagy is mechanistically distinct from lipolysis, there is cross-talk between the two.

THE STRUCTURE OF LIPASES

Lipases carry out hydrolysis reactions (nucleophilic substitution with water as the nucleophile), alcoholysis reactions (with an ROH as the nucleophile), and even transesterification reactions with other fatty acid esters. Lipases are critical enzymes yet there are no good structures available for hormone-sensitive lipase for example. Yet much is known about the structure and activity of lipases. Lipases have a catalytic Ser-His-Asp triad like serine proteases, which also carry out hydrolysis reactions. However, the substrates for lipase are nonpolar and could be part of larger structures such as lipid droplets. As such, the lipases work at the interface between aqueous and non-aqueous interfaces. They are **interfacially** activated.

Lipases appear to have low activity in aqueous solutions but more in solutions that are more nonpolar. Many appear to have two conformations, a closed one with lower activity and poor lipid substrate binding properties, and an open one with higher activity and activation by nonpolar solvents. An alpha-helical lid appears to close off the active site in the closed conformation. In a more nonpolar solvent or the presents of bulk lipids, such as lipid droplets, the alpha-helical lid moves exposing the enzyme's hydrophobic binding site to substrate interactions.

Figure 21.2.17 shows an interactive iCn3D model comparing the structure of a *Thermomyces lanuginosus* lipase in its closed (1DT3) and its open conformation (1EIN).

NCBI iCn3D Figure 21.2.17: The structure of a *Thermomyces lanuginosus* lipase in its closed (1DT3) and its open conformation (1EIN). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?C9yEwy6SR6MZ9Q1S7

The closed structure (1DT3) is shown in magenta and the open (1EIN) is in cyan. The active site triad Ser146, His258, and Asp201 is shown in CPK-colored sticks and labeled for each. The lid is shown in yellow with its side chains as sticks. Toggle between the two states using the "a" key. Note how the lid covers the active site in the closed (1DT3) structure. You will have to rotate the molecule to see the optimal orientation. Also, close out menu boxes for a larger and more centered display.

Endocrine functions

Not only does the adipocyte provide a store of energy but it is an endocrine organ that manages the flow of energy through the formation of the hormone **leptin**, which signals through the adiponectin receptors (AdipoR)1 and AdipoR2 and binds to the non-signaling interacting protein, T-cadherin. The result is the stimulation of signaling cascades to communicate with other tissues through the secretion of cytokines and other mediators. The synthesis of leptin is tightly controlled by adipocytes mainly, although it is also produced by the stomach, placenta, and mammary gland, and its principal function is the provision of information on the state of fat stores to other tissues. In so doing, it regulates food intake and energy expenditure. Leptin was initially described as an anti-obesity hormone, but it serves as an adiposity signal with a vital function in maintaining adipose tissue mass to ensure survival under conditions of negative energy balance and so protect against either a deficit or an excess of adiposity. Lipid droplets have a role in this process, sinc perilipin is required for the sensing function.

Adiponectin is a potent insulin sensitizer and suppressor of cell death and inflammation, directly promoting anti-diabetic and anti-atherosclerotic outcomes. Insulin is the main hormone that affects metabolism and its receptor at the plasma membrane is located in caveolae. The release of proinflammatory cytokines can stimulate lipolysis and cause insulin resistance, in turn leading to dysfunction of adipose tissue and systemic disruption of metabolism. Thus, adipose tissue metabolism has profound effects on whole-body metabolism, and defects in these processes can have severe implications for such serious pathological conditions as diabetes, obesity, cardiovascular disease, fatty liver disease, and cancer in humans. It is hoped that the development of specific inhibitors for hormone-sensitive lipase will improve the treatment of such metabolic complications. As **caveolae**, which contain the proteins caveolins (and presumably sphingolipids) and are particularly abundant in adipocytes, modulate the flux of fatty acids across the plasma membrane and are involved in signal transduction and membrane trafficking pathways, it is evident that they have a major role in this aspect of lipid metabolism.

White fat acts as an endocrine organ and can release a variety of hormones including **adipokines** (analogous to cytokines, chemokines, lymphokines) **adipsin,** and **leptin** as well as tumor necrosis factor α (TNF-α), adiponectin, resistin, and RBP4.

Functions other than energy management

Lipid droplets accumulate within many cell types other than adipocytes, including leukocytes, epithelial cells, hepatocytes, and even astrocytes, especially during infections, cancer, and other inflammatory conditions. They are important for the cellular storage and release of hydrophobic vitamins, signaling precursors, and other lipids that are not related to energy homeostasis while reducing the dangers of lipotoxicity. On the other hand, excessive fatty acid accumulation is associated with lipotoxicity, endoplasmic reticulum stress, and mitochondrial damage and dysfunction, so lipid storage in lipid droplets must be balanced for health.

A variety of enzymes are associated with lipid droplets, including protein kinases, which are involved in many different aspects of lipid metabolism, such as cell signaling, membrane trafficking, and control of the production of inflammatory mediators like the eicosanoids. Lipolysis enables the secretion of lipid species termed **lipokines** (more generally 'adipokines') from adipocytes that may signal in a hormone-like fashion to other tissues, thereby modulating gene expression and physiological function, including food intake, insulin sensitivity, insulin secretion, and related processes. These include **palmitoleic acid (9-16:1)** and **fatty acid esters of hydroxy fatty acids (FAHFA)**, though the circulating proteins adiponectin and leptin have been studied more intensively. Adiponectin is a powerful insulin sensitizer and suppressor of apoptosis and inflammation with anti-diabetic and anti-atherosclerotic functions, often operating through its effects on sphingolipids, while leptin exerts most of its effects on the brain to trigger behavioral, metabolic, and endocrine responses to control the body's fuel reserves.

Indeed, there are now suggestions that lipid droplets in all cell types are essential for the response mechanisms to cellular stress, including autophagy, inflammation, and immunity, and act as hubs to integrate metabolic and inflammatory processes. Via their lipolytic machinery, they regulate the availability of fatty acids for the activation of signaling pathways and the production of oxylipins from polyunsaturated fatty acids. For example, triacylglycerols in cytoplasmic lipid droplets of human mast cells, which are potent mediators of immune reactions and influence many inflammatory diseases, have a high content of arachidonic acid and this can be released by adipose triacylglycerol lipases as a substrate for production of specific eicosanoids when the cells are stimulated appropriately. Active metabolism in lipid droplets is important for the differentiation of monocytes, and it is essential for the sustained functional activity of differentiated macrophages, especially in relation to inflammation.

During apoptosis, triacylglycerols enriched in polyunsaturated fatty acids accumulate in lipid droplets, possibly as a protective mechanism against membrane damage caused by oxidative stress and hydroperoxide formation in this process. Triacylglycerols in lipid droplets of the skin are a highly specific source of linoleic acid that is required for the formation of the ***O*-acylceramides**, which are essential for epidermal barrier function. An organelle termed the midbody in dividing cells in humans and rodents contains a unique triacylglycerol that is a single molecular species consisting of three fatty acids 16:1-12:0-18:1 (12:0 especially is rarely detected in human lipids), but its function is not known.

Vitamin E (tocopherols) and vitamin A in the form of retinyl esters are stored in cytoplasmic lipid droplets, and the latter are present in appreciable concentrations in the stellate cells of the liver, for example. In endocrine cells of the gonads and adrenals, cholesterol esters stored in lipid droplets are an important source of cholesterol for the mitochondrial biosynthesis of various steroid hormones. In the nucleus

of the cell, in addition to providing a reservoir of fatty acids for membrane remodeling, lipid droplets can sequester transcription factors and chromatin components and generate the lipid ligands for certain nuclear receptors.

In addition to their role in lipid biochemistry, lipid droplets participate in protein degradation and glycosylation. Their metabolism can be manipulated by pathogenic viruses and bacteria such as *Mycobacterium tuberculosis* with unfortunate consequences for the host, but they also serve as reservoirs for proteins that fight intracellular pathogens. In consequence, such lipid droplets and their enzyme systems may be markers for disease states and are also considered to targets for pharmaceutical intervention.

Insects

In insects, the fat body is a multifunctional tissue that is the main metabolic organ. It integrates signals that control the immune system, molting, metamorphosis, and synthesis of hormones that regulate innumerable aspects of metabolism. In fat body cells, lipids, carbohydrates, and proteins are the substrates and products of many pathways for use in energy production or to act as reserves for mobilization at the appropriate stage of life (diapause, metamorphosis, and flight). In relation to innate and acquired humoral immunity, the fat body produces bactericidal proteins and polypeptides, i.e., lysozyme. It is also important in the early stages of an insect's life due to the production of vitellogenin, the yolk protein needed for the development of oocytes.

21.2.5: BROWN ADIPOSE TISSUE

Most adipose tissue depots ('white fat') serve primarily as storage and endocrine organs that provide a reservoir of nutrients for release when the food supply is low. However, a second specialized form of adipose tissue, brown fat, is multilocular, highly vascularized, and rich in mitochondria and the iron-containing pigments that transport oxygen and give the tissue its color and name. Brown adipocytes arise from progenitor cells that are closer to those of skeletal muscle than white adipocytes. In humans, these depots tend to be located in specific anatomical regions such as subcutaneous areas around the neck, where their function may be to supply warm venous blood directly to the spinal cord and brain, and elsewhere to the heart, kidney, pancreas, and liver. Brown fat can oxidize fat so rapidly that heat is generated (“non-shivering thermogenesis”), and it is especially important in young animals and those recovering from hibernation.

In brief, during cold exposure, the release of noradrenaline and stimulation of β-adrenergic receptors in the nervous system initiates a catabolic program that commences with a rapid breakdown of cellular triacylglycerol stores and release of unesterified fatty acids and transient activation of a co-activator of peroxisome proliferator-activated receptor gamma (PPARγ). These set in motion a signaling process that results in the efficient **β-oxidation** of fatty acids to produce heat. The key molecule is believed to be the **uncoupling protein-1 (UCP1)**, which acts as a valve to uncouple electron transport in the respiratory chain from ATP production with a highly exothermic release of chemical energy, i.e., as heat rather than as ATP. Its actions are illustrated in Figure 21.2.18.

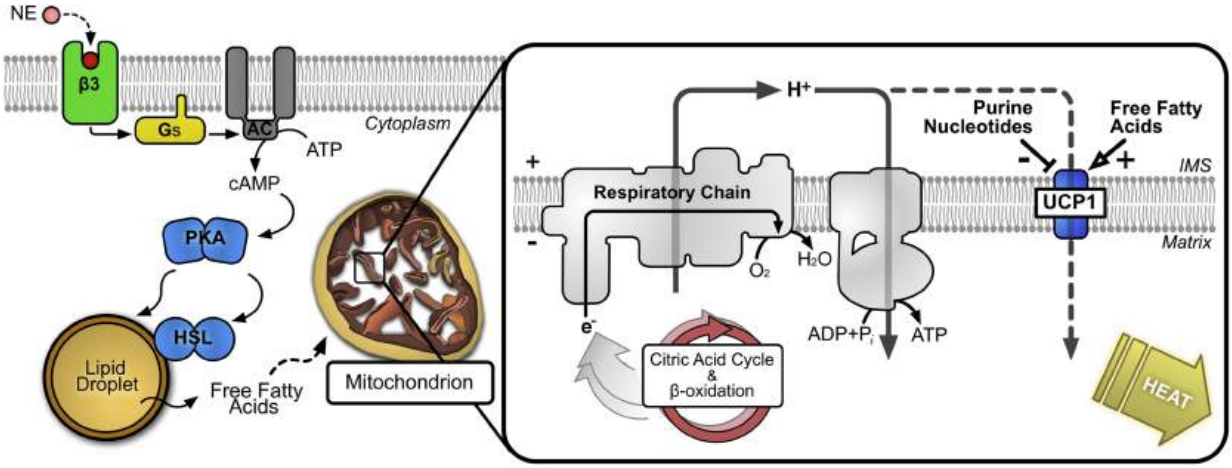


Figure 21.2.18: Role of UCP1 in generation of heat. https://commons.wikimedia.org/wiki/F...n_the_cell.jpg. Creative Commons Attribution-Share Alike 4.0 Internationa

This is an example of a futile cycle that releases heat. Although many aspects of the mechanism are uncertain, it is clear that proton conductance by UCP1 is highly regulated and inducible. It is activated by free long-chain fatty acids and inhibited by purine nucleotides, i.e., fatty acids are not only the substrate for thermogenesis but act also as self-regulating second messengers. The mitochondrial phospholipid **cardiolipin**, which is intimately involved in oxidative phosphorylation, is indispensable for stimulating and sustaining the function of thermogenic fat. Upon activation of brown adipose tissue, the dense vasculature increases the delivery of fatty acids and glucose to the brown adipocytes and warms the blood passing through the tissue.

Long-chain fatty acids bind to UCP1 and may be transported through the membrane along with a proton. Alternatively, it may just bind and activate the transport of a proton. Figure 21.2.19 shows an interactive iCn3D model of the AlphaFold predicted structure of human UCP1 (P25874).

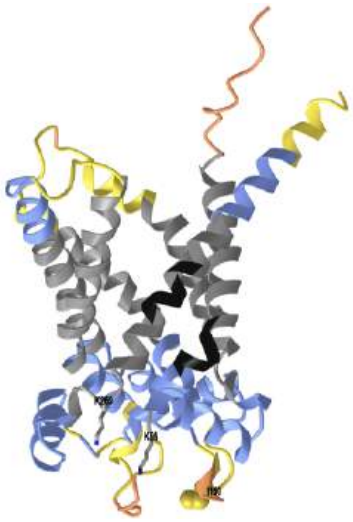


Figure 21.2.19: AlphaFold predicted structure of human UCP1 (P25874). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...FkfrEqWE3D2dX6

Imagine a bilayer orienting perpendicularly to the gray helices representing the transmembrane helical segments that imbed the protein into the inner mitochondrial membrane. The spacefill Ile150 denotes the matrix domain. Two key residues (K56, K29) involved in the binding of the carboxyl end of the fatty acids are shown on the matrix side of the protein is CPK-colored sticks and labeled. The black faces of the helices in the inner membrane region show the most probable binding site for the acyl chain of the fatty acid. The lighter yellow/orange regions have greater predicted uncertainty in the AlphaFold computational model.

There is evidence that **acylcarnitines** produced in the liver from fatty acids released from white adipose tissue in response to cold exposure are transported in plasma to brown adipose tissue and can serve as a substrate for thermogenesis. Indeed, a wide array of circulating lipids contributes to thermogenic potential, including free fatty acids and triacylglycerols.

The activities of acyl-CoA synthetases and acyl-CoA thioesterases determine the availability of substrates for β-oxidation and consequently the thermogenic capacity. Synthesis of the lipokine (or 'batokine') **octadecanoid** 12,13-dihydroxy-9*Z*-octadecenoic acid (12,13-diHOME), is induced by cold also, and this stimulates the activity of brown adipose tissue by promoting the uptake of fatty acids, acting via G-protein-coupled receptors. It increases cardiac function and cardiomyocyte respiration via enhanced calcium cycling. Similarly, there are suggestions that *n*-3 polyunsaturated fatty acids may promote adaptive thermogenesis, for example through the activity of the 12-lipoxygenase metabolite and batokine 12-hydroxyeicosapentaenoic acid (12-HEPE) by improving glucose metabolism via increased glucose uptake into adipocytes and skeletal muscle. (**FAHFA**) are relevant in this context. Peroxisomal synthesis of **plasmenyl-phospholipids** is believed to regulate adipose tissue thermogenesis by mediating mitochondrial fission.

In hibernating mammals, brown adipose tissue is especially important metabolically, and even in laboratory animals such as mice, it can consume about 50% of dietary lipids and glucose when the animals are exposed to cold. Similarly, in even humans, the activity of brown adipose tissue is induced acutely by cold and is stimulated via the sympathetic nervous system, and the relevance of this tissue to human metabolism is now becoming apparent. For example, there are suggestions that brown adipose tissue can behave as an endocrine system to secrete endocrine factors ('batokines') that may be favorable in reducing cardiovascular risk. For obvious reasons, there are efforts to determine whether sustained activation of brown fat by pharmaceutical means could be beneficial for human disease states, including obesity, diabetes, and cardiovascular disease.

Such research has been stimulated by the observation that clusters of distinct adipocytes with thermogenic capacity in addition to their storage function can be present in white adipose tissue and emerge in response to various physiological signals, especially reactive oxygen species. They are termed '**beige** or brite' adipocytes and arise from multipotent pre-adipocytes. Brown adipose tissue is a discrete organ in animals, but beige adipose tissue is interspersed with white, and the two forms have different developmental origins. In adult humans, most of the deposits once thought to be classical brown adipose tissue do not contain the genetic markers for this tissue and are now recognized to be beige/brite fat, although brown adipose tissue *per se* is present in significant amounts in human newborns and infants. However, beige adipocytes utilize the same machinery to release heat by oxidation of fatty acids under β-adrenergic stimulation.

The uncoupler protein 1 (UCP1) appears to be expressed only in brown and beige fat cells and is a marker for those cells. It can be regulated transcriptionally and posttranscriptionally by ROS modification of a cysteine to produce a sulfenyl group. The more classical brown adipocyte arises developmentally from the skeletal muscle-like cell, whereas the beige appears to develop to some extent from vascular smooth muscle cells. The beige cells can differentiate into more white adipocytes or more brown ones, by the factors described in Figure 21.2.20.

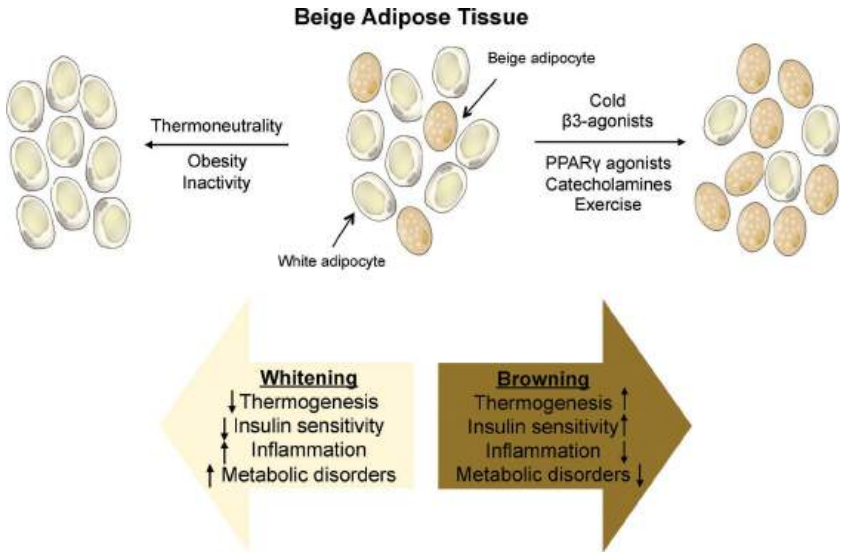


Figure (\PageIndex{20}\): Factors altering beige adipose tissue. Paul Cohen and Bruce M. Spiegelman. **Molecular Biology of the** Cell, 27, 2017. https://doi.org/10.1091/mbc.e15-10-0749. Attribution–Noncommercial–Share Alike 3.0 Unported Creative Commons License (http://creativecommons.org/licenses/by-nc-sa/3.0).

FAT CELL SIZE

In addition to the type of fat cell, the location and size of the adipocyte also affect health risks. Obesity, cardiovascular disease, type II diabetes, and nonalcoholic fatty liver disease all depend on and are correlated with fat cell dysfunction. The endocrine role of adipose tissue plays a key role in the development of dysfunction. The main depots of fat are the subcutaneous adipose tissue (SAT) and the visceral adipose tissue (VAT), the latter of which is associated with proinflammatory effects and associated greater health risks (type II diabetes, hypertension, metabolic syndrome, obesity, etc.). Figure (\PageIndex{21}\) below summarizes the health consequence as the size of the SAT and VAT adipocytes increase (in the direction of the red arrow).

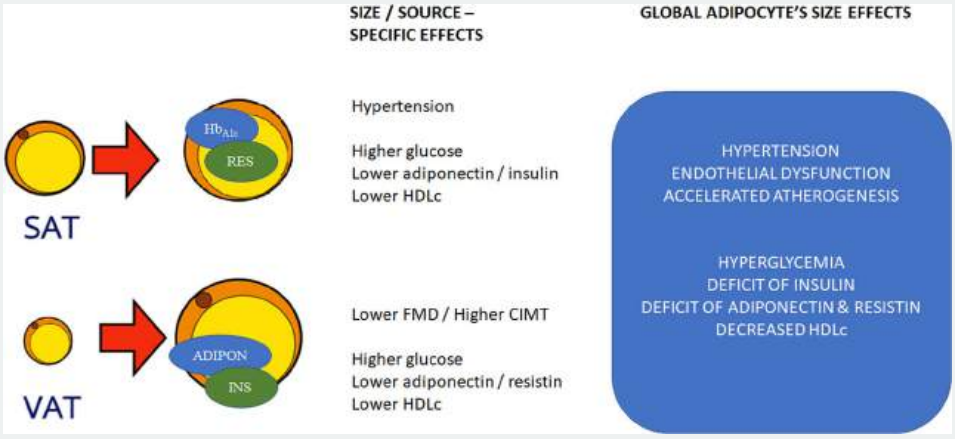


Figure (\PageIndex{2}\): Adipocyte characteristics and cardiovascular risk. The diagram describes a proposed model, based on the results of adipocyte size/source, potential interactions with metabolic mediators, and pathophysiological effects. SAT, subcutaneous adipose tissue; VAT, visceral adipose tissue; HbA1c, glycated hemoglobin (a marker of elevated blood sugar associated with diabetes); RES, resistin; ADIPON, adiponectin; INS, insulin; HDLc; high-density lipoprotein cholesterol; FMD, flow-mediated dilation; CIMT, carotid intima media thickness. Suárez-Cuenca et al. *Sci Rep* **11**, 1831 (2021). https://doi.org/10.1038/s41598-021-81289-2. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

It appears that the larger the adipocyte, the greater the inflammatory response. This condition is called adipocyte hypertrophy and it may be more important than obesity per se as a risk factor for inflammation, as indicated by inflammatory cytokine levels and macrophage infiltration into adipose tissue. The adipose-derived hormone leptin, when released under normal conditions, leads to

appetite suppression and also decreases liver and muscle fat. Higher fat storage is associated with increased blood leptin levels. It also correlates with greater lymphocyte size.

In contrast, the adipose hormone adiponectin leads to positive health consequences including increased sensitivity to insulin, decreased inflammation, and the formation of adipocytes from stem cells (adipogenesis). The larger the fat cell, the lower the serum levels of adiponectin. However, the link between fat cell size and adiponectin levels may be determined more by obesity.

21.2.6: OTHER FUNCTIONS OF TRIACYLGLYCEROL DEPOTS

Subcutaneous depots act as a cushion around joints and serve as insulation against the cold in many terrestrial animals, as is obvious in the pig, which is surrounded by a layer of fat, and it is especially true for marine mammals such as seals. Those adipocytes embedded in the skin differ from the general subcutaneous depots and support the growth of hair follicles and regenerating skin, and they may also have a defensive role both as a physical barrier and by responding metabolically to bacterial infection.

In marine mammals and fish, the fat depots are less dense than water and so aid buoyancy with the result that less energy is expended in swimming. More surprisingly perhaps, triacylglycerols together with the structurally related glyceryl ether diesters and wax esters are the main components of the sonar lens used in echo location by dolphins and toothed whales. The triacylglycerols are distinctive in that they contain two molecules of 3-methylbutyric (isovaleric) acid with one long-chain fatty acid. It appears that the relative concentrations of the various lipids in an organ in the head of the animals (termed the ‘melon’) are arranged anatomically in a three-dimensional topographical pattern to enable them to focus sound waves.

In cold climates, many insects do not feed over winter and must manage their energy stores to meet the energetic demands of development and reproduction in the spring. Some insect species that are tolerant of freezing produce triacylglycerols containing acetic acid, and these remain liquid at low temperatures; by interacting with water, they may play a role in cryoprotection.

21.2.7: TRIACYLGLYCEROL METABOLISM IN PLANTS AND YEASTS

Fruit and seed oils are major agricultural products with appreciable economic and nutritional value to humans. The mesocarp of fruits is a highly nutritious energy source that attracts animals that help to disperse the seeds, and in plants such as the oil palm and olive trees a high proportion of the fruit flesh contains triacylglycerols. Similarly in seeds, triacylglycerols are the main storage lipid and can comprise as much as 60% of their weight. Fruit lipids are not intended for use by the plant *per se* and are stored in lipid droplets in large irregular structures that break down readily, but seed lipids are required for the development of the plant embryo, so their metabolism is of particular importance. However, triacylglycerol biosynthesis and metabolism are required also for pollen viability and to maintain lipid homeostasis in chloroplasts (see the note on plastoglobules below).

Seed oils: In seeds and other plant tissues, biosynthesis of fatty acids takes place in plastids, and these are stored in the form of triacylglycerols in lipid droplets with a coherent surface layer of proteins and lipids in the embryo (e.g., Arabidopsis, soybean or sunflower) or endosperm (e.g., castor bean) tissues of seeds. In addition to the common range of fatty acids synthesized in plastids, mainly palmitate, and oleate, some plant species produce novel fatty acids, including medium- and very-long-chain components and those with oxygenated and other functional moieties. A specific means of diverting these to seeds for triacylglycerol production exists to prevent disruption of the plant membranes. Seed development occurs in three stages - rapid cell division with no accumulation of storage material, rapid deposition of triacylglycerols and other energy-rich metabolites, and finally desiccation.

During the period of oil accumulation in seeds, the newly formed ACP esters of fatty acids are first hydrolyzed by two different classes of acyl-ACP thioesterases at the inner plastid envelope membrane, before the unesterified fatty acids are transported to the endoplasmic reticulum (ER), by a family of fatty acid export proteins (FAX) of which there are seven isoforms in Arabidopsis, two of which (FAX2 and FAX4) are highly expressed during the early stage of seed development. The unesterified fatty acids are shuttled across the plastid outer envelope, probably by vectorial acylation by long-chain acyl-CoA synthases, which catalyze the formation of CoA esters.

In the ER, triacylglycerols and membrane lipids are synthesized by the Kennedy and other pathways described above. In yeast and plants, 1,2-diacyl-*sn*-glycerol esterification is the only committed step in triacylglycerol production, and this occurs by mechanisms that can be both dependent or independent of acyl-CoA esters. The acyl-CoA-dependent route is catalyzed by diacylglycerol:acyl-CoA acyltransferases (DGATs) with acyl-CoA and diacylglycerols as substrates, and two membrane-bound isoenzymes (DGAT1 and DGAT2) and a cytosolic isoenzyme (DGAT3) are known, DGAT1 is a key enzyme involved in triacylglycerol formation in developing seeds, while DGAT2 is especially important in those plant species with unusual fatty acid compositions. In Arabidopsis, DGAT3 has some specificity for polyunsaturated fatty acids in seed development. Figure 21.2.22 summarizes TAG synthesis in plant ER.

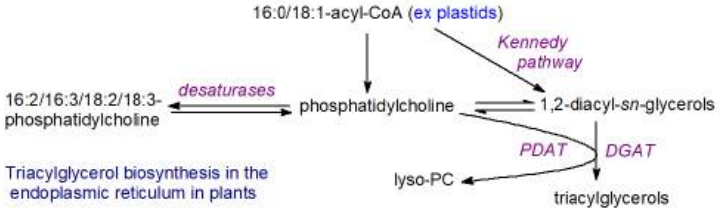


Figure 21.2.22 TAG biosynthesis in plant ER

In addition, a substantial proportion of triacylglycerol biosynthesis in some plant species is synthesized by flux through the membrane phospholipid **phosphatidylcholine**, produced by what are sometimes termed inaccurately the 'eukaryotic and prokaryotic pathways' with differing positional distributions, in which diacylglycerols are generated from phosphatidic acid by the action of a phosphatidate phosphatase as an intermediate. In the acyl-CoA-independent reaction, the direct transfer of one fatty acid from phosphatidylcholine to diacylglycerol by the action of the phospholipid:1,2-diacyl-*sn*-glycerol-acyltransferase (PDAT) enzyme also occurs, with the formation of lysophosphatidylcholine as a byproduct, which can be re-esterified for further reaction. As phosphatidylcholine undergoes extensive remodeling and its fatty acid components are subject to modification, for example by desaturation to form linoleic and linolenic acids, the compositions and especially the positional distributions of triacylglycerols produced in this way can be very different from those synthesized by the ‘classical’ pathways. Phosphatidylcholine may also function as a carrier for the trafficking of acyl groups between organelles and membrane subdomains, and it has been suggested that an assembly of interacting enzymes may facilitate the transfer of polyunsaturated fatty acids from this phospholipid to triacylglycerols in seeds. Both sterol and sphingolipid biosynthesis appear to be important factors for efficient seed oil production.

As triacylglycerol synthesis continues, oil droplets accumulate between the leaflets of the endoplasmic reticulum and are surrounded by a monolayer of phospholipids, sterols, and proteins, which in Arabidopsis include oleosins, a caleosin, a steroleosin, a putative aquaporin, and a glycosylphosphatidylinositol-anchored protein. Oleosins are the most abundant of these (~65%) and are small proteins (15-30 kDa) that contain cytosolic-facing N- and C-termini and a large hydrophobic domain necessary to target them to lipid droplets, where they are important for control of their size and stability. Eventually, lipid droplets “bud off” from the endoplasmic reticulum with their monolayer of phospholipids and proteins, and they are released into the cytosol by a yet-to-be-defined mechanism.

At the onset of germination, water is absorbed, and esterases/lipases are activated. The process of lipolysis begins at the surface of oil bodies, where the oleosins, which are the most abundant structural proteins, are believed to assist in the docking of lipases and to control the size and stability of lipid droplets in seeds. A number of esterases/lipases have been cloned from various plant species and possess a conserved catalytic triad of Ser, His, and Asp or Glu, somewhat different from the animal lipases, as in patatin (an especially abundant lipolytic protein in potatoes), which can hydrolyze triacylglycerols but not phospho- or galactolipids. The most important of these is believed to be the 'sugar-dependent lipase 1 (SDP1)', which is a patatin-like lipase similar in function to the mammalian adipose triacylglycerol lipase discussed above and is located on the surface of the oil body. This is active mainly against triacylglycerols to generate diacylglycerols but presumably works in conjunction with di- and monoacylglycerol lipases to generate free fatty acids and glycerol.

The lipid droplets in seeds exist near glyoxysomes (broadly equivalent to peroxisomes). These are the membrane-bound organelles that contain most of the enzymes required to oxidize fatty acids derived from the triacylglycerols via acetyl-CoA to four-carbon compounds, such as succinate, which are then converted to soluble sugars to provide germinating seeds with energy to fuel the growth of the seedlings and to produce shoots and leaves. In addition, they supply structural elements before the seedlings develop the capacity to photosynthesize. How the products of lipolysis are transported to the glyoxysomes for further metabolism has still to be determined, but a specific ‘ABC’ transporter is required to import fatty acids into the glyoxysomes in Arabidopsis. The free acids are converted to their coenzyme A esters by two long-chain acyl-CoA synthetases located on the inner face of the peroxisome membrane before entry into the β-oxidation pathway. All these processes are controlled by an intricate regulatory network, involving transcription factors that crosstalk with signaling events from the seed maturation phase through to embryo development. After about two days of the germination process, the glyoxysomes begin to break down, but β-oxidation can continue in peroxisomes in leaf tissue.

Lipid droplets - plastoglobules: Triacylglycerol-rich lipid droplets (LD) have been observed in most cell types in vegetative tissues of plants as well as in seeds, and although their origin and function are poorly understood, they contain all the enzymes required for triacylglycerol metabolism together with phospholipases, lipoxygenases, and other oxidative enzymes. Instead of oleosins, these lipid droplets in plants and algae contain a family of ubiquitously expressed 'LD-associated proteins' on the surface, together with a monolayer of phospholipids (mainly phosphatidylcholine), galactolipids such as sulfoquinovosyldiacylglycerol and in some species betaine lipids. As in yeast and humans, seipins (three in Arabidopsis) are necessary for normal LD biogenesis. Again, LD forms within the bilayer of the endoplasmic reticulum and pinch off into the cytoplasm. Abiotic stresses can induce the remodeling of lipid membranes through lipase action with the formation of toxic lipid intermediates, and these can be sequestered by triacylglycerols in lipid droplets to inhibit membrane damage and potentially prevent cell death. While they are believed to be involved mainly in stress responses, lipid droplets may have other specialized roles, for example in anther and pollen development, where triacylglycerols serve as a source of fatty acids for membrane

biosynthesis. Fatty acids derived from triacylglycerols in lipid droplets are believed to be subjected to peroxisomal β-oxidation to produce the ATP required for stomatal opening and no doubt many other purposes.

In addition, lipid droplets that have been termed 'plastoglobules' are produced by a localized accumulation of triacylglycerols and other neutral lipids between the membrane leaflets of the thylakoid cisternae and then pinch off into the stroma, where they are involved in a wide range of biological functions from biogenesis to senescence via the recruitment of specific proteins. During senescence, for example, lipid droplets accumulate rapidly in the leaves of *A. thaliana*. In reproductive tissues, may have a more direct function by recruiting and transporting proteins, both for organ formation and successful pollination. Antifungal compounds such as **2-hydroxy-octadecatrienoic acid** and other oxylipins are produced from α-linolenic acid in these organelles, and it has been suggested that the latter function as intracellular factories to produce stable metabolites via unstable intermediates by concentrating the enzymes and hydrophobic substrates efficiently. Plastoglobules are also implicated in the biosynthesis and metabolism of vitamins E and K.

Microalgae: Triacylglycerol metabolism in lipid droplets in microalgae is under intensive study because of their potential for nutraceutical and biodiesel production. It seems that similar processes occur in higher plants, but with a simpler genome encoding few redundant proteins. In the unicellular green model microalga *Chlamydomonas reinhardtii*, for example, key lipid droplet proteins, lipases, and enzymes of β-oxidation have been characterized.

Yeasts: Lipid droplets in yeast are a highly dynamic and functionally diverse hub that ensures stress resistance and cell survival by promoting membrane and organelle homeostasis. As most of the important biosynthetic and catabolic enzymes involved in triacylglycerol metabolism are conserved between yeasts and mammals, the former proving to be useful models for the study of triacylglycerol production. The size and triacylglycerol content of lipid droplets in yeasts change appreciably in different stages of growth and development, and *Saccharomyces cerevisiae* contains a single phosphatidic acid phosphatase (Pah1), which has an essential role in this process. During vegetative growth, Pah1 in the cytosol is phosphorylated by multiple protein kinases, and this enables the synthesis of phospholipids rather than triacylglycerols. As cells progress into stasis, the Pah1 is dephosphorylated and translocates to the endoplasmic reticulum, which ultimately leads to triacylglycerol synthesis for storage in lipid droplets. Some fatty acids derived from phospholipids are utilized for triacylglycerol biosynthesis at the inner nuclear membrane, and this is important for nuclear integrity.

21.2.8: TRIACYLGLYCEROL METABOLISM IN PROKARYOTES

The study of the biosynthesis of triacylglycerols in bacteria has been stimulated by an awareness of the role of this lipid class in the pathogenesis of *Mycobacterium tuberculosis* and the relationship with antibiotic biosynthesis by *Streptomyces coelicolor*. For example, triacylglycerols are believed to be an energy reserve for the long-term survival of *M. tuberculosis* during the persistence phase of infection as well as a means by which unesterified fatty acids are detoxified. Increasing numbers of bacterial species, for example from the genera *Mycobacterium*, *Nocardia*, *Rhodococcus*, *Micromonospora*, *Dietzia*, and *Gordonia*, are now known to produce triacylglycerols (sometimes wax esters), and these can be stored in lipid droplets in the organisms. The first three steps in triacylglycerol biosynthesis are catalyzed by GPAT, LPAT, and PAP enzymes comparable to those in other organisms. However, it has become apparent that the DGAT can be a dual-function CoA-dependent acyltransferase known as wax ester synthase/diacylglycerol acyltransferase, which accepts a broad diversity of acyl-CoA substrates for esterification of diacylglycerols or long-chain fatty alcohols for the synthesis of triacylglycerols or wax esters, respectively, depending on which intermediates are present in the organisms. Bacteria that lack such an enzyme are unable to produce these non-polar lipids.

21.3: BIOSYNTHESIS OF MEMBRANE GLYCEROLIPIDS

by William (Bill) W. Christie and Henry Jakubowski.

This section is an abbreviated and modified version of material from the Lipid Web, an introduction to the chemistry and biochemistry of individual lipid classes, written by William Christie.

21.3.1: INTRODUCTION

In this section will be explore the synthesis of the membrane glycerophospholipids and their metabolic derivatives. We will start with the simplest one, phosphatidic acids, and end with phosphatidylinositols.

21.3.2: PHOSPHATIDIC ACIDS AND DERIVATIVES

Phosphatidic acid or 1,2-diacyl-*sn*-glycero-3-phosphate is a key intermediate in the biosynthesis both of other glycerophospholipids and of triacylglycerols. It is structurally one of the simplest of the phospholipids and was long thought to be important only as a precursor of other lipids, where it is indeed a key molecule, but it is now known to have many other functions in animals, plants, and other organisms by its influence on membrane structure and dynamics, and by its interactions with various proteins. As a lipid mediator, it modulates various signaling and cellular processes, such as membrane tethering, conformational changes and enzymatic activities of specific proteins, and vesicular trafficking. Moreover, its metabolite **lysophosphatidic acid** is recognized as a key signaling molecule with a myriad of biological effects mediated through specific receptors.

21.3.2.1: PHOSPHATIDIC ACID – OCCURRENCE AND BIOSYNTHESIS

Phosphatidic acid is not an abundant lipid constituent of any living organism, seldom greater than picomolar concentrations in cells, but it is extremely important both as an intermediate in the biosynthesis of other glycerophospholipids and triacylglycerols and as a signaling molecule or a precursor of signaling molecules. Indeed, it is often over-estimated in tissues as it can arise by inadvertent enzymatic hydrolysis during inappropriate storage or extraction conditions during analysis. It is the simplest diacyl-glycerophospholipid, and the only one with a phosphomonoester as the head group. The molecule is acidic and carries a negative charge, i.e., it is an anionic lipid. The structure of phosphatidic acid is shown in Figure 21.3.1.

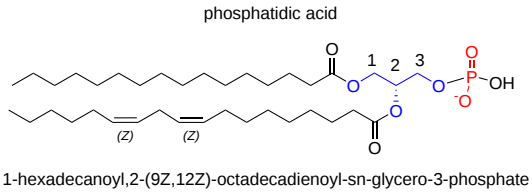


Figure 21.3.1: Phosphatidic acid

There are at least four important biosynthetic pathways for phosphatidic acid biosynthesis in different organelles under various stimuli, and possibly resulting in the formation of different molecular species. The main pathway involves sequential acylation of *sn*-glycerol-3-phosphate, derived from catabolism of glucose, by acyl-coA derivatives of fatty acids. First, one acyltransferases catalyses the acylation of position *sn*-1 to form lysophosphatidic acid (1-acyl-*sn*-glycerol-3-phosphate), and then a second specific acyltransferase catalyses the acylation of position *sn*-2 to yield phosphatidic acid. The synthesis of phosphatidic acid from glycerol-3-phosphate is shown in Figure 21.3.2.

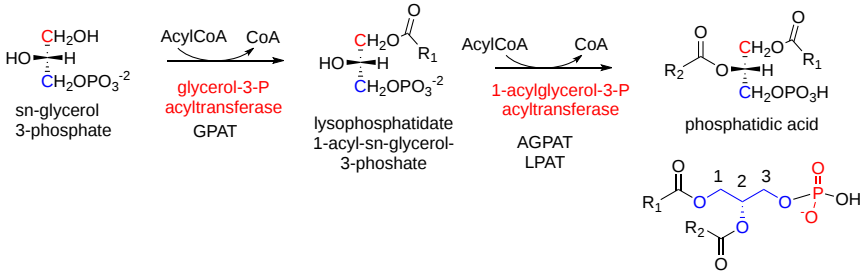


Figure 21.3.2: Synthesis of phosphatidic acid from glycerol-3-phosphate

In **mammals**, the glycerol-3-phosphate acyltransferase that catalyses the first step exists in four isoforms, two in the mitochondrial outer membrane (designated GPAT1 and 2) and two in the endoplasmic reticulum (GPAT3 and 4); all are membrane-bound enzymes, which are believed to span the membranes. GPAT1 is highly expressed in the liver and adipose tissue, where it is responsive to changes in feeding status via the sterol regulatory element binding protein-1 (SREBP-1), a master transcriptional regulator of lipogenic enzymes. It is essential in directing fatty acyl-CoA esters towards glycerolipid synthesis as opposed to β-oxidation. GPAT3 is especially important for

triacylglycerol storage in adipocytes, while GPAT4 is the main contributor to lysophosphatidic acid synthesis in liver and brown adipose tissue.

For the second step in phosphatidic acid biosynthesis, five mammalian acyl-CoA:lysophosphatidic acid acyltransferases are known of which three are in the endoplasmic reticulum (LPAAT or LPAT or AGPAT1, 2 and 3), with a further two (LPAT4 and 5) on the outer mitochondrial membrane. While LPAT1 and 2 have strict specificity for lysophosphatidic acid as acyl acceptor, other isoforms can esterify other lysophospholipids. Human LPAT1 showed higher activity with 14:0-, 16:0- and 18:2-CoAs, while LPAT2 prefers 20:4-CoA and LPAT3 produces phosphatidic acid containing docosahexaenoic acid (22:6(*n*-3)); the last is especially important in retina and testes. LPAT4 and 5 have a preference for oleoyl-CoA and polyunsaturated acyl-CoAs as the acyl donor, suggesting a dual role in glycerolipid synthesis and remodeling. The activity in the endoplasmic reticulum predominates in adipose tissue, but the mitochondrial forms are believed to be responsible for half the activity in liver. However, as there is traffic of phosphatidic acid between the mitochondria and endoplasmic reticulum for remodeling or for synthesis of other lipids, the relative contributions of the two can be difficult to assess.

In **plants**, the *sn*-glycerol-3-phosphate pathway exists both in plastids and at the endoplasmic reticulum with multiple isoforms of the two acyltransferases as well as differences in the acyl substrates. In brief most plant lipid biosynthesis begins with fatty acid biosynthesis in the chloroplasts. In plastids, the acyltransferase ATS1 transfers 18:1 acyl groups from acyl-acyl carrier protein (acyl-ACP) to position *sn*-1 of glycerol 3-phosphate, before ATS2 transfers a palmitoyl group from ACP to position *sn*-2, producing phosphatidic acid at the inner leaflet of the chloroplast inner envelope membrane (IEM). Fatty acids intended for the endoplasmic reticulum are released from ACP in the chloroplast stroma by IEM-associated thioesterases, exported and then activated by acyl-CoA synthetases of the outer envelope membrane to produce species with C_{18} fatty acids in both positions. Thus, acyl-CoAs are used for phosphatidic acid biosynthesis in the endoplasmic reticulum with marked differences in the specificity of the acyl substrates. Subsequently, phosphatidic acid in the plastids is utilized for biosynthesis of galactosyldiacylglycerols, while that in the endoplasmic reticulum is used for synthesis of triacylglycerols and phospholipids.

In **bacteria**, two families of enzymes are responsible for acylation of position *sn*-1 of glycerol-3-phosphate. One present in *Escherichia coli*, for example, utilizes the acyl-acyl carrier protein (acyl-ACP) products of fatty acid synthesis as acyl donors as well as acyl-CoA derived from exogenous fatty acids. In a second wider group of bacteria, including cyanobacteria, there are enzymes (PlsX and PlsY) that make use of the unique acyl donors, **acyl-phosphates** derived in part from acyl-ACP, to acylate position *sn*-1. Acylation of position *sn*-2 in this instance is performed by a further family of enzymes (PlsC) that uses acyl-ACP as the acyl donor, although some bacterial species may use acyl-CoA also.

In animals, a second biosynthetic pathway utilizes dihydroxyacetone phosphate (DHAP) as the primary precursor for the peroxisomal enzyme, DHAP acyltransferase, which produces acyl-DHAP. This intermediate is converted to lysophosphatidic acid in a NADPH-dependent reaction catalyzed by acyl-DHAP reductase, and this is in turn acylated to form phosphatidic acid by the same LPAT as in the previous mechanism. This pathway is of particular importance in the biosynthesis of **ether lipids**. The synthesis of phosphatidic acid from dihydroxyacetone phosphate is shown in Figure 21.3.3.

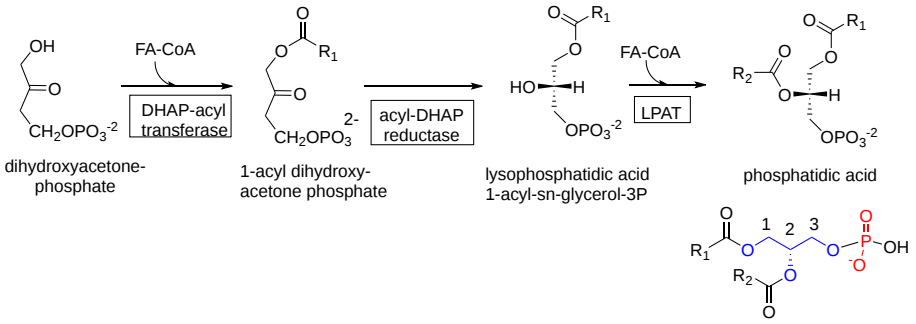


Figure 21.3.3:Synthesis of phosphatidic acid from dihydroacetone phosphate

A third important route to phosphatidic acid is via hydrolysis of other phospholipids, but especially phosphatidylcholine, by the enzyme phospholipase D (or by a family or related enzymes of this kind). The enzyme is readily available for study in plants, where the special functions of phosphatidic acid have long been known (see **below**), but it is now recognized that phospholipase D is present in bacteria, yeasts and most animal cells. In the last, it exists in two main isoforms with differing specificities and cellular locations; PLD1 is found mainly in the Golgi-lysosome continuum, while PLD2 is present mainly in the plasma membrane. They are phosphoproteins, the activity of which is regulated by kinases and phosphatases and by binding to phosphatidylinositol-4,5-bisphosphate. In mitochondria, a distinctive enzyme of this type utilizes cardiolipin as substrate. The mechanism involves the use of water as the nucleophile to catalyse the hydrolysis of phosphodiester bonds in phospholipids. Phospholipase D activity is dependent on and regulated by neurotransmitters, hormones, small monomeric GTPases and lipids. The hydrolysis of phosphatidylcholine to phosphatidic acids by phospholipase D is shown in Figure 21.3.4.

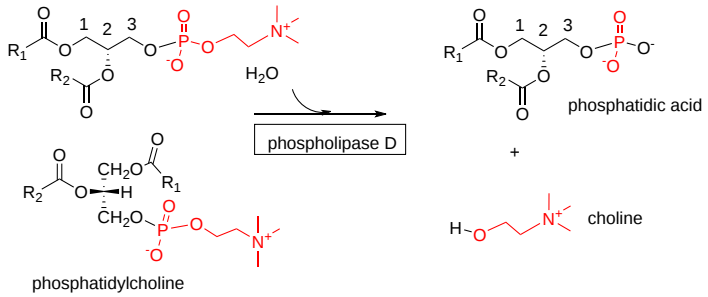


Figure 21.3.4: hydrolysis of phosphatidylcholine to phosphatidic acids by phospholipase D

In addition to its function in generating phosphatidic acid mainly for signaling purposes but also for the maintenance of membrane composition, phospholipase D is involved in intracellular protein trafficking, cytoskeletal dynamics, cell migration, and cell proliferation, partly through protein-protein interactions; it is considered to be important in inflammation and in cancer growth and metastasis as a downstream transcriptional target of proteins involved in the pathophysiology of these diseases. It also has an unusual activity as a guanine nucleotide exchange factor. By a transphosphatidylation reaction with ethanol, it generates **phosphatidylethanol**, a useful biomarker for ethanol consumption in humans.

Under some conditions, phosphatidic acid can be generated from 1,2-diacyl-*sn*-glycerols by the action of **diacylglycerol kinases**, for example those produced from other phospholipids by the action of phospholipase C. Such enzymes appear to be ubiquitous in nature, although those in bacteria and yeast are structurally different from the mammalian enzymes. Diacylglycerol kinases, of which at least ten isoforms (DGKα to DGKκ) exist with different sub-cellular locations and functions in animals, use ATP as the phosphate donor. While the epsilon isoform (DGKε) utilizes the 1-stearoyl-2-arachidonoyl species of diacyl-*sn*-glycerols preferentially to produce phosphatidic acid for the biosynthesis of **phosphatidylinositol**, other isoenzymes phosphorylate diverse diacylglycerol species. Aside from producing phosphatidic acid for phospholipid production or signaling , these enzymes may attenuate the signaling effects of diacylglycerols. For example, diacylglycerol kinases can contribute to cellular asymmetry and control the polarity of cells by regulating the gradients in diacylglycerol and phosphatidic acid concentrations. Figure 21.3.5 shows the synthesis of phosphatidic acid via diacylglycerols and the reverse reaction.

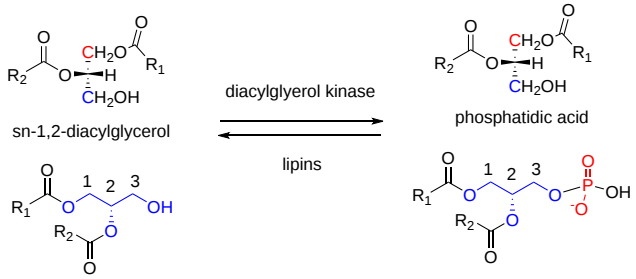


Figure 21.3.5: Synthesis of phosphatidic acid via diacylglycerols and the reverse reaction

The reverse reaction, hydrolysis, is catalyzed by lipins. These enzymes are of importance in regulating the local concentrations of phosphatidic acid and thence its biological activity.

A further possible route to phosphatidic acid production for signaling specifically is via acylation of lysophosphatidic acid, which can be produced independently for signaling purposes as discussed below. This pathway may be especially relevant in membranes, where the protein endophilin has LPAT activity and is believed to generate phosphatidic acid from lysophosphatidic acid in order to alter the curvature of the membrane bilayer.

21.3.2.2: PHOSPHATIDIC ACID - ROLE AS A LIPID PRECURSOR

In summary, phosphatidic acid generated via 1-acyl-*sn*-glycerol-3-phosphate is the primary precursor of other glycerolipids, although other pathways may be more important for generating the lipid for signaling functions. Whether separate pools of this lipid for specific purposes really exist is not certain since dynamic changes of intracellular distribution occur under various cellular conditions. These are attributed to inter-organelle transfer via vesicular transport or at membrane contact sites by lipid transfer proteins. Control of its concentration in membranes, especially in the endoplasmic reticulum, is therefore of great importance, and a transcriptional repressor 'Opi1', which binds specifically to phosphatidic acid in membranes, is a key regulatory factor. However, many other phosphatidic acid-binding proteins have been identified that influence how phosphatidic acid is used either as a biosynthetic precursor or for signaling purposes. The mechanisms for phosphatidic acid homeostasis differ among animals, plants, yeasts, and bacteria in response to the differing functional requirements in these organisms. Figure 21.3.6 shows the pathways for biosynthesis of complex glycerolipids.

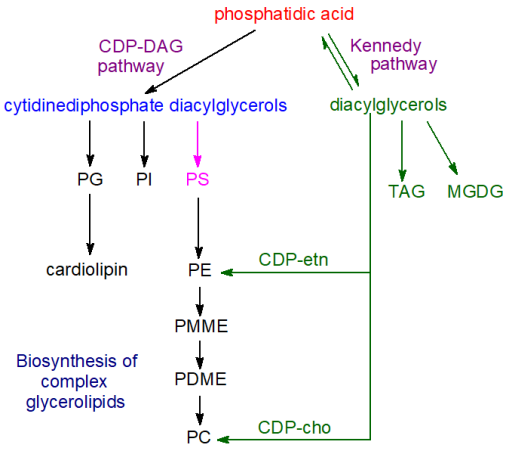


Figure 21.3.6: Biosynthesis of complex glycerolipids

In addition to dietary, hormonal and tissue-specific factors in animals, the extent to which fatty acids are channeling either into triacylglycerol synthesis for storage in lipid droplets and secretion in lipoproteins or into glycerophospholipids for membrane formation depends to a large extent upon the enzymes of glycerol-3-phosphate pathway, their isoform expression, activities and locations. On the other hand, phosphatidic acid is not only a biosynthetic precursor of other lipids but also a regulatory molecule in the transcriptional control of the genes for glycerolipid synthesis, and regulation of its concentration in cells for this purpose is similarly essential. For example, the local concentration of phosphatidic acid in the endoplasmic reticulum is an important factor in the biogenesis of lipid droplets.

The subsequent steps in the utilization of phosphatidic acid in the biosynthesis of triacylglycerols and of the various glycerophospholipids are described in separate documents of this website. Thus, hydrolysis of phosphatidic acid by phosphatidate phosphatase enzymes (including lipins 1, 2, and 3) is the source of most other glycerolipids, e.g. ***sn*-1,2-diacylglycerols** (DG), which are the precursors for the biosynthesis of **triacylglycerols** (TAG), **phosphatidylcholine** (PC) and **phosphatidylethanolamine** (PE) via the so-called Kennedy pathway (also of **monogalactosyldiacylglycerols** in plants). Via reaction with cytidine triphosphate, phosphatidic acid is the precursor of **cytidine diphosphate diacylglycerol**, which is the key intermediate in the synthesis of **phosphatidylglycerol** (PG), and thence of **cardiolipin** (CL), and of **phosphatidylinositol** (PI), and in prokaryotes and yeast but not animals **phosphatidylserine** (PS). Depending on the organism and other factors, phosphatidylserine can be a precursor for phosphatidylethanolamine, while the latter can give rise to phosphatidylcholine by way of mono- and dimethyl-phosphatidylethanolamine intermediates. The cytidine diphosphate diacylglycerol synthase is another enzyme that consumes phosphatidic acid and is important for modulating the concentration of phosphatidic acid in cells and for regulating processes mediated by this lipid.

While the fatty acid composition of phosphatidic acid can resemble that of the eventual products, the latter are generally much altered by re-modeling after synthesis via deacylation-reacylation reactions.

21.3.2.3: PHOSPHATIDIC ACID - BIOLOGICAL FUNCTIONS IN ANIMALS

In addition to its role as an intermediate in lipid biosynthesis, phosphatidic acid and especially that generated by the action of phospholipase D and by diacylglycerol kinases may have signaling functions as a second messenger, although it is not certain whether all the activities suggested by studies *in vitro* operate *in vivo*. Nonetheless, phosphatidic acid has been implicated in many aspects of animal cell biochemistry and physiology.

Some of the observed effects may be explained simply by the physical properties of phosphatidic acid, which has a propensity to form a hexagonal II phase, especially in the presence of calcium ions. Thus, hydrolysis of phosphatidylcholine, a cylindrical non-fusogenic lipid, converts it into cone-shaped phosphatidic acid, which promotes negative membrane curvature and fusion of membranes. It differs from other anionic phospholipids in that its small anionic phosphomonoester head group lies very close to the hydrophobic interior of the lipid bilayer. In model systems, phosphatidic acid can effect membrane fusion, probably because of its ability to form non-bilayer phases. For example, the phosphatidic acid biosynthesis is believed to favor intraluminal budding of endosomal membranes with the formation of exosomes, and in many cell types, vesicle trafficking, secretion and endocytosis may require phosphatidic acid derived by the action of phospholipase D.

Also of relevance in this context is its overall negative charge, and it is not always clear whether some of the observed biological effects are specific to phosphatidic acid or simply to negatively charged phospholipids in general. In contrast to phosphoinositide-interacting proteins, which have defined structural folds, the binding motifs of effector proteins with phosphatidic acid are not highly conserved. However, it has been demonstrated that the positively charged lysine and arginine residues on proteins can bind with some specificity to phosphatidic acid through hydrogen bonding with the phosphate group thus distinguishing it from other phospholipids. An ‘electrostatic-hydrogen bond switch model’ has been proposed in which the head group of phosphatidic acid forms a hydrogen bond to amino acid residues leading to de-protonation of the head group, increasing its negative charge from -1 to -2 and thus enabling stronger interactions with basic residues and

tight docking with the membrane interacting protein. In this way, phosphatidic acid can tether certain proteins to membranes, and it can simultaneously induce conformational changes, hinder ligand binding and/or oligomerize proteins to alter their catalytic activity, stability and interactions with other molecules. It functions as a cellular pH sensor in effect in that binding to proteins is dependent on intracellular pH and the protonation state of its phosphate head group.

One key target of the lipid is mTOR, a serine/threonine protein kinase with a signaling cascade that regulates cell growth, proliferation, motility and survival, together with protein synthesis and transcription, by integrating both nutrient and growth factor signals. This forms two distinct complexes of accessory proteins that regulate downstream targets. Of these, mTORC1 interacts directly with phosphatidic acid and this interaction allosterically activates the enzyme complex to regulate protein synthesis, mitochondrial metabolism and the transcription of enzymes involved in lipid synthesis. In contrast, phosphatidic acid appears to inhibit mTORC2 activity, for example in relation to insulin signaling .

Phosphatidic acid is believed to regulate membrane trafficking events, and it is involved in activation of the enzyme NADPH oxidase, which operates as part of the defence mechanism against infection and tissue damage during inflammation. By binding to targeted proteins, including protein kinases, protein phosphatases and G-proteins, it may increase or inhibit their activities. Effects on gene transcription have been observed that are linked to inhibition of peroxisome proliferator-activated receptor (PPAR) activity. In yeast, phosphatidic acid in the endoplasmic reticulum binds directly to a specific transcriptional repressor to keep it inactive outside the nucleus; when the lipid precursor inositol is added, this phosphatidic acid is rapidly depleted, releasing the transcriptional factor so that it can be translocated to the nucleus where it is able to repress target genes. The overall effect is a mechanism to control phospholipid synthesis.

In addition, phosphatidic acid regulates many aspects of phosphoinositide function. For example, the murine phosphatidylinositol 4-phosphate 5-kinase, the main enzyme generating the lipid second messenger **phosphatidylinositol-4,5-bisphosphate**, does not appear to function unless phosphatidic acid is bound to it; this lipid, generated by the action of phospholipase D, recruits the enzyme to the membrane and induces a conformational change that regulates its activity. It may have a role in promoting phospholipase A_2 activity, a key enzyme in eicosanoid production from phosphoinositide precursors.

In relation to signaling activities, it should be noted that phosphatidic acid can be metabolized to ***sn*-1,2-diacylglycerols** or to lysophosphatidic acid (see next section), both of which have distinctive signaling functions in their own right. Conversely, both of these compounds can be in effect be de-activated by conversion back to phosphatidic acid.

Phospholipase D isoforms and phosphatidic acid have been implicated in a variety of pathologies including neurodegenerative diseases, blood disorders, late-onset Alzheimer's disease and cancer, leading to attempts to develop specific inhibitors of the enzyme for therapeutic purposes. Similarly, the expression of LPAT isoforms can enhance the proliferation and chemoresistance of some cancer cells. Diacylglycerol kinase alpha (DGKα) is highly expressed in several refractory cancer cells, where it attenuates apoptosis, and promotes proliferation. In addition, DGKα is highly abundant in T cells and induces a nonresponsive state, which enables advanced cancers to escape immune action. Inhibition of this enzyme also is seen as a promising treatment strategy.

21.3.2.4: PHOSPHATIDIC ACID - BIOLOGICAL FUNCTIONS IN PLANTS

Phosphatidic acid is present at higher levels in roots of plants in comparison to leaves and is believed to have a function in root architecture. Similarly, its concentration is elevated in flowers and reproductive tissues, but the significance of this is not known. In addition to its role as one of the central molecules in lipid biosynthesis, it facilitates the transport of lipids across plant membranes, and it is also the key plant lipid second messenger, which is rapidly and transiently generated in response to many different biotic and abiotic stresses. In contrast to animal metabolism, the **diacylglycerol** signaling pathway is believed to be relatively insignificant in plants.

The main source of phosphatidic acid for these purposes is the action of phospholipase D (PLD) on membrane phospholipids, such as phosphatidylcholine and phosphatidylethanolamine. Plants contain numerous related enzymes of this type, 12 in Arabidopsis and 17 in rice, in comparison with two in humans and one in yeast, and individual iso-enzymes may elicit specific responses. In the former, the isoforms are grouped into six classes, based on the genic architecture, sequence similarities, domain structures and biochemical properties. These depend mainly on their lipid-binding domains, with some homologous to the human and yeast enzymes and with most containing a characteristic ‘C2’ (calcium- and lipid-binding) domain. The most widespread of these is PLDα, which does not require binding to phosphatidylinositol 4,5-bisphosphate, in contrast to other PLD isoforms and the mammalian enzyme, but millimolar levels of Ca^{2+} are necessary. Studies with fluorescent biosensors suggest that phosphatidic acid accumulates in the subapical region of the cytosolic leaflet of the plasma membrane.

Phosphatidic acid can also be produced by the sequential action of phospholipase C and diacylglycerol kinase on membrane inositol phospholipids, with diacylglycerols as an intermediate (there are 7 isoenzymes in *A. thaliana*). One difference from animal metabolism is that diacylglycerol pyrophosphate can be synthesized from phosphatidic acid in plants (see below).

Phosphatidic acid is required to bind and allosteric activate the **monogalactosyldiacylglycerol synthase (MGDG1)**, located in the inner envelope membrane of the chloroplast, and it may be a regulator of the biosynthesis of thylakoid membranes. Phospholipase D activity and the phosphatidic acid produced have long been recognized as of importance during germination and senescence, and they have an essential role in the response to stress damage and pathogen attack, both in higher plants and in green algae. A high content of phosphatidic acid

induced by phospholipase D action during wounding or senescence brings about a loss of the membrane bilayer phase, because of the conical shape of this negatively charged phospholipid in comparison to the cylindrical shape of structural phospholipids. This change in ionization properties has crucial effects upon lipid-protein interactions, "the electrostatic-hydrogen bond switch model" described above. By promoting negative curvature at the plasma membrane and binding to clathrin proteins, it is believed to facilitate the process of endocytosis. Similar phenomena may explain why phosphatidic acid is important in the response to other forms of stress, including osmotic stress (salinity or drought), cold and oxidation. Although much remains to be learned of the mechanism by which it exerts its effects, it is believed to promote the response to the plant hormone abscisic acid. In addition, phosphatidic acid may interact with salicylic acid to mediate defence responses.

In plants, phosphatidic acid is involved in many different cell responses induced by hormones, stress and developmental processes. In relation to cellular signaling, it often acts in concert with **phosphatidylinositol 4,5-bisphosphate** by binding to specific proteins rather than acting via a receptor. As in mammalian cells, targets for such signaling include protein kinases and phosphatases in addition to proteins involved in membrane trafficking and the organization of the cytoskeleton. It can both activate or inhibit enzymes. If the target protein is soluble, binding to phosphatidic acid can cause the protein to be sequestered into a membrane with effects upon downstream targets. For example, it is involved in promoting the growth of pollen-tubes and root hairs, decreasing peroxide-induced cell death, and mediating the signaling processes that lead to responses to ethylene and again to the hormone abscisic acid. Thus, in the 'model' plant *Arabidopsis*, phosphatidic acid interacts with a protein phosphatase to signal the closure of stomata promoted by abscisic acid; it interacts also with a further enzyme to mediate the inhibition of stomatal opening effected by abscisic acid. Together these reactions constitute a signaling pathway that regulates water loss from plants.

It is noteworthy that phosphatidic acid production can be initiated by opposing stress factors, such as cold and heat, as well as by hormones that are considered to be antagonistic, such as abscisic acid and salicylic acid. It is possible that phosphatidic acid molecules synthesized by the two main pathways differ in composition and cellular distributions and so may produce different responses, but this is an open question. Certainly, during low temperature stress, phosphatidic acid is generated by the action of diacylglycerol kinase. It also seems likely that these differing activities are controlled by the cellular environment where the lipid is produced and by the availability of target proteins or other molecules with which it can act synergistically. Genes encoding enzymes involved in phosphatidic acid metabolism have been manipulated to explore their potential application for crop improvements, based on effects on plant growth, development, and stress responses.

As in animals, phosphatidic acid is catabolized and its signaling functions are terminated by lipid phosphate phosphatases and phosphatidic acid hydrolases, and by acyl-hydrolases and lipoxygenases with the production of fatty acids and other small molecules, which are subsequently absorbed and recycled.

21.3.2.5: LYSOPHOSPHATIDIC ACID

Figure 21.3.7 shows the structure of a lysophosphatidic acid (note the absence of an acyl group at C2).

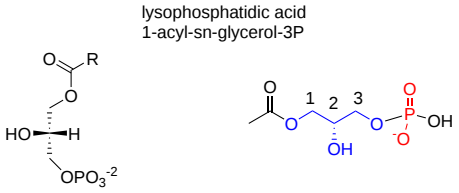


Figure 21.3.7: lysophosphatidic acid

Lysophosphatidic acid (LPA) or 1-acyl-*sn*-glycerol-3-phosphate differs structurally from phosphatidic acid in having only one mole of fatty acid per mole of lipid. As such, it is one of the simplest possible glycerophospholipids. It exists in the form of many different molecular species, i.e., esterified to 16:0 to 22:6 fatty acids, and there is preliminary evidence that saturated and polyunsaturated species may differ in their biological properties in some circumstances. As the *sn*-1-acylated form is more stable thermodynamically, facile isomerization ensures that this tends to predominate. As it lacks one fatty acid in comparison to phosphatidic acid, it is a much more hydrophilic molecule, while the additional hydroxyl group strengthens hydrogen bonding within membranes, properties that may be important for its function in cells.

Although lysophosphatidic acid is present at very low levels only in animal tissues, it is extremely important biologically, influencing many biochemical processes. It is a biosynthetic precursor of phosphatidic acid, but there is particular interest in its role as a lipid mediator with growth factor-like activities. For example, it is rapidly produced and released from activated platelets to influence target cells.

Biosynthesis: In the circulation, the most important source of lysophosphatidic acid is the activity of an enzyme with lysophospholipase D-like activity and known as '**autotaxin**' on **lysophosphatidylcholine** (200 μM in plasma) to yield LPA in an albumin-bound form mainly, although it is relatively soluble in aqueous media because of its polarity and small size. This lipid is more abundant in serum (1 to 5 μM) than in plasma (100 nM), because of the release of its main precursor, lysophosphatidylcholine, from activated platelets during coagulation. Autotaxin is a member of the nucleotide pyrophosphatase-phosphodiesterase family and is also present in cerebrospinal and seminal fluids and many other tissues including cancer cell lines from which it was first isolated and characterized. Indeed, the name derives from the

finding that it promoted chemo*taxis* on melanoma cells in an *auto*crine fashion. It binds to target cells via integrin and heparan sulfate proteoglycans and this may assist the delivery of lysophosphatidic acid to its receptors. Genetic deletion of the enzyme in mice results in aberrant vascular and neuronal development and soon leads to death of the embryos. However, the overexpression of autotaxin causes physical defects also and is eventually lethal to embryos.

Figure 21.3.8 shows the pathways for synthesis of lysophosphatidic acid.

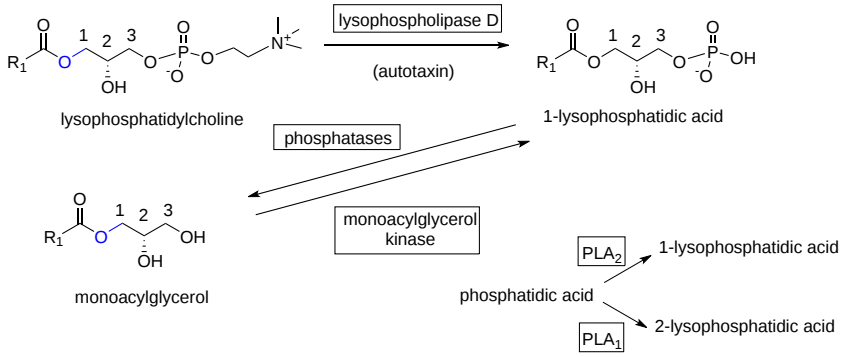


Figure 21.3.8: Pathways for synthesis of lysophosphatidic acid.

While autotaxin is the primary source of extracellular lysophosphatidic acid, it is now established that it is produced intracellularly by a wide variety of cell types by various mechanisms often with phosphatidic acid, derived from other phospholipids by the action of phospholipase D, as the primary precursor. For example, hydrolysis of phosphatidic acid by a phospholipase A_2 (PLA_2) is the main mechanism in platelets, but other cellular enzymes involved include a phosphatidic acid-selective phospholipase A_1 (PLA_1) producing *sn*-2-acyl-lysophosphatidic acid, a monoacylglycerol kinase (utilizing monoacylglycerols produced by the action of lipid phosphate phosphatases) and glycerol-3-phosphate acyltransferase (the first step in phosphatidic acid biosynthesis). In particular, secretory PLA_2-IIA (sPLA2-IIA) is able to induce the release of LPA from phosphatidic acid exposed on the surface of extracellular vesicles derived from platelets and Ca^{2+}-loaded erythrocytes upon stimulation by pro-inflammatory cytokines.

General function: Although lysophospholipids are relatively small molecules, they carry a high content of information through the nature of the phosphate head group, the positional distribution of the fatty acids on the glycerol moiety, the presence of ether or ester linkages to the glycerol backbone, and the chain-length and degree and position of saturation of the fatty acyl chains. Lysophosphatidic acid acts upon nearly all cell types, often as a proliferative and pro-survival signal, inducing cellular invasion, migration and differentiation, while stimulating smooth muscle and fibroblast contraction, cytoskeletal rearrangement, secretion of cytokines/chemokines and numerous other effects. Many of these activities are displayed also by the 1-*O*-alkyl- and alkenyl-ether forms, which can be derived from **platelet activating factor**. On the other hand, it is possible that much of the lysophosphatidic acid produced intracellularly is used for synthesis of other phospholipids rather than for signaling purposes.

Receptors: The informational content of the lysophosphatidic acid molecule leads to selectivity in the functional relationship with cell receptors. As most mammalian cells express receptors for lysophosphatidic acid, this lipid may initiate signaling in the cells in which it is produced, as well as affecting neighboring cells. Characterization of cloned lysophosphatidic acid receptors in combination with strategies of molecular genetics has allowed determination of both signaling and biological effects that are dependent on receptor mechanisms. At least six G protein-coupled receptors that are specific for lysophosphatidic acid have now been identified in vertebrates, each found in particular organs and coupled to at least one or more of the four heterotrimeric G_α proteins and designated $LPAR_1$ to $LPAR_6$, of which $LPAR_1$ is virtually ubiquitous in tissues. These vary appreciably amino acid sequences but are classified into two subgroups, the EDG (LPAR1-3) and P2Y (LPAR4-6) families, with differing tissue distributions. Most cell types express these receptors in different combinations. There is also some interaction with transient receptor potential cation channel V1 (TRPV1), peroxisome proliferator-activated receptor gamma (PPARγ) and other proteins. Plasma lysophosphatidic acid binds to its receptors while it is bound to albumin.

Experimental activation of the LPAR receptors has shown that a range of downstream signaling cascades are mediated by lysophosphatidic acid signaling via these various receptors. These include activation of adenylyl cyclase, cAMP production, intracellular Ca^{2+} and K^+ production (by activating ion channels), protein kinases, phospholipase C, phosphatidylinositol 3-kinase, small GTPases (Ras, Rho, Rac), release of arachidonic acid, and much more. In this way, lysophosphatidic acid regulates cell survival, proliferation, cytoskeleton re-arrangement, motility, cytokine secretion, cell differentiation and many other vital cellular processes. Sometimes, lysophosphatidic acid appears to function in contradictory ways, and there is evidence that it is involved in cell survival in some circumstances and in programmed cell death in others, for example.

Signaling by lysophosphatidic acid has regulatory functions in the mammalian reproductive system, both male and female, facilitating oocyte maturation and spermatogenesis through the action of the receptors $LPAR_1$ to $LPAR_3$. During early gestation, it regulates vascular remodeling at the maternal-fetal interface. There is also evidence that the lipid is involved in brain development, through its activity in

neural progenitor cells, neurons, and glia, and in vascular remodeling. In the central nervous system, these receptors are thought to play a central role in both triggering and maintaining neuropathic pain by mechanisms that may involve demyelination of damaged nerves.

Lysophosphatidic acid has been found in saliva in significant amounts, and it has been suggested that it is involved in wound healing in the upper digestive organs such as the mouth, pharynx, and oesophagus. When applied topically to skin wounds, it has similar effects probably by stimulating proliferation of new cells to seal the wound. Receptor $LPAR_6$ together with the phospholipase A_1 is required for the development of hair follicles, and this receptor is also involved in the regulation of endothelial blood-brain barrier function. The proliferation and survival of stem cells and their progenitors is regulated by lysophosphatidic acid signaling, while in bone cells, acting via $LPAR_1$, lysophosphatidic acid is important for bone mineralization and repair.

Disease: There is particular interest in the activity of lysophosphatidic acid in various disease states and **cancer** especially, as increased expression of autotaxin and the subsequent increased levels of lysophosphatidic acid have been reported in several primary tumors. For example, a finding that lysophosphatidic acid is markedly elevated in the plasma and peritoneal fluid (ascites) of ovarian cancer patients compared to healthy controls may be especially significant. Also, elevated plasma levels were found in patients in the first stage of ovarian cancer, suggesting that it may represent a useful marker for the early detection of the disease. It is believed that the secretory form of phospholipase A_2 acts preferentially on lipids from damaged membranes or microvesicles, such as those produced by malignant cells, and this eventually results in increased levels of this lipid. Lysophosphatidic acid has been shown to stimulate the expression of genes for many different enzymes that lead to the proliferation of ovarian and other cancer cells and may induce cell migration via receptors $LPAR_1$ to $LPAR_3$ and possibly $LPAR_6$, while $LPAR_4$ and $LPAR_5$ have opposing effects. Autotaxin and LPARs have been implicated in resistance to chemotherapy and radiation treatment in cancer therapy.

As lysophosphatidic acid has growth-factor-like activities for many cell types that induce cell proliferation and migration, changes in cellular shape and increasing of endothelial permeability, it is perhaps not surprising that it is relevant to tumor biology. Treatment of various cancer cell types with lysophosphatidic acid promotes the expression and release of interleukin 8 (IL-8), which is a potent angiogenic factor, and thus it has a critical role in the growth and spread of cancers by enhancing the availability of nutrients and oxygen. There is evidence that signaling by lysophosphatidic acid is causally linked to hyperactive lipogenesis in cancer. For example, it activates the sterol regulatory element-binding protein (SREBP) together with the fatty acid synthase and AMP-activated protein kinase–ACC lipogenic cascades leading to elevated synthesis of lipids *de novo*. Increased autotaxin expression has been demonstrated in many different cancer cell lines, and the expression of many of the surface receptors for lysophosphatidic acid in cancer cells is aberrant. Cancer cells must evade the immune system during metastasis, and lysophosphatidic acid facilitates this process by inhibiting the activation of T cells. Therefore, lysophosphatidic acid metabolism is a target of the pharmaceutical industry in the search for new drugs for cancer therapy, aided by a knowledge of the crystal structures of three of the receptors.

Signaling by lysophosphatidic acid has been implicated in many aspects of chronic inflammation, which it promotes by affecting the endothelium in several ways, for example by stimulating endothelial cell migration, the secretion of chemokines-cytokines and regulating the integrity of the endothelial barrier. Problems with lysophosphatidic acid signaling together with changes in autotaxin expression are believed to be factors in such metabolic and inflammatory disorders as obesity, insulin resistance, non-alcoholic fatty liver disease, rheumatoid arthritis, multiple sclerosis and cardiovascular disease. Further, there is evidence it contributes to neurological disorders, such as Alzheimer's disease and neuropathic pain, and to asthma, fibrosis and bone malfunction. Drugs that interact with the lysophosphatidic acid receptors are reported to be effective in attenuating symptoms of several diseases in animal models, and three have passed phase I and II clinical trials for idiopathic pulmonary fibrosis and systemic sclerosis in human patients. Drugs that target autotaxin production and catabolism of lysophosphatidic acid are also in development, and the steroidal anti-inflammatory agent, dexamethasone, appears to be especially useful.

Under certain conditions, lysophosphatidic acid can become athero- and thrombogenic and might aggravate cardiovascular disease. As oxidized low-density lipoproteins promote the production of lysophosphatidic acid, its content in atherosclerotic plaques is high, suggesting that it might serve as a biomarker for cardiovascular disease. Indeed, lysophosphatidic acid promotes pro-inflammatory events that lead to the development of atheroma as well encouraging progression of the disease. By mediating platelet aggregation, it could lead to arterial thrombus formation.

Related lipids: The sphingolipid analogue, **sphingosine-1-phosphate**, shows a similar range of activities to lysophosphatidic acid and the two lipids are often discussed together in the same contexts, although they may sometimes have opposing effects. Acute leukemia cells produce **methyl-lysophosphatidic acids** (the polar head-group is methylated). As these act as antigens to which a specific group of human T cells react strongly, it is possible that they might be a target for the immunotherapy of hematological malignancies. Other lysophospholipids are known to have distinctive biological functions.

Catabolism: Deactivation of lysophosphatidic acid is accomplished by dephosphorylation to produce monoacylglycerols by a family of three lipid phosphate phosphatases (LPP1, 2 and 3), which also de-phosphorylate sphingosine-1-phosphate, phosphatidic acid and ceramide 1-phosphate in a non-specific manner. These are integral membrane proteins with the active site in the plasma membrane facing the extracellular environment, enabling them to access and hydrolyse extracellular lysophosphatidic acid and other phospholipids. Mice with a constitutive LPP3 deficiency are not viable, but this is not true for LPP1 and LPP2 knockout mice. Lysophosphatidic acid can be converted

back to phosphatidic acid by a membrane-bound *O*-acyltransferase (MBOAT2) specific for lysophosphatidic acid (and lysophosphatidylethanolamine) with a preference for oleoyl-CoA as substrate.

21.3.3: PHOSPHATIDYLCHOLINE AND RELATED LIPIDS

21.3.3.1: PHOSPHATIDYLCHOLINE - STRUCTURE AND OCCURRENCE

Phosphatidylcholine or 1,2-diacyl-*sn*-glycero-3-phosphocholine (once given the trivial name 'lecithin') is a neutral or zwitterionic phospholipid over a pH range from strongly acid to strongly alkaline. It is usually the most abundant phospholipid in animals and plants, often amounting to almost 50% of the total complex lipids, and as such it is obviously a key building block of membrane bilayers. In particular, it makes up a very high proportion of lipids of the outer leaflet of the plasma membrane in animals. Virtually all the phosphatidylcholine in human erythrocyte membranes is present in the outer leaflet, for example, while in the plasma membranes of nucleated cells, 80 to 90% of this lipid is located on the outer leaflet. Phosphatidylcholine is also the principal phospholipid circulating in plasma, where it is an integral component of the **lipoproteins**, especially the HDL. On the other hand, it is less often found in bacterial membranes, perhaps ~10% of species, but there is none in the 'model' organisms *Escherichia coli* and *Bacillus subtilis*. In animal tissues, some of its membrane functions appear to be shared with the structurally related sphingolipid, **sphingomyelin**, although the latter has many unique properties of its own.

Figure 21.3.9 shows the structure of phosphatidylcholine

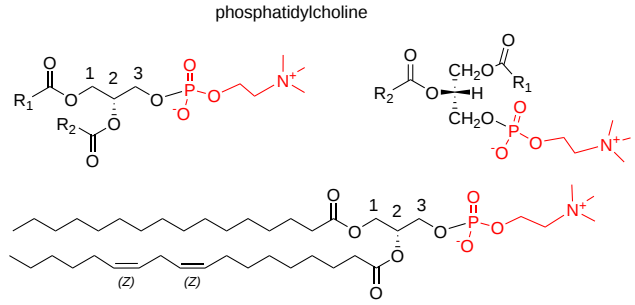


Figure 21.3.9: Phosphatidylcholine

In animal tissues, phosphatidylcholine tends to exist in mainly in the diacyl form, but small proportions (in comparison to phosphatidylethanolamine and phosphatidylserine) of alkyl,acyl and alkenylacyl forms may also be present. As a generalization, animal phosphatidylcholine tends to contain lower proportions of arachidonic and docosahexaenoic acids and more of the C_{18} unsaturated fatty acids than the other zwitterionic phospholipid, phosphatidylethanolamine. Saturated fatty acids are most abundant in position *sn*-1, while polyunsaturated components are concentrated in position *sn*-2. Indeed, C_{20} and C_{22} polyenoic acids are exclusively in position *sn*-2, yet in brain and retina the unusual very-long-chain polyunsaturated fatty acids (C_{30} to C_{38}) of the *n*-6 and *n*-3 families occur in position *sn*-1. Dietary factors obviously influence fatty acid compositions, but in comparing animal species, it would be expected that the structure of the phosphatidylcholine in the same metabolically active tissue would be somewhat similar in terms of the relative distributions of fatty acids between the two positions. Table 21.3.1 lists some representative data.

Table 21.3.1. **Positional distribution of fatty acids in the phosphatidylcholine of some animal tissues.**

Position	Fatty acid						
	16:0	16:1	18:0	18:1	18:2	20:4	22:6
				Rat liver [1]			
***sn*-1**	23	1	65	7	1	trace	
***sn*-2**	6	1	4	13	23	39	7
				Rat heart [2]			
***sn*-1**	30	2	47	9	11	-	-
***sn*-2**	10	1	3	17	20	33	9
				Rat lung [3]			
***sn*-1**	72	4	15	7	3	-	-
***sn*-2**	54	7	2	12	11	10	1
				Human plasma [4]			
***sn*-1**	59	2	24	7	4	trace	-
***sn*-2**	3	1	1	26	32	18	5
				Human erythrocytes [4]			
***sn*-1**	66	1	22	7	2	-	-
***sn*-2**	5	1	1	35	30	16	4
				Bovine brain (gray matter) [5]			
***sn*-1**	38	5	32	21	1	-	-
***sn*-2**	33	4	trace	48	1	9	4
				Chicken egg [6]			
***sn*-1**	61	1	27	9	1	-	-
***sn*-2**	2	1	trace	52	33	7	4

1, Wood, R. and Harlow, R.D. *Arch. Biochem. Biophys*., **131**, 495-501 (1969); DOI.
2, Kuksis, A. *et al*. *J. Lipid Res*., **10**, 25-32 (1969); DOI.
3, Kuksis, A. *et al. Can. J. Physiol. Pharm*., **46**, 511-524 (1968); DOI.
4, Marai, L. and Kuksis, A. *J. Lipid Res*., **10**, 141-152 (1969); DOI.
5, Yabuuchi, H. and O'Brien, J.S. J. Lipid Res., 9, 65-67 (1968); DOI.
6, Kuksis, A. and Marai, L. *Lipids*, **2**, 217-224 (1967); DOI.

There are some exceptions to the rule as the phosphatidylcholine in some tissues or organelles contains relatively high proportions of disaturated molecular species. For example, it is well known that lung phosphatidylcholine in most if not all animal species studied to date contains a high proportion (50% or more) of dipalmitoylphosphatidylcholine.

The positional distributions of fatty acids in phosphatidylcholine in representative plants and yeast are listed in Table 21.3.2. In the leaves of the model plant *Arabidopsis thaliana*, saturated fatty acids are concentrated in position *sn*-1, but monoenoic fatty acids are distributed approximately equally between the two positions, and there is a preponderance of di- and triunsaturated fatty acids in position *sn*-2; the same is true for soybean ‘lecithin’. In the yeast *Lipomyces lipoferus*, the pattern is somewhat similar except that much of the 16:1 is in position *sn*-1.

Table 21.3.2: **Composition of fatty acids (mol %) in positions *sn*-1 and *sn*-2 in the phosphatidylcholine from plants and yeast.**

Position	Fatty acid					
	16:0	16:1	18:0	18:1	18:2	18:3
			Arabidopsis thaliana (leaves) [1]			
***sn*-1**	42		4	5	23	26
***sn*-2**	1		trace	5	47	47
			Soybean 'lecithin' [2]			
***sn*-1**	24		9	14	47	4
***sn*-2**	5		1	13	75	6
			Lipomyces lipoferus [3]			
***sn*-1**	24	18	trace	37	16	4
***sn*-2**	4	5	trace	39	31	19
1, Browse, J., Warwick, N., Somerville, C.R. and Slack, C.R. *Biochem. J.*, **235**, 25-31 (1986); DOI. 2, Blank, M.L., Nutter, L.J. and Privett, O.S. *Lipids*, **1**, 132-135 (1966); DOI. 3, Haley, J.E. and Jack, R.C. *Lipids*, **9**, 679-681 (1974); DOI.						

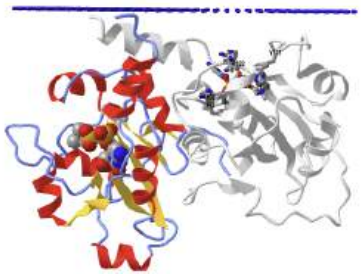
21.3.3.2: PHOSPHATIDYLCHOLINE – BIOSYNTHESIS

There are several mechanisms for the biosynthesis of phosphatidylcholine in animals, plants and micro-organisms. Choline itself is not synthesized as such by animal cells and is an essential nutrient, not only for phospholipid synthesis but also for cholinergic neurotransmission (acetylcholine synthesis) and as a source for methyl groups for numerous other metabolites. It must be obtained from dietary sources or by degradation of existing choline-containing lipids, for example those produced by the second pathway described below. Once taken across membranes and into cells by specific transporters, choline is immediately phosphorylated by a choline kinase (1) in the cytoplasm of the cell to produce phosphocholine, which is reacted with cytidine triphosphate (CTP) by the enzyme **CTP:phosphocholine cytidylyltransferase (CCT)** (2) to form cytidine diphosphocholine (CDP-choline).

$$CTP + PC \rightarrow CDP\text{-}choline + Pi$$

The latter enzyme exists in two isoforms of which CCTα is the more important and is a soluble protein found first in the nucleoplasm, but then in the nucleoplasmic reticulum. This is considered to be the rate-limiting step in phosphatidylcholine biosynthesis, and the activity of the enzyme is regulated by signals from a sensor in the membrane that reports on the relative abundance of the final product. However, choline kinase (ChoKα) also has regulatory functions.

Figure 21.3.10 shows an interactive iCn3D model of the mammalian (rat) CTP: Phosphocholine cytidylyltransferase catalytic domain (3HL4).



NCBI iCn3D Figure 21.3.10: Mammalian (rat) CTP-Phosphocholine cytidylyltransferase catalytic domain (3HL4). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...tBYRYk7vPKD7f6

The biologically active homodimer is shown. The B chain is colored by secondary structure and the A chain is shown in gray. Key active site residues (only shown in the A chain) are in CPK-colored sticks and labeled. A bound CDP-choline analog ((2-cytidylate-O'-phosphonyloxy)-ethyl-trimethylammonium) is shown in spacefill CPK colors.

This enzyme (CCT) catalyzes the key regulatory and rate-limiting step in PC synthesis. The C-terminal domain binds membrane lipids and regulates the enzyme. The iCn3D model above is for the catalytic domain with the regulatory domain deleted. Two nonconserved active site side chains, His 168 and Tyr 173 interact with and position the phosphocholine. Other active site residues include Arg-196 in L6, Lys-122 in L2, and Asp-94. Figure 21.3.11 shows a simplified mechanism for the reaction

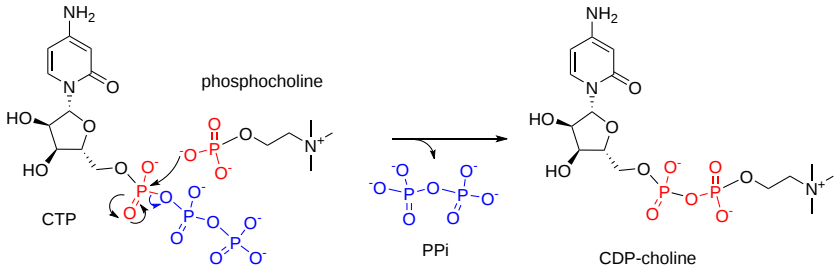


Figure 21.3.11: Simplified mechanism for CTP-Phosphocholine cytidylyltransferase

In plants, nematodes and certain parasites, most phosphocholine is synthesized by sequential methylation of phosphoethanolamine by phospho-base *N*-methyltransferases, but phosphatidylethanolamine is only methylated in this way in a few plant species. This is also the main route to free choline and betaine in plants.

The CDP-choline produced is acted upon by the membrane-bound enzyme CDP-choline:1,2-diacylglycerol choline/ethanolamine-phosphotransferase in the endoplasmic reticulum (CEPT1), and a related choline phosphotransferase 1 (CPT1) in the *trans*-Golgi, which catalyse the reaction with *sn*-1,2-diacylglycerols to form phosphatidylcholine. The first of these is responsible for most phosphatidylcholine biosynthesis but with a somewhat different molecular species composition from the second, which has a preference for 1-alkyl precursors. This is the main pathway for the synthesis of phosphatidylcholine in animals and plants, and it is analogous to that for a major route to **phosphatidylethanolamine**; it is also found in a few bacterial species (e.g. *Sinorhizobium meliloti*). Phosphatidylcholine in mitochondria is obtained by transfer from the endoplasmic reticulum.

Figure 21.3.12 shows the main pathways for PC synthesis in plants and animals.

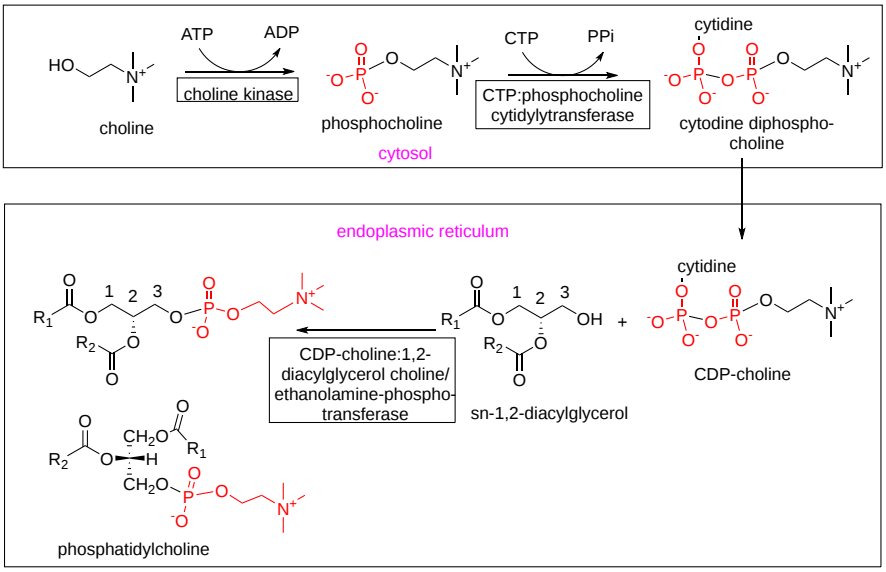


Figure 21.3.172: Main pathways for PC synthesis in plants and animals.

The discovery of the importance of this pathway depended a little on serendipity in that in experiments in the laboratory of Professor Eugene Kennedy, samples of adenosine triphosphate (ATP) contained some cytidine triphosphate (CTP) as an impurity. However, luck is of little value without receptive minds, and Kennedy and co-workers demonstrated that the impurity was an important metabolite that was essential for the formation of phosphatidylcholine.

The above reaction, together with the biosynthetic mechanism for phosphatidylethanolamine, is significantly different from that for phosphatidylglycerol, phosphatidylinositol and cardiolipin. Both make use of nucleotides, but with the latter, the nucleotide is covalently linked directly to the lipid intermediate, i.e., **cytidine diphosphate diacylglycerol**. However, a comparable pathway to the latter for biosynthesis of phosphatidylcholine occurs in bacteria (see below).

The source of the *sn*-1,2-diacylglycerol precursor, which is also a key intermediate in the formation of phosphatidylethanolamine and phosphatidylserine, and of triacylglycerols, is **phosphatidic acid**. In this instance, the important enzyme is phosphatidic acid phosphatase (also known as lipin or phosphatidate phosphatase' or 'lipid phosphate phosphatase' or 'phosphatidate phosphohydrolase').

Figure 21.3.13 shows the biosynthesis of the diacyl precursor of PC

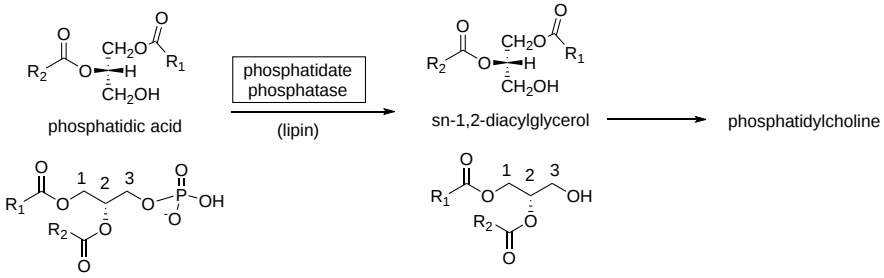


Figure 21.3.13: Synthesis of the diacyl precursor of PC

This enzyme is also important for the production of diacylglycerols as essential intermediates in the biosynthesis of triacylglycerols and of **phosphatidylethanolamine**. Yeasts contain two such enzymes, one of which is Mg^{2+}-dependent (PAP1) and the other Mg^{2+}-independent (PAP2). In mammals, much of the phosphatidic acid phosphatase activity resides in three related cytoplasmic proteins, termed lipins-1, -2, and -3. Lipin-1 is found mainly in adipose tissue, while lipin-2 is present mainly in liver. They are unique among biosynthetic enzymes for glycerolipids in that they can transit among cellular membranes rather than remain tethered to membranes. Of these lipin-1 is most important and exists in three isoforms, lipin-1α, lipin-1β and lipin-1γ with lipin-1α located mainly in the nucleus and lipin-1β in the cytoplasm. Lipin-1γ is present primarily in brain.

The second pathway for biosynthesis of phosphatidylcholine involves sequential methylation of phosphatidylethanolamine, with *S*-adenosylmethionine (SAM) as the source of methyl groups, with **mono- and dimethylphosphatidylethanolamine** as intermediates and catalyzed by the enzyme phosphatidylethanolamine *N*-methyltransferase. A single enzyme (~20 KDa) in two isoforms catalyses all three reactions in hepatocytes; the main form is located in the endoplasmic reticulum (ER) where it spans the membrane, while the second is found in the mitochondria-associated ER membrane. At least two *N*-methyltransferases are present in yeasts. This is a major pathway in the liver, generating one third of the phosphatidylcholine in this organ, but not in other animal tissues or in general in higher organisms. It may be the main route to phosphatidylcholine in those bacterial species that produce this lipid and in yeasts, but it appears to operate in only a few species of higher plants. When choline is deficient in the diet, this liver pathway is especially important.

Figure 21.3.14 shows the synthesis of PC via methylation of PE.

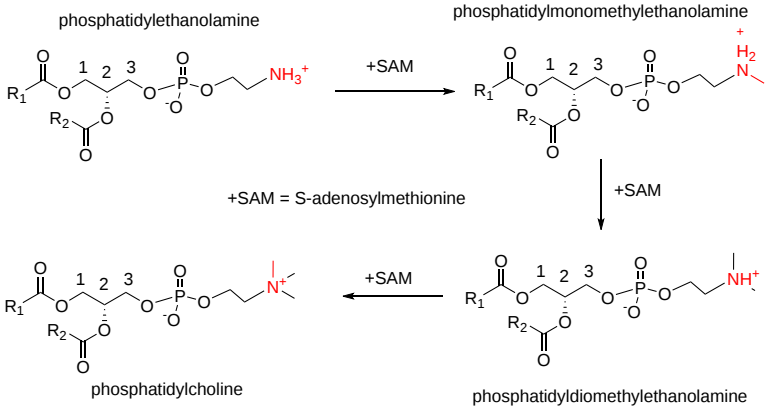


Figure 21.3.14: synthesis of PC via methylation of PE

A by-product of the biosynthesis of phosphatidylcholine from phosphatidylethanolamine is the conversion of *S*-adenosylmethionine to *S*-adenosylhomocysteine, which is hydrolyzed in the liver to adenosine and homocysteine. An elevated level of the latter in plasma is a risk factor for cardiovascular disease and myocardial infarction.

Phosphatidylcholine biosynthesis by both pathways in the liver is necessary for normal secretion of the plasma lipoproteins (VLDL and HDL), and it is relevant to a number of human physiological conditions. It should be noted that all of these pathways for the biosynthesis of diacylphosphatidylcholine are very different and are separated spatially from that producing **alkyl, acyl- and alkenylacyl-phosphatidylcholines** *de novo*. Also, synthesis of phosphatidylcholine does not occur uniformly throughout the endoplasmic reticulum but is located at membrane interfaces or where it meets other organelles, and especially where the membrane is expanding dynamically.

The enzymes in the endoplasmic reticulum responsible for the synthesis of all phospholipids are orientated in such a manner that their active sites are exclusively facing the cytosol. Problems would arise if there were a rapid expansion of the cytosolic leaflet while the luminal leaflet did not change, but a phospholipid transporter known as a scramblase enables a rapid bidirectional flip-flop of phospholipids between leaflets of the bilayer in an energy-independent manner. Compositional asymmetry in first seen in the *trans*-Golgi and is completed before the plasma membrane is formed with phosphatidylcholine and sphingolipids present mainly in the exofacial (outer) leaflet while phosphatidylethanolamine and phosphatidylserine are enriched in the cytosolic leaflet.

Dietary phosphatidylcholine is rapidly hydrolyzed in the proximal small intestine by pancreatic enzymes with formation of lysophosphatidylcholine (and free fatty acids). Further hydrolysis can occur in the jejuno-ileal brush-border by the action of the membrane phospholipases, with the release of glycerophosphocholine, but much of the lysophosphatidylcholine is reacylated by the lyso-PC-acyl-CoA-acyltransferase 3 for export in chylomicrons.

In **plant** cells, phosphatidylcholine biosynthesis occurs mainly in the endoplasmic reticulum, and it is a major component of most membranes other than the internal membranes of plastids; it is absent from the thylakoids and the inner envelope membrane, but is the main glycerolipid of the outer monolayer of the outer envelope membrane. Further complications arise in plants in that turnover or partial synthesis via lysophosphatidylcholine occurs in different organelles from different fatty acid pools or with enzymes with differing specificities, and in addition, fatty acids esterified to phosphatidylcholine serve as substrates for desaturases. The result is that an appreciable pool of the diacylglycerols for the biosynthesis of **triacylglycerols**, **galactosyldiacylglycerols,** and other glycerolipids pass through phosphatidylcholine as an intermediate so that the fatty acid compositions in different membranes change after the initial synthetic process. This mechanism has obvious differences from the remodeling of molecular species in animal tissues discussed next, although a comparable exchange of acyl groups does occur in part catalyzed by acyl transferases (see next section). Some transfer of phosphatidylcholine *per se* from the endoplasmic reticulum to plastids may occur via contact points between the two membranes or may be facilitated by specific transport proteins.

While phosphatidylcholine is a major lipid in yeasts, recent work suggests that it is not essential if suitable alternative growth substrates are available, unlike higher organisms where perturbation of phosphatidylcholine synthesis can lead to inhibition of growth or even cell death.

21.3.3.3: REMODELING OF PHOSPHATIDYLCHOLINE - THE LANDS' CYCLE

Whatever the mechanism of biosynthesis of phosphatidylcholine in animal tissues, it is apparent that the fatty acid compositions and positional distributions on the glycerol moiety are determined post synthesis by extensive re-modeling involving orchestrated reactions of hydrolysis (phospholipase A_2 mainly) to lysophosphatidylcholine, acyl-CoA synthesis and re-acylation by lysophospholipid acyltransferases or transacylases, a series of reactions that is sometimes termed the 'Lands' Cycle' after its discoverer W.E.M. (Bill) Lands. Similar processes occur with all glycerophospholipid classes.

The final composition of the lipid is achieved by a mixture of synthesis *de novo* and the remodeling pathway. There are at least fifteen different groups of enzymes in the phospholipase A_2 super-family, which differ in calcium dependence, cellular location and structure. All hydrolyze the *sn*-2 ester bond of phospholipids specifically, generating a fatty acid and lysophospholipid, both of which have important functions in their own right in addition to their role in the Lands cycle. There is also a phospholipase A_1 family of enzymes, which are esterases that are able to cleave the *sn*-1 ester bond but are less important in this context.

Figure 21.3.15 shows Land's cycle

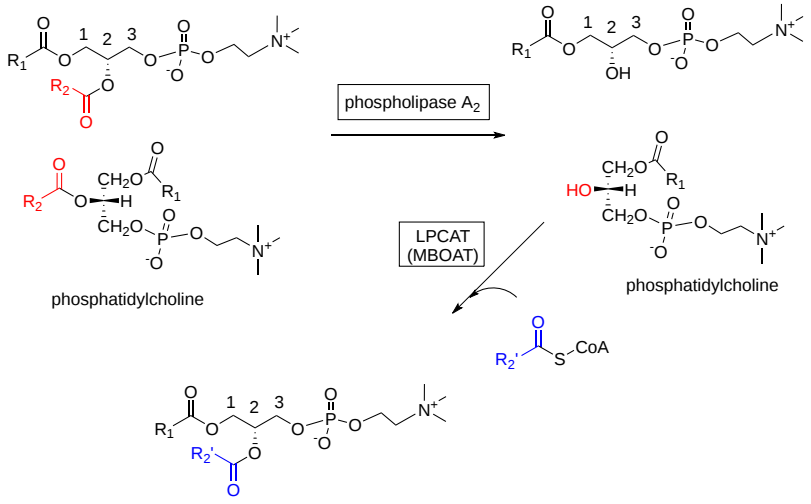


Figure 21.3.15 Land's cycle - 1

The re-acylation step is catalyzed by membrane-bound coenzyme A-dependent lysophosphatidylcholine acyltransferases such as LPCAT3 (also designated ‘MBOAT5’), which is located chiefly within the endoplasmic reticulum, though also in mitochondria and the plasma membrane in organs such as the liver, adipose tissue and pancreas. It maintains systemic lipid homeostasis by regulating lipid absorption and composition in the intestines, the secretion of lipoproteins, and lipogenesis *de novo* in liver, and is notable in that it incorporates linoleoyl and arachidonoyl chains specifically into lysophosphatidylcholine (as does a related enzyme LPCAT2). There is also a CoA-independent acyltransferase in inflammatory cells that transfers arachidonic acid from phosphatidylcholine to ethanolamine-containing phospholipids. While LPCAT3 prefers 1-acyl lysophosphatidylcholine as an acyl acceptor, LPCAT2 utilizes both 1-acyl and 1-alkyl precursors. LPCAT2 is highly expressed in inflammatory cells such as macrophages and neutrophils, which contain ether-phospholipids,

where it contributes to the production of eicosanoid lipid mediators. The highly saturated molecular species of phosphatidylcholine found in lung surfactant are formed from species with a more conventional composition by remodeling by an acyltransferase with a high specificity for palmitoyl-CoA acid (LPCAT1). In other tissues, those species containing high proportions of polyunsaturated fatty acids depend more on synthesis *de novo*. These and further related enzymes are involved in remodeling of all other phospholipids. Over-expression of the genes for these enzymes is associated with the progression of many different cancers and may be involved in other pathological conditions.

Phosphatidylcholine has a central role in glycerolipid metabolism in plants and remodeling occurs for reasons and by mechanisms that are rather different from those in animal cells as described briefly above. For example, there is extensive remodeling as a site of fatty acid desaturation and as the main entry point for acyl groups exported from the plastid into the endoplasmic reticulum. In addition, the remodeling of phosphatidylcholine provides fatty acids for triacylglycerol synthesis in developing seeds and diacylglycerols for the synthesis of thylakoid lipids such as **galactosyldiacylglycerols**. In Arabidopsis, two lysophosphatidylcholine acyltransferases, LPCAT1 and LPCAT2, are involved in remodeling in developing seeds and leaves, with some preference for position *sn*-2 using fatty acids exported from the plastid. In some plant species, there is a strong preference for C_{18}-unsaturated acyl chains over 16:0. However, the lipases that generate lysophosphatidylcholine from phosphatidylcholine for this purpose are not yet known. Some remodeling in plant membranes occurs in response to stress.

The yeast *Saccharomyces cerevisiae* is able to reacylate glycerophosphocholine, generated endogenously by the action of phospholipase B (an enzyme with both phospholipase A_1 and A_2 activities) on phosphatidylcholine, with acyl-CoA in the microsomal membranes by means of a glycerophosphocholine acyltransferase (Gpc1) to produces lysophosphatidylcholine, which can be converted back to phosphatidylcholine by the lysophospholipid acyltransferase (Ale1) with appreciable changes in the molecular species composition. The process is regulated in coordination with the other main lipid pathways and affects yeast growth. The enzyme Gpc1 does not affect other phospholipids in yeasts. A similar mechanism appears to operate in some plant species. Figure 21.3.16 shows variants of the Land's cycle.

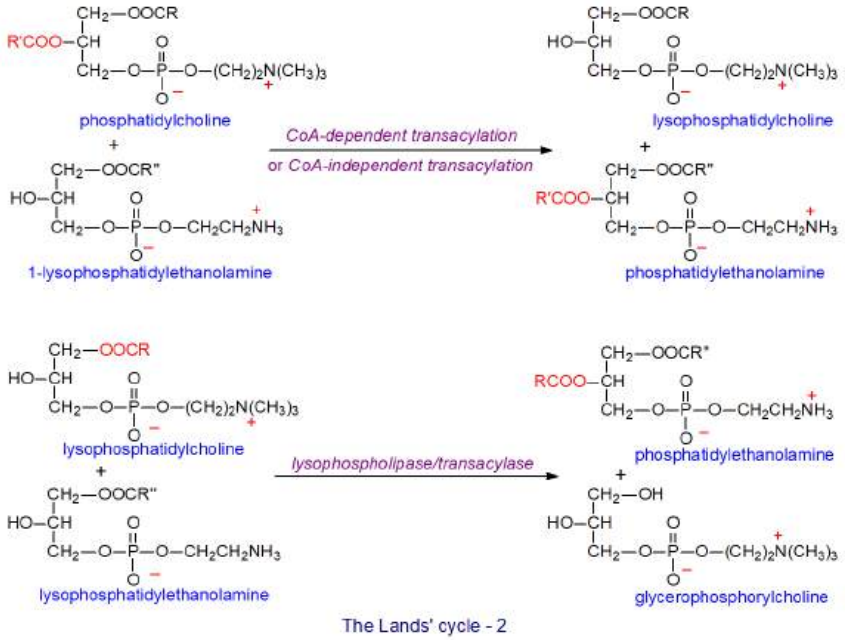


Figure 21.3.16: Variants of the Land's cycle - 2

21.3.3.4: CATABOLISM

Phosphatidylcholine (and most other glycerophospholipids) in membranes can be metabolized by lipolytic enzymes, especially phospholipases, some isoforms of which are specific for particular lipid classes in humans. For example, in addition to the action of phospholipase A (discussed above), phospholipase C (six families in mammals differing in expression and subcellular distribution) yields **diacylglycerols** together with phosphocholine by hydrolyzing glycerophospholipids at the phosphodiester bond, a process that is especially important in relation to **phosphoinositide** metabolism. The **sphingomyelin synthases** also have phospholipase C activity (in the absence of ceramide). Phospholipase D generates **phosphatidic acid** and choline, while phospholipase B removes both fatty acids to yield glycerophosphocholine.

Figure 21.3.17 shows the activities of phospholipases on phosphatidyl choline.

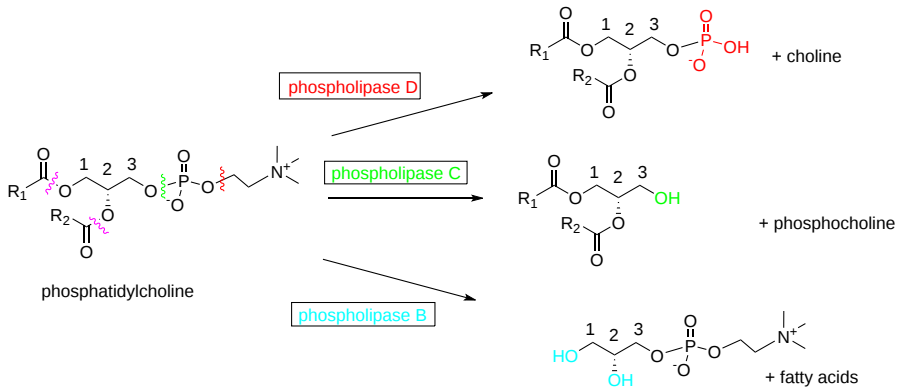


Figure 21.3.17: Activities of phospholipases on phosphatidyl choline

On catabolism in this way, the lipid components are re-cycled or they may have signaling functions, while much of the choline is re-used for phosphatidylcholine biosynthesis, often after being returned to the liver (the CDP-choline cycle). Some choline is oxidized in the kidney and liver to betaine, which serves as a donor of methyl groups for *S*-adenosylmethionine production, and some is lost through excretion of phosphatidylcholine in bile. A proportion is used in nervous tissues for production of acetylcholine, a neurotransmitter of importance to learning, memory and sleep. Phosphatidylcholine in the high-density lipoproteins of plasma is taken up by the liver, and perhaps surprisingly a high proportion of this is eventually converted to triacylglycerols via diacylglycerol intermediates.

21.3.3.5: PHOSPHATIDYLCHOLINE – BIOLOGICAL FUNCTIONS

Because of the generally cylindrical shape of the molecule, phosphatidylcholine organizes spontaneously into bilayers, so it is ideally suited to serve as the bulk structural element of biological membranes, and as outlined above it is makes up a high proportion of the lipids in the outer leaflet of the plasma membrane. The unsaturated acyl chains are kinked and confer fluidity on the membrane. Such properties are essential to act as a balance to those lipids that do not form bilayers or that form specific micro-domains such as **rafts**. While phosphatidylcholine does not induce curvature of membranes, as may be required for membrane transport and fusion processes, it can be metabolized to form lipids that do.

In contrast, dipalmitoyl phosphatidylcholine is the main surface-active component of human lung surfactant, although in other animals the lung surfactant can be enriched in some combination of short-chain disaturated and monounsaturated species, mainly palmitoylmyristoyl- and palmitoylpalmitoleoyl- in addition to the dipalmitoyl-lipid. This is believed to provide alveolar stability by decreasing the surface tension at the alveolar surface to a very low level during inspiration while preventing alveolar collapse at the end of expiration. Also, the internal lipids of the animal cell nucleus (after the external membrane has been removed) contain a high proportion of disaturated phosphatidylcholine. This is synthesized entirely within the nucleus, unlike phosphatidylinositol for example, and in contrast to other cellular lipids its composition cannot be changed by extreme dietary manipulation; it has been suggested that it may have a role in stabilizing or regulating the structure of the chromatin, as well as being a source of diacylglycerols with a signaling function. A further unique molecular species, 1-oleoyl-2-palmitoyl-phosphatidylcholine, is located specifically at the protrusion tips of neuronal cells and appears to be essential for their function, while 1-palmitoyl-2-arachidonoyl-phosphatidylcholine is important in the regulation of the progression of the cell cycle and cell proliferation, and this is independent of eicosanoid production.

Phosphatidylcholine is present bound non-covalently in the crystal structures of a number of membrane proteins, including cytochrome c oxidase and yeast cytochrome bc_1. The ADP/ATP carrier protein has two binding sites for phosphatidylcholine, one on each side. In addition, it is known that the enzyme 3-hydroxybutyrate dehydrogenase must be bound to phosphatidylcholine before it can function optimally. Both the head group and the acyl chains may be involved in the interactions depending on the protein.

As noted above, phosphatidylcholine is by far the most abundant phospholipid component in plasma and in all plasma **lipoprotein** classes. Although it is especially abundant in high density lipoproteins (HDL), it influences strongly the levels of all circulating lipoproteins and especially of the very-low-density lipoproteins (VLDL), which are surrounded by a phospholipid monolayer. Indeed, phosphatidylcholine with polyunsaturated fatty acids in position *sn*-2 is essential for the assembly and secretion of VLDLs and chylomicrons in liver and the intestines, and it must be synthesized *de novo* in the latter. Similarly, phosphatidylcholine synthesis is required to stabilize the surface of lipid droplets in tissues where **triacylglycerols** are stored.

Some of the phosphatidylcholine synthesized in the liver is secreted into bile by a specific flippase together with bile acids where it assists in the emulsification of dietary triacylglycerols in the intestinal lumen to facilitate their hydrolysis and uptake. Eventually, it is absorbed across the intestinal brush border membrane after hydrolysis to lysophosphatidylcholine, which may then be involved in the initiation of chylomicron formation in the endoplasmic reticulum of enterocytes by activation of a protein kinase. In addition, phosphatidylcholine produced in enterocytes is secreted into the intestinal lumen and forms part of the hydrophobic mucus layer that protects the intestinal surface.

Phosphatidic acid generated from phosphatidylcholine by the action of phospholipase D in plants has key signaling functions. Similarly, phosphatidic acid generated in this way from phosphatidylcholine in animals is involved in the metabolism and signaling function of phosphoinositides by activating phosphatidylinositol 4-phosphate 5-kinase, the main enzyme generating the lipid second messenger **phosphatidylinositol-4,5-bisphosphate**. The plasmalogen form of phosphatidylcholine may also have a signaling function, as thrombin treatment of endothelial cells activates a selective hydrolysis (phospholipase A_2) of molecular species containing arachidonic acid in the *sn*-2 position, releasing this fatty acid for eicosanoid production, while the diacyl form of phosphatidylcholine may have a related function in signal transduction in other tissues. In addition, phosphatidylcholine may have a role in signaling via the generation of **diacylglycerols** by phospholipase C, especially in the nucleus. Although the pool of the precursor is so great in many tissues that turnover is not easily measured, the presence of phospholipases C and D that are specific for phosphatidylcholine and are activated by a number of agonists suggests such a function especially in the cell nucleus. Diacylglycerols formed in this way would be much more saturated than those derived from phosphatidylinositol, and would not be expected to be as active in some functions.

Phosphatidylcholine is the biosynthetic precursor of **sphingomyelin** and as such must have some influence on the many metabolic pathways that constitute the sphingomyelin cycle. It is also a precursor for phosphatidic acid, lysophosphatidylcholine and **platelet-activating factor**, each with important signaling functions, and of **phosphatidylserine**.

Because of the increased demand for membrane constituents, there is enhanced synthesis of phosphatidylcholine in cancer cells and solid tumours; the various biosynthetic and catabolic enzymes are seen as potential targets for the development of new therapeutic agents. Impaired phosphatidylcholine biosynthesis has been observed in a number of pathological conditions in the liver in humans, including the development of non-alcoholic fatty liver disease, liver failure and impaired liver regeneration. Similarly, a deficiency in phosphatidylcholine or an imbalance in the ratio of phosphatidylcholine to phosphatidylethanolamine has negative effects upon insulin sensitivity and glucose homeostasis in skeletal muscle.

Plants and bacteria: In addition to its structural role in plant membranes, phosphatidylcholine levels at the shoot apex correlate with flowering time, and this lipid is believed to bind to the Flowering Locus T, a master regulator of flowering. Molecular species containing relatively low levels of α-linolenic acid are involved. Diacylglycerols formed by the action of a family of enzymes of the phospholipase C type on phosphatidylcholine, as opposed to phosphatidylinositol, may be more important in plants and especially during phosphate deprivation for the generation of precursors for galactolipid biosynthesis and perhaps for lipid re-modeling more generally. In prokaryotes, phosphatidylcholine is essential for certain symbiotic and pathogenic microbe-host interactions. For example, in human pathogens such as *Brucella abortus* and *Legionella pneumophila*, this lipid is necessary for full virulence, and the same is true for plant pathogens, such as *Agrobacterium tumefaciens*. Bacteria symbiotic with plants, e.g. the rhizobial bacterium *Bradyrhizobium japonicum*, require it to establish efficient symbiosis and root nodule formation.

21.3.3.6: LYSOPHOSPHATIDYLCHOLINE

Figure 21.3.18 shows the structure of lysophosphatidylcholine

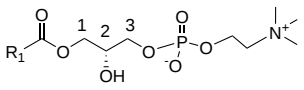


Figure 21.3.18 Lysophosphatidylcholine

Lysophosphatidylcholine (LPC), with one mole of fatty acid per mole of lipid in position *sn*-1, is found in trace amounts in most animal tissues, although there are relatively high concentrations in plasma (150–500μM). It is produced by hydrolysis of dietary and biliary phosphatidylcholine and is absorbed as such in the intestines, but it is re-esterified before being exported in the lymph. In addition, it is formed in most tissues by hydrolysis of phosphatidylcholine by means of the superfamily of **phospholipase A_2** enzymes as part of the de-acylation/re-acylation cycle that controls the overall molecular species composition of the latter, as discussed above. Much of the LPC in the plasma of animal species is secreted by hepatocytes into plasma in a complex with albumin, but an appreciable amount is formed in plasma by the action of the enzyme **lecithin:cholesterol acyltransferase (LCAT)**, which is secreted from the liver. This catalyses the transfer of fatty acids from position *sn*-2 of phosphatidylcholine to free cholesterol in plasma, with formation of cholesterol esters and of course of lysophosphatidylcholine, which consists of a mixture of molecular species with predominately saturated and mono- and dienoic fatty acid constituents. Some LPC is formed by the action of an endothelial lipase on phosphatidylcholine in HDL.

At high concentrations, lysophosphatidylcholine can disrupt membranes, while some biological effects at low concentrations may be simply due to its ability to diffuse readily into membranes, altering their curvature and indirectly affecting the properties of membrane proteins. In plasma, it is bound to albumin and lipoproteins so that its effective concentration is reduced to a relatively safe level.

Lysophosphatidylcholine is considered to be an important factor in cardiovascular and neurodegenerative diseases. It is usually considered to have pro-inflammatory properties and it is known to be a pathological component of oxidized lipoproteins (LDL) in plasma and of atherosclerotic lesions, when it is generated by over-expression or enhanced activity of phospholipase A_2. Its concentration is elevated in joint fluids from patients with rheumatoid arthritis. In addition, it is a major component of platelet-derived microvesicles and activates a

specific receptor in platelets that ultimately leads to vascular inflammation, increasing the instability of atherosclerotic plaques. The intracellular acyltransferase LPCAT cannot remove lysophosphatidylcholine directly from plasma or lipoproteins, nor do there appear to be any enzymes with lysophospholipase A_1 activity in the circulation. Lysophosphatidylcholine blocks the formation of early hemifusion intermediates required for cell-cell fusions. Lysophosphatidylcholine in insect bites attracts inflammatory cells to the site, enhances parasite invasion, and inhibits the production of nitric oxide, for example in Chagas disease. Elevated levels of 26:0-lysophosphatidylcholine in blood are reported to be characteristic of Zellweger spectrum disorders (the result of a defect in peroxisome biogenesis).

Elevated levels of lysophosphatidylcholine have been identified in cervical cancer and may be diagnostic for the disease. On the other hand, reduced concentrations of lysophosphatidylcholine are observed in some malignant cancers, and it has protective effects in patients undergoing chemotherapy. Stearoyl-lysophosphatidylcholine has an anti-inflammatory role in that it is protective against lethal sepsis in experimental animals by various mechanisms, including stimulation of neutrophils to eliminate invading pathogens through a peroxide-dependent reaction. Similarly, there are reports that lysophosphatidylcholine may have beneficial effects in rheumatoid arthritis and a number of other diseases. However, there are suggestions that some experimental studies *in vitro* of the activity of lysophosphatidylcholines may be flawed because insufficient levels of carrier proteins were used. A further point for consideration is that lysophosphatidylcholine is the precursor of the key lipid mediator **lysophosphatidic acid** via the action of the enzyme autotaxin in plasma, and this may be the true source of some of the effects described for the former, especially on cell migration and survival.

There is evidence to suggest that lysophosphatidylcholine containing docosahexaenoic (DHA) and eicosapentaenoic (EPA) acids, presumably in position *sn*-2, in plasma targets more of these fatty acids into the brain, via a specific receptor/transporter at the blood-brain barrier known as the sodium-dependent LPC symporter 1 (MFSD2A), than occurs from the corresponding fatty acids in unesterified form. Hepatic lipase is especially important for generation of these lipids. This finding is now being explored in relation to potential therapeutic applications for neurological diseases, cognitive decline and dementia. Similarly, at the maternal plasma/placental interface, phosphatidylcholine is taken up and hydrolyzed to *sn*-2-lysophosphatidylcholine, presumably by the endothelial lipase, to facilitate transfer of polyunsaturated fatty acids across the basal membrane into the fetal circulation with the aid of the same LPC transporter.

Lysophosphatidylcholine has been found to have some functions in cell signaling , and specific receptors (coupled to G proteins) have been identified, i.e., GPR119, GPR40 and GPR55. It activates the specific phospholipase C that releases diacylglycerols and inositol triphosphate with resultant increases in intracellular Ca^{2+} and activation of protein kinase C. Increased glucose-stimulated insulin secretion has been observed in different cell systems. Lysophosphatidylcholine also activates the mitogen-activated protein kinase in certain cell types, and it promotes demyelination in the nervous system. By interacting with the TRPV4 ion channels of skin keratinocytes, it causes persistent itching. Identification of a highly specific phospholipase $A_{2\gamma}$ in peroxisomes that is unique in generating *sn*-2-arachidonoyl lysophosphatidylcholine suggests that this may be of relevance to eicosanoid generation and signaling . For example, there is reportedly an enrichment of 2-arachidonoyl-lysophosphatidylcholine in carotid atheroma plaque from type 2 diabetic patients. In vascular endothelial cells, it induces the important pro-inflammatory mediator **cyclooxygenase-2 (COX-2)**, a key enzyme in prostaglandin synthesis. However, it has beneficial effects on the innate immune system as it is able to activate macrophages and increase their phagocytic activity in the presence of T lymphocytes.

As lysophospholipids in general and lysophosphatidylcholine in particular are important signaling molecules within mammalian cells, their levels are closely regulated, mainly by the action of the lysophospholipases A_1 and A_2 (LYPLA1 and LYPLA2), depending on the position to which the fatty acid is esterified; these are cytosolic serine hydrolases with esterase and thioesterase activity. The glycerophosphocholine produced can enter the Lands' cycle or be further degraded.

In relation to plants, amylose-rich starch granules of cereal grains contain lysophosphatidylcholine as virtually the only lipid in the form of inclusion complexes or lining channels in the starch macromolecules.

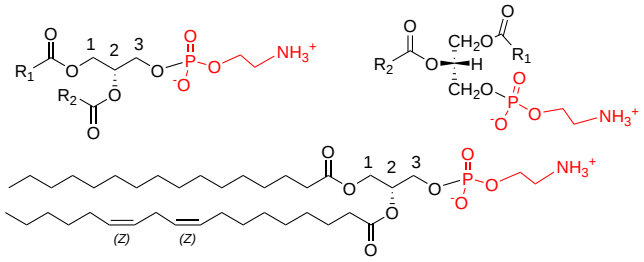
21.3.4: PHOSPHATIDYLETHANOLAMINE AND RELATED LIPIDS

21.3.4.1: PHOSPHATIDYLETHANOLAMINE – STRUCTURE AND OCCURRENCE

Phosphatidylethanolamine or 1,2-diacyl-*sn*-glycero-3-phosphoethanolamine (once given the trivial name 'cephalin') is usually the second most abundant phospholipid in animal and plant lipids, after phosphatidylcholine, and it is frequently the main lipid component of microbial membranes. It can amount to 20% of liver phospholipids and as much as 45% of those of brain; higher proportions are found in mitochondria than in other organelles. As such, it is obviously a key building block of membrane bilayers, and it is present exclusively in the inner leaflet of the plasma membrane in animal cells, for example. It is a neutral or zwitterionic phospholipid (at least in the pH range 2 to 7), with the structure shown (with one specific molecular species illustrated as an example).

Figure 21.3.19 shows the structure of phosphatidylethanolamine.

phosphatidylethanolamine



1-hexadecanoyl,2-(9Z,12Z)-octadecadienoyl-sn-glycero-3-phosphoethanolamine

Figure 21.3.19: Phosphatidylethanolamine

In animal tissues, phosphatidylethanolamine tends to exist in diacyl, alkyl,acyl and alkenyl, acyl forms. As much as 70% of the phosphatidylethanolamine in some cell types (especially inflammatory cells, neurons and tumor cells) can have an ether linkage, but in liver, the plasmalogen form of phosphatidylethanolamine, i.e., with an O-alk-1'-enyl linkage, accounts for only 0.8% of total phospholipids. Generally, there is a much higher proportion of phosphatidylethanolamine with ether linkages than of phosphatidylcholine. If biosynthesis of the plasmalogen form is inhibited by physiological conditions, it is replaced by the diacyl form so that the overall content of the phospholipid remains constant.

In general, animal phosphatidylethanolamine tends to contain higher proportions of arachidonic and docosahexaenoic acids than the other zwitterionic phospholipid, phosphatidylcholine. These polyunsaturated components are concentrated in position *sn*-2 with saturated fatty acids most abundant in position *sn*-1, as illustrated for rat liver and chicken egg in Table 21.3.3. In most other species, it would be expected that the structure of the phosphatidylethanolamine in the same metabolically active tissues would exhibit similar features.

Table 21.3.3: **Positional distribution of fatty acids in phosphatidylethanolamine in animal tissues.**

Position	Fatty acid						
	14:0	16:0	18:0	18:1	18:2	20:4	22:6
			Rat liver [1]				
***sn*-1**		25	65	8			
***sn*-2**	2	11	8	8	10	46	13
			Chicken egg [2]				
***sn*-1**		32	59	7	1		
***sn*-2**		1	1	25	22	29	12
1, Wood, R. and Harlow, R.D., *Arch. Biochem. Biophys.*, **131**, 495-501 (1969); 2, Holub, B.J. and Kuksis, A. *Lipids*, **4**, 466-472 (1969);							

The *O*-alkyl and *O*-alkenyl chains at the *sn*-1 position of the analogous ether lipids generally consist of 16:0, 18:0 or 18:1 chains, whereas arachidonic and docosahexaenoic acids are the most abundant components at the *sn*-2 position.

The positional distributions of fatty acids in phosphatidylethanolamine from the leaves of the model plant *Arabidopsis thaliana* are listed in Table 21.3.4. Here also saturated fatty acids are concentrated in position *sn*-1, and there is a preponderance of di- and triunsaturated in position *sn*-2. The pattern is somewhat different for the yeast *Lipomyces lipoferus*, where the compositions of the two positions are relatively similar.

Table 21.3.4: **Composition of fatty acids (mol %) in positions *sn*-1 and *sn*-2 in the phosphatidylethanolamine from leaves of *Arabidopsis thaliana* and from *Lipoferus lipoferus* .**

Position	Fatty acid					
	16:0	16:1	18:0	18:1	18:2	18:3
A. thaliana [1]						
***sn*-1**	58	trace	4	5	15	18
***sn*-2**	trace	trace	trace	2	60	38
L. lipoferus [2]						
***sn*-1**	29	18	4	28	13	6
***sn*-2**	23	15	3	34	17	6
1, Browse, J., Warwick, N., Somerville, C.R. and Slack, C.R. *Biochem. J.*, **235**, 25-31 (1986); DOI. 2, Haley, J.E. and Jack, R.C. *Lipids*, **9**, 679-681 (1974); DOI .						

21.3.4.2: PHOSPHATIDYLETHANOLAMINE – BIOSYNTHESIS

The two main pathways employed by mammalian cells for the biosynthesis of phosphatidylethanolamine are the CDP-ethanolamine pathway, i.e., one of the general routes to phospholipid biosynthesis *de novo* in plants and animals, and the phosphatidylserine decarboxylase pathway, which occur in two spatially separated organelles - the endoplasmic reticulum and mitochondria, respectively. Ethanolamine is obtained by decarboxylation of serine in plants, and in animals most must come from dietary sources and requires facilitated transport into cells. A small amount of ethanolamine phosphate comes from catabolism of **sphingosine-1-phosphate**, and this is essential for the survival of the protozoon *Trypanosoma brucei*. The initial steps in phosphatidylethanolamine biosynthesis occur in the cytosol with first the phosphorylation of ethanolamine by two specific ethanolamine kinases to produce ethanolamine phosphate; the reverse reaction can occur by means of the enzyme ethanolamine phosphate phosphorylase and this may have a regulatory function in some tissues. The second step is rate-limiting, i.e., reaction of the product with cytidine triphosphate (CTP) to form cytidine diphosphoethanolamine catalyzed by CTP:phosphoethanolamine cytidylyltransferase.

In the final step, a membrane-bound enzyme CDP-ethanolamine:diacylglycerol ethanolaminephosphotransferase catalyses the reaction of cytidine diphosphoethanolamine with diacylglycerol to form phosphatidylethanolamine. There are two such enzymes, ethanolamine phosphotransferase 1 (EPT1) in the Golgi and choline/EPT1 (CEPT1) in the endoplasmic reticulum, but EPT1 is more important for the biosynthesis of the **plasmalogen form**, 1-alkenyl-2-acyl-glycerophosphoethanolamine, and especially molecular species containing polyunsaturated fatty acids, while CEPT1 produced species with shorter-chain fatty acids. The diacylglycerol precursor is formed from phosphatidic acid via the action of the enzyme phosphatidic acid phosphohydrolase.

Figure 21.3.20 shows the synthesis of phosphatidyethanolamine in the ER/Golgi.

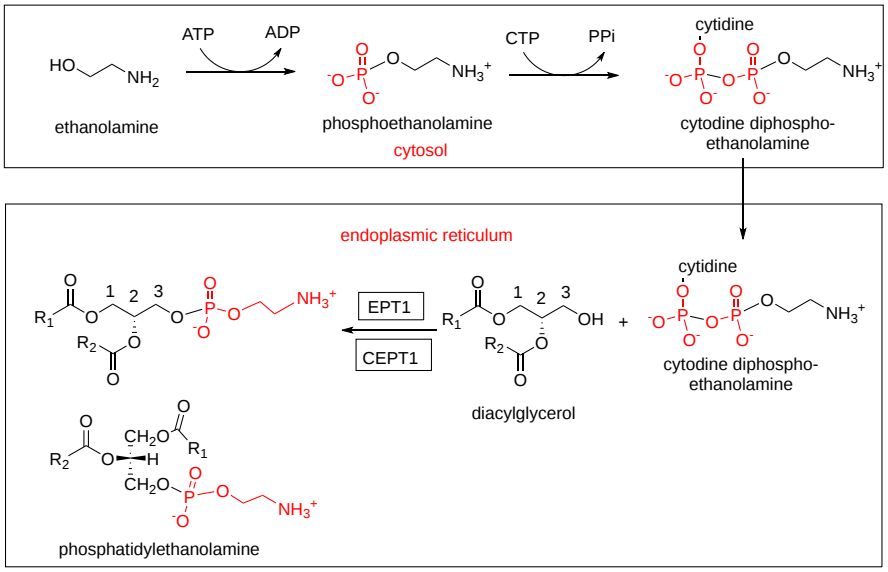


Figure 21.3.20: Synthesis of phosphatidyethanolamine in the ER/Golgi

The second major pathway is the conversion (decarboxylation) of phosphatidylserine to form phosphatidylethanolamine in mitochondria. Conservation of the this pathway from bacteria to humans suggests that it has been preserved to optimize mitochondrial performance. Mitochondria are not connected with the rest of the cell's membrane network by classical vesicular routes, but must receive and export small molecules through the nonvesicular transport at zones of close proximity with other organelles at membrane contact sites, such as a specific domain of the endoplasmic reticulum termed the mitochondria-associated membrane (MAM). Lipid transport of phosphatidylserine is then enabled by tethers that bridge two membranes, lipid transfer proteins and recruitment proteins. In this process, the lipid must traverse two aqueous compartments, the cytosol and the mitochondrial intermembrane space, to reach the inner mitochondrial membrane.

The phosphatidylserine decarboxylase is located on the external aspect of the mitochondrial inner membrane, and most of the phosphatidylethanolamine in mitochondria is derived from this pathway. While various isoforms of the phosphatidylserine decarboxylation exist in prokaryotes, yeasts and mammals, the main forms designated 'PSD1' are found only in mitochondria and are related structurally. An isoform designated 'PSD2' is located in the endosomal membranes of yeasts, and the phosphatidylethanolamine formed is the preferred substrate for phosphatidylcholine biosynthesis. It is evident that cellular concentrations of phosphatidylethanolamine and phosphatidylserine are closely related and tightly regulated.

Figure 21.3.21 shows the synthesis of phosphatidyethanolamine in mitochondria.

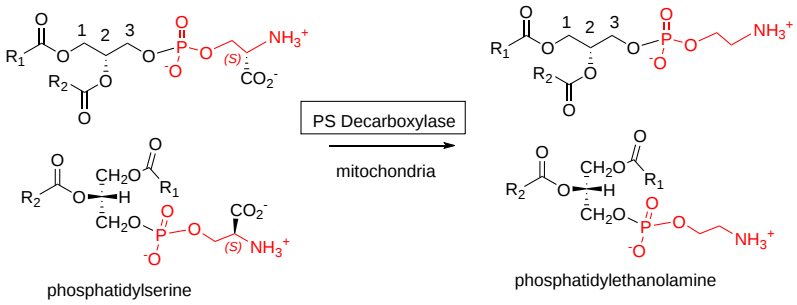


Figure 21.3.21: Synthesis of phosphatidyethanolamine in mitochondria

In prokaryotic cells, such as *E. coli*, in which phosphatidylethanolamine is the most abundant membrane phospholipid, all of it is derived from phosphatidylserine decarboxylation. In this instance, the enzyme undergoes auto-cleavage for activation and utilizes a pyruvoyl moiety to form a Schiff base intermediate with phosphatidylserine to facilitate decarboxylation. This pathway is also important in mammalian cells and yeasts, although the relative contributions of the two main pathways for phosphatidylethanolamine synthesis in mammalian cells appears to depend on the cell type. In cells in culture, more than 80% of the phosphatidylethanolamine is reported to be derived from the phosphatidylserine decarboxylase pathway, but in hamster heart and rat hepatocytes, only ~5% of phosphatidylethanolamine synthesis comes from this route and most is from the CDP-ethanolamine pathway. In yeasts, 70% of the phosphatidylethanolamine is generated by PSD1. The spatially distinct pools within the cell are functionally distinct, but both are essential to life. For example, disruption of the phosphatidylserine decarboxylase gene causes misshapen mitochondria and has lethal consequences in embryonic mice, although phosphatidylethanolamine synthesis continues for a time in other cellular regions. Similarly, elimination of the endoplasmic reticulum route is embryonically lethal.

Three additional minor biosynthetic pathways are known. Phosphatidylethanolamine can be formed by the enzymatic exchange reaction of ethanolamine with phosphatidylserine, or by re-acylation of lysophosphatidylethanolamine. The second of these is associated with the mitochondria-associated membrane where the phosphatidylserine synthase II is located. Finally, the bacterial plant pathogen *Xanthomonas campestris* is able to synthesize phosphatidylethanolamine by condensation of **cytidine diphosphate diacylglycerol** with ethanolamine.

Ether lipids: It should be noted that all of these pathways for the biosynthesis of diacyl-phosphatidylethanolamine are very different and are separated spatially from that producing **alkyl,acyl- and alkenyl,acyl-phosphatidylethanolamines**, suggesting that there may be functional differences. In the protozoon *T. brucei*, for example, it has been demonstrated that the diacyl and ether pools of phosphatidylethanolamine have separate functions and cannot substitute for each other.

Lands' cycle: The various mechanisms produce different pools of phosphatidylethanolamine species, which are often in different cellular compartments and have distinctive compositions. Studies with mammalian cell types *in vitro* suggest that the CDP-ethanolamine pathway produces molecular species with mono- or di-unsaturated fatty acids on the *sn*-2 position preferentially, while the phosphatidylserine decarboxylation reaction generates species with polyunsaturated fatty acids on the *sn*-2 position mainly. However, as with other phospholipids, the final fatty acid composition in animal tissues is attained by a process of remodeling known as the Lands' cycle . The first step is hydrolysis by a phospholipase A_2 to lysophosphatidylethanolamine, followed by reacylation by means of various acyl-CoA:lysophospholipid acyltransferases. The enzymes LPCAT1, 2 and 3, which are involved in phosphatidylcholine biosynthesis, are also active with phosphatidylethanolamine, while LPEAT1 utilizes lysophosphatidylethanolamine mainly and is specific for oleoyl-CoA. Some of these isoforms appear to be confined to particular tissues. There is also a CoA-independent acyltransferase in macrophages that transfers arachidonic acid from phosphatidylcholine to ethanolamine-containing phospholipids.

21.3.4.3: PHOSPHATIDYLETHANOLAMINE – BIOLOGICAL FUNCTION

Physical properties: Although phosphatidylethanolamine has sometimes been equated with phosphatidylcholine in biological systems, there are significant differences in the chemistry and physical properties of these lipids, and they have different functions in biochemical processes. Both are key components of membrane bilayers and the ratio of the two may be important to many cellular functions. However, phosphatidylethanolamine has a smaller head group, which gives the lipid a cone shape. On its own, it does not form bilayers but inverted hexagonal phases. With other lipids in a bilayer, it is believed to exert a lateral pressure that modulates membrane curvature and stabilizes membrane proteins in their optimum conformations. It can hydrogen bond to adjacent lipids and to proteins through its polar head group. For example, the phosphate can be stabilized at the binding site by interactions with lysine and arginine side chains, or with hydroxyls from tyrosine side chains. There can also be strong interactions with acyl chains of the phospholipid.

In contrast to phosphatidylcholine, phosphatidylethanolamine is concentrated with phosphatidylserine in the inner leaflet of the plasma membrane. On the other hand, it is present in both the inner and outer membranes of mitochondria, but especially the former, where it is present at much higher levels than in other organelles and facilitates membrane fusion and protein movement across membranes in addition to many other essential mitochondrial functions, including oxidative phosphorylation via the electron transport chain. It is an important functional component of membrane contact sites between the endoplasmic reticulum and mitochondria. In eukaryotic cells, especially those

of insects, studies suggest that it has a similar function to cholesterol in membranes in that it increases the rigidity of the bilayer to maintain membrane fluidity. In bacterial membranes, it appears that a primary role for phosphatidylethanolamine is simply to dilute the high negative charge density of the anionic phospholipids.

Protein Interactions: Membrane proteins amount to 30% of the genome, and they carry out innumerable biochemical functions, including transport, energy production, biosynthesis, signaling and communication. Within a membrane, most integral proteins consist of hydrophobic α-helical trans-membrane domains that zigzag across it and are connected by hydrophilic loops. Of those parts of the proteins outwith the bilayer, positively charged residues are much more abundant on the cytoplasmic side of membrane proteins as compared to the trans side (the positive-inside rule). Phosphatidylethanolamine is believed to have a key function in that it inhibits location of negative amino acids on the cytoplasmic side, supporting the positive-inside rule, and it has an appropriate charge density to balance that of the membrane surface and the protein. However, it can also permit the presence of negatively charged residues on the cytosolic surface in some circumstances in support of protein function.

Phosphatidylethanolamine is vital for mitochondrial functionality, as demonstrated by defects in the operation of oxidative phosphorylation in the absence of PSD1, and by the role of the lipid in the regulation of mitochondrial dynamics in general and the biogenesis of proteins in the outer mitochondrial membrane. Synthesis of phosphatidylethanolamine in the inner membrane of mitochondria is critical for the function of the cytochrome bc_1 complex (III), where there is a conserved binding site for the lipid in a specific subunit.

Phosphatidylethanolamine binds non-covalently to a superfamily of cytosolic proteins with multiple functions termed 'phosphatidylethanolamine-binding proteins'. While four members have been identified in mammals (PEBP1-4), more than 400 members of the family that are conserved during evolution are known from bacteria to higher eukaryotes. These have many different functions including lipid binding, neuronal development, serine protease inhibition, and the regulation of several signaling pathways; in plants, they control shoot growth and flowering. PEBP4 is of particular interest as it is highly expressed in many different cancers and can increase their resistance to therapy. In animal tissues, phosphatidylethanolamine is especially important in the sarcolemmal membranes of the heart during ischemia, it is involved in the secretion of the nascent very-low-density lipoproteins from the liver and it has functions in membrane fusion and fission. It has a functional role in the Ca^{2+}-ATPase in that one molecule of phosphatidylethanolamine is bound in a cavity between two transmembrane helices, acting as a wedge to keep them apart. This is displaced when Ca^{2+} is bound to the enzyme. Many other important proteins bind non-covalently to phosphatidylethanolamine in a similar way, including rhodopsin and aquaporins.

The content of phosphatidylethanolamine in newly secreted VLDL particles and in apoB-containing particles isolated from the lumen of the Golgi is much higher than that in circulating VLDLs, suggesting that this lipid is involved in VLDL assembly and/or secretion. However, it is rapidly and efficiently removed from the VLDL in the circulation. With lipid droplets in cells, phosphatidylethanolamine is believed to promote coalescence of smaller droplets into larger ones.

Although the mechanism has yet to be fully elucidated, effects on protein conformation are believed to be behind a finding that phosphatidylethanolamine is the primary factor in the brain required for the propagation and infectivity of mammalian prions. Host defense peptides are antimicrobial agents produced by both prokaryotic and eukaryotic organisms, and many of these have a high affinity for phosphatidylethanolamine as a lipid receptor to modulate their activities. For example, the peptide antibiotics cinnamycin and duramycins from Streptomyces have a hydrophobic pocket that fits around phosphatidylethanolamine such that the binding is stabilized by ionic interaction between the ethanolamine group of the lipid and the carboxylate moiety of the peptide; this complex exhibits activity against other Gram-positive organisms, such as *Bacillus* species.

Much of the evidence for the unique properties of phosphatidylethanolamine comes from studies of the biochemistry of the bacterium *E. coli*, where this lipid is a major component of the membranes. Gram-negative bacteria have two membrane bilayers in the cell wall, and as much as 90% of the phospholipid in the inner leaflet of the outer membrane is phosphatidylethanolamine, with a high proportion in the cytoplasmic leaflet of the inner membrane also. In particular, phosphatidylethanolamine has a specific involvement in supporting the active transport of lactose by the lactose permease, and other transport systems may require or be stimulated by it. There is evidence that phosphatidylethanolamine acts as a 'chaperone' during the assembly of this and other membrane proteins to guide the folding path for the proteins and to aid in the transition from the cytoplasmic to the membrane environment, although in contrast it inhibits the folding of some multi-helical proteins. In the absence of this lipid, the transport membranes may not have the correct tertiary structure and so will not function correctly. Whether the lipid is required once the protein is correctly assembled is not fully understood in all cases, but it may be needed to orient enzymes correctly in the inner membrane. Some studies suggest that life in this organism can be maintained without phosphatidylethanolamine, but that life processes are inhibited.

Autophagy and ferroptosis: A covalent conjugate of phosphatidylethanolamine with a protein designated 'Atg8' is formed by the action of cysteine protease ATG4 (belonging to the caspase family) and various other proteins, and is involved in the process of autophagy (controlled degradation of cellular components) in yeast by promoting the formation of membrane vesicles containing the components to be degraded (**phosphatidylinositol 3-phosphate** is also essential to this process). Similarly, oxidatively modified phosphatidylethanolamine is an important factor in **ferroptosis**, a form of apoptosis in which disturbances to iron metabolism lead to an accumulation of hydroperoxides.

Precursor of other lipids: Phosphatidylethanolamine is a precursor for the synthesis of *N*-acyl-phosphatidylethanolamine (see below) and thence of **anandamide** (*N*-arachidonoylethanolamine), and it is the donor of ethanolamine phosphate during the synthesis of the

glycosylphosphatidylinositol anchors that attach many signaling proteins to the surface of the plasma membrane. In bacteria, it functions similarly in the biosynthesis of **lipid A** and other lipopolysaccharides. It is also the substrate for the hepatic enzyme phosphatidylethanolamine *N*-methyltransferase, which provides about a third of the **phosphatidylcholine** in the liver.

Miscellaneous other functions: Phosphatidylethanolamine is the precursor of an ethanolamine phosphoglycerol moiety bound to two conserved glutamate residues in eukaryotic elongation factor 1A, which is an essential component in protein synthesis. This unique modification appears to be of great importance for the resistance of plants to attack by pathogens. *Francisella tularensis* bacteria, the cause of tularemia, suppresses host inflammation and the immune response when infecting mouse cells. The effect is due to a distinctive phosphatidylethanolamine species containing 10:0 and 24:0 fatty acids, and the synthetic lipid produces the same effects *in vitro* in human cells infected with dengue fever virus. It is hoped that this lipid will prove to be a potent anti-inflammatory therapeutic agent.

Plants: In the seeds of higher plants, a deficiency of phosphorylethanolamine cytidylyltransferase, a rate-limiting enzyme in the biosynthesis of phosphatidylethanolamine, has profound effects on the viability and maturation of embryos.

21.3.4.4: LYSOPHOSPHATIDYLETHANOLAMINE

Figure 21.3.22 shows the the structure of lysophosphatidyethanolamine

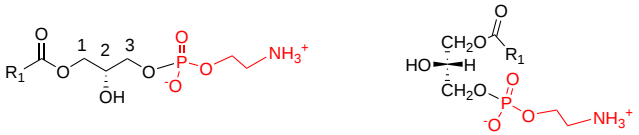


Figure 21.3.22: Lysophosphatidyethanolamine

Lysophosphatidylethanolamine (LPE), with one mole of fatty acid per mole of lipid, is found in trace amounts only in animal tissues, other than plasma (10 to 50μM, or ~1% of total serum phospholipids). It is formed by hydrolysis of phosphatidylethanolamine by the enzyme phospholipase A_2, as part of a de-acylation/re-acylation cycle that controls its overall molecular species composition as discussed above. A membrane-bound *O*-acyltransferase (MBOAT2) specific for LPE (and lysophosphatidic acid) has been characterized with a preference for oleoyl-CoA as substrate. There are reports of the involvement of LPE in cellular functions, such as differentiation and migration of certain neuronal cells, but also of various cancer cells. For example, oleoyl-LPE in the brain stimulates neurite outgrowth and protects against glutamate toxicity.

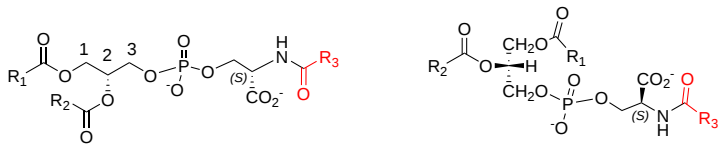
In plants, lysophosphatidylethanolamine is a specific inhibitor of phospholipase D, a key enzyme in the degradation of membrane phospholipids during the early stages of plant senescence. Through this action, it retards the senescence of leaves, flowers, and post-harvest fruits. Indeed, it has a number of horticultural applications when applied externally, e.g., to stimulate ripening and extend the shelf-life of fruit, delay senescence, and increase the vase life of cut flowers. In bacteria, lysophosphatidylethanolamine displays chaperone-like properties, promoting the functional folding of citrate synthase and other enzymes. Some biological properties have been reported in animal tissues *in vitro*, but a specific receptor has yet to be identified.

Lysophospholipids and especially lysophosphatidylethanolamines are produced in the envelope membranes of bacteria by many different endogenous and exogenous factors and must be transported back into the bacterial cell by flippases for conversion back to the diacyl forms by the action of a peripheral enzyme, acyl-ACP synthetase/LPL acyltransferase. Lysophosphatidylethanolamines produced by certain bacteria act synergistically with the **sulfonolipid** rosette-inducing factors (RIFs) to maximize the activity of the latter to induce choanoflagellates to move from a unicellular to a multicellular state.

21.3.4.5: *N*-ACYL PHOSPHATIDYLETHANOLAMINE

In *N*-Acyl phosphatidylethanolamine, the free amino group of phosphatidylethanolamine is acylated by a further fatty acid. This lipid has been detected in rather small amounts in several animal tissues (~0.01%), but especially brain, nervous tissues, and the epidermis, when the *N*-acyl chain is often palmitic or stearic acid (human plasma: N16:0-PE (40%), N18:1-PE (23.3%), N18:0-PE (19%), N18:2-PE (16.6%) and N20:4-PE (1.4%)). Under conditions of degenerative stress, it can accumulate in significant amounts, for example as the result of ischemic injury, infarction, or cancer. It is present in plasma after feeding a high-fat diet to rats, and then it can cross into the brain where it accumulates in the hypothalamus.

Figure 21.3.23 shows the structures of *N*-Acyl Phosphatidylethanolamine.



N-acylphosphatidylethanolamine

Figure 21.3.23: *N*-Acyl Phosphatidylethanolamine

In animals, *N*-Acyl phosphatidylethanolamine is of particular importance as the precursor of **anandamide**, and of other biologically important ethanolamides (e.g., ***N*-oleoylethanolamide**) in brain and other tissues, but especially the intestines. In brief, it is formed

biosynthetically by the action of a transferase (cytosolic phospholipase $A_2\varepsilon$) exchanging a fatty acid from the *sn*-1 position of a phospholipid (probably phosphatidylcholine) to the primary amine group of phosphatidylethanolamine (without a hydrolytic step). Both diacyl- and alkenylacyl-species of phosphatidylethanolamine can serve as acceptors. In addition, some transfer can also occur from phosphatidylethanolamine *per se* by an intramolecular reaction. However, it should be noted that some *N*-acyl phosphatidylethanolamine can be formed artefactually as a result of faulty extraction procedures during analysis.

In plants, *N*-acyl phosphatidylethanolamine is a common constituent of cereal grains (e.g., wheat, barley and oats) and of some other seeds (1.9% of the phospholipids of cotton seeds, but 10-12% of oats). In other plant tissues, it is detected most often under conditions of physiological stress. In contrast to animals, synthesis involves direct acylation of phosphatidylethanolamine with a free fatty acid, catalyzed by a membrane-bound transferase in a reverse serine-hydrolase catalytic mechanism. Activation of *N*-acyl phosphatidylethanolamine metabolism in plants with the release of *N*-acylethanolamines and phosphatidic acid formation seems to be associated with cellular stresses, but research is at an early stage. However, both *N*-acyl lipid classes have been implicated in such physiological processes as the elongation of main and lateral roots, regulation of seed germination, seedling growth, and defense from attacks by pathogens.

N-Acyl phosphatidylethanolamine has been found in a number of microbial species, while *N*-acetyl phosphatidylethanolamine was detected in a filamentous fungus, *Absidia corymbifera*, where it comprised 6% of the total membrane lipids. It was accompanied by an even more unusual lipid 1,2-diacyl-*sn*-glycero-3-phospho(*N*-ethoxycarbonyl)-ethanolamine.

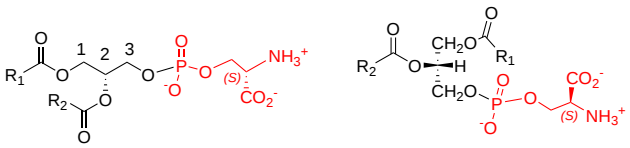
21.3.5: PHOSPHATIDYLSERINE AND RELATED LIPIDS

Phosphatidylserine or 1,2-diacyl-*sn*-glycero-3-phospho-L-serine is an important anionic phospholipid, which brings essential physical properties to membranes in both eukaryotes and prokaryotes. Independently of this, it has many biological functions in cells, including effects on blood coagulation and apoptosis, and it is the biosynthetic precursor for phosphatidylethanolamine in prokaryotes and in eukaryote mitochondria. Its metabolite lysophosphatidylserine has signaling functions and operates through specific receptors. Also, there is increasing interest in a structurally related lipid phosphatidylthreonine, and other phospholipids linked to amino acids.

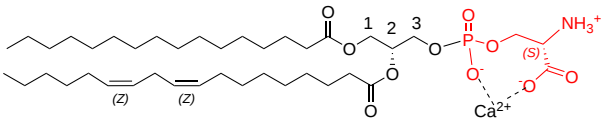
21.3.5.1: PHOSPHATIDYLSERINE - STRUCTURE AND OCCURRENCE

Although phosphatidylserine is distributed widely among animals, plants, and microorganisms, it is usually less than 10% of the total phospholipids, the greatest concentration being in myelin from brain tissue. For example, mouse brain and liver contain 14 and 3% phosphatidylserine, respectively. However, it may comprise 10 to 20 mol% of the total phospholipids in the plasma membrane, where under normal conditions it is concentrated in the inner leaflet, and in the endoplasmic reticulum of cells. In the yeast *Saccharomyces cerevisiae*, it is a minor component of most cellular organelles other than the plasma membrane, where surprisingly it can amount to more than 30% of the total lipids. In most bacteria, it is a minor membrane constituent, although it is important as an intermediate in phosphatidylethanolamine biosynthesis. The 1-octadecanoyl-2-docosahexaenoyl molecular species, which is especially important in brain tissue, is illustrated here.

Figure 21.3.24 shows the structure of phosphatidylserine



phosphatidylserine



1-hexadecanoyl,2-(9Z,12Z)-octadecadienoyl-sn-glycero-3-phosphoserine

Figure 21.3.24 Phosphatidylserine

Phosphatidylserine is an acidic (anionic) phospholipid with three ionizable groups, i.e., the phosphate moiety, the amino group, and the carboxyl function. As with other acidic lipids, it exists in nature in salt form, but it has a high propensity to chelate to calcium via the charged oxygen atoms of both the carboxyl and phosphate moieties, modifying the conformation of the polar head group. This interaction may be of considerable relevance to the biological function of phosphatidylserine, especially during bone formation for example.

In animal cells, the fatty acid composition of phosphatidylserine varies from tissue to tissue, but it does not appear to resemble the precursor phospholipids, either because of selective utilization of specific molecular species for biosynthesis or because of the re-modeling of the lipid via deacylation-reacylation reactions with lysophosphatidylserine as an intermediate (see below). In human plasma, 1-stearoyl-2-oleoyl and 1-stearoyl-2-arachidonoyl species predominate, but in the brain (especially grey matter), retina and many other tissues 1-stearoyl-2-docosahexaenoyl species are especially abundant and appear to be essential for normal functioning of the nervous system. Indeed, the ratio

of *n*-3 to *n*-6 fatty acids in brain phosphatidylserine is much higher than in most other lipids. The positional distribution of fatty acids in phosphatidylserine from rat liver and bovine brain are listed in Table 21.3.5. As with most phospholipids, saturated fatty acids are concentrated in position *sn*-1 and polyunsaturated in position *sn*-2.

Table 21.3.5: **Positional distribution of fatty acids in phosphatidylserine from rat liver and bovine brain**

Position	Fatty acid					
	16:0	18:0	18:1	18:2	20:4	22:6
			Rat liver [1]			
***sn*-1**	5	93	1			
***sn*-2**	6	29	8	4	32	19
			Bovine brain [2]			
***sn*-1**	3	81	13			
***sn*-2**	2	1	25	trace	1	60

1. Wood, R. and Harlow, R.D. *Arch. Biochem. Biophys*., **135**, 272-281 (1969); DOI.
2. Yabuuchi, H. and O'Brien, J.S. *J. Lipid Res*., **9**, 65-67 (1968); DOI.

In leaves of *Arabidopsis thaliana*, used as a 'model' plant in many studies, the fatty acid composition of phosphatidylserine resembles that of **phosphatidylethanolamine**. There is an intriguing report that the chain lengths of the acyl groups increase with age and stress in phosphatidylserine quite specifically, and 22:0 and 24:0 fatty acids have been reported to occur in this lipid in the plasma membrane of some plant species.

In marked contrast to phosphatidylethanolamine, phosphatidylserines with ether-linked moieties (alkyl and alkenyl) are not common in animal tissues, although they are reported to be relatively abundant in human retina and macrophages (they were first found in rat lung). As a generality, the concentration of phosphatidylserine is highest in plasma membranes and endosomes but is very low in mitochondria. As it is located entirely on the inner monolayer surface of the plasma membrane (and of other cellular membranes) and it is the most abundant anionic phospholipid, it may make the largest contribution to interfacial effects in membranes involving non-specific electrostatic interactions. This normal distribution is disturbed during platelet activation and cellular apoptosis.

***N*-Acylphosphatidylserine** is reportedly present in the frontal cortex of patients with schizophrenia, as a minor component of the lipids of sheep erythrocytes, bovine brain, and the central nervous system of freshwater fish, and Bryozoans amongst others. The *N*-arachidonoyl form may be the precursor of the endocannabinoid ***N*-arachidonoylserine**.

21.3.5.2: BIOSYNTHESIS OF PHOSPHATIDYLSERINE

L-Serine is a non-essential amino acid that is actively synthesized by most organisms. In animals, it is produced in nearly all cell types, although in brain it is synthesized by astrocytes but not by neurons, which must be supplied with this amino acid for the biosynthesis of phosphatidylserine (and of sphingoid bases).

In animal tissues, phosphatidylserine is synthesized solely by calcium-dependent base-exchange reactions in which the polar head-group of an existing phospholipid is exchanged for L-serine. There are two routes involving distinct enzymes (PS synthase I and II) with 30% homology and several membrane-spanning domains that can utilize different substrates. Phosphatidylserine is synthesized by both enzymes on the cytosolic face of the endoplasmic reticulum (ER) of the cell, but mainly in a specific domain of this termed the mitochondria-associated membrane ('MAM'), because it is tethered transiently to the mitochondrial outer membrane, presumably by a protein bridge. In yeast, a complex of integrated proteins ('ERMES') has been characterized with a similar function. The reaction involves the exchange of L-serine with either phosphatidylcholine or phosphatidylethanolamine, catalyzed by PS synthase I (although it was long thought that only phosphatidylcholine was a substrate for this enzyme), while PS synthase II catalyzes a similar exchange with diacyl-phosphatidylethanolamine and its the plasmalogen form. Both enzymes are subject to feedback regulation by their product phosphatidylserine, thereby maintaining the correct amounts of this lipid. Figure 21.3.25 shows the synthesis and metabolism of phosphatidylserine in animals.

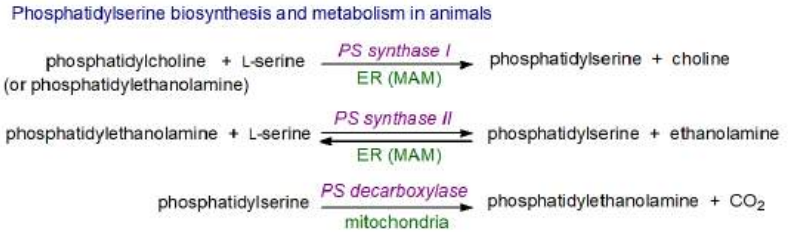


Figure 21.3.25: Synthesis and metabolism of phosphatidylserine in animals.

Phosphatidylserine synthase I is expressed in all mouse tissues, but especially the kidney, liver, and brain, while phosphatidylserine synthase II is most active in the brain and testis and much less so in other tissues. The latter enzyme has a high specificity for molecular species containing docosahexaenoic acid. It is not known why such a complex series of coupled reactions is necessary, or why there should be two enzymes, but one virtue is that the free ethanolamine and choline formed are rapidly re-utilized for phospholipid synthesis. Thus, both phosphatidylserine and phosphatidylethanolamine are produced without a reduction in the amount of phosphatidylcholine. Elimination of both enzymes is embryonically lethal in knock-out mice, but each of them can be knocked out separately and the mice survive, even though they have substantially reduced levels of phosphatidylserine and phosphatidylethanolamine.

As with other phospholipids, the final fatty acid composition in animal tissues is attained by a process of remodeling known as the Lands' cycle. The first step is hydrolysis by a phospholipase A_2 to lysophosphatidylserine, followed by the reacylation by various acyl-CoA:lysophospholipid acyltransferases. One membrane-bound *O*-acyltransferase (LPCAT4 or MBOAT2) with a preference for oleoyl-CoA has been characterized, while a second (LPCAT3 or MBOAT5) incorporates linoleoyl and arachidonoyl chains (and also utilizes lysophosphatidylcholine).

Following synthesis, phosphatidylserine molecules can diffuse laterally in a concentration-dependent manner to different regions of the membrane to fulfill their physiological functions. In humans, cytosolic transport proteins transfer phosphatidylserine and other acidic phospholipids between membranes, and this can also occur by a vesicular transport mechanism.

Some of the newly synthesized phosphatidylserine is transferred to the plasma membrane, while a proportion is transported to the mitochondria, probably again via transient membrane contact (MAM), where it is decarboxylated to produce **phosphatidylethanolamine** by a specific decarboxylase in the inner mitochondrial membrane. In yeast, there is a preference for molecular species containing two monoenoic fatty acids for transport and metabolism; this process occurs also at the Golgi/endosome membranes. All the phosphatidylethanolamine in mitochondria is formed in this way, but some can return to the endoplasmic reticulum where it may be converted back to phosphatidylserine by the action of the PS synthases. Mitochondrial production of phosphatidylethanolamine from phosphatidylserine is not fully complemented by the CDP-ethanolamine pathway, as mice lacking the enzyme do not survive for long. Evidently, cellular concentrations of these two lipids are intimately related and tightly regulated. Figure 21.3.26 shows the mitochondrial metabolism of phosphatidylserine

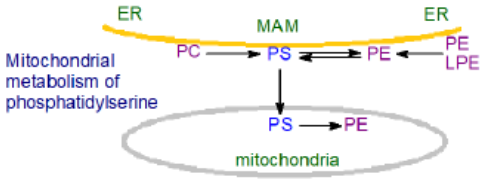


Figure 21.3.26: Mitochondrial metabolism of phosphatidylserine

Much of the phosphatidylserine thus formed is decarboxylated to phosphatidylethanolamine, and this may be the major route to the latter in bacteria. As phosphatidylcholine in yeast is produced via methylation of phosphatidylethanolamine, phosphatidylserine is the primary precursor for this phospholipid in these organisms.

Bacteria and plants: In bacteria and other prokaryotic organisms and in yeast, phosphatidylserine is synthesized by a mechanism comparable to that of most other phospholipids, i.e., by reaction of L-serine with CDP-diacylglycerol, and depends on Mg^{2+} or Mn^{2+}. Phosphatidylserine synthases belong to two different families: type I (non-integral membrane form) in the phospholipase D-like family as in *E. coli*, and type II (integral membrane form) in the CDP-alcohol phosphotransferase family as in *Bacillus* sp. and the yeast *S. cerevisiae*, although the latter shows no homology with the bacterial enzymes.

In many plants, including in the model plant Arabidopsis, much of the phosphatidylserine is produced by a calcium-dependent base-exchange reaction in which the head group of an existing phospholipid is exchanged for L-serine in the luminal leaflet of the endoplasmic reticulum (i.e., mechanistically similar to PS synthase I). It is transferred to the cytoplasmic membrane leaflet by flippases and thence to the post-Golgi compartments before eventually accumulating at the plasma membrane. However, some vesicular transport may occur or there may be direct transfer at membrane contact sites. A CDP-diacylglycerol (prokaryotic-like) biosynthetic pathway exists in some species, e.g. wheat.

Let's explore the mechanism of the *Methanocaldococcus jannaschii* phosphatidylserine synthase (**MjPSS**). The organism is a hyperthermophilic methanogen. Figure 21.3.27 shows substrate binding by MjPSS.

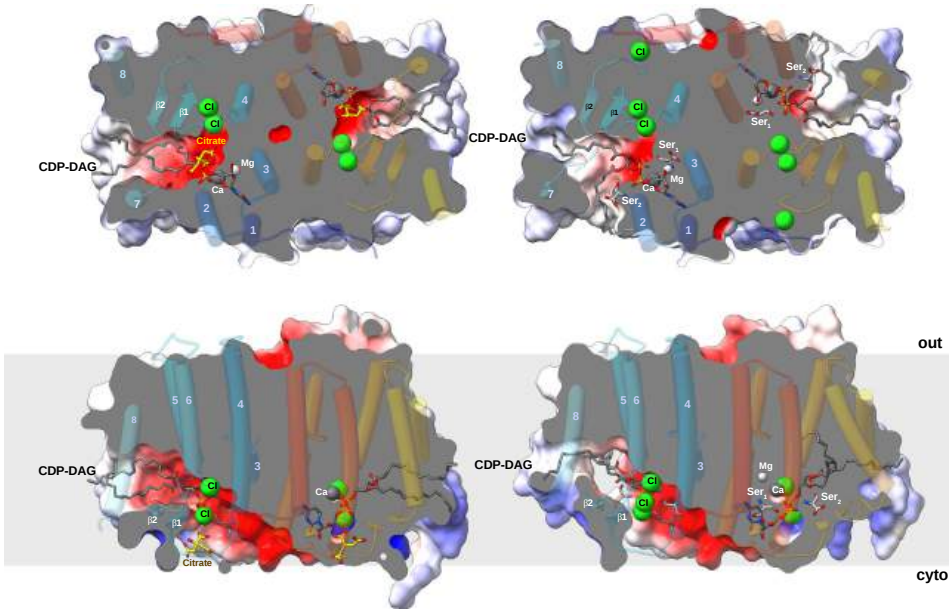


Figure 21.3.27: Substrate binding by MjPSS. Centola, M., van Pee, K., Betz, H. *et al.* Crystal structures of phosphatidyl serine synthase PSS reveal the catalytic mechanism of CDP-DAG alcohol O-phosphatidyl transferases. *Nat Commun* **12,** 6982 (2021). https://doi.org/10.1038/s41467-021-27281-w. Creative Commons Attribution 4.0 International License, http://creativecommons.org/licenses/by/4.0/.

The large binding pocket for CDP-DAG in MjPSS extends from the hydrophobic membrane core to the active site near the cytoplasmic surface in the center of the dimer. In the closed structures (left), both CDP-DAG alkyl chains adopt similar conformations within the binding pocket, whereas the positions of helix 7 and 8 in the open structures (right) allow one alkyl chain to reach the membrane via a different path. Serine molecules are only found in open structures (right). In one closed conformation, there is a citrate near the substrate-binding site, whereas in the other closed structures this position is empty. A chain of three chloride ions extends parallel to the dimer interface from the active site to the cytoplasmic interface with the N-terminal helix hH from the other protomer.

Figure 21.3.28 shows the reaction cycle of MjPSS during the synthesis of PS from CDP-DAG and serine in the presence of Mg^{2+} and Ca^{2+}.

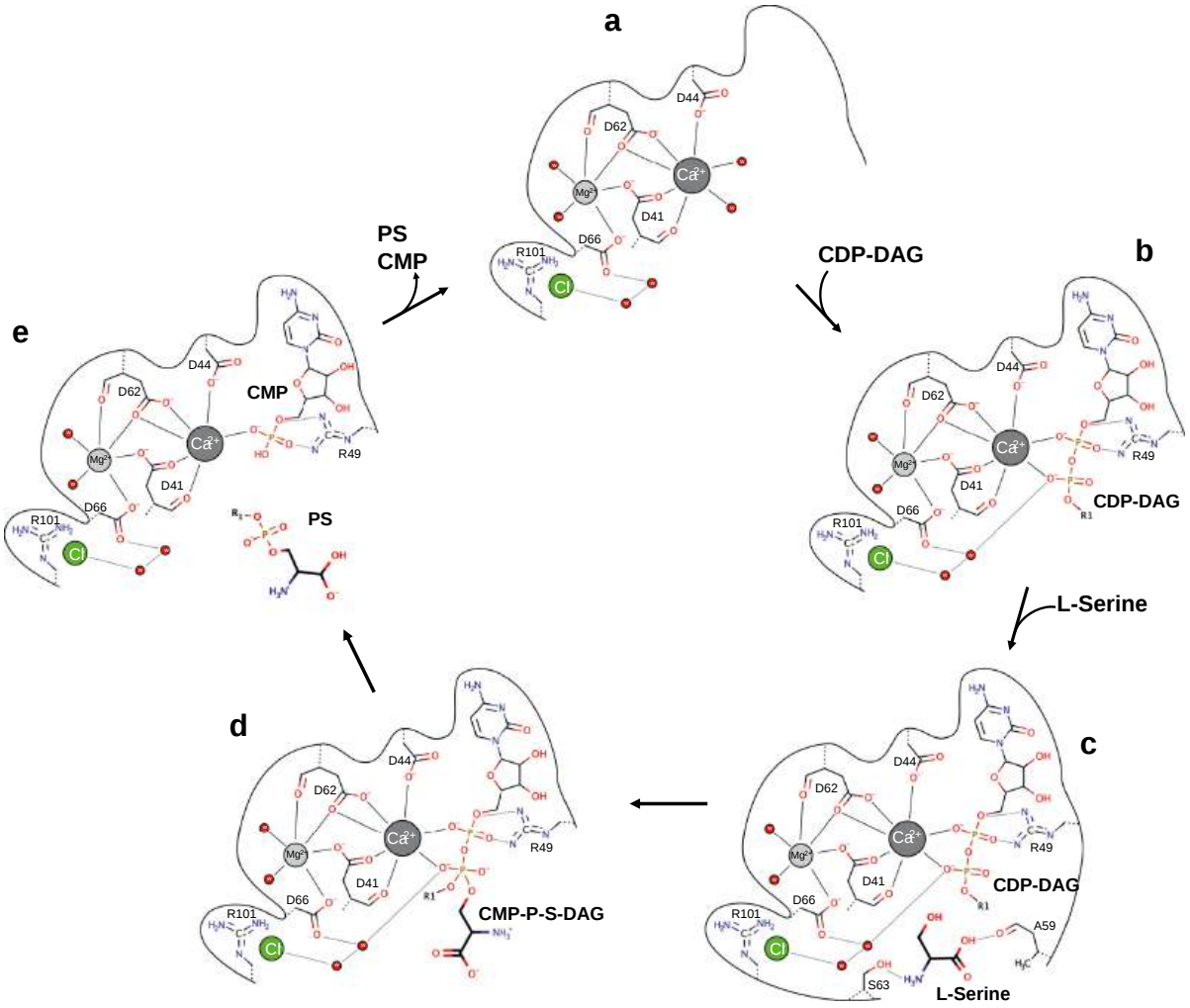


Figure 21.3.28 Reaction mechanism for MjPss. Centola, M. et al., ibid

The CDPDAG binding site in MjPSS of is formed and stabilized by the divalent cations Mg2+ and Ca2+ (a). In the absence of CDP-DAG, Ca2+ most likely is coordinated by water molecules, as shown in a, or by residues in nearby loops that would be flexible in the absence of CDP-DAG. The binding of CDP-DAG (b) is driven by the coordination of Ca2+ by the negatively charged phosphates. Serine binds to the binding pocket after CDP-DAG (c). For the nucleophilic attack, the serine molecule is positioned with its hydroxyl group near the β-phosphate of CDP-DAG. The serine molecule is activated by deprotonation, attacks the β-phosphate, and forms the penta-coordinated transition state (d). The proton of serine is probably removed by one of the water molecules located in the interaction network of Asp66, Arg101, and the nearby chloride ions. Hydrolysis of the CDP-DAG/serine complex from the transition state leads to the complex of MjPSS with the products PS and CMP (e). The next cycle starts after release of the products and binding of CDP-DAG. Structural data are available for the state with bound CDP-DAG (b), CDP-DAG, and serine (c), and for the transition state of the CDP-DAG/serine complex (d).

Figure 21.3.29 shows an interactive iCn3D model of the Methanocaldococcus jannaschii phosphatidyl serine synthase (PSS) in the open state with bound CDP-DAG and serine (7B1L). The enzyme is also named CDP-diacylglycerol--serine O-phosphatidyltransferase.

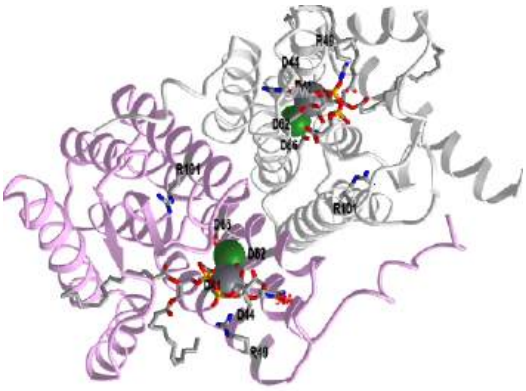


Figure 21.3.29: Bacterial phosphatidyl serine synthase (PSS) in the open state with bound CDP-DAG and serine (7B1L) (Note the actual PDB file title names state that this is the closed state which it is not.) Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...B3Rn3gVN4c4ge6

The A chain of the homodimer is shown in gray and the B in plum. The active site residues in the A chain in the above mechanism are shown in stick, CPK colors, and labeled. Hover over the large ligand (58A) in the gray subunit. Two free serines are shown near it but only one is probably the substrate.

21.3.5.3: PHOSPHATIDYLSERINE – BIOLOGICAL FUNCTION

Membrane location: Phosphatidylserine modulates membrane charge locally, enabling the recruitment of soluble cations and proteins, and so it contributes to the organization of processes within cell membranes. Its distribution within membranes is tightly controlled as it facilitates signaling within the various cellular compartments. Thus, it undergoes a transition from the lumenal leaflet of the endoplasmic reticulum to the cytosolic leaflet in the trans-Golgi network, probably by the activity of flippases and scramblases in the Golgi, and it is highly enriched on the inner leaflet of the plasma membrane. Transport to the plasma membrane against a concentration gradient is aided in part by proteins designated 'ORP5' and 'ORP8' in humans (Osh 6 and Osh7 in yeast) with a 'PH' binding domain for **phosphatidylinositol 4,5-bisphosphate** and an 'ORD' domain for phosphatidylserine. At a membrane contact site between the endoplasmic reticulum and plasma membrane, phosphatidylserine is exchanged for phosphatidylinositol 4-phosphate. Such transfer requires an input of energy, which can be supplied in the form of ATP or by phosphoinositides Although it does not take part in membrane **raft** formation, phosphatidylserine is present in **caveolae**, where it is believed to interact with caveolin-1. It is also present in appreciable amounts in the endosomal compartment.

The asymmetric structure of the plasma membrane with high concentrations of anionic lipids such as phosphatidylserine in the cytosolic leaflet with zwitterionic lipids in the extracellular leaflet generates two surfaces with greatly different electrostatic potentials that influence the association of proteins with the membrane surface and the activities of integral membrane proteins. This distribution is maintained and can be altered, after specific activation, by various flippases (transfer back into the cytoplasmic leaflet), floppases (transfer out of the cytoplasmic leaflet), and scramblases (bidirectional transfer), including ATP-dependent translocases selective for phosphatidylserine. Phosphatidylserine is highly enriched in the cytosolic leaflet of the membranes of recycling endosomes, which replenish the lipids and proteins of the plasma membrane, and it is essential for their function.

Enzyme activation: In addition to its function as a component of cellular membranes and as a precursor for other phospholipids, phosphatidylserine is an essential cofactor that binds to and activates a large number of proteins, especially those with signaling activities. The negative charge on the lipid facilitates the binding to proteins through electrostatic interactions or Ca^{2+} bridges. For example, the presence of appreciable amounts of phosphatidylserine on the cytosolic leaflet of endosomes and lysosomes enables these compartments to dock with proteins that possess specific phosphatidylserine-binding domains including several important signaling and fusogenic effectors. The cytoskeletal protein spectrin binds to phosphatidylserine in this way, and it is also required by enzymes such as the neutral sphingomyelinase and the Na^+/K^+ ATPase, where the 18:0/18:1 molecular species is especially important. It is believed that the fatty acyl components of this species in the inner leaflet of the plasma membrane (and potentially other intracellular membranes) may interact (interdigitation or "hand-shake") with the very-long chains of sphingolipids in the outer leaflet in raft microdomains, resulting in a high local concentration of the anionic phospholipid and an accumulation of negative surface charge to which specific poly-cationic proteins in the membranes can bind. This may then enable the transfer of signals across the membrane to the cytosol.

Similarly, phosphatidylserine participates directly in key signaling pathways in the brain by binding to the cytosolic proteins involved in neuronal signaling and thereby activating them. At least three major pathways are affected, including those involving phosphatidylinositol 3-kinase and protein kinase C. For example, most enzymes of the protein kinase C family contain a 'C2' calcium-dependent cysteine-rich region that recognizes phosphatidylserine, and in coordination with the 'C1' domain that binds to **diacylglycerols**, is essential for activating and locating them to the plasma membrane of appropriately stimulated cells. Phosphatidylserine is not involved in cell signaling through the formation of metabolites, as is the case with phosphatidylinositol.

Blood coagulation: Phosphatidylserine is an important element of the blood coagulation process in platelets, where it is transported from the inner to the outer surface of the plasma membrane in platelets activated by exposure to fibrin-binding receptors, for example. Here, the exposed phosphatidylserine enhances the activation of prothrombin to thrombin (the key molecule in the blood clotting cascade) by triggering a cascade of reactions and providing the negatively charged platform that enables calcium ions to form bridges with γ-carboxyglutamic acid-containing domains on the coagulation factors. Membrane vesicles with phosphatidylserine exposed on the surface can also be released from platelets and promote the coagulation process. Apolipoprotein A-1 in high-density lipoproteins has a controlling function in that it neutralizes these procoagulant properties by arranging the phospholipid in surface areas that are too small to accommodate the prothrombinase complex. Blood coagulation is beneficial when it prevents the loss of blood from the circulatory system, but it is detrimental when it causes thrombosis, and the action of phosphatidylserine is essential to the regulation of the process.

Apoptosis: In addition in response to particular calcium-dependent stimuli, phosphatidylserine is known to have an important role in the regulation of apoptosis or programmed cell death, the natural process by which aged or damaged cells are removed from tissues before they can exert harmful effects. When cells are damaged, a mechanism is initiated in which the normal distribution of this lipid on the inner leaflet of the plasma membrane bilayer is disrupted by stimulation of scramblases, which are ATP-independent and can move the lipid across the membrane to the outer leaflet. This occurs together with the inhibition of aminophospholipid translocases, which return the lipid to the inner side of the membrane. In erythrocytes, phosphatidylserine is located in the inner leaflet of the membrane bilayer under low Ca^{2+} conditions when a phospholipid scramblase is suppressed by membrane cholesterol, but it is exposed to the outer leaflet under elevated Ca^{2+} concentrations which activate the scramblase. After the collapse of this asymmetry and transfer of phosphatidylserine to the outer leaflet of an effete cell, it is believed that it is recognized by a cohort of receptors, either directly or indirectly, through bridging ligands on the surface of macrophages and related scavenger cells. These activate a family of cysteine-dependent aspartate-specific proteases, the caspases, and other enzymes to facilitate the engulfment of the apoptotic cells and their potentially toxic or immunogenic contents in a non-inflammatory manner. It is noteworthy that the transition from a pro-inflammatory to an anti-inflammatory state is defined by phagocytosis of neutrophils by macrophages via this phosphatidylserine-dependent process.

During apoptosis, the generation of reactive oxygen species occurs, mainly hydrogen peroxide, which together with the enzyme cytochrome c brings about rapid oxidation of the fatty acids in phosphatidylserine before this lipid is externalized. Indeed, it is now apparent that molecular species of phosphatidylserine with an oxidatively truncated *sn*-2 acyl group that incorporates terminal γ-hydroxy(or oxo)-α,β-unsaturated acyl moieties are especially potent signals for scavenger receptors in macrophages as a prerequisite for engulfment of apoptotic cells.

This has been described as "a dominant and evolutionarily conserved immunosuppressive signal that promotes tolerance and prevents local and systemic immune activation" or more succinctly as an "eat-me signal" (externalized phosphatidylinositol 3,4,5-trisphosphate (PI(3,4,5)P) may have a similar function). The binding of phosphatidylserine to specific proteins, such as apolipoprotein H (β2-glycoprotein 1), enhances the recognition and clearance. This process is essential for the development of the lung and brain, and it is also relevant to clinical situations where apoptosis plays an important part, such as cancer, chronic autoimmunity, and infections. For example, phosphatidylserine is a necessary component of the TAM family of receptor tyrosine kinases and the receptor-ligand complex of particular importance in cancer cells, where phosphatidylserine-TAM signaling regulates many aspects of inflammation and immune resolution and is seen as a target for therapeutic intervention. Exposure of phosphatidylserine is increased substantially on the surface of tumor cells or tumor cell-derived microvesicles, which have innate immunosuppressive properties and facilitate tumor growth and metastasis. Targeting phosphatidylserine is considered to be a promising strategy in cancer immunotherapy. In relation to atherosclerosis, phosphatidylserine is believed to have anti-inflammatory and protective effects as a component of the high-density lipoproteins, probably mediated by the apoptosis mechanism. In contrast, as this mechanism is important for the turnover of erythrocytes, it is relevant to thrombus formation and the stabilization of blood clots. The innate immunosuppressive effect of externalized phosphatidylserine has been hijacked by numerous viruses and bacteria to facilitate infection.

A similar apoptotic mechanism operates in retinal pigment epithelial cells to remove the large amounts of photoreceptor cell debris that are generated continuously. In addition, appreciable amounts of phosphatidylserine are translocated by an analogous mechanism to the surface of T lymphocytes that express low levels of the trans-membrane enzyme tyrosine phosphatase. This change in distribution acts then as a signaling mechanism to modulate the activities of several membrane proteins. The anti-coagulant protein annexin V binds with high specificity to phosphatidylserine and is used as a probe to detect apoptotic cells. It is noteworthy that phosphatidylserine is a major component of the membranes of microvesicles in animal cells, and translocation to the outer leaflet upon cellular activation is essential for their biogenesis. In addition, exposure of phosphatidylserine on the cell surface is reported to be a factor in non-apoptotic forms of regulated inflammatory cell death, such as necroptosis.

Role in infections: Unfortunately, viruses such as Ebola and HIV viruses can hijack this apoptosis machinery by incorporating phosphatidylserine into their viral envelopes so conning cells into engulfing them; the viral glycoprotein/cellular receptor complex may then facilitate the entry of foreign organisms into other cells. Similarly, parasites ingested in this manner, including *Leishmania* and *Trypanosoma* species, utilize host phosphatidylserine to establish infections and facilitate disease progression as they do not then elicit the production of proinflammatory cytokines. This mechanism has been termed 'apoptotic mimicry' and is critical for the survival of parasites within the macrophage.

Other activities: Phosphatidylserine is required for the transmembrane movement of excess cholesterol, derived initially from the lysosomal degradation of low-density lipoproteins, from the plasma membrane to the endoplasmic reticulum thereby maintaining membrane integrity and ensuring cell survival. It is therefore an important element in **cholesterol homeostasis**. The mechanism is believed to involve proteins known as GRAMD1s embedded in the endoplasmic reticulum membrane at sites in contact with the plasma membrane. These have two functional domains: the StART-like domain that binds cholesterol and the GRAM domain that binds anionic lipids, such as phosphatidylserine, and so forms a link between the two membranes that enables the transfer of cholesterol.

A further unusual function of phosphatidylserine is that it is a key component of the lipid-calcium-phosphate complexes that act as nucleation centers for hydroxyapatite formation and initiate mineral deposition during the formation of bone. It has been established that phosphatidylserine and inorganic phosphate must be present, before calcium ions are introduced, when the high affinity of phosphatidylserine for calcium ions becomes important. Nucleation is facilitated by the protein annexin V. Similarly, during bone repair and maintenance, the fusion of osteoclasts requires the non-apoptotic exposure of phosphatidylserine at the surface of fusion-committed cells with the aid of a transmembrane protein (DC-STAMP) expressed in dendrocytes. This activity is relevant to cardiovascular disease and in particular to the phenomenon of "hardening of the arteries," where atherosclerotic plaques can undergo mineralization with the deposition of hydroxyapatite.

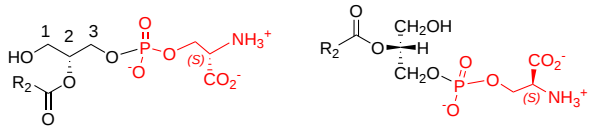
Among many other functions of phosphatidylserine, it is believed to be an essential surface membrane component for the fusion of cell types other than osteoclasts, including during the formation of fibers in muscle cells, and the fusion of macrophages into inflammatory giant cells and myoblasts into myotubes. Such cell fusions require the non-apoptotic exposure of phosphatidylserine at the surface of fusing cells, where it interacts with phosphatidylserine-recognizing proteins to regulate the time and place of cell fusion. Phosphatidylserine provides stable membrane domains in spermatozoa that are essential for fertilization, and it is also an essential component of the plasma membrane microdomains known as **caveolae**, where it is required both for their formation and stability possibly through specific binding to the cavin proteins.

The high concentrations of docosahexaenoic acid (DHA) in the brain and retinal phosphatidylserine are certainly important for the development and function of these tissues. Accumulation of phosphatidylserine in neuronal membranes is promoted by DHA, and this is important for the maintenance of neuronal survival. Phosphatidylserine may also be a reservoir of DHA for **protectin** formation in neuronal tissue. On the other hand, the Food and Drug Administration in the USA considers that there is little scientific evidence to support claims that dietary supplements of phosphatidylserine reduce the risk of dementia or cognitive dysfunction in the elderly, and other nutritional claims appear to be dubious also. Antibodies to phosphatidylserine are formed in some disease states, including thrombosis and recurrent spontaneous pregnancy loss. The rare genetic disease Lenz-Majewski syndrome is caused by a mutation in the phosphatidylserine synthase I gene that greatly increases the activity of the enzyme while preventing feedback inhibition, and abnormal metabolism of phosphatidylserine has been implicated in other diseases.

In yeasts such as *Candida albicans*, phosphatidylserine and the enzyme phosphatidylserine decarboxylase, which generates phosphatidylethanolamine, are both essential for the virulence of the organism towards a host species.

21.3.5.4: LYSOPHOSPHATIDYLSERINE

Figure 21.3.30 shows the structure of lysophosphatidylserine



lysophosphatidylserine

Figure 21.3.30 lysophosphatidylserine

Lysophosphatidylserine, i.e., with a fatty acid in one position only, is known to be a mediator of a number of biological processes, especially in the context of the immune system in animal tissues. It has been found in the thymus, peripheral lymphoid tissues, central nervous system, and colon, but is barely detectable in plasma. Deacylation of the diacyl lipid by phospholipases is the primary source. For example, a secreted isoform that is phosphatidylserine-specific (PLA1A) removes the *sn*-1 acyl group of phosphatidylserine to generate *sn*-2-lysophosphatidylserine containing unsaturated fatty acids, and this is upregulated greatly by various inflammatory stimuli. This extracellular enzyme utilizes phosphatidylserine exposed on the cell membrane as a substrate, although other phospholipases may operate intracellularly and produce *sn*-1-lysophosphatidylserine. In addition, platelets in some species (not significantly in humans) secrete a phospholipase A_2 group IIA (ABHD16A), which generates saturated *sn*-1-lysophosphatidylserine (and other lysophospholipids).

Lysophosphatidylserine has been detected after injury to animal tissues (tumor growth, graft rejection, burns), and it may have a similar function to **lysophosphatidic acid** in cell signaling, for example in regulating calcium flux and stimulating immune cells through G protein-coupled receptors of which three (GPR34, P2Y10 and GPR174, LPS1-3) have been detected in mice and humans. For example, GPR174 mediates the suppression of T-cell proliferation induced *in vitro* by lysophosphatidylserine. When cells are damaged, lysophosphatidylserine can be generated by a reaction dependent on the activation of the NADPH oxidase. It can diffuse and transmit the information to other cells,

especially mast cells, and it is produced to enhance the clearance of activated and dying neutrophils. It thus has a role in the resolution of inflammation. One specific molecular species, i.e., 1-(11*Z*-eicosenoyl)-glycero-3-phosphoserine, is reported to be a true agonist of the Toll-like receptor 2/6 heterodimer of importance to the immune response to pathogens; both its polar head group and the length of the acyl chain are required for this activity. On the other hand, *sn*-2-lysophosphatidylserine has proinflammatory reactions in that it augments mast cell degranulation and mast cell-dependent anaphylactic shock; most other lysophospholipids have no such activity.

Deregulated lysophosphatidylserine metabolism has been linked to certain cancers, cardio-metabolic disorders, night blindness, and the human genetic neurological disorder PHARC. High serum levels of PLA1A are associated with such autoimmune disorders as Graves' disease and systemic lupus erythematosus, and there is increased expression of the enzyme in metastatic melanomas. It is necessary for the assembly of the hepatitis C virus, and it can play a role in the antivirus innate immune response. In Schistosome infections, lysophosphatidylserine from the parasite is believed to be a key activator molecule in the host.

Negatively charged lysophosphatidylserine species tend to organize in non-bilayer structures and are believed to facilitate the folding of certain membrane proteins *in situ* better than bilayer-forming lipids.

21.3.6: PHOSPHATIDYLINOSITOL AND RELATED PHOSPHOINOSITIDES

Although it had long been recognized that phosphatidylinositol or 1,2-diacyl-*sn*-glycero-3-phospho-(1'-*myo*-inositol) was a key membrane constituent, it was initially something of a surprise when the manifold biological activities of this lipid, and then of the derived phosphatidylinositol phosphates and their hydrolysis products, were discovered in animals, plants, and microorganisms. Many years after the initial discoveries in the 1950s, these lipids continue to be a major focus for research efforts around the world with considerable relevance to human health. Phosphatidylinositol and its various metabolites and relevant enzymes can be located and function within different membrane regions in cells, and they form part of what have been termed phosphoinositide and phosphatidylinositol cycles, their versatility stemming from the inositol head group, a six-carbon the hexahydroxy ring, which can be reversibly phosphorylated on the 3, 4 and 5 positions. In addition to their structural role in membranes, these lipids are intimately involved in innumerable aspects of membrane trafficking and signaling in eukaryotic cells, functions that are essential to cell growth and metabolism. Only a brief overview of such a highly complex topic is possible here. **Glycosyl-phosphatidylinositol (GPI)** is a related lipid that serves as an anchor for proteins.

21.3.6.1: PHOSPHATIDYLINOSITOL

Structure and Occurrence: Phosphatidylinositol is an important lipid, both as a membrane constituent and as a participant in essential metabolic processes in all plants and animals, both directly and via a number of its metabolites. It is an acidic (anionic) phospholipid that in essence consists of a phosphatidic acid backbone, linked via the phosphate group to inositol (hexahydroxycyclohexane). In most organisms, the stereochemical form of the last is *myo*-D-inositol (with one axial hydroxyl in position 2 with the remainder equatorial, i.e. a chair-like structure), although other forms (*scyllo*- and *chiro*-) have been found on occasion in plants. The 1-stearoyl,2-arachidonoyl molecular species, which is of considerable biological importance in animals, is illustrated.

Figure 21.3.31 shows the structure of phosphatidylinositol.

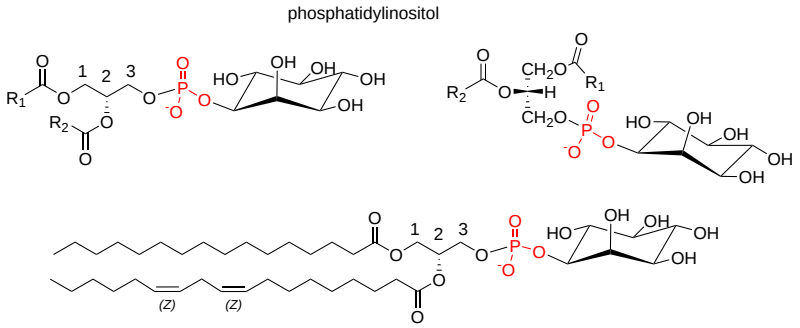


Figure 21.3.31: Phosphatidylinositol

Phosphatidylinositol is especially abundant in brain tissue, where it can amount to 10% of the phospholipids, but it is present in all tissues, cell types, and membranes at relatively low levels in comparison to many other phospholipids. In rat liver, it amounts to 1.7 micromoles/g., i.e. less than phosphatidylcholine, phosphatidylethanolamine, and phosphatidylserine. Under normal conditions, it is present entirely in the inner leaflet of the erythrocyte membrane and of the plasma membrane in nucleated cells. Phosphatidylinositol *per se* is rarely found in prokaryotes other than the *Actinomycetales*, although the thermophilic α-proteobacterium *Rhodothermus marinus* contains dialkylether glycerophosphoinositides.

The fatty acid composition of phosphatidylinositol is rather distinctive as shown in Table 21.3.6. Thus, in almost all animal tissues, the characteristic feature is a high content of stearic and arachidonic acids. All the stearic acid is linked to position *sn*-1 and all the arachidonic acid to position *sn*-2, and as much as 78% of the total lipid may consist of the single molecular species *sn*-1-stearoyl-*sn*-2-arachidonoyl-

glycerophosphorylinositol (see Table 21.3.7 below). Although 1-alkyl- and alkenyl- forms of phosphatidylinositol are known, they tend to be much less abundant than the diacyl form. In plant phosphatidylinositol, e.g. *Arabidopsis thaliana* as listed, palmitic acid is the main saturated fatty acid in position *sn*-1, while linoleic and linolenic acids are the main unsaturated components in position *sn*-2. Similarly in yeast, palmitic acid is in position *sn*-1 with oleic and palmitoleic acids in position *sn*-2 predominantly; the Amoebozoa have a C_{16} alkyl group in position *sn*-1 and *cis*-vaccenic acid in position *sn*-2.

Table 21.3.6: **Fatty acid composition of phosphatidylinositol (wt % of the total) in animal and plant tissues.**

Tissue	Fatty acids									
	16:0	18:0	18:1	18:2	18:3	20:3	20:4	22:3	22:5	22:6
Bovine brain [1]	8	38	10	1	-	5	34	2	tr.	1
Bovine liver [2]	5	32	12	6	1	7	23	4	3	5
Rat liver [3]	5	49	2	2		4	35			1
A. thaliana [4]	48	3	2	24	24					

[1] = Holub, B.J. *et al.*. *J. Lipid Res.*., **11**, 558-564 (1970); DOI. [2] = Thompson, W. and MacDonald, G., *J. Biol. Chem.*, **250**, 6779-6785 (1975); DOI. [3] = Wood, R. and Harlow, R.D. *Arch. Biochem. Biophys*., **135**, 272-281 (1969); DOI. [4] = Browse, J. *et al. Biochem. J.*, **235**, 25-31 (1986); DOI.

Biosynthesis: The basic mechanism for the biosynthesis of phosphatidylinositol and phosphatidylglycerol is sometimes termed a branch point in phospholipid synthesis, as phosphatidylcholine and phosphatidylethanolamine are produced by a somewhat different route.

Phosphatidylinositol is found in all eukaryotes, which are in general able to synthesize inositol *de novo* via glucose-6-phosphate. As with **phosphatidylglycerol** (and hence cardiolipin), phosphatidylinositol is formed biosynthetically from **phosphatidic acid** via the intermediate **cytidine diphosphate diacylglycerol**, which is produced by the action of a CDP-diacylglycerol synthase believed to be the rate-limiting enzyme in phosphatidylinositol biosynthesis. Then, the enzyme CDP-diacylglycerol inositol phosphatidyltransferase ('phosphatidylinositol synthase' or 'PIS') catalyzes a reaction with *myo*-inositol to produce phosphatidylinositol.

Figure 21.3.32 shows the synthesis of phosphatidylinositol in eukaryotes.

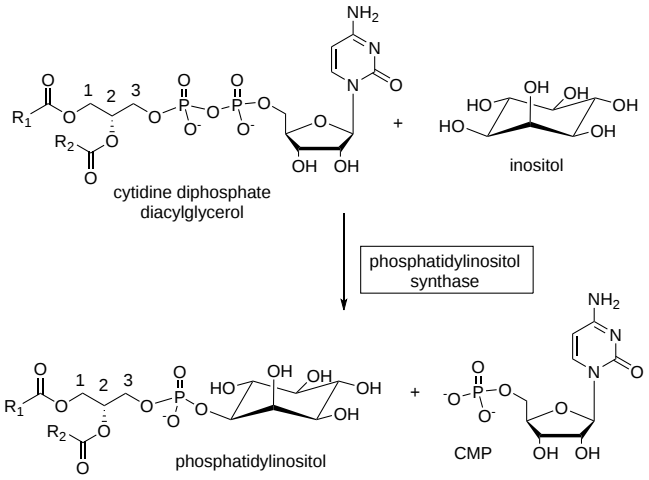


Figure 21.3.32 Synthesis of phosphatidylinositol in eukaryotes

Only isoform of PIS exists in mammals and it is located in the endoplasmic reticulum, in part in a subcompartment of this associated with mitochondria (mitochondria-associated membranes - MAM) and in mitochondria *per se*. Indeed, it is reported that PIS is present in a mobile ER-derived subcompartment that makes transient contacts with other organelles, including the plasma membrane, and facilitates the distribution of phosphatidylinositol to other subcellular compartments. The other product of the reaction is cytidine monophosphate (CMP). As PIS catalyzes the reverse reaction also, the rate of phosphatidylinositol synthesis is determined by the relative concentrations of the precursors and product, and the latter must be transported away from the site of synthesis for the reaction to continue. Much of the phosphatidylinositol is delivered to other membranes by vesicular transport, but a family of soluble phosphatidylinositol transfer proteins (PITPα, PITPβ and PITPNC1) provides phosphatidylinositol from the ER to kinases for phosphorylation (see below).

Molecular species specificity: The phosphatidylinositol synthase *per se* does not exhibit the fatty acyl specificity observed in the final product, but earlier in the biosynthetic process 1-stearoyl-2-arachidonoyl species of diacyl-*sn*-glycerols are converted preferentially into **phosphatidic acid** by the epsilon isoform of **diacylglycerol kinase** (DGKε), anchored to the membrane via its N-terminal hydrophobic helix segment; ATP is the phosphate donor. In addition, one of the CDP-diacylglycerol synthases (CDS2) has similar specificity in the generation of the immediate precursor CDP-diacylglycerols from phosphatidic acid, while some specificity may be introduced via lysophosphatidylinositol, formed as a by-product of eicosanoid formation (see below) or as an intermediate as part of the normal cycle of

deacylation-acylation of phosphatidylinositol in tissues in which the fatty acid composition is remodeled to give the final distinctive composition. A membrane-bound *O*-acyltransferase (MBOAT7 or LPIAT1) specific for position *sn*-2 of lysophosphatidylinositol with a marked preference for arachidonoyl-CoA is ubiquitously expressed in animal tissues, and this may be one means by which free arachidonic acid and eicosanoid levels are regulated.

In macrophages subjected to inflammatory stimuli, phosphatidylinositol containing two molecules of arachidonate is produced by remodeling reactions, and there is evidence that it is a novel bioactive phospholipid regulating innate immune responses in these cells. Further specificity may be introduced by lysocardiolipin acyltransferase (LYCAT; also known as LCLAT1 or ALCAT1), which exhibits a preference for lysophosphatidylinositol and lysophosphatidylglycerol over other phospholipids *in vitro*, and incorporates 18:0 rather than shorter chain fatty acids into position *sn*-1 of phosphatidylinositol and other phosphoinositides, especially phosphatidylinositol-4,5-bisphosphate and phosphatidylinositol-3-phosphate; this enzyme may be located adjacent to the phosphatidylinositol synthase in the endoplasmic reticulum. Some of the phosphatidylinositol in membranes is derived from recycling of polyphosphoinositides via the phosphatidylinositol cycle, and this could influence the molecular species composition (see below).

The highly specific distribution of fatty acids on the glycerol moiety of phosphatidylinositol breaks down in some cancer cells, especially those with a mutation on the transcription factor p53 gene, which is one of the most highly mutated genes in cancers.

Plants and bacteria: In contrast to animals, **plants** have two phosphatidylinositol synthase isoforms, PIS1 and PIS2, which display specificities for particular species of the CDP-diacylglycerol substrate. PIS1 generates phosphatidylinositol with saturated or monounsaturated fatty acids preferentially, while PIS2 generates polyunsaturated species, the two forms possibly having different functions. In protozoan parasites, such as *Trypanosoma brucei*, the active site of phosphatidylinositol synthase may be the lumen of the endoplasmic reticulum and Golgi. There is evidence for two distinct pools of product in this organism, the bulk membrane form derived from inositol imported from the environment, and a second used for the synthesis of GPI anchors, which uses *myo*-inositol synthesized *de novo*. In yeasts, some biosynthesis may occur on the cytosolic side of the plasma membrane.

The enzyme is a transmembrane protein.and use CDP-diacylglycerol as a donor and either inositol (eukaryotes) or inositol phosphate (prokaryotes) as the acceptor alcohol. The structure of a similar enzyme, phosphatidylinositol-phosphate synthase from *Renibacterium salmoninarum*, is shown in Figure 21.3.33.

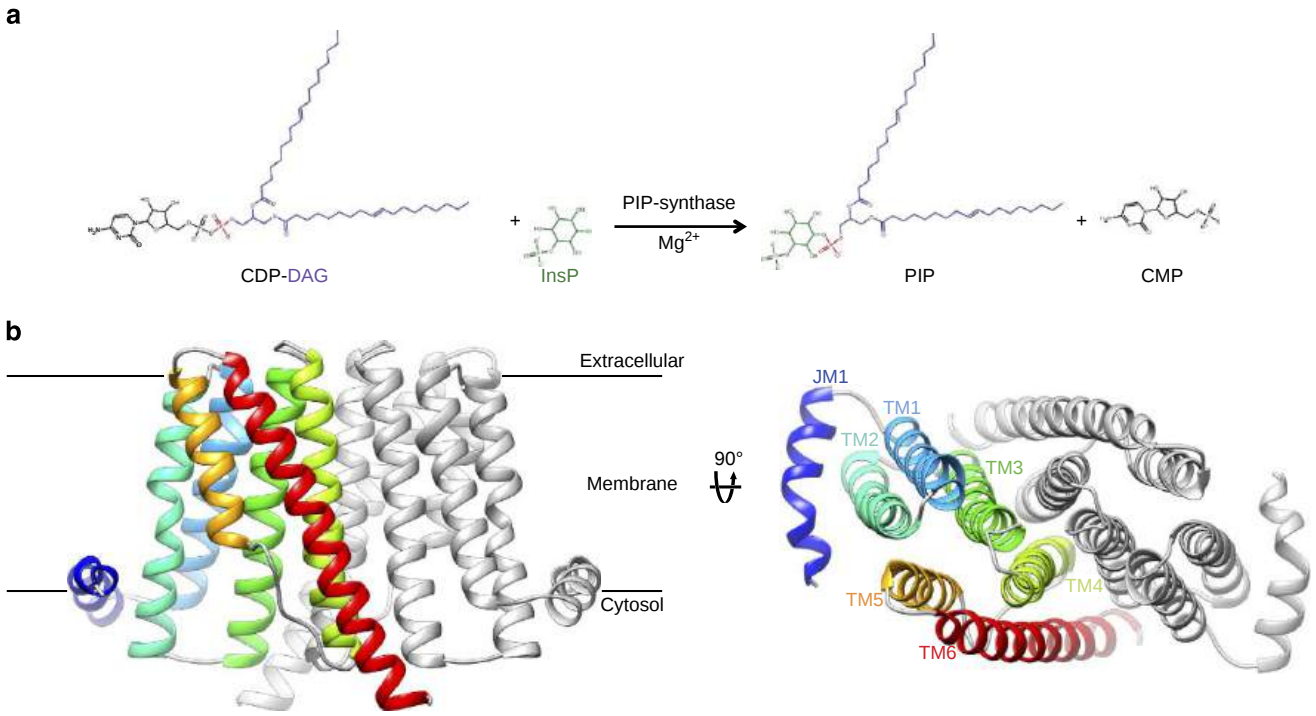


Figure 21.3.33: Structure and reaction of phosphatidylinositol-phosphate synthase from *Renibacterium salmoninarum.* Clarke, O., Tomasek, D., Jorge, C. *et al.* Structural basis for phosphatidylinositol-phosphate biosynthesis. *Nat Commun* **6,** 8505 (2015). https://doi.org/10.1038/ncomms9505. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/

Panel A shows the reaction for PIP synthases which involves the transfer of a diacylglycerol-substituted phosphate group (purple/red) from the CDP-DAG donor to the inositol phosphate acceptor (green), generating PIP and CMP. Panel B shows the structure of the *Rs*PIPS-Δ6N homodimer in ribbon representation viewed from two orthogonal orientations (in the plane of the membrane on the left; towards the cytosol

down the dimer axis on the right). One protomer is colored grey, and the helices of the other are depicted in spectral coloring, from blue (JM1) to red (TM6). The *Af*2299 extramembrane domain used to facilitate crystallization is not shown here.

Figure 21.3.34 shows a large cavity that contains the active site of RsPIPS.

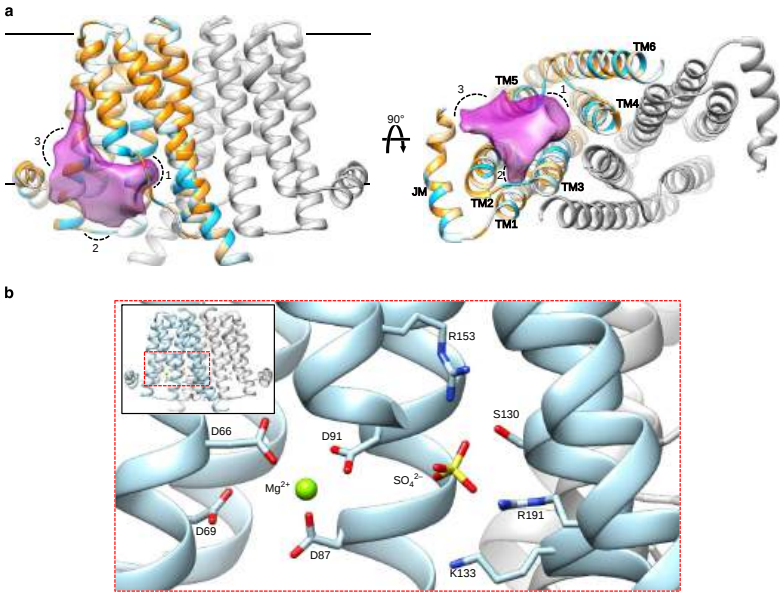
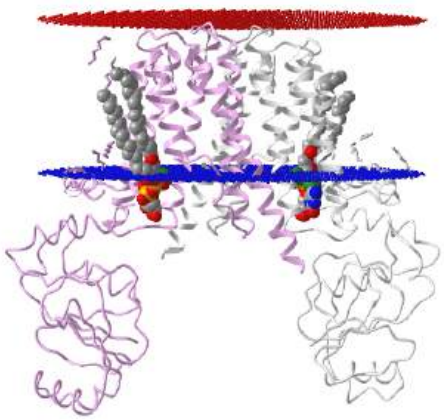


Figure 21.3.34 Large cavity containing the active site of RsPIPS . Clarke, O., Tomasek, D., Jorge, C. *et al.* ibid.

Panel (**a**) shows the structure of *Rs*PIPS-Δ6N is shown in ribbon representation, with one protomer colored grey and the other colored by the Kyte–Doolitle hydrophobicity scale, from −4.5 (most polar, light blue) to 4.5 (most hydrophobic, orange). Two orthogonal representations are shown, on the left is a view in the plane of the membrane, and on the right is a view from the cytosol along the dimer axis. A transparent purple surface delineates the borders of the interfacial cavity, which contains three subregions as follows: 1, the inositol phosphate acceptor-binding pocket; 2, the nucleotide-binding pocket between TM2 and TM3; and 3, a hydrophobic groove between TM2 and JM1. (**b**) Detail of the active site viewed in the plane of the membrane, with side chains that contact the bound Mg^{2+} and SO_4^{2-} ions labeled and depicted in stick representation.

A nucleotide-binding site formed from transmembrane segments 1, 2, and 3 contains 8 conserved residues (D_1xx$D_2$$G_1$xxAR… G_2xxxD_3xxxD_4). The first 3 aspartic acid side chains coordinate a metal ion while the 4th is likely a general base in catalysis.

Figure 21.3.35 shows an interactive iCn3D model of the phosphatidylinositolphosphate (PIP) synthase with bound CDP-DAG from Renibacterium salmoninarum (5D92).



NCBI iCn3D Figure 21.3.35: Phosphatidylinositolphosphate (PIP) synthase with bound CDP-DAG from Renibacterium salmoninarum (5D92).. Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...i9nhSw9U32Man8

The two identical subunits of the homodimer are shown in gray and plum. The two CDP-DAGs are shown in spacefill and CPK colors

Function: In addition to functioning as negatively charged building blocks of membranes, the inositol phospholipids (including the phosphatidylinositol phosphates or 'polyphosphoinositides' discussed below) have crucial roles in the interfacial binding of proteins and in the regulation of protein activity at the cell interface. As phosphoinositides are polyanionic, they can be very effective in non-specific

electrostatic interactions with proteins. However, they are especially efficient in specific binding to so-called 'PH' domains of cellular proteins. At least three phosphatidylinositol molecules are present in the crystal structure of human erythrocyte glycophorin, for example, and they are believed to influence binding to other proteins via their head groups. The lipid is a structural component of yeast cytochrome bc_1.

In animal tissues, phosphatidylinositol is the primary source of the arachidonic acid required for the biosynthesis of **eicosanoids**, including prostaglandins, via the action of the enzyme phospholipase A_2, which releases the fatty acids from position *sn*-2. The reverse reaction also occurs.

Figure 21.3.36s shows the generation of arachidonic acid and eicosanoids from PI by means of phospholipase A2

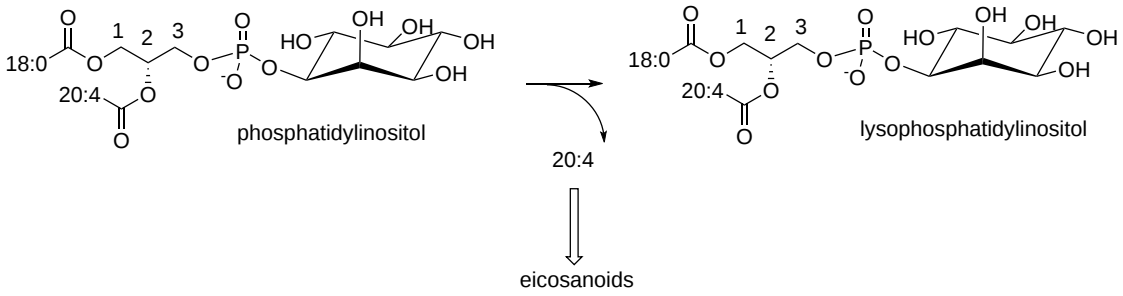


Figure 21.3.36: Generation of arachidonic acid and eicosanoids from PI by means of phospholipase A2.

Similarly, phosphatidylinositol and the phosphatidylinositol phosphates are the main sources of diacylglycerols that serve as signaling molecules in animal and plant cells, via the action of a family of highly specific enzymes collectively known as phospholipase C. In brief, diacylglycerols regulate the activity of a group of at least a dozen related enzymes known as protein kinase C, which in turn control many key cellular functions, including differentiation, proliferation, metabolism, and apoptosis. Indeed, the biological actions of the various components released have been the subject of intensive study over many years. **2-Arachidonoylglycerol**, an endogenous cannabinoid receptor ligand, may also be a product of phosphatidylinositol catabolism.

21.3.6.2: PHOSPHATIDYLINOSITOL PHOSPHATES (POLYPHOSPHOINOSITIDES) IN ANIMALS

Structure and Occurrence: The pioneering work of Mable and Lowell Hokin in the 1950s led to the discovery that phosphatidylinositol was converted to polyphosphoinositides with important signaling and other functional activities, including cell communication via signal transduction, cell survival and proliferation, membrane trafficking and modulation of gene expression. Phosphatidylinositol is now known to be phosphorylated by a number of substrate-selective kinases that place the phosphate moiety on positions 3, 4, and 5 of the inositol ring with the balance among them maintained by distinct phosphatases and phospholipases. Seven different isomers are known (mono-, bis-, and tris-phosphorylated), which are produced in a tightly coordinated manner, and all of these have characteristic biological activities. They each turn over much more rapidly than the parent phosphatidylinositol molecule. In addition, there can be an array of molecular species of each of these isomers that differ in the nature of the fatty acyl groups. Although the most significant in quantitative and possibly biological terms were long thought to be **phosphatidylinositol 4-phosphate** and **phosphatidylinositol 4,5-bisphosphate**, it is now recognized that **phosphatidylinositol 3-phosphate** and its metabolites are as important biologically at least.

Figure 21.3.37 shows the structures of phosphatidylinositol phosphates

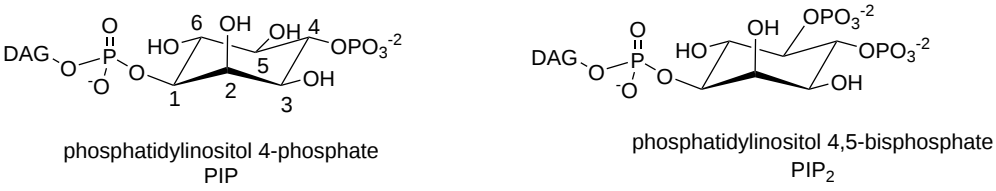


Figure 21.3.37: Structures of phosphatidylinositol phosphates in animals

These lipids are usually present at low levels only in tissues, typically at about 0.5 to 1% of the total lipids of the inner leaflet of the plasma membrane, so they are unlikely to have an appreciable structural role. On the other hand, static measurements of lipids that turn over very rapidly do not provide a meaningful assessment of their cellular functions. The positional distributions of fatty acids in the phosphatidylinositol, phosphatidylinositol 4-phosphate, and phosphatidylinositol 4,5-bisphosphate of ox brain are listed in Table 21.3.7. In each the saturated fatty acids are concentrated in position *sn*-1 and polyunsaturated, especially arachidonate, in position *sn*-2; there are few differences among the three lipids in this instance.

Table 21.3.7: **Distribution of fatty acids (mol % of the total) in positions *sn*-1 and *sn*-2 in phosphatidylinositol (PI) and the phosphatidylinositol mono- and diphosphates of ox brain.**

Fatty acids	PI		PI monophosphate		PI diphosphate	
	sn-1	*sn*-2	*sn*-1	*sn*-2	*sn*-1	*sn*-2
16:0	15		9		7	
18:0	74		69		69	
18:1	10	10	20	13	21	10
18:2	1	2	trace	1	1	1
20:3(*n*-9)		5		10		10
20:3(*n*-6)		5		11		12
20:4(*n*-6)		67		49		52
22:3		7		10		7
22:6(*n*-3)		trace		trace		trace

Data from Holub, B.J. *et al.*, *J. Lipid Res.*, **11**, 558-564 (1970); DOI.
Molecular species data, see Traynor-Kaplan, A. *et al.*, *Biochim. Biophys. Acta*, **1862**, 513-522 (2017); DOI.

Biosynthesis: Phosphatidylinositol *per se* is the ultimate precursor of all phosphoinositides, the head groups of which have different charges and structures that impact directly on membrane properties and via metabolic interactions can function as chemical switches. The individual phosphoinositides are maintained at steady state levels in membranes by a continuous and sequential series of phosphorylation and dephosphorylation reactions by specific kinases, phosphatases, and phospholipase C enzymes, which are regulated and/or relocated through cell surface receptors for extracellular ligands, the phosphoinositide cycle. While this has been termed a 'futile cycle', which can consume a significant proportion of cellular ATP production, it is only part of a wider pattern of reactions - the **phosphatidylinositol cycle** (see below). Controlled synthesis of these different phosphoinositides occurs in different intracellular compartments for distinct and independently regulated functions with spacially distinct target enzymes or receptors. In mammals, the complexity is such that 18 phosphoinositide inter-conversion reactions have been identified to date, and these are mediated by at least 20 phosphoinositide kinases and 34 phosphoinositide phosphatases that span 8 and 10 classes, respectively; some have yet to be characterized. Most of these enzymes are conserved across all of the eukaryota, and each has distinct functions and specificities that cannot be replaced by the activity of related isoforms.

As a generality, most mono-phosphorylations occur in endomembranes, such as the endosomes and the Golgi network, while second and third phosphorylations occur primarily at the plasma membrane, and this is reflected in the lipid composition of each membrane. While these enzymes are believed to work independently and sequentially to produce a specific product, there remains a possibility that some participate in protein complexes to coordinate their activities. Specific transporters, especially the 'Nir2' protein, facilitate the exchange of phosphoinositides between membranes. It should be noted that there are links to the metabolism of **phosphatidylcholine**, which can be hydrolyzed by phospholipase D to phosphatidic acid, an important activator of key kinases. Figure 21.3.38 shows an overview of polyphosphoinositide metabolism in animal tissues.

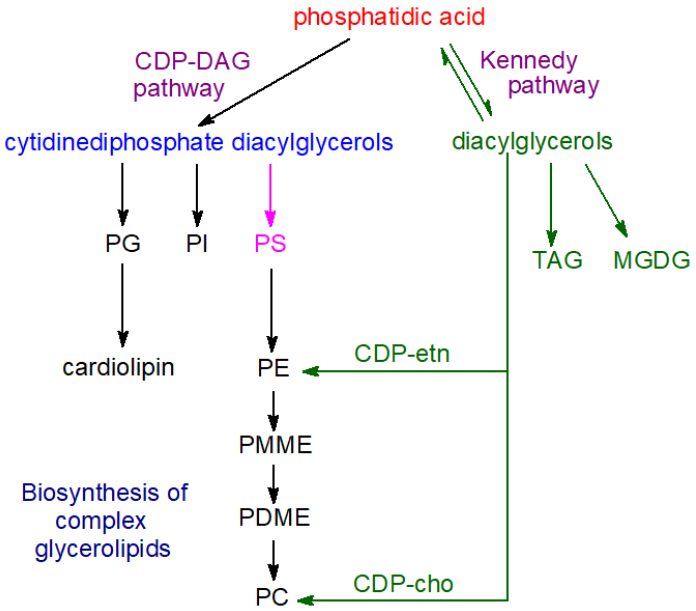


Figure 21.3.38 Overview of polyphosphoinositide metabolism in animal tissues

Thus as an example, phosphatidylinositol 4-phosphate (PI(4)P) is produced by the action of a phosphatidylinositol 4-kinase (PI4K) in the Golgi, and is in turn phosphorylated by a phosphatidylinositol phosphate 5-kinase (PIPK I) to form phosphatidylinositol 4,5-bisphosphate (PI(4,5)P) at the plasma membrane, although this can also be formed by phosphorylation of phosphatidylinositol 5-phosphate by a specific 4-kinase (PIPK II). Four isoforms of PI4K in two structural families are known that each operate in different subcellular membrane compartments to produce phosphatidylinositol 4-phosphate for particular signaling functions. Some selectivity in the formation of molecular species or remodeling may occur to further enrich the arachidonic acid content.

Subsequently, it was discovered that phosphatidylinositol is also phosphorylated by a 3-kinase (PI3K III or the VPS 34 complex) to produce phosphatidylinositol 3-phosphate (PI(3)P) in the early endosomes. In fact, three phosphatidylinositol 3-kinases families (eight isoforms) have been described, each with distinct substrate specificities. A second phosphoinositide signaling pathway involves activation of two of these 3-kinases, stimulated by growth factors and hormones, which phosphorylate phosphatidylinositol 4,5-bisphosphate (by PI3K I - four isoforms) and phosphatidylinositol 4-phosphate (by PI3K II - three isoforms) to produce phosphatidylinositol 3,4,5-trisphosphate (PI(3,4,5)P) and phosphatidylinositol 3,4-bisphosphate (PI(3,4)P), respectively. While phosphatidylinositol 3-phosphate and other 3-phosphorylated metabolites amount to only about 0.5% of the total phosphoinositides in resting mammalian cells, they are now recognized to be of profound importance for cellular metabolism.

In addition to the activity of kinases, the amounts of these various metabolites are regulated by the activities of specific phosphoinositide phosphatases, which are highly conserved in eukaryotes and dephosphorylate phosphoinositides at the 3, 4, and 5 positions of the inositol ring. For example, so-called ‘SHIP’ phosphatases convert phosphatidylinositol 4,5-bisphosphate back to phosphatidylinositol 4-phosphate by hydrolysis of the 5-phosphate group. 3-Phosphorylated phosphoinositides are only degraded by phosphatases, especially those of the PTEN family, and not by phospholipase C (see below).

The various organelles in cells have membranes with distinct functions and molecular compositions. Yet, all the phosphatidylinositol precursor is formed primarily at the endoplasmic reticulum, and the different membrane lipids must be transported between membrane sites via specific trafficking processes/proteins. There is selective recruitment of effector proteins to particular membranes by binding only to a single type of phosphoinositide, and this is followed by interactions between the phosphoinositide-binding proteins and various enzymes to channel phosphoinositide production to the required biological outcomes and to regulate signaling. For example, much of the phosphatidylinositol 4-phosphate and phosphatidylinositol 4,5-bisphosphate involved in signaling is believed to be formed at contact sites between the endoplasmic reticulum and plasma membrane.

A concept has emerged in which each phosphoinositide has its own role – **the ‘lipid code’ hypothesis**, in which defined lipids act as labels for each cellular membrane to organize cells into dynamic and responsive membrane-bound compartments and maintain the orderly flow required for the complexities of membrane trafficking and spatio-temporal signaling reactions. Thus, phosphatidylinositol 4-phosphate, phosphatidylinositol 4,5-bisphosphate, phosphatidylinositol 3-phosphate and phosphatidylinositol 3,5-bisphosphate are found mainly on the Golgi, plasma membrane, early endosomes, and late endocytic organelles, respectively, where they are sometimes regarded as landmarks for these compartments. For example, phosphatidylinositol 4,5-bisphosphate is present throughout the plasma membrane and is considered a general marker for this, while phosphatidylinositol 3,4,5-triphosphate, is a characteristic component of the basolateral region of this membrane in a polarized cell but is absent from the apical part. On the other hand, it should be noted that this map of phosphoinositides to specific organelles is derived from their steady state distributions, but the highly dynamic generation and consumption of different phosphoinositides in response to different stimuli in the various sub-cellular compartments in living cells by the action of kinases and phosphatases together with lipase reactions, may lead to the formation of transient pools of distinct molecular forms. There must be a continuous replenishment of the precursors by new synthesis.

Function: The distinctive subcellular location of the different phosphoinositide species, together with the rapid and reversible nature of phosphorylation, gives them a central and general position in the fields of cell signaling cascades and intracellular membrane trafficking. The precise locations of particular phosphoinositides are factors that contribute a specific identity to each organelle and sometimes even to each face of an organelle, such as the *cis* and *trans* faces of the Golgi apparatus, and this enables directional transport of cellular constituents between organelles or membranes. Phosphoinositides are able to achieve signaling effects directly by binding to specific cytosolic domains of membrane proteins via their polar head groups, thereby triggering downstream signaling cascades, often in conjunction with an acidic phospholipid, such as phosphatidylserine or phosphatidic acid at an adjacent-binding site. The term 'lipidon' has been coined to describe the unique collection of co-located lipids that distinguish biological membrane nano-environments and which provide the context for PI recognition *in vivo*. In this way, they can regulate the function of innumerable proteins integral to membranes, for example by relocating a protein from one area of the cell to another, e.g., from the cytosol to the inner leaflet of the plasma membrane, or they can attract cytoskeletal and signaling components to the membrane. Amongst the proteins that bind to phosphoinositides in this way are phospholipases, protein kinases, regulators of membrane trafficking, and cytoskeletal, scaffold, and ion channel proteins. Dysregulation of phosphoinositide metabolism and signaling is a factor in a number of diseases, including cancer.

Binding usually involves electrostatic interactions with the negative charges of the phosphate groups on the inositol ring with characteristic clusters of basic amino acid residues in proteins to recruit them to intracellular membranes, while often leading to specific folding and thence increased activity of unstructured peptides. At least 70 distinct types of binding sites for phosphoinositides have been identified in

proteins. In particular, a binding region termed the pleckstrin homology (PH) domain, consisting of ~100 amino acids, is the most abundant lipid-binding domain with more than 225 examples identified, and this can exhibit great specificity for particular polyphosphoinositides, often binding simultaneously with other proteins. While the interaction is driven by non-specific electrostatic interactions initially, it is followed by specific binding to increase the membrane residence time. The phox homology (PX) domain family with 49 members in humans is unique in that it can recognize all seven phosphoinositide forms, while proteins with a FYVE domain, which is enriched in cysteine and is stabilized by two zinc atoms, bind specifically to phosphatidylinositol 3-phosphate (PI(3)P). The protein kinase C family have C1 or C2 domains which recognize phosphatidylinositol 4,5-bisphosphate and phosphatidylinositol 3,4,5-trisphosphate specifically (and sometimes other lipids). The distinctive phosphoinositide composition of membranes in different organelles adds strength and specificity to the interactions by cooperative binding with other membrane proteins.

Phosphatidylinositol 3-phosphate and the other phosphatidylinositol monophosphates are present in cells at low levels only, although their levels do not appear to fluctuate greatly. PI(3)P has been implicated in membrane trafficking through its interactions with certain proteins in endosomes. In particular, it plays a pivotal role in the initiation of autophagy, i.e. the controlled internal degradation and turnover of cellular constituents, while PI(3,5)P_2 is important in the autophagosome–lysosome fusion step and in the subsequent acidification of this organelle. After sorting of the lysosomal contents, components of the internalized cargo are recycled to the plasma membrane and PI(3)P is dephosphorylated to phosphatidylinositol by a specific phosphatase, and this is in turn phosphorylated to PI(4)P. Thus the processes of internalization, sorting, and trafficking of membrane proteins depend on the interconversion of phosphoinositide species by coordinated phosphorylation-dephosphorylation reactions.

In general, PI(3)P controls cellular processes by recruiting effector proteins through low to moderate affinity interaction with specific PI(3)P binding domains. A protein designated Akt (protein kinase B) is recognized as a direct effector of the PI3K signaling cascade with receptor tyrosine kinases as the main upstream activators, for example, but it is now known that every phosphatidylinositol phosphate has a specific set of effector proteins that are recruited to target membranes or are allosterically regulated by the specific receptors; each function may require a different effector. A further function of PI(3)P is in the regulation of the final stage of cell division (cytokinesis), and the lipid is known to accumulate where cells divide. As the class I PI3K isoforms especially have been implicated in the etiology and maintenance of various diseases and metabolic disorders, including cancer, inflammation, and autoimmunity, drug companies are actively pursuing the development of inhibitors. In particular, they mediate insulin-independent glucose transport and many of the physiological actions of insulin. In relation to lung cancer especially, RAS proteins, which are key signaling switches essential for the control of proliferation, differentiation, and survival of eukaryotic cells, regulate the activity of type I phosphatidylinositol 3-kinase (PI3K); this is essential for tumor initiation and maintenance.

Phosphatidylinositol 4-phosphate is the precursor for the 4,5-bisphosphate, but it binds to a protein on the cytoskeleton of the cell and has its own characteristic functions. It is the most widely distributed of the phosphoinositides, and in addition to the Golgi and the plasma membrane, it is present in late endosomes, lysosomes, secretory vesicles, and autophagosomes. As a part of protein-lipid complexes, it is believed to have a role in essential nuclear processes. In yeast, it has a function in the anterograde transport from the trans-Golgi and the retrograde transport from the Golgi to the endoplasmic reticulum; it is also necessary for the formation of secretory vesicles in the Golgi that are targeted to the plasma membrane. Some of that in the plasma membrane is exchanged for phosphatidylserine by the action of specific transport proteins at junctions with the endoplasmic reticulum.

In addition, PI(4)P is essential for the structure and function of the late endosomes, where it is required for the recruitment of specific proteins that control cargo exit (following hydrolysis of PI(3)P). Some of these participate in vesicle formation, while others like the oxysterol binding protein (OSBP) are involved in lipid transfer. After initiation of the process by PI(3)P, PI(4)P, PI(4,5)P_2 and their binding proteins are modulators of autophagy at most stages of the process. PI(4)P has been called the 'fuel' that drives cholesterol transport, as its hydrolysis provides the energy that enables the establishment of active sterol concentration gradients across membrane-bound compartments with the aid of OSBP, which is a key regulator of cholesterol, oxysterol, and PI(4)P concentrations in membranes. In the plasma membrane, PI(4)P can support the functions of ion channels, and it contributes to the anchoring of proteins with polybasic domains, although it is not utilized for the synthesis of PI(4,5)P_2 in this membrane. On the other hand, PI(4)P derived from PI(4,5)P_2 in the membrane of primary cilia in the retina is important for vision. PI(4)P has an important influence on the progression of many diseases, especially virus replication, cancer, and various inflammatory diseases, and inhibitors of PI4-kinase are under study for their therapeutic potential.

While the biological properties of **phosphatidylinositol 5-phosphate** have taken longer to unravel, because of the difficulties of separation of this isomer, it is now apparent that it is involved in osmoregulation both in plants and animals. It also has signaling functions, and although it is the least abundant phosphatidylinositol monophosphate, it is involved in signaling at the nucleus and in the cytoplasm, modulating cellular responses to various stresses, hormones and growth factors. In the endosomes, it is a regulator of protein sorting.

Although **phosphatidylinositol 4,5-bisphosphate** (PI(4,5)P_2) is found primarily in the inner leaflet of the plasma membrane, where it may define membrane identity in eukaryotic cells, it is also present in endosomes, the endoplasmic reticulum and nucleus. It is an essential precursor of lipid second messengers such as diacylglycerols with vital signaling functions that operate through plasma membrane G-protein coupled receptors, receptor tyrosine kinases, and immune receptors. Because of its large head group and multivalent negative charge, PI(4,5)P_2 has been described as an "electrostatic beacon" that interacts in various ways with membrane proteins, other lipids and

cellular cations. In consequence and in spite of its relatively low concentration, it is a key regulator of innumerable events at the plasma membrane, including cell adhesion and motility, vesicle endocytosis and exocytosis, and the function of ion channels, especially those for potassium, calcium, and sodium. With ion channels, for example, it appears to be an obligatory factor, increasing their activity by activating key proteins, while its hydrolysis by phospholipase C reduces such activity.

$PI(4,5)P_2$ interacts with cationic residues of a large array of proteins in concert with cholesterol to form localized membrane domains that are distinct from the sphingolipid-enriched **rafts**. Indeed, it has a much higher concentration than other phosphoinositide species in cells, although most of this is in effect sequestered by binding proteins. Also, phosphatidylinositol 4,5-bisphosphate and its diacylglycerol metabolites are important for vesicle formation in membranes. For example, a major pathway in cells for the internalization of cell surface proteins such as transferrin is the clathrin-coated vesicle pathway. $PI(4,5)P_2$ is essential to this process in that it binds to the machinery involved in the membrane, increasing the number of clathrin-coated pits and permitting the internalization of proteins. It has a related function in **caveolae**, where it is concentrated at the rim.

Through its attachment to the apical plasma membrane, phosphatidylinositol 4,5-bisphosphate is intimately involved in the development of the actin cytoskeleton and thereby controls cell shape, motility, and many other processes. In particular, it binds with high specificity to effectors such as vinculin, a membrane-cytoskeletal protein that is involved in the linkage of integrin adhesion molecules to the actin cytoskeleton. Dysregulation of this function has been implicated in the migration and metastasis of tumor cells. In yeasts, it appears that the presence of stearic acid in position *sn*-1 is essential for this function. In the cell nucleus, this lipid is believed to be involved in maintaining chromatin, the complex combination of DNA, RNA, and protein that makes up chromosomes in a transcriptionally active conformation, as well as being a precursor for further signaling molecules. It has a role in gene transcription, and RNA processing, especially in the modulation of RNA polymerase activity, and in other nuclear processes.

Via its binding to specific proteins, the lipid is an essential component of the immune response of animal tissues to toxic bacterial **lipopolysaccharides**. It is also involved in the pathophysiology of the HIV virus via an interaction with the Tat protein secreted by infected cells.

$PI(4,5)P_2$ is the primary precursor of the endocannabinoid **2-arachidonoylglycerol** in neurons, and it is also an essential cofactor for phospholipase D and so affects the cellular production of **phosphatidic acid** with its specific signaling functions. By binding specifically to ceramide kinase, the enzyme responsible for the synthesis of **ceramide-1-phosphate**, it has an influence on sphingolipid metabolism. Like ceramide-1-phosphate, it binds to and activates the Ca^{2+}-dependent phospholipase A_2, which generates the arachidonate for eicosanoid production. One molecule of phosphatidylinositol 4,5-bisphosphate is bound to each subunit of the protein in the X-ray crystal structures of mammalian GIRK2 potassium channel, where it enables a conformational change that assists the transport function of the protein.

Perhaps, the best characterized of the phosphoinositide signaling functions results from the hydrolysis of phosphatidylinositol phosphates by phospholipase C isoforms, in this instance to produce ***sn*-1,2-diacylglycerols** and inositol 3,4,5-trisphosphate (see below), which act as second messengers. Only those polyunsaturated diacylglycerol species derived from $PI(4,5)P_2$ are able to bind and activate protein kinase C (α, ε, δ) isoforms both *in vitro* and *in vivo*. This lipid is doubly important as it binds strongly to these enzymes via a basic patch distal to a Ca^{2+} binding site, and this targets them selectively to the plasma membrane. Aberrant expression of phospholipase Cγ2 may be a factor in neurodegenerative diseases. Via the action of PI3 kinase, $PI(4,5)P_2$ is the precursor of $PI(3,4,5)P_3$ with its own distinctive signaling properties.

Phosphatidylinositol 3,4-bisphosphate can be produced by two routes and regulates a variety of cellular processes with relevance to health and disease that include B cell activation and autoantibody production, insulin sensitivity, neuronal dynamics, endocytosis, and cell migration. It is known to bind selectively to a number of proteins, and it acts as a secondary messenger by recruiting the protein kinases Akt (protein kinase B) and so may influence the cell cycle, cell survival, angiogenesis, and glucose metabolism. During endocytosis in the endolysosomal system, it is produced from $PI(4,5)P_2$ and controls the maturation of endocytic-coated pits. Its synthesis and turnover of are spatially segregated within the endocytic pathway. In epithelial cells, it is located on the apical membrane, i.e. facing the lumen, as opposed to the basolateral membranes, and it is believed to be is a determinant of the identity and function of the apical membrane.

Phosphatidylinositol 3,5-bisphosphate is present at low levels only in cells (0.04-0.1% of the total phosphatidylinositides) unless stimulated by growth factors, but it is important in membrane and protein trafficking, especially in the late endosomes in eukaryotes and in yeast vacuoles. For example, the conversion of PI(3)P to $PI(3,5)P_2$ promotes endosomal maturation and degradative sorting. It is involved in the mediation of signaling in response to stress and hormonal cues and in the control of ion transport in membranes, while genetic studies confirm that it is essential for healthy embryonic development, especially in the nervous system.

Phosphatidylinositol 3,4,5-trisphosphate ($PI(3,4,5)P_3$) is almost undetectable in quiescent cells, but its intracellular level rises very rapidly from synthesis at the plasma membrane in response to agonists such as extracellular growth factors and hormonal stimuli. By recruiting proteins with pleckstrin homology (PH) domains to the plasma membrane, it has been implicated in a variety of cellular functions that include growth, cell survival, proliferation, cytoskeletal rearrangement, intracellular vesicle trafficking, and cell metabolism. In particular, it is an important component of a signaling pathway in the cell nucleus. In epithelial cells, it is located on the basolateral membrane, i.e. facing adjacent cells, where it may be a determinant of the identity and function of this membrane. In contrast to phosphatidylinositol 3-phosphate, it opposes autophagy by binding to and activating the PH domain of Akt, so inducing cell proliferation. During feeding, various

physiological responses lead to the secretion of insulin, which stimulates the phosphorylation of phosphatidylinositol 4,5-bisphosphate to phosphatidylinositol 3,4,5-trisphosphate and triggers a signaling cascade that leads to the suppression of autophagy. When this pathway is impaired it has deleterious effects on insulin resistance associated with various metabolic diseases including obesity and diabetes. It has been implicated in tumor cell migration and metastasis. PI(3,4,5)P_3 is also present in the nucleus and nucleoli of cells where it is believed to have functions in RNA processing/splicing, cytokinesis, protein folding, and DNA repair. In complete contrast, like **phosphatidylserine**, it is reportedly transferred to the outer leaflet of the plasma membrane in aged or damaged cells as an 'eat-me' signal for phagocytes and apoptosis.

The human immune system utilizes neutrophils, which are highly mobile cells, to eliminate pathogens from infected tissue. The first step is to track and then pursue molecular signals, such as cytokines, emitted by pathogens. It has been established that two phospholipids operate in sequence to point the neutrophils in the correct direction. The first of these is phosphatidylinositol 3,4,5-trisphosphate, which binds to a specific protein DOCK2 and enables it to translocate to the plasma membrane. Then phosphatidic acid, generated by the action of phospholipase D on phosphatidylcholine, takes over and directs the DOCK2 to the leading edge of the plasma membrane. This causes the polymerization of actin within the cell and in effect reshapes the neutrophil and points it in the direction from which the pathogens signals are coming. On the other hand, *Mycobacterium tuberculosis* is able to subvert phosphoinositide signaling to arrest phagosome maturation by dephosphorylation of phosphatidylinositol 3-phosphate.

21.3.6.3: WATER-SOLUBLE INOSITOL PHOSPHATES

As mentioned briefly above, hydrolysis of phosphatidylinositol phosphates by calcium-dependent phospholipase C (or 'phosphoinositidase C') leads to the generation of ***sn*-1,2-diacylglycerols**, which act as second messengers in animal cells and are of enormous metabolic importance. There are many different enzymes of this type, but the activity of the phosphoinositide-specific phospholipase C constitutes an essential step in the inositide signaling pathways. The enzyme exists in six families consisting of at least 13 isoenzymes, all of which have conserved regions such as the plekstrin homology (PH) binding domain. Each one has a distinctive role and can have a characteristic cell distribution that is linked to a specific function. The activity of these enzymes is stimulated by signaling molecules such as G-protein coupled receptors, receptor tyrosine kinases, Ras-like GTPases, and calcium ions, thus linking the hydrolysis of phosphatidylinositol phosphates to a wide range of other cellular signals. As phospholipase C is a soluble protein located mainly in the cytosol, translocation to the plasma membrane is a crucial step in signal transduction. Regulation of these isoenzymes and the form PLCγ1 in particular is vital for health as they are associated with the activation or inhibition of important pathophysiological processes, especially in relation to cancer.

Some phosphatidic acid is synthesized from the diacylglycerols produced within the plasma membrane through the activity of diacylglycerol kinases, and this is transported back to the endoplasmic reticulum and ultimately can be re-utilized for phosphatidylinositol biosynthesis.

The other products of the phospholipase C reaction that are of special relevance because of their many essential functions are **water-soluble inositol phosphates**. Up to 60 different compounds of this type are possible, and at least 37 of these have been found in nature at the last count, all of which are extremely important biologically. However, polyphosphoinositides with a phosphate in position 3 are not substrates for phospholipase C.

Figure 21.3.39 shows the generation of inositol phosphates by phospholipase C.

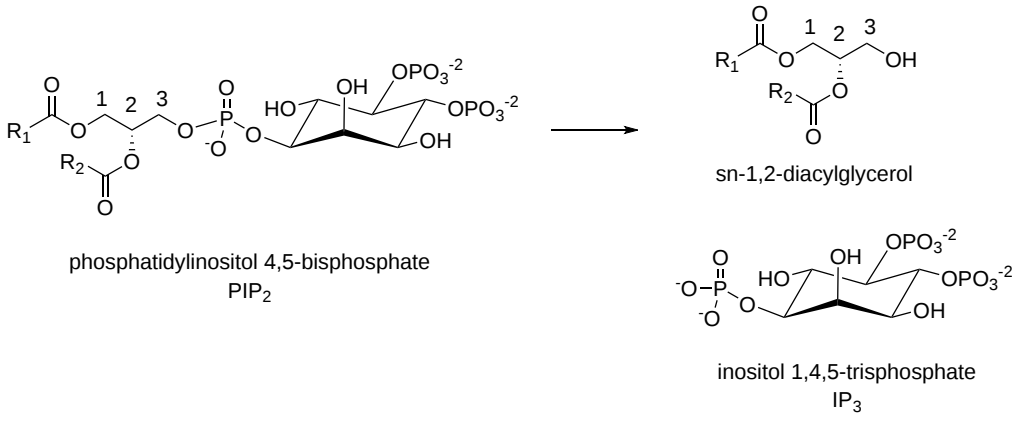


Figure 21.3.39: Generation of inositol phosphates by phospholipase C

For example, under the action of various physiological stimuli in animals, including sphingosine-1-phosphate, and acting via various G-protein-coupled receptors, phosphatidylinositol 4,5-bisphosphate in the plasma membrane is hydrolyzed to release inositol 1,4,5-trisphosphate, an important cellular messenger that diffuses into the cytosol and stimulates calcium release from an ATP-loaded store in the endoplasmic reticulum via ligand-gated calcium channels (the diacylglycerols remain in the membrane to recruit and activate members of the protein kinase C family). The increase in calcium concentration, together with the altered phosphorylation status, activates or de-

activates many different protein targets, enabling cells to respond in an appropriate manner to the extracellular stimulus. To enable rapid replenishment of the phosphatidylinositol 4,5-bisphosphate used in this way, a cycle of reactions - the phosphatidylinositol cycle - must occur (see below). On the other hand, a recent publication suggests that phosphatidylinositol 4-phosphate in the plasma membrane may be a more important source of diacylglycerols following stimulation of G protein–coupled receptors.

All of the various inositol phosphates appear to be involved in the control of cellular events in very specific ways, but especially in the organization of key signaling pathways, the rearrangement of the actin cytoskeleton, or intracellular vesicle trafficking. They have been implicated in gene transcription, RNA editing, nuclear export, and protein phosphorylation. As these remarkable compounds can be rapidly synthesized and degraded in discrete membrane domains or even sub-nuclear structures, they are considered to be ideal regulators of dynamic cellular mechanisms. From structural studies of inositol polyphosphate-binding proteins, it is believed that the inositides may act in part at least by modifying protein function by acting as structural cofactors, ensuring that proteins adopt their optimum conformations. In addition, phosphoinositides and the inositol polyphosphates are key components of the nucleus of the cell, where they have many essential functions, including DNA repair, transcription regulation, and RNA dynamics. It is believed that they may be activity switches for the nuclear complexes responsible for such processes, with the phosphorylation state of the inositol ring being of primary importance. As different isomers appear to have specific functions at each level of gene expression, extracellular events must coordinate the production of these compounds in a highly synchronous manner.

In organisms from plants to mammals, an extra tier of regulatory mechanisms is produced by kinases that generate energetic diphosphate (pyrophosphate)-containing molecules from inositol phosphates. Conversely, these can be dephosphorylated by polyphosphate phosphohydrolase enzymes to regenerate the original inositol phosphates. These inositol pyrophosphates and the enzymes involved in their metabolism are also involved in the regulation of cellular processes by modulating the activity of proteins by a variety of mechanisms.

It should be noted that the phospholipase C isoenzymes regulate the concentration of phosphatidylinositol 4,5-bisphosphate and related lipids and thence their activities in addition to the generation of new biologically active metabolites.

21.3.6.4: PHOSPHATIDYLINOSITIDES IN PLANTS

In plants as in animals, phosphatidylinositol and polyphosphoinositides have essential biological functions, exerting their regulatory effects by acting as ligands that bind to protein targets via specific lipid-binding domains and so alter the location of proteins and their enzymatic activities. However, it appears that polyphosphoinositide metabolism developed in different ways after the divergence of the animal and plant kingdoms so the details of the processes in each are very different, not least because the subcellular locations of phosphoinositides differ appreciably between plants and animals. Phosphatidylinositol *per se* is of course the precursor of the phosphorylated forms and determines their fatty acid compositions. It also has a role in inhibiting programmed cell death by acting as the biosynthetic precursor of the sphingolipid **ceramide phosphoinositol** and so reducing the levels of ceramide.

As in animals, the various phosphoinositides (five in total) are produced and inter-converted rapidly by a series of kinases and phosphatases (in many isoforms) in different cellular membranes in response to environmental or developmental cues. For example, phosphatidylinositol is generated mainly in the endoplasmic reticulum, while PI 4-kinases and their product are located in the trans-Golgi network and nucleus, and PI4P 5-kinases and product are present in the plasma membrane. During the biosynthesis of polyphosphoinositides, the first phosphorylation occurs at the hydroxyl group at positions 3 or 4 of the inositol ring, catalyzed by the appropriate kinases, while the second phosphorylation then takes place at position 5; PI 5-phosphate is produced by the action of a phosphatase on PI 3,5-bisphosphate. Most other metabolites are produced via phosphatidylinositol 3-phosphate, and reports that some phosphatidylinositol 3,4,5-trisphosphate may be produced from phosphatidylinositol 4,5-bisphosphate require confirmation. In contrast to mammalian phosphatidylinositol 3-kinases, which accept both phosphatidylinositol and its monophosphates as substrates, the plant enzyme acts only on the former.

Figure 21.3.40 shows polyphosphoinositide metabolism in plants

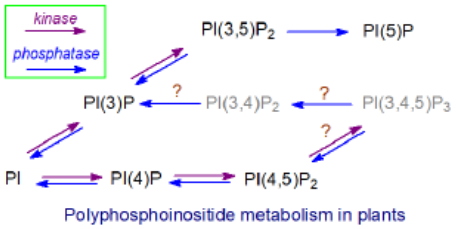


Figure 21.3.40: Polyphosphoinositide metabolism in plants

The reverse reaction in plants is accomplished by phosphoinositide phosphatases, which can be grouped into three main families, the phosphatase/tensin (PTEN) family, 5-phosphatases (5-PTases) and phosphatases containing Suppressor of Actin (SAC) domains, each with differing subcellular locations, substrate specificities and regulatory mechanisms.

Although what might be considered normal levels of phosphatidylinositol 4-phosphate are present, the concentrations of phosphatidylinositol 4,5-bisphosphate and other phosphoinositides are extremely low in plants (10 to 20-fold lower than in mammalian cells), although they still have vital functions. There are differences between cell types, but in Arabidopsis epidermal root cells, $PI(4,5)P_2$ is

present at the highest concentration in the plasma membrane (apex region) and nucleus, while PI4P slowly distributes between the plasma membrane and Golgi, with the highest concentration in the former. Multivesicular bodies/late endosomes accumulate both PI3P and PI(3,5)P_2, and the tonoplast and autophagosomes contain PI3P. How the various metabolites are transported between membranes has yet to be determined, but non-vesicular transport is believed to occur at membrane contact sites and vesicular transport probably occurs also.

Highly polarized distributions of phosphoinositides are found within membranes, generally oriented toward the cytosolic leaflet, and they are believed to be organized in nanoclusters together with other lipids and proteins. For example, phosphatidylinositol-4-phosphate is an important constituent of the plasma membrane in plant cells, where it controls the electrostatic state and is involved in cell division. It may control the location and function of many membrane proteins, including those required for development, reproduction, immunity, nutrition, and signaling. PI(4)P is the only phosphoinositide present at the cell plate, i.e. the membrane separating two daughter cells during cell division. In addition, PI(4)P may interact with salicylic acid in the plant immune response, and it is produced during salt stress. However, specific functions are now being discovered for each of the plant phosphoinositides, which are produced rapidly in response to osmotic and heat stress, and it has become evident that a continuous turnover is essential for cell growth and development. For example, they have marked effects on the growth of many cell types and on guard cell function. In the nucleus, proteins have been identified that bind to phosphoinositides via the acyl chains, leaving the head group exposed for enzymatic modifications and signal transduction.

Phosphoinositides are of special importance in microdomains at the tip of growing tissues such as the shoot apical meristem, pollen tubes, and root hairs where phosphatidylinositol 4,5-bisphosphate functions in stem cell maintenance and organogenesis. In the plasma membrane, it is enriched in the detergent-resistant component commonly equated with '**rafts**'. Although its concentration is low, PI(4,5)P_2 has been shown to have signaling functions by binding to a number of different target proteins, which have characteristic binding domains. For example, together with phosphatidic acid, PI(4,5)P_2 regulates the activity of a number of actin-binding proteins, which in turn control the activity of the actin cytoskeleton. This has a key role in plant growth, the movement of subcellular organelles, cell division and differentiation, and plant defense. In addition, this lipid exerts control over ion channels, ATPases, and phospholipase C-mediated lipid degradation and the production of further second messengers. It is an important factor in both clathrin-mediated endocytosis and exocytosis. The specificity of the interactions may be dependent on the fatty acid composition of the lipid and on the activity of phosphatidylinositol 4-phosphate 5-kinase.

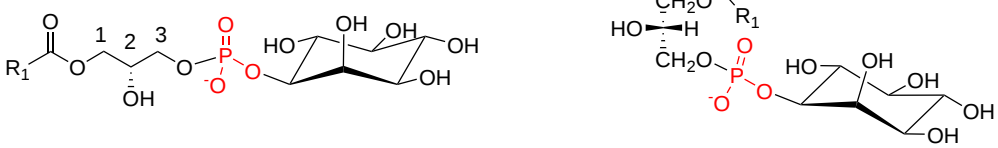
As in animals, phosphoinositides have a role in endosomal sorting but through the central vacuole, which is a plant-specific organelle with both lytic and storage functions. Phosphatidylinositol 3,5-bisphosphate is the least abundant of the phosphoinositides, but it is a crucial lipid for membrane trafficking systems. The PI to PI(3)P to PI(3,5)P_2 cascade, the second step requiring a kinase designated FAB1, is required for endosomal sorting events leading to membrane protein degradation or retrieval, vacuolar morphogenesis and autophagy. PI(3,5)P_2 is involved in stomatal closure and the growth of root hairs, and it is also induced in salt stress.

A number of different enzymes of the phospholipase C type that are specific for polyphosphoinositides have been isolated from higher plants; they are activated by Ca^{2+} and unlike their mammalian counterparts, they are not regulated by G proteins. It is not certain whether phosphatidylinositol is itself a substrate for these enzymes *in vivo*. Less is known of the metabolism of the water-soluble inositol phosphates produced in comparison to animals, and plants appear to lack a receptor for inositol 1,4,5-trisphosphate (IP_3), although it is the most abundant metabolite of this type and is reported to induce the release of calcium ions to trigger stomatal closure. However, there is increasing evidence for lipid signaling mediated by phospholipase C in abiotic stress tolerance and development in plants. There is a general if contested belief that inositol hexakisphosphate (phytic acid or IP_6), produced at least in part by sequential phosphorylation of inositol 1,4,5-trisphosphate, is a more important cellular messenger in plants and mobilizes an endomembrane store of calcium ions. Inositol-1,2,4,5,6-pentakisphosphate (IP_5) is a structural co-factor of the **jasmonic acid** receptor coronatine insensitive 1, linking phosphoinositide signaling with phytohormone-controlled pathways.

In plants in contrast to animals, **diacylglycerols**, the other product of phospholipase C hydrolysis of phosphoinositides, are rapidly converted to **phosphatidic acid** by diacylglycerol kinases and have not been considered important in signal transduction. Plants lack protein kinase C but they do have proteins with related properties that appear to be influenced by diacylglycerols. Via the action of phospholipase D, inositol phospholipids are a source of phosphatidic acid with its well-characterized signaling functions in plants, especially in defence.

21.3.6.5: LYSO-PHOSPHOINOSITIDES

Figure 21.3.41 shows the structure of lysophosphatidylinositol



lysophosphatidylinositol

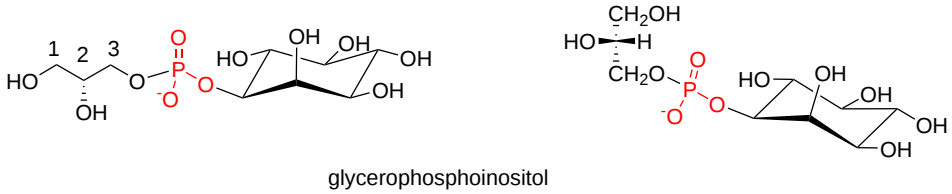
Figure 21.3.41: Lysophosphatidylinositol

Lysophosphatidylinositols: Lysophosphatidylinositols (LPI), i.e. with a single fatty acid only linked to the glycerol moiety, are formed as intermediates in the remodeling of the fatty acid compositions of the lipids by the action of phospholipase A_1 or phospholipase A_2 (e.g. $cPLA_2\alpha$), and when arachidonic acid is released for eicosanoid biosynthesis (see above). In ovarian cancer, LPI is elevated appreciably to around 15µM in ascites, and it is also present at high levels in obese subjects.

It has become apparent relatively recently that like other lysophospholipids, phosphatidylinositol, and polyphospho-analogues may have messenger functions. For example, it has long been known to stimulate the release of insulin from pancreatic cells, suggesting a role in glucose homeostasis. *sn*-2-Arachidonoyl-lysophosphatidylinositol, in particular, is an endogenous ligand for a G protein-coupled receptor GPR55, and thereby can induce rapid phosphorylation of certain enzymes, including a protein kinase, which promote cancer cell proliferation, migration, and metastasis. Indeed, lysophosphatidylinositol is a biomarker for poor prognosis in cancer patients, and its concentration is elevated significantly in highly proliferative cancer cells *in vitro*. GPR55 is expressed in many regions of the brain, the intestines, endocrine pancreas and islets (where it may stimulate insulin release). It has been implicated in macrophage activation and inflammation. In addition to its role in cancer, lysophosphatidylinositol has been implicated in a number of metabolic diseases. It is reported to be a precursor of the endocannabinoid **2-arachidonoylglycerol** by the action of human glycerophosphodiesterase 3 as a lysophospholipase C. This enzyme suppresses the receptor for lysophosphatidylinositol, and so acts as a switch between GPR55 and endocannabinoid (CB2) signaling.

Glycerophosphoinositol: Sequential removal of both fatty acids from phosphatidylinositol by a specific phospholipase A_2 ($PLA_2IV\alpha$) with both phospholipase A_2 and lysophospholipase activities releases water-soluble glycerophosphoinositol. While this can be hydrolyzed by a glycerophosphodiester phosphodiesterase to inositol 1-phosphate, glycerophosphoinositol *per se* has distinctive biological activities and functions, as do related compounds derived from the phosphatidylinositol phosphates. In particular, glycerophosphoinositol has anti-inflammatory activity in that it inhibits the inflammatory and thrombotic responses induced by bacterial lipopolysaccharides (endotoxins).

Figure 21.3.42 shows the structure of glycerophosphoinositol



glycerophosphoinositol

Figure 21.3.42 Glycerolphosphoinositol

21.3.6.6: THE PHOSPHATIDYLINOSITOL CYCLE

Phosphatidylinositol can be considered to be at the center of a cycle of reactions and intermediates that are involved in innumerable aspects of cellular signaling in animals (a similar cycle could be described for plants). These are discussed individually at length above, but it is useful to point out how each component forms part of a larger pattern. In brief as illustrated, the various synthetic and hydrolytic reactions involved in phosphoinositide metabolism can be considered to constitute a **phosphatidylinositol cycle** with enzymes located both in the endoplasmic reticulum and plasma membrane, so lipids have to be transferred across the cytosol in both directions between the two to complete the cycle, probably via adjacent membrane structures and facilitated by proteins of the phosphatidylinositol transfer protein membrane-associated family (PITPNM or nir2), which may channel phosphoinositide production to specific biological outcomes. Phospholipase C and phosphatidylinositol-4-phosphate 5-kinase (PI4P 5K) are located in the plasma membrane, while the cytidine diphosphate-diacylglycerol synthase (CDS2) and phosphatidylinositol synthase are in the endoplasmic reticulum. The epsilon isoform of diacylglycerol kinase (DGKε) is located at contact sites between the endoplasmic reticulum and plasma membrane, but there are nine further isoforms with differing cellular and subcellular locations that may be involved in the cycle. Each turn of the cycle uses a great deal of energy and consumes three moles of ATP, together with cytidine triphosphate and inositol. If it is assumed that the pyrophosphate is hydrolyzed by endogenous pyrophosphatases to inorganic phosphate, the cycle can proceed in one direction only.

Figure 21.3.43 shows the phosphatidylinositol cycle.

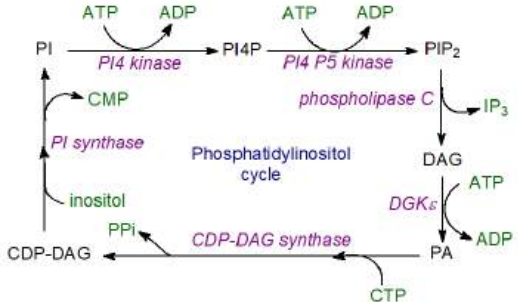


Figure 21.3.43 Phosphatdylionositol cycle

Factors such as membrane curvature must be taken into account, and the diagram is of necessity a considerable over-simplification. In addition to participating in this cycle, many of the lipid intermediates can be precursors for other lipids, and for example, diacylglycerols are potential precursors for triacylglycerols, while phosphatidic acid is a precursor for phosphatidylcholine and phosphatidylethanolamine. Each lipid intermediate is subject to remodeling of the acyl chains via the Lands cycle, and polyunsaturated fatty acids released can be utilized for eicosanoid production. A further by-product of the cycle is inositol triphosphate, which contributes to the regulation of intracellular calcium levels.

It has been suggested that the unique molecular species composition of phosphoinositides (18:0-20:4) could influence their selective recycling back into phosphatidylinositol as many of the enzymes involved have a preference for this substrate. A further proposal is that the phosphatidylinositol cycle could act to enrich this species through multiple passages around the cycle.

21.4: BIOSYNTHESIS OF MEMBRANE SPHINGOLIPIDS

by William (Bill) W. Christie and Henry Jakubowski.

This section is an abbreviated and modified version of material from the Lipid Web, an introduction to the chemistry and biochemistry of individual lipid classes, written by William Christie.

21.4.1: SPHINGOLIPIDS

21.4.1.1: INTRODUCTION:

The sphingolipids comprise a wide range of complex lipids in which the defining component is a long-chain or sphingoid base, which in living tissues is usually linked to a fatty acid via an amide bond. J.L.W. Thudichum, a German chemist working in London, first coined the root term “sphingo-” in 1884 following his discovery of the first glycosphingolipids, because the enigmatic nature of the molecules reminded him of the riddle of the sphinx. Regretfully, the importance of his work was not recognized until 25 years after his death, and it was 1947 before the term “sphingolipide” was introduced by Herbert Carter and colleagues. While they are much less enigmatic than they once were, sphingolipids are extremely versatile molecules that continue to fascinate as new knowledge is gained of their functions in healthy (and diseased) animal and plant tissues. They are found in only a few bacterial genera, but they are present in *Sphingomonas*, *Sphingobacterium* and a few other species, and many pathogenic species utilize host sphingolipids to promote infections. Novel sphingolipid structures continue to be reported, and as an example at the last count, 188 of the complex sphingolipids classified as **gangliosides**, with variations in the complex carbohydrate component alone, had been characterized in vertebrates.

Long-chain or sphingoid bases, of which sphingosine is typical, are the basic elements and are the simplest possible functional sphingolipids. They vary in chain length and in the presence of various functional groups including double bonds of both the *cis*- and *trans*-configuration at different locations in the aliphatic chain. **Ceramides**, which contain sphingoid bases linked to fatty acids by amide bonds, vary appreciably in the compositions of both aliphatic components, depending on their biological origins. The structure of sphingosine and ceramide, the sphingolipid building blocks, are shown in Figure 21.4.1 .

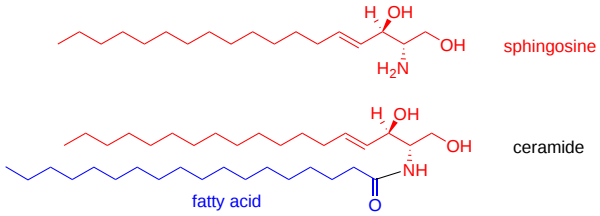


Figure 21.4.1: Sphingolipid building blocks

Long-chain bases and ceramides have important biological properties in their own right, for example in relation to intra- and inter-cellular molecular signaling, especially in animal cells, while another relatively simple sphingolipid, **sphingosine-1-phosphate**, is now recognized as a key factor in countless aspects of animal metabolism. The concentrations of these bioactive lipids respond rapidly to the action of specific stimuli and then regulate downstream effectors and targets.

Ceramides are the precursors of a multitude of sphingo-phospho- and sphingo-glycolipids with an immense range of functions in tissues. The properties and functions of these complex sphingolipids are quite distinct from those of the comparable glycerophospho- and glyceroglycolipids. For example in animals, **sphingomyelin** has structural similarities to phosphatidylcholine, but has very different physical and biological properties, while the complex **oligoglycosylceramides** and **gangliosides**(glycosphingolipids, of which **glucosylceramide** is the precursor, have no true parallels among the glyceroglycolipids. Figure 21.4.2 shows the structure of the complex sphingolipids sphingomyelin and glucosylceramide.

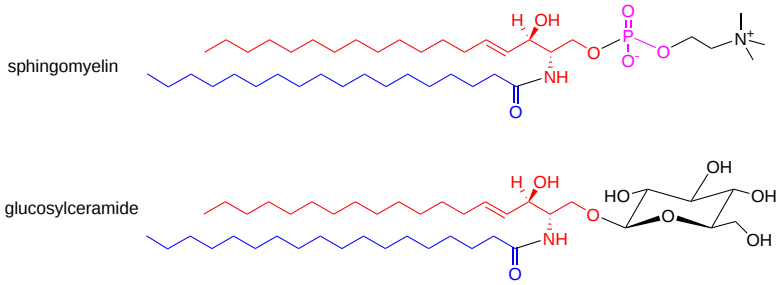


Figure 21.4.2: Complex Sphingolipids

Complex sphingolipids are synthesized in the endoplasmic reticulum and Golgi, but are located mainly in the plasma membrane of most mammalian cells where they have a structural function and also serve as adhesion sites for proteins from the extracellular tissue. The glycosphingolipids are especially important for myelin formation in the brain. However, sphingolipids have intracellular functions in all

cellular compartments, including the nucleus. The first five carbon atoms of the sphingoid base in sphingolipids have a highly specific stereochemistry and constitute a key feature that has been termed the ‘sphingoid motif’, which in comparison to other lipid species facilitates a relatively large number of noncovalent interactions with other membrane lipids, via hydrogen-bonding, ion-ion interactions and induced dipole-induced dipole interactions. A distinctive property of sphingolipids in membranes is that they spontaneously form transient nanodomains termed '**rafts**', usually in conjunction with cholesterol, where such proteins as enzymes and receptors congregate to carry out their signaling and other functions. Thus, in addition to their direct effects on metabolism, sphingolipids affect innumerable aspects of biochemistry indirectly via their physical properties.

While it may be obvious that a well-balanced sphingolipid metabolism is important for health in animals, increasing evidence has been acquired to demonstrate that impaired sphingolipid metabolism and function are involved in the pathophysiology of many of the more common human diseases. These include diabetes, various cancers, microbial infections, Alzheimer's disease and other neurological syndromes, and diseases of the cardiovascular and respiratory systems. In humans, a number of important genetic defects in sphingolipid metabolism or sphingolipidoses have been detected, especially storage diseases associated with the lysosomal compartment where sphingolipids are catabolized. Sphingolipids and their metabolism are therefore likely to prove of ever increasing interest to scientists.

There are appreciable differences in sphingolipid compositions and metabolism between animal and plant cells, both with respect to the aliphatic components and especially the polar head groups, although there are also some important similarities. While sphingomyelin is the most abundant sphingolipid in animals, it does not occur in plants and fungi. Although less is known of the role they play in plants, it has become apparent that complex sphingolipids are much more abundant in plant membranes than was once believed, and it is now recognized that they are key components of the plasma membrane and endomembrane system.

21.4.1.2: SOME GENERAL COMMENTS ON SPHINGOLIPID METABOLISM

The biosynthesis and catabolism of sphingolipids involves a large number of intermediate metabolites, all of which have distinctive biological activities of their own. In animals, the relationships between these metabolites have been rationalized in terms of a ‘sphingomyelin, sphingolipid or ceramide cycle’, as shown in Figure 21.4.3.

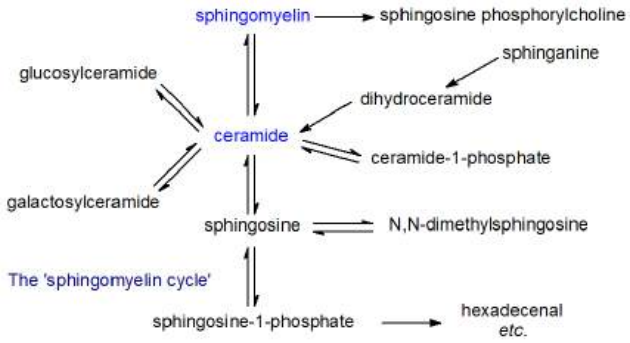


Figure 21.4.3: Sphingomyelin, sphingolipid or ceramide cycle

Many different enzymes (and their isoforms) are involved, and their activities depend on a number of factors, including intracellular locations and mechanisms of activation. Each of the various compounds in these pathways has characteristic metabolic properties. Thus, free sphingosine and other **long-chain bases**, which are the primary precursors of ceramides and thence of all the complex sphingolipids, function as mediators of many cellular events, for example by inhibiting the important enzyme protein kinase C. **Ceramides** are involved in cellular signaling, and especially in the regulation of apoptosis, and cell differentiation, transformation and proliferation, and most stress conditions. In contrast, **sphingosine-1-phosphate** and **ceramide-1-phosphate** promote cellular division (mitosis) as opposed to apoptosis, so that the balance between these lipids and ceramide and/or sphingosine levels in cells is critical and necessitates exquisite control in each cellular compartment.

Similarly, the ‘structural’ sphingolipids, such as **sphingomyelin**, **monoglycosylceramides**, **oligoglycosylceramides** and **gangliosides**, all have unique and characteristic biological functions, some of which are due to their physical properties and location within **rafts,** nanodomains of membranes. Most of the reactions in the sphingomyelin cycle are reversible and the relevant enzymes are located in the endoplasmic reticulum, Golgi, plasma membrane, and mitochondria, but the more complex sphingolipids are catabolized in the lysosomal compartment. Sphingolipids are especially important in providing the permeability barrier in the **skin**, where they are characterized by the presence of ultra-long fatty acyl components as well as fatty acyl groups linked to a hydroxyl group at the terminal end of the *N*-linked fatty acids (thereby generating a three-chain rather than a two-chain molecule).

Metabolic pathways that are comparable to those of the sphingomyelin cycle are believed to occur in plants, as shown in Figure 21.4.4, although they have not been studied as extensively as those in animals.

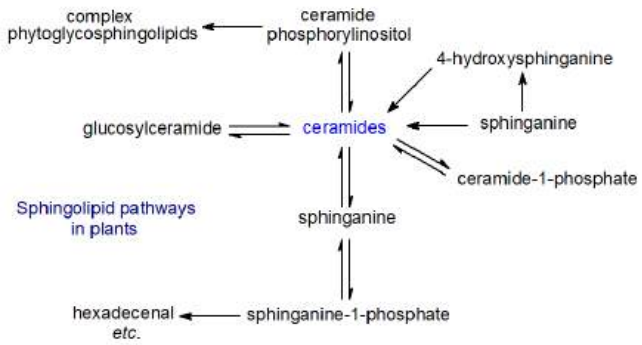


Figure 21.4.4: Sphingomyelin, sphingolipid or ceramide cycle in plants

However, sphingolipid metabolites such as sphingosine-1-phosphate (or analogues) have been linked to programmed cell death, signal transduction, membrane stability, host-pathogen interactions and stress responses, for example. Plants also have a unique range of complex sphingolipids in their membranes, such as **ceramide phosphorylinositol and the phytoglycosphingolipids**, and these are now known to constitute a higher proportion of the total lipids than had hitherto been supposed, although their functions have hardly been explored. While sphingolipids are produced by relatively few bacterial species, sulfono-analogues of long-chain bases and ceramides (capnoids) are produced by some specie.

21.4.1.3: FATTY ACID COMPONENTS OF SPHINGOLIPIDS

The **fatty acids** of sphingolipids are very different from those of glycerolipids, consisting of very-long-chain (up to C_{26}) odd- and even-numbered saturated or monoenoic and related **2(*R*)-hydroxy** components, while even longer fatty acids (C_{28} to C_{36}) occur in spermatozoa and the epidermis. The dienoic acid 15,18-tetracosadienoate (24:2(*n*-6)), derived from elongation of linoleic acid, is found in the ceramides and other sphingolipids of a number of different tissues, but at relatively low levels. Polyunsaturated fatty acids are only rarely present, although sphingomyelins of testes and spermatozoa are exceptions in that they contain such fatty acids, which are even longer in chain-length (up to 34 carbon atoms) and include 28:4(*n*-6) and 30:5(*n*-6). Skin ceramides also contain unusual very-long-chain fatty acids, while yeast sphingolipids are distinctive in containing mainly C_{26} fatty acids. In plants and yeasts, a similar range of chain-lengths occur as in animals, but 2-hydroxy acids predominate sometimes accompanied by small amounts of 2,3-dihydroxy acids; saturated fatty acids are most abundant, but monoenes are present in higher proportions in the Brassica family (including Arabidopsis) and a few other species. Some fungal species contain monoenoic fatty acids with a *trans*-3 double bond and/or a hydroxyl group. Figure 21.4.5 shows typical sphingolipid fatty acids.

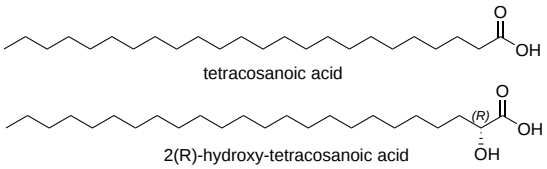


Figure 21.4.5: Typical sphingolipid fatty acids

Very-long-chain saturated and monoenoic fatty acids for sphingolipid biosynthesis are produced from medium-chain precursors by **elongases (ELOVL)** in the endoplasmic reticulum of cells in mammals, and there is increasing evidence that specific isoforms are involved in the biosynthesis of certain ceramides. For example, ELOVL1 has been linked to the production of ceramides with C_{24} fatty acids (saturated and unsaturated), while ELOVL4 is responsible for the ultra-long-chain fatty acids in skin. Yeasts possess three elongation enzymes: Elo1 (for medium to long-chain fatty acids), Elo2 (up to C_{22}) and Elo3 (up to C_{26}).

The hydroxyl group is believed to add to the hydrogen-bonding capacity of the sphingolipids, and it helps to stabilize membrane structures and strengthen the interactions with membrane proteins. Hydroxylation is effected by a fatty acid 2-hydroxylase in mammals, i.e. an NAD(P)H-dependent monooxygenase, which is an integral membrane protein of the endoplasmic reticulum. It converts unesterified long-chain fatty acids to 2-hydroxy acids *in vitro* and probably also *in vivo*. For example, experimental evidence has been obtained that is consistent with 2-hydroxylation occurring at the fatty acid level prior to incorporation into ceramides in the brain of mice where the enzyme is expressed at high levels. A second enzyme of this kind is known to exist but has yet to be characterized, and it is possible that a proportion of the odd-chain fatty acids in brain are synthesized by Peroxisomal α-oxidation of the 2-hydroxy acids. Similarly, in skin, 2-hydroxy and non-hydroxy fatty acids as their CoA esters are used with equal facility for ceramide biosynthesis by ceramide synthases. As mutations in the fatty acid 2-hydroxylase in humans and mice give rise to demyelination disorders, such as leukodystrophy, it is evident that sphingolipids containing 2-hydroxy acids have unique functions in membranes that cannot be substituted by non-hydroxy analogues.

In plants, it appears that 2-hydroxyl groups are inserted into fatty acyl chains while they are linked to ceramide, as ceramide synthase does not accept hydroxy fatty acids *in vitro* at least. Two fatty acid 2-hydroxylases (di-iron-oxo enzymes) have been found in Arabidopsis, with

one specific for very-long-chain fatty acids and one for palmitic acid. In fungi, a hydroxyl group is inserted at C2 of the fatty acid in a dihydroceramide intermediate.

Although the fatty acids are only occasionally considered in terms of the biological functions of sphingolipids, their influence is considerable, especially but not only in relation to their physical properties and function in membranes. For example, very-long-chain fatty acids may play a role in stabilizing highly curved membrane domains as is required during cell division. The hydrophobic nature of the fatty acyl groups (together with the long-chain bases) enables the hydrogen bonding that is essential for the formation of **raft nanodomains** in membranes. As a general rule, lipid bilayers containing sphingolipids with 2-hydroxy-fatty acyl or 4-hydroxy-sphingoid base moieties, tend to generate condensed and more stable gel phases with higher melting temperatures than their non-hydroxylated equivalents, because they have a more extended and strengthened intermolecular hydrogen bonding network. Changes in fatty acid composition are seen in some disease states, and for example increased concentrations of fatty acids $>C_{24}$ are a feature of adrenoleukodystrophy, an X-linked genetic disorder.

Removal of very-long-chain fatty acids from sphingolipids in mutants of the model plant Arabidopsis inhibits completely the development of seedlings. As example of a more specific interaction, it has been demonstrated that synthetic glycerolipids must contain very-long-chain fatty acids (C_{26}) to allow growth in yeast mutants lacking sphingolipids, probably by stabilizing the proton-pumping enzyme H^+-ATPase. Similarly, ceramides containing different fatty acids can be used in highly specific ways. Thus in fungi, C_{16} or C_{18} hydroxy acids are used exclusively for synthesis of glucosylceramide, while those containing very-long-chain C_{24} and C_{26} hydroxy acids are used only for synthesis of **glycosyl inositol phosphorylceramide** anchors for proteins. In plants, sphingolipids containing 2-hydroxy acids are protective against oxidative and other biotic stresses.

21.4.1.4: LINKS BETWEEN GLYCEROLIPID AND SPHINGOLIPID METABOLISM

Sphingolipid metabolism and glycerolipid metabolism have been widely treated as separate sciences until relatively recently, partly for historical reasons and partly because the analysis of the two lipid groups required different approaches and skills. However, there are many areas where the two overlap, not least because phosphatidylcholine is the biosynthetic precursor of **sphingomyelin** in animal cells, while in plants and fungi, phosphatidylinositol is the biosynthetic precursor of **ceramide phosphorylinositol**. In contrast, ethanolamine phosphate derived from the catabolism of sphingolipids via **sphingosine 1-phosphate** is recycled for the biosynthesis of **phosphatidylethanolamine**, and this is essential for survival in the protozoan parasite *Trypanosoma brucei*. In studies in vitro, sphingosine 1-phosphate has been shown to be an activator of the phospholipase C involved in the hydrolysis of the lipid mediator **phosphatidylinositol 4,5-bisphosphate** with formation of diacylglycerols and inositol triphosphate. The location and functions of glycerophospholipids in membranes is influenced both positively and negatively by sphingolipid-rich domains or **rafts** in membranes.

In addition, there are several examples of phosphoinositides and other complex lipids binding to enzymes of sphingolipid metabolism, either as part of a regulatory function that controls their activity or to facilitate their location to various membranes. Thus, sphingosine kinase 2, one of the enzymes responsible for the biosynthesis of sphingosine 1-phosphate, binds to phosphatidylinositol monophosphates, while the ceramide kinase responsible for the biosynthesis of **ceramide 1-phosphate** requires phosphatidylinositol 4,5-bisphosphate to function. Similarly, the CERT protein involved in **ceramide** transport has a binding site for phosphatidylinositol 4-phosphate. Sphingomyelin production at the trans-Golgi network triggers a signaling pathway leading to dephosphorylation of phosphatidylinositol 4-phosphate, interrupting transport of cholesterol and sphingomyelin. Again, the interactions are not solely in one direction as ceramide 1-phosphate (with phosphatidylinositol 4,5-bisphosphate) binds to the specific **phospholipase A_2 ($cPLA_2\alpha$)** responsible for the hydrolysis of phosphatidylinositol and thence the release arachidonic acid for eicosanoid production. Other than the phosphoinositides, phosphatidylserine activates the neutral sphingomyelinase in brain.

21.4.2: LONG-CHAIN (SPHINGOID) BASES

Long-chain/sphingoid bases are the characteristic and defining structural unit of the sphingolipids, which are important structural and signaling lipids of animals and plants and of a few bacterial species. These are long-chain aliphatic amines, containing two or three hydroxyl groups, and often a distinctive *trans*-double bond in position 4. To be more precise, they are 2-amino-1,3-dihydroxy-alkanes or alkenes with (2*S*,3*R*)-*erythro* stereochemistry, often with various further structural modifications in the alkyl chain. They are important for the physical and biological properties of all of the more complex sphingolipids, but free sphingoid bases are also bioactive and interact with specific receptors and target molecules. As discussed below, the mechanisms for biosynthesis of sphingoid bases and of the *N*-acylated form (ceramides) are intimately linked.

21.4.2.1: STRUCTURES AND OCCURRENCE

In animal tissues, the most common or abundant of the sphingoid bases is **sphingosine** ((2*S*,3*R*,4*E*)-2-amino-4-octadecene-1,3-diol) or sphing-4*E*-enine, i.e., with a C_{18} aliphatic chain, hydroxyl groups in positions 1 and 3 and an amine group in position 2; the double bond in position 4 has the *trans* (or *E*) configuration. This was first characterized in 1947 by Professor Herbert Carter, who was also the first to propose the term “sphingolipides” for those lipids containing sphingosine. It is usually accompanied by the saturated analogue dihydrosphingosine (or sphinganine). Sphingoid bases are illustrated in Figure 21.4.6.

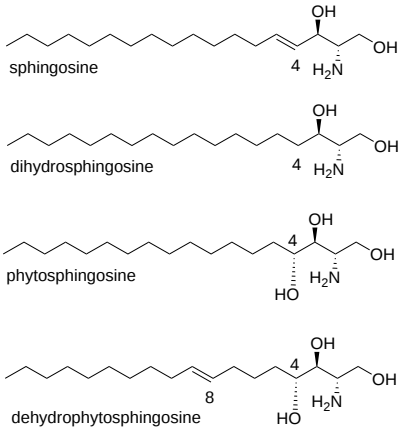


Figure 21.4.6: Sphingoid Bases

For shorthand purposes, a nomenclature similar to that for fatty acids can be used; the chain length and number of double bonds are denoted in the same manner with the prefix 'd' or 't' to designate di- and trihydroxy bases, respectively. Thus, sphingosine is denoted as d18:1 and phytosphingosine is t18:0. The position of the double bond may be indicated by a superscript, i.e., 4-sphingenine is d18:1$^{\Delta 4t}$ or 4*E*-d18:1. While alternative nomenclatures are occasionally seen in publications, they are not recommended.

The number of different long-chain bases that has been found in animals, plants and microorganisms now amounts to over one hundred, and many of these may occur in a single tissue or organism, but almost always as part of a complex lipid with an *N*-acyl-linked fatty acid and often phosphate or carbohydrate functional groups, as opposed to in the free form. The aliphatic chains can contain from 14 to as many as 28 carbon atoms, and most often they are saturated, monounsaturated or diunsaturated, with double bonds of either the *cis* or *trans* configuration. For example, the main dienoic long-chain base (sphingadienine) in human plasma is D-erythro-1,3-dihydroxy-2-amino-4-*trans*,14-*cis*-octadecadiene, and this is especially abundant in kidney, with more in women than in men. It is not present in zebra fish, widely used as a model species. Forms with three double bonds, such as sphinga-4*E*,8*E*,10*E*-trienine, sometimes with a methyl group in position 9, have been found the sphingolipids of some marine invertebrates and in a dinoflagellate. In addition, long-chain bases can have branched chains with methyl substituents in the *omega*-1 (*iso*), *omega*-2 (*anteiso*) or other positions, hydroxyl groups in positions 4, 5 or 6, ethoxy groups in position 3, and even a cyclopropane ring in the aliphatic chain in some organisms. *N*-Methyl, *N*,*N*-dimethyl and *N*,*N*,*N*-trimethyl derivatives of sphingoid bases have been detected in mouse brain.

The main C_{18} components of long-chain bases of sphingomyelins of some animal tissues are accompanied by small amounts of C_{16} to C_{19} dihydroxy bases, although the latter attain higher proportions in tissues of ruminant animals. In gangliosides from human brain and intestinal tissues, eicosasphingosine (2*S*,3*R*,4*E*-d20:1) occurs in appreciable concentrations with variable amounts in different regions and membranes. However, human skin contains an especially wide range of isomers, including saturated, monoenoic and 6-hydroxy bases and phytosphingosines from C_{16} to C_{28} in chain-length. Shorter-chain bases are found in many insect species, and in the fruit fly, *Drosophila melanogaster*, which is widely used as a model species in genetic and metabolic experiments, the main components are C_{14} bases. In contrast to higher animals, nematodes such as *Caenorhabditis elegans* produce C_{17} *iso*-methyl-branched sphingoid bases, which are essential for normal sphingolipid function in the organism.

The long-chain base composition of individual lipids can vary markedly between species, tissues, organelles and even different membranes within a single organelle. For example, the data in Table 21.4.1 is perhaps from an extreme example, but it illustrates that remarkable differences that can exist among lipids in one cellular component (rat liver mitochondria). Only part of the data from the paper cited is listed, but it illustrates that 3-keto-sphinganine, produced in the first step of sphingosine biosynthesis (see below) and normally a minor component of sphingolipids - often not detectable, can vary from 28 to 100% of the sphingoid bases depending on the lipid class and membrane within the organelle.

Table 21.4.1: **Long chain base composition of some lipid components of mitochondria from rat liver.**

Type	Base (%)			
	d18:1	d18:0-3keto	t21:1 (phyto)	Unidentified
Ceramides[a]	18	28	53	-
Glucosylceramides[a]	3	95	-	3
Lactosylceramides[b]	100			

[a] whole mitochondria; [b] mitochondrial inner membrane
Data from Ardail, D. *et al. FEBS Letts*, **488**, 160-164 (2001).

Phytosphingosine or 4D-hydroxy-sphinganine ((2*S*,3*R*,4*R*)-2-amino-octadecanetriol) is a common long-chain base of mainly plant origin. It is a saturated C_{18}-trihydroxy compound, although unsaturated analogues, for example with a *trans* (or occasionally a *cis* (*Z*)) double bond in position 8, i.e., dehydrophytosphingosine or 4D-hydroxy-8-sphingenine, tend to be much more abundant. In many plant species, there are

lipid class preferences also, and dihydroxy long-chain bases are more enriched in glucosylceramides than in glycosylinositolphosphoceramides, for example. This is true in the model plant *Arabidopsis thaliana*, where the data listed for whole tissue is probably representative largely of the latter lipid, as shown in Table 21.4.2 below.

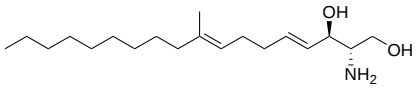
Table 21.4.2: **Sphingolipid long-chain base composition of whole tissue and glucosylceramides from *Arabidopsis thaliana*.**

	Base (%)					
	t18:1 (8*Z*)	t18:1 (8*E*)	t18:0	d18:1 (8*Z*)	d18:1 (8*E*)	d18:0
Whole tissue	12	70	13		4	1
Glucosylceramides	44	22		5	28	2
Data from Sperling, P. *et al*. *Plant Physiol. Biochem*., **43**, 1032-1038 (2005)						

Other plant long-chain bases have double bonds in position 4, which can be of either the *cis* or *trans* configuration, although *trans*-isomers are by far the more common, while the base d18:2$\Delta^{4E,8Z/E}$ is relatively abundant in most plant species. In *A. thaliana* and related species, Δ^4 long-chain bases are found mainly in the flowers and pollen and then exclusively as a component of the glucosylceramides. In general outwith Brassica species, the composition is dependent on species, but typically it is composed of up to eight different C_{18}-sphingoid bases, with variable geometry of the double bond in position 8, i.e., (*E/Z*)-sphing-8-enine (d18:1Δ^8), (4*E*,8*E/Z*)-sphinga-4,8-dienine (d18:2$\Delta^{4,8}$) and (8*E/Z*)-4-hydroxy-8-sphingenine (t18:1Δ^8); d18:1Δ^4, d18:0 and t18:0 tend to be present in small amounts only.

Phytosphingosine is not restricted to plants, but is found in significant amounts in intestinal cells and skin of animals, with much smaller relative proportions in kidney. Although non-mammalian sphingoid bases in general tend to be poorly absorbed from the intestines, a small proportion of the phytosphingosine and related sphingoid bases found in animal tissues may enter via the food chain.

Yeasts and fungi tend to have distinctive and characteristic long-chain base compositions. For example, filamentous fungi have 9-methyl-4*E*,8*E*-sphingadienine as the main sphingoid base in the glucosylceramides, as shown in Figure 21.4.7, but not in the ceramide phosphoinositol glycosides, while yeasts contain mainly the saturated C_{18} bases sphinganine and phytosphingosine, although some *trans*-4/8-unsaturated forms are usually present. Only a few bacterial species synthesize sphingolipids, but the family Bacteroidetes, which is abundant in the human gut is an important exception; they usually contain saturated (and branched) long-chain bases. Other pathogenic bacteria may utilize sphingolipids and sphingoid bases from their hosts.



(4E,8E)-9-methylsphinga-4,8-dienine

Figure 21.4.7: Main sphingoid base in filamentous fungi

Sphingoid bases are surface-active amphiphiles with critical micellar concentrations of about 20 μM in aqueous solutions; they probably exist in the gel phase at physiological temperatures. In that they bear a small positive charge at neutral pH, they are unusual amongst lipids, although their pK_a (9.1) is lower than in simple amines as a consequence of intra-molecular hydrogen bonding. Together with their relatively high solubility (> 1μM), this enables them to cross membranes or move between membranes with relative ease. In so doing, they increase the permeability of membranes to small solutes. In esterified form in complex lipids, they participate in the formation of ordered lipid domains in membranes such as **rafts**.

In the complex sphingolipids, the sphingoid base is linked via the amine group to a fatty acid, including very-long-chain saturated or monoenoic and 2-hydroxy components, i.e., to form **ceramides**, which can be attached a polar head group, such as phosphate or a carbohydrate, via the primary hydroxyl moiety. An important exception is **sphingosine-1-phosphate**, which is not acylated and has signaling functions in cells akin to those of lysophospholipids.

21.4.2.2: BIOSYNTHESIS AND METABOLISM

Sphinganine biosynthesis

The basic mechanism for the biosynthesis of **sphinganine** involves condensation of palmitoyl-coenzyme A with L-serine, catalyzed by the membrane-bound enzyme **serine palmitoyltransferase**, requiring pyridoxal 5'-phosphate as a cofactor, which binds to a specific lysine residue on the enzyme. The reaction occurs on the cytosolic side of the endoplasmic reticulum in animal, plant and yeast cells with formation of 3-keto-sphinganine as illustrated in Figure 21.4.8.

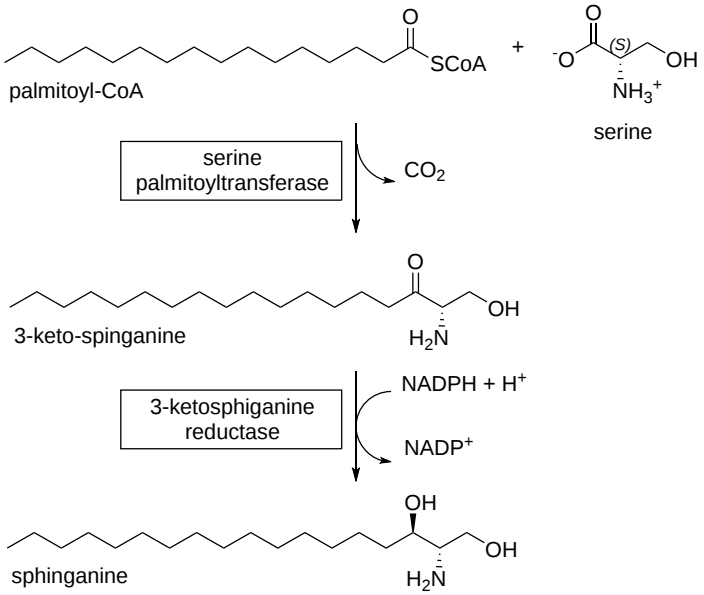


Figure 21.4.8: Sphinganine synthesis

This is believed to be the key regulatory or rate-limiting step in sphingolipid biosynthesis and is conserved in all organisms studied to date. Elimination of this enzyme is embryonically fatal in mammals and fruit flies. In mammals, serine palmitoyltransferase is a heterotrimer composed of two main subunits, designated SPTLC1 with either SPTLC2 or SPTLC3 (sometimes termed SPTLC2a and SPTLC2b, respectively). SPTLC1 is essential for activity, and it is ubiquitously expressed as is SPTLC2, while SPTLC3 is present in a relatively limited range of tissues and is most abundant in skin and placental tissue. In addition, there are two small subunits ssSPTA and ssSPTB (again other nomenclatures exist), which differ in a single amino acid residue, and may have regulatory functions; the active site is at the interface between the two main subunits. ssSPTA is essential for serine palmitoyltransferase function during development and hematopoiesis.

A possible mechanism for the 1st step in the pathway, catalyzed by serine palmitotyltransferase, is shown in Figure 21.4.9.

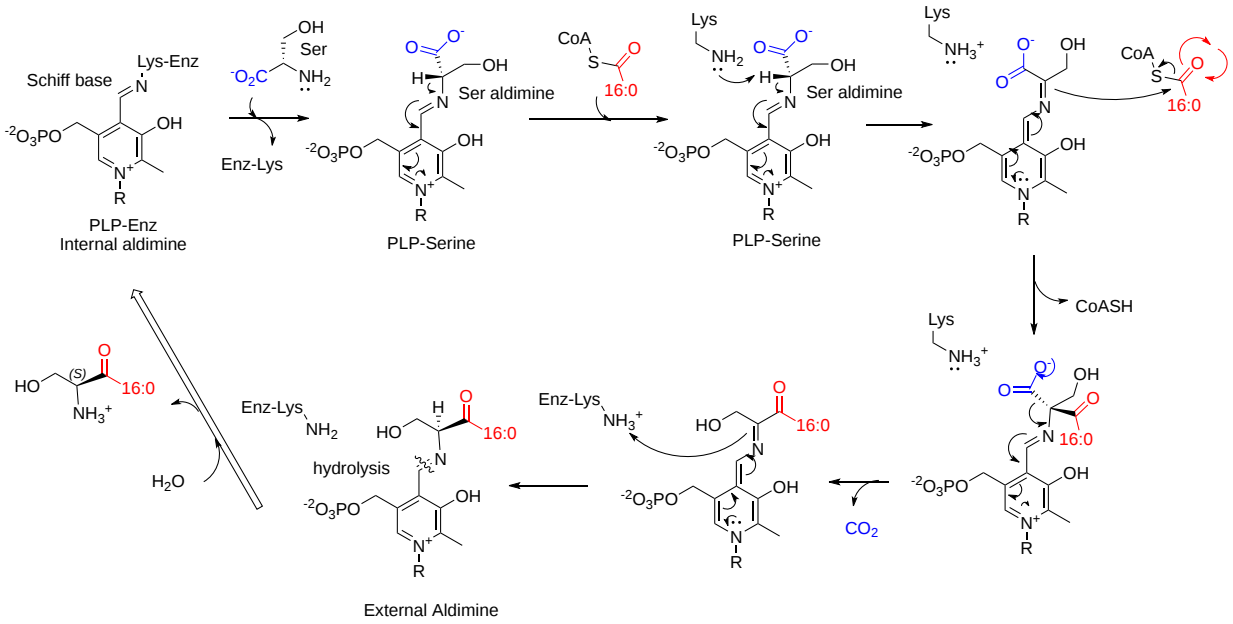
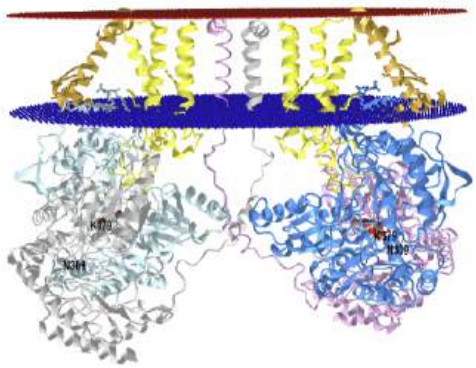


Figure 21.4.9: Mechanism for serine palmitotyltransferase, the key regulatory step in sphingolipid metabolism

The addition of either of the two small subunits to the complexes changes the substrate preferences substantially and enables the synthesis of the wide range of homologs found in nature. In mammals, the SPTLC1-SPTLC2 complex forms C_{18} sphingoid bases specifically (with some C_{19}, and C_{20}), while the combination of SPTLC1 and SPTLC3 gives a broader product spectrum, including an *anteiso*-methylbranched-C_{18} isomer (from *anteiso*-methyl-palmitate as the precursor). Such branched bases are synthesized to a limited extent in human skin, but they are the main forms in lower invertebrates such as *C. elegans*. The activity of the serine palmitoyltransferase is governed by negative feedback and partly by orosomucoid (ORM-like or ORMDL) proteins, three in mammals (ORMDL1 to 3) and two in

yeast (Orm1/2), which are ubiquitously expressed trans-membrane proteins located in the endoplasmic reticulum. The availability of serine is also an important factor.

Figure 21.4.10 shows an interactive iCn3D model of the human serine palmitoyltransferase complex (7K0M).



NCBI iCn3D Figure 21.4.10: Human serine palmitoyltransferase complex (7K0M).. Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...tuGThTq9Vi4J58

- Gray and Plum (SPT1 A and E Chains
- Light Blue and Blue (SPT2 B and F Chains)
- Light Brown (small subunit A - ssA, C and G chains)
- Yellow (ORM D and H)

The second step in sphinganine biosynthesis is reduction of the keto group to a hydroxyl in an NADPH-dependent manner by a specific 3-ketodihydrosphingosine reductase ('3KSR'), also on the cytosolic side of the endoplasmic reticulum, a step that must occur rapidly as the intermediate is rarely encountered in tissues. The enzymes are presumed to be in similar subcellular locations in plant cells.

In plants, serine palmitoyltransferase is a heterodimer composed of LCB1 and LCB2 subunits with some homology to the mammalian enzymes, while in the yeast *Saccharomyces cerevisiae*, there are three subunits: Lcb1, Lcb2, and Tsc3. In the few bacteria that synthesize sphingoid bases, serine palmitoyltransferase is a water-soluble homodimer. The enzyme in the apicomplexan parasite *Toxoplasma gondii* is a homodimer also in contrast to other eukaryotes, but it is located in the endoplasmic reticulum.

Free sphinganine formed in this way is rapidly *N*-acylated by acyl-coA to form dihydroceramides by dihydroceramide synthases, which in animals are located primarily on the endoplasmic reticulum, presumably on the cytoplasmic surface. Animals and plants have multiple isoforms of this enzyme, for which the abbreviated term ‘ceramide synthase’ is now widely applied as they utilize most other sphingoid bases, such as those produced by hydrolysis of sphingolipids, as substrates. They are unique gene products with each located on a different chromosome and with considerable variation in the expression of the enzymes in different cell types within each tissue. Each isoenzyme has distinct specificities for the chain-length of the fatty acyl-CoA moieties but to a limited extent only for the base, suggesting that ceramides containing different fatty acids have differing roles in cellular physiology. All of these enzymes have six membrane spanning regions, but the only substantial difference is in an 11-residue sequence in a loop between the last two putative transmembrane domains. **Ceramides** are central to all elements of sphingolipid biochemistry. These steps are illustrated in Figure 21.4.11.

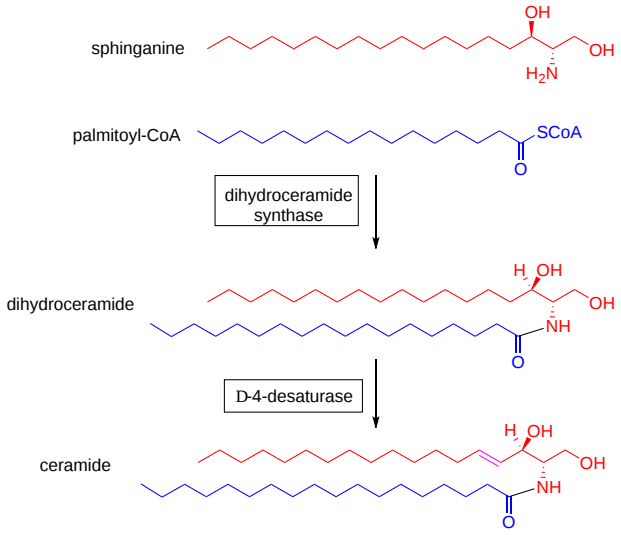


Figure 21.4.11: Synthesis of ceramide from sphinganine.

Humans and mice have six ceramide synthases, which utilize subsets of acyl-CoAs and thus producing ceramides with specific acyl chain lengths. Of these, ceramide synthase 2 is most abundant and is specific for coA esters of very-long-chain fatty acids (C_{20} to C_{26}); it is most active in lung, liver and kidney. Ceramide synthase 1 is specific for 18:0 and is located mainly in brain with lower levels in skeletal muscle and testes. Ceramide synthase 3 is responsible for the unusual ceramides of skin and testes and uses C_{26}-CoA and higher including polyunsaturated-CoAs with the latter tissue, while ceramide synthase 4 (skin, liver, heart, adipose tissue and leukocytes) uses C_{18} to C_{22}-CoAs. Ceramide synthases 5 (lung epithelia and brain gray and white matter) generates C_{16} (mainly) and C_{18} ceramides, and ceramide synthase 6 (intestine, kidney and lymph nodes) produces C_{14} and C_{16} ceramides. However, hydroxylation and the presence or otherwise of double bonds in the acyl-coAs do not appear to influence the specificity of the ceramide synthases. Also, the expression of mRNA expression for ceramide synthases does not always correlate with the fatty acid composition of sphingolipids in a particular tissue, suggesting that other factors are involved in determining which molecular species are formed. One such is acyl-coenzyme A-binding protein (ACBP), which facilitates the synthesis of ceramides containing very-long fatty acids and stimulates ceramide synthases 2 and 3 especially.

Insertion of the *trans*-double bond in position 4 to produce sphingosine occurs only **after** the sphinganine has been esterified in this way to form a ceramide as illustrated in Figure 21.4.11, with desaturation occurring at the cytosolic surface of the endoplasmic reticulum also. The desaturases were first characterized in plants, and this subsequently simplified the isolation of the appropriate enzymes in humans and other organisms. Two dihydroceramide desaturases have now been identified in animals and designated 'DEGS1 and DEGS2'. Both enzymes insert *trans* double bonds in position 4, but DEGS2 is a dual function enzyme that also acts as a hydroxylase to generate phytoceramides, i.e., to add a hydroxyl group on position 4. Distribution of the enzymes in tissues is very different, with DEGS1 expressed ubiquitously but highest in liver, Harderian gland, kidney and lung. DEGS2 expression is largely restricted to skin, intestine and kidney, where phytoceramides are more important. A considerable family of Δ^4-sphingolipid desaturases has now been identified, and an early study by Stoffel and colleagues demonstrated that Δ^4-desaturation involves first *syn*-removal of the C(4)- H_R and then the C(5)-H_S hydrogens. This appears to have been the first evidence that desaturases in general operate in this stepwise fashion.

The enzyme responsible for the insertion of the *cis*-14 double bond into sphinga-4-*trans*,14-*cis*-dienine is the fatty acid desaturase 3 (FADS3), which utilizes ceramides containing sphingosine as the precursor. The only other known activity of this enzyme is to insert a *cis*-double bond in position 13 of the CoA ester of vaccenic acid (11*t*-18:1) to produce the conjugated diene 11*t*,13*c*-18:2.

Synthesis of sphingoid bases *de novo* is essential in most organisms and inhibition of the biosynthetic pathways affects growth and viability. However, this can be tissue specific, as deletion of the liver-specific SPTLC2 in mice, was found to have no effect on liver function, while a comparable deletion of adipocyte-specific SPTLC1 caused major tissue defects. Presumably, the latter tissue is unable to take up enough sphingolipid from the circulation to remedy the problem. Deficiencies in SPTLC3 are related to dermal pathologies, and genetic variant of SPTLC3 are associated with dyslipidemia and atherosclerosis. The essentiality of sphingoid base synthesis in plants has been demonstrated in a similar manner in studies with mutants in which specific enzymes have been deleted.

Phytosphingosine and plant ceramides: Phytosphingosine is formed from sphinganine, produced as above, by hydroxylation in position 4, possibly via the free base in plants, although it can be formed both from sphinganine and a ceramide substrate in yeasts. A single sphinganine C4-hydroxylase is present in yeast, but Arabidopsis has two such enzymes (SBH1 and 2), which are critical for growth and viability. Much remains to be learned of the processes involved, but it is known that the enzyme responsible is closely related to a Δ^4 desaturase. Indeed, it has been shown that there are bifunctional Δ^4-desaturase/Δ^4-hydroxylases in *Candida albicans* and mammals, especially in keratinocytes (DEGS2 discussed above) with which either 4-hydroxylation or Δ^4-desaturation is initiated by removal of the pro*R* C-4 hydrogen. Sphinganine linked to ceramide is the substrate for 4-hydroxylation in intestinal cells.

In *Arabidopsis thaliana* leaves, 90% of the sphingoid bases are phytosphingosine with a Δ^8-double bond. In plants in general, in addition to Δ^4-desaturation, two distinct types (20 gene products) of sphingoid Δ^8-desaturase have been characterized that catalyse the introduction of a double bond at position 8,9 of phytosphingosine. These are evolutionarily distinct from the Δ^4-desaturases. One type produces the *trans (E)*-8 isomer mainly and the other mostly the *cis (Z)*-8 isomer, with overall the *trans*-isomer tending to predominate but dependent upon plant species. It appears that the *trans* isomer is formed when the hydrogen on carbon 8 is removed first, and the *cis* when carbon 9 is the point of attack. While the main group of Δ^8-desaturases requires a 4-hydroxysphinganine moiety as substrate, the second does not.

In Arabidopsis, three different isoforms of ceramide synthase have been identified and denoted LOH1, LOH2 and LOH3. Phytosphingosine is used efficiently by LOH1 and LOH3 (class II synthases), but only LOH2 (class I synthase) uses sphinganine efficiently; LOH2 and 3 prefer unsaturated long-chain bases. Marked fatty acid specificity is also observed with LOH2 showing almost completely specific for palmitoyl-CoA and dihydroxy bases, while LOH1 shows greatest activity for 24:0- and 26:0-CoAs and trihydroxy bases; none utilize unsaturated acyl-CoA esters efficiently. In plants, fatty acid desaturases and hydroxylases are also closely related, and sphingolipid fatty acid α-hydroxylation is believed to occur on the ceramide, as opposed to the free acyl chain. It is believed that the Δ^8-desaturase utilizes ceramide as the substrate and the channels the products selectively into the synthesis of complex sphingolipids, while Δ^4-desaturation channels ceramides for synthesis of glucosylceramide.

It has been established that long-chain bases with 4-hydroxyl groups are necessary for the viability of the filamentous fungus *Aspergillus nidulans* and for growth in plants such as *A. thaliana*. The presence of an 8*E* double bond confers aluminium tolerance to yeasts and plants, and it is important for chilling resistance in tomatoes. However, a *trans*-4 double bond in the sphingoid base does not appear to be essential for growth and development in *Arabidopsis*.

Fungal sphingoid bases: Fungi produce *trans* Δ^8-isomers only, but Δ^4- and Δ^8-desaturases do not occur in the widely studied yeast *S. cerevisiae*. In the biosynthesis of sphingoid bases in fungi, the double bonds in positions 4 and 8 and the methyl group in position 9 are inserted sequentially into the sphinganine portion of a ceramide, the last by means of an *S*-adenosylmethionine-dependent methyltransferase similar to plant and bacterial **cyclopropane fatty acid synthases**. In *S. cerevisiae* the ceramide synthase is a heteromeric protein complex, containing three subunits, Lag1, Lac1, and Lip1, of which the first two are homologous proteins that feature eight transmembrane domains. In the yeast *Pichia pastoris*, there is a distinct ceramide synthase, which utilizes dihydroxy sphingoid bases and C_{16}/C_{18} acyl-coenzyme A as substrates to produce ceramides. The long-chain-base components of the ceramide are then desaturated *in situ* by a Δ^4-desaturase and the fatty acid components are hydroxylated in position 2. Further desaturation of the long-chain base component by a Δ^8-(*trans*)- desaturase occurs before the methyl group in position 9 is introduced by an *S*-adenosylmethionine-dependent sphingolipid C-9 methyltransferase. As a final step a *trans*-double bond may be introduced into position 3 of the fatty acid component. These ceramides are used exclusively for the production of **glucosylceramides**, and it is believed that a separate ceramide synthase encoded by a different gene produces the ceramide precursors for **ceramide phosphorylinositol mannosides**.

Viral sphingoid bases: The genome of an important marine virus (EhV) encodes for a novel serine palmitoyltransferase, which hijacks the metabolism of algal hosts to produce unusual hydroxylated C_{17} sphingoid bases; these accumulate in lytic cells of infected algae such as the important bloom-forming species *Emiliania huxleyi*. While this may seem a rather esoteric topic, viruses constitute a high proportion of the marine biome, and their control of the growth of algal blooms has global consequences.

Unesterified sphingosine: A cycle of reactions occurs in tissues by which sphingoid bases are incorporated via ceramide intermediates into sphingolipids, which are utilized for innumerable functions, before being broken down again to their component parts. It is worth noting that all the free sphingosine in tissues must arise by this route, in particular by the action of ceramidases on **ceramides**. Five such ceramidases are known with differing pH optima and varying subcellular locations. The levels of free sphingoids and their capacities to function as lipid mediators, as shown in Figure 21.4.12, are controlled mainly by enzymic re-acylation to form ceramides, although some is acted upon by sphingosine kinases to produce sphingosine-1-phosphate.

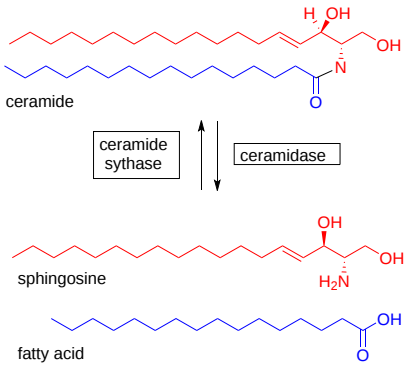


Figure 21.4.12 Ceramide hydrolysis and resynthesis

Free sphingoid bases are absorbed by enterocytes following digestion of dietary sphingolipids in animals (including some from gut microorganisms), and while some of this is converted to complex sphingolipids, much is catabolized with the eventual formation of palmitic

acid.

Catabolism of sphingosine and other long-chain bases occurs after conversion to **sphingosine-1-phosphate** and analogues. In yeasts, an alternative means of detoxification has been reported in which an excess of phytosphingosine is first acetylated and then converted to a vinyl ether prior to export from the cells.

21.4.2.3: BIOLOGICAL FUNCTIONS OF UNESTERIFIED SPHINGOID BASES

The primary function of sphingoid bases is to serve as a basic component of ceramides and complex sphingolipids, where variations in their compositions can influence the physical and biological properties of these lipids. Independently of this in their free (unesterified) form, they are important mediators of many cellular events even though they are present at low levels only in tissues (typically 25 and 50 nM in plasma), with intracellular levels determined by hydrolysis by ceramidases or by the action of sphingosine kinases (sphingosine-1-phosphate production). In animal cells, they inhibit protein kinase C indirectly, possibly by a mechanism involving interference with the binding of activators of the enzyme, such as diacylglycerols or phorbol esters. In addition, sphingoid bases are known to be potent inhibitors of cell growth, although they stimulate cell proliferation and DNA synthesis. They are involved in the process of apoptosis in a manner distinct from that of ceramides by binding to specific proteins and regulating their phosphorylation. While sphingosine does not appear to participate in raft formation in membranes, it may rigidify pre-existing gel domains in mixed bilayers, although any such effects will be dependent on local concentrations and pH. It should be noted that some of the biological effects observed experimentally may be due to conversion to sphingosine-1-phosphate.

Free sphingosine has been implicated in various pathological conditions, and for example, plasma sphingosine levels are increased in hyperthyroidism and in patients with type 2 diabetes. Lysosomal storage of the lipid is an initiating factor in Niemann Pick type C disease, a neurodegenerative disorder, where it causes a change in calcium release leading to a buildup of cholesterol and sphingolipids. In the human adrenal cortex, sphingosine produced *in situ* by the acid ceramidase has a function in steroid production by serving as a ligand for steroidogenic factor 1 at the cell nucleus, which controls the transcription of genes involved in the conversion of cholesterol to steroid hormones. Unesterified sphingoid bases may have a protective role against cancer of the colon in humans. Thus, *N*,*N*-dimethylsphingosine and dihydrosphingosine, like the deoxysphingoid bases, are known to induce cell death in a variety of different types of malignant cells. There is evidence that sphingadienes of plant and animal origin inhibit colorectal cancer in mouse models by reducing sphingosine-1-phosphate levels. In consequence, synthetic analogues of long-chain bases are being tested for their pharmaceutical properties.

Free sphingosine is believed to have a signaling role in plants by controlling pH gradients across membranes. In addition, free long chain bases (and the balance with the 1-phosphate derivatives) are essential for the regulation of apoptosis in plants.

21.4.3: CERAMIDES

21.4.3.1: STRUCTURE AND OCCURRENCE

The structure of ceramide is shown again in Figure 21.4.13.

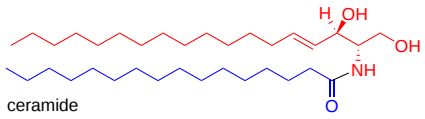


Figure 21.4.13: Structure of ceramides (with varying fatty acids in ester link)

Ceramides consist of a long-chain or **sphingoid base** linked to a fatty acid via an amide bond. They are essential intermediates in the biosynthesis and metabolism of all sphingolipids including the complex sphingolipids in which the terminal primary hydroxyl group is linked to carbohydrate, phosphate, and so forth (sphingomyelin, glycosphingolipids and gangliosides) as shown in Figure 21.4.14.

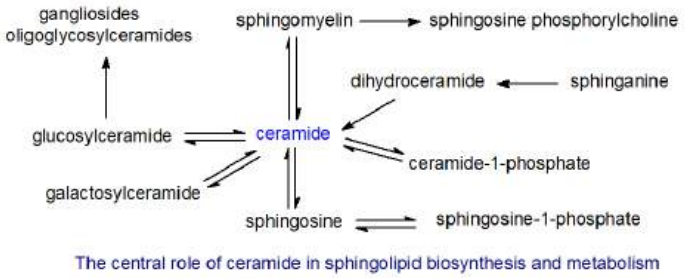


Figure 21.4.14 Central role of ceramkin in sphingolipid biosynthesis and metabolism

They are also the primary source of unesterified sphingoid bases and of the important biological mediators sphingosine-1-phosphate and ceramide-1-phosphate. At the last count, 33 different enzymes were known to participate in ceramide metabolism. While ceramides are rarely found as such at greater than trace amounts in tissues other than skin, they can exert important biological effects of their own at these low levels. They are present in membranes where they participate in the formation of raft domains.

Each organism and indeed each tissue may synthesize ceramides in which there are a variety of di- and trihydroxy long-chain bases linked to fatty acids. As discussed previously, the fatty acids consisting mainly of longer-chain (up to C_{24} or greater) saturated and monoenoic (mainly (*n*-9)) components, sometimes with a hydroxyl group in position 2. Other than in certain testicular cells, polyunsaturated fatty acids do not occur. More than 200 structurally distinct molecular species of ceramides have been characterized from mammalian cells. In plants, 2-hydroxy acids predominate sometimes accompanied by small amounts of 2,3-dihydroxy acids. Although small amounts of free ceramides are produced in all tissues as required for the specific biological functions described below, most is converted rapidly to more complex sphingolipids, including sphingomyelin (in animals) and the various glycosylceramides. The ceramides in skin are a remarkable exception to this rule, and as such they are discussed separately below.

A shorthand nomenclature simply combines those used conventionally for fatty acids and long-chain bases to denote molecular species of ceramides, including those as components of more complex lipids, e.g. *N*-palmitoyl-sphingosine is d18:1-16:0. Ceramides containing sphinganine are sometimes termed 'dihydroceramides'.

21.4.3.2: CERAMIDE BIOSYNTHESIS

Ceramide production is complex and involves at least three pathways. Biosynthesis *de novo* takes place in the endoplasmic reticulum with palmitoyl-CoA and serine as the precursors for the long-chain base component, which is subsequently converted to ceramide. Biosynthesis of the very specific fatty acids in ceramides involving various chain elongases (ELOVL) requires consideration also. Alternative routes for ceramide production involve regeneration from complex sphingolipids. For example, in animals in the sphingomyelinase pathway, conversion of sphingomyelin into ceramides (and *vice versa*) occurs in the plasma membrane, Golgi and mitochondria. Finally, the polar moieties of complex glycosphingolipids can be removed by various hydrolytic enzymes in the lysosomal compartment to recover the ceramides (or their component parts) in a re-cycling/catabolic process. As these biosynthetic or metabolic pathways are located in different organelles, specific pools of ceramide and sphingolipids result with differing biological properties and functions.

Ceramide synthesis *de novo*: The first of these pathways is described in mechanistic. In brief in animals, sphinganine is coupled to a long-chain fatty acid to form dihydroceramide by means of one of six ceramide synthases in the endoplasmic reticulum mainly, before the double bond is introduced into position 4 of the sphingoid base. Of these, ceramide synthase 2 is most abundant and is specific for CoA esters of very-long-chain fatty acids (C_{20} to C_{26}); it is most active in the central nervous system. Ceramide synthase 1 is specific for 18:0 and is located exclusively in brain and skeletal muscle, ceramide synthases 5 and 6 generate 16:0-containing ceramides, and ceramide synthase 3 is responsible for the unusual ceramides of skin and testes.

Figure 21.4.15 shows again the synthesis of ceramide from sphinganine and palmitoyl-CoA (a repeat of Figure 21.4.11

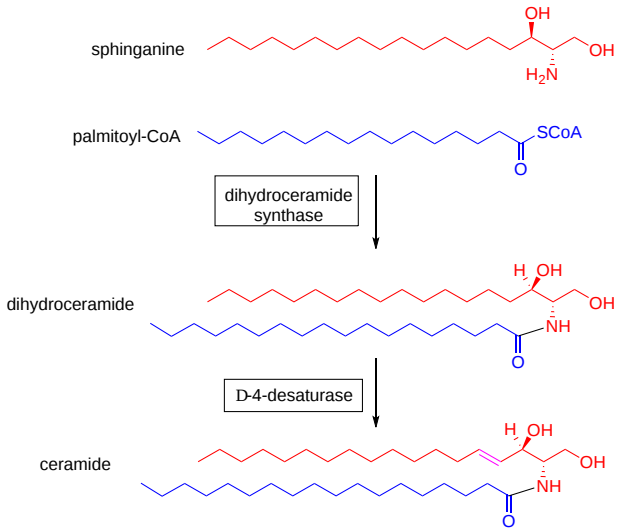


Figure 21.4.15: Synthesis of ceramide from sphinganine and palmitoyl-CoA

Each synthase has six membrane-spanning domains and contains a characteristic motif with the specific structures required for catalysis and substrate binding that are essential for its activity, and they have been shown to differ primarily in an 11-residue sequence in a loop between the last two putative transmembrane domains. In addition to separate transcriptional regulation of each of these enzymes, ceramide synthase activity is modulated by many different factors including reversible dimerization, while ceramide synthase 2 has a **sphingosine-1-phosphate** binding motif and this lipid may inhibits its activity. Acyl-coenzyme A-binding protein (ACBP) facilitates the synthesis of ceramides containing very-long fatty acids and stimulates ceramide synthases 2 and 3 especially.

Most of the ceramides generated in this way are rapidly utilized for synthesis of complex sphingolipids, especially sphingomyelin and hexosylceramides, to ensure that cellular ceramide concentrations are regulated to control their biological activities. In mammalian cells, most complex glycerolipids are synthesized in the endoplasmic reticulum prior to their transport to their final subcellular locations, but the

process is rather different for sphingolipids. Ceramide is synthesized on the cytoplasmic leaflet of the endoplasmic reticulum, but subsequent formation of complex sphingolipids occurs in the Golgi apparatus, and a key cytoplasmic protein, ceramide transporter or '**CERT**' (CERamide Trafficking), mediates the transport of ceramide between these organelles in a non-vesicular manner. It has a number of distinct functional domains, including an N-terminal **phosphatidylinositol-4-monophosphate (PI(4)P)**-binding or Pleckstrin homology (PH) domain, which targets the Golgi apparatus, and a C-terminal 'START' domain, which can recognize ceramide species with the natural D-*erythro* stereochemistry, including dihydroceramide and phytoceramide (but not sphingosine), and holds them within in a long amphiphilic cavity by hydrogen bonding with all three polar atoms of the sphingoid motif. There is also a short peptide motif (FFAT) that recognizes a specific protein in the endoplasmic reticulum. There is sufficient flexibility in the body of the protein to enable transfer of ceramide from the endoplasmic reticulum to the Golgi without free movement through the cytosol.

Very-long-chain ceramides containing 24:0 or 24:1 fatty acids turn over much more rapidly in animal cells than those containing 16:0 or 18:0 fatty acids, because of the more rapid conversion of the former into complex sphingolipids, where they may regulate the levels and perhaps the biological functions of the latter. In contrast, ceramides containing d16:1 and d18:1 sphingoid bases turnover at similar rates so do not affect the flux of ceramides through these pathways. The CERT protein is a major factor in this specificity, as it extracts ceramides from membrane bilayers with a preference for those required for synthesis of complex sphingolipids. Removal of ceramide by this process provides the gradient that enables the process to continue, and prevents an accumulation of ceramide in the endoplasmic reticulum that might otherwise be disruptive to the membrane and even cause cell death. While the transfer process itself is not dependent on ATP, the overall process requires ATP, possibly to keep PI(4)P in a phosphorylated form, and the multiple factors that control the biosynthesis of this lipid must also influence sphingolipid metabolism.

As a neutral lipid, ceramide can flip readily across membrane leaflets, and this is also necessary for the synthesis of sphingomyelin, which occurs on the lumen of the Golgi. The pool of ceramide utilized for the synthesis of **glycosylceramide** is delivered to the Golgi by a separate transport mechanism that also does not require ATP. In addition, some ceramide synthesis occurs in mitochondria although this has the potential to lead to cell death. Regulation of ceramide and subsequent sphingolipid biosynthesis is crucial as an excess of sphingolipids can be toxic, while reduced synthesis can inhibit cell proliferation.

Some ceramides are transported from the liver to other tissues in plasma **lipoproteins**, but especially subclasses HDL_2 and HDL_3, i.e. those containing apolipoprotein B. There is a suggestion that transport of ceramides via lipoproteins could be a paracrine mechanism to regulate the metabolism of other cells.

Ceramides are also produced during the catabolism of other complex sphingolipids, and especially by the action of one or other of the sphingomyelinases or of phospholipase C on **sphingomyelin** in animal tissues as part of the 'sphingomyelin cycle' as shown in Figure 21.4.16.

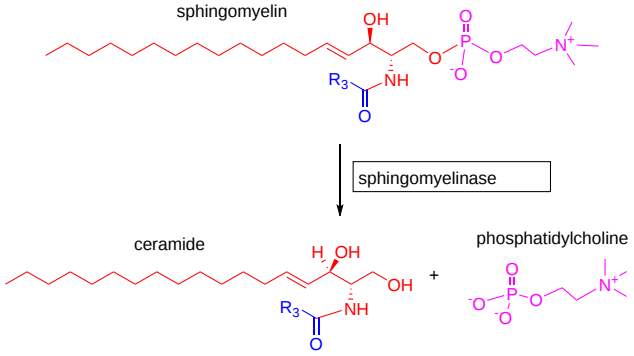


Figure 21.4.16: Ceramide formation from sphinogmyelin

Many agonists including chemotherapeutic agents, tumor necrosis factor-alpha, 1,25-dihydroxy-vitamin D_3, endotoxin, gamma-interferon, interleukins, nerve growth factor, ionizing radiation and heat stimulate hydrolysis of sphingomyelin to produce ceramide. In addition, reversal of the sphingomyelin synthesis reaction may generate ceramide, and some may be produced by operation of the enzyme ceramidase in reverse (see next section). Such reactions are much more rapid than synthesis *de novo*, so they are of special relevance in relation to the signaling functions of ceramides, especially when they occur at the plasma membrane. For example, in this context, the acid sphingomyelinase may be especially important by generating the ceramides that initiate the train of events that leads to apoptosis (see below).

Glycosphingolipids can be hydrolyzed by glycosidases to ceramides also in tissues, but the process tends to be less important in quantitative terms (other than in skin). The key enzymes of sphingolipid metabolism were first characterized from the yeast *Saccharomyces cerevisiae*, and these were found to be sufficiently similar to the corresponding enzymes in mammals to facilitate their study in the latter.

As discussed above, there are specific ceramide synthases that utilize specific fatty acids for ceramide biosynthesis in animals, and knowledge is slowly being acquired of how these are compartmentalized and regulated within cells. Thus, the synthesis and subsequent catabolism of ceramides involves a complex web of at least 28 distinct enzymes, including six ceramide synthases and five

sphingomyelinases, which are all products of different genes. Each of these enzymes may produce distinctive molecular species of ceramides with their own characteristic biological properties. It has been determined that ceramide species containing very-long-chain fatty acids (C_{24}) turnover more rapidly than those containing $C_{16/18}$ components.

21.4.3.3: CERAMIDE CATABOLISM

In animals, ceramide metabolism and function are controlled in part by the action of **ceramidases**, which cause hydrolysis forming sphingoid bases and free fatty acids, and indeed this is the only route to the formation of unesterified sphingosine. This is illustrated in Figure 21.4.17.

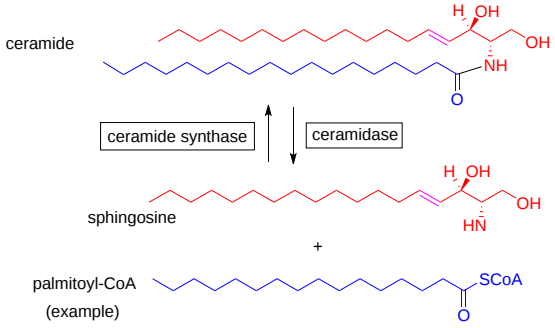


Figure 21.4.17: Ceramide hydrolysis-resynthesis

Five such enzymes are known in humans, classified according to their pH optima, i.e. acid ('ASAH1'), neutral ('ASAH2', which differs between humans and animals), and alkaline (three enzymes - 'ACER1 to ACER3'), with differing cellular locations and fatty acid specificities and with the potential to affect distinct signaling and metabolic events. The acid ceramidase is of particular importance, and aberrations in its synthesis or activity is involved in several human disease states, including the rare autosomal-recessive **Farber disease** where there is a deficiency in the enzyme so ceramide accumulates; ceramide containing 26:0 in the blood is considered to be a biomarker for diagnosis of the disease. ASAH1 is located in the lysosomes and hydrolyses ceramides with small to medium-chain fatty acid components (C_6 to C_{18}) most efficiently. The neutral ceramidase is located in the plasma membrane and Golgi, especially of intestinal epithelial cells and colorectal tissues, and prefers long-chain components (C_{16} to C_{18}); it also catalyzes the reverse reaction, and this may be a means of ceramide synthesis in mitochondria. ACER1 and ACER2 are found in the endoplasmic reticulum and Golgi, respectively, and they prefer species with very-long-chain acyl groups. ACER3 is present in both the endoplasmic reticulum and Golgi; it has a marked specificity for ceramides, dihydroceramides, and phytoceramides linked to unsaturated long-chain fatty acids (18:1, 20:1 or 20:4) *in vitro* at least. Neutral/alkaline ceramidase activity has also been found in mitochondria and nuclei.

In Arabidopsis, an alkaline ceramidase (AtACER) can hydrolyze phytosphingosine-containing ceramides, and a related enzyme from rice has a preference for d18:1Δ4-ceramide; the latter can function in reverse to increase the content of C_{26}- and C_{28}-phytoceramides. Several neutral ceramidases (AtNCERs) have been identified, but there does not appear to be an equivalent to the acid ceramidase in plants. Ceramidases are also present in lower organisms such as *Pseudomonas aeruginosa* and slime molds, where they are secreted proteins rather than integral membrane enzymes. A neutral ceramidase only is found in prokaryotes, including some pathogenic bacteria.

Sphingoid bases released by the action of acid ceramidase can escape from the lysosomes and be re-utilized for ceramide biosynthesis through the action of a ceramide synthase. This has been termed the 'salvage' pathway and is important in both quantitative and biological terms. For example, it has been estimated that it contributes from 50 to 90% of sphingolipid biosynthesis. The biological functions of ceramides are discussed below, but there are reasons to believe that ceramides derived from the salvage pathway are spacially and thence functionally distinct from those synthesized *de novo*. In addition, sphingoid bases released in this way have their own biological functions, which includes utilization for the synthesis of the biologically important metabolite **sphingosine-1-phosphate**. Therefore, regulation of ceramidase action is central to innumerable biological processes in animals.

21.4.3.4: BIOLOGICAL FUNCTIONS OF CERAMIDES

The role of ceramides in the biosynthesis of complex glyco- and phospho-sphingolipids are discussed elsewhere in this text. Ceramides, like other lipid second messengers in signal transduction, are produced rapidly and transiently in response to specific stimuli in order to target specific proteins, for example to activate certain serine/threonine protein kinases or phosphatases. They may also regulate cellular processes by influencing membrane properties. While they can be produced by synthesis *de novo* for such functions, activation of one of the sphingomyelinases under physiological stress or other agents is a more rapid means of generation in animal tissues at least. In fact, ceramides appear to be formed under all conditions of cellular stress by a multiplicity of activators in eukaryotic organisms. However, it should be noted that ceramides with different fatty acid and long-chain base (molecular species) compositions are formed in different compartments or membranes of the cell by various mechanisms over different time scales and potentially with distinct functions. The biological functions of those ceramides containing medium-chain (up to C_{14}), long-chain (C_{16} and C_{18}), and very-long-chain (C_{20} and longer) fatty acids, in particular, may have to be considered separately.

Physical properties: Unsaturation in the sphingoid backbone augments intramolecular hydrogen bonding in the polar region, which permits a close packing of ceramide molecules and a tight intramolecular interaction in membranes. A further important factor in this context is the length of the fatty acyl moiety, as shorter-chain ceramides tend to produce a positive curvature in a lipid monolayer, while long-chain molecules have the opposite effect and possess a marked intrinsic negative curvature that facilitates the formation of inverted hexagonal phases as well as increasing the order of the acyl chains in bilayers. By their interactions with ion channels, ceramides influence the permeability of membranes and render bilayers and cell membranes permeable to solutes that vary from small- up to protein-size molecules.

While ceramides are minor components of membranes in general, their physical properties ensure that they are concentrated preferentially into lateral liquid-ordered microdomains (a distinct form of '**raft**' termed 'ceramide-rich platforms'), although these effects are again chain-length specific. These domains differ appreciably in composition from those rafts enriched in **sphingomyelin** and cholesterol, and ceramides containing C_{12} to C_{18} fatty acids can in fact displace cholesterol from rafts to modify their physical properties. Ceramides are generated within rafts by the action of acid sphingomyelinase, causing small rafts to merge into larger units and modifying the membrane structure in a manner that is believed to permit oligomerization of specific proteins such as cytokines and death receptors. Ceramides are also essential for the formation and/or secretion of exosomes by facilitating or inducing membrane curvature. In contrast, sphingosine, sphingosine-1-phosphate and ceramide-1-phosphate do not facilitate raft formation.

Through the medium of these modified rafts, ceramides are able to function in signal transduction. Specific receptor molecules and signaling proteins are recruited and cluster within such domains, thereby excluding potential inhibitory signals, while initiating and greatly amplifying primary signals. It is believed that ceramide-rich platforms amplify both receptor- and stress-mediated signaling events and thence may influence various disease states. Ceramide-enriched membrane domains formed in response to sphingomyelinase activity are sites for endocytic uptake of pathogens because of a concentration of pathogen receptors and signaling complexes, and in particular these can enhance viral infections, including Norovirus, Japanese encephalitis virus, Ebola and possibly SARS-CoV-2. However, elevated levels of ceramide inhibit cellular uptake of the HIV virus.

Although ceramides and diacylglycerols have structural similarities, their occurrence, location, and behavior in membranes are different. Ceramides cross synthetic lipid bilayers relatively quickly *in vitro*, but it is not clear whether they can flip across more complex biological membranes equally readily, especially in the ceramide-rich platforms. Restricted flipping could have important effects on the signaling role of ceramides in that those generated by different enzymes on each side of a membrane could have distinct functions.

Enzyme activation: In general, ceramides tend to modify intracellular signaling pathways to slow anabolism and promote catabolism. Amongst a wide range of biological functions in relation to cellular signaling, ceramides are especially important in triggering apoptosis, and they have also been implicated in the activation of various protein kinase cascades, dependent on the site of generation. The mechanism of these interactions is the subject of intensive study at present, but in relation to the latter, two intracellular targets for ceramide action of special importance have been discovered – at least two protein phosphatases (ceramide-activated protein phosphatases) and a family of protein kinases (ceramide-activated protein kinases). For example, the phosphatase may be involved in the regulation of glycogen synthesis, insulin resistance, and response to apoptotic stimuli. Ceramides generated by the action of sphingomyelinase and by synthesis *de novo* are both important to the process, while ceramidases have contrasting effects in these and other biological effects of ceramides.

Apoptosis: The role of ceramides in the regulation of **apoptosis**, and cell differentiation, transformation, and proliferation has received special attention. Apoptosis is a normal process, which occurs in response to oxidative stress in particular, in which a cell can be considered to actively 'commit suicide'. It is essential for many aspects of normal development and is required for maintaining tissue homeostasis. There are two pathways - 'extrinsic' initiated in the plasma membrane by ligation of so-called 'death factors', such as the tumor necrosis factor-α (TNF-α), and 'intrinsic' induced by external actions in mitochondria, e.g. by DNA damage, oxidation or radiation injury. Although the mechanism of the ceramide interaction with these pathways is uncertain, it is clear that a cascade of reactions is initiated that culminates in the release of intracellular proteases of the caspase family to promote apoptosis. In dysfunctional mitochondria, one mechanism involves the formation of channels in the membrane that enable the release of specific mitochondrial proteins that include caspases. Ceramides with fatty acids of differing chain lengths are believed to function in different ways, and 16:0-ceramide generated by ceramide synthase 6 is especially pro-apoptotic, for example, while ceramides with very-long-chain fatty acids accumulate in necroptosis, a form of apoptosis. On the other hand, ceramides containing 2-hydroxy acids in keratinocytes appear to be protective against apoptosis. Ceramides induce the related process of cellular senescence also.

Failure to properly regulate apoptosis can have catastrophic consequences, and many disease states, including cancer, diabetes, neuropathies, Alzheimer's disease, Parkinson's disease, and atherosclerosis, are thought to arise from the deregulation of apoptosis. For example, ceramides have been implicated in the actions of TNF-α and in the cytotoxic responses to amyloid Aβ peptide, which are involved in Alzheimer's disease and neurodegeneration. In addition, ceramides appear to be involved in many aspects of the biology of aging and of male and female fertility. These effects may hold implications for diseases associated with obesity and insulin resistance, including again diabetes and cardiovascular disease.

Similarly, ceramides are intimately involved in the induction of autophagy, the 'maintenance' process by which cellular proteins and excess or damaged organelles are removed from cells by engulfing them in a membrane-enclosed cellular compartment called the phagosome. In

particular, maturing phagosomes are enriched in very-long-chain ceramides. While this process is beneficial in that it aids the recycling of cellular nutrients, the presence of excess ceramide can lead to unnecessary apoptosis.

As animals and plants have multiple isoforms of ceramide synthase that are specific for the chain length of the base and fatty acid, it has been suggested that ceramides containing different fatty acids have distinct roles in cellular physiology. In particular, C_{16} ceramide appears to be especially important in apoptosis in non-neuronal tissues, while C_{18} ceramide has growth-arresting properties and may be involved in apoptosis in some carcinomas treated with chemotherapy agents. In addition, a transferase has been identified that transfers the acetyl group from **platelet-activating factor** to sphingosine with a high specificity. The product, *N*-acetylsphingosine - the simplest of all ceramide molecules, has signaling functions that are distinct from those of the parent lipids or of other ceramides; it does not enter the salvage pathway in cancer cells *in vitro* and is cytotoxic.

In contrast, the ceramide metabolite, **sphingosine-1-phosphate**, has opposing effects on cell survival and proliferation. As ceramide and sphingosine-1-phosphate are inter-convertible via sphingosine as an intermediate, which also has pro-apoptopic activity, the balance between these lipids and with ceramide-1-phosphate is obviously of great metabolic importance. It has been termed the 'sphingolipid-rheostat', as illustrated in Figure 21.4.18.

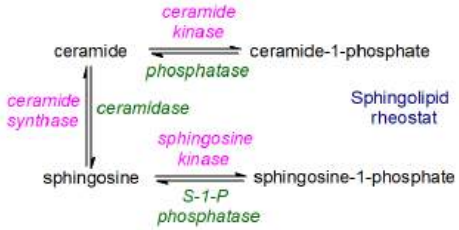


Figure 21.4.18 Sphingolipid rheostat

Plants: Comparatively little information is available on the role of ceramides in cell signaling in plants, but there are suggestions that sphingolipid catabolic products may be linked to programmed cell death, signal transduction, membrane stability, host-pathogen interactions, and stress responses. For example, there is evidence that enhanced synthesis of ceramides with very-long-chain fatty acids and trihydroxy sphingoid bases by ceramide synthases LOH1 and LOH3 promotes cell division and growth, while in contrast, accumulation of the ceramide species C_{16} fatty acid with a dihydroxy sphingoid base, due to LOH2 overexpression, leads to plant dwarfing and programmed cell death. Ceramides aggregate in rafts in plant membranes, together with other sphingolipids and sterols, as in animal tissues. Similarly, in the yeast *S. cerevisiae*, widely used as a model organism, it has been reported that ceramide species with different *N*-acyl chains and sphingoid bases are involved in the regulation of different sets of functionally related genes.

21.4.3.5: SKIN CERAMIDES

The mammalian skin forms the protective barrier between the internal tissues of the host and the hostile external environment, which can include chemicals, ultraviolet light, mechanical damage, and pathogenic microorganisms, while preventing the loss of water and electrolytes. It consists of stratified layers of increasingly differentiated cells or keratinocytes of which the basal layer is responsible for the renewal of the tissue but begins to migrate upwards and differentiate, while accumulating specific lipids and proteins that change the cellular architecture. Eventually, the keratinocytes lose their nucleus and become flattened structures of insoluble protein surrounded by lipids termed 'corneocytes' in the outermost impermeable layer or stratum corneum. By secreting peptides and proteins that possess antimicrobial activity, keratinocytes add to the defensive capability of skin against commensal microorganisms and opportunistic pathogens, and this is reinforced by lipid mediators such as free sphingoid bases and eicosanoids in the stratum corneum and free fatty acids in sebum.

The stratum corneum contains high levels of ceramides (as much as 50% of the total lipids), including *O*-acylceramides, which exist both in the free form and linked by ester bonds to structural proteins. They are present mainly in the extracellular domains (interstices) and are accompanied by nearly equimolar amounts of cholesterol and free fatty acids, a ratio that is believed to be essential for the normal organization of the tissue into the membrane structures that are responsible for the functioning of the epidermal barrier. In contrast to other biological membranes, the lipid organization in the membranes of skin consists of two lamellar phases, which form crystalline lateral phases mainly, with repeat distances of approximately 6 and 13 nm. Small sub-domains of lipids in a liquid phase may also exist.

Some of these skin ceramides have distinctive structures not seen in other tissues, and many different forms are commonly recognized. They can contain the normal range of longer-chain fatty acids (a), e.g. formula 1 in the figure, some with hydroxyl groups in position 2 (a*), e.g. formula 2, linked both to dihydroxy bases with *trans*-double bonds in position 4 or to trihydroxy bases. This is illustrated in Figure 21.4.18.

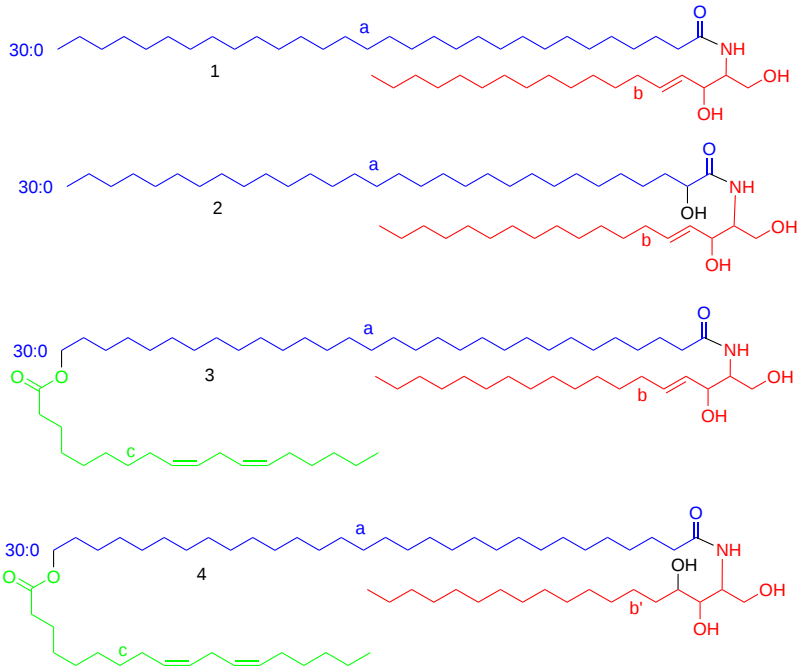
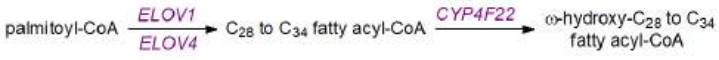


Figure 21.4.18 Skin ceramides

In addition, there are *O*-acyl ceramides in which a unique very-long-chain fatty acid component (typically C_{30} or C_{32}) has a terminal hydroxyl group, and this may be in the free form or esterified with linoleate (**c**), e.g., formulae 3 and 4; the sphingoid base can be either di- (**b**) or trihydroxy (**b***), e.g., formula 4; the latter is not a common feature in sphingolipids of animal origin, and can include both phytosphingosine and the unique 6-hydroxy-4-sphingenine in human epidermis. Ceramides of type 1 in which the 1-*O*-hydroxyl group of the sphingoid base is acylated by a very-long-chain fatty acid are also present (1-*O*-acylceramides - illustrated above); these comprise 5% of the total ceramides in the epidermis of mice and humans and comprise as much as 700 molecular species. In all, 15 classes of free ceramides and 3 classes of covalently bound ceramides with up to 1700 distinct molecular species have been identified. Such lipids were first studied in detail in the skin of the pig as a convenient experimental model, but they have been characterized in humans and rats. In addition, several molecular forms of **glucosylceramide**, based on similar ceramide structures, have been characterized in skin, and these are also essential for its proper function.

Depending on the particular layer of the skin (keratinocytes, stratum corneum, *etc*.), the lipid composition can vary. These lipids have an obvious role in the barrier properties of the skin, limiting the loss of water and solutes and at the same time preventing the ingress of harmful substances. As the aliphatic chains in the ceramides and the fatty acids are mainly non-branched long-chain saturated compounds with a high melting point and a small polar head group, the lipid chains are mostly in a solid crystalline or gel state, which exhibits low lateral diffusional properties and low permeability at physiological temperatures. There is a report that the stratum corneum layer of the skin has a water permeability only one-thousandth that of other biomembranes, for example. Natural and synthetic ceramides are now commonly added to cosmetics and other skin care preparations.

Most steps in the biosynthesis of ceramides linked to ω-*O*-acylated fatty acids occur in the endoplasmic reticulum of keratinocytes. First, fatty acid synthesis of very-long-chain (and ultra-long-chain, $\geq C_{26}$) acyl-CoA *de novo* must take place, requiring the chain-elongation enzymes **ELOVL1 and ELOVL4**. Desaturation can occur, and importantly oxidation in the 2 (α) and terminal (ω) positions. The ω-hydroxylation step requires an enzyme of the cytochrome P450 family, designated CYP4F22, of the kind involved in the synthesis of **hydroxy-eicosatetraenoic acids** (HETE). Mutations are a cause of lamellar ichthyosis, and knockout mice deficient in the equivalent enzyme were found to die within 8 hours of birth.



Biosynthesis of very-long chain ω-hydroxy fatty acids

Ceramides are first synthesized by ceramide synthase 3 (CERS3), which has a high specificity for very-long-chain fatty acids ($>C_{26}$) with the incorporation of the ω-hydroxy fatty acid. This is acylated with linoleate by the action of an unusual enzyme related to the phospholipase A family, PNPLA1, which catalyzes esterification by first releasing linoleate from triacylglycerols in the skin while acting as an acyltransferase to link the linoleate directly to the ω-hydroxyl moiety of the ultra-long chain fatty acid. PNPLA1 is unique among phospholipases in that it is involved in the metabolism of sphingolipids rather than glycerophospholipids and catalyzes transacylation rather than hydrolysis. In addition, some linoleate for this purpose is released from triacylglycerols by the action of the adipose tissue lipase aided

by a protein ABHD5. This process is vital for proper skin barrier function and keratinocyte differentiation, as mice with defective triacylglycerol biosynthesis and metabolism, including a deficiency of the acyl-CoA synthase ACSL1, are unable to synthesis ω-*O*-acylceramides and have an impaired skin barrier. Mutations in the human *PNPLA1* gene are believed to be the cause of autosomal recessive disease congenital ichthyosis.

The resulting ceramides are converted to the complex sphingolipids **sphingomyelin** and especially **glucosylceramide**, which are transferred with the aid of ATP-binding cassette (ABC) transporters together with degradative enzymes into the stratum corneum via specific organelles termed 'lamellar bodies.' These organelles must fuse with the apical plasma membrane of the outermost cell layer of the epidermis in order that their contents can be secreted. It is only then that the final step of hydrolysis of the lipid precursors occurs in the extracellular spaces of the stratum corneum, i.e. ceramides are generated from sphingomyelin by the action of acid sphingomyelinase and from glucosylceramides by β-glucocerebrosidase. This mechanism ensures that ceramides, with their potentially harmful biological activities, never accumulate within nucleated cells.

Eventually, ceramides with a terminal ω-hydroxyl group in the fatty acyl moiety are bound covalently to the proteins of the cornified envelope, especially to involucrin. This is illustrated in Figure 21.4.19.

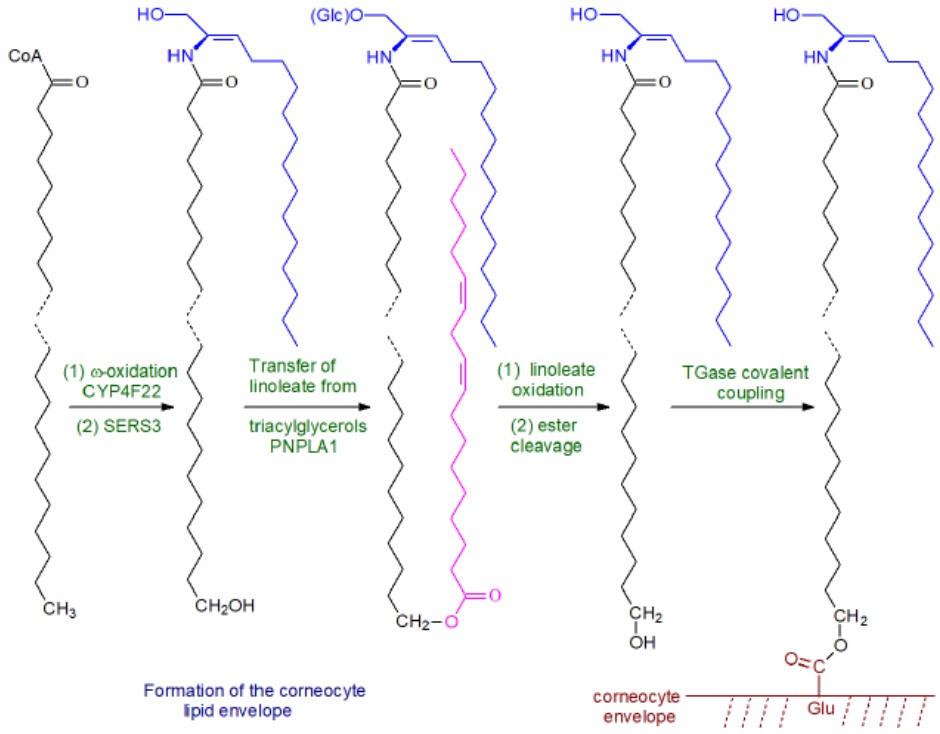
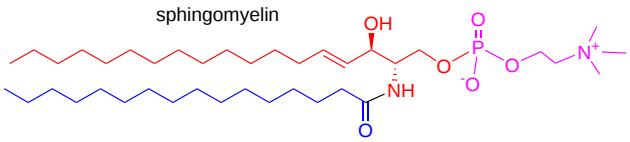


Figure 21.4.19: Formation of the corneocyte lipid envelope

21.4.4: SPHINGOMYELIN AND RELATED SPHINGOPHOSPHOLIPIDS

21.4.4.1: STRUCTURE AND OCCURRENCE OF SPHINGOMYELIN

Sphingomyelin or ceramide 1-phosphocholine consists of a **ceramide** unit with a phosphorylcholine moiety attached to position 1 of the **sphingoid base** component. It is thus the sphingolipid analog of **phosphatidylcholine**, and like that lipid it is zwitterionic. The d18:1/16:0 molecular species is illustrated as an example in Figure 21.4.20.



N-(hexadecanolyl)-sphing-4-enine-1-phosphocholine

Figure 21.4.20: Example of a sphingolipid

Sphingomyelin is primarily of animal origin and is a ubiquitous component of all animal cell membranes, from mammals to nematodes (and in a few protozoa), where it is by far the most abundant sphingolipid. Indeed, it can comprise as much as 50% or more of the lipids in certain tissues, though it is usually lower in concentration than phosphatidylcholine. For example, it makes up about 10% of the lipids of the brain, where it is a key constituent of myelin, but 70% of the phospholipids of the human lens. Like phosphatidylcholine, sphingomyelin

tends to be in greatest concentration in the plasma membrane of cells (up to 20%), and in the endocytic recycling compartment and *trans* Golgi network. It is also abundant in the nucleus where it is the main phospholipid associated with chromatin, but there is very little in the endoplasmic reticulum (2 to 4%) and even less in mitochondria. All the sphingomyelin in human erythrocyte membranes is in the outer leaflet, and ~90% of that in the plasma membrane of nucleated cells is in the outer leaflet. All lipoprotein fractions in plasma contain appreciable amounts of sphingomyelin with a higher proportion in the VLDL/LDL. Sphingomyelin is the single most abundant lipid in erythrocytes of most ruminant animals, where it replaces phosphatidylcholine entirely. In this instance, there is known to be a highly active phospholipase A that breaks down the glycerophospholipids, but not sphingomyelin.

Sphingomyelin is not synthesized in plants or fungi, which produce the sphingophospholipid **ceramide phosphoinositol** and related lipids instead, or in bacteria, and its evolutionary significance is a matter for speculation. However, a number of bacteria and viruses utilize sphingomyelin or its metabolism in their hosts for growth and viability.

Sphingosine is usually the most abundant **long-chain base** constituent, together with sphinganine and C_{20} homologues, although other bases can be present, especially in ruminant animals. In contrast, sphinganine is the major sphingoid base in the sphingomyelin of human lens membranes, linked mainly to 16:0. Typically, the fatty acids are very-long-chain saturated and monounsaturated, including odd-numbered components. In comparison to the glycosphingolipids, 2-hydroxy acids are only rarely detected and then in small amounts, but they are found in testes, spermatozoa, kidney and skin sphingomyelin, for example. The absolute proportions of each fatty acid and sphingoid base can vary markedly between tissues and species, and some of the variability in compositions can be seen from the data in Table 21.4.3 and Table 21.4.4 below.

Table 21.4.3: Fatty acid compositions of sphingomyelin (wt % of the total) in some animal tissues.

Source	Fatty acids									
	16:0	18:0	18:1	20:0	22:0	22:1	23:0	23:1	24:0	24:1
Egg	66	10	1	4	6	1	2	-	5	3
Bovine brain	3	42	-	6	7	3	3	3	6	27
Cow's milk	14	3	1	1	22	-	32	-	19	5

Adapted from Ramstedt, B. *et al*. Analysis of natural and synthetic sphingomyelins using high-performance thin-layer chromatography. *Eur. J. Biochem*., **266**, 997-1002 (1999); DOI.

Table 21.4.4: Long-chain base compositions of sphingomyelin (wt % of the total) in some animal tissues.

Source	Sphingoid base						
	d16:0*	d17:0	d17:1	d17:1-methyl	d18:0	d18:1	d19:0
Egg					7	93	
Bovine brain					19	81	
Cow's milk	9	15	8	11	10	44	3

Also from Ramstedt, B. *et al. Eur. J. Biochem*., **266**, 997-1002 (1999); DOI.
* d = dihydroxy base

Palmitic acid (16:0) is the most common fatty acid component of sphingomyelin in peripheral cells of mammals, while stearic acid (18:0) is more abundant in neural tissue, but this only hints at the potential complexity as there can be variability within tissues. For example, about 60% of the fatty acids of the sphingomyelin of the grey matter of the human brain consist of stearic acid (18:0), while lignoceric (24:0) and nervonic (24:1) acids make up 60% of the corresponding lipid of white matter, although this is dependent on the stage of development. During the first two years of life, the 18:0 concentration in sphingomyelin of white matter decreases from 82% to 33%, while the proportions of 24:0 and 24:1 increase. This pronounced shift from long-chain to very-long-chain sphingomyelins is not observed in the cerebral cortex. Approximately 100 molecular species of sphingomyelin have been detected in human plasma. Although polyunsaturated fatty acids such as arachidonic acid are rarely present, they have sometimes been mistakenly identified in the literature. Exceptions are the sphingomyelins of testes and spermatozoa, which contain very-long-chain polyunsaturated fatty acids (up to 34 carbon atoms), the major components being 28:4(*n*-6) and 30:5(*n*-6) with a proportion having hydroxyl groups in position 2.

21.4.4.2: BIOSYNTHESIS, METABOLISM AND FUNCTION OF SPHINGOMYELIN

The biosynthesis of sphingomyelin is distinct from that of **phosphatidylcholine** and indeed depends upon it, as it involves the transfer of phosphorylcholine from phosphatidylcholine to ceramide, synthesized in the endoplasmic reticulum, with the liberation of 1,2-diacyl-*sn*-glycerols. as illustrated in Figure 21.4.21.

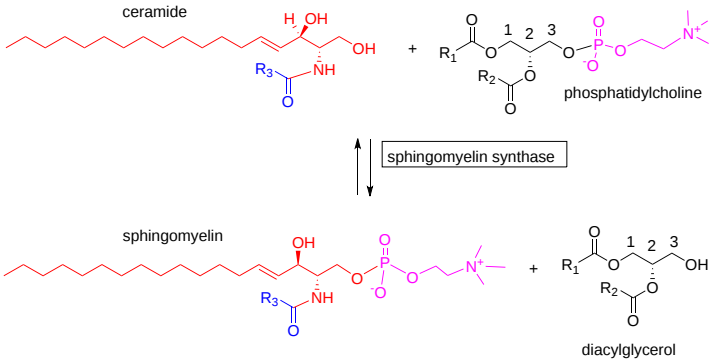
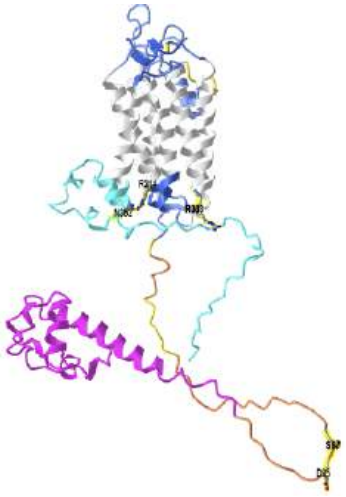


Figure 21.4.21: Synthesis of sphingomyelin

The reaction is catalyzed by a ceramide choline-phosphotransferase (sphingomyelin synthase or SMS) and takes place primarily on the luminal side of the *trans*-Golgi but also in the plasma membrane, with two related enzymes each with six transmembrane domains and their N- and C-termini facing the cytosol, i.e., SMS1 and SMS2. Both enzymes are present in the Golgi, but only SMS2 is in the plasma membrane (facing the extra-cellular space in this instance) and may be necessary for the formation of raft domains (see below). SMS2 is also present in the membranes of nuclei from rat liver cells. It is noteworthy that in the absence of ceramide, both SMS1 and 2 have phospholipase C activity, and so may regulate the steady-state levels of phosphatidylcholine and diacylglycerols as well as that of sphingomyelin. The reaction does not use free phosphorylcholine or CDP-choline as a donor.

Figure 21.4.22 shows an interactive iCn3D model of the AlphaFold model of human Golgi membrane phosphatidylcholine:ceramide cholinephosphotransferase 1, also called sphingomyelin synthase 1 (Q86VZ5).



NCBI iCn3D Figure 21.4.22: . Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...v6s5wHRgPXKJC6

The gray helices are the transmembrane helices. A cytoplasmic protein:protein interaction domain called SAM (sterile alpha motif) is shown in magenta. The other cytoplasmic C-terminal domain is shown in cyan Obviously much of the cytoplasmic domain is disordered in this computational structure. Side chains involved in binding phosphatidylcholine are shown as sticks colored CPK. Note that two, D95 and S97 are located in a disordered section in this model but would close in the actual active site in the actual structure.

It has been proposed that Asp101 (95 in the AlphaFold structure) deprotonates Arg220 (214 in the model), which then acts as a nucleophile which attacks the phosphate group of phosphatidylcholine. Given the very high pKa of arginine, this mechanism, if true, is somewhat unique. Phosphocholine is linked to ceramide to produce sphingomyelin. Three key and extremely conserved amino acids in the active site are Asp101, Arg220 and Asn358.

A specific ceramide transport molecule (CERT) is important to the reaction with SMS1 in that it transfers ceramide from the cytosolic surface of the endoplasmic reticulum to the *trans*-Golgi in an ATP-dependent and non-vesicular manner. Much of the sphingomyelin produced in the Golgi is then delivered to the apical plasma membrane by a vesicular transport mechanism. Sphingomyelin synthesis is regulated in part by phosphatidylinositide metabolism and is connected to sterol homeostasis through the oxysterol binding protein (OSBP).

SMS2 in the plasma membrane is not dependent on CERT-mediated ceramide delivery, but is believed to convert ceramide produced locally by a sphingomyelinase back to sphingomyelin; this may be an important protective mechanism for the cell. The location of the enzymes explains the enrichment of sphingomyelin in specific membranes and the sidedness, i.e., the luminal *trans*-Golgi and the outer leaflet of the plasma membrane, while ceramide reaching the *cis*-Golgi is utilized for the synthesis of **glucosylceramide**. As the nature of the molecular

species of sphingomyelins produced differs appreciably from that of the ceramide precursors, the sphingomyelin synthases must have considerable substrate specificity. The reaction can be reversible, using sphingomyelin to generate ceramide for specific signaling functions. It is evident that sphingomyelin biosynthesis forms a link between the sphingolipid signaling pathway (pro-apoptotic - see below) and that of glycerolipids via the mitogenic **diacylglycerol** by-products. Although the importance of this production relative to that via **phosphatidylinositol** is not known, it is possible that it is significant locally at the external leaflet of the plasma membrane.

An alternative pathway of sphingomyelin synthesis has been demonstrated in the endoplasmic reticulum in which ceramide is first converted to ceramide phosphoethanolamine via transfer of the head group from phosphatidylethanolamine, followed by stepwise methylation of the ethanolamine moiety. However, the physiological significance of this pathway has yet to be established.

It was long thought that the only function of sphingomyelin was to serve as a substitute for phosphatidylcholine as a building block of membranes, i.e., by forming a stable and chemically resistant outer leaflet of the plasma membrane lipid bilayer. For example, it may limit the ingress of oxygen and thence oxidation of adjacent unsaturated acyl chains. While this is certainly one of its functions, the apparent similarity between phosphatidylcholine and sphingomyelin is superficial, and there are great differences in the hydrogen bonding capacities and physical properties of the two lipids. For example, sphingomyelin has an amide bond at position 2 and a hydroxyl at position 3 of the sphingoid base, both of which can participate in hydrogen bonding, while the *trans* double bond also appears to assist intermolecular interactions in membranes. Indeed, the first five carbon atoms of the sphingoid base in sphingolipids constitute a key feature that has been termed the ‘sphingoid motif’, which facilitates a relatively large number of molecular interactions with other membrane lipids, via hydrogen-bonding, charge-pairing, hydrophobic and van der Waals forces. With phosphatidylcholine, in contrast, the two ester carbonyl groups can act only as hydrogen acceptors. The degree of unsaturation of the alkyl moieties in each lipid is very different, and this gives them dissimilar packing properties in membranes.

It is now recognized that sphingomyelin and other sphingolipids have a strong tendency to interact with proteins and cholesterol, often via strong van der Waals interactions and hydrogen bonding, to form transient nano-domains in membranes known as '**rafts**' and on the surface of lipoprotein particles. Initially, there was a view that saturated sphingomyelin formed a liquid-ordered phase with cholesterol or a gel phase with saturated ceramides to lead to lateral segregation within the membrane, and that sphingomyelin and cholesterol metabolism were closely integrated, even that the sphingomyelin concentration might control the distribution of cholesterol in cells. On the other hand, the understanding of the mechanism of raft formation in membranes has changed substantially in recent years, and while an interaction with cholesterol is certainly important, it may not be the major factor *in vivo*. Ceramide can displace cholesterol from its association with sphingomyelin, when formed in membranes by hydrolysis of the latter.

Other functions: Sphingomyelin *per se* is generally considered to be a relatively inert molecule, although modern molecular biology methods are uncovering potential regulatory functions via interactions with particular proteins. For example, it has been shown to inhibit the activity of **phospholipase $A_2\alpha$**, a key enzyme in eicosanoid production. Sphingomyelin in the plasma membrane may be essential for the internalization of transferrin and thence of iron into cells, and it appears to be required for the activity of a number of membrane-bound proteins, including those of certain ion channels and receptors. As the most abundant sphingolipid in the nucleus, it is intimately involved in chromatin assembly and dynamics as well as being an integral component of the nuclear matrix. A single molecular species of sphingomyelin with a C_{18} acyl chain binds specifically to a coat protein designated 'p24' to enable it to form membrane vesicles. In addition, sphingomyelin is selectively recognized and acts as a receptor for the actinoporins, which are pore-forming toxins produced by sea anemones.

There is a specific binding site for sphingomyelin on the amyloid beta-peptide (Aβ) in brain, and there is evidence from studies *in vitro* that this may promote the aggregation of these proteins in Alzheimer's disease. In turn, this leads to depletion of brain sphingomyelin by activation of acid sphingomyelinase with disruption of many protein–lipid interactions and thence of downstream signaling pathways. In contrast, the ganglioside GM1 may have a protective role towards Aβ aggregation

As well as its role in membranes, it serves as a precursor for ceramides, long-chain bases, sphingosine-1-phosphate, and many other biologically important sphingolipids, as part of the 'sphingomyelin cycle' (also termed the ‘sphingolipid’ or ‘ceramide’ cycles depending on the context). Some of these metabolites are intra- and inter-cellular messengers, and others are essential membrane constituents. The sphingomyelin cycle extends to other sphingolipids via the action of sphingomyelinases and enzymes such as glycosylhydrolases and glycosyltransferases in cells to produce innumerable new oligoglycosylceramides. It can also give rise to ***sn*-1,2-diacylglycerols**, which are central to many metabolic and signaling pathways. These molecular relationships are illustrated only briefly in Figure 21.4.23.

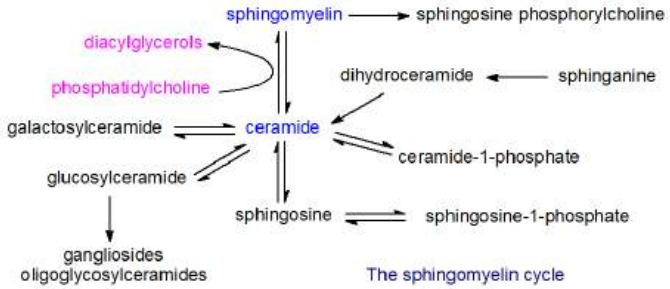


Figure 21.4.23: The sphingomyelin cycle

In particular, sphingomyelin is a major source of **ceramides** in most cellular organelles, including the nucleus and even mitochondria, via the action of sphingomyelinases (see next section), and in addition to being a source of other sphingolipids these are required to trigger apoptosis and other metabolic changes. As ceramides do not mix well with glycerophospholipids and cholesterol, this conversion results in the formation of new membrane domains enriched in ceramide that exclude cholesterol and so differ in composition from other sphingolipid rafts. This has profound effects on membrane function, especially of the plasma membrane, in that different proteins may be recruited or excluded depending on their relative affinities for cholesterol and ceramides. It may also influence disease states such as cancer.

Chlamydiae, widespread bacterial pathogens, acquire sphingomyelin from the Golgi apparatus and plasma membrane of their hosts and this is necessary for the viability and growth of the organisms. Other pathogenic bacteria, notably *Pseudomonas aeruginosa* and *Neisseria gonorrhoeae*, can hijack sphingomyelin catabolic enzymes with deleterious effects upon the host. Likewise, human immunodeficiency virus (HIV) and the hepatitis C virus utilize host sphingomyelin for their own nefarious purposes.

Nutrition: Although there is no known nutritional requirement for sphingomyelin and other sphingolipids, they are a component of any diet containing egg, meat or dairy products. Thus, it has been estimated that *per capita* sphingolipid consumption in the United States, for example, is of the order of 0.3-0.4 g/d. As sphingolipids constitute an appreciable proportion of the polar lipid constituents of milk, they may be significant if minor nutrients for infants and beneficial effects upon their development have been claimed. From animal experiments, there is evidence that dietary sphingolipids can reduce the intestinal absorption of cholesterol and other lipids, leading to reductions in serum lipid concentrations. Feeding sphingolipids inhibits colon carcinogenesis and may alleviate some of the symptoms of inflammatory bowel disease. 2-Hydroxyoleic acid suppresses the growth and induces autophagy in cancer cells by stimulating the synthesis of sphingomyelin and increasing the amount of this lipid in the plasma membrane. On the other hand, plasma sphingomyelin levels are considered to be an independent risk factor for atherosclerosis, possibly as a result of its ability to retain cholesterol in cells and the arterial wall with consequent diminished reverse cholesterol transfer via HDL.

21.4.4.3: SPHINGOMYELIN CATABOLISM

In contrast to the glycerolipids, dietary sphingolipids are not hydrolyzed by pancreatic enzymes only. Rather, most of the sphingomyelin in the diet is hydrolyzed in the brush border of the intestines by an **alkaline sphingomyelinase** (at a pH of 8.5–9 optimally) to ceramide and thence by a neutral ceramidase to free fatty acids and sphingosine. Some of this enzyme is also present in liver from which it is secreted in bile into the intestinal lumen where it can hydrolyze sphingomyelin and other phospholipids with the aid of bile salts. The sphingosine released at the brush border is absorbed, some is re-*N*-acylated to form ceramides, and the remainder is converted via sphingosine-1-phosphate to palmitic acid, which is esterified into the triacylglycerol component of chylomicrons. In the process, some of these sphingolipid intermediates may have signaling functions and anti-inflammatory properties in intestinal cells. The alkaline sphingomyelinase is unusual in that is very different in its structure and other properties from intracellular enzymes with a related function; it is part of the (ecto)nucleotidepyrophosphatase-phosphodiesterase protein family (NPP) that includes **autotaxin**. The enzyme is believed to have a role in the production of sphingolipid metabolites within the intestines and colon especially, which may influence a number of disease states. For example, it appears to inhibit colon cancer by generating ceramides. In addition, alkaline sphingomyelinase has phospholipase C activity towards the pro-inflammatory metabolite **platelet-activating factor** and towards lysophosphatidylcholine with potentially further beneficial effects. By reducing the level of endogenous sphingomyelin and increasing that of ceramides in the membranes of intestinal cells, it is believed to reduce the uptake of dietary cholesterol.

Catabolism in other tissues: The key enzymes for the degradation of sphingomyelin to ceramides in most tissues are also sphingomyelinases (phosphodiesterases), as shown in Figure 21.4.24.

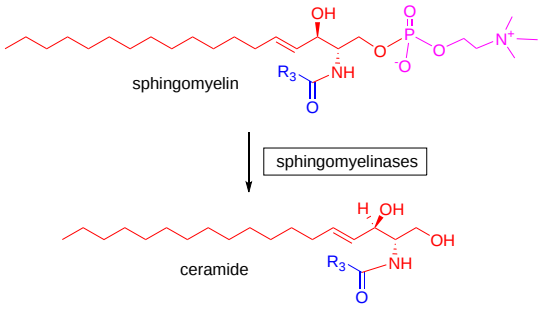


Figure 21.4.24: Sphingomyelin catabolism

These are similar in function to phospholipase C and generate ceramides with their innumerable and important signaling properties as the main product. There are many such enzymes with different pH optima and metal ion requirements that operate in different regions of the cell with potentially distinct biochemical roles. Thus, there is an acid sphingomyelinase in the endo-lysosomes, and different neutral sphingomyelinases in the plasma membrane, endoplasmic reticulum, Golgi, and mitochondria in addition to the alkaline sphingomyelinase in the intestines. It should not be forgotten that the other product of the reaction is phosphocholine, which has importance as a nutrient. Bacterial sphingomyelinases are known to lyse red blood cells, although intriguingly, there is a sphingomyelinase in the bacterium *Pseudomonas aeruginosa* that can also act as a sphingomyelin synthase *in vitro* at least.

The lysosomal **acid sphingomyelinase** (pH optimum ca. 5) is expressed ubiquitously and has a key housekeeping role in maintaining normal membrane turnover and remodeling of the sphingolipid constituents, especially those of lipoproteins. While other lysosomal sphingolipid hydrolases require a **saposin** activator protein for full activity, the acid sphingomyelinase incorporates a built-in N-terminal saposin domain so does not require an external activator. Under resting conditions, acid sphingomyelinase is stored inside lysosomes, but upon stimulation, it undergoes vesicular transport to the plasma membrane where it docks with a specific protein and is exposed on the outer leaflet. It then generates ceramide by hydrolysis of sphingomyelin and initiates the train of events that leads to apoptosis. There are reports that acid sphingomyelinase, by acting at the plasma membrane to produce ceramides, regulates the localization and trafficking of palmitoylated proteins from the Golgi, and it may also facilitate bacteria-host interactions. Experiments *in vitro* have demonstrated that the enzyme can be considered as a phospholipase C that is active against a wide range of phospholipids, including **ceramide-1-phosphate** and the unique lysosomal phospholipid **bis(monoacylglycero)phosphate**.

There is a related secreted acid sphingomyelinase (Zn^{2+}-dependent), which can be transported to the outer membrane of the cell and is especially important in endothelial cells of the human coronary artery. This enzyme is produced by the same gene but differs from the lysosomal enzyme as it requires Zn^{2+} ions for activation and has a different glycosylation pattern. It can also operate at neutral pH and has multiple functions in that it is involved in many aspects of cellular signaling as well as in membrane sphingomyelin turnover. By acting at the plasma membrane to produce ceramides, it is believed to regulate the trafficking of **palmitoylated proteins** from the Golgi to their new location.

Neutral sphingomyelinases (pH optima 7.4), of which four quite distinct enzymes are known, are located in membranes of the endoplasmic reticulum, Golgi, and plasma membrane with one in mitochondria (MA-NSM), where they have signaling functions by generating ceramides and thence other biologically active sphingolipids. Human NSM-1 has 423 amino acid residues and a molecular weight of 47.6 kDa; it has two putative transmembrane domains in the C-terminus and resides mainly in the nucleus and endoplasmic reticulum. It has a broad specificity for choline phospholipids, but it is most active with sphingomyelin and may not have a significant role in cellular signaling. In contrast, NSM-2 which is located in the Golgi apparatus and plasma membrane is activated by phosphatidylserine and is important for ceramide signaling. It is especially important in brain and nervous tissue, where it is required for the secretion of hypothalamic-sssreleasing hormones, although it is relevant to many cellular functions and physiological processes in most other tissues. Dysregulation of NMS-2 is reported to be a factor in many inflammation-related pathologies. Neutral sphingomyelinases-3 is found mainly in the plasma membrane of bone and cartilage, where it is vital for the process of mineralization; it is also important in striated and cardiac muscle. Little seems to be known of the function of the mitochondrial enzyme. Losses, mutation, and poor expression of the gene encoding neutral sphingomyelinase have been observed in several cancers, but exposure to ionizing irradiation led to rapid hydrolysis of sphingomyelin to ceramide by this enzyme, and thence to cancer cell death.

A diverse range of factors activates the enzymes, including chemotherapeutic agents, tumor necrosis factor-alpha, 1,25-dihydroxy-vitamin D_3, endotoxin, gamma-interferon, interleukins, nerve growth factor, and most conditions known to induce cellular stress, especially in relation to inflammation. As they utilize by far the most abundant sphingolipid in animal tissues to generate ceramides and other sphingolipid metabolites that have important signaling functions, sphingomyelinases are believed to function as regulators of signaling mechanisms, especially in the nucleus of the cell. Thus, they have a much wider metabolic role than simply catabolism of sphingomyelin.

The **type A and B forms of Niemann-Pick disease** are lysosomal lipid storage disorders that are a consequence of a deficiency of acid sphingomyelinase with a resulting accumulation of sphingomyelin and smaller amounts of other sphingolipids, including gangliosides, in cells and tissues and especially in the monocyte/macrophage system to form the so-called “foam cells” that characterize the disease. A

consequent lack of ceramide production may be involved in the pathology of the disease. Increasing sphingomyelin levels in turn result in elevated cholesterol concentrations. It is noteworthy that membranes containing ceramides have a much lower binding capacity for cholesterol, so sphingomyelin degradation may play a part in cholesterol homeostasis. **Type C Niemann-Pick** disease differs from the A and B forms and is caused by defects in two distinct cholesterol-binding proteins (NPC1 and NPC2).

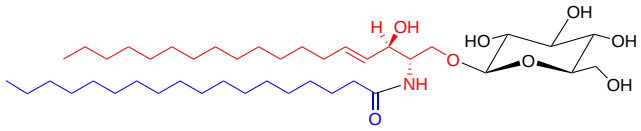
21.4.5: GLUCOSYL- AND GALACTOSYLCERAMIDES (CEREBROSIDES)

There are two natural monoglycosylceramides of special importance in animals, i.e., glucosylceramide and galactosylceramide. Both have biological functions in their own right, but especially as structural components of membranes, as in the brain, for example, where galactosylceramide is required for the maintenance of the structure and stability of myelin and the differentiation of oligodendrocytes. Glucosylceramide is a vital component of all cell types, and is most abundant in human skin; it is the key intermediate in the biosynthesis of lactosylceramide and thence of complex oligoglycosphingolipids, including gangliosides. This monoglycosylceramide is also a major component of the membranes of plants and fungi. Although the two lipids have very similar structures in that D-galactose is an epimer of D-glucose and they differ only in the configuration at C4, they have very different biological properties. A few other monoglycosylceramides are produced in nature, for example by some bacteria of the order *Sphingomonadales* of α-proteobacteria.

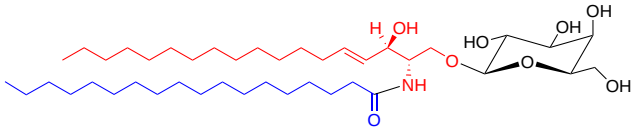
21.4.5.1: STRUCTURE AND OCCURRENCE

β-D-Galactosylceramide (Galβ1-1'Cer) is the principal glycosphingolipid in brain tissue, hence the trivial name "cerebroside", which was first conferred on it in 1874, although it was much later before it was properly characterized. In fact, galactosylceramides are found in all nervous tissues and indeed at low levels in all organs, but in they brain they can amount to 2% of the dry weight of grey matter and 12% of white matter or 23% of myelin lipids, where they insulate the axons of neuronal cells and constitute a substantial component of the extended plasma membrane of oligodendrocytes. It is also present in some fungal species. While galactosylceramide can be sulfated to form a sulfatide or sialylated to form ganglioside GM_4, only a small proportion is subjected to further galactosylation to form Gal_2Cer as the precursor for the limited gala-series of oligoglycosphingolipids.

β-D-Glucosylceramide (Glcβ1-1'Cer), with the trivial name "glucocerebroside", is a major constituent of skin lipids, where it is essential for the maintenance of the water permeability barrier of the skin. Otherwise, it is most abundant in animal tissues such as the spleen and erythrocytes as well as in nervous tissues, especially in the neurons if at low levels, and it is also found in plants. Higher than normal concentrations of this glycosphingolipid have been reported for the apical plasma membrane domain of epithelial cells from the intestines (especially the absorptive villous cells) and urinary bladder. The d18:1/16:0 molecular species of the two lipids are illustrated in Figure 21.4.25.



glucosylceramide



b- galactosylceramide

Figure 21.4.25: gluco- and galactoceramides (cerebrosides)

However, of equal or greater importance to the natural occurrence of glucosylceramide *per se* is its role as the biosynthetic precursor of lactosylceramide in animals, and thence of most of the complex **neutral oligoglycolipids** and **gangliosides**. In contrast, glucosylceramide is the end-product of the biosynthetic pathway in plants and fungi.

Interestingly, the proportion of galactosylceramides relative to glucosylceramides in myelin glycolipids increases greatly in the ascending phylogenic tree, and the ratio of hydroxy- to nonhydroxy fatty acids in cerebrosides increases with the complexity of the central nervous system. There is also an intriguing sex difference in the kidney, where it has been shown that galactosylceramide rather than glucosylceramide occurs in male mice only (or androgen-treated adult females). Only glucosylceramide is present in the nerves of the most primitive animals (protostomes).

In the brain, the galactosylceramides are enriched in very-long-chain fatty acids (C_{22}–C_{26}). The fatty acid and long-chain base compositions of cerebrosides from the intestines of the Japanese quail are listed in Table 21.4.5 for illustrative purposes. The fatty acid components resemble those of other sphingolipids, although the percentage of 2-hydroxy acids is higher than that in sphingomyelin, for example. They are exclusively saturated in this instance, though a small proportion of monoenoic components may also be found in other tissues.

Glucosylceramides tend to contain mainly non-hydroxylated fatty acids that are of relatively shorter chain length. The proportion of trihydroxy bases listed is perhaps higher than in other many other tissues or species studied, probably reflecting the diet. Usually, sphingosine is the main long-chain base in cerebrosides of animal tissues.

Table 21.4.5: **Composition of fatty acids and long-chain bases (wt % of the total) in cerebrosides of intestines from the Japanese quail.***

Long-chain bases		Fatty acids	Non-hydroxy acids	2-Hydroxy acids
Type	**%**		**%**	**%**
t18:0	43	16:0	5	6
d18:0	9	18:0	3	trace
d18:1	27	20:0	2	4
t20:0	6	21:0	trace	2
d20:0	3	22:0	4	43
d20:1	11	23:0	1	13
		24:0	3	12

* The cerebrosides comprised 81% galactosylceramide and 19% glucosylceramide.
From Hirabayashi, Y. *et al*., *Lipids*, **21**, 710-714 (1986); DOI.

Plants: Glucosylceramide is the only glycosphingolipid common to plants, fungi, and animals. It has often been described incorrectly as the main sphingolipid in plants, but this has been because the more polar complex **glycosylinositol phosphoceramides** are not easily extracted and until relatively recently were missed in conventional analyses. Nonetheless, glucosylceramide is abundant in photosynthetic tissues and constitutes approximately a third of the total sphingolipids, where the main long-chain bases are C_{18} 4,8-diunsaturated (*Z*/*Z* and *E*/*Z*) (not sphingosine as illustrated above); it is a major component of the outer layer of the plasma membrane and is also enriched in the late endosomes and plant tonoplast. Small amounts of monoglycosylceramides containing a β-D-mannopyranosyl unit may be present in non-photosynthetic tissues, but galactosylceramides have not been found in plants. Glucosylceramide is a common component of the lipids of yeast and other fungi, including most fungal pathogens. However, it does not occur in the yeast *Saccharomyces cerevisiae*, which is widely used as an experimental model, although trace levels of galactosylceramide have been detected.

The fatty acid and long-chain base compositions of glucosylceramides from two plant sources are listed in Table 21.4.6. Perhaps surprisingly, the fatty acid components are not very different in nature from those in animal tissues, comprising mainly longer-chain saturated and monoenoic acids, with a high proportion being saturated and having a hydroxyl group in position 2. In the examples selected for the table here, both di- and tri-hydroxy long-chain bases were found, mainly diunsaturated (*Z*/*Z* and *E*/*Z*) and almost entirely C_{18} in chain length. Much higher concentrations of glucosylceramides are found in pollen than in leaves, with substantial compositional differences. For example, the long-chain bases in Arabidopsis leaves consist mainly of t18:1, with relatively little d18:1, t18:0 and d18:0 (with 16:0, 24:0 and 24:1 hydroxy fatty acids mainly), but no d18:2 base although this is 50% of those in pollen. While saturated 2-hydroxy acids predominate in most plants, some cereal glucosylceramides contain high proportions of mono-unsaturated very-long-chain fatty acids of the *n*-9 family. Glucosylceramides from algae tend to resemble those from higher plants, although some novel structures have been reported from microalgae.

Table 21.4.6: **Composition of fatty acids and long-chain bases (wt % of the total) in glucosylceramides of seeds from scarlet runner beans and kidney beans.**

	Fatty acids[a]			**Long-chain bases[b]**	
Type	**Runner beans**	**Kidney beans**		**Runner beans**	**Kidney beans**
	%	**%**		**%**	**%**
16:0	4	5	t18:0	trace	trace
Other non-hydroxy	1	2	t18:1-8*t*	13	11
14:0-OH	1	1	t18:1-8*c*	10	9
15:0-OH	1	1	d18:0	trace	trace
16:0-OH	58	58	d18:1-8*c*/*t*	1	3
18:0-OH	trace	trace	d18:1-4*t*	trace	trace
20:0-OH	trace	trace	d18:2-4*t*,8*t*	45	60
22:0-OH	7	6	d18:2-4*t*,8*c*	31	17
23:0-OH	2	1			
24:0-OH	23	23			
25:0-OH	1	1			
26:0-OH	1	1			

From Kojima, M. *et al*., *J. Agric. Food. Chem.*, **39**, 1709-1714 (1991); DOI, but see also Yamashita, S. *et al*. for further data: DOI
[a] including 2-hydroxy acids; [b] di- and tri-hydroxy bases with *cis* or *trans* double bonds in the positions indicated.

21.4.5.2: BIOSYNTHESIS

Ceramides synthesized both *de novo* and by catabolism of sphingomyelin are used for the biosynthesis of monoglycosylceramides in animal tissues. The biosynthetic mechanism resembles that for **glycosyldiacylglycerols**, i.e., there is a direct transfer of the carbohydrate moiety from a sugar-nucleotide, e.g. uridine 5-diphosphate(UDP)-galactose, UDP-glucose, etc, to a ceramide unit synthesized in the endoplasmic reticulum. This is illustrated in Figure 21.4.26.

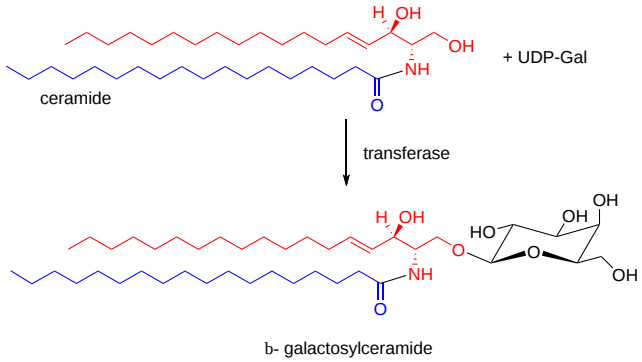


Figure 21.4.25: Synthesis of galactosylceramide

During the transfer, which is catalyzed by specific glycosyl-transferases, inversion of the glycosidic bond occurs from the *alpha* to *beta* configuration. Synthesis of β-D-galactosylceramide takes place on the lumenal surface of the endoplasmic reticulum, although it has free access to the cytosolic surface by an energy-independent flip-flop process. Expression of the UDP-galactose:ceramide galactosyl transferase (galactosylceramide synthase) is restricted to oligodendrocytes, Schwann cells, kidneys, and testes. Prior to **sulfation**, galactosylceramide is transported to the *trans*-Golgi compartment.

In contrast, after the transfer of the precursor **ceramides** from the endoplasmic reticulum to the cytosolic side of the early Golgi membranes with the aid of the CERT protein, glucosylceramide is produced by a glucosylceramide synthase present in this membrane (with the possible exception of neuronal tissues). If it is to be converted to more complex **oligoglycosylceramides**, this must be translocated to the luminal leaflet of the *trans*-Golgi membranes, a process that occurs both by vesicular and by non-vesicular transport. The latter is mediated by a conserved clade of integral membrane proteins, i.e., phospholipid flippases (P4-ATPases) designated ATP10A and ATP10D, together with the four phosphate adapter protein-2 (FAPP2) and glycolipid transfer protein (GLTP) in humans with related enzymes in fungi, which utilize the energy from ATP catalysis to translocate lipids across cellular membranes. The human enzymes are entirely specific for glucosylceramide and not galactosylceramide. Indeed, the galactosyl- and glycosylceramide synthases have no significant sequence homology, indicating different evolutionary origins.

For their functions in protein interactions and signaling, both galactosyl- and glucosylceramide must be transported to and then across the plasma membrane. Some glucosylceramide is carried by lipoproteins (VLDL, LDL, and HDL) in the circulation and presumably requires active transport for absorption and distribution across the membranes of target tissues.

In plants, glucosylceramides are also formed by an evolutionarily conserved glucosylceramide synthase involving UDP-glucose in the endoplasmic reticulum, although an alternative mechanism has been described that utilizes sterol glucoside as the immediate glucose donor to ceramide. There is also evidence for a requirement for ceramides containing Δ^4 *trans*-double bonds for synthesis of glucosylceramides but not other sphingolipids in some plant and fungal tissues. However, there is a distinct ceramide synthase in the yeast *Pichia pastoris*, which produces ceramides of defined composition exclusively for the production of glucosylceramides. A separate ceramide synthase with different specificities produces the ceramide precursors for **ceramide phosphorylinositol**, which contains only phytosphingosine as the long-chain base. In fungi, glucosylceramide synthases have been characterized, but a galactosylceramide synthase has yet to be identified. Enzymes responsible for the biosynthesis of glucuronosylceramide and α-galactosylceramide in some bacterial species have been characterized.

21.4.5.3: FUNCTION

Galactosylceramides: A remarkable property of cerebrosides is that their 'melting point' is well above physiological body temperature, so that glycolipids have a para-crystalline structure at this temperature. Each cerebroside molecule may form up to eight inter- or intramolecular hydrogen bonds by lateral interaction between the polar hydrogens of the sugar and the hydroxy and amide groups of the sphingosine base of the ceramide moiety, and this dense network of hydrogen bonds is believed to contribute to the high transition temperature and the compact alignment of cerebrosides in membranes. As with sphingomyelin, monoglycosylceramides tend to be concentrated in the outer leaflet of the plasma membrane together with cholesterol and thence in myelin in the specific membrane domains termed '**rafts**'. Indeed, the latter appear to facilitate segregation to a greater extent than sphingomyelin via the combination of hydrogen bonds and hydrophobic interactions, and these forces are also of great importance for binding to the wide range of proteins, including enzymes and receptors, which are found in raft domains.

Galactosylceramide is essential to myelin structure and function and it is involved in oligodendrocytes differentiation. While molecular species with 2'-hydroxy fatty acid constituents are not essential for myelin formation, they are critical for the long-term stability of myelin, presumably because increased hydrogen bonding with neighboring lipids in membranes stabilizes the phase structure. Galactosylceramide is important as a precursor of **3'-sulfo-galactosylceramide**, which is also essential to brain development in addition to numerous functions in other tissues. By interacting with sulfatide located in the membrane of opposing layers in the myelin sheath by carbohydrate-carbohydrate interaction, it forms what is known as a glycosynapse, which provides a necessary contribution to the long-term stability of myelin.

Glucosylceramides: Glucosylceramides have similar physical properties in membranes to the galactose analog, and they are also concentrated in raft domains in the outer leaflet of the plasma membrane. As mentioned briefly above, they are the primary precursor for most of the more complex oligoglycosphingolipids in animal tissues, especially in brain, where synthesis is vital for the production of most neuronal oligoglycosphingolipids, while glucosylceramide *per se* is essential for axonal growth. They are major constituents of skin lipids, where they are essential for lamellar body formation in the stratum corneum and to maintain the water permeability barrier of the skin. In addition, the epidermal glucosylceramides (together with sphingomyelin) are the source of the unusual complex **ceramides** that are found in the stratum corneum including those with terminal hydroxyl groups and estolide-linked fatty acids. Some of the glucosylceramide in the skin is linked covalently to proteins via terminal hydroxyl groups, presumably to strengthen the epidermal barrier.

Much of the evidence for the function of glycosylceramides in animals has been derived from cell lines in which synthesis of the lipid has been suppressed by various means *in vitro*. It appears that glucosylceramide is not essential for the viability of certain cell lines in culture, but disruption of the global synthase gene in mice results in the death of embryos. It is essential for the survival of cancer cells, and deletion from other cell types can lead to abnormalities. In addition to being an intermediate in the biosynthesis of more complex glycosphingolipids and its role in the permeability barrier of the skin (discussed above), glucosylceramide is believed to be required for intracellular membrane transport, cell proliferation, and survival, and for various functions in the immune system. In contrast, there are indications that it may have adverse implications for various disease states. For example, over-expression of glucosylceramide synthase in cancer cells has been linked to tumour progression with a reduction in ceramide concentration, resulting in increased resistance to chemotherapy. The lipid has also been associated with drug resistance in a wider context. In the nematode *Caenorhabditis elegans*, glucosylceramide containing the fatty acid 22:0 is reported to be a longevity metabolite that functions through the membrane localization of clathrin, a protein that regulates membrane budding.

In Arabidopsis, glucosylceramides are critical for cell differentiation and organogenesis, but not necessarily for the viability of cells. It has been proposed that glycosphingolipids could impose positive curvature to membranes, thereby facilitating vesicle fusion. There is evidence that glycosylceramides (but not glycosyldiacylglycerols) together with sterols are located in 'rafts' in plant membranes in an analogous manner to sphingolipids in animal tissues, and that they are associated with specific proteins. Correlative studies suggest that glucosylceramides help the plasma membrane in plants to withstand stresses brought about by cold and drought. For example, glycosylceramides containing 2-hydroxy monounsaturated very-long-chain fatty acids and long-chain bases with 4-*cis* double bonds appear to be present in higher concentrations in plants that are more tolerant of chilling and freezing. While fungal glucosylceramides with a 9-methyl group within the sphingosine backbone elicit defence responses in rice, cerebrosides with double bonds in positions 4 and/or 8 of the long-chain base appear to be involved in the defense of some plant species against fungal attack.

Less is known of the function of glucosylceramides in fungi, although they are certainly major constituents of the plasma membrane and cell wall. They are believed to be involved in such processes as cell wall assembly, cell division and differentiation, and signaling. The presence of the methyl branch in the long-chain base is essential for cell division and alkali tolerance. In the case of fungal pathogens, glucosylceramides are recognized by the host immune system and regulate virulence, often after export into the external environment as extracellular vesicles. In contrast to animals, ceramide monohexosides are not precursors for oligoglycosylceramides in fungi. Some molecular species of this lipid from plants (a Δ^8 double bond in the long-chain base is essential) show fruiting-inducing activity in the fungus *Schizophyllum commune*.

α-D-Galactosylceramides: Cerebrosides linked to an α-D- rather than a β-D-galactosyl unit such as that found in the marine sponge *Agelas mauritianus*, in human gut microflora, and even in cow's milk are potent stimulators of mammalian immune systems by binding to the protein CD1d on the surface of antigen-presenting cells and activating invariant natural killer T cells. Indeed this was one of the first pieces of evidence to show that glycolipids, like glycoproteins, could invoke an immune response. Subsequently, it was demonstrated that α-galactosylceramide with a 24:1 fatty acid, though present in very small amounts, is loaded onto the CD1d or CD40 protein and is presented as the natural endogenous ligand for NKT cells in the thymus and the periphery. Once activated, NKT cells secrete a range of pro-and anti-inflammatory cytokines to modulate innate and adaptive immune responses. The α-glucosyl and α-psychosine analogs show similar activity.

It is not certain whether α-galactosylceramide is synthesized in animal tissues, and it is likely that is derived primarily from members of the gut microbiome such as *Bacteroides fragilis* and related species (although in general, few bacterial species produce sphingolipids). Ceramide-galactosyltransferases responsible for the synthesis of this lipid in two species of bacteria from the intestinal microbiome have been identified. In mouse gut, the main molecular form consisted of a 2-(*R*)-hydroxylated hexadecanoyl chain linked to C_{18}-sphinganine, while that in *B. fragilis* contained longer-chain components with *iso*-methyl-branches in the sphingoid base and often fatty acid moieties. The sphinganine chain branching is a critical determinant of NKT cell activation by the bacterial enzyme. A decrease in the production of

this lipid was observed in mice exposed to stress conditions that alter the composition of the gut microbiota, including Western-type diet, colitis, and influenza A virus infection with potential consequences upon the systemic immune responses. Its concentration within animal tissues is controlled by catabolic enzymes in a two-step mechanism: removal of the acyl chain by an acid ceramidase followed by hydrolysis of the sugar residue by an α-glycosidase. Initial studies with animal models suggest that treatment with α-D-galactosylceramides is effective against lung and colorectal cancers, melanomas and leukemia, and pre-clinical trials of this lipid and synthetic analogs so far have shown that these are safe and effective as an anti-tumour immunotherapeutic agents and vaccine adjuvants. Indeed, a phase I trial with high-risk melanoma patients has given promising preliminary results.

21.4.5.4: CATABOLISM OF GLYCOSPHINGOLIPIDS

In animal tissues, the main sites for the degradation of **all** glycosphingolipids, including the monoglycosylceramides, oligoglycosphingolipids and gangliosides, are the lysosomes. These are membrane-bound organelles that comprise a limiting external membrane and internal lysosomal vesicles, which contain soluble digestive enzymes that are active at the acidic pH of this organelle. All membrane components are actively transported to the lysosomes to be broken down into their various primary components. In the case of glycosphingolipids, this means to fatty acids, sphingoid bases, and monosaccharides, which can be recovered for re-use or further degraded. Thus, sections of the plasma membrane enter the cell by a process of endocytosis, and they are then transported through the endosomal compartment to the lysosomes. The compositional and physical arrangement of the lysosomal membranes is such that they are themselves resistant to digestion with **bis(monoacylglycero)phosphate** (lysobisphosphatidic acid) as a characteristic component of the inner membrane. A glycocalyx of highly *N*-glycosylated integral membrane proteins protects the perimeter membrane with the aid of the ganglioside GM3, which is resistant to degradation. This glycocalix forms an efficient hydrophilic barrier at the luminal surface of the lysosomal perimeter membrane to protect it from degradation by proteases and hydrolases, and to prevent lipids and their hydrolysis products from escaping from the lumen of the lysosome.

Degradation of oligoglycosylceramides and gangliosides occurs by sequential removal of monosaccharide units via the action of specific exohydrolases from the non-reducing end until a monoglycosylceramide unit is reached when glucosylceramide β-glucosidases or an analogous β-galactosidase (one isoform) removes the final carbohydrate moiety. Several glucosylceramidases are known; GBA1 is a lysosomal hydrolase, GBA2 is a ubiquitous non-lysosomal enzyme and GBA3 is a cytosolic β-glucosidase. The last is found in the kidney, liver, spleen and a few other tissues of mammals, but its function is not clear.

As glycolipids with fewer than four carbohydrate residues are embedded in intralysosomal membranes, while the degradative enzymes are soluble, the process requires the presence of negatively charged lipids and specific activator proteins, which are water-soluble glycoproteins of low molecular weight. These are not themselves active catalytically but are required as cofactors either by directing the enzyme to the substrate or by activating the enzyme by binding to it in some manner. Five such proteins are known, the GM_2-activator protein (specific for gangliosides) and Sphingolipid Activator Proteins or **saposins A, B, C and D**, which perturb the membranes sufficiently to enable the degradative enzymes to reach the glycolipid substrates. The four saposins are derived by proteolytic processing from a single precursor protein, prosaposin, which is synthesized in the endoplasmic reticulum, transported to the Golgi for glycosylation, and then to the lysosomes. Of these, saposin C is essential for the degradation of galactosyl- and glucosylceramide, while saposin B is required for the hydrolysis of sulfatide, globotriaosylceramide, and digalactosylceramide. The products of the hydrolysis reaction with monoglycosylceramides are ceramides and monosaccharides with net retention of the stereochemistry of the latter in the process. This is illustrated in Figure 21.4.26.

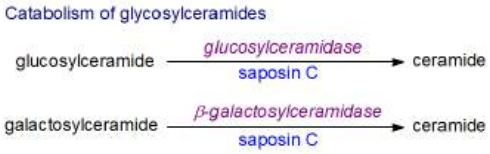


Figure 21.4.26: Catabolism of glycosylceramide

The reactions are aided by the presence of anionic lipids such as bis(monoacylglycero)phosphate. In particular, this increases the ability of the GM2-activator to solubilize lipids and stimulates the hydrolysis of membrane-bound GM1, GM2, and some of the kidney sulfatides. Saposin D stimulatesthe degradation of lysosomal ceramide by acid ceramidase, and it is also involved in the solubilization of negatively charged lipids at an appropriate pH. Eventually, the ceramides can in turn be hydrolyzed by an acid ceramidase to fatty acids and sphingoid bases.

β-Glucosylceramidase and saposin C are also required for the generation of the structural **ceramides** from glucosylceramide in the outer region of the skin, a process essential for optimal skin barrier function and survival. Some glucosylceramide is hydrolyzed by the enzyme GBA2 at the plasma membrane, where the ceramide formed is rapidly converted to sphingomyelin by the sphingomyelinase 2, which may be co-located with the glucosidase. In addition, it has been established that cellular β-glucosidases are able to transfer the glucose moiety from glucosylceramide to and from other lipids as in the formation of **cholesterol glucoside**.

Small but significant amounts of glucosyl- and galactosylceramides are ingested as part of the human diet. They are not hydrolyzed by pancreatic enzymes but are degraded in the brush border of the intestines by the enzyme lactase-phlorizin hydrolase (which also hydrolyses

the lactose in milk) to ceramides and thence to sphingosine.

An Arabidopsis homolog of human glucosylceramidase (AtGCD3) preferentially hydrolyses glucosylceramides that contain long acyl chains, and three further isoforms may exist based on sequence homology.

21.4.5.5: GENETIC DISORDERS AND DISEASE

Harmful quantities of glucosylceramide accumulate in the spleen, liver, lungs, bone marrow, and, in rare cases, the brain of patients with **Gaucher disease**, the most common of the inherited metabolic disorders (autosomal recessive) involving storage of excessive amounts of complex sphingolipids. Three clinical forms (phenotypes) of the disease are commonly recognized of which by far the most dangerous are those affecting the brain (Types 2 and 3). All of the patients exhibit a deficiency in the activity of the lysosomal glucosylceramide-β-glucosidase (GBA1), which catalyzes the first step in the catabolism of glucosylceramide. The enzyme may be present, but a mutation prevents it from forming its correct conformation, although other factors may be involved as patients with a defective saposin C, the lysosomal activator protein, develop similar symptoms.

In the brain, glucosylceramide accumulates when complex lipids turn over during brain development and during the formation of the myelin sheath of nerves. Other than in the brain, the excess glucosylceramide arises mainly from the biodegradation of old red and white blood cells. The result is that the glucosylceramide remains stored within the lysosomes of macrophages, i.e., the specialized cells that remove worn-out cells by degrading them to simple molecules for recycling, thus preventing them from functioning normally and often leading to chronic inflammation. The enlarged macrophages containing undigested glucosylceramide are termed Gaucher cells. They over-express and secrete certain proteins into the circulation, and some of these are used as biomarkers. In addition, glucosylceramide is converted more rapidly to **gangliosides** in these cells, leading to an increase in ganglioside GM3 in the plasma and spleen of patients with Gaucher disease. Fortunately, there are now effective enzyme replacement therapies for patients with the milder (non-neurological or Type 1) form of Gaucher disease that successfully reverse most manifestations of the disorder, including decreasing liver and spleen size and reducing skeletal abnormalities. Two oral drugs that inhibit glucosylceramide synthesis have also been approved.

Defective GBA1 enzyme activity in humans has been implicated in an increased risk of multiple myeloma and other cancers. Oligoglycosylceramides and gangliosides in particular are known to be involved in the pathology of a number of cancers, and glucosylceramide is an important precursor of these. Inhibition of glucosylceramide synthase, which is overexpressed in many human tumors lead to a marked arrest of cell growth in cancer cells *in vitro*, so this is believed to have the potential for the treatment of colorectal and other cancers.

A deficiency in glucocerebrosidase activity may predispose individuals to more common disorders such as Parkinson's disease and Lewy body dementia. Excess glucosylceramide production and thence of more complex glycosphingolipids is a factor in polycystic kidney disease. It appears to be a general rule that the mere process of lysosomal substrate accumulation in all lysosomal storage disorders impairs lysosome integrity and results in more general disruptions to lipid metabolism and membrane structure and function. On the other hand, inhibition of glucosylceramidases may be of benefit in cystic fibrosis. Krabbe disease is discussed in the next section.

Galactosylceramide is believed to function as an initial receptor for the human immunodeficiency virus (HIV) in mucosal epithelial cells and controls the early infection-independent phase of HIV transfer to T cells. Glucosylceramide levels regulate the uptake of viruses that rely upon the late endosomal compartment for fusion, including the influenza A and Ebola viruses.

21.4.6: GANGLIOSIDES

The name **ganglioside** was first applied by the German scientist Ernst Klenk in 1942 to a mixture of complex glycosphingolipids newly isolated from ganglion cells of brain. Subsequently, he demonstrated that as part of an oligosaccharide chain, they contained an acidic carbohydrate component, which he named "neuraminic acid" - later termed "sialic acid" from the Greek "sialon" for saliva, from which they were first isolated. However, it was not until 1963 that the first ganglioside species was fully characterized. Innumerable sphingolipids are now known that differ in the nature of both the glycan (glucose, galactose, *N*-acetylgalactosamine, and sialic acid residues) and **ceramide** structures. They are present throughout the animal kingdom, from echinoderms up to higher animals, but not in plants. Such highly polar, acidic and relatively hydrophilic molecules have distinctive physical properties, which are essential for the vital functions of gangliosides in the membranes of the central nervous system and other tissues.

21.4.6.1: SIALIC ACIDS AND GANGLIOSIDES

Sialic acids: Gangliosides are oligoglycosylceramides derived as a first step from lactosylceramide, and they are defined by the presence of one to as many as five **sialic acid** residues, i.e. carbohydrate molecules with a nine-carbon backbone and a carboxylic acid group, a subclass of the superfamily of naturally occurring non-2-ulosonic acids. Of the many forms that have been characterized, only a few are relevant to gangliosides, and the most important of these is *N*-acetylneuraminic acid (‘NANA’ or ‘SA’ or 'Neu5Ac' or 'NeuAc'). Less often the sialic acid component is *N*-glycolylneuraminic acid (Neu5Gc), which differs by only one oxygen atom at the C-5 *N*-acetyl group, or it can be a Neu5Ac analogue in which the amide group is replaced by a hydroxyl group, i.e. 3-deoxy-D-glycero-D-galacto-nonulosonic acid (ketodeoxynonulosonic acid or ‘KDN’). The sialic acids are joined via α-glycosidic linkages to one or more of the monosaccharide units,

e.g. via the hydroxyl group on position 2, or to another sialic acid residue. The polar head groups of the lipids carry a net-negative charge at pH 7.0 and they are acidic. Their structures are shown in Figure 21.4.27.

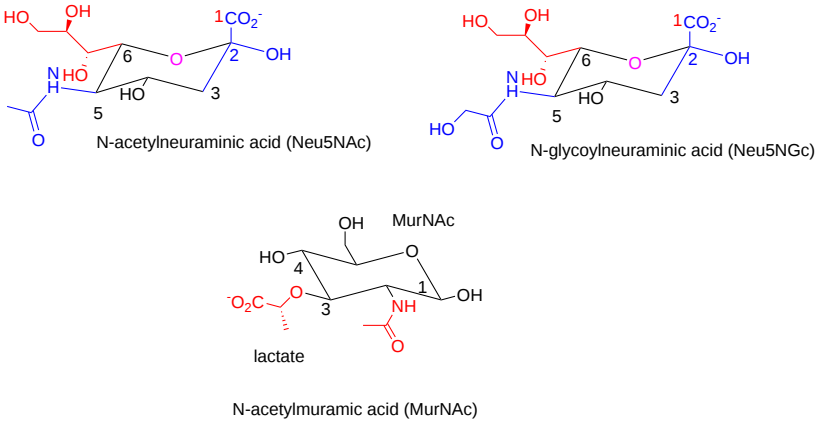


Figure 21.4.27: Sialic acids

Humans lack Neu5Gc: Neu5Ac is the biosynthetic precursor of Neu5Gc, a component of gangliosides from most animal species, including mice, horse, sheep, and goats, via the action of the enzyme CMP–*N*-acetylneuraminic acid hydroxylase (CMAH). However, NeuGc is not synthesized in humans (or birds and New World monkeys), although it is present in other primates such as the great apes, and indeed as it is a xeno-antigen, anti-NeuGc antibodies are produced normally in healthy humans (and especially after injection of NeuGc-containing glycoconjugates). The absence or irreversible inactivation of a number of relevant genes, but especially a critical exon in the *CMAH* gene, both for sialolipids and peptides in humans suggests that this may have been a major biochemical branch-point in human evolution that occurred ~2 to 3 million years ago after the divergence of humans and chimpanzees from a common ancestor. It may even be a factor in the superior performance of the human brain as the overexpression of Neu5Gc in the brains of transgenic mice was found to result in abnormal development. It could also mean that there might have been a fertility barrier between us and other hominids during evolution.

While these are speculations, there is some evidence that the loss of Neu5Gc in humans had complex effects on immunity, providing greater capabilities to clear sublethal bacterial challenges. Some NeuGc may be obtained from the diet in meat and milk, for example, and this may be incorporated into human gangliosides to a limited extent, especially in fetal tissues and some cancers. In the latter, preferential expression of dietary Neu5Gc has been ascribed to their higher metabolic rate.

21.4.6.2: 2. STRUCTURE AND OCCURRENCE OF GANGLIOSIDES

Most of the common range of gangliosides are derived from the ganglio- and neolacto-series of **neutral oligoglycosphingolipids (Table 1)**, and they should be named systematically in the same way with the position of the sialic acid residue(s) indicated as for branched structures. However, they are more conveniently defined by a short-hand nomenclature system proposed by Svennerholm in which M, D, T and Q refer to mono-, di-, tri- and tetrasialogangliosides, respectively, and the numbers 1, 2, 3, etc refer to the order of migration of the gangliosides on thin-layer chromatography. For example, the order of migration of monosialogangliosides is GM3 > GM2 > GM1 (sometimes defined by subscripts, e.g. G_{M1} or GM_1). To indicate variations within the basic structures, further terms are added, e.g. GM1a, GD1b, etc. Although alternatives have been proposed that are more systematic in structural terms, the Svennerholm nomenclature is that approved by IUPAC-IUB. Ganglio-series glycosphingolipids having 0, 1, 2 and 3 sialic acid residues linked to the inner galactose unit are termed asialo- (or 0-), a-, b- and c-series gangliosides, respectively, while gangliosides having sialic acid residues linked to the inner *N*-galactosamine residue are classified as α-series gangliosides. The structures for these groups are illustrated in the section on ganglioside **biosynthesis** below, for reasons of practical convenience.

As of 2020, more than 200 gangliosides with variations in the carbohydrate chain had been characterized in vertebrates alone. One of the most studied monosialo-gangliosides and the first to be fully characterized is ganglioside GM1a (Neu5Acα2-3(Galβ1-3GalNAcβ1-4)Galβ1-4Glcβ1Cer), a major brain ganglioside of mammals and the preferred ligand of cholera toxin, illustrated in Figure 21.4.28.

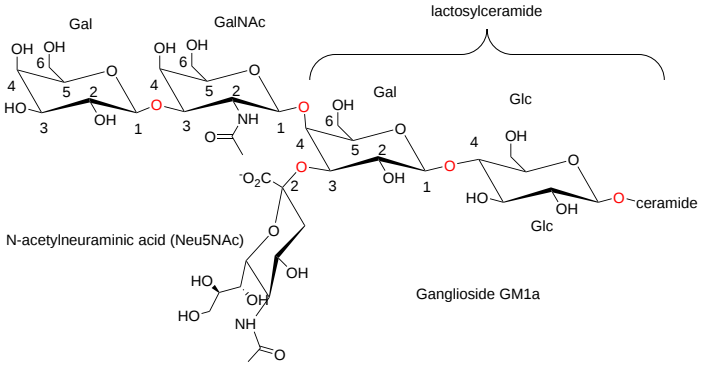


Figure 21.4.28: Ganglioside GM1a

It can also be depicted using the abbreviated structure shown in Figure 21.4.29.

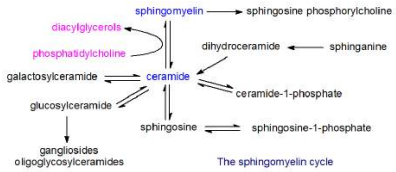


Figure 21.4.29: Abbreviated structure for Ganglioside GM1a

An alternative nomenclature, which is less used, is recommended by IUPAC-IUB and is based upon the ganglio (Gg) root structure; it employs Roman numerals to designate each hexose unit and the location of the Neu5Ac along the carbohydrate chain with Arabic superscripts to designate the hydroxyl group to which this is linked. By this system, GM1a is defined as II^3-α-Neu5Ac-Gg_4Cer.

Brain gangliosides: Gangliosides can amount to 6% of the weight of lipids from the brain (20 to 500 times more than in other tissues), where they constitute 10 to 12% of the total lipid content (20-25% of the outer layer) of neuronal membranes, for example. Aside from this, they are synthesized and are present at low levels (1 to 2% of the total lipids) in all animal tissues, where like the neutral oligoglycosphingolipids they are concentrated in the outer leaflet of the plasma membrane in the nanodomains known as '**rafts**' or in related structures. Mammalian neurons actively synthesize gangliosides of the ganglio-series primarily, but oligodendrocytes in the brain produce instead myelin-forming glycosphingolipids, such as galactosylceramide and sulfatide together with a minor amount of ganglioside GM4.

The brain contains as much as 20 to 500 times more gangliosides than most non-neural tissues, with three times as much in grey as in white matter. As the brain develops, there is an increase in the content of gangliosides and in their degree of sialylation. There are large differences between species and tissues. For example, during embryogenesis and the postnatal period in the human central nervous system, the total amount of gangliosides increases approximately threefold, while that of GM1 and GD1a increases 12 to 15-fold. During the same period, the hemato-series gangliosides GM3, GD3, and 9-OAc-GD3, which lack a hexosamine residue, are the predominant ganglioside species, but they are present in much lower amounts in adults and then in some areas of the brain only. In the mouse brain, the total amount of gangliosides is almost 8-fold greater in adults than in embryos, with a similar shift in composition from simple (GM3 and GD3) to more complex gangliosides. It is evident that the ganglioside changes during brain maturation are correlated with many neuro-developmental milestones, and there is no doubt that gangliosides play a crucial role in neuronal function and brain development, especially during infancy when there is high nutrient demand as the brain undergoes rapid restructuring.

The main gangliosides (~95%) of adult mammalian brain are ganglio series GM1, GD1a, GD1b, and GQ1b, while lactosyl series gangliosides such as GM3 (sialyllactosylceramide) are found mainly in the extra-neural tissues. The remaining ~5% consists of minor components in the brain include gangliosides GM4, GM3, GD3, GM2, GD2, Fuc-GM1, Fuc-GD1b, GT1a and GP1c, the proportions of which vary depending on species. On the other hand, modern mass spectrometric methodology (electrospray ionization ion mobility MS) has revealed a much higher degree of sialylation than was previously recognized, including a complete series of mono- to octasialylated gangliosides in fetal frontal lobe. Subsequently, many previously unknown acetylated gangliosides were found in fetal hippocampus by this methodology. The content and composition of gangliosides in the brain also change with aging, with a substantial fall in the content of lipid-bound sialic acid but an increase in the proportion of the more complex forms in terms of carbohydrate structures in the elderly.

Gangliosides in other tissues and species: Among the extraneural tissues, lactosyl series gangliosides such as GM3 (sialyllactosylceramide) and monosialogangliosides, in general, tend to predominate. Relatively high concentrations of ganglioside GD1a are present in erythrocytes, bone marrow, testis, spleen, and liver, while GM4 is more abundant in kidney, GM2 in bone marrow, GM1 in erythrocytes and GM3 in intestine. In germ cells of mice, there is a switch between gangliosides of the a- and 0-series upon differentiation when they are crossing the blood-testis barrier. Skin fibroblasts and many cells of visceral organs generate gangliosides of the globo series mainly. Similarly, glob-o and lacto series gangliosides are characteristic components of the stage-specific embryonic antigens (SSEA), which underlie the development and differentiation of human embryonic stem cells. A sialyl-lactotetraosylceramide is present in the latter

and in the brains of children under the age of two, but not in tissues of adult humans. Gangliosides can cross the placental barrier into the fetus and those in milk, derived from the apical plasma membrane of secretory cells of the mammary gland, may be of nutritional importance for the newborn. GD3 is the main ganglioside in human breast milk at an early stage of lactation, whereas GM3 is more abundant in the later stages (and in bovine milk). Unfortunately, gangliosides are poorly characterized and quantified in foods in general.

A 5-*N*-deacetylated form of ganglioside GM3 has been detected in human melanoma tumors. In addition, *O*-acetylation or lactonization of the sialic acid residue adds to the potential complexity. Gangliosides containing *O*-acetylated sialic acids, such as 9-*O*Ac-GD3, are expressed during embryonic development and in the retina and cerebellum of adult rats, but not other brain regions. They occur also in certain tumors and may protect them from apoptosis. It is possible that such gangliosides are even more widespread, but they are missed after treatment with mild alkali during the isolation procedure, a common analytical practice. A further complexity is the occurrence of gangliosides with sulfate groups, and these have been isolated from human, mouse, and monkey kidney cells. KDN-containing gangliosides are minor components of egg, ovarian fluid, sperm and testis of fish and of some mammalian tissues

Gangliosides from marine invertebrates (echinoderms), such as starfish and sea cucumbers, are very different in structure from those in vertebrates and do not have a shorthand nomenclature. They include forms with distinctive ceramide compositions, untypical carbohydrate residues, sialic acids within the oligosaccharide chain, or with glycosyl inositol-phosphoceramide structures. The mollusc, *Aplysia kurodai*, lacks gangliosides but produces complex oligoglycosylceramides with 2-aminoethylphosphonic acids and/or phosphoethanolamine groups attached that may serve as ganglioside surrogates.

Ceramide structures: In general, the ceramide structures of gangliosides tend to be relatively simple. Sphingosine is usually the main sphingoid base, accompanied by the C_{20} analog in gangliosides of the central nervous system. Stearic acid (18:0) can be 80 to 90% of the fatty acid constituents in the brain, accompanied by small amounts of 16:0, 20:0 and 22:0, but with little or no polyunsaturated or 2-hydroxy acids, other than in some exceptional circumstances (e.g. some carcinomas). Palmitic acid is more abundant in gangliosides of the intestines and liver, while 2-hydroxylated fatty acids are relatively abundant in the last and in the kidney. There are also differences in the composition of the base and fatty acid components in different cells or regions of the brain. During development, the nature and concentrations of these constituents change markedly, and for example, the ratio of C_{20}/C_{18}-sphingosine in ganglioside GD1a of cerebellum increases 16-fold from 8-day-old to 2-year-old rats. In gangliosides outwith the nervous system, C_{20}-sphingosine is barely detectable, and there is often a much wider range of fatty acid constituents (C_{14} to C_{24}).

The nature of the ceramide component is relevant to the biological function of gangliosides, and changing the fatty acid component to α-linolenic acid by synthetic means alters the biological activity of gangliosides dramatically *in vitro*. However, it is the carbohydrate moiety that has the primary importance for most of their functions, and detailed discussion of these structures would take us into realms of chemistry best left to carbohydrate experts (see the reading list below). In any given cell type, the number of different gangliosides may be relatively small, but their nature and compositions may be characteristic and in some way related to the function of the cell. It is noteworthy that some terminal glycan structures of gangliosides are also present in glycoproteins of membranes.

21.4.6.3: 3. BIOSYNTHESIS

There is evidence that the pool of **glucosylceramide** and thence of **lactosylceramide** that is utilized for ganglioside biosynthesis is different from that for the other neutral **oligoglycosylceramides**. This may explain some of the differences between the two groups in the fatty acid and sphingoid base components, which will also be dependent upon cell type. It is an open question how the ganglioside precursors enter the Golgi and trans-Golgi network where synthesis occurs at the luminal leaflet, but it appears that the regulation of intracellular sphingolipid traffic may be as important as the control of enzyme expression and activity in determining the final compositions of the various glycosphingolipid types.

In humans, sialic acid biosynthesis occurs by a series of reactions in the cytosol, but the Neu5Ac produced is transferred to the nucleus and activated by the cytosine 5'-monophosphate *N*-acetylneuraminic acid synthetase (CMAS) to form CMP-Neu5Ac, which is transported to the Golgi apparatus by a family of sialyltransferases specific for particular glycosidic linkages (α2,3, α2,6, α2,8, and α2,9).

Thereafter, the pathways for the biosynthesis of the common series of gangliosides of the ganglio-series, for example, involve sequential activities of distinct membrane-spanning sialyltransferases and glycosyltransferases as illustrated in Figure 21.4.29 for the four main 0-, a-, b- and c-series of gangliosides.

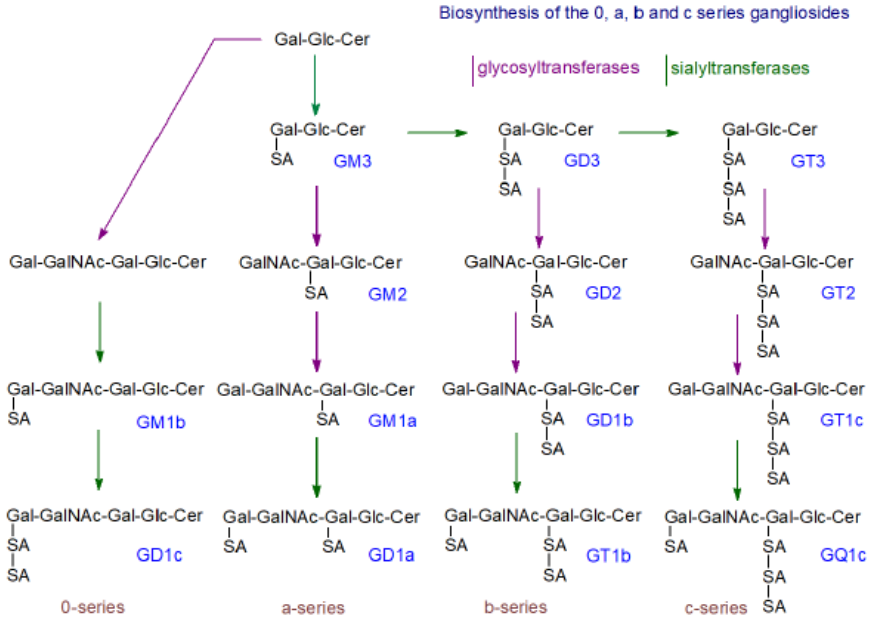


Figure 21.4.29: Biosynthesis of gangliosides

The required enzymes are bound to the membranes of the Golgi apparatus in a sequence that corresponds to the order of addition of the various carbohydrate components. Thus, the sialyltransferase that catalyzes the synthesis of the relatively simple ganglioside GM3 is located in the *cis*-region of the Golgi, while those that catalyse the terminal steps of ganglioside synthesis are located in the distal or *trans*-Golgi region. The GM3 synthase in particular, which catalyzes the transfer of Neu5Ac from cytidine monophosphate (CMP)-Neu5Ac onto the terminal galactose residue of lactosylceramide, has a unique specificity.

The simple ganglioside GM3 is synthesized by the addition of sialic acid to lactosylceramide by CMP:LacCer α2-3 sialyltransferase (or GM3 synthase), before GD3 and GT3 are produced in turn by the action of appropriate synthases. Subsequently, GM3, GD3 and GT3 serve as precursors of more complex gangliosides by the action of further glycosyl- and sialyl-transferases. An alternative theory with some supporting evidence proposes that a multiglycosyl-transferase complex is responsible for the synthesis of each individual ganglioside rather than a series of individual enzymes. Further sialylation of each of the a, b, and c series and in different positions in the carbohydrate chain can occur to give an increasingly complex and heterogeneous range of products, such as the α-series gangliosides with sialic acid residue(s) linked to the inner *N*-acetylgalactosamine residue (not illustrated). GM4 or NeuAcα2,3Gal-Ceramide, a minor component of the brain and present in a few other tissues at low levels, is an exception in that galactosylceramide is its precursor. Finally, the newly synthesized gangliosides are transferred to the external leaflet of the plasma membrane via the lumenal surface of transport vesicles. Gangliosides are also important constituents of nuclear membranes.

The changes that occur in ganglioside compositions of brain and other tissues in the embryonic and post-natal stages are governed mainly by changes in the expression level and activity of the glycosyl- and sialyl-transferases, although the former can also be regulated by glycosylation and phosphorylation.

The presence of distinctive sialidases that differ from the catabolic lysosomal enzymes (see below) in raft-like regions of the plasma membrane bring about further changes in the composition of the cell surface gangliosides that can be specific to particular cell types, causing a shift from poly-sialylated species involving a decrease of GM3 and formation of GM2 then GM1 by hydrolysis of terminal sialosyl residues linked either α2-8 on another sialic acid or α2-3 on galactose. As GM1 is resistant to most sialidases, it tends to increase in concentration relative to oligosialo species as developmental and other GM1-requiring processes come into play. This may have consequences for important cellular events, such as neuronal differentiation and apoptosis. Conversely, sialylation may occur in some neuronal membranes, increasing the proportions of poly-sialylated species. In particular, a CMP-NeuAc:GM3 sialyltransferase is able to sialylate GM3. Gangliosides GM1 and GD1a have been identified in both membranes of the nuclear envelope together with two neuraminidases.

Ganglioside lactones, where the sialic acids are linked together with ester linkages, have been detected as minor components in brain tissues, where lactonization occurs at the plasma membrane. As the process of lactonization profoundly influences the shape and biological properties of the original ganglioside, it is possible that lactonization-delactonization in a membrane might be a trigger for specific cellular reactions. Similarly, GD3 ganglioside can undergo *O*-acetylation at C9 of the outer sialic acid with important metabolic implications.

Gangliosides added to many types of cell preparations *in vitro* are rapidly taken up by the cells, while gangliosides injected into animals *in vivo* are rapidly internalized by tissues. They can cross the blood-brain barrier, and via the placenta, they can enter the fetus. Similarly, dietary gangliosides are absorbed intact by intestinal cells but are broken down to their lipid and carbohydrate constituents for re-use. The

sialic acids released by an intestinal sialidase are transported in plasma to the brain and other tissues where they influence ganglioside expression. Indeed, there is some experimental evidence that dietary gangliosides may improve cognitive functions in animals and humans.

21.4.6.4: CATABOLISM

Degradation of gangliosides takes place at the surface of intralysosomal luminal vesicles, generated by an inward budding of the endosomal membrane, and these are reached by a process of endocytosis. In brief in relation to gangliosides, soluble sialidases (neuraminidases) and exoglycohydrolases remove individual sialic acid and sugar residues sequentially from the non-reducing terminal unit, as illustrated for ganglioside GM1, with the eventual formation of ceramide, which is then split into long-chain base and fatty acids by ceramidases. This degradation occurs through the endocytosis-endosome-lysosome pathway with a requirement for an acidic pH inside the organelle. In addition to the sialidases and exoglycohydrolases, the various reactions have an absolute requirement for effector molecules, termed 'sphingolipid activator proteins', including saposins (Sap), and the specific GM2-activator protein (GM2-AP). Ganglioside GM3 is a component of the lysosomal perimeter membrane, but is protected from degradation by a glycocalix of the membrane facing the lysosol. Anionic lipids and especially **bis(monoacylglycero)phosphate** in the membranes stimulate ganglioside degradation while cholesterol is inhibitory. The catabolic pathway is shown in Figure 21.4.30.

Figure 21.4.30 Catabolism of gangliosides

This process constitutes a salvage mechanism that is important to the overall cellular economy since a high proportion of the various hydrolysis products are recycled for glycolipid biosynthesis. By generating ceramide and sphingosine, it may also be relevant to the regulatory and signaling functions of these lipids. In addition, some partial hydrolysis of gangliosides occurs in the plasma membrane as part of a biosynthetic remodeling process discussed above. Defects in catabolism lead to the gangliosidoses discussed later.

21.4.6.5: GANGLIOSIDE FUNCTION

Cell surface effects: In their natural biological environment, gangliosides have a negative charge because of the presence of sialic acids, which also add to the hydrophilicity of the polysaccharide constituent. This is balanced somewhat by the hydrophobic character of the ceramide moiety, so that over all the molecules are amphiphilic in nature, but very different from the glycerophospholipids, which are essential for the formation of membrane bilayers. Indeed, a ganglioside such as GM1 is virtually soluble in water, where it can form large aggregates though hydrophilic effects. The nature of the ceramide unit with its capacity to form hydrogen bonds with glycerophospholipids is important in ensuring that gangliosides are inserted in a stable manner into the outer layer of the plasma membrane.

Thus, gangliosides are anchored in membranes by their ceramide units with the double-tailed sialoglycan components extending out from the cell surface, where they can participate in intermolecular interactions by a network of hydrogen bonds and hydrophobic interactions. For example, the glucose-ceramide bond of GM1 is oriented in the outer leaflet of the plasma membrane such that the glycan extends perpendicularly to the plane of the lipid bilayer. All gangliosides, but especially the simplest GM3 or Neu5Acα2-3Galβ1-4Glcβ1Cer, have a structural role, and they a natural propensity to laterally segregate and to associate with each other and with other sphingolipids, phospholipids and cholesterol into **raft** nano-domains or in related structures, such as the **caveolae**, where the very large surface area occupied by the oligosaccharide chain imparts a strong positive curvature to the membrane. In this environment, gangliosides can interact with each other through side-by-side hydrogen bonds mediated by water molecules that act as bridges between the chains.

Further, molecules of GM3 and other gangliosides self-aggregate into clusters on the surface of lymphocytes of human peripheral blood, and there is evidence that the density of these clusters in membranes governs their reactivity as antigens. In addition, it is believed that gangliosides and other oligoglycosylceramides can cluster together through hydrogen donor-acceptor (*cis*) interactions because of the presence of hydroxyl and acetamide groups to form glycosynaptic domains, which are related to but functionally distinct from raft signaling platforms (with lower cholesterol concentrations). Many of the biological functions of gangliosides are mediated through their location in these nanodomains, where they may have specialized functions in cell adhesion, growth, and motility through interactions with specific proteins and signal transduction pathways. However, not all gangliosides are present in such raft-like structures.

Receptor/signaling functions: Gangliosides can bind to membrane proteins directly by carbohydrate-carbohydrate or carbohydrate-amino acid interactions, usually involving specific ganglioside head groups, resulting in changes to the location of proteins within membrane microdomains for recruitment of signaling partners, or to dimerization or other effects upon receptors. In rafts and caveolae especially, gangliosides can modulate cell signaling processes by their interactions with specific receptors, adhesion molecules, and ion channels. Cell–cell (*trans*) interactions occur by sialoglycans on one cell binding to complementary binding proteins (lectins) on adjacent cells, bringing

about adhesion of cells and enabling regulation of intracellular signaling pathways, e.g. myelin-associated glycoprotein on myelin sheaths binds to gangliosides present on axonal membranes.

In addition, gangliosides act as receptors of interferon, epidermal growth factor, nerve growth facto,r and insulin, and they may regulate cell signaling and control growth and differentiation of cells in this way. While intact gangliosides inhibit growth by rendering cells less sensitive to stimulation by epidermal growth factor, removal of the *N*-acetyl group of sialic acid enhances this reaction and stimulates growth. Gangliosides function as antigens or receptors by recognizing specific molecules (lectins), including bacterial toxins, at the cell surface and by modulating the charge density at the membrane surface (see the section on Gangliosides and Disease below). They also regulate the activities of proteins within the plasma membrane and especially receptor-type tyrosine kinases. For example, the phosphorylation state and activity of insulin receptors in caveolae and thence the insulin resistance of cells is controlled by the concentration of GM3, the main ganglioside in plasma and other extraneural tissues. GM3 interacts also with the epidermal growth factor receptor leading to cell growth inhibition. GM1 strongly influences specific neuronal functions by interacting with specific receptors such as the tropomyosin receptor kinase (Trk) A (TrkA) receptor by altering its conformation to enable interaction with the nerve growth factor (NGF) ligand.

GM3 (SA-Gal-Glc-Cer) is a serum ganglioside that is highly enriched in a type of membrane microdomain termed a 'glycosynapse', and it forms complexes with co-localized cell signaling molecules. It has a function in the innate immune function of macrophages and it has been demonstrated that molecular species of GM3 with differing acyl-chain structures and modifications can operate as pro- and anti-inflammatory modulators of Toll-like receptor 4 (TLR4); very-long-chain and α-hydroxy GM3 species increase TLR4 activation, while long-chain and unsaturated GM3 species have the opposite effect. In addition, gangliosides have been shown to be cell-type specific antigens that have key functions in immune defense. For example, a major immunological function of gangliosides and sialic acids is to protect cells from attack by our own immune system and from autoimmunity. They recognize and protect host organs and tissues from complement attack by binding to the complement regulatory protein factor H, which has the potential to exert strong cytotoxic and inflammation-inducing activity. In particular, sialic acids protect against complement killing of autologous cells by binding to this protein via the α2–3 linked sialic acid glycans of the GD3 ganglioside. On the other hand, the breakdown of this system can lead to autoimmune diseases.

Brain function: One of the first examples of a ganglioside influencing a signaling event to be studied in some detail concerns the simple ganglioside GD3, which has a central role in early neurogenesis. GD3 binds to the epidermal growth factor receptor (EGFR) via a protein-carbohydrate interaction involving its terminal *N*-acetylneuraminic acid and a lysine residue in the transmembrane domain of the receptor and also by a carbohydrate-carbohydrate interaction thereby maintaining the latter in its inactive monomeric state. EGFR then binds to the epidermal growth factor and stimulates the transition of the receptor from an inactive monomeric to an active homodimeric form, and this in turn triggers receptor auto-phosphorylation and activation of a signaling cascade that promotes cell proliferation. This has proven to be essential for the regulation of the stem cell self-renewal capacity in the brain. In contrast, the neutral oligoglycosphingolipid Gb4 exerts the opposite effect on EGFR by interacting directly with it to potentiate its auto-phosphorylation with activation of the downstream cascade.

The techniques of molecular biology such as targeted gene deletion, which enable specific enzymes to be eliminated from experimental animals, are now leading to a better understanding of the function of each ganglioside. It is evident that they are essential to central myelination, to maintain the integrity of axons and myelin, and for the transmission of nervous impulses. These effects may be mediated by interactions of the negatively charged sialic acid residues of gangliosides with calcium ions, which are critical for neuronal responses. For example, a variant of GD3, 9-*O*-acetyl GD3, appears to be involved in glial-guided neuronal migration during brain development in the rat, while GM1 may have a similar function in humans; it determines which growth cone of unpolarized neurons becomes the axon. By stabilizing neuronal circuits, gangliosides have a function in memory, and conversely, disturbances in ganglioside synthesis can lead to neurodegenerative disorders (see below). Ganglioside GM3 in raft domains has been shown to have an indispensable role for the development, function, and viability of cochlear hair cells and thence it is essential for hearing. On the other hand, mice that express GM3 primarily and are devoid of the typical complex gangliosides of the brain suffer weight loss, progressive motor and sensory dysfunction, and deterioration in spatial learning and memory with aging. GD3 is important for retinal structure and visual function in mice.

Changes in ganglioside composition can be induced by nerve stimulation, environmental factors, or drug treatments. The various interconvertible ganglioside types in the plasma membrane of neurons are particularly important for its development in that they regulate such processes as axonal determination and growth, signaling, and repair. In addition, gangliosides are believed to be functional ligands for the maintenance of myelin stability and the control of nerve regeneration by binding to a specific myelin-associated glycoprotein. The occurrence of gangliosides in cell nuclei suggests a possible involvement of gangliosides in the expression of genes relevant to neuronal function. For example, the monosialoganglioside **GM1** has been shown to promote the differentiation of various neuronal cell lines in culture. It has protective effects on the neural system by encouraging neural stem cell survival and proliferation, while facilitating the stability and regeneration of axons, and by inhibiting neurodegeneration through autophagy, for example after ischemic stroke. Within membrane rafts, this ganglioside has key roles in several signaling systems through association with specific proteins that have glycolipid-binding domains, including those that modulate mechanisms such as ion transport, neuronal differentiation, G protein-coupled receptors (GPCRs), immune system reactivities and neuroprotection. It is important for Ca^{2+} and Na^{+} homeostasis in the nucleus and plasma

membrane and in regulating the effects of platelet-derived growth factor. However, there have been unpleasant complications when GM1 has been administered for therapeutic purposes. GD1a is sometimes considered to be a reserve pool for GM1.

After nerve injury, toll-like receptor 2 (TLR2) signaling is important for the induction of neuropathic pain; ganglioside GT1b functions as a TLR2 agonist to produce mechanical and thermal hypersensitivity.

Other functions: The ganglioside GD3 is essential for the process of apoptosis by blocking the activation of specific transcription factors and thence disabling the induction of antiapoptotic genes. 9-*O*-Acetylation of the GD3 molecule prevents ganglioside oxidation and blocks its pro-apoptotic effects. Similarly, GD3 is a regulator of autophagy, i.e. the degradation and/or recycling of cellular components. Gangliosides are also important in reproduction, and in mice, GD1a has been shown to be important to oocyte maturation, monospermic fertilization, and embryonic development, while GM1 is important in sperm-oocyte interactions and sperm maturation processes. Deletion of the GM2/GD2 synthase leads to infertility in male mice and the production of a novel fucosylated ganglioside containing very-long-chain polyunsaturated fatty acids. Related studies with gene knockout mice have revealed that b-series gangliosides are important in leptin secretion from adipocytes, while a-series gangliosides interact with the leptin receptor in the hypothalamus to influence the balance of energy.

21.4.6.6: GANGLIOSIDES AND DISEASE

Bacterial toxins and viruses: In relation to adaptive immunity, a-series and o-series gangliosides in the plasma membrane are involved in the function and stimulation of receptors on certain subsets of T cells by acting as pattern-recognition receptors for invading pathogens. In particular, certain gangliosides bind specifically to viruses and to various bacterial toxins, such as those from botulinum, tetanus and cholera, and to blood merozoites of the deadliest malaria parasite *Plasmodium falciparum*, and they mediate interactions between microbes and host cells during infections, with NeuAc as the main recognition module. The best known example is cholera toxin, which is an enterotoxin produced by *Vibrio cholerae* where the specific cell surface receptor is ganglioside GM1; the five B-chains of cholera toxin each bind one molecule of GM1. Interestingly, the subsequent metabolism of the ganglioside-toxin complex is dependent on the nature of the fatty acid components of the ganglioside. It is believed that toxins utilize the gangliosides to hijack an existing retrograde transport pathway from the plasma membrane to the endoplasmic reticulum. For example, the passage of the cholera toxin through the epithelial barrier of the intestine is mediated by GM1, possibly by endocytosis of the toxin-GM1 complex via caveolae into the apical endosome and thence into the Golgi/endoplasmic reticulum, where the complex dissociates. The consequence is persistent activation of adenylate cyclase by the toxin and continuous production of cAMP that leads to the severe fluid loss typical of cholera infections. As a further example, the botulinus toxin binds to a complex of a polysialoganglioside with the protein synaptotagmin, which together act as a high-affinity receptor complex to enable the neurotoxic effects. Similarly, ganglioside GM2 binds to a toxin secreted by *Clostridium perfringens*.

Influenza viruses have two glycoproteins in their envelope membranes, hemagglutinins, which bind to cellular receptors such as gangliosides, and after entry into respiratory epithelial cells, the sialidase (neuraminidase) of the virus cleaves the sialic acid from the receptors to prevent entry of further viruses to the cell. Variations in the structure of these proteins force the development of new vaccines The carbohydrate moiety of gangliosides is essential for the initial binding of viruses, but the lipid moiety is believed to be important for controlling their intracellular transport.

Some gangliosides and GD1a especially have anti-inflammatory properties in that they inhibit the effects of bacterial lipopolysaccharides by preventing the activation of tumor necrosis factor (TNF) and other cytokines. In contrast, GM2 may increase cytokine production in similar circumstances, while the heat-labile toxins of *Escherichia coli* bind to several gangliosides in macrophages, thus activating an inflammatory response.

Gangliosidoses and other neurodegenerative diseases: It appears to be a general rule that the mere process of lysosomal substrate accumulation in all lysosomal storage disorders impairs lysosome integrity and results in more general disruptions to lipid metabolism and membrane structure and function, inevitably triggering pathologic mechanisms. Endogenous generation of antibodies to gangliosides is often a factor, and it has been argued that gangliosides and their sialic acids components are at the border of immune tolerance.

As with the neutral **oligoglycosylceramides** and **ceramide monohexosides**, a number of unpleasant lipidoses have been identified that involve the storage of excessive amounts of gangliosides in tissues because of failures in the catabolic mechanism. The most important of these are the **GM2 gangliosidoses**, i.e. **Tay-Sachs disease** (and the similar Sandhoff disease), a fatal genetic disorder found mainly in Jewish populations in which harmful quantities of ganglioside GM2 accumulate in the nerve cells in the brain and other tissues. Lyso-GM2 (non-acylated) in plasma may serve as a marker. A modified GM2 derivative that contains taurine in amide linkage to the sialic acid carboxyl group has been identified in the brain of such patients. As infants with the most common form of the disease develop, the nerve cells become distended and a relentless deterioration of mental and physical abilities occurs. The condition is caused by insufficient activity of specific enzymes, i.e. β-*N*-acetylhexosaminidase, which catalyzes the degradation of gangliosides by removing the terminal *N*-acetylgalactosamine residue from GM2, or the GM2 activator protein.

In addition, a generalized **GM1 gangliosidosis** (an autosomal recessive and neurodegenerative disease) has been characterized in which ganglioside GM1 accumulates in the nervous system leading to mental retardation and enlargement of the liver. The condition is a consequence of a deficiency of the lysosomal β-galactosidase enzyme, which hydrolyses the terminal β-galactosyl residues from GM1

ganglioside to produce GM2. It appears that storage of substantial amounts of unwanted lipids in the lysosomal system leads to a state of cellular starvation, so that essential elements such as iron are depleted in brain tissue. The presence of lyso-GM1 in plasma is now seen as a useful aid to diagnosis. Small amounts of some gangliosides accumulate as secondary storage compounds in **Niemann–Pick** disease. The **Guillain–Barré syndrome** is an acute inflammatory disorder, usually triggered by a severe infection, which affects the peripheral nervous system. Antibodies to gangliosides are produced by the immune system, leading to damage of the axons, which can result in paralysis of the patient. Huntington's disease is believed to involve disruption of the metabolic pathways between glycosylceramides and gangliosides, and there is a human autosomal recessive infantile-onset epilepsy syndrome caused by a mutation to a sialyl transferase. Impaired ganglioside metabolism is also relevant to Alzheimer's disease, because complexation with ganglioside GM1 may cause aggregation of the amyloid β-protein deposits that characteristically accumulate in the brain in this condition (this explanation does not appear to be universally accepted). In general, in ganglioside deficiencies, natural or induced, it appears that progressive inflammatory reactions take place, leading to neurodegeneration in part because of the deterioration of the architecture of lipid rafts.

On the other hand at normal tissue concentrations, gangliosides such as GM1 are believed to have an anti-inflammatory and neuroprotective role in certain types of neuronal injury, Parkinsonism, and some related diseases. For example in relation to Parkinson's disease, GM1 binds to α-synuclein and inhibits or eliminates fibril formation. It may have a protective role by preventing sphingomyelin-induced aggregation, although as the overall level of GM1 decreases during aging, its beneficial effect decreases. For these reasons, the therapeutic properties of ganglioside GM1, the most accessible species, and derived molecules are under clinical investigation. However, there is no approved therapy for any gangliosidosis, although a number of different therapeutic strategies are being studied, including hematopoietic stem cell transplantation and gene therapy. For the moment, the blood-brain barrier remains a challenge.

Cancer: Gangliosides have important functions in cancer, especially in the regulation of signal transduction induced by growth-factor receptors in a specific microdomain termed a 'glycosynapse' in the cancer cell membranes, and in interactions with glycan recognition molecules involved in cell adhesion and immune regulation. In particular, depending on tissue, certain distinctive gangliosides are expressed at much higher levels in tumors than in normal healthy tissues, mainly by aberrant expression of glycosyltransferases and glycohydrolases. This enables tumor cells to escape immune surveillance and retain their malignancy. GM3 is not expressed in melanocytes normally, but is detected in 60% of primary melanomas and in 75% of metastatic melanomas, for example. Gangliosides can be shed from the surface of tumor cells into the local environment where they can influence interactions between cancer cells, including the transition of tumors from a dormant to a malignant state (angiogenesis); when present in the circulation they can be useful diagnostic aids. For example, the ganglioside GM3 is elevated in the serum of patients with breast cancer and may be a biomarker for the disease, while disialylated gangliosides GD2 and GD3 (Figure 21.4.31) are considered to be markers of neuroectoderm origin in tumors (neuroblastoma).

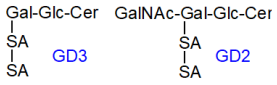


Figure 21.4.31: Structure GD2 and GD3

Specific gangliosides can have either positive or negative effects upon the regulation of the malignant properties of cancer cells. As a generality, disialyl glycosphingolipids or tandem-repeated sialic acid-structures confer malignant properties in various cancer systems; they are not merely markers. For example, the disialo-gangliosides GD2 and GD3 are present in trace amounts only in normal tissues, but are found at much higher concentrations in cancer cells, especially melanomas and neuroblastomas, with GD2 especially elevated in triple-negative breast cancer. These b-series gangliosides play a substantial part in the malignant properties of gliomas by mediating cell proliferation, migration, invasion, adhesion, and angiogenesis, and in preventing immunosuppression. They are considered to be tumor-associated antigens, and the GD2 and GD3 synthases are seen as important drug targets. In contrast, monosialyl gangliosides, such as GM1, GM2 and GM3, may suppress the malignant properties of various cancer cells. The mechanism is believed to involve complex formation at the cell surface with membrane proteins, such as growth factor receptors and adhesion receptors like those of the integrin family, leading to the modification of cell signals mediated by these receptors. Metastatic melanoma cells have high levels of GD3 in comparison to poorly metastatic cells or the normal counterpart, suggesting that GD3 may promote metastasis possibly by suppressing the anti-tumor immune response.

Ganglioside GM3(Neu5Gc), i.e. containing an abnormal sialic acid, is sometimes considered to be a tumor-specific antigen and a target for cancer immunotherapy. Aberrant sialylation is found in many malignant cancers, where the levels of neuraminidases are key factors for metastasis and survival of cancer cells, and there can be a significant accumulation of unusual gangliosides containing *N*-glycolyl sialic acid in some cancers. *N*-Glycolyl-GM3, normally absent from human tissues, is present in all stage II breast cancers, and it is accompanied by a number of other less common complex gangliosides. Similarly, the 5-*N*-deacetylated form of GM3 is expressed in metastatic melanomas, but not in healthy tissue or even in primary melanomas; it is considered to be a specific marker for the metastatic condition and a target for potential therapy. Increased synthesis of 9-*O*-acetyl-GD3, dependent on a sialyl-*O*-acetyltransferase - CAS1 Domain-Containing Protein 1, occurs in acute lymphoblastic leukemia and in malignant melanomas, and this appears to limit apoptosis, while *O*-acetylated GD2 (OAcGD2) is expressed in breast cancer and other tumors. A unique fucosyl-GM1 in which the terminal galactose is α-1,2-fucosylated at the non-reducing end is found circulating in the serum of patients with a number of cancers and especially with small-cell lung cancer but rarely in normal conditions, and it is also considered to be a potential indicator of cancer and a candidate for immunotherapy.

Clinical trials with an antibody to GD2 have been carried out successfully against the rare childhood cancer neuroblastoma, and the USDA has approved the use of this in combination with other drugs to treat this often lethal cancer. However, this antibody can have painful side effects due to an interaction with GD2 on neurons, and modified antibodies, which may be safer, are now being tested in multiple clinical trials. A phase I clinical trial with an antibody to GD3 has shown promising results in patients with malignant melanoma. Similarly, antibodies to OAcGD2 and fucosyl-GM1 have shown anti-tumour effects *in vitro*, and studies with human patients are underway.

Other diseases: Aberrant production of the ganglioside GM3 has been linked to pathophysiological changes associated with obesity, metabolic disorders, and type 2 diabetes mellitus through its effects on insulin receptors. It has a role in autoimmune disorders such as multiple sclerosis. In epilepsy, it is believed that a deficiency in the enzyme ceramide synthase 1, which produces 18:0 ceramides, leads to reduced ganglioside formation. By their presence in certain subsets of T cells, gangliosides influence allergic responses and auto-immune diseases. As gangliosides are present on the surface of vascular, vascular-associated, and inflammatory cells, they may have a role in atherosclerosis and in aging.

21.5: BIOSYNTHESIS OF CHOLESTEROL AND STEROIDS

By William (Bill) W. Christie and Henry Jakubowski.

This section is an abbreviated and modified version of material from the Lipid Web, an introduction to the chemistry and biochemistry of individual lipid classes, written by William Christie.

21.5.1: STEROLS: CHOLESTEROL AND CHOLESTEROL ESTERS

In animal tissues, **cholesterol** (cholest-5-en-3β-ol) is by far the most abundant member of a family of polycyclic compounds known as **sterols**. It can also be described as a polyisoprenoid or a triterpene from its biosynthetic origin. Cholesterol was first recognized as a component of gallstones as long ago as 1769, while the great French lipid chemist Chevreul isolated it from animal fats in 1815. However, it was well into the 20th century before the structure was fully defined by the German Chemist Heinrich Wieland, who received the Nobel Prize in Chemistry for his work in 1927, the first of thirteen so honored for research on cholesterol and its metabolism.

Cholesterol plays a vital role in animal life, and it is essential for the normal functioning of cells both as a structural component of cell membranes and as a precursor of steroid hormones and other key metabolites including vitamin D and bile acids. It is also important for cell signaling, transport processes, nerve conduction and the regulation of gene transcription. Every cell in vertebrates is able both to synthesize cholesterol and to metabolize it, and there is evidence that synthesis *de novo* is essential whatever the dietary intake; this is vital in the brain. However, excess cholesterol can contribute to the pathology of various diseases, notably cardiovascular disease, so cholesterol levels must be balanced to ensure an adequate but not excessive supply.

21.5.1.1: CHOLESTEROL – STRUCTURE, OCCURRENCE, AND FUNCTION IN MEMBRANES

The struture of cholesterol is shown below in Figure 21.5.1.

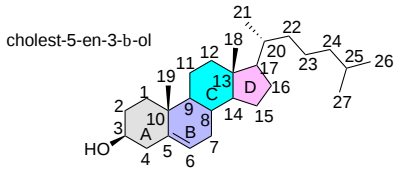


Figure 21.5.*xx*:

Cholesterol consists of a tetracyclic cyclopenta[a]phenanthrene structure with an *iso*-octyl side-chain at carbon 17. The four rings (A, B, C, D) have *trans* ring junctions, and the side chain and two methyl groups (C-18 and C-19) are at an angle to the rings above the plane with β stereochemistry (as for the hydroxyl group on C-3 also); there is a double bond between carbons 5 and 6. Thus, the molecule has a rigid planar four-ring nucleus with a flexible tail. Of the two recognized numbering systems in use, one originally described by Fieser and Fieser in 1959 and a second by IUPAC-IUB in 1989, the first appears to be preferred by most current authors.

Most of the atoms in cholesterol can be placed on the diamond lattice, which is the structure showing the position of each carbon atom in the network solid diamond. In that structure, each carbon is connected to four other carbons using sp^3-hybridized atomic orbitals and tetrahedral geometry. A representation of the diamond lattice is shown below. It consists of a series of interconnected boat conformations of cyclohexane propagating in the xyz direction. The structures of two different fused cyclohexanes, trans- and cis-decalins as well as the structure of three fused cyclohexanes, adamantane, are shown in Figure 21.5.2. Superimposing sp^3-connected atoms onto a diamond lattice allows improved visualization of the real and allowed structures of more complicated molecules.

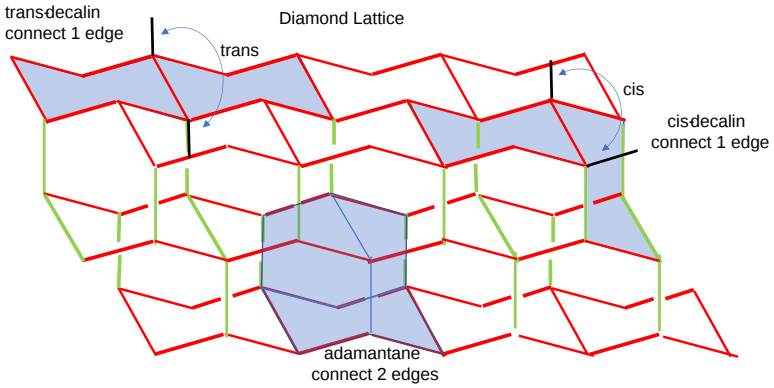


Figure 21.5.2: The diamond lattice showing decalins and adamantane

Figure 21.5.3 shows a superposition of 5-α-cholestane, a reduced and non-hydroxylated form of cholesterol, onto the diamond lattice. The pink cyclopentane D ring of 5-α-cholestane, which is distorted from the 109^0 bond angle for sp^3-hybridized carbon atoms, does not fit

exactly onto the lattice. Cholesterol, with its double bond, would also deviate from the ideal position on the diamond lattice given the two sp^2-hybridized carbons in the double bond.

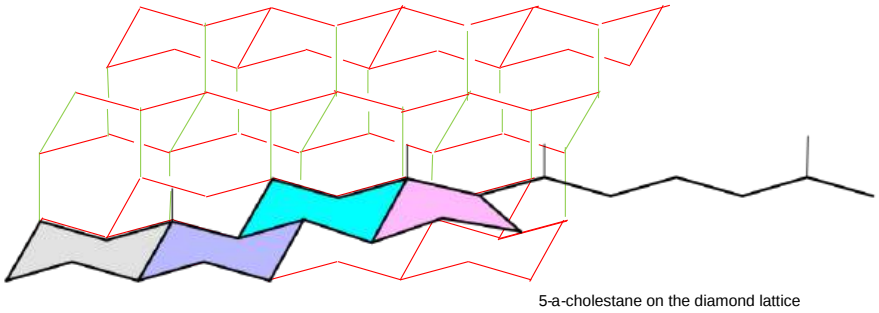


Figure 21.5.3: The diamond lattice showing 5-α-cholestane

Cholesterol is a ubiquitous component of all animal tissues (and of some fungi), produced by every nucleated animal cell, where much of it is located in the membranes, although it is not evenly distributed. The highest proportion of unesterified cholesterol is in the plasma membrane (roughly 30-50% of the lipid in the membrane or 60-80% of the cholesterol in the cell), while mitochondria and the endoplasmic reticulum have much less (~5% in the latter), and the Golgi contains an intermediate amount. Cholesterol is also enriched in early and recycling endosomes, but not in late endosomes. It may surprise some to learn that the brain contains more cholesterol than any other organ, where it comprises roughly a quarter of the total free cholesterol in the human body, 70-80% of which is in the myelin sheath. Of all the organic constituents of blood, only glucose is present in a higher molar concentration than cholesterol. In animal tissues, it occurs in the free form, esterified to long-chain fatty acids (**cholesterol esters**), and in other covalent and non-covalent linkages, including an association with the **plasma lipoproteins**. In plants, it tends to be a minor component only of a complex mixture of structurally related '**phytosterols**', although there are exceptions but it is nevertheless importance as a precursor of some plant hormones.

Animals in general synthesize a high proportion of their cholesterol requirement, but they can also ingest and absorb appreciable amounts from foods. On the other hand, many invertebrates, including insects, crustaceans and some molluscs cannot synthesize cholesterol and must receive it from the diet; for example, spiny lobsters must obtain exogenous cholesterol to produce essential sex hormones. Similarly, it must be supplied from exogenous sources to the primitive nematode *Caenorhabditis elegans*, where it does not appear to have a major role in membrane structure, other than perhaps in the function of ion channels, although it is essential the production of steroidal hormones required for larval development; its uptake is regulated by the novel lipid **phosphoethanolamine glucosylceramide**. Some species are able to convert dietary plant sterols such as β-sitosterol to cholesterol. Prokaryotes lack cholesterol entirely with the exception of some pathogens that acquire it from eukaryotic hosts to ensure their intracellular survival (e.g., *Borrelia sp.*); bacterial **hopanoids** are often considered to be sterol surrogates.

Cholesterol has vital structural roles in membranes and in lipid metabolism in general with an extraordinary diversity of biological roles, including cell signaling, morphogenesis, lipid digestion and absorption in the intestines, reproduction, stress responses, sodium and water balance, and calcium and phosphorus metabolism, and we can only touch on a few of these functions in this web page. It is a biosynthetic precursor of bile acids, vitamin D, and steroid hormones (glucocorticoids, estrogens, progesterones, androgens, and aldosterone), and it is found in covalent linkage to specific membrane proteins or **proteolipids** ('hedgehog' proteins), which have vital functions in embryonic development. In addition, it contributes substantially to the development and working of the central nervous system. On the other hand, excess cholesterol in cells can be toxic, and a complex web of enzymes is essential to maintain the optimum concentrations. Because plasma cholesterol levels can be a major contributory factor to atherogenesis, media coverage has created what has been termed a 'cholesterophobia' in the population at large.

One of the main functions of cholesterol is to modulate the fluidity of membranes by interacting with their complex lipid components, specifically the phospholipids such as phosphatidylcholine and sphingomyelin. As an amphiphilic molecule, cholesterol is able to intercalate between phospholipids in lipid bilayers to span about half a bilayer. In its three-dimensional structure, it is in essence a planar molecule that can interact on both sides. The tetracyclic ring structure is compact and very rigid. In addition, the location of the hydroxyl group facilitates the orientation of the molecule in a membrane bilayer, while the positions of the methyl groups appear to maximize interactions with other lipid constituents. The structure of cholesterol as it would appear on the diamond lattice is shown in Figure 21.5.4.

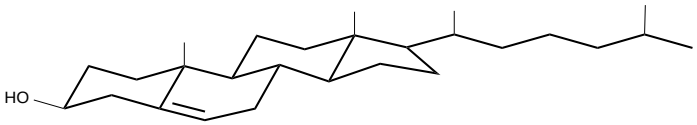


Figure 21.5.4: "Diamond" lattice structure of cholesterol

As the α-face of the cholesterol nucleus (facing down) is 'smooth', it can make good contact with the saturated fatty-acyl chains of phospholipids down to about their tenth methylene group; the β-face (facing up) is made 'rough' by the projection of methyl groups from carbons 10 and 13. The interaction is mainly via van der Waals and hydrophobic forces with a contribution from hydrogen bonding of the cholesterol hydroxyl group to the polar head group and interfacial regions of the phospholipids, especially sphingomyelin. Intercalated cholesterol may also disrupt electrostatic interactions between the ionic phosphocholine head groups of nearby membrane phospholipids, leading to increased mobility of the head groups. Indeed, there is evidence that cholesterol forms stoichiometric complexes with the saturated fatty acyl groups of sphingomyelin and to a lesser extent of phosphatidylcholine.

Experiments with mutant cell lines and specific inhibitors of cholesterol biosynthesis suggest that an equatorial hydroxyl group at C-3 of sterols is essential for the growth of mammalian cells. The Δ5 double bond ensures that the molecule adopts a planar conformation, and this feature also appears to be essential for cell growth, as is the flexible *iso*-octyl side-chain. The C-18 methyl group is crucial for the proper orientation of the sterol. While plant sterols appear to be able to substitute for cholesterol in supporting many of the bulk properties of membranes in mammalian cells *in vitro*, cholesterol is essential for other purposes.

In the absence of cholesterol, a membrane composed of unsaturated lipids is in a fluid state that is characterized by a substantial degree of lipid chain disorder, i.e., it constitutes a *liquid-disordered* phase. The function of cholesterol is to increase the degree of order (cohesion and packing) in membranes, leading to the formation of a *liquid-ordered* phase. In contrast, it renders bilayers composed of more saturated lipids, which would otherwise be in a solid gel state, more fluid. Thus, cholesterol is able to promote and stabilize a liquid-ordered phase over a substantial range of temperatures and sterol concentrations. Further, high cholesterol concentrations in membranes reduce their passive permeability to solutes. These effects enable membranes to bend or withstand mechanical stresses, and they permit the fine-tuning of membrane lipid composition and organization, and regulate critical cell functions. Simplistically, the higher cholesterol concentrations in the plasma membrane support its barrier function by increasing membrane thickness and reducing its permeability to small molecules. In contrast, the endoplasmic reticulum has increased membrane flexibility because of its lower cholesterol concentrations and thus enables the insertion and folding of proteins in its lipid bilayer. While mitochondrial membranes have a low cholesterol content in total, this may be concentrated in nanodomains at regions of high curvature in the inner mitochondrial membrane with links to nucleoprotein complexes (nucleoids).

In comparison to other lipids, it has been reported that cholesterol can flip rapidly between the leaflets in a bilayer, although this does not appear to be accepted universally, leading to doubts as to the trans-bilayer distribution of cholesterol in some biological membranes. However, much recent evidence suggests that the concentration of cholesterol in the inner leaflet of the plasma membrane is much lower than that in the outer leaflet in a range of mammalian cells. This distribution is important in that cholesterol promotes negative curvature of membranes and may be a significant factor in bringing about membrane fusion in the process of exocytosis. It may also be relevant for the regulation of various cellular signaling processes at the plasma membrane.

Cholesterol also has a key role in the lateral organization of membranes and their free volume distribution, factors permitting more intimate protein-cholesterol interactions that may regulate the activities of membrane proteins. Many membrane proteins bind strongly to cholesterol, including some that are involved in cellular cholesterol homeostasis or trafficking, and contain a conserved region termed the ‘sterol-sensing domain’. Some proteins bind to cholesterol deep within the hydrophobic core of the membrane via binding sites on the membrane-spanning surfaces or in cavities or pores in the proteins, driven by hydrogen bond formation. Cholesterol has an intimate interaction with G-protein-coupled receptors (GPCRs) to affect ligand binding and activation, either by direct high-affinity binding to the receptor, by changing their oligomerization state, or by inducing changes in the properties of the membrane. For example, it is essential for the stability and function of the β2-adrenergic, oxytocin and serotonin receptors by increasing the agonist affinities, while the inactive state of rhodopsin is stabilized both through indirect effects on plasma membrane curvature and by a direct interaction between lipid and protein. The GPCR neurotransmitter serotonin$_{1A}$ receptor has ten closely bound cholesterol molecules, and these control its organization and positioning; the receptor senses membrane cholesterol via a lysine residue in a so-called 'CRAC' motif in transmembrane helix 2.

Ion pumps such as the (Na^+-K^+)-ATPase, which have specific binding sites for cholesterol molecules, are the single most important consumer of ATP in cells and are responsible for the ion gradients across membranes that are essential for many cellular functions; depletion of cholesterol in the plasma membrane deactivates these ion pumps. In the brain in addition to being essential for the structure of the myelin sheath, cholesterol is a major component of synaptic vesicles and controls their shape and functional properties. In the nucleus of cells, cholesterol is intimately involved in chromatin structure and function.

The role of cholesterol together with sphingolipids in the formation of the transient membrane nano-domains known as **rafts** (see the specific web page for detailed discussion), is of crucial importance for the function of cells, while the interaction of cholesterol with **ceramides** is essential for the barrier function of the skin.

21.5.1.2: CHOLESTEROL BIOSYNTHESIS

Cholesterol biosynthesis involves a highly complex series of at least thirty different enzymatic reactions, which were unraveled in large measure by Konrad Bloch and Fyodor Lynen, who received the Nobel Prize for their work on the topic in 1964. When the various regulatory, transport, and genetic studies of more recent years are taken into account, it is obvious that this is a subject that cannot be treated in depth here. The bare bones of mechanistic aspects are therefore delineated, which with the references listed below should serve as a guide

to further study. In **plants**, cholesterol synthesis occurs by a somewhat different pathway with cycloartenol as the key intermediate. We'll explore the reaction mechanisms for several of the enzymes on this complicated pathway given its medical importance.

Almost all nucleated cells are able to synthesize their full complement of cholesterol. The first steps involve the synthesis of the important intermediate mevalonic acid from acetyl-CoA and acetoacetyl-CoA, both of which are in fact derived from acetate, in two enzymatic steps. These precursors are in the cytosol as is the first enzyme, **3-hydroxy-3-methyl-glutaryl(HMG)-CoA synthase**. The second enzyme **HMG-CoA reductase** is a particularly important control point, and is widely regarded as the rate-limiting step in the overall synthesis of sterols; its activity is regulated at the transcriptional level and by many more factors including a cycle of phosphorylation-dephosphorylation. This and subsequent enzymes are membrane-bound and are located in the endoplasmic reticulum. The enzyme HMG-CoA reductase is among the targets inhibited by the drugs known as 'statins' so that patients must then obtain much of their cholesterol from the diet. The first two reactions in the synthesis of cholesterol are shown in Figure 21.5.5.

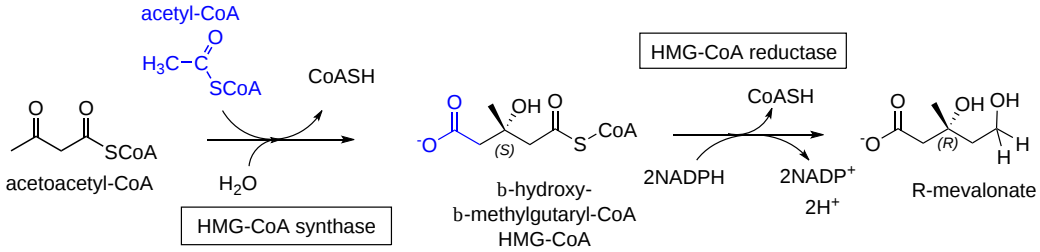


Figure 21.5.5: First two reaction in the synthesis of cholesterol

HMG-CoA Synthase

We saw this reaction in the synthesis of ketone bodies in Chapter 17. This enzyme catalyzes the condensation of acetoacetyl–CoA (AcAc–CoA) and acetyl–CoA (Ac-CoA) to form 3-hydroxy-3-methylglutaryl (HMG)–CoA, and requires the formation of a C-C bond. HMG-CoA synthase forms a C-C bond by activating the methyl group of **acetyl-cysteine**. The acetyl group comes from an acyl-CoA "donor". The enzyme catalyzes the first committed step in the formation of complex isoprenoids (like cholesterol) and ketone bodies. The product, 3-hydroxy-3-methylglutaryl HMG–CoA, can either be reduced by HMG-CoA reductase to form mevalonate, which leads to cholesterol synthesis, or cleaved by the enzyme HMG-CoA lyase, to produce acetoacetate, a ketone body.

HMG-CoA synthase catalyzes a bisubstrate reaction that displays ping-pong kinetics, characteristic of a covalent enzyme intermediate. The first substrate binds to the enzyme and transfers an acetyl group to a nucleophilic Cys 111 in the active site to form an acetyl-Cys 111 intermediate. Free CoA departs. Next the second substrate, acetoacetyl-CoA binds, and condenses with the acetyl group donated by acetyl-Cys 111. This condensation involves an enolate. A plausible reaction mechanism is shown in Figure 21.5.6.

Figure 21.5.6: Reaction mechanism for HMG-CoA synthase

Hence there are three parts of the reaction: acetylation/deacetylation, condensation/cleavage with an enolate intermediate, and C-C formation and hydrolysis/dehydration. Figure 21.5.7 shows an interactive iCn3D model of the Staphylococcus aureus HMG-COA Synthase with bound HMG-CoA and acetoacetyl-CoA (1XPK)

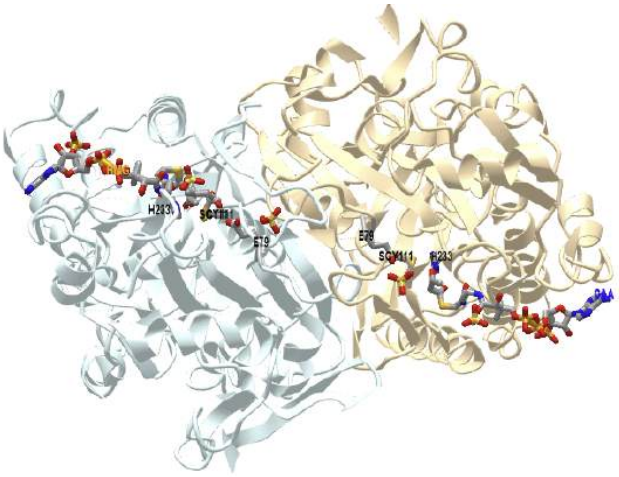


Figure 21.5.7: Staphylococcus aureus HMG-COA Synthase with bound HMG-CoA and acetoacetyl-CoA (1XPK). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...deiFs7JceE5H76

The biologically active unit (homodimer) is shown with each monomer shown in a different color. The A chain (light cyan) has bound HMG-CoA (HMG) while the B chain (light gold) has acetoacetyl-CoA (CAA) bound. The Glu 79, Cys 111, and His 233 in each subunit are shown in CPK sticks and labeled. Note that the Cys 111 is covalently modified in each subunit.

Figure 21.5.8 shows an interactive iCn3D model of the human 3-hydroxy-3-methylglutaryl CoA synthase I with bound CoASH (2P8U)

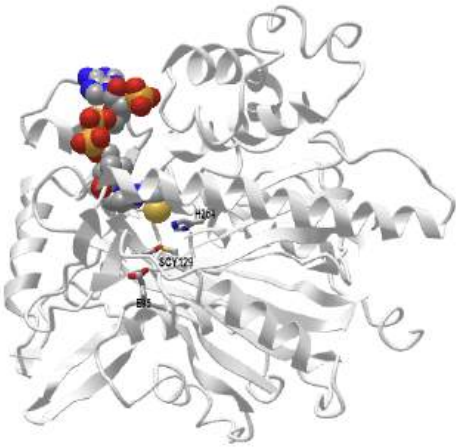


Figure 21.5.8: Human 3-hydroxy-3-methylglutaryl CoA synthase I (monomer) with bound CoASH (2P8U). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...xxEoUfecTk3eL6

The active site side chains are numbered differently (Glu 95, Cys 129, and His 264) compared to the S. aurus protein. Only the monomer is shown in this model.

HMG-CoA Reductase

In this key rate-limiting step that commits HMG to the sterol and isoprenoid synthetic pathways, HMG is converted to mevalonate. A reducing agent (NADPH) is required for this biosynthetic reaction. A plausible mechanism is shown in Figure 21.5.9.

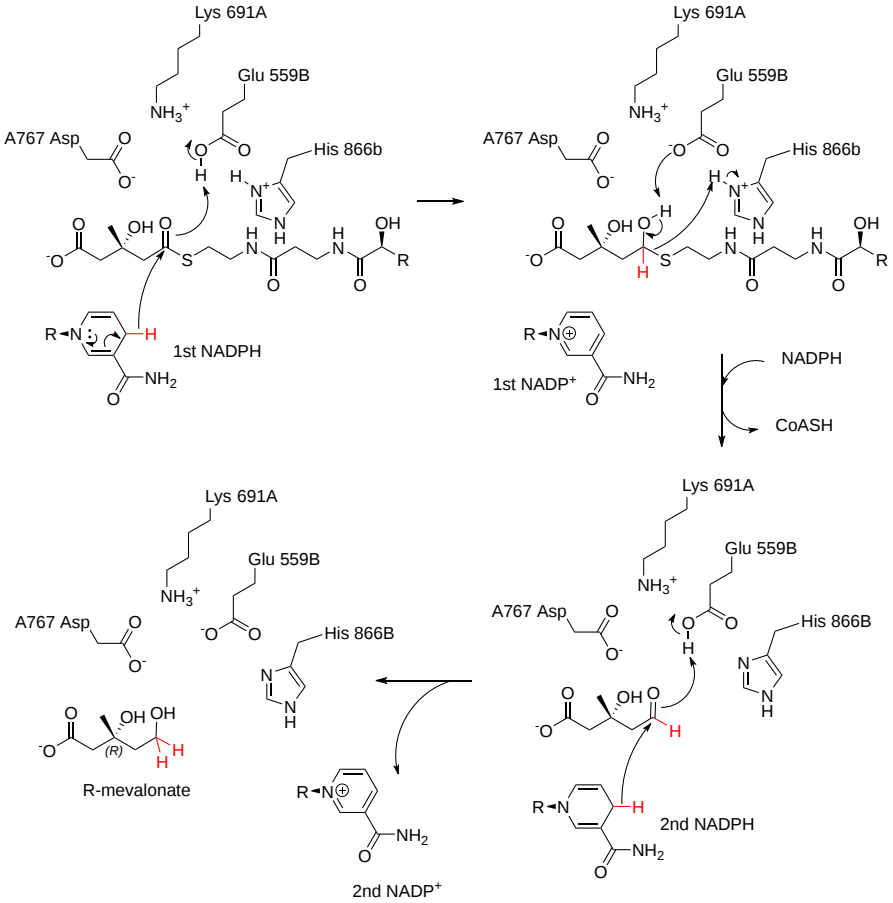
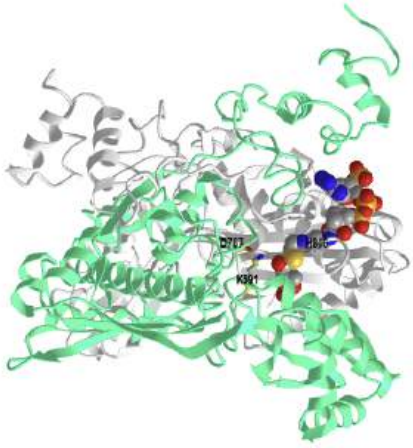


Figure 21.5.9: Reaction mechanism for HMG-CoA reductase. Adapted from https://www.ebi.ac.uk/thornton-srv/m-csa/entry/93/. Creative Commons Attribution 4.0 International (CC BY 4.0) License.

Figure 21.5.10 shows an interactive iCn3D model of the catalytic domain of human HMG-CoA reductase with bound HMG-CoA (1DQ9).



NCBI iCn3D Figure 21.5.10: Catalytic domain of human HMG-CoA reductase with bound HMG-CoA (1DQ9). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...dZvWzTocKhUzt8

Two of the four identical monomers in the biological unit, the C (gray) and D (cyan) chains are shown. HMG-CoA is shown in spacefill and CPK colors. Three key catalytic residues, His 866 in the C chain (b chain in the mechanism above) and Lys 691 and Asp 767 in the D chain (A chain in the mechanism above) are shown in sticks, CPK colors and labeled.

Synthesis of 5-isopentenyl and 2-isopentenyl pyrophosphate

The next sequence of reactions involves first the phosphorylation of mevalonic acid by a mevalonate kinase to form the 5-monophosphate ester, followed by further phosphorylation to yield an unstable pyrophosphate, which is rapidly decarboxylated to produce 5-isopentenyl

pyrophosphoric acid, the universal isoprene unit. An isomerase converts part of the latter to 3,3-dimethylallyl pyrophosphoric acid. These reactions are shown in Figure 21.5.11.

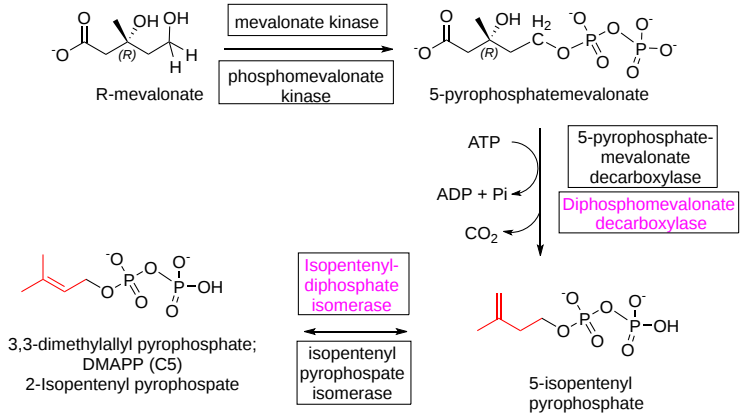


Figure 21.5.11: Reactions for the synthesis of 5-isopentenyl and 2-isopentenyl pyrophosphate for the synthesis of sterols and isoprenoids

Two phosphorylations are required, one by mevalonate kinase, which proceeds by an ordered sequential binding of mevalonate as the first reactant and its phosphomevalonate as the first product released. The enzyme is inhibited by two downstream products of the reaction pathway, farnesyl pyrophosphate, and geranyl pyrophosphate. A mechanism for mevalonate kinase is shown in Figure 21.5.12.

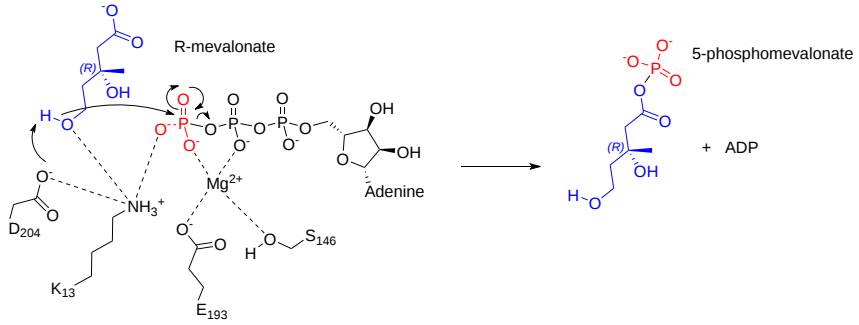


Figure 21.5.12: Mechanism of mevalonate kinase (adapted from Fu et al. JBC. 277 (2002). https://doi.org/10.1074/jbc.M200912200

A reaction mechanism for the second kinase, phosphomevalonate kinase, showing progression through the transition state, is shown in Figure 21.5.13. The reaction proceeds through direct phosphorylation through a dissociative mechanism.

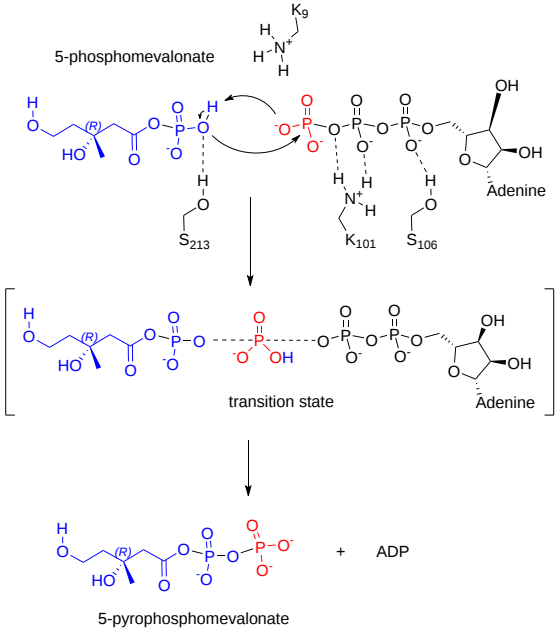


Figure 21.5.13: Reaction mechanism for phosphomevalonate kinase (after Huang et al. J. Phys. Chem. B 2016, 120. DOI: 10.1021/acs.jpcb.6b08480)

Mevalonate diphosphate decarboxylase

This enzyme catalyzes the decarboxylation of 5-pyrophosphate mevalonate to 5-isopentenylpyrophosphate as shown in Figure 21.5.14.

Figure 21.5.14: Conversion of 5-pyrophosphaemevalonate to 5-isopentyl pyrophosphate

The binding of the two substrates, pyrophosphatemevalonate (or mevalonate pyrophosphate - MVAPP) and ATP to the enzyme (A), and a reaction mechanism (B, are shown in Figure 21.5.15.

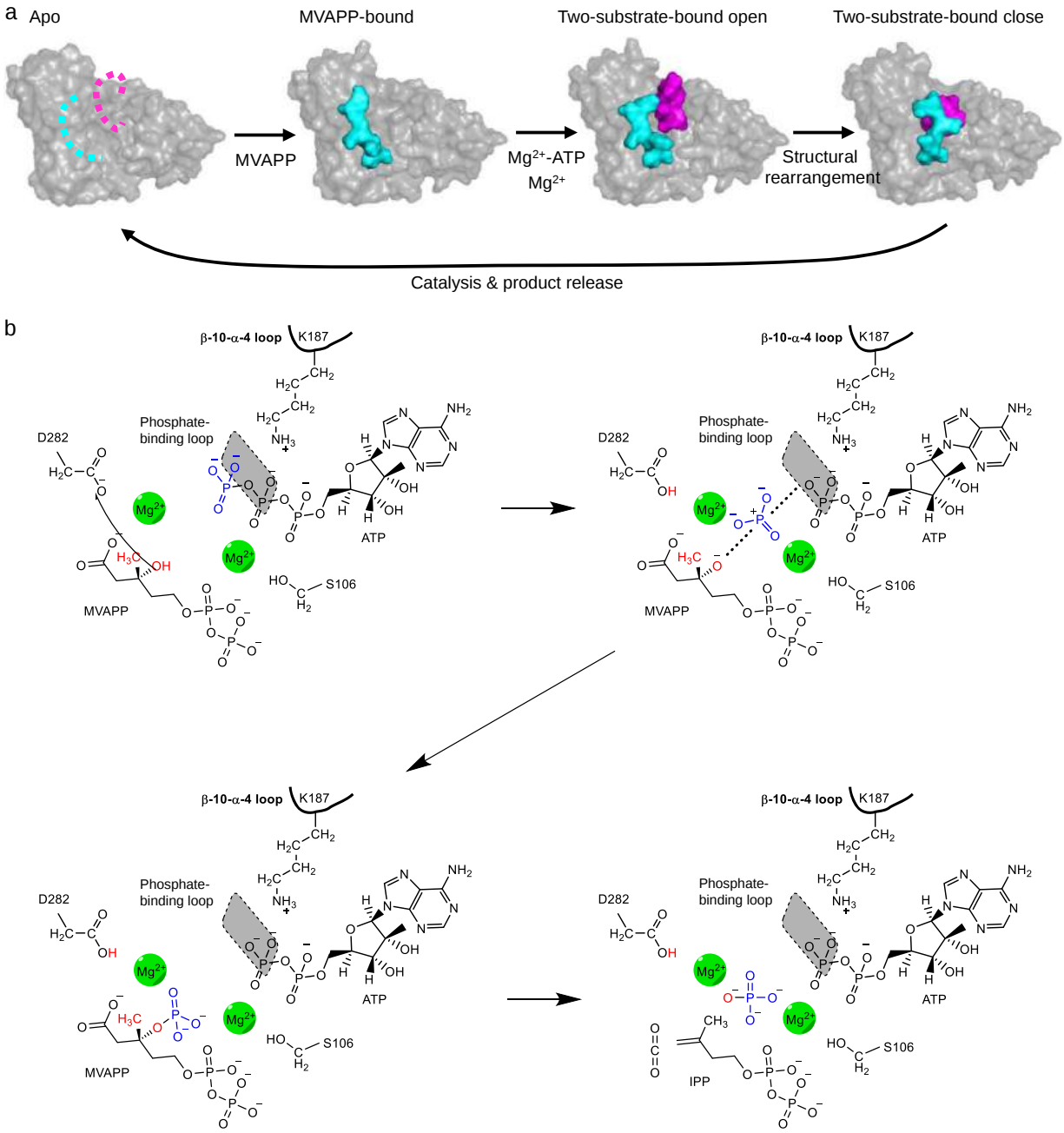


Figure 21.5.15: Proposed model for the detailed MDD enzyme mechanism. Chen et al. Nature Communications. (2020) 11:3969 |https://doi.org/10.1038/s41467-020-17733-0. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Panel (a) shows changes in the enzyme structure upon substrate binding. The apo-enzyme is shown on the left. The middle structure shows the enzyme after binding of the first substrate, MVAPP. A key loop (β10-α4) is shown as a cyan surface. The enzyme is then shown in an open conformation with the second substrate (ATP) also bound. An additional phosphate-biding loop is shown in magenta. At the far right, the enzyme is shown in a closed conformation after conformational changes in both loops which traps substrates in the active site. These changes enable catalysis. Product release follows.

Panel (b) shows the dissociative phosphoryl transfer catalytic mechanism. At the top-left, D282 is shown interacting with the 3′-OH group of MVAPP (red). The top-right shows a dissociative phosphoryl (blue) transition state. In the bottom-left, the phosphate attaches to the 3′ oxygen (red) of MVAPP. The bottom-right shows the products after the dephosphorylation and decarboxylation to produce IPP, ADP, phosphate, and CO2. K187 from the β10-α4 loop and metal ions in the active site are involved in neutralizing the negatively charged environment and assists catalysis.

Figure 21.5.16 shows an interactive iCn3D model of mevalonate diphosphate decarboxylase with mevalonate-5-diphosphate, AMPPCP and Magnesium (6E2U).

NCBI iCn3D Figure 21.5.16: Mevalonate diphosphate decarboxylase with mevalonate-5-diphosphate, AMPPCP and Magnesium (6E2U) . (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...SukmP1BTyNhH26

The correctly positioned substrates interact with two Mg^{2+} ions in the initial steps of the reaction. The conserved lysine facilitates the phosphoryl transfer.

21.5.1.3: POLYMERIZATION OF ISOPRENE

Isoprene, a small branched alkadiene, which can polymerize into larger molecules containing isoprene monomer to form isoprenoids, often called terpenes. Instead of using isoprene as the polymerization monomer, either dimethylally pyrophosphate (DMAPP) or isopentenylpyrophosphate (IPP) are used biologically.

Figure 21.5.x below shows how DMAPP and IPP (both containing 5Cs) are used in a polymerization reaction to form geranyl-pyrophosphate (C10), farnesyl pyrophosphate (C15) and geranyl-geranyl pyrophosphate (C20). Figure 21.5.17:

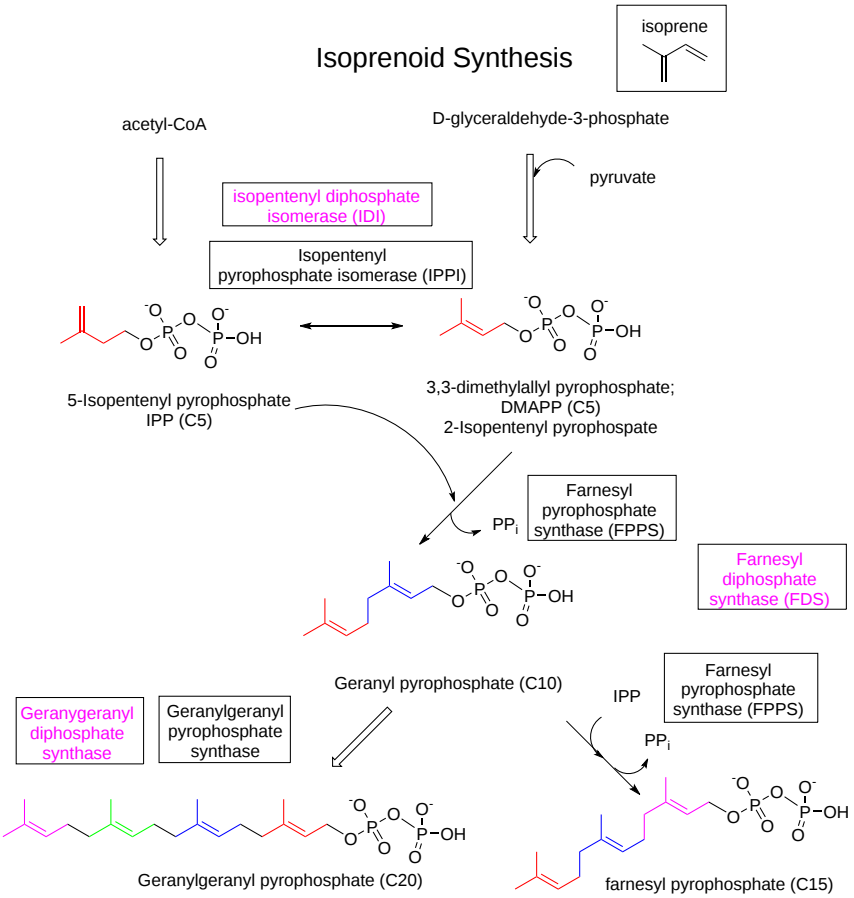


Figure 21.5.17: Synthesis of geranyl-pyrophosphate (C10), farnesyl pyrophosphate (C15) and geranyl-geranyl pyrophosphate (C20)

The condensation of IPP and DMAPP is a head-to-tail condensation reaction. Another IPP reacts with geranylpyrophosphate using the same enzyme to produce farnesylpyrophosphate. DMAPP is first formed by the isomerization of an IPP to DMAPP catalyzed by isopentenyl-diphosphate delta-isomerase. It catalyzes the 1,3-allylic rearrangement of the homoallylic substrate isopentenyl (IPP). 5-isopentenyl pyrophosphate is a nucleophile, but its isomer, DMAPP, is highly electrophilic, which promotes the condensation of the two molecules.

A mechanism for the next reaction, the first condensation of DMAPP and IPP to form geranylpyrophosphate by farnesyl pyrophosphate (diphosphate) synthase reaction, is shown in Figure 21.5.18.

Figure 21.5.18 Mechanism for the condensation of DMAPP and IPP to form geranylpyrophosphate by farnesyl pyrophosphate (diphosphate) synthase

The reaction appears to proceed using a carbocation transition state, followed by the transfer of a hydrogen atom (not ion) from IPP to pyrophosphate. Figure 21.5.19 shows an interactive iCn3D model of E. Coli farnesyl pyrophosphate synthase bound to isopentyl pyrophosphate and dimethylallyl S-thiolodiphosphate (1RQI).

Figure 21.5.19: Farnesyl pyrophosphate synthase Bound to isopentyl pyrophosphate (IPP) and dimethylallyl S-thiolodiphosphate (1RQI). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...kQBU5dNnKrCVt6

DST in the structure is dimethylallyl S-thiolodiphosphate, an analog of dimethylallyl diphosphate (DMAPP). Key conserved amino acids involved in substrate binding, transition state stabilization, and catalysis are shown as sticks, CPK colors. Thr203, Gln241, and Lys202 presumably stabilized the carbocation intermediate/transition state. Lys202 also forms a hydrogen bond with DMAPP. Two arginines (69, 116) form salt bridges with the pyrophosphates of IPP and DMAPP.

Many isoprenoid lipids are made from farnesyl pyrophosphate. For membrane purposes, the most important of these is cholesterol. Figure 21.5.20 shows an overview of the synthesis of cholesterol from two farnesyl pyrophosphates linking together in a "tail-to-tail" reaction to form squalene, a precursor of cholesterol. Each isoprene unit (5Cs) is shown in different colors to make it easier to see.

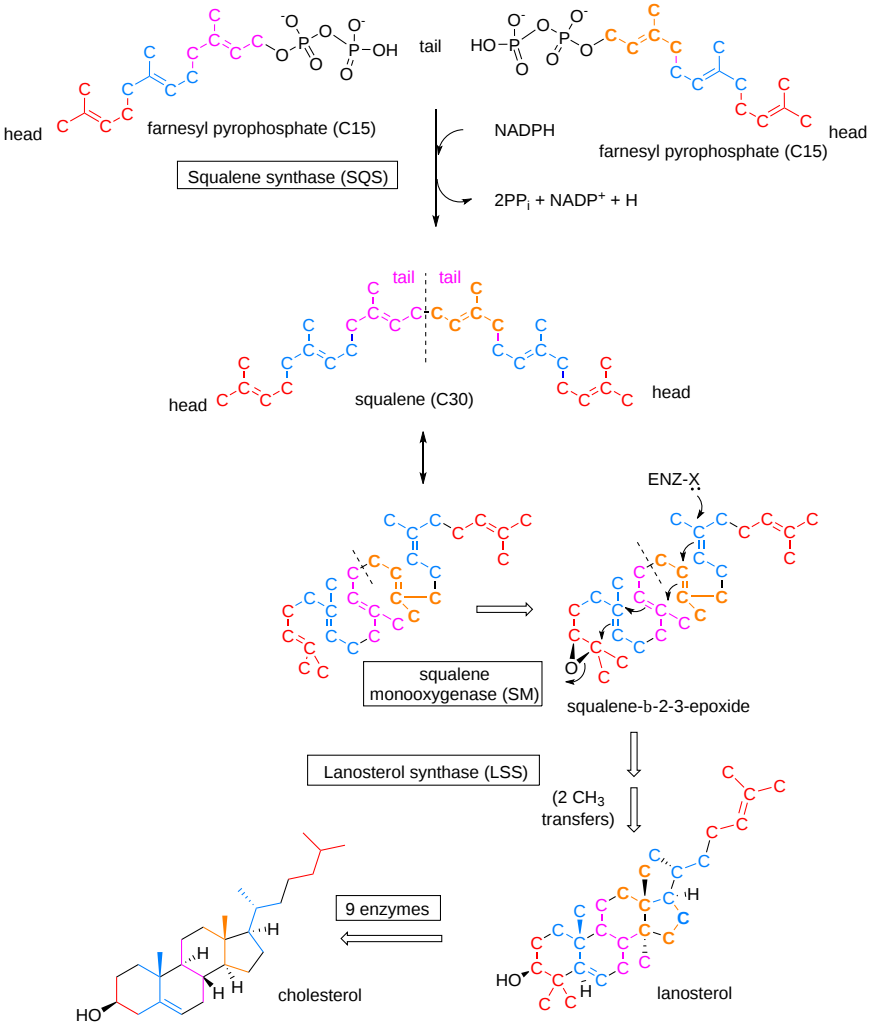


Figure 21.5.20: Synthesis of squalene from isoprene units Figure 21.5.21 shows reactants (two farnesyl pyrophosphate), intermediate (presqualene diphosphate) and product (squalene) in the reaction catalyzed by squalene synthase.

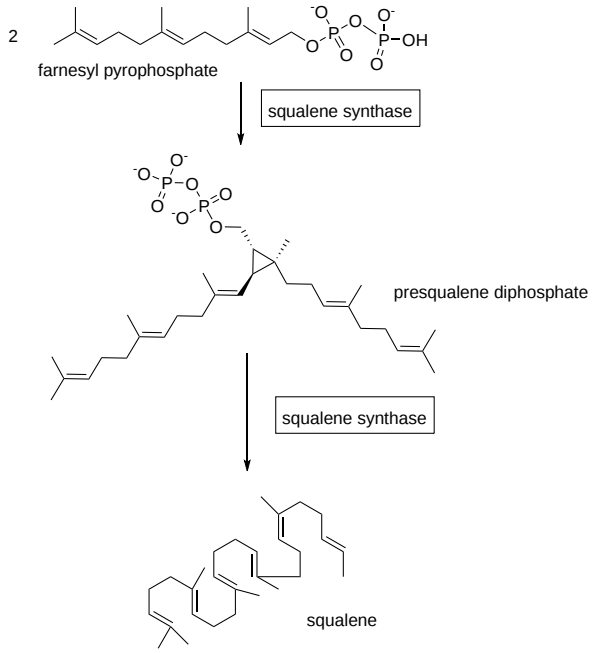


Figure 21.5.21:

In the squalene synthase reaction, two molecules of farnesyl pyrophosphate condense to yield presqualene pyrophosphate. In turn, this is reduced by NADPH to produce the key intermediate **squalene**. The enzyme squalene synthase, which regulates the flow of metabolites into either the sterol or non-sterol pathways (with farnesyl pyrophosphate as the branch point), is considered to be the first committed enzyme in cholesterol biosynthesis.

Given the importance of this reaction, we will explore the unique mechanism of squalene synthase in some detail.

Part 1: Formation of the cyclopropyl presqualene intermediate

Figure 21.5.22 shows the mechanism for the first part of the reaction in which the cyclopropyl intermediate presqualene form. The reaction proceeds through a series of carbocation intermediates.

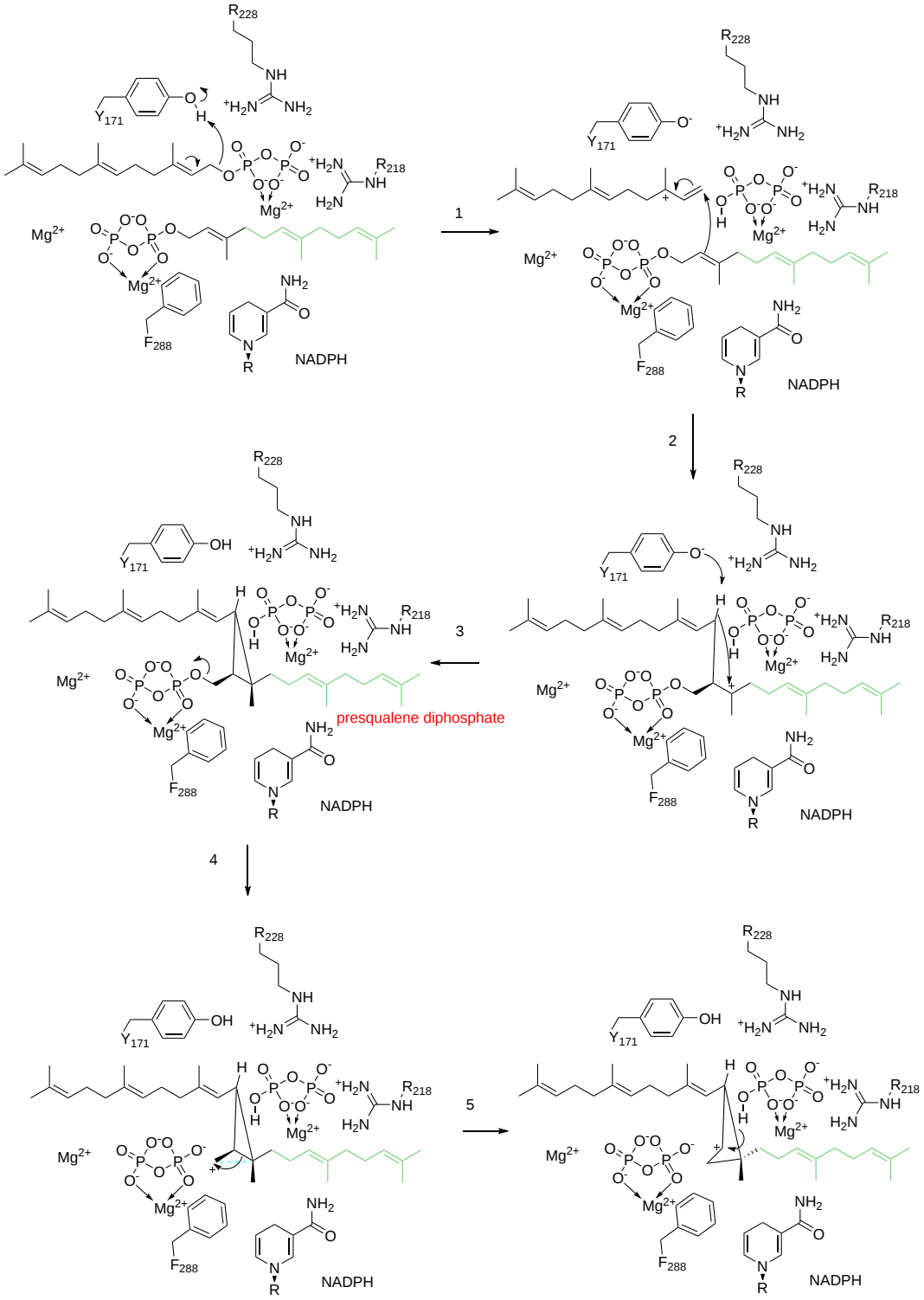


Figure 21.5.22 Mechanism for the synthesis of presqualene by squalene synthase. https://www.ebi.ac.uk/thornton-srv/m-csa/entry/264/. Creative Commons Attribution 4.0 International (CC BY 4.0) License

Part 2: Conversion of cyclopropyl intermediate to squalene - reduction

The next part of the reaction involves the reductive formation of squalene, as shown in Figure 21.5.23.

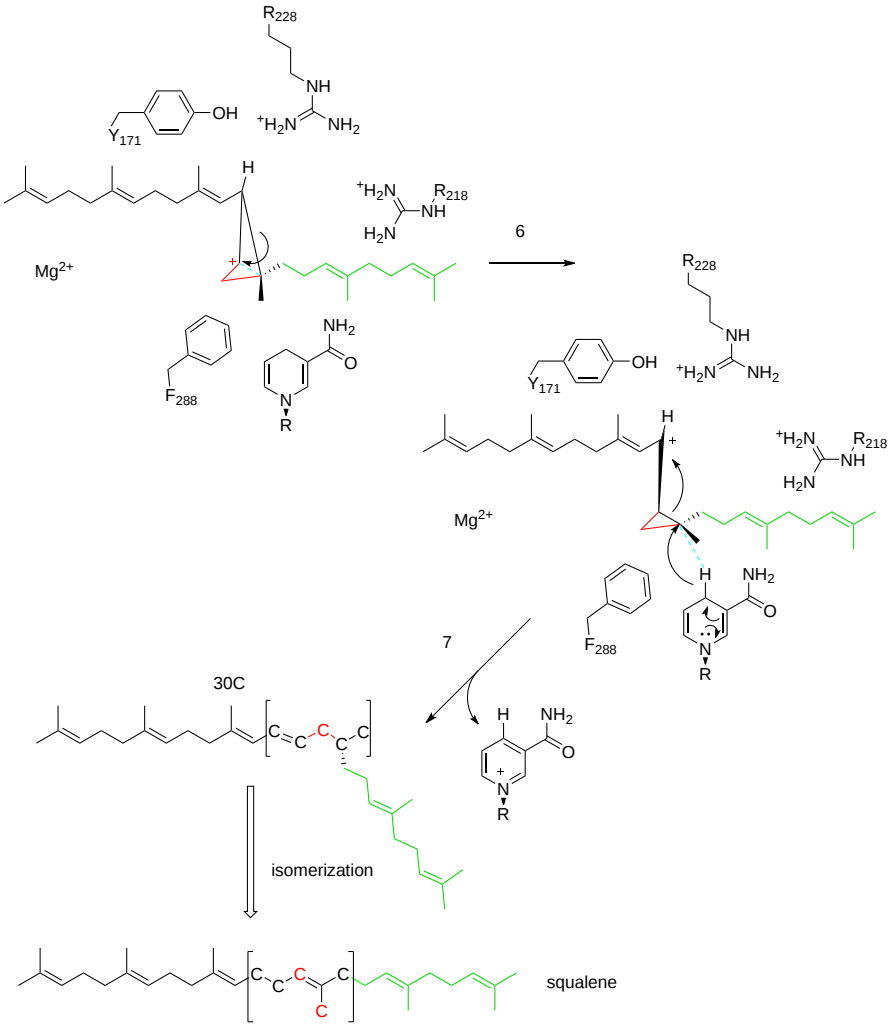
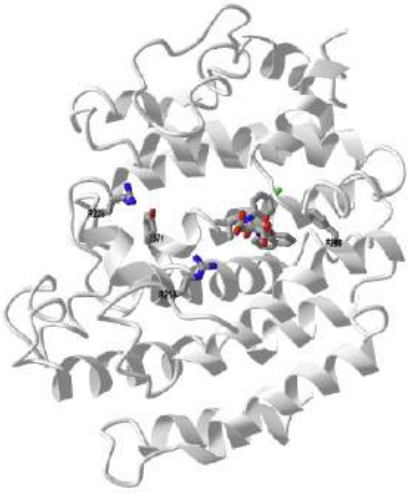


Figure 21.5.23: Mechanism for conversion of presqualene to squalene by squalene synthase. https://www.ebi.ac.uk/thornton-srv/m-csa/entry/264/. Creative Commons Attribution 4.0 International (CC BY 4.0) License

Figure 21.5.24 shows an interactive iCn3D model of human squalene synthase with bound inhibitor (1EZF).



NCBI iCn3D Figure 21.5.24: Human squalene synthase with bound inhibitor (1EZF). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...jLreHkvbftoFPA

Tyr 171 acts as a general acid/base in the first half of the reaction. Arg 218 and 228 stabilize the diphosphate in the transition state as it leaves. Phe 288 stabilizes the reactive carbocation.

The enzyme has a single domain that contains a large channel in one face of the enzyme which leads from a solvent-exposed to the hydrophobic interior. Two FPPs bind in the beginning of the channel where key side chains for the first reaction are located. The cyclopropyl intermediate then moves into the hydrophobic end of the channel where it reacts in the second half of the reaction without exposure to water.

In the next important step, squalene is oxidized by a squalene monooxygenase to squalene 2,3-epoxide, a key control point in the cholesterol synthesis pathway. This introduces the oxygen atom to squalene that becomes the signature oxygen of the hydroxyl group in cholesterol. The epoxide then undergoes cyclization catalyzed by the enzyme squalene epoxide lanosterol-cyclase to form the first steroidal intermediate lanosterol (or cycloartenol en route to **phytosterols** in photosynthetic organisms). This is illustrated in Figure 21.5.26.

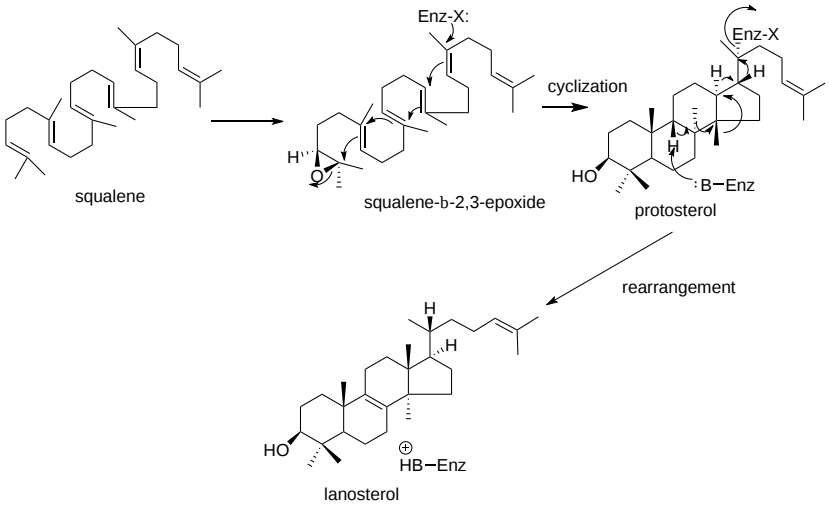


Figure 21.5.26: Conversion of squalene to lanosterol

In this remarkable reaction, there is a series of concerted 1,2-methyl group and hydride shifts along the chain of the squalene molecule to bring about the formation of the four rings. No intermediate compounds have been found. This is believed to be one of the most complex single enzymatic reactions ever to have been identified, although the enzyme involved is only 90 kDa in size. Again, the reaction takes place in the endoplasmic reticulum, but a cytosolic protein, sterol carrier protein 1, is required to bind squalene in an appropriate orientation in the presence of the cofactors NADPH, flavin adenine dinucleotide (FAD) and O_2; the reaction is promoted by the presence of phosphatidylserine.

The ring closure reaction starting with the epoxide involves a concerted flow of electrons from a source to the epoxide oxygen atom electron sink. This brings to mind a reaction known to all students who have ever studied chemistry, the reaction of the pH indicator phenolphthalein with a base to produce a pink colored-solution. That reaction which produces a more conjugated molecule that absorbs at 553 nm (green) which causes a magenta solution color, is illustrated for comparison (and fun) in Figure 21.5.27.

Figure 21.5.27: Reaction of phenolphthalein with base

In subsequent steps, lanosterol is converted to cholesterol by a series of demethylations, desaturations, isomerizations, and reductions, involving nineteen separate reactions as illustrated in Figure 21.5.28.

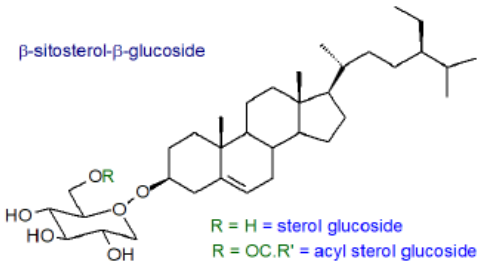


Figure 21.5.28: Production of cholesterol from lanosterol

Thus, demethylation reactions produce zymosterol as an intermediate, and this is converted to cholesterol via a series of intermediates, all of which have been characterized, and by at least two pathways that utilize essentially the same enzymatic machinery but differ in the order of the various reactions, mainly at the point at which the Δ24 double bond is reduced. Desmosterol is the key intermediate in the so-called 'Bloch' pathway, while 7-dehydrocholesterol is the immediate precursor in the 'Kandutsch-Russell' pathway. While some tissues, such as adrenal glands and testis, use the Bloch pathway mainly, the brain synthesizes much of its cholesterol by the 'Kandutsch-Russell' pathway. This may enable the production of a variety of other minor sterols for specific biological purposes in different cell types/locations.

The energy cost of the synthesis of one cholesterol molecule is roughly one hundred ATP equivalents, and eleven oxygen molecules are required. Synthesis occurs mainly in the liver, although the brain (see below), peripheral nervous system, and skin synthesize their own considerable supplies. Cholesterol is exported from the liver and transported to other tissues in the form of low-density **lipoproteins** (LDL) for uptake via specific receptors. In animals, cells can obtain the cholesterol they require either from the diet via the circulating LDL, or they can synthesize it themselves as outlined above. Cholesterol biosynthesis is highly regulated with rates of synthesis varying over hundreds of fold depending on the availability of any external sources of cholesterol, and cholesterol homeostasis requires the actions of a complex web of enzymes, transport proteins, and membrane-bound transcription factors, as discussed below.

21.5.1.4: REGULATION OF CHOLESTEROL HOMEOSTASIS

In humans, only about a third of the cholesterol is of dietary origin (mainly eggs and red meat), the remainder is produced by synthesis *de novo* in the endoplasmic reticulum. The latter must be tightly regulated as it is an energetically expensive process that requires appreciable amounts of acetyl-CoA, ATP, oxygen, and the reducing factors NADPH and NADH, especially since cholesterol cannot be catabolized for energy purposes (see below).

Many factors are involved in maintaining the large differences in cholesterol concentrations among the various membranes and organelles in cells within precise limits. In order to explain how cholesterol in the plasma membrane, where it is most abundant, can regulate cholesterol biosynthesis and uptake through enzymes in the endoplasmic reticulum, where it is least abundant, it has been suggested that a key to the process is that there are three pools of cholesterol in the plasma membrane with distinct functional roles. The first of these is “accessible” to receptor proteins for transport to the endoplasmic reticulum, while the second pool is sequestered by sphingomyelin and can be released by the action of sphingomyelinase if required. The third residual pool of cholesterol is essential for plasma membrane integrity. These correspond to about 16, 15, and 12 mol % of total plasma membrane lipids, respectively, in cholesterol-replete cells. Simplistically, when cholesterol in the plasma membrane is in excess for any reason, e.g., after LDL uptake by receptor-mediated endocytosis, there is a rise in accessible cholesterol, which is then transported to the endoplasmic reticulum to switch off cholesterol biosynthesis and expression of the LDL receptor. This process requires a host of regulatory proteins and mechanisms that can involve either vesicle formation or non-vesicular pathways that utilize specific transport proteins, such as the ABC transporters.

Ultimately, post-translational control of the many different enzymes involved provides a rapid means for modifying flux through the biosynthetic pathway in the endoplasmic reticulum; some are rapidly degraded in response to tissue levels of cholesterol and its intermediates, while others have their activity altered through phosphorylation or acetylation mechanisms. For example, the second rate-limiting enzyme in cholesterol biosynthesis is squalene monooxygenase, which undergoes cholesterol-dependent proteasomal degradation when cholesterol is in excess, guided by a 12-amino acid hydrophobic sequence on the enzyme that can serve as a degradation signal. When the cholesterol concentration in the endoplasmic reticulum is high, the degradation sequence detaches from the membrane and is exposed to provide the signal for the enzyme to be degraded. Similarly, HMG-CoA reductase is recognized as the key enzyme in the regulation of cholesterol biosynthesis, and this can be regulated by a feedback mechanism involving ubiquitin–proteasomal degradation. Further regulation of cholesterol biosynthesis is exerted by sterol intermediates in cholesterol biosynthesis, such as lanosterol and 24,25-dehydrolanosterol (dimethyl-sterols) by accelerating degradation of the biosynthetic enzymes such as HMG-CoA reductase. It is noteworthy that ceramide down-regulates cholesterol synthesis – another link between cholesterol and sphingolipid metabolism.

The regulatory element-binding proteins (mainly SREBP-1c and SREBP-2), which contain an N-terminal membrane domain and a C-terminal regulatory domain, are essential to the maintenance of cholesterol levels. Each is synthesized as an inactive precursor that is inserted into the endoplasmic reticulum where it can encounter an escort protein termed **SREBP cleavage-activating protein (SCAP)**, which is the cellular cholesterol sensor. When the latter recognizes that cellular cholesterol levels are inadequate, it binds to the regulatory domain of SREBP. The SCAP-SREBP complex then moves to the Golgi, where two specific proteases (designated site-1 and site-2 proteases) cleave the SREBP enabling the *C*-terminal regulatory domain to enter the nucleus. There it activates transcription factors, such as the nuclear liver X receptor (LXR), which stimulate the expression of the genes coding for the LDL receptor in the plasma membrane and for the key enzyme in cholesterol biosynthesis, HMG-CoA reductase. This in turn stimulates the rate of cholesterol uptake and synthesis. Conversely, when cholesterol in the endoplasmic reticulum exceeds a threshold, it binds to SCAP in such a way that it prevents the SCAP-SREBP complex from leaving the membrane for the nucleus, cholesterol synthesis and uptake are thereby repressed, and cholesterol homeostasis is restored. In effect, cholesterol exerts feedback inhibition by suppressing its own production by preventing the proteolytic cleavage and maturation of SREBP-2. **Oxysterols**, especially 25-hydroxycholesterol, are also inhibitors of this process.

Cholesterol in the endoplasmic reticulum is transferred to the Golgi and eventually to the plasma membrane by vesicular and non-vesicular transport mechanisms involving in part soluble sterol transport proteins, including the so-called 'START' domain proteins, and partly by binding to those proteins that are intimately involved in the transport and metabolism of polyphosphoinositides such as **phosphatidylinositol 4-phosphate (PI(4)P)**. In the latter mechanism, cholesterol is transported by binding to the ORD domain of oxysterol binding protein (OSBP) or Osh4 in yeast, before OSBP binds to PI(4)P in the plasma membrane to transfer its cargo. The key to this process is that cholesterol and PI(4)P are synthesized at two different locations, i.e., the endoplasmic reticulum for sterols and the trans-Golgi network and plasma membrane for PI(4)P, so the two lipids do not compete but rather can be exchanged. OSBP carries cholesterol in the forward direction to the trans-Golgi network and plasma membrane and PI(4)P, which binds to a C-terminal PH domain in the protein, in the

reverse direction. The subsequent hydrolysis of PI(4)P is the energy source for the reaction, and indeed PI(4)P has been termed "lipid ATP". As this reaction is irreversible, a gradient of cholesterol along organelles of the secretory pathway is established. OSBP is thus a lipid transfer protein that enables two organelles to exchange cholesterol rapidly between them at membrane contact sites in a cycle of reactions involving membrane tethering, cholesterol transport, PI(4)P counter transport, and PI(4)P hydrolysis. A similar mechanism is involved in the transport of phosphatidylserine from the endoplasmic reticulum to the inner leaflet of the plasma membrane.

Subsequently, the ATP binding cassette (ABC) transporters ABCA1 and ABCG1 in the plasma membrane, which contains much of the cellular cholesterol, are activated to export the excess. Nuclear factor erythroid 2 related factor-1 or NRF1 in the endoplasmic reticulum binds directly to cholesterol and senses when its level is high to bring about a de-repression of genes involved in cholesterol removal, also with mediation by the liver X receptor. Also, side-chain oxysterols, especially 25-hydroxycholesterol, can suppress the activation of SREBP by binding to an oxysterol-sensing protein in the endoplasmic reticulum.

Within cells, cholesterol derived initially from the lysosomal degradation of low-density lipoproteins is transferred first to the plasma membrane and thence to the endoplasmic reticulum, the latter step by a mechanism involving proteins known as GRAMD1s embedded in the endoplasmic reticulum membrane at sites in contact with the plasma membrane. These have two functional domains: the START-like domain that binds cholesterol and the GRAM domain that binds anionic lipids, such as phosphatidylserine, and so are able to form a link between the two membranes that enable the transfer of cholesterol.

In peripheral tissues, excess cholesterol is exported to high-density lipoproteins (HDL) in the circulation and returned to the liver, a process known as **reverse cholesterol transport**. The liver is important for cholesterol synthesis, but it is essential for its elimination from the body in bile. Also, some lipoproteins with their content of cholesterol and cholesterol esters are delivered to lysosomes by endocytosis for degradation. The cholesterol is transported to the inner surface of the lysosomal membrane through the glycocalyx, via a transglycocalyx tunnel, with the aid of Niemann-Pick C1, C2, and other proteins, and thence via contact sites between membranes to other organelles. Cholesterol in cellular membranes in excess of the stoichiometric requirement can escape back into the cell, where it may serve as a feedback signal to down-regulate cholesterol accumulation, while some is converted to the relatively inert storage form, i.e., cholesterol esters, and some is used for steroidogenesis.

The intestines play a major part in cholesterol homeostasis via the absorption of dietary cholesterol and fecal excretion of cholesterol and its metabolites. A specific transporter (Niemann-Pick C1-like 1 or NPC1L1)in the brush border membrane of enterocytes in the proximal jejunum of the small intestine is involved in the uptake of cholesterol from the intestinal contents, while the metabolism of sterols in the intestines is controlled mainly by an acetyl-CoA acetyltransferase (ACAT2), which facilitates intracellular cholesterol esterification, and the microsomal triglyceride transfer protein (MTTP), which is involved in the assembly of chylomicrons for export into lymph. Some cholesterol can be transferred in the opposite direction (trans-intestinal cholesterol excretion), but the quantitative importance of this process is not clear. There is evidence that dietary or synthesized cholesterol is necessary to maintain intestinal integrity, as cholesterol derived from circulating lipoproteins is not sufficient for the purpose.

In the intestines and especially the colon, the intestinal microflora are able to hydrogenate cholesterol from bile, diet, and desquamated cells to form coprostanol with an efficiency that is dependent on the composition of microbial species. Coprastanol is not absorbed by the intestinal tissue to a significant extent, and it may inhibit the uptake of residual cholesterol. There are two mechanisms for this conversion in bacteria, one involving direct reduction and another via cholestenone and coprostanone as intermediates, and as the relevant genes have now been identified the therapeutic potential is under investigation.

Brain: There are substantial differences in cholesterol synthesis and metabolism in brain in comparison to the liver and peripheral tissues. Trace amounts only of cholesterol are able to cross the blood-brain barrier via transport in low-density lipoproteins. Therefore, virtually all the cholesterol in the brain must be synthesized *de novo*, mainly in astrocytes (glial cells). During the perinatal and adolescent years especially, cholesterol is synthesized in large amounts to form the myelin that surrounds the axons, before this rate begins to decline to eventually reach about 10% of earlier values.

Cholesterol is transported to neurons in the form of Apo E complexes in discoidal HDL-like particles, for which seven main receptors have been identified in brain cells that take up cholesterol from these lipoproteins. Apo E is synthesized in the brain, and its transcription is regulated by 24-hydroxy-cholesterol concentrations. Similarly, in the brain and central nervous system, cholesterol synthesis is regulated independently of that in peripheral tissues, mainly by forms of the liver X receptor (LXR). As cholesterol and oxysterols are involved in providing neuroprotective effects and lowering neuroinflammation, dysregulation of their concentrations has been noted in many neurodegenerative disorders. Most of the lipoproteins in cerebrospinal fluid differ from the nascent poorly-lipidated HDL secreted by astrocytes, suggesting that the latter are modified during maturation.

21.5.1.5: CHOLESTEROL CATABOLISM

Cholesterol is not readily degraded in animal tissues so does not serve as a metabolic fuel to generate ATP. Only the liver possesses the enzymes to degrade significant amounts, and then via pathways that do not lead to energy production. Cholesterol and oxidized metabolites (oxysterols) are transferred back from peripheral tissues in lipoprotein complexes to the liver for catabolism by conversion to oxysterols and bile acids. The latter are exported into the intestines to aid in digestion, while leading to some loss that is essential for cholesterol

homeostasis. Until recently, it was believed that approximately 90% of cholesterol elimination from the body occurred via bile acids in humans. However, experiments with animal models now suggest that a significant amount is secreted directly into the intestines by a process known as trans-intestinal cholesterol efflux. How this occurs and its relevance to humans are under active investigation.

Gut bacteria reduce some of the cholesterol in the diet to highly insoluble 5β-cholestan-3β-ol (coprostanol), which is excreted and can be used as a biomarker for sewage in the environment. Certain bacterial species contain a 3β-hydroxysteroid:oxygen oxidoreductase (EC 1.1.3.6), commonly termed cholesterol oxidase, a flavoenzyme that catalyzes the oxidation of cholesterol to cholest-5-en-3-one which is then rapidly isomerized to cholest-4-en-3-one as the first essential step in the catabolism of sterols. The enzyme is widespread in organisms that degrade organic wastes, but it is also present in pathogenic organisms where it influences the virulence of infections (see below). In biotechnology, it has been used for the production of a number of steroids, and it is employed in a clinical procedure for the determination of cholesterol levels in serum.

21.5.1.6: CHOLESTEROL ESTERS

Cholesterol esters, i.e., with long-chain fatty acids linked to the hydroxyl group, are much less polar than free cholesterol and appear to be the preferred form for transport in plasma and as a biologically inert storage or detoxification form to buffer an excess. They do not contribute to membrane structures but are packed into intracellular lipid droplets. Cholesterol esters are major constituents of the adrenal glands, and they accumulate in the fatty lesions of atherosclerotic plaques. Similarly, esters of steroidal hormones are also present in the adrenal glands, where they are concentrated in cytosolic lipid droplets adjacent to the endoplasmic reticulum; 17β-estradiol, the principal estrogen in fertile women, is transported in lipoproteins in the form of a fatty acid ester.

Because of the mechanism of synthesis (see below), plasma cholesterol esters tend to contain relatively high proportions of the polyunsaturated components typical of phosphatidylcholine as shown in Table 21.5.1 below. Arachidonic and “adrenic” (20:4(n-6)) acids can be especially abundant in cholesterol esters from the adrenal gland.

Table 21.5.1. **Fatty acid composition of cholesterol esters (wt % of the total) from various tissues.**

Form	Fatty acids						
	16:0	18:0	18:1	18:2	18:3	20:4	22:4
Human							
plasma	12	2	27	45		8	
liver	23	10	28	22		6	
Sheep							
plasma	10	2	27	35	7	-	-
liver	17	9	29	7	4	3	-
adrenals	13	7	35	18	2	4	2

Data from - Christie, W.W. *et al. Lipids*, **10**, 649-651 (1975); DOI. Nelson, G.J. *Comp. Biochem. Physiol.*, **30**, 715-725 (1969); Horgan, D.J. and Masters, C.J. *Aust. J. Biol. Sci.*, **16**, 905-915 (1963); Nestel, P.J. and Couzens, E.A. *J. Clin. Invest.*, **45**, 1234-1240 (1966); DOI.

In plasma and in the high-density lipoproteins (HDL) in particular, cholesterol esters are synthesized largely by the transfer of fatty acids to cholesterol from position *sn*-2 of phosphatidylcholine (‘lecithin’) catalyzed by the enzyme **lecithin:cholesterol acyl transferase (LCAT)**; the other product is 1-acyl lysophosphatidylcholine. This is illustrated in Figure 21.5.29.

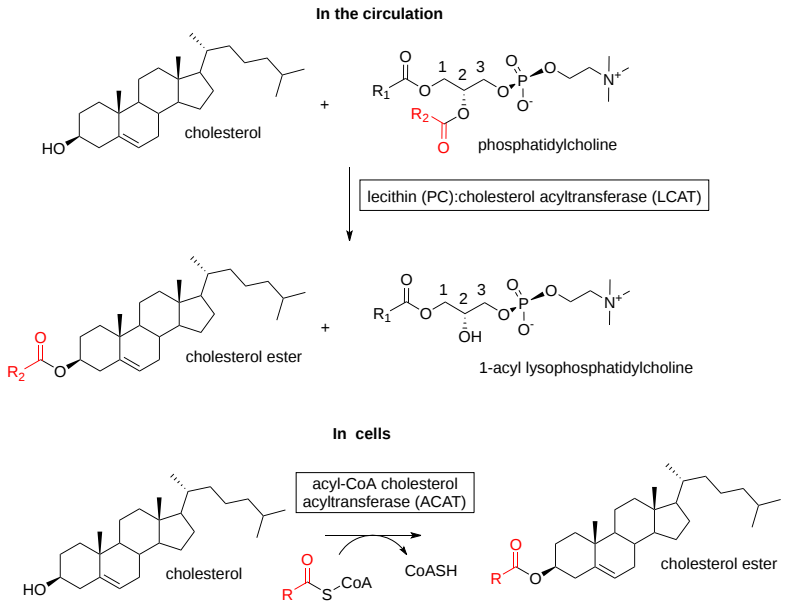


Figure 21.5.29 Synthesis of cholesterol esters

In fact, the reaction occurs in several steps. First, apoprotein A1 in the HDL acts to concentrate the lipid substrates near LCAT and present it in the optimal conformation; at the same time, it opens a lid on the enzyme that activates it by opening up the site of transesterification. Then, cleavage of the *sn*-2 ester bond of phosphatidylcholine occurs via the phospholipase activity of LCAT with the release of a fatty acyl moiety. This is transacylated to the sulfur atom of a cysteine residue forming a thioester, and ultimately it is donated to the 3β-hydroxyl group of cholesterol to form the cholesterol ester. Some LCAT activity has also been detected in apolipoprotein B100-containing particles (β-LCAT activity as opposed to α-LCAT with HDL).

It has been established that human LCAT is a relatively small glycoprotein with a polypeptide mass of 49 kDa, increased to about 60 kDa by four *N*-glycosylation and two *O*-glycosylation moieties. Most of the enzyme is produced in the liver and circulates in the bloodstream bound reversibly to HDL, where it is activated by the main protein component of HDL, apolipoprotein A1. As cholesterol esters accumulate in the lipoprotein core, cholesterol is removed from its surface thus promoting the flow of cholesterol from cell membranes into HDL. This in turn leads to morphological changes in HDL, which grow and become spherical. Subsequently, cholesterol esters are transferred to the other lipoprotein fractions LDL and VLDL, a reaction catalyzed by cholesterol ester transfer protein. This process promotes the efflux of cholesterol from peripheral tissues (‘reverse cholesterol transport’), especially from macrophages in the arterial wall, for subsequent delivery to the liver. LCAT is often stated to be the main driving force behind this process, and it is of great importance for cholesterol homeostasis and a suggested target for therapeutic intervention against cardiovascular disease.

The stereospecificity of LCAT changes with molecular species of phosphatidylcholine containing arachidonic or docosahexaenoic acid, when 2-acyl lysophosphatidylcholines are formed. This reaction may be especially important for the supply of these essential fatty acids to the brain in that such lysophospholipids are believed to cross the blood-brain barrier more readily than the free acids.

In other animal tissues, a further enzyme acyl-CoA:cholesterol acyltransferase (ACAT) synthesizes cholesterol esters from CoA esters of fatty acids and cholesterol. ACAT exists in two forms, both of which are intracellular enzymes found in the endoplasmic reticulum and are characterized by multiple transmembrane domains and a catalytic histidine residue in a hydrophobic domain; they are members of the *O*-acyltransferase (MBOAT) superfamily. ACAT1 is present in many tissues, but especially in macrophages and adrenal and sebaceous glands, which store cholesterol esters in the form of cytoplasmic lipid droplets; it is responsible for the synthesis of cholesterol esters in arterial foam cells in human atherosclerotic lesions. ACAT2 is found only in the liver and small intestine, and it is believed to be involved in the supply of cholesterol esters to the nascent lipoproteins. Analogous enzymes are found in yeast where ergosterol is the main sterol, but a very different process occurs in plants.

Oxidized Cholesterol Esters: All lipid classes containing polyunsaturated fatty acids are susceptible to oxidation. Under normal circumstances, cholesterol esters are considered to be relatively inert. However, when they contain oxidized polyunsaturated fatty acids, their properties change and they acquire biological activity. Such oxidized cholesterol esters may be formed by a reaction with 15-lipoxygenase, but they can be produced also through free radical-induced lipid peroxidation, and they have been detected in lipoproteins, LDL especially, in human blood and atherosclerotic lesions. Those oxidized cholesterol esters in plasma are trafficked into cells and metabolized by the same mechanisms as the corresponding unoxidized lipids.

Such "minimally oxidized LDL" do not bind to CD36 but rather to CD14, a receptor that recognizes bacterial lipopolysaccharides. The result is stimulation of toll-like receptor 4 (TLR4), although the response differs from that of lipopolysaccharides. In addition, oxidized

metabolites of cholesteryl arachidonate of this kind stimulate macrophages to express inflammatory cytokines of relevance to atherosclerosis among other effects. Oxidized cholesterol esters can be hydrolyzed to release their fatty acids, which can then be incorporated into phospholipids with a different repertoire of activities.

Hydrolysis of cholesterol esters: Cholesterol ester hydrolases in animals liberate cholesterol and free fatty acids when required for membrane and lipoprotein formation, and they also provide cholesterol for hormone synthesis in adrenal cells. Many cholesterol ester hydrolases have been identified, including a carboxyl ester hydrolase, a lysosomal acid cholesterol ester lipase, hormone-sensitive lipase, and hepatic cytosolic cholesterol ester hydrolase. These are located in many different tissues and organelles and have multiple functions. A neutral cholesterol ester hydrolase has received special study, as it is involved in the removal of cholesterol esters from macrophages so reducing the formation of foam cells and thence the development of fatty streaks within the arterial wall, a key event in the progression of atherosclerosis.

21.5.1.7: OTHER ANIMAL STEROLS

Cholesterol will oxidize slowly in tissues or foods to form a range of different products with additional hydroperoxy, epoxy, hydroxy or keto groups, and these can enter tissues via the diet. There is increasing interest in these from the standpoint of human health and nutrition since the accumulation of oxo-sterols in plasma is associated with inhibition of the biosynthesis of cholesterol and bile acids and with other abnormalities in plasma lipid metabolism.

A number of other sterols occur in small amounts in tissues, most of which are intermediates in the pathway from lanosterol to cholesterol, although some of them have distinct functions in their own right. **Lanosterol**, the first sterol intermediate in the biosynthesis of cholesterol, was first found in wool wax, both in free and esterified form, and this is still the main commercial source. It is found at low levels only in most other animal tissues (typically 0.1% of the cholesterol concentration). As oxygen is required, lanosterol cannot be produced by primitive organisms, hence its absence from prokaryotes, leading to some speculation on its evolutionary significance. When sterols became available to eukaryotes, much greater possibilities opened for their continuing evolution. The production of cholesterol from lanosterol is then seen as ‘molecular streamlining’ by evolution, removing protruding methyl groups that hinder the interaction between sterols and phospholipids in membranes.

Desmosterol (5,24-cholestadien-3β-ol), the last intermediate in the biosynthesis of cholesterol by the Bloch pathway, may be involved in the process of myelination, as it is found in relative abundance in the brains of young animals but not in those of adults, other than astrocytes. It is also found in appreciable amounts in testes and spermatozoa together with another cholesterol intermediate, testis meiosis-activating sterol. In addition, there is evidence that desmosterol activates certain genes involved in lipid biosynthesis in macrophages, and may deactivate others associated with the inflammatory response. There is a rare genetic disorder in which there is an impairment in the conversion of desmosterol to cholesterol, desmosterolosis, with serious consequences in terms of mental capacity. These and related sterols appear to be essential for human reproduction.

In human serum, the levels of **lathosterol** (5α-cholest-7-en-3β-ol) were found to be inversely related to the size of the bile acid pool, and in general, the concentration of serum lathosterol is strongly correlated with the cholesterol balance under most dietary conditions. The isomeric saturated sterols, **cholestanol,** and **coprastanol**, which differ in the stereochemistry of the hydrogen atom on carbon 5, are formed by microbial biohydrogenation of cholesterol in the intestines, and together with cholesterol are the main sterols in feces. Further examples of animal sterols include **7-dehydrocholesterol** (cholesta-5,7-dien-3β-ol) in the skin, which on irradiation with UV light is converted to **vitamin D_3** (cholecalciferol). These sterols are shown in Figure 21.5.30.

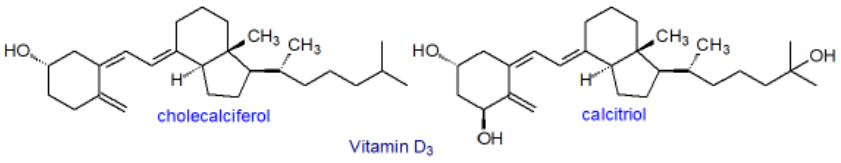


Figure 21.5.30: Other animal sterols

Marine invertebrates produce a large number of novel sterols, with both unusual nuclei and unconventional sidechains, some derived from cholesterol and others from plant sterols or alternative biosynthetic intermediates. For example, at least 80 distinct sterols have been isolated from echinoderms and 100 from sponges.

21.5.1.8: CHOLESTEROL AND DISEASE

Elevated cholesterol and cholesterol ester levels are associated with the pathogenesis of cardiovascular disease (atherosclerotic plaques, myocardial infarctions, and strokes), as is well known, and this is considered briefly on this website together with the metabolism of the **plasma lipoproteins**. The rate-limiting enzyme in the synthesis of cholesterol HMG-CoA reductase is the target of statins, but drugs that target other steps in the biosynthetic pathway, especially the squalene monooxygenase and lanosterol synthase, are under investigation. Further discussion of such a complex nutritional and clinical topic is best left to others better qualified than myself.

Cholelithiasis or the presence of 'stones' in the gallbladder or bile ducts, which consist largely of cholesterol (~85%), is one of the most prevalent and costly digestive diseases in developed countries. The primary cause is the excessive excretion of cholesterol from the liver. Excess accumulation of cholesterol associated with the metabolism of **bis(monoacylglycero)phosphate** and causing disturbances in glycosphingolipid trafficking in cell membranes is involved in the pathogenesis of **Niemann-Pick C disease**, a lysosomal storage disease in which endocytosed cholesterol becomes sequestered in late endosomes/lysosomes because of gene mutations affecting two binding proteins (NPC1 and NPC2) thereby causing neuronal and visceral atrophy. In addition, deficiencies in cholesterol transport and metabolism are associated with many forms of neurodegeneration, including Alzheimer's disease, Huntington's disease, and related conditions associated with old age. These proteins are also key virulence factors for several viral and bacterial pathogens.

Several genetic disorders of cholesterol biosynthesis have been identified in recent years that can result in developmental malformations including neurologic defects. As there is limited cholesterol transport across the placenta, the human fetus is highly dependent upon endogenous synthesis. While the molecular basis for the altered developmental pathways is not fully understood, impaired synthesis of the **hedgehog** family of signaling proteins, which require covalently linked cholesterol to function in membranes, is believed to be involved in many cases. In others, there are confirmed enzyme defects. For example, the recessive Smith-Lemli-Opitz syndrome in infants born with a decreased concentration of the enzyme 7-dihydrocholesterol reductase, produces symptoms varying from mild autism to severe mental and often fatal physical problems. The effects are due to a lack of cholesterol and the accumulation of 7-dehydrocholesterol and its 27-hydroxy metabolite, as brain tissue cannot utilize dietary cholesterol or that produced peripherally. In fact, at least eight different inherited disorders of cholesterol biosynthesis lead to congenital abnormalities in those afflicted. In animal models, deficiencies in SREBP-2 and genes encoding sterol biosynthetic enzymes display embryonic lethality. Deficiencies in the enzymes responsible for the hydrolysis of cholesterol esters, such as the lysosomal acid lipase, occur in Wolman disease and cholesterol ester storage disease.

Cholesterol and other sterols bind directly to several immune receptors, especially in macrophages and T cells, and dynamic changes in cholesterol biosynthesis impact directly upon innate and adaptive immune responses, such that functional coupling between sterol metabolism and immunity has implications for health and disease. For example, cholesterol binds directly to the αβ T cell antigen receptor (αβTCR) and has at least two opposing functions in its activation. By binding to the trans-membrane domain of this receptor, it is kept in an inactive, non-signaling conformation, but when required it can stimulate the formation of receptor nanoclusters to increase their avidity for the antigen. In cancer, there is a high demand for cholesterol in order to support the inherent nature of tumor cells to divide and proliferate, and perturbations of reverse cholesterol transport can have negative consequences. Drugs that lower cholesterol levels in cancer cells by inhibiting the mevalonate pathway are undergoing clinical trials.

When increased levels of sterols other than cholesterol are found in plasma, they usually serve as markers for abnormalities in lipid metabolism associated with disease states. For example, premature atherosclerosis and xanthomatosis occur in two rare lipid storage diseases, cerebrotendinous xanthomatosis, and sitosterolemia. In the former, cholestanol is present in all tissues, while in the latter, the dietary plant sterols campesterol and sitosterol accumulate in plasma and red blood cells. Inhibition of cholesterol biosynthesis may be associated with the appearance of some of the precursor sterols in the plasma.

In infections with *Mycobacterium tuberculosis*, the organism uses host cholesterol as the major carbon and energy source and thereby promotes persistent infection with appreciable effects on pathogenicity. Similarly, *Chlamydia trachomatis*, a gram-negative obligate intracellular bacterium and a major cause of sexually transmitted infections, requires host cholesterol for growth. Many viruses use cholesterol as part of their life cycle, and reduction in cellular cholesterol is sometimes seen as an anti-viral strategy, although this may not always be helpful. For example, an HIV protein has a binding site for cholesterol, which it utilizes to facilitate the fusion with raft regions in the membranes of the host cell.

21.5.2: STEROLS: 2. OXYSTEROLS AND OTHER CHOLESTEROL DERIVATIVES

Oxysterols as defined and discussed here are oxygenated derivatives of **cholesterol** and its precursors, i.e., with additional hydroxyl, epoxyl, or keto groups, that are found in all animal tissues. Many of these have vital functions in animals, while others are important as short-lived intermediates or end products in the catabolism or excretion of cholesterol or in the biosynthesis of steroid hormones, bile acids, and 1,25-dihydroxy-vitamin D_3. They are normally present in biological membranes and lipoproteins at trace levels only, though they can exert profound biological effects at these concentrations. However, they are always accompanied by a great excess (as much as 10^6-fold) of cholesterol *per se*.

A multiplicity of different oxysterols are synthesized in cells by sequential reactions with specific oxygenases. However, because of the presence of the double bond in the 5,6-position, oxysterols can also be formed rapidly by non-enzymatic oxidation (autoxidation) of cholesterol and cholesterol esters within tissues with the formation of many different oxygenated derivatives. Simplistically, non-enzymatic oxidation leads mainly to the generation of products in which the sterol ring system is oxidized, while enzymatic processes usually produce metabolites with an oxidized side chain (7-hydroxylation is an important exception). Oxidized cholesterol molecules can also be generated by the gut microflora and be taken up through the enterohepatic circulation. Once an oxygen function is introduced into cellular cholesterol, the product can act as a biologically active mediator by interacting with specific receptors before it is metabolized to **bile acids** or is degraded further, processes assisted by the fact that oxysterols are able to diffuse much more rapidly through membranes than is cholesterol

itself. Cholesterol metabolites of this kind are especially important in the brain, which is a major site for cholesterol synthesis *de novo*, and they are crucial elements of cholesterol homeostasis.

21.5.2.1: ENZYMATIC OXIDATION OF CHOLESTEROL

Within animal cells, the oxidation of sterols is mainly an enzymic process that is carried out by several enzymes that are primarily from the cytochrome P450 family of oxygenases (named for a characteristic absorption at 450 nm). These comprise a disparate group of proteins that contain a single heme group and have a similar structural fold, though the amino acid sequences can differ appreciably. They are all mono-oxygenases. Oxysterol biosynthesis can be considered in terms of different pathways that depend on the position of the initial oxidation, but these pathways tend to overlap and lead to a complex web of different oxysterols (and eventually to **bile acid** formation). As these enzymes, which include cytochrome P450, cholesterol hydroxylase, hydroxysteroid dehydrogenases, and squalene epoxidase, are specific to particular tissues and indeed animal species, there is considerable variation in oxysterol distributions between organs. A few examples only of the first steps in some of these pathways are illustrated in Figure 21.5.31.

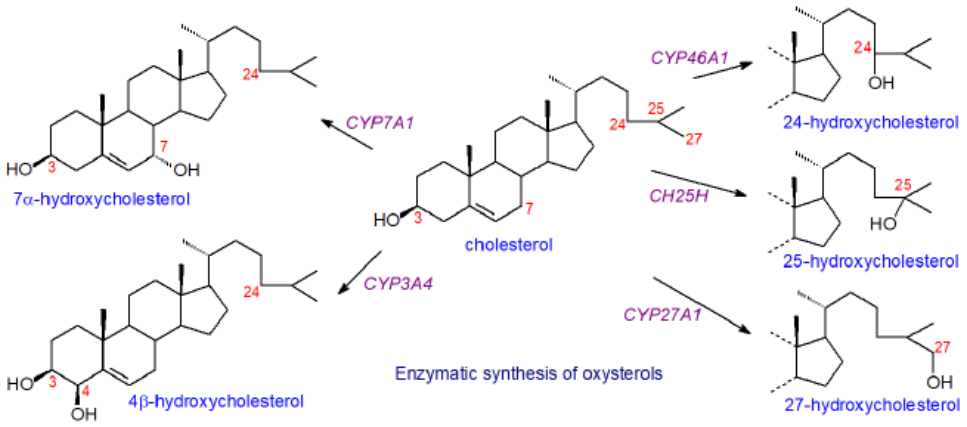


Figure 21.5.31: Enzymatic synthesis of oxysterols

As an example, a primary product is **7α-hydroxycholesterol**, which is an important intermediate in the biosynthesis of bile acids by the 'neutral' pathway and of many other oxysterols, and it is produced in the liver by the action of cholesterol 7α-hydroxylase (CYP7A1), an enzyme that has a critical role in cholesterol homeostasis. The reaction is under strict regulatory control, and the expression of CYP7A1 is controlled by the farnesoid X receptor (FXR) and is activated by cholic and chenodeoxycholic acids. Any circulating 7α-hydroxycholesterol represents leakage from the liver. Further oxidation of 7α-hydroxycholesterol can occur, and the action of CYP3A4 in humans generates 7α,25-dihydroxycholesterol as an important metabolite, for example, while oxidation by CYP27A1 yields 7α,27-dihydroxycholesterol; the latter is regarded as a key step in a further pathway to oxysterols and bile acids. On the other hand, the epimer **7β-hydroxycholesterol** is produced in the brain by the action of the toxic β-amyloid peptide and its precursor on cholesterol, but whether this is involved in the pathology of Alzheimer's disease has yet to be determined.

The hydroxysteroid 11-β-dehydrogenase 1 (HSD11B1) is responsible for the conversion of 7β-hydroxy-cholesterol to the important metabolite **7-keto-cholesterol**, while HS11B2 catalyzes the reverse reaction; 7-keto-cholesterol is also formed by autoxidation (see below). HSD11B1 is better known as the oxidoreductase that converts inactive cortisone to the active stress hormone cortisol in glucocorticoid target tissues.

An alternative ('acidic') pathway to bile acids starts with the synthesis of **27-hydroxycholesterol** (or more systematically named (25*R*)26-hydroxycholesterol), which is produced by the cytochrome P450 enzyme (CYP27A1) and introduces the hydroxyl group into the terminal methyl carbon (C27 or C26 - used interchangeably). While this enzyme is present in the liver, it is found in many extra-hepatic tissues and especially the lung, which provides a steady flux of 27-oxygenated metabolites to the liver. As a multifunctional mitochondrial P450 enzyme in the liver, it generates both 27-hydroxycholesterol and 3β-hydroxy-5-cholestenoic acid, the bile acid precursor, which occurs in small but significant amounts in plasma. 27-Hydroxycholesterol is the most abundant circulating oxysterol, and its concentration in plasma correlates with that of total cholesterol. It can be oxidized to 7α,27-dihydroxycholesterol by the enzyme CYP7B1. **4β-Hydroxycholesterol** is also abundant in plasma and is relatively stable; it is produced in humans by the action of the cytochromes CYP3A4 and CY3A5.

In humans, the specific cytochrome P450 that produces **24*S*-hydroxycholesterol** (cholest-5-ene-3β,24-diol) is cholesterol 24*S*-hydroxylase (CYP46A1) and is located almost entirely in the smooth endoplasmic reticulum of neurons in the brain, including those of the hippocampus and cortex, which are important for learning and memory. It is by far the most abundant oxysterol in the brain after parturition, but during development, many more many oxysterols are produced. 24*S*-hydroxycholesterol is responsible for 98-99% of the turnover of cholesterol in the central nervous system, which is the source of most of this oxylipin found in plasma. A small amount of it is converted in the brain directly into to 7α,24*S*-dihydroxycholesterol by the cytochrome CYP39A1 and thence via side-chain oxidation in peroxisomes to bile acids,

such as cholestanoic acid. It is evident that the blood-brain barrier is crossed by constant passive fluxes of oxysterols, but not of cholesterol *per se*, as a result of their permissive chemical structures and following their concentration gradients. In plasma, it is transported via high-density lipoproteins, as discussed further below. In contrast to humans, CYP46A1 is present in the liver of rodents as well as the brain.

25-Hydroxycholesterol is a relatively minor but biologically important cholesterol metabolite, which is produced rapidly by immune cells during the inflammation resulting from bacterial or viral infections. The dioxygenase enzyme cholesterol 25-hydroxylase (CH25H in humans), which utilizes a diiron cofactor to catalyze hydroxylation, is the most important route to this metabolite *in vivo*, although at least two cytochrome P450 enzymes, CYP27A1 and CYP3A4, can catalyze this conversion to a limited extent. Further oxidation by CYP7B1 is a second route to 7α,25-dihydroxycholesterol, and hence to further oxysterols.

24(*S*),25-Epoxycholesterol is not produced by the pathways described above but is synthesized in a shunt of the mevalonate pathway using the same enzymes that produce cholesterol, specifically squalene mono-oxygenase and lanosterol synthase, by means of which a second epoxy group is introduced on the other end of squalene from the initial epoxidation. A further mechanism in the brain is the action of CYP46A1 on desmosterol, another intermediate in cholesterol biosynthesis.

The oxysterols formed by both autoxidation and enzymatic routes can undergo further oxidation-reduction reactions, and they can be modified by many of the enzymes involved in the metabolism of cholesterol and steroidal hormones, such as esterification and sulfation of position 3, as illustrated for 7-keto-cholesterol as an example in Figure 21.5.32.

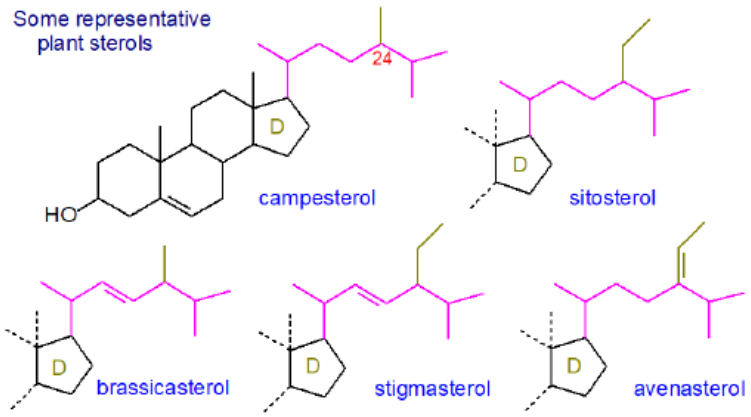


Figure 21.5.32 Metabolism of 7-keto-cholesterol

In most tissues, esterification of the 3β-hydroxyl group only occurs and requires the activity of sterol *O*-acyltransferases 1/2 (SOAT1/2 or ACAT1/2) with the participation of cytosolic phospholipase A_2 (cPLA$_2$α) to liberate the required fatty acids from phospholipids. In plasma, oxysterols can be esterified by the lecithin–cholesterol acyltransferase (LCAT) for transport in lipoproteins, but in this instance, a diester can be produced from 27-hydroxycholesterol specifically. Whether such esters are an inert storage form for oxysterols to be liberated on demand by esterases remains to be determined.

It is noteworthy that the important human pathogen, *Mycobacterium tuberculosis*, utilizes a cytochrome P450 enzyme (CYP125) to catalyze C26/C27 hydroxylation of cholesterol as an essential early step in its catabolism as part of the infective process.

Catabolism: Because of their increased polarity relative to cholesterol, oxysterols produced by both enzymatic and non-enzymatic means can exit cells relatively easily. A proportion is oxidized further and converted to **bile acids**, and some are converted to sulfate esters (especially at the 3-hydroxyl group) or glucuronides (see below) for elimination via the kidneys.

21.5.2.2: NON-ENZYMATIC OXIDATION OF CHOLESTEROL

In biological systems in which both cholesterol and fatty acids are present, it would be expected that autoxidation of polyunsaturated fatty acids by free radical mechanisms would be favored thermodynamically with the formation of isoprostanes from arachidonic acid in phospholipids. However, there are circumstances that can favor cholesterol oxidation *in vivo*, and, for example, the concentration of cholesterol in low-density lipoprotein particles (LDL) is about three times higher than that of phospholipids, and the rate of cholesterol-hydroperoxide formation can be higher than that of phospholipid hydroperoxides. The rate and specificity of the reaction can depend also on whether it is initiated by free radical species, such as those arising from the superoxide/hydrogen peroxide/hydroxyl radical system (Type I autoxidation), or whether it occurs by non-radical but highly reactive oxygen species such as singlet oxygen, HOCl or ozone (Type II autoxidation). As examples of the main types of products of non-enzymatic oxidation, the structures of a few of the more important of these oxysterols are illustrated in Figure 21.5.33.

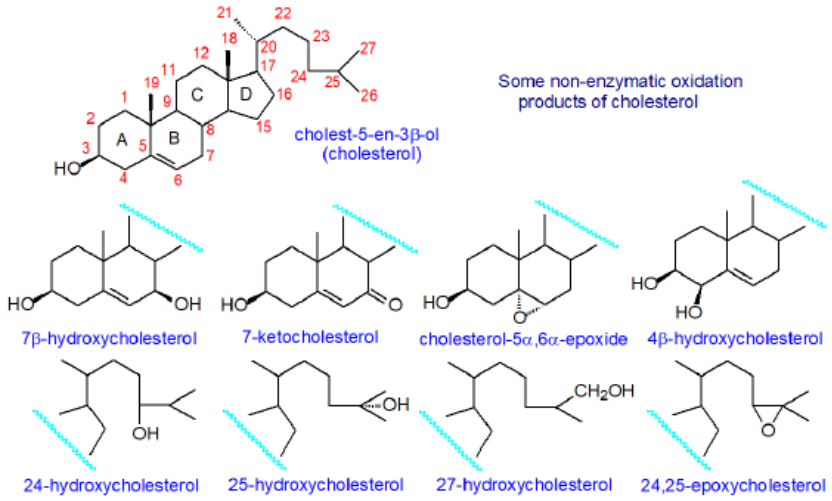


Figure 21.5.33: Some non-ezymatic oxidation products of cholesterol

Oxysterols produced by this means can vary in the type (hydroperoxy, hydroxy, keto, epoxy), number and position of the oxygenated functions introduced, and nature of their stereochemistry. Derivatives with the A and B rings and the *iso*-octyl side-chain oxidized are illustrated, but compounds with oxygen groups in position 15 (D ring) are also important biologically. Many are similar to those produced by enzymatic means, although the stereochemistry will usually differ. Like the enzymic products, they are named according to their relationship to cholesterol, rather than by the strict systematic terminology.

Mechanisms of autoxidation have been studied intensively in terms of unsaturated fatty acids, and it appears that similar mechanisms operate with sterols. The first event in lipid peroxidation by a radical mechanism is an initiation reaction in which a carbon with a labile hydrogen undergoes hydrogen abstraction by reaction with a free radical, which can be a non-lipid species such as a transition metal or hydroxyl or peroxynitrile radical, and this is followed by oxygen capture. The resulting reactive species recruits further non-oxidized lipids and starts a chain reaction termed the propagation phase. Finally, the reaction is terminated by the conversion of hydroperoxy intermediates to more stable hydroxy products by reaction with endogenous antioxidants such as tocopherols.

As an example, the reaction mechanism leading to the production of 7-oxygenated cholesterol derivatives is illustrated in Figure 21.5.34.

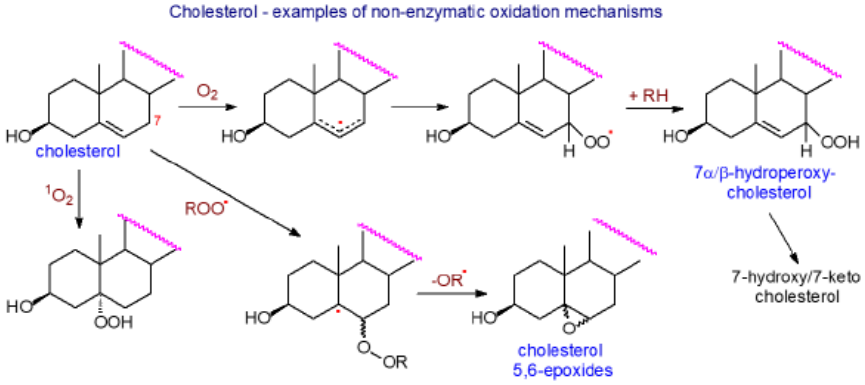


Figure 21.5.34: Cholesterol non-enzymatic oxidation mechanisms

In aqueous dispersions, oxidation is initiated by a radical attack from a reactive-oxygen species such as a hydroxyl radical with the abstraction of hydrogen from the C-7 position to form a delocalized three-carbon allylic radical, which reacts with oxygen to produce 7α-hydroperoxycholesterol, which gradually isomerizes to the more thermodynamically stable 7β-hydroperoxycholesterol. Subsequent enzymic and non-enzymic reactions lead to the 7-hydroxy and 7-keto analogs, which tend to be the most abundant non-enzymatically generated oxysterols in tissues, often accompanied by epoxy-ene and ketodienoic secondary products. Reaction with singlet oxygen ($^1O_2$) produces 5α-hydroperoxycholesterol mainly together with some 6α- or 6β-hydroperoxycholesterol. The reaction does not occur readily at the other allylic carbon 4, presumably because of steric hindrance. When cholesterol is in the solid state, the reaction occurs primarily at the tertiary carbon-25, though some products oxygenated at C-20 may also be produced.

Cholesterol hydroperoxides can be converted to stable diols only by the phospholipid hydroperoxide glutathione peroxidase - type 4 (GPx4) and then relatively slowly, but not by the type 1 glutathione peroxidase (GPx1) when in a membrane-bound state. However, in mammalian cells, monomeric GPx4 (~20 kDa), although present in several cellular compartments, including mitochondria, is much less abundant than tetrameric GPx1. Phospholipid-hydroperoxides are reduced most rapidly followed by cholesterol 6β-OOH > 7α/β-OOH >> 5α-OOH. The

result is that cholesterol hydroperoxides are expected to have a relatively long half-life and so can potentially be rather dangerous in biological systems. Of these, 5α-OOH with the lowest reduction rate is the most cytotoxic of the hydroperoxides, unfortunately.

Epimeric 5,6-epoxy-cholesterols may be formed by a non-radical reaction involving the non-enzymatic interaction of a hydroperoxide with the double bond, a process that is believed to occur in macrophages especially and in low-density lipoproteins (LDL). In this instance, the initial peroxidation product is a polyunsaturated fatty acid; the hydroperoxide transfers an oxygen atom to cholesterol to produce the epoxide, and in so doing is reduced to a hydroxyl. Other non-radical oxidation processes include reaction with singlet oxygen, which is especially important in the presence of light and photosensitizers and can generate 5-hydroxy- as well as 6- and 7-hydroxy products. In addition, the reaction with ozone in the lung can generate a family of distinctive oxygenated cholesterol metabolites.

Similarly, a diverse range of oxidation products are generated by peroxidation of the cholesterol and vitamin D precursor 7-dehydrocholesterol, which has the highest propagation rate constant known for any lipid toward free-radical chain oxidation, and these metabolites have important biological properties.

Oxysterols occur in tissues both in the free state and esterified with long-chain fatty acids. For example, in human atherosclerotic lesions, 80–95% of all oxysterols are esterified. Appreciable amounts of oxysterols can be present in foods, especially those rich in cholesterol such as meat, eggs, and dairy products, where they are most probably generated non-enzymically during cooking or processing when such factors as temperature, oxygen, light exposure, the associated lipid matrix, and the presence of antioxidants and water all play a part. Those present in foods can be absorbed from the intestines and transported into the circulation in chylomicrons, but the extent to which dietary sources contribute to tissue levels either of total oxysterols or of individual isomers is not known and is probably highly variable but relatively lower than of cholesterol *per se*.

21.5.2.3: OXYSTEROLS – BIOLOGICAL ACTIVITY

General Functions: In tissues *in vivo*, the very low oxysterol:cholesterol ratio means that oxysterols have little impact on the primary role of cholesterol in cell membrane structure and function, although it has been claimed that oxysterols could cause packing defects and thence atheroma formation in vascular endothelial cells. It is often argued that there are few reliable measurements of cellular or subcellular oxysterol concentrations, because of the technical difficulties in the analysis of the very low concentrations of oxysterols in the presence of a vast excess of native cholesterol; the average levels of 26-, 24- and 7α-hydroxy-cholesterol in human plasma that are often quoted are 0.36, 0.16 and 0.14 μM, respectively. Autoxidation products of cholesterol, especially 7-keto- and 7-hydroxy-cholesterol, are cytotoxic and may be useful markers of oxidative stress or for monitoring of the progression of various diseases. However, experts in the field caution that it can be difficult to extrapolate from experiments *in vitro* to the situation *in vivo*, because of the rapidity with which cholesterol can autoxidize in experimental systems and because of the difficulty of carrying out experiments with physiological levels of oxysterols.

Nonetheless, aside from their role as precursors of bile acids and some steroidal hormones, it is evident that oxysterols have a variety of roles in terms of maintaining cholesterol homeostasis and perhaps in signaling, where those formed enzymatically are most important. They can exert potent biological effects at physiologically relevant concentrations by binding to various receptors to elicit transcriptional programs, i.e., to regulate gene and hence protein expression. Among many cell membrane receptors for oxysterols to have been identified, nuclear receptors are especially important and include the liver X receptors (LXRs), retinoic acid receptor-related orphan receptors (RORs), estrogen receptors (ERs), and glucocorticoid receptors (GRs). In addition, *N*-methyl-D-aspartate receptors (NMDARs) are expressed in nerve cells and work over a short time scale to regulate excitatory synaptic function, while G protein-coupled receptors operate at cell membranes and are activated by molecules outside the cell to activate signaling pathways within the cell. As various isoforms of these receptors exist in different tissues, and these can interact with several oxysterols, only a brief summary of this complex topic is possible here.

A family of oxysterol-binding proteins (OSBP) transports and regulates the metabolism of sterols and targets oxysterols to specific membranes and especially to contact sites between organelles with the transport of **phosphatidylinositol 4-phosphate** in the reverse direction (see our web page on the latter). In this way, they can enable oxysterols to regulate membrane composition and function and mediate intracellular lipid transport. As with cholesterol, oxysterols can be eliminated from cells by transporters such as the ATP-binding cassette proteins ABCA1 and ABCG1, and they are transported in the blood-stream within lipoproteins, especially in association with HDL and LDL and mainly in the esterified form.

Cholesterol homeostasis: While cholesterol plays a key role in its own feedback regulation, there is some evidence that oxysterols are **regulators of cholesterol concentration** in cell membranes, and that **25-hydroxycholesterol** and **24(*S*),25-epoxycholesterol** may be especially effective, although the other side-chain oxysterols 22-, 24- and 27-hydroxycholesterol have been implicated. Several mechanisms appear to be involved, and it is suggested that 24(*S*),25-epoxycholesterol especially acts as a ligand for the liver X receptor, which forms a heterodimer with the retinoic X receptor, to inhibit the transcription of key genes in cholesterol biosynthesis, as well as directly inhibiting or accelerating the degradation of such important enzymes in the process as HMG-CoA reductase and squalene synthase. Similarly, both 26-hydroxylanosterol and 25-hydroxycholesterol inhibit HMG-CoA reductase. 25-Hydroxycholesterol inhibits the transfer of the 'sterol regulatory element binding protein' (SREBP-2) to the Golgi for processing to its active form as a transcription factor for the genes of the cholesterol biosynthesis pathway, and it stimulates the enzyme acyl-CoA:cholesterol acyl transferase in the endoplasmic reticulum to

esterify cholesterol. By such mechanisms, these oxysterols fine-tune cholesterol homeostasis and ensure smooth regulation rather than substantial fluctuations in tissue concentration.

Oxysterols and the immune system: Oxysterols and especially are known to have vital and diverse roles in immunity by regulating both the adaptive and innate immune responses to infection. For example, infection with viruses leads to the production of type I interferon, and in macrophages, this induces synthesis of **25-hydroxycholesterol**, which in general is regarded as anti-inflammatory and exerts broad antiviral activity by activating integrated stress response genes and reprogramming protein translation again via its interaction with LXR receptors. It is a potent inhibitor of SARS-CoV-2 replication, for example, possibly by a mechanism involving the blocking of cholesterol export from the late endosome/lysosome compartment and depletion of membrane cholesterol levels. However, the formation of 25-hydroxycholesterol may be harmful in the case of influenza infections, as it can lead to over production of inflammatory metabolites. Similarly, the biosynthesis of 25-hydroxycholesterol in macrophages is stimulated by the endotoxin **Kdo2-lipid A**, the active component of the lipopolysaccharide present on the outer membrane of Gram-negative bacteria, which acts as an agonist for Toll-like receptor 4 (TLR4). There is enhanced expression of the oxygenase CH25H in immune cells in response to bacterial and viral infection.

Many oxysterol species are active in a range of immune cells subsets, mediated through the control of LXR and SREBP signaling, but also by acting as ligands for RORs, and for the cell surface receptors G protein-coupled receptor 183 (GPR183) or CXCR2. Activation of LXR tends to dampen the immune response. In response to various stimuli, they can operate through ion channels to effect rapid changes in intracellular ion concentrations, especially of Ca^{2+}, to bring about changes in membrane potential, cell volume, cell death (apoptosis, autophagy, and necrosis), gene expression, secretion, endocytosis, or motility. For example, 27-hydroxycholesterol in human milk is reported to be active against the pathogenic human rotavirus and rhinovirus of importance in pediatrics, and 7-Dehydrocholesterol has anti-viral properties also. While they can exert their immune functions within the cell in which they are generated, oxysterols can also operate in a paracrine fashion towards other immune cells.

25-Hydroxycholesterol in particular can have either pro- or anti-inflammatory effects, depending upon the conditions, but the enzyme CH25H responsible for its biosynthesis is induced markedly in macrophages activated by inflammatory agents. It is reported to have a regulatory effect on the biosynthesis of sphingomyelin, which is required with cholesterol for the formation of **raft** sub-domains in membranes, where signaling molecules are concentrated, and together with other oxysterols, such as 24*S*,25-epoxycholesterol, to regulate the activities of the **hedgehog proteins** involved in embryonic development. Metabolites of 25-hydroxycholesterol, such as 7α,25-dihydroxycholesterol, and further oxidation products, are pro-inflammatory act as chemoattractants to lymphocytes; they have a role in the regulation of immunity in secondary lymphoid organs by interactions with the receptor GPR183.

Oxysterols in brain: Oxysterols are especially important for cholesterol homeostasis in the brain, which contains 25% of the total body cholesterol, virtually all of it in unesterified form, in only about 2% of the body volume. Cholesterol is a major component of the plasma membrane especially, where it serves to control fluidity and permeability. This membrane is produced in large amounts in the brain and is the basis of the compacted myelin, which is essential for the conductance of electrical stimuli and contains about 70% of brain cholesterol. While this pool is relatively stable, the remaining 30% is present in the membranes of neurons and glial cells of gray matter and is more active metabolically. Even in the fetus and the newborn infant, all the cholesterol required for growth is produced by synthesis *de novo* in the brain, not by transfer from the circulation across the blood-brain barrier, which consists of tightly opposed endothelial cells lining the extensive vasculature of the tissue. The fact that this pool of cholesterol in the brain is independent of circulating levels must reflect a requirement for constancy in the content of this lipid in membranes and myelin. In adults, although there is a continuous turnover of the membrane, the cholesterol is efficiently re-cycled and has a remarkably high half-life (up to 5 years). The rate of cholesterol synthesis is a little greater than the actual requirement so net movement of cholesterol out of the central nervous system must occur. An important component of this system is apolipoprotein E (Apo E), a 39-kDa protein, which is highly expressed in the brain and functions in the cellular transport of cholesterol and in cholesterol homeostasis. Apo E complexes with cholesterol are required for transport from the site of synthesis in astrocytes to neurons.

Hydroxylation by CYP46A1 of cholesterol to **24(*S*)-hydroxycholesterol** (cerebrosterol) is responsible for 50–60% of all cholesterol metabolism in the adult brain. If cholesterol itself cannot cross the blood-brain barrier, this metabolite is able to do so with relative ease. When the hydroxyl group is introduced into the side chain, this oxysterol causes a local re-ordering of membrane phospholipids such that it is more favorable energetically to expel it at a rate that is orders of magnitude greater than that of cholesterol *per se*, though still only 3-7 mg per day. There is a continuous flow of the metabolite from the brain into the circulation, much of it in the form of the inactive sterol ester, where it is transported by lipoprotein particles to the liver for further catabolism, i.e., it is hydroxylated in position 7 and then converted to **bile acids**. This is illustrated in Figure 21.5.35.

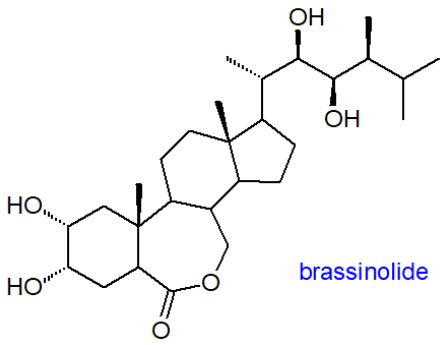


Figure 21.5.35: Brain/Liver 27-carboxy and 27-OH steroid metabolism.

Both 24(*S*)-hydroxycholesterol and 24(*S*),25-epoxycholesterol are believed to be important in regulating cholesterol homeostasis in the brain. They interact with the specific nuclear receptors involved in the expression and synthesis of proteins involved in sterol transport, and for example, 24-hydroxy-cholesterol regulates the transcription of Apo E. In particular, it is an agonist of the nuclear liver X receptors (LXRs), influencing the expression of those LXR target genes involved in cholesterol homeostasis and inflammatory responses. It is also a high-affinity ligand for the retinoic acid receptor-related orphan receptors α and γ (RORα and RORγ). In this way, it can act locally to affect the functioning of neurons, astrocytes, oligodendrocytes, and vascular cells.

24(*S*)-Hydroxycholesterol down-regulates the trafficking of the amyloid precursor protein and may be a factor in preventing neurodegenerative diseases. Especially high levels of 24(*S*)-hydroxycholesterol are observed in the plasma of human infants and in patients with brain trauma, while reduced levels are found in the plasma of patients with neurodegenerative diseases, including Parkinson's disease, multiple sclerosis, and Alzheimer's disease. In contrast, there are elevated levels in the brain and especially cerebrospinal fluid in patients with these conditions, where it may be a marker of neurodegeneration. Increased expression of cholesterol 24-hydroxylase (CYP46A1) is believed to improve cognition, while a reduction leads to poor cognitive performance, as occurs at advanced stages of the disease, probably reflecting a selective loss of neuronal cells, and it may be a factor in age-related macular degeneration. An excess of 24(*S*)-hydroxycholesterol and especially of its ester form can lead to neuronal cell death, and elevated levels in plasma are reported to be a potential marker for Autism Spectrum Disorders in children. On the other hand, it may be protective against glioblastoma, the most common primary malignant brain tumour in adults via activation of LXRs.

27-Hydroxy-cholesterol diffuses across the blood-brain barrier from the bloodstream into the brain (in the reverse direction to 24-hydroxycholesterol), where it does not accumulate but is further oxidized and then exported as steroidal acids. This flux may regulate certain key enzymes within the brain, and there are suggestions that the balance between the levels of 24- and 27-hydroxy-cholesterol, especially excess of the latter, may be relevant to the generation of β-amyloid peptides in Alzheimer's disease by reducing insulin-mediated glucose uptake by neurons. While 7β-hydroxycholesterol is pro-apoptotic, any links with Alzheimer's disease are unproven although there is a school of thought that other oxidized cholesterol metabolites may be major factors behind the development of this disease. For example, seco-sterols such as 3β-hydroxy-5-oxo-5,6-secocholestan-6-al and its stable aldolization product, the main ozonolysis metabolites derived from cholesterol, have been detected in brain samples of patients who have died from Alzheimer's disease and Lewy body dementia; they are also found in atherosclerotic lesions. Oxidation products of the cholesterol precursor 7-dehydrocholesterol and especially 3β,5-dihydroxycholest-7-ene-6-one are involved in the pathophysiology of the human disease Smith-Lemli-Opitz syndrome.

Cell differentiation: Oxysterols can influence the differentiation of many cell types and this was first studied in the skin, where 22(*R*)- and 25(*R*)-hydroxycholesterol were shown to induce human keratinocyte differentiation. Subsequently, by stimulating nuclear binding receptors, oxysterols were found to have similar effects on mesenchymal stem cells. There have been many reports of the involvement of oxysterols in disease processes, especially atherosclerosis and the formation of human atherosclerotic plaques, but also cytotoxicity, necrosis, inflammation, immuno-suppression, phospholipidosis and gallstone formation. They have been implicated in the development of cancers, especially those of the breast, prostate, colon, and bile duct. For example, **27-hydroxycholesterol** is an element in cholesterol elimination from macrophages and arterial endothelial cells, but it is also an endogenous ligand for the human nuclear estrogen receptor (ERα) and the liver X receptor, and it modulates their activities with effects upon various human disease states, including cardiovascular dysfunction and progression of cancer of the breast and prostate, as well as having an involvement in the regulation of bone mineralization (osteoporosis). It has been linked to cancer metastasis through effects on immune cells, and there is hope that pharmacological inhibition of CYP27A1 and thence the formation of 27-hydroxycholesterol may be a useful strategy in the treatment of breast cancer; *CYP7A1* gene polymorphism has been associated with colorectal cancer. In contrast, oxysterols can interfere in the proliferation of several types of cancer cell (glioblastoma, leukemia, colon, breast, and prostate cancer).

Cholesterol 5,6-epoxide (with either 5α or 5β stereochemistry) is formed non-enzymatically in tissues, but it is also believed to be produced by an unidentified cytochrome P450 enzyme in the adrenal glands. While it was for some time believed to be a causative agent in cancer, it is now recognized that downstream metabolites are responsible. Thus, cholesterol epoxide hydrolase converts cholesterol 5,6-epoxide into cholestane-3β,5α,6β-triol, which is transformed by 11β-hydroxysteroid-dehydrogenase-type-2 into the oncometabolite 3β,5α-

dihydroxycholestan-6-one (oncosterone). By binding to the glucocorticoid receptor, this oncosterone stimulates the growth of breast cancer cells, and it also acts as a ligand to the LXR receptors, which may mediate its pro-invasive effects. In contrast, in normal breast tissue, cholesterol 5,6-epoxide is metabolized to the tumor suppressor metabolite, a steroidal alkaloid designated dendrogenin A that is a conjugation product with histamine and controls a nuclear receptor to trigger lethal autophagy in cancers; its synthesis is greatly reduced in cancer cells. Tamoxifen, a drug that is widely used against breast cancer, binds to the cholesterol 5,6-epoxide hydrolase, which is also a microsomal anti-estrogen binding site (AEBS), to inhibit its activity.

7-Ketocholesterol is a major oxysterol produced during the oxidation of low-density lipoproteins, and is one of the most abundant in plasma and atherosclerotic lesions; it accumulates in erythrocytes of heart failure patients. It has a high pro-apoptotic potential and associates preferentially with membrane lipid **raft** domains. As it is not readily exported from macrophages, it impairs cholesterol efflux and promotes the foam cell phenotype. In cardiomyocytes, this accumulation can lead to cell hypertrophy and death, and it has been suggested that oxysterols are a major factor precipitating morbidity in atherosclerosis-induced cardiac diseases and inflammation-induced heart complications. Photoxidation in the retina via the action of free radicals or singlet oxygen generates unstable cholesterol hydroperoxides, which may be involved in age-related macular degeneration. For example, these compounds can quickly be converted to highly toxic 7α- and 7β-hydroxycholesterols and 7-ketocholesterol, depending on the status of tissue oxidases and reductases. Three separate enzymatic pathways have developed in the eye to neutralize their activities. These sterols are metabolized by novel branches of the acidic pathway of bile acid biosynthesis.

Those oxysterols formed non-enzymatically can be most troublesome in relation to disease in general. For example, they are enriched in pathologic cells and tissues, such as macrophage foam cells, atherosclerotic lesions, and cataracts. They may regulate some of the metabolic effects of cholesterol, but as cautioned above, effects observed *in vitro* may not necessarily be of physiological importance *in vivo*. Various oxysterols have been implicated in the differentiation of mesenchymal stem cells and the signaling pathways involved in this process. High levels of 7-hydroxycholesterol and cholestane-3β,5α,6β-triol are characteristic of the lysosomal storage diseases Niemann-Pick types B and C and of lysosomal acid lipase deficiency.

Cholesterol hydroperoxides: With the aid of START domain proteins, cholesterol hydroperoxides can translocate from a membrane of origin to another membrane such as mitochondria. Such transfer of free radical-generated 7-hydroperoxycholesterol, for example, has adverse consequences in that there is impairment of cholesterol utilization in steroidogenic cells, and of anti-atherogenic reverse-cholesterol transport in vascular macrophages. The antioxidant activity of GPx4 may be crucial for the maintenance of mitochondrial integrity and functionality in these cells.

21.5.2.4: VITAMIN D

Vitamin D encompasses two main sterol metabolites that are essential for the regulation of calcium and phosphorus levels and thence for bone formation in animals. However, these have many other functions, including induction of cell differentiation, inhibition of cell growth, immunomodulation, and control of other hormonal systems. Vitamin D (with calcium) deficiency is responsible for the disease rickets in children in which bones are weak and deformed, and it is associated with various cancers and autoimmune diseases. Ultraviolet light mediates cleavage of 7-dehydrocholesterol, an important intermediate in the biosynthesis of cholesterol, with the opening of the second (B) ring in the skin to produce pre-vitamin D, which rearranges spontaneously to form the secosteroid vitamin D_3 or cholecalciferol. Its structure is shown in Figure 21.5.36.

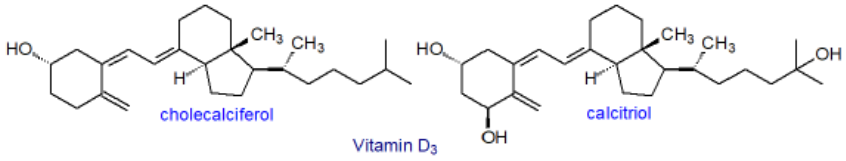


Figure 21.5.36: Structure of Vitamin D_3

The newly generated vitamin D_3 is transported to the liver where it is subject to 25-hydroxylation and thence to the kidney for 1α-hydroxylation to produce the active form 1α,25-dihydroxyvitamin D_3 (calcitriol); this is a true hormone and serves as a high-affinity ligand for the vitamin D receptor in distant tissues. For transportation in plasma, it is bound to a specific glycoprotein termed unsurprisingly, the 'vitamin D binding protein (BDP)'. Vitamin D_2 or ergocalciferol is derived from ergosterol, which is obtained from plant and fungal sources in the diet.

Vitamin D_3 functions by activating a cellular receptor (vitamin D receptor or VDR), a transcription factor binding to sites in the DNA called vitamin D response elements. There are thousands of such binding sites, which together with co-modulators regulate innumerable genes in a cell-specific fashion. In this way, it enhances bone mineralization by promoting dietary calcium and phosphate absorption, as well as having direct effects on bone cells. In addition, it functions as a general development hormone in many different tissues, while together with Vitamin D_2 it has profound effects on immune responses in the defense against microbes.

21.5.2.5: STEROIDAL HORMONES AND THEIR ESTERS

Steroidal hormones cannot be discussed in depth here as their structures, biosynthesis, and functions comprise a rather substantial and specialized topic. In brief, animal tissues produce small amounts of vital steroidal hormones from cholesterol as the primary precursor with 22*R*-hydroxycholesterol, produced by hydroxylation by the cholesterol side-chain cleavage enzyme (P450scc), as the first of its metabolites in the pathway. This step involves the 'STAR' protein which enables the transport of cholesterol into mitochondria where conversion to pregnenolone is rate-limiting and involves first hydroxylation and then cleavage of the side-chain. After export from the mitochondria, this can be converted directly to progesterone or in several steps to testosterone. 17β-Estradiol, for example, is the most potent and important of the endogenous estrogens; it is made mainly in the follicles of the ovaries and regulates menstrual cycles and reproduction, but it is also present in testicles, adrenal glands, fat, liver, breasts, and brain. Testosterone is the primary male sex hormone and an anabolic steroid, and it is produced mainly in the testes; it has a key function in the development of male reproductive tissues such as testes and prostate, in addition to promoting secondary sexual characteristics. Pregnane neurosteroids are synthesized in the central nervous system. Cholesterol homeostasis is therefore vital to fertility and a host of bodily functions. The structures of key steroidal hormones are shown in Figure 21.5.37.

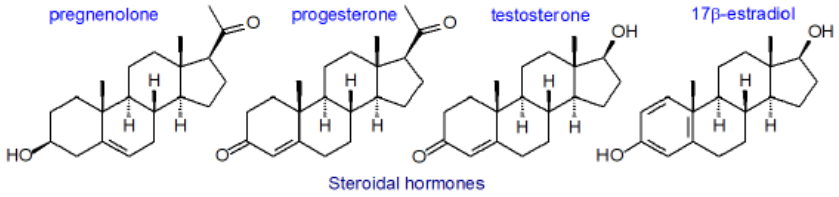


Figure 21.5.37: Structures of key steroidal hormones

Steroidal esters accumulate in tissues such as the adrenal glands, which synthesize corticosteroids such as cortisol and aldosterone and are responsible for releasing hormones in response to stress and other factors. It is also apparent that fatty acyl esters of estradiols, such as dehydroepiandrosterone, accumulate in adipose tissue in post-menopausal women. Small amounts of estrogens acylated with fatty acids at the C-17 hydroxyl are present in the plasma lipoproteins. In each instance, they appear to be biologically inert storage or transport forms of the steroid. Eventually, esterified steroids in low-density lipoproteins (LDL) particles are taken up by cells via lipoprotein receptors, and then are hydrolyzed to release the active steroid. Pharmaceutical interest in oleoyl-estrone, a naturally occurring hormone in humans, which was found to induce a marked loss of body fat while preserving protein stores in laboratory animals, has declined as clinical trials with humans were not successful.

21.5.3: STEROLS 3. STEROLS AND THEIR CONJUGATES FROM PLANTS AND LOWER ORGANISMS

21.5.3.1: PLANT STEROLS - STRUCTURES AND OCCURRENCE

Plants, algae, and fungi contain a rather different range of sterols from those in animals. Like **cholesterol**, to which they are related structurally and biosynthetically, plant sterols form a group of triterpenes with a tetracyclic cyclopenta[*a*]phenanthrene structure and a side chain at carbon 17, sometimes termed the $C_{30}H_{50}O$ triterpenome. The four rings (A, B, C, D) have *trans* ring junctions, and the side chain and two methyl groups (C-18 and C-19) are at an angle to the rings above the plane with β stereochemistry (as for a hydroxyl group commonly located on C-3 also). The basic sterol from which other sterol structures are defined is 5α-cholestan-3β-ol with the numbering scheme recommended by IUPAC as shown in Figure 21.5.38.

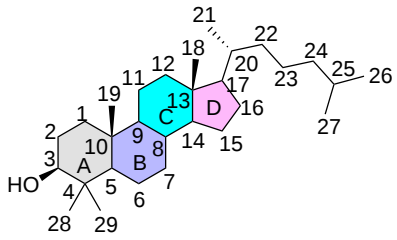


Figure 21.5.38: IUPAC sterol numbing system

The phytosterols (as opposed to zoosterols) include campesterol, β-sitosterol, stigmasterol, and Δ5-avenasterol, some of which are illustrated in Figure 21.5.39.

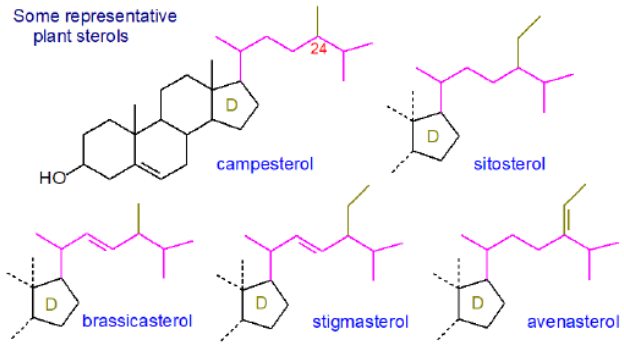


Figure 21.5.39: Some common plant sterols

These more common plant sterols have a double bond in position 5, and a definitive feature – a one- or two-carbon substituent with variable stereochemistry in the side chain at C-24, which is preserved during subsequent metabolism. For example, campesterol is a 24-methylsterol, while β-sitosterol and stigmasterol are 24-ethylsterols. Occasionally, there is a double bond in this chain that can be of the *cis* or *trans* configuration as in stigmasterol (at C22) or fucosterol (C24), the main sterol in green algae.

Phytosterols can be further classified on a structural or biosynthetic basis as 4-desmethyl sterols (i.e. with no substituent on carbon-4), 4α-monomethyl sterols and 4,4-dimethyl sterols. The most abundant group is the 4-desmethyl sterols, which may be subdivided into Δ5-sterols (illustrated above), Δ7-sterols (e.g. α-spinasterol) and Δ5,7-sterols depending on the position of the double bonds in the B ring. As the name suggests, brassicasterols (24-methyl-cholesta-5,22-dien-3β-ol and related sterols) are best known from the brassica family of plants, but they are also common constituents of marine algae (phytoplankton). Phytostanols (fully saturated) are normally present at trace levels only in plants, but they are relatively abundant in cereal grains.

Many different sterols may be present in photosynthetic organisms, and the amounts and relative proportions are dependent on the species. Over 250 different phytosterols have been recorded with 60 in corn (maize) alone, for example. As a rough generality, a typical plant sterol mixture would be 70% sitosterol, 20% stigmasterol, and 5% campesterol (or >70% 24-ethyl-sterols and <30% 24-methyl-sterols), although this will vary with the stage of development and in response to stress. **Table 1** contains data on the main components from some representative commercial seed oils.

Table 21.5.2. **Sterol composition in some seed oils of commercial importance (mg/Kg).**

	corn	cottonseed	olive	palm	rapeseed	safflower	soybean	sunflower
cholesterol	-	-	-	26	-	-	-	-
campesterol	2691	170	28	358	1530	452	720	313
stigmasterol	702	42	14	204	-	313	720	313
β-sitosterol	7722	3961	1310	1894	3549	1809	1908	2352
Δ5-avenasterol	468	85	29	51	122	35	108	156
Δ7-stigmasterol	117	-	58	25	306	696	108	588
Δ7-avenasterol	-	-	-	-	-	104	36	156
brassicasterol	-	-	-	-	612	-	-	-
other	-	-	-	-	-	69	-	39

Data adapted from Gunstone, F.D. *et al. The Lipid Handbook (Second Edition)* (Chapman & Hall, London) (1994).

Cholesterol is usually a minor component only of plant sterols (<1%), but it is unwise to generalize too much as it can be the main sterol component of red algae and of some families of higher plants such as in the Solanaceae, Liliaceae and Scrophylariaceae. It can also be a significant constituent sterol of chloroplasts, shoots, pollen and leaf surface lipids in other plant families; wheat roots contain 10% and Arabidopsis cells 19% of the sterols as cholesterol. Yeasts and fungi tend to contain ergosterol as their main sterol (see below). **Ecdysteroids** (phytoecdysteroids) are polyhydroxylated plant sterols that can occur in appreciable amounts in some species. Sterols are also found in some bacterial groups but not in archaea, and hopanoids in bacteria are considered to be functional triterpenic counterparts.

Sterols can occur in plants in the 'free' state, i.e. in which the sterol hydroxyl group is not linked to any other moiety, but they are usually present also as conjugates with the hydroxyl group covalently bound via an ester bond to a fatty acid, for example, i.e. as sterol esters, or via a glycosidic linkage to glucose (and occasionally other sugars), i.e. as steryl glycosides.

21.5.3.2: PLANT STEROLS - BIOSYNTHESIS

The biosynthetic route to plant sterols resembles that to cholesterol in many aspects in that it follows an isoprenoid biosynthetic pathway with isopentenyl pyrophosphate, derived primarily from mevalonate, as the key building block in the cytoplasm (but not plastids) at least. The main pathway for the biosynthesis of isopentenyl pyrophosphate and dimethylallyl pyrophosphate, the isoprene units, is described previously and so need not be repeated here. It is known as the 'mevalonic acid (MVA) pathway' and functions in the cytosol, endoplasmic reticulum and mitochondria.

However, an alternative pathway that does not use mevalonic acid as a precursor was established first for bacterial **hopanoids**, but has since been found in plant chloroplasts, algae, cyanobacteria, eubacteria, and some parasites (but not in animals). This route is variously termed the 'non-mevalonate', '1-deoxy-D-xylulose-5-phosphate' (DOXP) or better the **2*C*-methyl-D-erythritol 4-phosphate (MEP)** pathway as the last compound is presumed to be the first committed intermediate in sterol biosynthesis by this route. In the first step, pyruvate and glyceraldehyde phosphate are combined to form deoxyxylose phosphate, which is in turn converted to 2C-methyl-D-erythritol 4-phosphate. The pathway then proceeds via various erythritol intermediates until isopentenyl pyrophosphate and dimethylallyl pyrophosphate are formed. This pathway is illustrated in Figure 21.5.40.

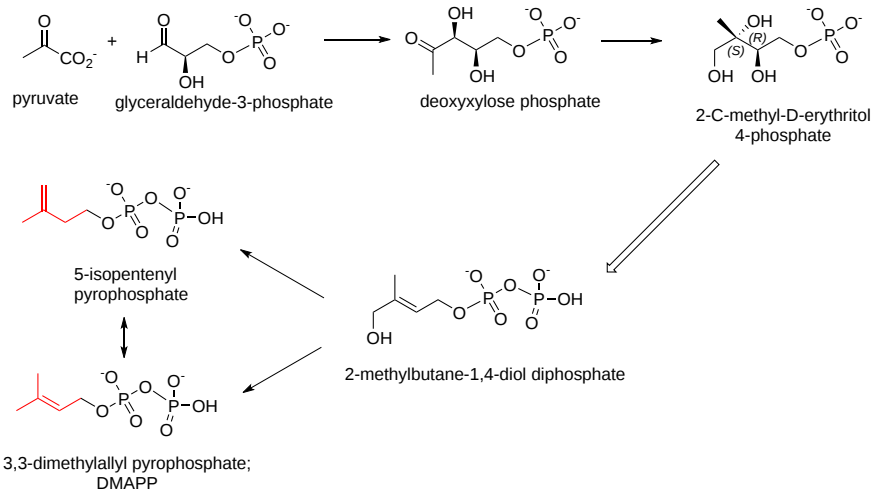


Figure 21.5.40: Alternative 2*C*-methyl-D-erythritol 4-phosphate (MEP) pathway for synthesis of sterol and isoprenoid precursors in chloroplasts.

There is evidence that some of the isoprene units are exchanged between the cytoplasm and plastids. In much of the plant kingdom, both the MVA and MEP pathways operate in parallel, but green algae use the MEP pathway only. Thereafter, sterol biosynthesis continues via squalene and (3*S*)-2,3-oxidosqualene.

In photosynthetic organisms (as opposed to yeast and fungi), the subsequent steps in the biosynthesis of plant sterols differ from that for cholesterol in that the important intermediate in the route from squalene via 2,3-oxidosqualene is **cycloartenol**, rather than lanosterol, and this is produced by the action of a 2,3(*S*)-oxidosqualene-cycloartenol cyclase (cycloartenol synthase). Then, the enzyme sterol methyltransferase 1 is of special importance in that it converts cycloartenol to 24-methylene cycloartenol, as the first step in introducing the methyl group onto C-24, while the enzyme cyclopropyl sterol isomerase is required to open the cyclopropane ring. Animals lack the sterol C24-methyltransferase gene. While this pathway is in essence linear up to the synthesis of 24-methylene lophenol, a bifurcation then occurs that results in two alternative pathways, one of which leads to the synthesis of sitosterol and stigmasterol and the other to that of campesterol. This pathway is shown in Figure 21.5.41.

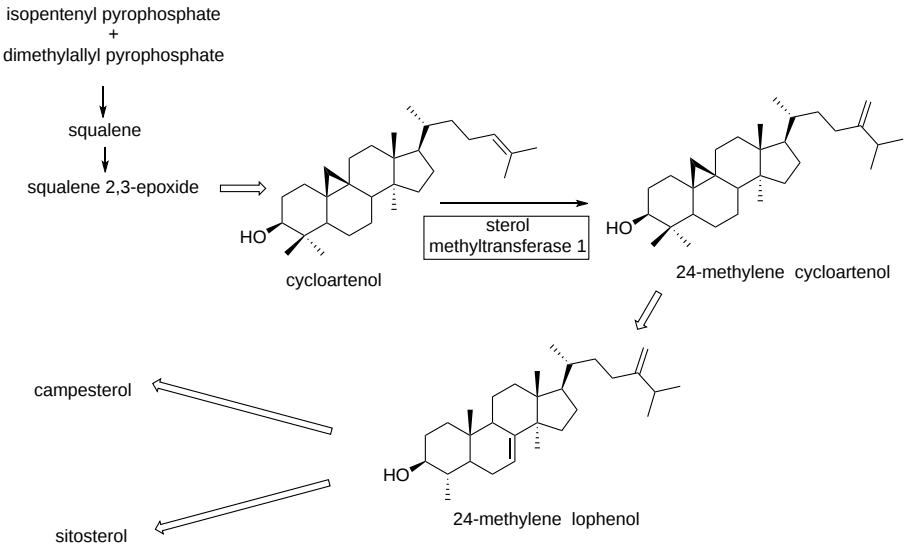


Figure 21.5.41: Synthesis of plant sterols via the cycloartenol intermediate

In fact, there are more than thirty enzyme-catalyzed steps in the overall process of plant sterol biosynthesis, each associated with membranes, and detailed descriptions are available from the reading list below. The 4,4-dimethyl- and 4α-methylsterols are part of the biosynthetic pathway, but are only minor if ubiquitous sterol components of plants. New biosynthetic pathways are now being discovered by genome analysis that reveal the complexity of sterol biosynthesis in different plant species.

Dinoflagellates produce a characteristic 4-methylsterol termed dinosterol and others like gorgosterol via lanosterol as precursor. Protozoans synthesize many different sterols related to those in plants. For example, some species of *Acanthamoeba* and *Naegleria* produce both lanosterol and cycloartenol, but only the latter is used for the synthesis of other sterols, especially ergosterol, but in other protozoan species, sterol biosynthesis occurs via lanosterol. The best studied bacterial pathway is that of the methylotroph *Methylococcus capsulatus*, which produces a number of unique Δ18(14)-sterols and is known to possess a squalene epoxidase and a lanosterol-14-demethylase.

Cholesterol in plants is produced from cycloartenol as the key intermediate with the Sterol Side Chain Reductase 2 (SSR2) as the key enzyme. It is now established that the cholesterol biosynthetic pathway in tomato plants comprises 12 enzymes acting in 10 steps. Of these, half evolved through gene duplication and divergence from phytosterol biosynthetic enzymes, whereas others act reciprocally in both cholesterol and phytosterol metabolism. Algae can also produce cholesterol in a multi-step process from cycloartenol, and many more sterols via 24-methylene lophenol as the key intermediate. It is hoped that genetic manipulation of these enzymes will lead to plants that synthesize high-value steroidal products.

Oxidation: Phytosterols can be subjected to non-enzymatic oxidation with the formation of **oxysterols** in a similar manner to that of cholesterol in animals, resulting in ring products such as hydroxy-, keto-, epoxy- and triol-derivatives, and further enzymic reactions can oxidize the side chain. However, photosensitized oxidation is more common in plants and is much faster (>1500 times); it starts with the *ene*-addition of singlet oxygen ($^1O_2$) on either side of the double bond in the B ring to generate 5α-/6α-/6β-hydroperoxysterols, of which 5α-OOH is the most abundant and rearranges to form the more stable 7α-OOH isomer. This is the main reaction in foods stored under LED lighting in food retailers.

21.5.3.3: PLANT STEROLS - FUNCTION

Like cholesterol, plant sterols are amphiphilic and are vital constituents of all membranes, and especially of the plasma membrane, the mitochondrial outer membrane and the endoplasmic reticulum. The three-dimensional structure of the plant sterols is such that there are planar surfaces at both the top and the bottom of the molecules, which permit multiple hydrophobic interactions between the rigid sterol and the other components of membranes. Indeed, they must determine the physical properties of membranes to an appreciable extent. It is believed that campesterol, β-sitosterol, and 24-methylcholesterol (in this order) are able to regulate membrane fluidity and permeability in plant membranes by restricting the mobility of fatty acyl chains in a similar manner to cholesterol in mammalian cells, but stigmasterol has much less effect on lipid ordering and no effect on the permeability of membranes. In the plasma membrane, plant sterols associate with the glycosphingolipids such as **glucosylceramide,** and **glycosylinositolphosphoceramides** in **raft**-like sub-domains, analogous to those in animal cells, and these support the membrane location and activities of many proteins with important functions in plant cells. The sterol glycosides are especially important in this context (see below).

Sterols (and their conjugates) are involved in plant membrane adaptations to changes in temperature and other biotic and abiotic stresses. For example, β-sitosterol is a precursor of stigmasterol via the action of a C22-sterol desaturase, and the ratio of these two sterols is important to the resistance of *A. thaliana* plants to low and high temperatures. In addition, plant sterols can modulate the activity of membrane-bound enzymes. Thus, stigmasterol and cholesterol regulate the activity of the Na^+/K^+-ATPase in plant cells, probably in a manner analogous to that of cholesterol in animal cells. Stigmasterol may be required specifically for cell differentiation and proliferation. As well as being the precursor of plant steroidal hormones, campesterol, is a signaling molecule that regulates growth, development, and stress adaptation.

Perhaps surprisingly, cholesterol is a precursor for the biosynthesis of some steroidal saponins and alkaloids in plants, for example, the well-known steroidal glycoalkaloid in potato (α-solanine), as well as of other steroids including the phytoecdysteroids (in some species they are derived from lathosterol). While the physiological roles of ecdysteroids in plants yet to be been confirmed, they are believed to enhance stress resistance by promoting health and vitality. Withanolides are complex oxysterols, which are believed to be defense compounds against insect herbivores.

21.5.3.4: STEROIDAL PLANT HORMONES

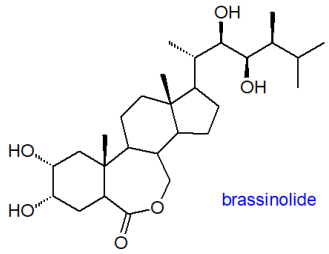


Figure 21.5.42 Structure of brassinolide

They have crucial importance for plant growth processes, including cell elongation, division, differentiation, immunity, and development of reproductive organs, and they are involved in the regulation of innumerable aspects of metabolism. Via signal transduction pathways, they interact with transcription factors through phosphorylation cascades to regulate the expression of target genes. Brassinosteroids are also signaling molecules in abiotic stress responses such as drought, salinity, high temperature, low temperature, and heavy metal stresses. Outwith plants, they may have biomedical applications as anticancer drugs for endocrine-responsive cancers to induce apoptosis and inhibit growth. Some plant species produce small amounts of steroid hormones that are often considered to be of animal origin only, including progesterone and testosterone, and these may have physiological roles in plants.

21.5.3.5: STEROL ESTERS IN HIGHER PLANTS

Sterol esters are present in all plant tissues, but they are most abundant in tapetal cells of anthers, pollen grains, seeds, and senescent leaves. In general, they are minor components relative to the free sterols other than in waxes. Usually, the sterol components of sterol esters are similar to the free sterols, although there may be relatively less of stigmasterol. The fatty acid components tend to resemble those of the other plant tissue lipids, but there can be significant differences on occasion. Sterol esters are presumed to serve as inert storage forms of sterols, as they are often enriched in the intermediates of sterol biosynthesis and can accumulate in **lipid droplets** within the cells. However, they have been found in some membranes, especially in microsomes and mitochondrial preparations, although their function there is uncertain. They may also have a role in transport within cells and between tissues, as they can be present in the form of soluble lipoprotein complexes.

Biosynthesis of sterol esters in *A. thaliana* is known to occur in the endoplasmic reticulum by the action of a phospholipid:sterol acyltransferase, which catalyzes the transfer of a fatty acyl group to the sterol from position *sn*-2 of phospholipids - mainly phosphatidylethanolamine; the enzyme is very different from those in animals and yeasts. However, an acyl CoA:sterol acyltransferase closer in structure to the animal enzyme has been characterized also; it prefers saturated fatty acyl-CoAs as acyl donors and cycloartenol as the acceptor molecule. The enzymes responsible for the hydrolysis of sterol esters in *A. thaliana* are not yet known. Certain distinctive phytosterol esters occur in the aleurone cells of cereal grains, including *trans*-hydroxycinnamate, ferulate (4-hydroxy-3-methoxycinnamate), and *p*-coumarate esters. Similarly, rice bran oil is a rich source of esters of ferulic acid and a mixture of sterols and triterpenols, termed 'γ-orizanol'', and an example of one of these compounds is illustrated in Figure 21.5.43.

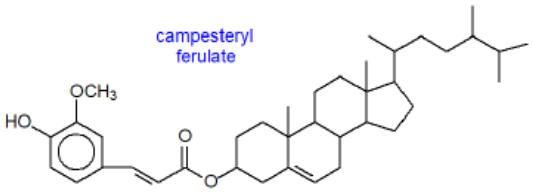


Figure 21.5.43 Structure of campesteryl ferulate

This is sold as a health food supplement, because of the claimed beneficial effects, including cholesterol-lowering and antioxidant activities, while enhancing muscle growth and sports performance. However, none of these effects have been confirmed by rigorous clinical testing.

21.5.3.6: STEROL GLYCOSIDES

Leaf and other tissue in plants contain a range of **sterol glycosides** and **sterol acyl-glycosides** in which the hydroxyl group at C3 on the sterol is linked to the sugar by a glycosidic bond. Other than in the genus *Solanum*, where they can represent up to 85% of sterol fraction in tomato fruit as an example, they tend to be minor components relative to other lipids. Typical examples (glucosides of β-sitosterol) are illustrated in Figure 21.5.44.

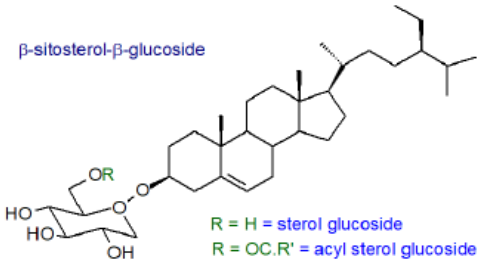


Figure 21.5.44 Structure of β-sitosterol-β-glucoside

Most of the common plant sterols occur in this form, but Δ5 sterols are preferred (Δ7 in some genera). Glucose is the most common carbohydrate moiety but galactose, mannose, xylose, arabinose can also be present depending on plant species; occasionally, complex carbohydrates with up to five hexose units linked in a linear fashion are present. Algae also contains sterol glycosides with a wide range of sterol and carbohydrate components. Plant, animal, fungal, and most bacterial steryl glycosides have a β-glycosidic linkage, but in a few bacterial species there is an α-linkage.

Similarly, the nature of the fatty acid component in the acyl sterol glycosides can vary as well as the hydroxyl group to which they are linked, although it is usually position 6 of the glucose moiety. In potato tubers, for example, the 6'-palmitoyl-β-D-glucoside of β-sitosterol is the major species, while the corresponding linoleate derivative predominates in soybeans. Usually, the sterol acyl-glycosides are present at concentrations that are two- to tenfold greater than those of the non-acylated forms. They are known to be located in the plasma membrane, tonoplasts and endoplasmic reticulum.

Biosynthesis involves the reaction of free sterols with a glucose unit catalyzed by a sterol glycosyltransferase, or by the reaction of the sterol with uridine diphosphoglucose (UDP-glucose) and UDP-glucose:sterol glucosyltransferase on the cytosolic side of the plasma membrane. The acyl donor for acyl sterol glycoside synthesis is not acyl-coenzyme A but is believed to be a glycerolipid. Steryl β-D-glycoside hydrolases have been characterized from plants that reverse this reaction, but no fatty acyl hydrolase activity for sterol acyl-glycosides is yet known. One route to the biosynthesis of **glucosylceramides** in plants involves the transfer of the glucose moiety of sterol glycosides to ceramide.

The functions of sterol glycosides and sterol acyl-glycosides are slowly being revealed, and they are believed to be significant components of the plasma membrane that associate with sphingolipids in **raft-like domains**; the esterified form especially may be involved in the adaptation of plant membranes to low temperatures and other stresses. It is possible that they have a role in signal transmission through membranes, and they are reported to be beneficial in the response to pathogens. It seems probable that sterol glycosides are oriented with the sterol moiety buried in the hydrophobic core of the lipid bilayer with the sugar located in the plane of the polar head groups of the membrane, while with sterol acyl-glycosides both the sterol moiety and the fatty acid chain are embedded in the hydrophobic core of the membrane. Sitosterol-β-D-glucoside in the plasma membrane is believed to be the primer molecule for cellulose synthesis in plants, as in cotton (*Gossypium arboreum*) fiber, where it may be required for the initiation of glucan polymerization. The sterol is eventually removed from the polymer by a specific cellulase enzyme (the multimeric cellulase synthase is believed to be stabilized by sterols in the plasma membrane).

Sterol glycosides appear to be essential for the pathogenicity of certain fungi and for some bacteria, and ergosterol glycosides especially are especially troublesome components of plant fungal pathogens. Sterol glycosides have only rarely been reported from organisms other than plants and fungi, although some bacteria, such as the gram-negative bacterium *Helicobacter pylori* and *Borrelia burgdorferi*, the causative agent of Lyme disease produce **cholesterol glucoside** from host cholesterol. On the other hand, cholesteryl glucoside has been found as a natural component of a few animal tissues, and through acting as immunoadjuvants, sterol glycosides are reported to be efficacious in protecting animal hosts against lethal Cryptococcal infections. In the human diet, sterol glycosides have potential benefits in that like free sterols they inhibit the absorption of cholesterol from the gut and reduce the plasma cholesterol levels. The fatty acids are removed from sterol acyl-glycosides by enzymes in the intestine.

A number of species of monocotyledons contain complex steroidal **saponins**, which consist of an aglycone based on a triterpenoid furostanol or spirostanol skeleton (derived from cholesterol) and an oligosaccharide chain of two to five hexose or pentose moieties attached to the 3β-hydroxyl group of the sterol. These can interact with cholesterol in plant membranes to form insoluble complexes, which increase membrane permeability.

21.5.3.7: ERGOSTEROL AND OTHER STEROLS IN YEASTS AND FUNGI

Yeasts and fungi, together with microalgae and protozoa, can contain a wide range of different sterols. However, **ergosterol** ((22*E*)-ergosta-5,7,22-trien-3β-ol) is the most common sterol in fungi and yeast, and is accompanied by other sterols not normally abundant in higher plants including cholesterol, 24-methyl cholesterol, 24-ethyl cholesterol, and brassicasterol, depending upon species. In *Saccharomyces cerevisiae*, which is widely studied as a model species of yeasts, ergosterol is the most abundant sterol (ca. 12% of all lipids), with the highest levels in the plasma membrane (up to 40% of the lipids or 90% of the total cell sterols). Its structure is shown in Figure 21.5.45.

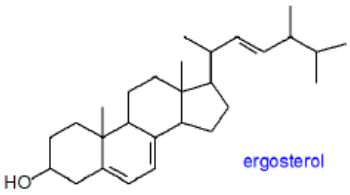


Figure 21.5.45: Strurture of ergosterol

Like cholesterol and in contrast to the plant sterols, it is synthesized in the endoplasmic reticulum via lanosterol as the key intermediate and then zymosterol, but the pathway diverges at this stage to produce fecosterol on the way to ergosterol (see the reading list below for further details). Ergosterol is transported to other organelles within the cell in a non-vesicular manner by two families of evolutionarily conserved sterol-binding proteins - 'Osh' and 'Lam', which are able to optimize the sterol composition of cell membranes rapidly under conditions of stress. As in humans, a Niemann-Pick protein NCR1 integrates sterols into the lysosomal membrane prior to further distribution as part of the mechanism of sterol homeostasis. Some antifungal drugs are targeted against ergosterol, either by binding to it to cause damaging cellular leakage, or to prevent its synthesis from lanosterol.

Many mutants defective in ergosterol biosynthesis have been isolated, and these have yielded a great deal of information on the features of the sterol molecule required for its structural role in membranes of yeast and fungi. Ergosterol stabilizes the liquid-ordered phase in the same manner as cholesterol and also forms raft microdomains with sphingolipids in membranes, whereas lanosterol does not. It is also evident that ergosterol has a multiplicity of functions in the regulation of yeast growth.

Under some conditions, especially those that retard growth, a high proportion of the sterols in yeasts can be in esterified form, where they are stored in lipid droplets. Ergosterol esters are synthesized in yeast by enzymes (ARE1 and ARE2), which are related to ACAT-1 and ACAT-2 that perform this function in animals, and both transfer an activated fatty acid to the hydroxyl group at the C3-position of a sterol molecule. In addition, specific sterol ester hydrolases that catalyze the reverse reaction have been characterized from yeasts, two in lipid droplets and one at the plasma membrane. Many fungal species and slime molds contain steryl glycosides (ergosteryl β-monoglucopyranosides in the former), but they are present at very low levels only in the widely studied yeast *Saccharomyces cerevisiae*.

Most fungi conjugate the 3β-hydroxyl group of ergosterol with aspartate in an RNA-dependent reaction catalyzed by an ergosteryl-3β-*O*-L-aspartate synthase, with the reverse reaction using a dedicated hydrolase. A phylogenomic study has shown that this pathway is conserved across higher fungi (except *S. cerevisiae*), including pathogens, and it has been suggested that these reactions constitute a homeostasis system with a potential impact upon membrane remodeling, trafficking, antimicrobial resistance, and pathogenicity.

21.5.3.8: BACTERIAL STEROLS

Hopanoids take the place of sterols in many species of bacteria, but it has long been recognized that some bacteria take up cholesterol and other sterols from host animals for use as membrane constituents. Indeed, an external source of sterols is required for growth in species of *Mycoplasma*. In addition, there have been a number of reports of the biosynthesis of sterols by various bacterial species, although a high proportion of these appears now to have been discounted because of fungal contamination. In particular, the possibility of sterol biosynthesis in cyanobacteria has been controversial, and molecular biology studies have yet to detect the presence of the required enzyme squalene epoxide cyclase.

That said, there is good evidence that a few species of prokaryotes at least have the capacity to synthesize sterols *de novo*. Among the eubacteria, certain methylotrophs (*Methylobacterium* and *Methylosphaera* species) produce mono- and dimethyl sterols, including lanosterol. Similarly, some soil bacteria produce 4-desmethylsterols. It has now been established from gene sequence studies that a few bacteria contain enzymes of the sterol biosynthesis pathway such as oxidosqualene cyclase, but as these have no obvious evolutionary link it seems probable that they were acquired via lateral transfer from eukaryotes.

21.5.3.9: PLANT STEROLS IN THE HUMAN DIET

The absorption of dietary plant sterols and stanols in humans is low (0.02-3.5%) compared to cholesterol (35-70%), although there are similar amounts in an average Western diet. The explanation is believed to be that the Niemann-Pick C1-like protein 1 (NPC1L1), which is responsible for cholesterol absorption in enterocytes does not take up plant sterols efficiently, while two transporters (ABCG5 and ABCG8) redirect any that are absorbed back into the intestinal lumen. In some rare cases, increased levels of plant sterols in plasma serve as markers for an inherited lipid storage disease (phytosterolemia) caused by mutations in the enterocyte transporters. Among many symptoms, accelerated atherosclerosis is often reported although the reasons for this are not clear. There is evidence that while plant sterols can substitute for cholesterol in maintaining membrane function in mammalian cells, they can exert harmful effects by disrupting cholesterol homeostasis. This may be relevant to the brain especially, since phytosterols are able to cross the blood-brain barrier, although they cannot be oxidized enzymatically because of the alkyl moiety on C24. In contrast, dietary supplements of plant sterols have been reported to have anti-cancer effects.

Substantial amounts of phytosterols are available as by-products of the refining of vegetable oils and of tall oil from the wood pulp industry. As it appears that they can inhibit the uptake of cholesterol from the diet and thereby reduce the levels of this in the plasma low-density lipoproteins, there is an increasing interest in such commercial sources of plant sterols to be added as "nutraceuticals" to margarines and other foods, Hydrogenated phytosterols or "**stanols**" are also used for this purpose, and studies suggest they are as effective as sterols in reducing LDL cholesterol. The consensus amongst experts in the field (including the FDA in the USA) is that such dietary supplements do indeed have the effects claimed and such claims can be used in advertising of commercial products, with the important caveat that there are no randomized, controlled clinical trial data that establish ensuing benefits to health, especially with respect to cardiovascular disease. Other pharmacological effects are under investigation, and there may be beneficial effects for the development of the human fetus and newborn, and for the treatment of non-alcoholic steatohepatitis, inflammatory bowel diseases ,and allergic asthma.

It is not clear whether oxy-phytosterols are generated in animal tissues, but those produced by enzymatic or non-enzymatic means can enter the food chain, especially when they are produced during cooking. Although they are not efficiently absorbed, 7-keto-sitosterol and 7-keto-campesterol have been detected in human plasma and have the potential to exert a variety of biological effects. For example, they have pro-atherogenic and pro-inflammatory properties in animal models.

21.6: BIOSYNTHESIS OF ISOPRENOIDS

By William (Bill) W. Christie and Henry Jakubowski.

This section is an abbreviated and modified version of material from the Lipid Web, an introduction to the chemistry and biochemistry of individual lipid classes, written by William Christie.

21.6.1: ISOPRENOIDS: 1. TOCOPHEROLS AND TOCOTRIENOLS (VITAMIN E)

Tocopherols and tocotrienols constitute a series of related benzopyranols (or methyltocols) that are synthesized in plants and other photosynthetic organisms, where they have many important functions but especially as part of a complex web of antioxidants that protect plants from the activities of reactive oxygen species (ROS). First described in 1922 as a dietary factor essential to prevent fetal reabsorption in rats, it was soon understood that plants contained a fat-soluble vitamin (vitamin E) that is essential for innumerable aspects of animal development. Of these many related molecules, only one form, i.e., **α-tocopherol**, is recognized as having vitamin E activity in humans in that it prevents a spectrum of human deficiency diseases termed 'ataxia', which is characterized by very low concentrations of α-tocopherol in plasma. The progression of the disease can be prevented by the administration of α-tocopherol only, although the pathogenic mechanism appears to be uncertain. While all tocopherols are known to be powerful lipid-soluble antioxidants *in vitro* at least, α-tocopherol has indirect roles in signal transduction and gene expression in animal tissues. On the other hand, specific functions (non-vitamin E) for other tocopherol forms and their metabolites in animal tissues are now being revealed. Vegetable oils are a major dietary source of vitamin E for humans.

21.6.1.1: STRUCTURE AND OCCURRENCE

In the tocopherols, the C_{16} side chain is saturated, and in the **tocotrienols** it contains three *trans* double bonds. Together, these two groups are termed the **tocochromanols.** In essence, the tocopherols have a 20-carbon phytyl tail (including the pyranol ring) and four chiral centers in total, with variable numbers of methyl groups attached to the benzene ring, while the tocotrienols have a 20-carbon geranylgeranyl tail with double bonds at the 3', 7' and 11' positions relative to the ring system. Tocopherols contain three chiral carbons, one at C2 in the chromanol ring and two in the side chain at C4′ and C8′ with *R,R,R* stereochemistry. The four main constituents of the two classes are termed - *alpha* (5,7,8-trimethyl), *beta* (5,8-dimethyl), *gamma* (7,8-dimethyl) and *delta* (8-methyl). In contrast to the tocopherols, the tocotrienols have only one chiral center. **Plastochromanol-8** is an analog of γ-tocotrienol with a much longer side-chain. Their structures are shown in Figure 21.6.1.

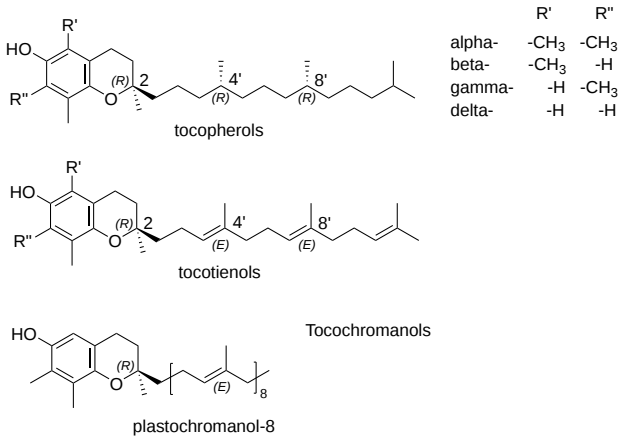


Figure 21.6.1: Tocochromanols

The tocochromanols are only synthesized by plants and other oxygenic photosynthetic organisms, such as algae and some cyanobacteria, but they are essential components of the diet of animals. Of these, only natural *R,R,R*-α-tocopherol is now designated ‘vitamin E’, as explained below, although the other tocopherols are sometimes termed ‘vitamers’ (some claim incorrectly - nor should all forms be termed isomers strictly speaking). In the USA, the current recommended dietary allowance (RDA) is 15 mg α-tocopherol daily for adults. In plants, there is a great range of tocochromanol contents and compositions, and photosynthetic plant tissues contain from 10 to 50 μg tocochromanols per g fresh weight. α-Tocopherol only is present in photosynthetic membranes of plant leaves, while γ-tocopherol and other forms are found principally in fruits, seeds, and nuts. While tocopherols are present in all photosynthetic organisms, the tocotrienols are found only in certain plant families.

Seed oils are a major source for the human diet and the compositions of tocopherols in some unrefined oils are listed in **Table 1**. Sunflower and olive oils are good sources of α-tocopherol and palm oil of the tocotrienols. In general, tocotrienols tend to be abundant only in seeds and fruits, especially of monocots such as wheat, rice, and barley, though a major commercial source is palm oil. In leaf tissue, α-tocopherol is often the main form, while γ-tocopherol is the primary tocopherol of many seeds. Plastochromanol-8 was first found in leaves of the rubber tree (*Hevea brasiliensis*) but has since been found in many other plants including rapeseed and maize, but usually at lower levels

than of the tocopherols. In addition, tocopherol esters of fatty acids occur in plant tissues, where they may be an inert storage form, but unesterified tocopherols are not released during digestion in animals so they may not make a contribution to vitamin E activity.

Table 21.6.1: **Tocopherol and tocotrienol contents (mg/Kg) in some seed oils.**

	α-T*	β-T	γ-T	δ-T	α-TT*	β-TT	γ-TT	δ-TT
palm	89	-	18	-	128	-	323	72
soybean	100	8	1021	421	-	-	-	-
maize	282	54	1034	54	49	8	161	6
sunflower	670	27	11	1	-	-	-	-
rapeseed	202	65	490	9	-	-	-	-

* Abbreviations: T, tocopherol; TT, tocotrienol

Data from: Gunstone, F.D., Harwood, J.L. and Padley, F.B. *The Lipid Handbook (Second Edition)* (Chapman & Hall, London)(1994).

An unusual tocopherol that has been termed marine-derived α-tocomonoenol is found together with α-tocopherol in a wide range of marine fish species, where it appears to be a more efficient scavenger of free radicals at low temperatures. A related isomer with a Δ11 double bond has been found in palm oil and kiwi fruit. While pumpkin seeds contain both α- and γ-tocomonoenols, other plant species contain β, γ- and δ-tocomonoenols with unsaturation in the terminal isoprene unit of the side chain. Tocochromenols or 3,4-dehydrotocopherols, i.e., with a double bond in the pyranol ring, are also known in addition to more complex tocopherol-like molecules. The structures of tocomonoenols are shown in Figure 21.6.2.

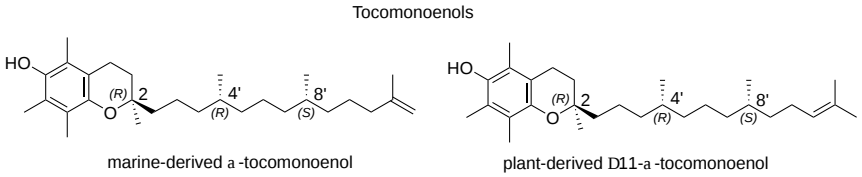


Figure 21.6.2: Tocomonoenols

α-Tocopherol is a minor but ubiquitous component of the lipid constituents of animal cell membranes (non-raft domains), with estimates ranging from one molecule of tocopherol to from 100 to 1000 molecules of phospholipids, depending on the membrane. The hydrophobic tail lies within the membrane, as might be expected, and the polar head group is orientated towards the surface but below the level of the phosphate moieties of the phospholipids. There may be some limited hydrogen bonding between the hydroxyl groups and phosphate depending on the degree of hydration of the membrane. On the other hand, there is a strong affinity of α-tocopherol for polyunsaturated fatty acids, where the chromanol unit may interact with the double bonds, suggesting that tocopherol is located deep within the membrane.

α-Tocopheryl phosphate has recently been detected at low levels in plasma, liver, and adipose tissue. Its structure is shown in Figure 21.6.3. Together with catabolic tocopherol metabolites, it has important biological properties.

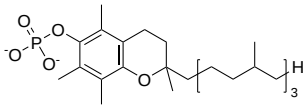


Figure 21.6.3: α-Tocopheryl phosphate

During the refining of vegetable oils, much of the natural tocopherols is lost or destroyed. Most commercial vitamin E is therefore prepared by chemical synthesis with trimethylhydroquinone and phytyl bromide as the precursors. The resulting product is a mixture of eight stereoisomers (from *R*,*R*,*R*- to *S*,*S*,*S*-methyl groups) of α-tocopherol, with the various stereoisomers differing by a factor of two in biologic activity, as a consequence of the stereochemistry of position 2 in the chromanol ring (i.e., 2*S*-α- compared to 2*R*-α-tocopherol). It is usually administered as the acetate derivative *in vivo*. Tocopherols are not usually regarded as effective antioxidants in the polyunsaturated seed oils of commerce, and at higher concentrations can even act as pro-oxidants, although the reasons for this are not understood.

21.6.1.2: BIOSYNTHESIS AND FUNCTIONS OF TOCOCHROMANOLS IN PLANTS

The mechanism of biosynthesis of tocopherols has been elucidated and involves the coupling of phytyl diphosphate with homogentisic acid (2,5-dihydroxyphenylacetic acid), followed by cyclization and methylation reactions. The plant chloroplast is the site of biosynthesis, and most of the enzymes are located on the inner membrane of the chloroplast envelope, although there is increasing evidence that **plastoglobules** associated with the thylakoid membrane may be involved. The pathway for biosynthesis of tocopherols is shown in Figure 21.6.4.

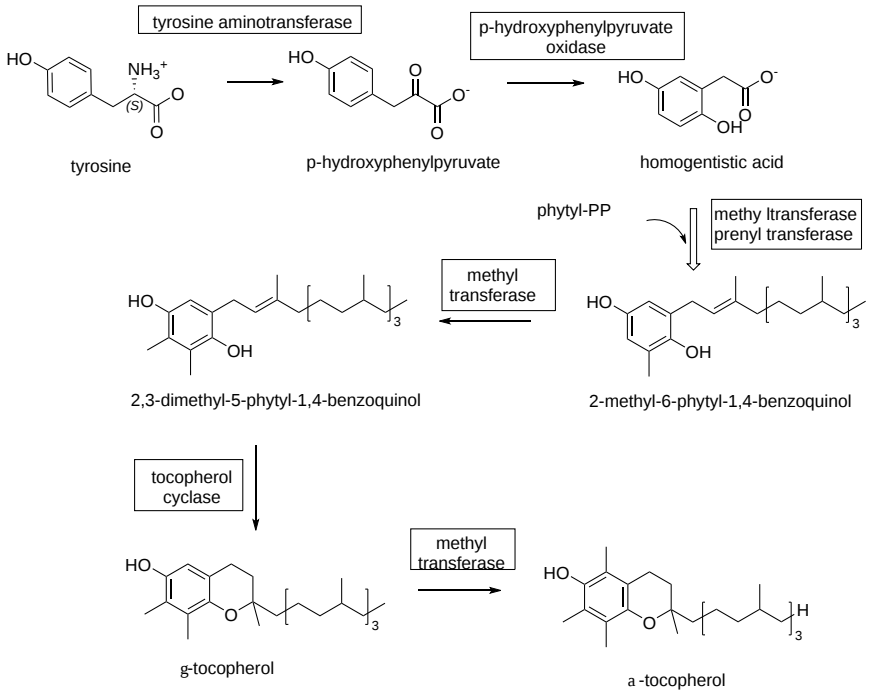


Figure 21.6.4: Synthesis of tocopherols

The aromatic amino acid tyrosine can be considered the basic precursor, and this is oxidized to *p*-hydroxypyruvic acid, which in the first committed step is converted to homogentisic acid by the enzyme *p*-hydroxyphenylpyruvate dioxygenase. Homogentisic acid is condensed with phytyl diphosphate, derived from **phytol** obtained from hydrolysis of chlorophyll, in a reaction catalyzed by a prenyl transferase to yield 2-methyl-6-phytyl-plastoquinol, which is first methylated to form 2,3-dimethyl-5-phytyl-1,4-benzoquinol and then converted by the enzyme tocopherol cyclase to γ-tocopherol. A further methylation reaction produces α-tocopherol, while modifications to the pathway produce β- and δ-tocopherols, together with **plastoquinones** and thence plastochromanol-8. Tocotrienols and tocomonoenols result from a similar series of reactions but with geranylgeranyl diphosphate and tetrahydro-geranylgeraniol diphosphate, respectively, as substrates in the condensation step. The isoprenoid precursors are synthesized in the plastid also by the non-mevalonate or 'MEP' pathway.

In plants, tocopherols are found mainly in the chloroplasts of green tissues, but they are also present in seeds, fruits, roots and tubers. They are especially important as antioxidant molecules, limiting the damage from photosynthesis-derived reactive oxygen species during conditions of oxidative stress, including high-intensity light stress, and the mechanisms for this antioxidant activity are discussed below. However, recent studies seem to suggest that they are just one of a number of different components that are involved in photoprotection. Certainly, any tocochromanol peroxy radicals formed must be converted back to the original compounds by the concerted action of other plant antioxidants, for example by ascorbate, glutathione, ubiquinol or lipoic acid, and antioxidant enzymes, including superoxide dismutase, catalase, and peroxidases. Tocopherols are essential for the control of non-enzymatic lipid peroxidation during seed dormancy and germination of seedlings. In their absence, elevated levels of malondialdehyde and phytoprostanes are formed, and there can be inappropriate activation of plant defense responses.

There is evidence that tocopherols play a part in intracellular signaling in plants in that they regulate the amounts of **jasmonic acid** in leaves, via modulating the extent of lipid peroxidation and gene expression, and so influence plant development and stress responses. Thus, by controlling the degree of lipid peroxidation in chloroplasts (redox regulation), they limit the accumulation of lipid hydroperoxides required for the synthesis of jasmonic acid, which in turn regulates the expression of genes that affect a number of abiotic stress conditions, including drought, salinity and extremes of temperature. The translocation of enzymes to the plasma membrane is regulated by tocopherols, possibly by modulating protein-membrane, altering membrane microdomains (lipid **rafts**), or by competing for common binding sites within lipid transport proteins. In addition, tocopherols are required for the development of the cell walls in phloem transfer cells under cold conditions. It appears that α- and γ-tocopherol and the tocotrienols may each have distinct functions. For example, γ-tocopherol is reportedly more potent than α-tocopherol in protecting plants from the harmful effects of osmotic stresses and is important for the longevity of seeds. Efforts are underway to increase the tocopherol levels in plants by selective breeding and genetic manipulation with the aim of producing crops with greater potential health benefits to consumers and perhaps for the plants *per se*.

21.6.1.3: TOCOPHEROLS METABOLISM IN ANIMALS

In animals, the first step in the digestion of tocopherols is their dissolution with other lipids in mixed micelles in the intestines. All tocopherol forms are absorbed to a similar extent in the intestines by means of transporters in the enterocyte apical membrane that have a broad specificity for hydrophobic molecules, such as cholesterol, vitamin D, and carotenoids. These include scavenger receptor class B type

I (SR-BI), the CD36 protein, and NPC1-like intracellular cholesterol transporter 1 (NPC1L1). However, some passive diffusion cannot be ruled out. Transport across the enterocyte may involve cytoplasmic transporters or clathrin-coated vesicles before the tocopherols are incorporated into chylomicrons in free form in the Golgi apparatus for release into the lymph. At the liver, α-tocopherol specifically is taken up from the chylomicrons by a receptor-mediated mechanism with the aid of a specific tocopherol-binding protein (the α-tocopherol transfer protein (α-TTP)), i.e., a 30,500 Da cytosolic protein that has a marked affinity for α-tocopherol and can enhance its transfer between membranes. This recognizes α-tocopherol by the three methyl groups and hydroxyl on the chromanol ring and by the structure and orientation of the phytyl side chain. It is the chief regulator of whole body α-tocopherol status and is expressed primarily in the cytosol of hepatocytes in the liver, but has been reported (in much lower concentrations) in other tissues, such as the placenta.

α-TTP ultimately regulates the egress of α-tocopherol selectively from hepatocytes with the aid of the ATP-binding cassette proteins ABCA1 and ABCG for conveyance in the plasma **lipoproteins**, mainly in the very-low-density lipoproteins or VLDL (and thence to LDL) and HDL in humans, to the peripheral tissues (together with much smaller amounts of γ-tocopherol). Most of the other tocopherol forms are directed toward catabolism. Once in the circulation, tocopherol can exchange spontaneously between membranes and lipoproteins, and no specific transport protein for vitamin E in plasma has yet been described. Transfer of tocopherols from the VLDL to peripheral tissues occurs as triacylglycerols are hydrolyzed by the enzyme lipoprotein lipase, while that in LDL is processed via the LDL receptor-mediated uptake pathway. Within cells of peripheral tissues, including the central nervous system, α-TTP functions in transporting α-tocopherol to wherever it is required in membranes, a process that appears to be aided by **phosphatidylinositol** metabolites. In the brain, tocopherol is transported by apo-E rich lipoproteins. Concentrations of tocopherols can vary appreciably amongst tissues, with most in adipose tissue and adrenals, less in kidney, heart, and liver, and least in the erythrocytes.

The "α-tocopherol salvage pathway" is partly due to this process and partly to selective oxidation (see below), and the result is a 20- to 30-fold enrichment of α-tocopherol in plasma (average concentration 22-28 μM) relative to the other tocopherols. Thus, the process of conservation of one specific tocopherol appears to determine the relative vitamin E activities of the tocopherols and tocotrienols *in vivo*, rather than their individual potencies as antioxidants as measured in model systems *in vitro*. Only α-tocopherol (including synthetic material) or natural mixtures containing this can be sold under the label 'Vitamin E'. γ-Tocopherol is the second most abundant form in plasma, and it is present in relatively greater proportions in the skin, adipose tissue, and skeletal muscle, where it has some specific biological properties that are distinct from those of α-tocopherol. Although tocotrienols are more potent antioxidants *in vitro*, they are not usually detected in tissues, although they are believed to have some important functions.

Catabolism: The unwanted surplus of tocochromanols other than α-tocopherol may be excreted in the urine and feces in the form of carboxy-chromanols, including the so-called 'Simon metabolites' - tocopheronic acids (carboxyethylhydroxychromans, CEHC) and tocopheronolactones, after oxidative cleavage of much of the phytyl tail, although these are normally detected in the form of conjugates as sulfate or glucuronidate esters, the forms in which they are excreted in feces and urine. For example for illustrative purposes in liver cells, the first step in the catabolism of γ-tocopherol is ω-hydroxylation by cytochrome P450 (CYP4F2) at the 13' carbon to form γ-13'-hydroxychromanol in the endoplasmic reticulum, followed by ω-oxidation in the peroxisomes to produce γ-13'-carboxychromanol, and finally by stepwise β-oxidation in the mitochondria to cut off two or three carbon moieties from the phytyl chain in each cycle. These steps are shown in Figure 21.6.5.

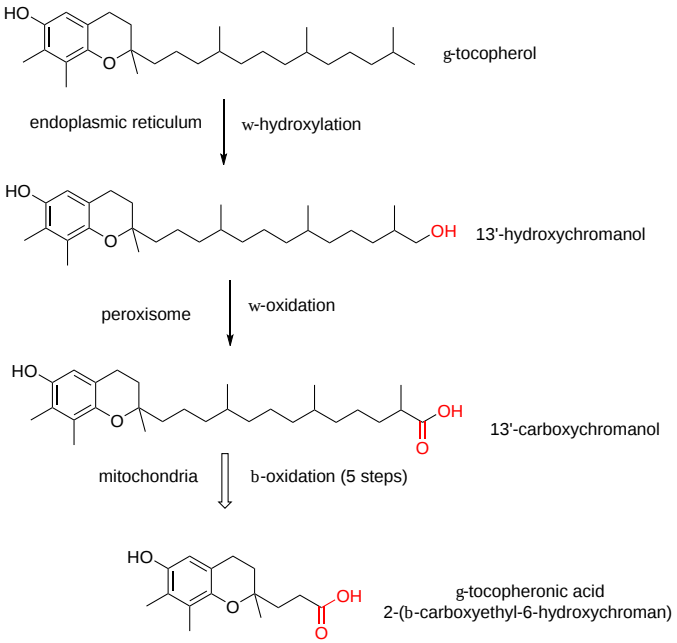


Figure 21.6.5: Catabolism of γ-tocopherol

Various carboxychromanol intermediates have been identified for all of the tocopherols together with forms in which the hydroxyl group is sulfated in human cell cultures *in vitro*; sulfated carboxychromanols are the main tocopherol metabolites in the plasma of rodents. As the vitamin E ω-hydroxylase has a high affinity for the tocopherols other than the α-form and does not attack that bound to the α-tocopherol transfer protein, this provides a further specific enhancement of the α-tocopherol concentration in plasma relative to the others. Some of these catabolic metabolites may have some biological activity in their own right. For example, carboxyethylhydroxychromans derived from γ-tocopherol were reported to induce apoptosis in cancer cells and to have anti-inflammatory effects by inhibition of cyclooxygenases and 5-lipoxygenase (see below). Tocotrienols are catabolized in a similar manner, but with additional steps in which the double bonds are reduced prior to oxidation; the final carboxyethylchromanols are the same as for tocopherols.

21.6.1.4: TOCOPHEROLS AS ANTIOXIDANTS

Although the syndrome associated with a lack of vitamin E in the diet of animals has been known for decades, the mode of action and specific location of tocopherols in cell membranes are not clearly understood. Several theories have been proposed to explain the functions of vitamin E in animal cells. From studies *in vitro*, it has long been believed that a major task is to act as an **antioxidant** to inhibit, decrease, delay, or prevent oxidative damage to unsaturated lipids or other membrane constituents and thence to tissues by scavenging free radicals. For example, vitamin E administration can prevent lipid peroxidation and hepatotoxicity upon exposure to the free radical-generating agent carbon tetrachloride. Lipid peroxidation is also a cause of **ferroptosis**, an iron-dependent form of nonapoptotic cell death. However, tocopherols have functions other than as antioxidants. In non-biological systems such as foods, cosmetics, pharmaceutical preparations, etc., tocopherols are invaluable as antioxidant additives.

Because of their lipophilic character, tocopherols are located in the membranes or with storage lipids where they may be available immediately to interact with lipid hydroperoxides, such as those described in more detail in our web pages on **isoprostanes**, **reactive aldehydes,** and **oxidized phospholipids**. In brief, **Reactive Oxygen Species (ROS)**, of which innumerable forms, exist can be derived by enzymatic or non-enzymatic means and produce superoxide anions and other peroxyl radicals. Superoxide radicals ($O_2^{\bullet-}$) ultimately generate highly toxic hydroxyl ($^{\bullet}OH$) or alkoxyl radicals, which can abstract a hydrogen atom from *bis*-allylic methylene groups of polyunsaturated fatty acids under aerobic conditions *in vivo* in animals and plants to generate lipid peroxyl radicals ($LOO^{\bullet}$) and hydroperoxy-fatty acids. **Singlet oxygen** ($^1O_2$ or O=O) is an especially important ROS (non-radical) in photosynthetic tissues of plants. As radical generation is not enzymatic, all methylene groups between two *cis* double bonds can potentially be involved in the reaction, although not necessarily to the same degree. Tocopherols react rapidly in a non-enzymic manner unlike many other cellular antioxidants, which are dependent on enzymes, to scavenge lipid peroxyl radicals, i.e., the chain-carrying species that propagate lipid peroxidation. In model systems *in vitro*, all the tocopherols (α > γ > β > δ) and tocotrienols are good antioxidants, with the tocotrienols being the most potent.

In general, the oxidation of lipids is known to proceed by a chain process mediated by such free radicals, in which the lipid peroxyl radical serves as a chain carrier. In the initial step of chain propagation, a hydrogen atom is abstracted from the target lipid by the peroxyl radical as shown in Figure 21.6.6.

$$LOO^{\bullet} + LH \longrightarrow LOOH + L^{\bullet} \quad (1)$$

$$L^{\bullet} + O_2 \longrightarrow LOO^{\bullet} \quad (2)$$

- where LH is a lipid, $LOO^{\bullet}$ is the lipid peroxyl radical and LOOH is the lipid hydroperoxide

Figure 21.6.6: Formation of lipid peroxyl radical

The main function of α-tocopherol is to scavenge the lipid peroxyl radical before it is able to react with the lipid substrate as shown in Figure 21.6.7.

$$LOO^{\bullet} + TOH \longrightarrow LOOH + TOO^{\bullet} \quad (3)$$

- where TOH is tocopherol and $TOO^{\bullet}$ is the tocopheroxyl radical

Figure 21.6.7: Scavenging of free radicals by tocopherol

The potency of an antioxidant is determined by the relative rates of reactions (1) and (2). When a tocopheroxyl radical is formed, it is stabilized by the delocalization of the unpaired electron about the fully substituted chromanol ring system rendering it relatively unreactive, thus preventing propagation of the chain reaction. This also explains the high first-order rate constant for hydrogen transfer from α-tocopherol to peroxyl radicals, as studies of the relative rates of chain propagation to chain inhibition by α-tocopherol in model systems have demonstrated that α-tocopherol is able to scavenge peroxyl radicals much more rapidly than the peroxyl radical can react with a lipid substrate.

In biological systems, oxidant radicals can spring from a number of sources, including singlet oxygen, alkoxyl radicals, superoxide, peroxynitrite, nitrogen dioxide, and ozone. α-Tocopherol is most efficient at providing protection against peroxyl radicals in a membrane environment.

The reaction of the tocopheroxyl radical with a lipid peroxyl radical, as illustrated, yields 8α-substituted tocopherones, which are readily hydrolyzed to 8α-hydroxy tocopherones that rearrange spontaneously to form α-tocopherol quinones. In an alternative pathway, the

tocopheroxyl radical reacts with the lipid peroxyl radical to form epoxy-8α-hydroperoxytocopherones, which hydrolyze and rearrange to epoxyquinones. Tocopherol dimers and trimers may also be formed as minor products. These reactions are shown in Figure 21.6.8.

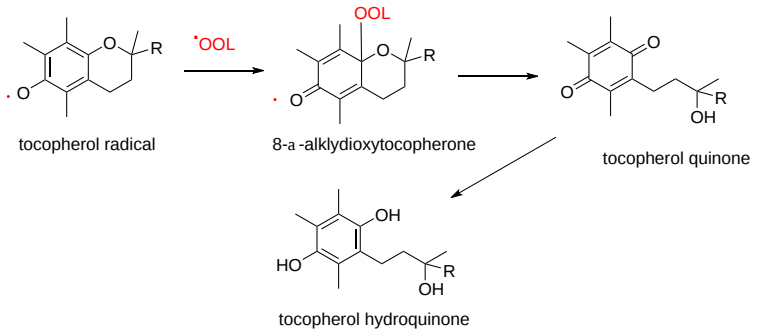


Figure 21.6.8: Reaction of the tocopherol radical

Free radical-mediated lipid peroxidation is the major pathway of lipid oxidation taking place in humans, and α-tocopherol is a major antioxidant, but it does not scavenge the nitrogen dioxide radical, carbonate anion radical, and hypochlorite efficiently. Vitamin E forms with an unsubstituted 5-position, such as γ-tocopherol, are an exception to the rule that the various tocopherols have similar antioxidant properties in that they are able to trap electrophiles, including **Reactive Nitrogen Species (RNS)**, which are enhanced during inflammation. The enzyme nitric oxide synthase is capable of continuously producing a large amount of nitric oxide ($NO^{\bullet}$), which can react with superoxide to produce peroxynitrite ($ONOO^-$), a potent and versatile oxidant that can attack a wide range of biological targets. It induces lipid peroxidation and nitrates aromatic compounds and unsaturated fatty acids while isomerizing *cis*-double bonds in fatty acids to the *trans*-configuration. γ-Tocopherol is superior to α-tocopherol in detoxifying the NO_2 radical and peroxynitrite with formation of 5-nitro-γ-tocopherol, as shown in Figure 21.6.9.

Figure 21.6.*xx*:

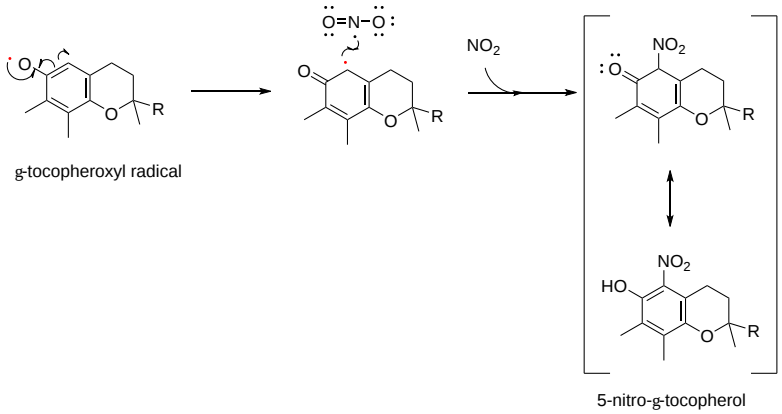


Figure 21.6.9: Radical reaction with γ-Tocopherol

This occurs *in vivo*, and the concentrations of 5-nitro-γ-tocopherol have been shown to be elevated in the plasma of subjects with coronary heart disease and in carotid-artery atherosclerotic plaque.

In plant and animal tissues, tocopherols can be regenerated from the tocopheroxyl radicals in a redox cycle mediated by a number of endogenous antioxidants, including vitamins A and C and coenzyme Q, and this must greatly extend their biological potency. Vitamin C (ascorbic acid) may be especially important in aqueous systems, although it may also act at the surface of membranes, to regenerate α-tocopherol, while in turn being oxidized to dehydroascorbic acid. This can be regenerated to the reduced form by glutathione (GSH) with the production of glutathione disulfide (GSSG), which can subsequently be enzymatically reduced by glutathione reductase with NAD(P)H as a cofactor. In plants, an NAD(P)H-dependent quinone oxidoreductase is involved at an early stage of the regeneration process, while tocopherol cyclase, an enzyme involved in the biosynthesis of tocopherols, re-introduces the chromanol ring. These linked cycles (the antioxidant network) are shown in Figure 21.6.10.

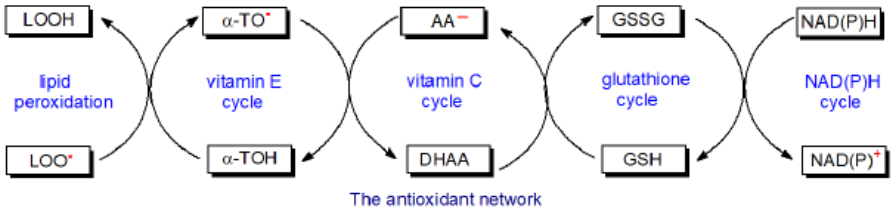


Figure 21.6.10: The antioxidant network

Thus, tocopherols are only one component of a complex web of metabolites and enzymes in tissues that have antioxidant activities and act by various mechanisms, including the stimulation of genes involved in signaling responses to environmental stresses. One antioxidant mechanism involves the removal of free radicals and reactive species by enzymes such as superoxide dismutase, catalase, and glutathione peroxidase, while electron donors, such as glutathione, tocopherols, ascorbic acid, vitamin K, coenzyme Q, and thioredoxin, scavenge free radicals also. Metal-binding proteins such as transferrin, metallothionein, haptoglobin, and ceruloplasmin have antioxidant activity by sequestering pro-oxidant metal ions, such as iron and copper, although some metals such as selenium and zinc are in fact antioxidants. Other antioxidants, including flavonoids, carotenoids, and phenolic acids in addition to tocopherols, enter animal tissues via the food chain. Although the discussion here has been limited to the effects upon lipids, free radicals can cause damage to proteins, DNA, and indeed virtually any native substance in living organisms.

21.6.1.5: BIOLOGICAL FUNCTIONS OF TOCOCHROMANOLS IN ANIMALS

Vitamin E deficiency has been detected in patients with fat malabsorption, cystic fibrosis, Crohn's disease, liver disease, and pancreatic insufficiency, and in premature infants. Impairment of the normal functions of the immune system has been demonstrated in animals and humans in vitamin E deficiency, and this can be corrected by vitamin E repletion. It also displays activity against nonalcoholic hepatosteatosis. Although there are various proposals for the pathogenic mechanism, none as yet appears to be generally accepted. After the discovery of the effects of vitamin E on fertility in studies with laboratory animals, its importance was documented for the development of tissues and organs such as brain and nerves, muscle and bones, skin, bone marrow, and blood, most of which are specific to α-tocopherol. However, there is no evidence for an effect of vitamin E on fertility in humans, as was originally found in the rat. The rare genetic disorder “Ataxia with Isolated Vitamin E Deficiency” or “AVED” is the result of mutations in the gene coding for α-TTP. It is caused by the death of cerebellar Purkinje cells, but administration of α-tocopherol prevents this and the subsequent development of clinical symptoms of the disease.

There appears to be little doubt that tocopherols inhibit many of the enzymes associated with inflammation *in vitro* in animals, and may contribute to the amelioration and treatment of some chronic diseases. However, it has been argued that data on the effects of vitamin E on biomarkers of oxidative stress *in vivo* are inconsistent. Oxidized metabolites of vitamin E, i.e., that have reacted as antioxidants, are barely detectable in tissues, and vitamin E maintenance *in vivo* does not appear to have been clearly associated with its regeneration. There appear to be significant differences between results obtained in studies with laboratory animals in comparison to those in humans. Thus, suggestions that dietary supplements of vitamin E may reduce the rate of oxidation of lipids in low-density lipoproteins in humans and thence the incidence or severity of atherosclerosis have not been confirmed by clinical intervention studies, although benefits in some conditions have been claimed. Indeed, there are suggestions that excessive vitamin E supplementation may even be harmful. One study has suggested that relatively high doses of natural α-tocopherol over a long period are required to demonstrate a significant reduction in the levels in the urine of F_2 **isoprostanes**, which are considered to be the most reliable marker for oxidative stress *in vivo*. While there are many fat-soluble antioxidants in the diet, only α-tocopherol is a vitamin. It has even been suggested that tocopherol may be protected from functioning as an antioxidant in some tissues *in vivo* through a network of cellular antioxidant defenses, such that tocopherols are utilized only when other antioxidants are exhausted, although there is no experimental proof of this hypothesis.

At the cellular level, *RRR*-α-tocopherol has been shown to inhibit protein kinase C, and in the process, it inhibits the assembly and radical-producing activity of NADPH oxidase in monocytes. Similarly, vitamin E suppresses the expression of xanthine oxidase, a source of reactive oxygen species, in the liver. It is thus possible that α-tocopherol is able to diminish the levels of free radicals by preventing their production and not by scavenging them. Its physical presence in membranes adjacent to polyunsaturated fatty acids may thus limit autoxidation.

With the discovery that the antioxidant effects of various tocopherols and tocotrienols have little relation to their vitamin E activities *in vivo* has come the realization that they have other functions in tissues, most of which are specific to α-tocopherol. Most current research is concerned with how vitamin E and its metabolites act in signaling and controversially in the regulation of gene activity. While it is certainly true that most other vitamins are essential cofactors for specific enzymes or transcription factors, no receptor that binds specifically to vitamin E has yet been discovered. By preventing the increase of peroxidized lipids that alter both metabolic pathways and gene expression profiles within tissues and cells, it may act indirectly as a regulator of genes connected with tocopherol catabolism, lipid uptake, collagen synthesis, cellular adhesion, inflammation, the immune response and cell signaling. Vitamin E affects a number of transcription factors in

this manner, including peroxisome proliferator-activated receptor gamma (PPARγ), nuclear factor erythroid-derived 2 (NRF2), nuclear factor kappa B (NFκB), RAR-related orphan receptor alpha (RORα), estrogen receptor beta (ERβ), and the pregnane X receptor (PXR).

α-Tocopherol and its metabolites are believed to modulate the activity of several enzymes involved in signal transduction, including protein kinases and phosphatases, lipid kinases and phosphatases, and other enzymes involved in lipid metabolism, but especially those with inflammatory properties such as lipoxygenases, cyclooxygenase-2, and phospholipase A_2. While the credentials of tocopherols as antioxidants *in vivo* have been doubted, this does not preclude a role in the inhibition of oxidative enzymes, especially in relation to the function of the immune system. For example, vitamin E regulates T cell function directly by its effects upon T cell membrane integrity, signal transduction, and cell division, and it also functions indirectly by affecting eicosanoids and related inflammatory mediators generated from other immune cells. Various tocopherols and tocotrienols have been shown to suppress COX-2 involvement in prostaglandin (PGD_2 and PGE_2) synthesis in lipopolysaccharide-activated macrophages.

In addition, it has been established that the 13'-carboxy metabolite of α-tocopherol (α-T-13'-COOH) and other tocopherol ω-carboxylates are potent allosteric inhibitors of 5-lipoxygenase, a key enzyme in the biosynthesis of the inflammatory **leukotrienes**. α-T-13'-COOH accumulates in immune cells and inflamed exudates both *in vitro* and *in vivo* in mice, and it has even been suggested that the immune regulatory and anti-inflammatory functions of α-tocopherol depend on this endogenous metabolite. The structure of α-T-13'-COOH is shown in Figure 21.6.11.

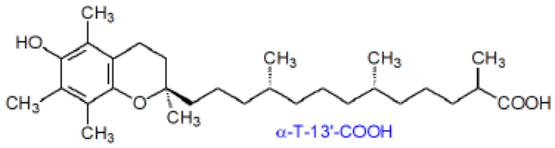


Figure 21.6.11: Structure of α-T-13'-COOH

α-Tocopherol has a stimulatory effect on the dephosphorylation enzyme, protein phosphatase 2A, which cleaves phosphate groups from protein kinase C, leading to its deactivation. The mechanism may involve the binding of vitamin E directly to enzymes in order to compete with their substrates, or it may change their activities by redox regulation. It may also compete for common binding sites within lipid transport proteins, and so may alter the traffic of lipid mediators indirectly with effects upon their signaling functions and enzymatic metabolism. For example, it binds to albumin as well as to a specific α-tocopherol-associated protein (TAP), and in the latter form especially it inhibits the phosphoinositide 3-kinase. It has been suggested that vitamin E may have a secondary role in stabilizing the structure of membranes, or it may interact with enzymes in membranes to interfere with binding to specific membrane lipids, or it may affect membrane microdomains such as lipid **rafts**.

Evidence suggests that the biological activities of β-, γ- and δ-tocopherols do not reflect their behavior as chemical antioxidants, but anti-inflammatory, antineoplastic, and natriuretic actions have been reported. Some non-antioxidant effects of **γ-tocopherol** in tissues in relation to reactive nitrogen oxide species have been observed, but the specificity of these *in vivo* is not yet certain. In addition, anti-inflammatory properties have been described that have been attributed to a chain-shortened metabolite. Beneficial effects against cancer cells *in vitro* have been observed that have been ascribed to scavenging of reactive nitrogen species, since such effects are not seen with α-tocopherol. On the other hand, vitamin E and its derivatives are believed to regulate tumor cells by activating the mitogen-activated protein kinase (MAPK) signaling pathway.

Tocotrienols have been shown to have neuroprotective effects and to inhibit cholesterol synthesis. They reduce the growth of breast cancer cells *in vitro*, possibly by influencing gene expression by interaction with the estrogen receptor-β. When administered in combination with either standard antitumor agents as in chemotherapy or with natural compounds with anticancer activity, they are reported to exert a synergistic antitumor effect on cancer cells. γ-Tocotrienol is reported to be an inducer of apoptosis via endoplasmic reticulum stress, while α-tocotrienol may be neuroprotective by inhibition of lipoxygenase activity. Although anti-obesity and anti-diabetic effects have been observed in mice, clinical trials with humans appear to have given inconclusive results. These properties are largely distinct from those of the tocopherols, and the pharmaceutical potential of tocotrienols against cancer, bone resorption, diabetes, and skin, cardiovascular and neurological diseases are currently being studied.

The biological functions of **α-tocopheryl phosphate** are slowly being revealed. In addition to being a possible storage or a transport (water-soluble) form of tocopherol, it is involved in cellular signaling and regulates a number of genes, including those involved in angiogenesis and vasculogenesis, in a different manner from α-tocopherol *per se*. As it lacks the free hydroxyl group, it cannot act directly as an antioxidant, and some consider it to be the biologically active form of the vitamin. It is certainly more active in a number of biological systems *in vitro* than α-tocopherol, so these effects cannot be ascribed to the hydrolyzed molecule, and in some instances, it is antagonistic to α-tocopherol, for example in its activity towards phosphatidylinositol 3-kinase. On the other hand, activation requires a kinase, while a phosphatase is needed to make the system reversible, but neither has yet been identified. Synthetic phosphate derivatives of γ-tocopherol and α-tocopheryl succinate are known to have potent anti-cancer properties.

21.6.2: ISOPRENOIDS: 2. RETINOIDS (VITAMIN A)

That a dietary factor was involved in visual acuity was known to the ancient Egyptians and Greeks, but it was the 1930s before the importance of the carotenoids and their metabolites was recognized, and β-carotene and retinol were fully characterized. It is now recognized that vitamin A activity now resides in the metabolites retinol, retinal and retinoic acid, and in several provitamin A carotenoids, most notably β-carotene. A share in the Nobel Prize for Medicine in 1967 was awarded to George Wald, who over many years showed how retinol derivatives (named for their function in the retina) constituted the chemical basis of vision. Now, it is recognized that retinol, retinoic acid, and their many metabolites have innumerable other functions in human metabolism from embryogenesis to adulthood, including growth and development, reproduction, cancer, and resistance to infection. They are important natural antioxidants with benefits to health, although some potentially harmful properties have been reported.

Carotenoids are a class of highly unsaturated terpenoids that occur in innumerable molecular forms (>1000). They are common colorful pigments of plants, fungi, and bacteria, of vital importance to photosynthesis, and as dietary constituents, they can add ornament to some animal species. Apart from acting as precursors of retinoids, carotenoids *per se* appear to have a relatively limited range of functions in animal tissues, but they are important to vision and as antioxidants, especially in the skin. They do of course have important functions in plants and lower organisms where they originate, but this topic can only be dealt with briefly here. Other fat-soluble vitamins **tocopherols (vitamin E)**, **vitamin K,** and **vitamin D** are discussed elsewhere.

21.6.2.1: OCCURRENCE AND BASIC METABOLISM OF CAROTENOIDS AND RETINOIDS

The term ‘vitamin A’ is used to denote **retinol** (or all-*trans*-retinol, sometimes termed 'vitamin A_1'), together with a family of biologically active C_{20} **retinoids** derived from this ('vitamers'). The structure of retinol is shown in Figure 21.6.12.

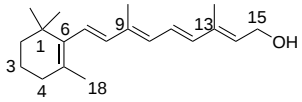


Figure 21.6.12: All-trans-retinol

These are only found in animal tissues, where they are essential to innumerable biochemical processes. However, they cannot be synthesized *de novo* in animals and their biosynthetic precursors are plant carotenoids with a β-ionone ring (provitamin A), C_{40} tetraterpenes of which β-carotene is most the efficient; it is an orange-red pigment that occurs in the photosynthetic tissues of plants and in seed oils. In the human diet in the developed world, plant sources tend to be less important than those from dairy products, meat, fish oils, and margarines, which provide vitamin A *per se*, although carrots and spinach are good sources of the provitamin. In the U.K., for example, all vegetable spreads must be supplemented with the same level of vitamin A (synthetic retinol or β-carotene) as is found in butter. While most research effort has been focused on retinoids, there is increasing interest in the biological activities of intact carotenoids in animal tissues.

The biosynthesis of carotenoids in plants via isopentenyl diphosphate and dimethylallyl diphosphate has much in common with that of the **plant sterols**, but this is too specialized a topic to be treated at length here. They have many important functions in plants, for example during photosynthesis or as precursors of plant hormones, and these are discussed below. Some crop plants with increased carotene levels are available with the aim of preventing vitamin A deficiency in the populations of developing countries, and further efforts are underway. Non-photosynthetic bacteria produce a different range of carotenoids, some with chain lengths other than C_{40} (C_{30} to C_{50}).

In animals (including humans), dietary carotenoids such as β-carotene are solubilized with other dietary lipids in mixed micelles with the aid of bile acids, and they are absorbed in the intestines in intact form by a process facilitated by specific receptor proteins. Dietary retinol and retinol esters are absorbed similarly in the intestines, but the latter are first hydrolyzed by pancreatic lipase. Conversion to retinoids leading ultimately to retinol esters occurs in the enterocytes, where dietary β-carotene is subjected to oxidative cleavage at its center, the first step of which is catalyzed by a cytosolic enzyme β-carotene-15,15'-oxygenase-1 (BCO1), which is specific for carotenes with a β-ionone ring, to yield two molecules of all-*trans*-retinal, which is reversibly reduced by a retinol reductase to retinol. Xanthophyll carotenoids are absorbed without cleavage mainly. These reactions are shown in Figure 21.6.13.

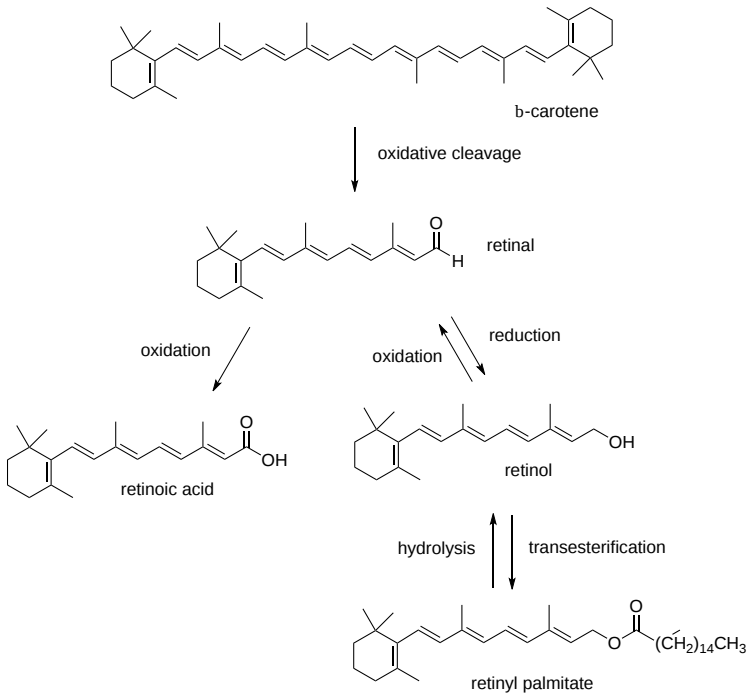


Figure 21.6.13: Synthesis of all-trans retinoids in the intestines and liver

jbc: https://www.jbc.org/article/S0021-92...842-6/fulltext

Based on the incorporation of ^{18}O into the products. it appears that the enzyme that introduces the oxygen atom, β-carotene-15,15'-oxygenase-1 (BCO1), into the cleavage products is a **dioxygenase** as both atoms of oxygen in dioxygen are incorporated into products. This is in contrast to a possible mechanism in which only one oxygen atom from dioxygen is added, with the other coming from H_2O if the enzyme acted as a monooxygenase. Figure 21.6.14 shows how oxygen could be introduced in the reaction catalyzed by BCO1 through both possible mechanisms.

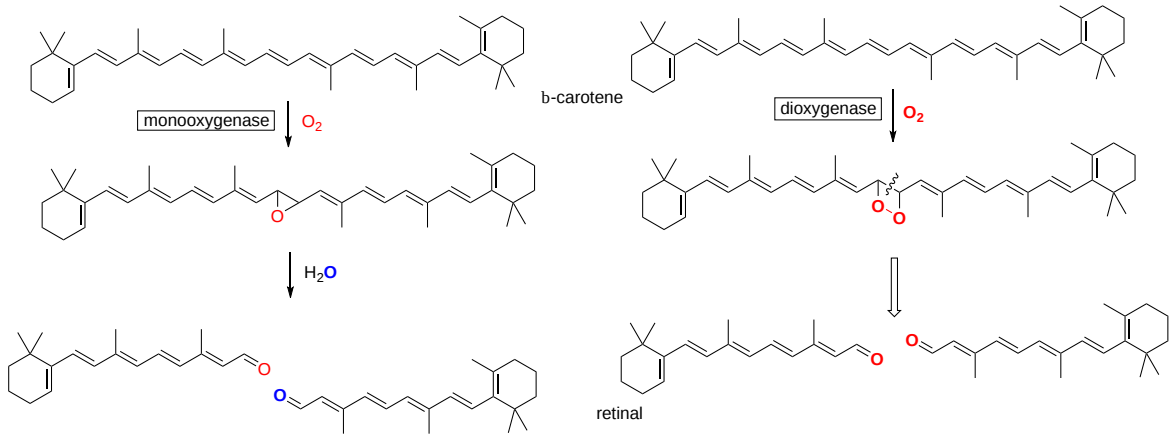


Figure 21.6.14: Introduction of oxygen atoms into retinal through either mono- and dioxygenase actions of β-carotene-15,15'-oxygenase-1 (BCO1).

The dioxygenase mechanism on the right best accounts for the incorporation of ^{18}O isotope of oxygen in dioxygen.

Figure 21.6.15 shows an interactive iCn3D model of the apocarotenoid cleavage oxygenase from Synechocystis, a Retinal-Forming Carotenoid Oxygenase (2BIW)

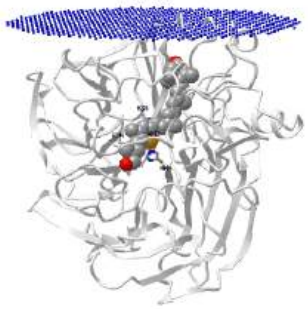


Figure 21.6.15: Retinal-Forming Carotenoid Oxygenase from Synechocystis (2BIW). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov...dttF958wFPAUP7

The enzyme has a bound carotene analog, (3R)-3-hydroxy-8'-apocarotenol. It contains an active site Fe^{2+} ion at the end of a hydrophobic tunnel. The ion is ligated by 4 histidine side chains. On binding, three of the trans C=C bonds convert to a distorted cis-trans-cis conformation. The middle trans bond is proximal to the Fe^{2+} which is ligated by dioxygen, the source of the incorporated oxygen atoms on cleavage.

Any unchanged β-carotene and newly formed retinol esters in the enterocytes are incorporated into **chylomicrons** and released into the lymphatic system and thence into the bloodstream, where some is taken up by peripheral tissues before most is absorbed by the liver. Some intact carotene and other carotenoids are transferred to lipoproteins (LDL and HDL) for transport in plasma, with assistance from specific binding and transport proteins, and for example, carotene can be absorbed at the placental barrier and transferred to the fetus for conversion to retinoids that are essential for development. Within the hepatocytes, retinol esters are hydrolyzed in the late endosomes with the release of free retinol into the cytosol, from which it can be released back into the circulation, converted to retinoids or transferred to hepatic stellate cells for storage in lipid droplets, the main body reservoir of vitamin A. In these specialized cells, retinol is esterified to form retinyl palmitate by the transfer of fatty acids from position *sn*-1 of phosphatidylcholine, mainly via the action of a membrane-bound lecithin:retinol acyltransferase (LRAT) in the endoplasmic reticulum. There are also lesser acyl-CoA dependent pathways, including an acyl CoA:retinol acyltransferase and even the enzyme diacylglycerol acyltransferase 1 (DGAT1); esterification is facilitated by binding to cellular retinol-binding protein type II (CRBP2).

Both retinol and retinoic acid are precursors of a number of metabolites (retinoids), which are required for specific purposes in tissues, by enzymatic modification of the functional groups and geometrical isomerization of the polyene chains. In the liver, activation of the retinol pathway involves first mobilization of the ester, followed by hydrolysis by retinol ester hydrolases, which includes carboxylesterase ES-10. Then, the reversible oxidation of retinol to retinal is carried out by one of several enzymes that include dehydrogenases and various cytochrome P450s, before some retinal is oxidized irreversibly to retinoic acid by enzymes with retinal dehydrogenase activity. On-demand conversion of retinol to retinoic acid occurs by the same mechanisms in other tissues, although for vision, retinol esters serve directly as the substrate for the formation of the visual chromophore 11-*cis*-retinal (see below). Retinyl-β-D-glucoside, retinyl-β-D-glucuronide, and retinoyl-β-D-glucuronide are naturally occurring and biologically active metabolites of vitamin A, which are found in fish and mammals. Indeed, the last has similar activity to all-*trans*-retinoic acid without any of the unwanted side effects in some circumstances.

Cleavage of β-carotene at double bonds other than that in the center or of a wider range of other carotenoids occurs by the action of a related enzyme β-carotene-9',10'-dioxygenase (β-carotene-oxygenase-2 or BCO2) in mitochondria, which leads to the formation of similar molecules, i.e. β-apocarotenals and β-apocarotenones of variable chain-length. While these may exert distinctive biological activities in their own right, there is evidence that they can also be metabolized to form retinal.

In the aqueous environment within cells, as well as in plasma, retinol, retinal and retinoic acid are bound to retinoid-binding proteins (RBP), which solubilize, protect and in effect detoxify them. These proteins also have a role in facilitating retinoid transport and metabolism; some are present only in certain tissues, and many are specific for particular retinoids and metabolic pathways. To prevent infiltration through the kidneys, retinol, and holo-RBP form an association in blood with a protein termed transthyretin (TTR), which also serves as a thyroid hormone carrier and is essential for secretion. Normally, vitamin A circulates in plasma as a retinol:RBP:TTR complex with a 1:1:1 molar ratio. Unesterified retinol is the main form of the vitamin that is exported from the liver upon demand, and it is transported in the blood in this bound form in VLDL, LDL, and HDL, with some directly from the diet in the chylomicrons and their remnants. Peripheral tissues have specific receptors to take up what they require, probably after hydrolysis of any esters to retinol by means of the enzyme lipoprotein lipase. Then, retinol dissociates from the protein as it forms a complex with a receptor (STRA6) at a target cell and diffuses through the plasma membrane, a process driven by retinol esterification.

The RBP-TTR complex does not bind to retinal and retinoic acid, although these do bind to RBP on its own, and most of the low levels of retinoic acid transported in the blood are bound to albumin. Local levels of retinoic acid are the result of an interplay between enzymes of

synthesis, binding, and catabolism. For example, within cells retinoic acid binding proteins (CRABP1 and CRABP2) bind to the newly synthesized retinoic acid, increase its rate of metabolism and protect cells from an excess.

In skin, 3,4-dehydroretinoids are synthesized from all-*trans* retinoids by the desaturase Cytochrome P450 27C1 with the assistance of cellular retinol-binding proteins (CRBPs). 3,4-Dehydroretinol, which is sometimes termed vitamin A_2, is shown in Figure 21.6.16.

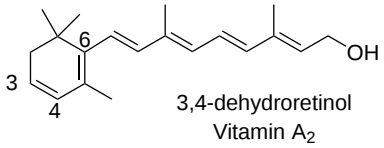


Figure 21.6.16: 3,4-dehydroretinol (Vitamin A_2)

Its derivative 3,4-dehydroretinal is used as a visual chromophore in many cold-blooded vertebrates including lampreys, fish, amphibians, and some reptiles (see below). Geranylgeranoic acid has structural similarities to retinoic acid and has been termed an acyclic retinoid, although it has no vitamin A activity. It is synthesized in animal tissues from mevalonate, and together with its 2,3-dihydro metabolite, has potent anticancer properties.

Retinol esters: A relatively small proportion of the cellular retinoids is located in membranes in tissues. Rather, retinol esters, mainly retinyl palmitate, are the main storage form of vitamin A, and they occur in many different organs, including adipose tissue and testes, but chiefly in stellate cells of the liver and pancreas. How the retinol is directed specifically to these cells and enters them prior to esterification is not known. Although hepatic stellate cells are much smaller and less abundant than hepatocytes (only 5 to 8% of all liver cells), they are characterized by cytoplasmic lipid droplets that contain 90-95% of the hepatic retinoids (and up to 80% of the body pool) in addition to other non-retinoid lipids; the lecithin:retinol acyltransferase is the only retinol ester synthase in this instance. In addition, specialized cells in the eye store retinoids essential for vision in the form of lipid droplets. When the supply of retinol in the diet is limited, hepatic stores of retinol esters are mobilized as retinol ester hydrolases are activated to maintain constant circulating retinol levels; hormone-sensitive lipase is the most important of these enzymes, although the adipose tissue triacylglycerol lipase and the lysosomal acid lipase are also involved.

Catabolism: All-*trans*-retinoic acid formation is irreversible, so its synthesis and degradation must be tightly regulated. As a first step in catabolism, the excess is cleared by conversion to more polar metabolites through oxidation by various enzymes of the cytochrome P450 family. Secondly, the water-soluble retinoic acid metabolites, including 4-hydroxy-, 4-oxo- and 18-hydroxy-retinoic acids, conjugate with glucuronic acid and then can be rapidly removed from circulation and eliminated from the body via the kidney.

21.6.2.2: RETINOIDS AND VISION

The structure of 11-cis-retinal, which we discussed in Chapter 11, is shown in Figure 21.6.17

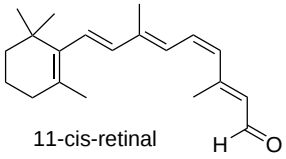


Figure 21.6.17: 11-cis-retinal

Retinoids are essential for vision, and there is now a good appreciation of how this works at the molecular level. In the eye, uptake of retinol from the circulation is mediated by the transmembrane cell-surface STRA6 receptor of the retinal pigment epithelium, a pigmented monolayer of cells located between the photoreceptors and choroid that nourishes retinal visual cells and catalyzes the release of retinol from retinol-binding proteins and transports it to the cytosol. The process by which light is converted to a signal recognized by the brain, sometimes termed the '**retinoid (visual) cycle**', requires a two-cell system beginning in the retinal pigment epithelium and continuing in photoreceptor cells, i.e. retinal rod and cone cells in the eye that contain membranous vesicles that serve as light receptors. Roughly half of the proteins in these vesicles consist of the protein conjugate, rhodopsin, which consists of a protein – opsin – with the retinoid **11-*cis*-retinal**. Each step in the visual process requires specific binding or transport proteins, and especially the interphotoreceptor retinoid-binding protein (IRBP).

All-*trans*-retinol is first converted to its ester by the enzyme lecithin:retinol acyltransferase as described above in the RPE, and the products coalesce into lipid droplets, i.e. dynamic organelles termed 'retinosomes'. The next step involves a dual-purpose enzyme (RPE65) in the endoplasmic reticulum, which cleaves the *O*-alkyl bond (not a conventional hydrolysis reaction) in the retinol ester and at the same time causes a change in the geometry of the double bond in position 11 of retinol from *trans* to *cis*. The 11-*cis*-retinol is then oxidized to 11-*cis*-retinal by 11-*cis*-retinal dehydrogenase (RDH5). The full cycle is shown in Figure \(\PageIndex{18}\) below.

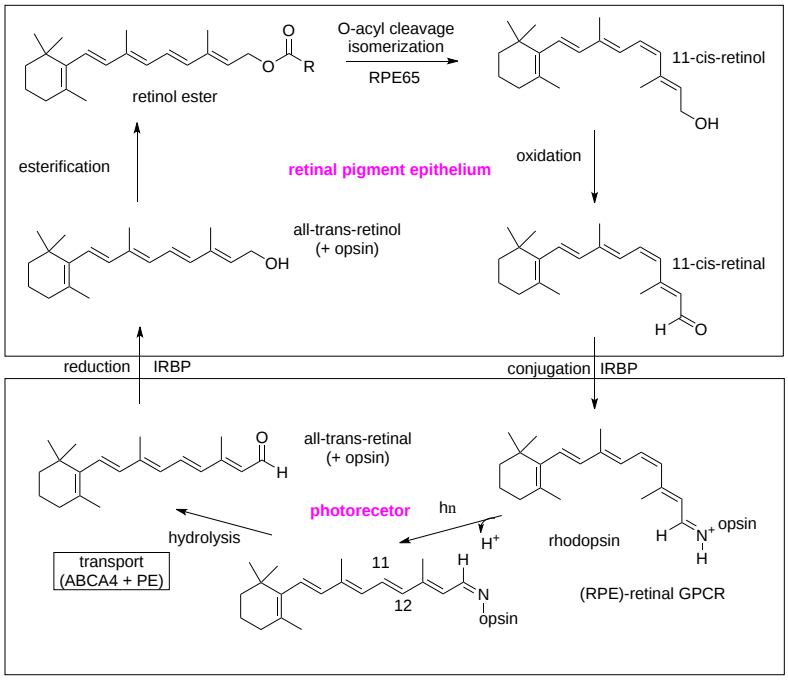


Figure 21.6.18: Retinoid visual cycle

The final part of the cycle occurs in the photoreceptor, where first the 11-*cis*-retinal is reacted with opsin to produce the protein conjugate rhodopsin in a protonated form. When rhodopsin is activated by light, the *cis*-double bond in the retinoid component is isomerized non-enzymatically by the energy of a photon to the 11-*trans* form with a change of conformation that in turn affects the permeability of the membrane and influences calcium transport. This results in further molecular changes that culminate in the release of opsin and all-*trans*-retinal, which is the trigger that sets off the nerve impulse so that the light is perceived by the brain.

A second mechanism for 11-*cis*-retinal formation that may function to ensure continuous visual responsiveness in bright light involves the (RPE)-retinal G protein-coupled receptor (RGR), which can function as a retinaldehyde photoisomerase. As the enzyme RPE65 functions optimally under low light conditions, it is believed that RGR prevents the saturation of photoreceptors under high light levels, and in this way facilitates vision in daylight. The isomerase, RPE65, and the photoisomerase, RGR, operate together to provide a sustained supply of the visual chromophore under different levels of illumination.

The all-*trans*-retinal is removed from the photoreceptor either by reduction to all-*trans*-retinol by all-*trans*-retinol dehydrogenase 8 expressed in the outer segments of photoreceptors or after transport by means of a specific transporter (ABCA4), which provides phosphatidylethanolamine (PE) for conversion to the Schiff-base adduct, i.e. *N*-retinylidene-phosphatidylethanolamine, as shown in Figure 21.6.19.

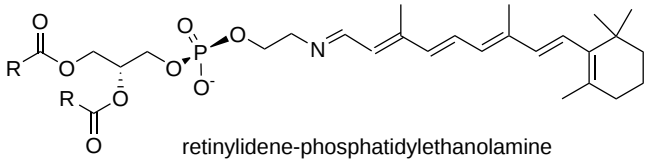


Figure 21.6.19: *N*-retinylidene-phosphatidylethanolamine

It flips from the lumen to the cytosolic leaflet of the disc membrane. This process prevents non-specific aldehyde activity with the effect of removing potentially toxic retinoid compounds from the photoreceptors. The adduct is a transient sink that dissociates so the retinal can be reduced back to all-*trans*-retinol by the cytoplasmic retinol dehydrogenase. All-trans-retinol exits the photoreceptor and enters the retinal pigment epithelium with the aid of binding to the retinoid-binding protein (IRBP) where it is converted back to a retinyl ester to complete the cycle and restore light sensitivity.

As a side-reaction, some troublesome bis-retinoid adducts of PE (and further byproducts) may be produced by non-enzymatic mechanisms, and these can accumulate with age to affect vision. **Lower organisms:** Bacteriorhodopsin is the best studied of a family of opsins, found in archaea, eubacteria, fungi, and algae. It is a protein with seven transmembrane helices that acts as an opto-electrical transducer or light-gated active ion pump to capture photon energy via its covalently bound chromophore, all-*trans*-retinal, converting it to 13-*cis*-retinal, and moves protons against their electrochemical gradient from the cytoplasm to the extracellular space. In Archaea, it is known as the "purple membrane" and can occupy a high proportion of the surface area of the organism.

21.6.2.3: OTHER FUNCTIONS OF RETINOIDS IN HEALTH AND DISEASE

In addition to their function in vision, it is now realized that retinoids have essential roles in growth and development, reproduction and resistance to infection. They are particularly important for the function of epithelial cells in the digestive tract, lungs, nervous system, immune system, skin, and bone at all stages of life. They are required for the regeneration of damaged tissues, including the heart, and they appear to have some potential as chemo-preventive agents for cancer and for the treatment of skin diseases such as acne. Under pathological conditions, stellate cells lose their retinoid content and transform into fibroblast-like cells, contributing to the fibrogenic response. Cirrhosis of the liver is accompanied by a massive loss of retinoids, but it is not clear whether this is a cause or a symptom, and there appears to be confusion as to when supplementation may be helpful in this and other diseases of the liver. Like retinol and retinoic acid, the metabolite 9-*cis*-retinoic acid also has valuable pharmaceutical properties.

With such a large number of double bonds in conjugation, it is not surprising that carotenoids in general, and retinoids in particular are efficient quenchers of singlet oxygen and scavengers of other reactive oxygen species. However, any direct antioxidant properties are not believed to be important in terms of general health *in vivo*, and it is not clear how relevant the physical properties of retinoids are to specific biochemical processes in comparison to their effects on signaling and gene transcription. There is a caveat that retinoids may stimulate some antioxidant genes and so have an indirect antioxidant function. In fact, nutritional studies with dietary supplements of carotenoids have sometimes suggested pro-oxidant activity. One explanation for detrimental effects may be that regeneration of the parent carotenoid or retinoid from the corresponding radical cation may be limited when concentrations of reductants such as ascorbic acid are low.

Many of the retinol metabolites function as ligands to activate specific transcription factors for particular receptors in the nucleus of the cell, and thus they control the expression of a large number of genes (>500), including those essential to the maintenance of normal cell proliferation and differentiation, embryogenesis, for a healthy immune system, and for male and female reproduction. In the innate immune system, vitamin A is required for the differentiation of cells such as macrophages, neutrophils and natural killer cells, while all-*trans*-retinoic acid is involved in differentiating the precursors of dendritic cells. **Retinoic acid** and its 9-*cis*-isomer are especially important in this context, and they are often considered the most important retinoids in terms of function other than in the eye. This structure of the cis-isomer is shown in Figure 21.6.20.

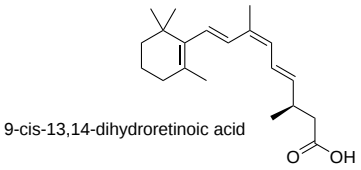


Figure 21.6.20: Cis isomer of retinoic acid

In essence, retinoic acid moves to the nucleus with the aid of small intracellular lipid-binding proteins (CRABP2 and FABP5), which channel it to specific nuclear receptors, the **retinoic acid receptors (RAR)** of which there are three, RAR-α, β and γ. These are ligand-dependent regulators of transcription and they function *in vivo* as heterodimers with **retinoid X receptors (RXR)** to process the retinoic acid signal by acting through polymorphic retinoic acid response elements (RAREs) within the promoter regions of responsive genes. Similarly, 9-*cis*-retinoic acid and 9-*cis*-13,14-dihydroretinoic acid are high-affinity ligands for RXR in mice. Together with retinoic acid, these are also ligands for the farnesoid X receptor (FXR), which forms a heterodimer with RXR. The latter receptor complex is involved primarily in **bile acid** homeostasis, and conversely, there are suggestions that bile acids may have regulatory effects on vitamin A homeostasis.

In addition, other nuclear receptors, such as the peroxisome proliferator-activated receptor PPARγ forms a heterodimer with the retinoid X receptor and is activated by retinoic acid to recruit cofactors. This complex in turn binds to the peroxisome proliferator response element (PPRE) gene promoter, leading to regulation mainly of those genes involved in lipid and glucose metabolism, including some involved in inflammation and cancer. To add to the complexity, retinoic acid has extra-nuclear, non-transcriptional effects, such as the activation of protein kinases and other signaling pathways.

It has also become evident that many of the functions of retinoids are mediated via the action of specific binding proteins (as discussed briefly above), which control their metabolism *in vivo* by reducing the effective or free retinoid concentrations, by protecting them from unwanted chemical attack, and by presenting them to enzyme systems in an appropriate conformation. With some tissues, retinol-bound RBP in the blood is recognized by the membrane protein STRA6, which transports retinol into cells where it binds to an intracellular retinol acceptor, cellular retinol-binding protein 1 (CRBP1), and is then able to activate a signaling cascade that targets specific genes. In addition, a specific retinol-binding protein secreted by adipose tissue (RPB4) is involved in the development of insulin resistance and type 2 diabetes, possibly by affecting glucose utilization by muscle tissue, with obvious application to controlling obesity. In the eye, the activity of retinoic acid during development is controlled by binding to apolipoprotein A1.

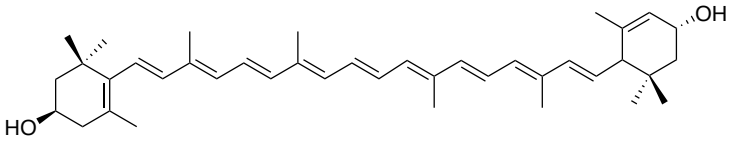
All-*trans*-retinoic acid has been shown to be effective against many different types of human cancers, especially in model systems but also in some clinical trials, because of its specific effects on cell proliferation, differentiation, and apoptosis (where its relatively low toxicity at normal tissue levels is a virtue). For example, it induces complete remission in most cases of acute promyelocytic leukemia when administered in combination with other chemotherapy techniques. Similarly, 13-*cis*-retinoic acid has been used successfully in the treatment

of children with high-risk neuroblastoma to reduce the risk of recurrence and increase long-term survival rates. However, the efficacy of similar treatments against other types of acute myeloid leukemia and solid tumors appears to be poor. It is hoped that current efforts to obtain a better understanding of the mechanism of the anti-cancer activities will lead to improved treatments. Synthetic analogs of retinoic acid, termed **rexinoids**, which activate retinoic X receptors, also hold promise as anti-cancer agents.

Vitamin A deficiency in children and adult patients is usually accompanied by impairment of the immune system, leading to a greater susceptibility to infection and an increased mortality rate, often with growth retardation and congenital malformations. However, vitamin A deficiency in malnourished children is the major reason for childhood mortality in the underdeveloped world, causing over 650,000 early childhood deaths annually and pediatric blindness. This is doubly tragic in that it is so easily prevented. In adults, vitamin A deprivation affects the reproductive system, inhibiting spermatogenesis in males and ovulation in females. Unfortunately, it is not always easy to distinguish between the effects of vitamin A deficiency and primary defects of retinoid signaling.

21.6.2.4: FUNCTIONS OF XANTHOPHYLLS AND OTHER CAROTENOIDS IN HUMANS

Xanthophylls are plant C_{40} tetraterpenes that differ from the carotenoids in having oxygen atoms in the ring structures (hydroxyl, oxo, or epoxyl). Lutein, zeaxanthin, and *meso*-zeaxanthin from dietary sources, such as green leafy vegetables and yellow and orange fruits and vegetables, are found specifically in the macula of the eye in humans and other primates, i.e. the functional center of the retina in a small central pit known as the *macula lutea*, where they enhance visual acuity and protect the eye from high-intensity, short-wavelength visible light. They are powerful antioxidants in a region vulnerable to light-induced oxidative stress. Binding proteins specific for lutein- and zeaxanthin mediate the highly selective uptake of these carotenoids into the retina, but *meso*-zeaxanthin is mainly a metabolite of dietary lutein. Macular xanthophylls decrease the risk of age-related macular degeneration. In the brain, they may stimulate and maintain cognitive function in the elderly, and assist with brain development in infants. Hydroxylated xanthophylls such as lutein, shown in Figure 21.6.21, occur both in the free form and esterified to fatty acids; the latter are hydrolyzed in the intestines when consumed by animals.



lutein

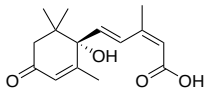
Figure 21.6.21: Leutin

Many other carotenoids are absorbed from the diet, and are subject to oxidative cleavage or other catabolic processes, partly in the intestines and partly in other tissues after transport in the lipoproteins. Some carotenoids remain intact and are believed to act as antioxidants, and some may have specific anti-inflammatory actions. For example, carotenoids accumulate in the skin of mammals, where they may have an antioxidant and photo-protective role as well as effects on the moisture content, texture, and elasticity. Lycopene may have protective effects against atherogenesis, coronary heart disease, and prostate cancer.

21.6.2.5: FUNCTIONS OF CAROTENOIDS IN PLANTS

As carotenoids have a polyene chain of 9 to 11 double bonds in conjugation, they are able to absorb light in the gap of chlorophyll absorption, and so function as additional light-harvesting pigments in plants. Their distinctive arrangement of electronic levels gives them the capacity to transfer excitation energy from the carotenoid excited state to chlorophyll in the light-harvesting complex (photosystem II). Energy can also be transferred back from chlorophyll to carotenoids as a photoprotection mechanism. During photosynthesis, damaging species are produced by both light and oxygen with reactive oxygen species (ROS) of special concern. The energy is transferred from chlorophyll to the polyene tail of the carotenoid where electrons are moved between the carotenoid bonds until the most balanced or lowest energy state (state) is reached. While there is therefore appreciable potential for carotenoids to act as antioxidants in plants, it is uncertain how important this is from a practical functional standpoint. The length of the polyene tail of carotenoids determines which wavelengths of light will be absorbed by the plant, and those not absorbed are reflected and so determine coloration. F

Carotenoids are precursors for two plant hormones and a diverse set of apocarotenoids. For example, **abscisic acid**, shown in Figure 21.6.21, is a C_{15} isoprenoid plant hormone, which is synthesized in plastids from the C_{40} carotenoid zeaxanthin.



abscisic acid

Figure 21.6.21:

A series of enzyme-catalyzed epoxidations and isomerizations is involved followed by cleavage of the intermediate product by a dioxygenation reaction and further oxidations to yield eventually abscisic acid. Functioning via signaling cascades, abscisic acid regulates innumerable biological effects in plants, especially in relation to developmental processes that include plant growth, seed and bud dormancy, embryo maturation and germination, cell division and elongation, floral growth and the control of stomatal closure. It is critical for the responses to environmental stresses that include drought, cold and heat stress, salinity, and tolerance of heavy metal ions. Similarly,

strigolactones are C_{15} oxidation products of carotenoids that are involved in the regulation of symbiosis between plants and arbuscular mycorrhizal fungi and in interactions with plant parasites.

21.6.3: ISOPRENOIDS: 3. OTHER MEMBRANE-ASSOCIATED ISOPRENOIDS

Terpenes (isoprenoids) are one of the most varied and abundant natural products produced by animals, plants, and bacteria. They are generally defined on the basis of their biosynthetic derivation from isoprene units (C_5H_8), with 55,000 different types characterized to date according to a recent estimate. By most definitions, all isoprenoids should be classified as 'lipids', from simple monoterpenes such as geraniol, which is derived from two prenol units, to complex polymers such as natural rubber. Only those isoprenoids that have a **functional** role in cellular membranes will be discussed here. These include plastoquinone, ubiquinone (coenzyme Q), phylloquinone and menaquinone (vitamin K), dolichol and polyprenols, undecaprenyl phosphate and lipid II, and farnesyl pyrophosphate, together with some key biosynthetic precursors. The nature and function of **tocopherols and tocotrienols (vitamin E)** and **retinoids (vitamin A)** are relevant here, but have been discussed previously. Of course, sterols are also isoprenoids.

There are two basic mechanisms for the biosynthesis of the isoprene units that are the precursors for the biosynthesis of isoprenoids, i.e., isopentenyl pyrophosphate and dimethylallyl pyrophosphate. These are the mevalonate pathway, which is located in the cytosol of the cell, and the non-mevalonate pathway, found mainly in the plastids of plants. These have been discussed previously.

21.6.3.1: PHYTOL

Phytol or (2*E*,7*R*,11*R*)-3,7,11,15-tetramethyl-2-hexadecen-1-ol, i.e., with 20 carbons in a 16-carbon chain and one double bond, is an acyclic diterpene alcohol, which is synthesized in large amounts in plants as an essential component of chlorophyll, the most important photosynthetic pigment in plants and algae. Geranylgeranyl-diphosphate synthesized in chloroplasts via the 4-methylerythritol-5-phosphate (non-mevalonate) pathway is the primary precursor of phytol following reduction of three double bonds by geranylgeranyl reductase, and this can occur before or after attachment to chlorophyll, depending upon species. Chlorophyll dephytylase (CLD1) is the enzyme in plants responsible for chlorophyll hydrolysis and the release of phytol. The reactions are shown in Figure 21.6.22.

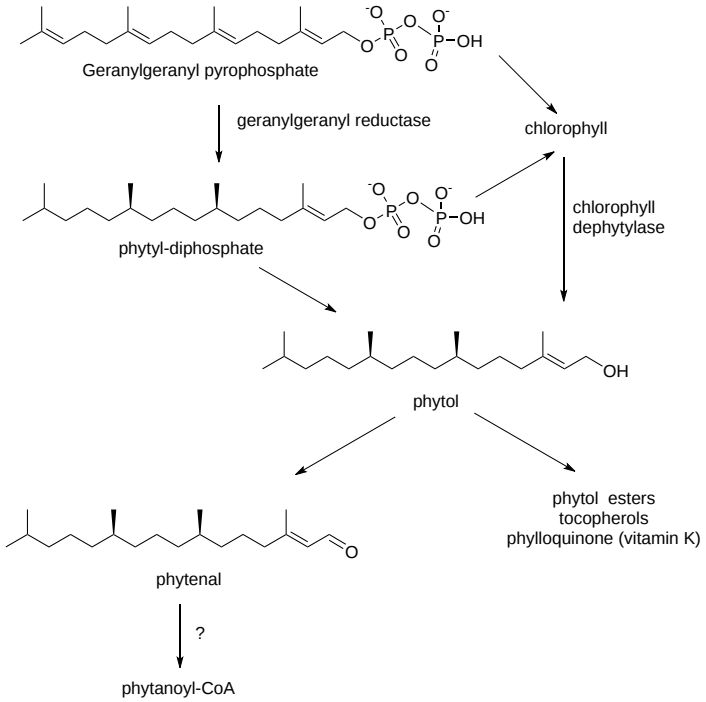


Figure 21.6.22 Phytol synthesis and metabolism in plants

Little free phytol is present in plant tissues, although some phytol esters of fatty acids may occur, especially when plants are stressed during nitrogen deprivation or in senescence, when chlorophyll is degraded, fatty acids are released from glycerolipids and a phytol ester synthase is induced as part of a detoxification and recycling process. Bell peppers and rocket salad are especially rich sources under normal conditions. In addition, phytol as its diphosphate is utilized for the synthesis of **tocopherols** (vitamin E) and phylloquinol (vitamin K - see below), and the precursor geranylgeraniol and its fatty acid ester occur in small amounts in some plant species. It is the biosynthetic precursor of tocotrienols and the highly unsaturated carotenoids (and hence of **retinoids**). Phytenal has been isolated as an intermediate in the catabolism of phytol in plants, but further steps are uncertain although phytanoyl-CoA has been detected in stressed plants. As phytenal is highly reactive and potentially toxic via its interaction with proteins, its accumulation must be kept at a low level by competing pathways.

In ruminant animals, chlorophyll is hydrolyzed by rumen microorganisms with the release of free phytol. This does not occur in humans, but some phytol may be ingested with plant foods either in free form or as phytol esters and can be absorbed from the intestines. Within animal tissues, phytol is oxidized to **phytanic acid**. Phytol and/or its metabolites have been reported to activate the transcription factors PPARα and retinoid X receptor. In mice, oral phytol induces a substantial proliferation of peroxisomes in many organs.

21.6.3.2: PLASTOQUINONE

A molecule that is related to the tocopherols, **plastoquinone**, is found in cyanobacteria and plant chloroplasts, and it is produced in plants by analogous biosynthetic pathways to those of tocopherols in the inner chloroplast envelope with solanesol diphosphate as the biosynthetic precursor of the side chain; there appears to be a somewhat different mechanism in cyanobacteria. The molecule is sometimes designated - 'plastoquinone-n' (or PQ-n), where 'n' is the number of isoprene units, which can vary from 6 to 9. It's structure is shown in Figure 21.6.23.

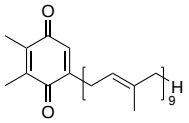


Figure 21.6.23: Plastoquinone

Plastoquinone has a key role in photosynthesis, by providing an electronic connection between photosystems I and II, generating an electrochemical proton gradient across the thylakoid membrane. This provides energy for the synthesis of adenosine triphosphate (ATP). The reduced dihydroplastoquinone (plastoquinol) that results in the transfers further electrons to the photosynthesis enzymes before being re-oxidized by a specific cytochrome complex; the redox state of the plastoquinone pool regulates the expression of many of the genes encoding photosystem proteins. X-Ray crystallography studies of photosystem II from cyanobacteria show two molecules of plastoquinone forming two membrane-spanning branches. In addition, plastoquinone has antioxidant activity comparable to that of the tocopherols, protecting especially against excess light energy and photooxidative damage. Similarly, in thylakoid membranes, plastoquinol is able to scavenge superoxide with the production of H_2O_2. Plastoquinone is a cofactor participating in desaturation of phytoene in carotenoid biosynthesis, and the biosynthetic precursor of plastochromanols. With these many different functions, plastoquinone connects photosynthesis in plants with metabolism, light acclimation, and stress tolerance.

Plastoquinone-9, together with phylloquinone, tocopherol, and plastochromanol-8, is stored in plastoglobuli, lipoprotein-like micro-compartments, which enable exchange with the thylakoid membrane and are also involved in chlorophyll catabolism and recycling. It has been suggested that the redox state of the plastoquinone pool is the main redox sensor in chloroplasts that initiates many physiological responses to changes in the environment and in particular to those related to light intensity by regulating the expression of chloroplast genes.

21.6.3.3: UBIQUINONE (COENZYME Q)

The **ubiquinones**, which are also known as coenzyme Q (CoQ) or mitoquinones, have obvious biosynthetic and functional relationships to plastoquinone and they are found in all the domains of life (hence the name). They have a 2,3-dimethoxy-5-methylbenzoquinone nucleus and a side chain of six to ten isoprenoid units; the human form illustrated below has ten such units (coenzyme Q_{10}), i.e., it is 2,3-dimethoxy-5-methyl-6-decaprenyl-1,4-benzoquinone, while that of the rat has nine, *Escherichia coli* has eight and *Saccharomyces cerevisiae* has six. In plants, ubiquinones tend to have nine or ten isoprenoid units. Forms with a second chromanol ring, resembling the structures of tocopherols, are also produced (ubichromanols), but not in animal tissues. They are generated on an industrial scale for pharmaceutical purposes by yeast fermentation. Because of their hydrophobic properties, ubiquinones are located entirely in membrane bilayers in most eukaryote organelles, probably at the mid-plane.

Ubiquinones are synthesized *de novo* in mitochondria in most cells in animal, plant, and bacterial tissues by a complex sequence of reactions from the essential amino acid phenylalanine and then tyrosine to generate *p*-hydroxybenzoic acid, which is the key precursor that is condensed with the polyprenyl unit (from the cholesterol synthesis pathway) via a specific transferase; this is followed by decarboxylation, hydroxylation, and methylation steps, depending on the specific organism, although some of the required enzymes have yet to be fully characterized. In *Escherichia coli*, biosynthesis does not occur in a membrane environment as had been thought. Rather, the seven proteins that catalyze the last six reactions of the biosynthetic pathway, following the attachment of the isoprenoid tail, form a stable complex or metabolon in the cytoplasm so enabling modification of the hydrophobic substrates in a hydrophilic environment.

In mitochondria, coenzyme Q is present both as the oxidized (ubiquinone) and reduced (ubiquinol) forms as shown in Figure 21.6.24.

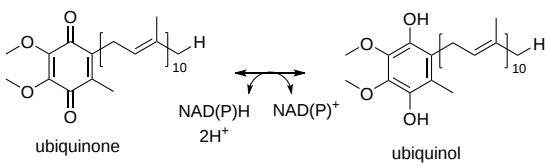


Figure 21.6.24: Ubiquinone-ubiquinol interconversion

Ubiquinones are essential components of the respiratory electron transport system, possibly as part of supramolecular complexes, taking part in the oxidation of succinate or NADH via the cytochrome system to generate the protonmotive force used by the mitochondrial ATPase to synthesize ATP. In this process, coenzyme Q transfers electrons from the various primary donors, including complex I, complex II, and the oxidation of fatty acids and branched-chain amino acids, to the oxidase system (complex III), while simultaneously transferring protons to the outside of the mitochondrial membrane with the result of a proton gradient across the membrane. As a consequence, it is reduced to ubiquinol. Thus, it is an essential component of the cycle that generates the proton motive force driving ATP production via oxidative phosphorylation. In yeast, one coenzyme Q binding protein (COQ10), and in humans two related proteins (COQ10A and COQ10B) may serve as chaperones or transporters during this process. Mitochondrial coenzyme Q is also implicated in the production of reactive oxygen species by a mechanism involving the formation of superoxide from ubisemiquinone radicals, and in this way is responsible for causing some of the oxidative damage behind many degenerative diseases. In this action, it is a pro-oxidant. It is most abundant in organs with a high metabolic rate such as the heart, kidneys, and liver.

In complete contrast in its reduced form (ubiquinol) in non-mitochondrial cellular membranes and plasma lipoproteins, it acts as an endogenous antioxidant, the only lipid-soluble antioxidant to be synthesized endogenously. It inhibits lipid peroxidation in biological membranes and serum low-density lipoproteins, and it may also protect membrane proteins and DNA against oxidative damage. The ferroptosis suppressor protein 1 (FSP1) replenishes ubiquinol, and this acts protectively by combating the lipid peroxidation that drives **ferroptosis**. The mechanism involves the recruitment of FSP1 to the plasma membrane following myristoylation, where this functions as an oxidoreductase that reduces ubiquinone to ubiquinol, which acts as a lipophilic radical-trapping antioxidant that halts the propagation of lipid peroxides. In this manner, it regulates cellular redox status and cytosolic oxidative stress, and thereby is a controlling factor in apoptosis. Other NAD(P)H dehydrogenases with CoQ reductase activity include cytochrome b5 reductase and NQo1 (NAD(P)H:quinone oxidoreductase).

Although ubiquinone has only about one-tenth of the antioxidant activity of **vitamin E** (α-tocopherol), it is able to stimulate the effects of the latter by regenerating it from its oxidized form back to its active fully reduced state (similarly with vitamin C). However, ubiquinones and tocopherols appear to exhibit both cooperative and competitive effects under different conditions. Similarly, in bacteria and other prokaryotes, ubiquinones participate a large number of redox reactions, notably in the respiratory electron transport system but also in other enzyme reactions that require electron donation, including the formation of disulfide bonds.

Coenzyme Q has many other functions that are not related directly to its antioxidant function. Some coenzyme Q is used in mitochondria by enzymes that link the mitochondrial respiratory chain to other metabolic pathways, including fatty acid β-oxidation as an electron acceptor, nucleotide biosynthesis *de novo*, amino acid oxidation (glycine, proline, glyoxylate, and arginine), and detoxification of sulfide. Via these activities, coenzyme Q may modulate metabolic pathways located outside the mitochondria indirectly. There are also suggestions that coenzyme Q may be involved in redox control of cell signaling and gene expression, and in particular to repress the expression of inflammatory genes. In relation to its anti-inflammatory properties, clinical studies suggest that supplementation with coenzyme Q_{10} reduces the levels of the inflammatory mediators C-reactive protein, interleukin-6, and tumor necrosis factor alpha (TNFα) to a significant degree. In addition, it is a regulator of mitochondrial permeability, and in relation to pyrimidine nucleotide biosynthesis, it is required for DNA replication and repair.

Dietary ubiquinone, i.e., that in food or dietary supplements, is absorbed by enterocytes via a process of “passive facilitated diffusion”, probably requiring a carrier molecule, before incorporation into chylomicrons for transport to the liver. This eventually leads to elevated levels of ubiquinol in blood, especially in the LDL and VLDL lipoproteins, presumably because of the reduction of the oxidized form in the lymphatic system. In consequence, there is reported to be enhanced protection against lipid peroxidation with beneficial effects on health, especially in relation to cardiac function, sperm motility, and neurodegenerative diseases. For example, CoQ levels in both plasma and the heart correlate with heart failure in patients, and clinical trials of dietary supplementation have shown promising results. This may be of particular importance in the elderly or in patients on statins, when endogenous synthesis declines. A CoQ_{10} deficiency syndrome is associated with inherited pathological diseases, defined by a decrease of the CoQ_{10} content in muscle and/or cultured skin fibroblasts. Early clinical trials with CoQ10 and a synthetic analog, idebenone, against various neurodegenerative diseases, including Alzheimer's disease, Parkinson's disease, Huntington's disease, and others, are encouraging.

21.6.3.4: PHYLLOQUINONE AND MENAQUINONES (VITAMIN K)

Phylloquinone or 2-methyl-3-phytyl-1,4-naphthoquinone is synthesized in the inner chloroplast envelope of cyanobacteria, algae, and higher plants by a mechanism analogous to that of the **tocopherols**, i.e., from chorismate in the shikimate pathway with a prenyl side chain derived from phytyldiphosphate. In this membrane, it is a key component of the photosystem I complex where it receives an electron from the chlorophyll a acceptor molecule and then donates an electron to the membrane-associated iron-sulfur protein acceptor cluster in the complex. In an obvious parallel to the plastoquinones (above), two molecules of phylloquinone form two membrane-spanning branches, as demonstrated by X-ray crystallography studies of photosystem I from cyanobacteria. Plastoglobules associated with the thylakoid membrane are believed to function as a reservoir for excess phylloquinone, and may also function in its metabolism. Edward A. Doisy and Henrik Dam received the Nobel Prize in Physiology or Medicine in 1943 for their discovery of vitamin K and its chemical structure. The structure of vitamin K is shown in Figure 21.6.25.

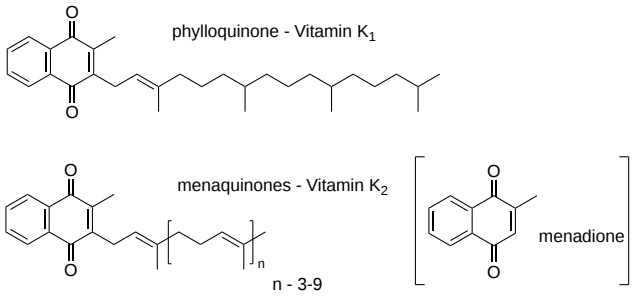


Figure 21.6.25: Vitamin K

The menaquinones are related bacterial products, which function in the respiratory and photosynthetic electron transport chains of bacteria. They have a variable number (4 to 10) of isoprenoid units in the tail, and they are sometimes designated ‘MK-4’ to ‘MK-10’. In contrast to phylloquinone, these are usually highly unsaturated. In some species, there are methyl or other groups attached to the naphthoquinone moiety. Remarkably high concentrations of menaquinones are present in membranes of some extremophiles such as the haloarchaea, where it has been suggested that they act as ion permeability barriers and as a powerful shield against oxidative stress in addition to their functions as electron and proton transporters.

Phylloquinone is an essential component of the diet of animals and has been termed '**vitamin K_1**'. It must be supplied by green plant tissues, where it occurs in the range 400-700 µg/100 g, or seed oils. The menaquinones, the main source of which in the human diet is cheese and yogurt, also have vitamin K activity and are termed '**vitamin K_2**'. They account for about 10-25% of the vitamin K content of the Western diet. A synthetic saturated form of this, which is used in animal feeds, is known as 'vitamin K_3 or menadione', though strictly speaking it is not a vitamin but a pro-vitamin in that it can be converted to the menaquinone MK-4 in animal tissues by addition of a phytyl unit; it is too toxic for human nutrition. Vitamin K forms are absorbed from the intestines and transported in plasma in the form of lipoproteins in a similar manner to the other fat-soluble vitamins. Different tissues have differ storage capacities and presumably requirements for the various forms of vitamin K. As a high proportion is excreted, there appears to be a requirement for a constant intake. Vitamin K_1 is taken up rapidly by the liver, but vitamin K_2 remains in the plasma for much longer and maybe the main source of the vitamin in peripheral tissues. In addition, it has been established that some dietary phylloquinone is converted to menaquinone-4 in animals, although the quantitative significance has still to be established; the mechanism involves conversion to menadione in the intestines followed by transport to tissues where a geranylgeranyl side-chain is attached by a specific prenyl transferase.

The primary role of vitamin K in animal tissues is to act as a cofactor specific to the vitamin K-dependent enzyme γ-glutamyl carboxylase in the endoplasmic reticulum in the liver mainly. Its function is the post-translational carboxylation of glutamate residues to form γ-carboxyglutamic acid in proteins, such as prothrombin. In this way, prothrombin and three related proteins are activated to promote blood clotting. The γ-carboxyglutamic acid residues are located at the binding site for Ca^{2+}, and are vital for the activity of the enzyme. Phylloquinone must first be converted to the reduced form, phylloquinol, which is the actual cofactor for the enzyme; molecular oxygen and carbon dioxide are both required also. Phylloquinol donates hydrogen to the glutamic acid residue and is oxidized in the process to 2,3-epoxyphylloquinone. A further enzyme, vitamin K epoxide reductase, regenerates phylloquinone by reduction of the epoxide in a dithiol-dependent reaction so that this can be re-utilized many times ("the vitamin K cycle", shown below in Figure 21.6.26).

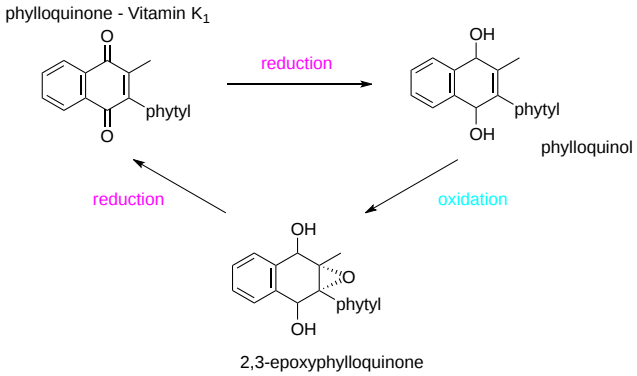


Figure 21.6.26: The Vitamin K cycle

Menaquinones undergo the same cycle of reaction. By interfering with the last step in the metabolic cycle, warfarin, the rodenticide, prevents blood clotting. In the same way, a deficiency in vitamin K results in the inhibition of blood clotting and can lead to brain hemorrhaging in malnourished newborn infants, though this is not seen in adult humans, presumably because intestinal bacteria produce sufficient for our needs. Vitamin K-dependent proteins are also known to have important functions in the central and peripheral nervous systems, and vitamin K influences sphingolipid biosynthesis in the brain.

Unsaturated isoprene units rather than phytol are used for the biosynthesis of menaquinones, and they differ from phylloquinone with respect to their chemical structure and pharmacokinetics. It is now apparent that vitamin K_2 (MK-4 especially) has a number of different actions, some with specificities for particular tissues. For example, osteocalcin is a γ-carboxyglutamic acid-containing protein, which forms a strong complex with the mineral hydroxyapatite (calcium phosphate) of bone; it must be carboxylated to function properly and vitamin K_2 appears to be of particular importance in this instance. Vitamin K_2 also regulates bone remodeling by osteoclasts to remove old or damaged bone and its replacement by new bone. In addition, vitamin K_2 is involved in vascular calcification, cell growth, and apoptosis. Side effects of the use of anticoagulants that bind to vitamin K can be osteoporosis and increased risk of vascular calcification. Although careful control of the dosage is necessary, vitamin K_2 may be a useful adjunct for the treatment of osteoporosis, and it may reduce morbidity and mortality in cardiovascular health by reducing vascular calcification.

Excess vitamin K_1 and the menaquinones are catabolized in the liver by a common degradative pathway in which the isoprenoid side chain is shortened to yield carboxylic acid aglycones such as menadiol, which can be excreted in bile and urine as glucuronides or sulfates.

21.6.3.5: DOLICHOLS AND POLYPRENOLS

Polyisoprenoid alcohols, such as dolichols, are ubiquitous if minor components, relative to the glycerolipids, of membranes of most living organisms from bacteria to mammals. They are hydrophobic linear polymers, consisting of up to twenty isoprene residues or a hundred carbon atoms (or many more in plants especially), linked head-to-tail, with a hydroxy group at one end (α-residue) and a hydrogen atom at the other (ω-end). In **dolichols** (or dihydropolyprenols), the double bond in the α-residue is hydrogenated, and this distinguishes them from the **polyprenols** with a double bond in the α-residue. Their structures are shown in Figure 21.6.27.

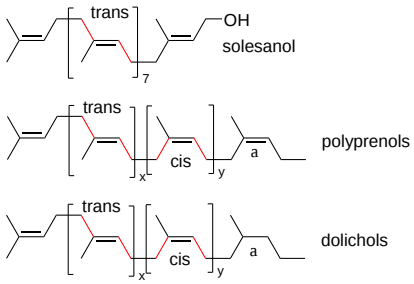


Figure 21.6.27: Polyprenols and dolichols

Polyisoprenoid alcohols are further differentiated by the geometrical configuration of the double bonds into three subgroups, i.e., di-*trans*-poly-*cis*, tri-*trans*-poly-*cis*, and all-*trans*. For many years, it was assumed that polyprenols were only present in bacteria and plants, especially photosynthetic tissues, while dolichols were found in mammals or yeasts, but it is now known that dolichols can also occur at low levels in bacteria and plants, while polyprenols have been detected in animal cells. **Solanesol** is a related and distinctive plant product with *trans* double bonds only and is a precursor of plastoquinone.

Within a given species, components of one chain length may predominate, but other homologs are usually present. The chain length of the main polyisoprenoid alcohols varies from 11 isoprene units in eubacteria, to 16 or 17 in Drosophila, 15 and 16 in yeasts, 19 in hamsters, and 20 in pigs and humans. In plants, the range is from 8 to 22 units, but some species of plant have an additional class of polyprenols with up to 40 units. In tissues, polyisoprenoid alcohols can be present in the free form, esterified with acetate or fatty acids, phosphorylated or monoglycosylated phosphorylated (various forms), depending on species and tissue. Polyisoprenoid alcohols *per se* do not form bilayers in aqueous solution, but rather a type of lamellar structure. However, they are found in most membranes, especially the plasma membrane of liver cells and the chloroplasts of plants.

Dolichoic acids, i.e., related molecules with a terminal carboxyl group and containing 14–20 isoprene units, have been isolated from the substantia nigra of the human brain. However, they were barely detectable in the pig brain.

Biosynthesis of the basic building block of dolichols, e.g. isopentenyl diphosphate, follows either the mevalonate pathway or a more recently described methylerythritol phosphate pathway discussed in relation to **cholesterol biosynthesis** previously. Farnesyl pyrophosphate is the primary precursor in the biosynthesis of polyprenols and is the branch point in sterol/isoprene biosynthesis, depending on the nature of the organism (see a further note below). Subsequent formation of the linear prenyl chain is accomplished by *cis*-prenyl transferases that catalyze the condensation of isopentenyl pyrophosphate and farnesyl pyrophosphate and then the growing the allylic prenyl diphosphate chain. The end products are polyprenyl pyrophosphates, which are dephosphorylated first to polyprenol phosphate and thence to the free alcohol. Finally, a specific reductase has been identified from human tissues that catalyzes the reduction of the double bond in position 2 to produce dolichols. There is a family of *cis*-prenyl transferases, present in both eukaryotes and bacteria, that in addition to the synthesis of dolichols can catalyze the formation of isoprenoid carbon skeletons from neryl pyrophosphate (C_{10}) to natural rubber ($C_{>10,000}$), including the polyisoprenoid phosphates involved in protein glycosylation as discussed in the next section. The reactions in the synthesis of dolichol are shown in Figure 21.6.28.

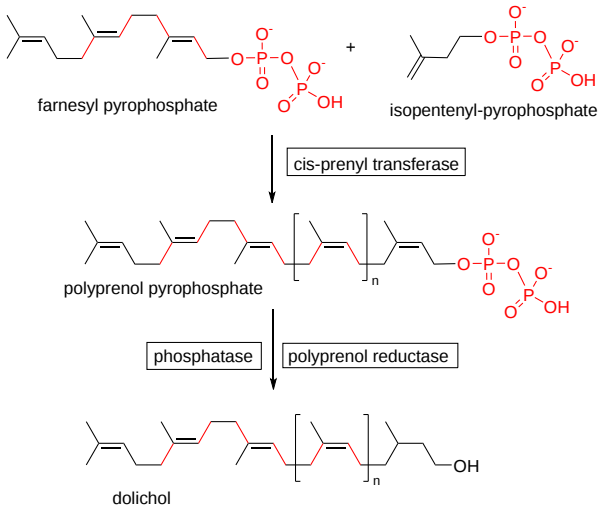


Figure 21.6.28: Synthesis of dolichol

Although polyprenols and dolichols were first considered to be simply secondary metabolites, they are now known to have important biological functions. Glycosylation of asparagine residues is the main protein modification in all three domains of life, and phosphorylated polyisoprenoids, including dolichols, are essential to this process (next section). There is also a suggestion that free dolichol may have a beneficial antioxidant function in cell membranes.

21.6.3.6: POLYISOPRENOID PHOSPHATES AND GLYCOSYLATION OF PROTEINS

Glycosylated phosphopolyisoprenoid alcohols are the carriers of oligosaccharide units for transfer to proteins and as glycosyl donors, i.e., they are substrates for glycosyl transferases for the biosynthesis of glycans in a similar manner to the cytosolic sugar nucleotides. They differ from the latter in their intracellular location, with the lipid portion in the membrane of the endoplasmic reticulum and the oligosaccharide portion specifically located either on the cytosolic or lumenal face of the membrane. The degree of unsaturation and chain length of the product is important for recognition by the enzymes in the next stage of the pathway.

Dolichol phosphates: In eukaryotes, *N*-glycosylation begins on the cytoplasmic side of the endoplasmic reticulum with the transfer of carbohydrate moieties from nucleotide-activated sugar donors, such as uridine diphosphate *N*-acetylglucosamine, onto dolichol phosphate. Then, *N*-acetylglucosamine phosphate is added to give dolichol-pyrophosphate linked to *N*-acetylglucosamine, to which a further *N*-acetylglucosamine unit is added followed by five mannose units, the last catalyzed by dolichol phosphate mannose synthase, which is also essential for **GPI-anchor** biosynthesis. The dolichol-PP glycan is shown in Figure 21.6.29.

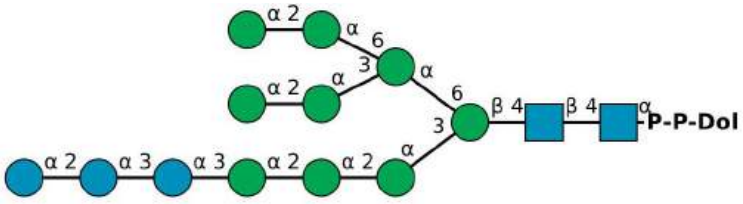


Figure 21.6.29: Dolichol-pyrophosphate glycan. Blue squares GlcNAc, Green circles Man, blue circles Glc. Cheng, K. et al., 'DrawGlycan-SNFG: a robust tool to render glycans and glycopeptides with fragmentation information', Glycobiology 27(3): 200-205, 2017

The resulting dolichol-pyrophosphate heptasaccharide is then flipped across the endoplasmic reticulum membrane to the luminal face with the aid of a “flippase”. Four further mannose and three glucose residues are added to the oligosaccharide chain by means of glycosyltransferases, which utilize as donors dolichol-phospho-mannose and dolichol-phospho-glucose, which are also synthesized on the cytosolic face of the membrane and flipped across to the luminal face. In humans, the final lipid product is a C_{95}-dolichol pyrophosphate-linked tetradecasaccharide, the oligosaccharide unit of which is transferred from the dolichol carrier onto specific asparagine residues on a developing polypeptide in the membrane. The carrier dolichol-pyrophosphate is dephosphorylated to dolichol-phosphate then diffuses or is flipped back across the endoplasmic reticulum to the cytoplasmic face.

The **Archaea** use dolichol in their synthesis of lipid-linked oligosaccharide donors with both dolichol phosphate (Euryarchaeota) and pyrophosphate (Crenarchaeota) as carriers; these can have variable numbers of isoprene units many of which can be saturated. In the haloarchaeon *Haloferax volcanii*, for example, a series of C_{55} and C_{60} dolichol phosphates with saturated isoprene subunits at the α- and ω-positions is involved in the glycosylation reaction of target proteins, while similar lipid carriers of oligosaccharide units appear to be present in methanogens. Archaea of course use **isoprenyl ethers linked to glycerol** as major membrane lipid components in addition to unusual carotenoids such as the C_{50} bacterioruberins. In many of these species, isoprenoid biosynthesis is via the 'classical' mevalonate pathway, but in other species, some aspects of this pathway differ.

Undecaprenylphosphate: Most other bacteria use undecaprenyl-diphosphate-oligosaccharide as a glycosylation agent in a similar way for the biosynthesis of peptidoglycan, the main component of most bacterial cell walls and a structure unique to bacteria, of many other cell-wall polysaccharides, including lipopolysaccharides, *O*-antigenic polysaccharides and capsular polysaccharides, and of *N*-linked protein glycosylation in both in Gram-negative and Gram-positive bacteria. **Undecaprenyl phosphate** (a C_{55} isoprenoid), also referred to as bactoprenol, is the essential lipid intermediate. Its structure is shown in Figure 21.6.30.

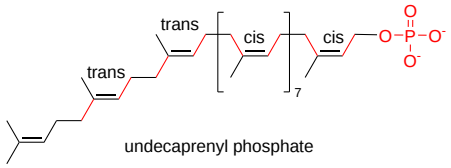


Figure 21.6.30: Undecaprenyl phosphate

It differs from the dolichol phosphates mainly in that the terminal unit is unsaturated, and is synthesized by the addition of eight units of isopentenyl pyrophosphate to farnesyl pyrophosphate, a reaction catalyzed by undecaprenyl pyrophosphate synthase in the cytoplasm, followed by the removal of a phosphate group. Undecaprenyl phosphate is required for the synthesis and transport of glycans for external polymer formation. Thus, glycans are covalently transferred to the carrier lipid by membrane-embedded or membrane-associated enzymes using nucleotide-activated precursors. For example, the carrier lipid with GlcNAc-MurNAc-peptide monomers, i.e., as lipid II, is hydrophilic and is then transported across the cytoplasmic membrane to external sites for peptidoglycan formation. Similarly, it is required for the synthesis of **lipoteichoic acids** and **lipopolysaccharide *O*-antigens**.

Lipid II: Undecaprenyl diphosphate-MurNAc-pentapeptide-GlcNAc, often simply termed **lipid II**, is the last significant lipid intermediate in the construction of the peptidoglycan cell wall in bacteria (Lipid I is the biosynthetic precursor lacking the *N*-acetylgluosamine residue). Its structure is shown below in Figure 21.6.31.

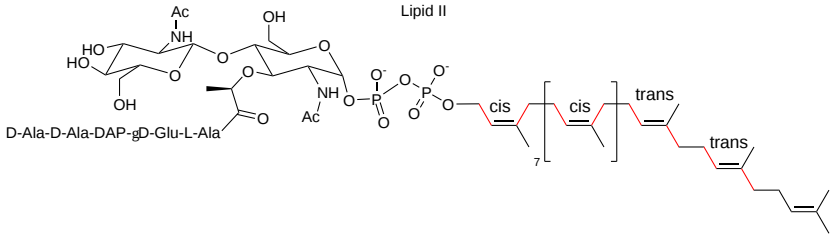


Figure 21.6.31: Lipid II

This molecule must be translocated from the cytosolic to the exterior membrane of the organism, and three different protein classes have been identified that can accomplish this of which ATP-binding cassette (ABC) transporters are best characterized. Once across the membrane, lipid II is cleaved to provide the MurNAc-pentapeptide-GlcNAc monomer, which undergoes polymerization and crosslinking to form the complex peptidoglycan polymer that provides strength and shape to bacteria. The undecaprenyl-pyrophosphate remaining is hydrolyzed to undecaprenyl phosphate by a membrane-integrated member of the type II phosphatidic acid phosphatase family and is recycled back to the interior of the membrane by an as yet unidentified transport mechanism. The turnover rate is very high so the lipid II cycle is considered to be the rate-limiting step in peptidoglycan biosynthesis. There is also some evidence that lipid II has a function on the inner leaflet of the cytoplasmic membrane in organizing the proteins of the cytoskeleton. Because of its highly conserved structure and accessibility on the surface membrane, the proteins involved in the synthesis and transport of lipid II are considered important targets for the development of novel antibiotics. Figure 21.6.*xx*:

In a few prokaryotes, the membrane intermediate has a polyprenyl-monophosphate-glycan structure instead of lipid II, and undecaprenyl-phosphate-L-4-amino-4-deoxyarabinose is involved in **lipid A** modification in Gram-negative bacteria, for example. There are obvious parallels with the involvement of glycosylated phosphopolyisoprenoid alcohols as carriers of oligosaccharide units for transfer to proteins and as glycosyl donors in higher organisms (see above).

β-D-Mannosyl phosphomycoketide is an isoprenoid phosphoglycolipid found in the cell walls of *Mycobacterium tuberculosis*, the lipid component of which is a C_{32}-mycoketide, consisting of a saturated oligoisoprenoid chain with five chiral methyl branches. It acts as a potent antigen to activate T-cells upon presentation by CD1c protein. Its structure is shown in Figure 21.6.32.

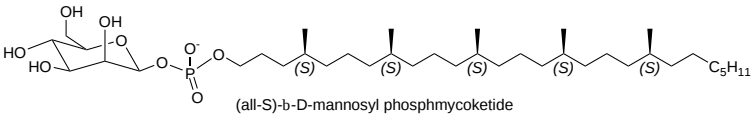


Figure 21.6.32 β-D-Mannosyl phosphomycoketide

21.6.3.7: FARNESYL PYROPHOSPHATE AND RELATED COMPOUNDS

Farnesyl pyrophosphate is a key intermediate in the biosynthesis of sterols such as **cholesterol**, and it is the donor of the farnesyl group in the biosynthesis of dolichols and polyprenols (see above) as well as for the isoprenylation of many proteins (see the web page on **proteolipids**). However, it is also known to mediate various biological reactions in its own right via interaction with a specific receptor. It is synthesized by two successive phosphorylation reactions of farnesol.

Presqualene diphosphate is unique among the isoprenoid phosphates in that it contains a cyclopropylcarbinyl ring. In addition to being a biosynthetic precursor of squalene, and thence of cholesterol, it is a natural anti-inflammatory agent, which functions by inhibiting the activity of phospholipase D and the generation of superoxide anions in neutrophils.

Their structures are shown in Figure 21.6.32.

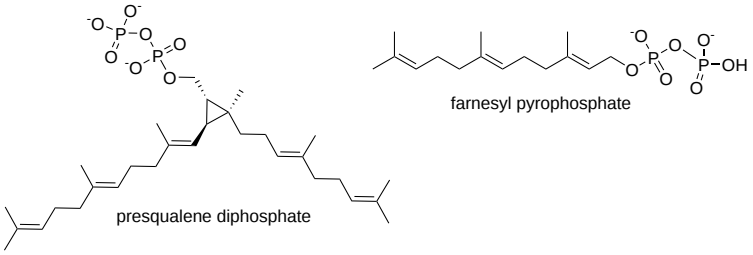


Figure 21.6.32 Structures of farnesyl pyrophosphate and presqualene diphosphate

CHAPTER OVERVIEW

22: BIOSYNTHESIS OF AMINO ACIDS, NUCLEOTIDES, AND RELATED MOLECULES

22.1: Overview of Nitrogen Metabolism
22.2: Biosynthesis of Amino Acids
22.3: Molecules Derived from Amino Acids
22.4: Biosynthesis and Degradation of Nucleotides

22.1: OVERVIEW OF NITROGEN METABOLISM

22.1.1: INTRODUCTION

Organic chemistry is usually described as the chemistry of carbon-containing molecules. But isn't that definition a bit carbon centric, especially since the prevalence of oxygen-containing molecules is staggering? What about nitrogen? We live in a dinitrogen-rich atmosphere (80%), and all classes of biomolecules (lipids, carbohydrates, nucleic acids, and proteins) contain nitrogen. Dinitrogen is very stable, given its triple bond and nonpolarity. We rely on a few organisms to fix N_2 from the atmosphere to form ammonium (NH_4^+), which through nitrification and denitrification can form nitrite (NO_2^-), nitrate (NO_2^-), nitric oxide (NO), and nitrous oxide (N_2O), the latter being a potent greenhouse gas. We'll concentrate on the metabolic fate of amino groups in amino acids and proteins in the next section. Before exploring their fates, look at Figure 22.1.1 which shows an overall view of the biological nitrogen cycle. The study of biochemistry should encompass more than homo sapiens and expand to the ecosystem in which we are such a small but damaging part.

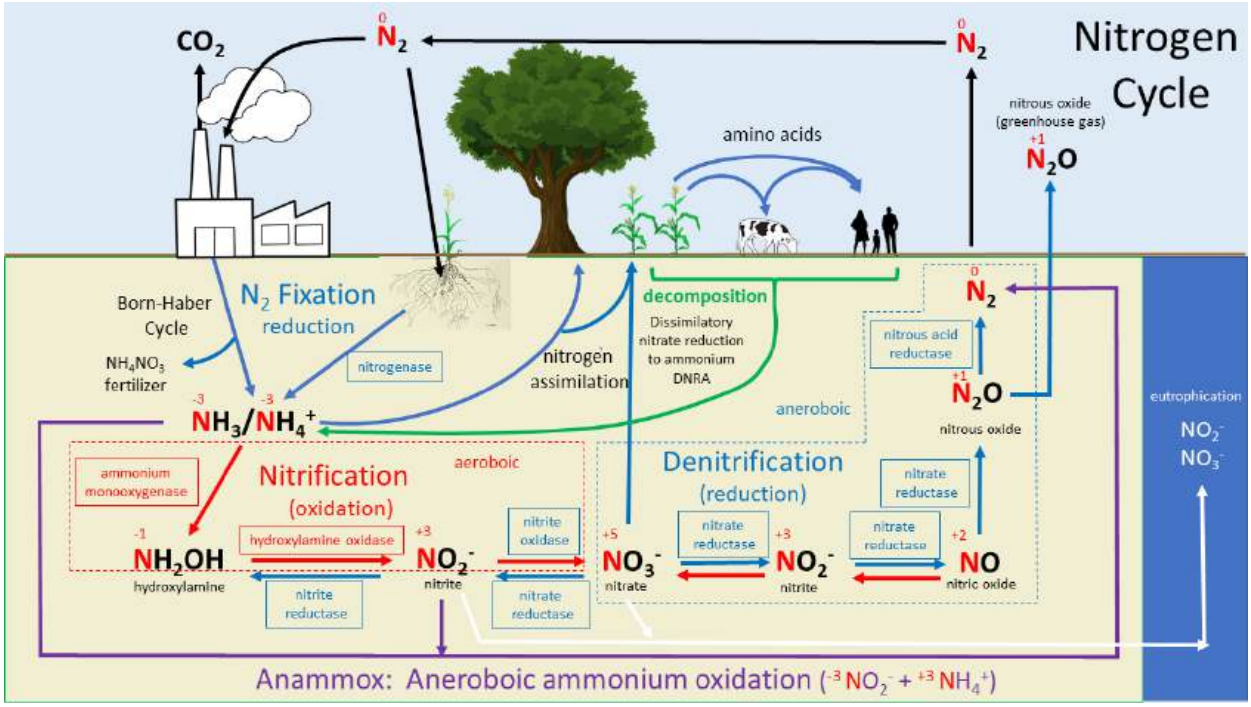


Figure 22.1.1: Nitrogen Cycle

Let's break down the diagram from a biochemical perspective. There are aerobic and anaerobic processes (conducted by bacteria). Nitrogen-containing substances include both inorganic (ammonium, nitrate, nitrite) and organic (amino acids, nucleotides, etc) molecules. The reactions shown are oxidative and reductive (note: the oxidation number of the nitrogen atoms in the molecules is shown in red). Most of the reactions are carried out underground by bacterial and Archaeal microorganisms.

Here are some of the major reactions:

- N_2 fixation (a reduction): N_2 from the air is converted by bacteria to ammonium (NH_4^+) by the enzyme nitrogenase of soil prokaryotes. The energetically disfavored reaction requires lots of ATPs. Ammonium once made can then be taken up by primary producers like plants and incorporated into biomolecules such as amino acids, which animals consume. For those who may still believe that people have marginal effects on our biosphere, consider this. We may soon fix more N_2 to NH_3 through the industrial Born-Haber reaction (used for fertilizer and explosive productions) that is all made by the biosphere. Much of the nitrogen in use comes from the Born-Haber reaction. The excess NH_4^+ (upwards of 50%) produced industrially and which enters the soil in fertilizers (mostly as NH_4NO_3) has overwhelmed nature's ability to balance the nitrogen cycle and is not taken up by plants. It is metabolized by microorganisms to nitrite and nitrate.
- Nitrification: Ammonium is converted to nitrite by ammonia-oxidizing aerobic microorganisms and further to nitrate by a separate group of nitrite-oxidizing aerobic bacteria. Here are the reactions (Rx 1 and 2) to produce nitrate through a hydroxylamine intermediate, followed by the formation of nitrate (Rx 3).

$$NH_3 + O_2 + 2e^- \rightarrow NH_2OH + H_2O \text{ Rx 1}$$

$$NH_2OH + H_2O \rightarrow NO_2^- + 5H^+ + 4e^- \text{ Rx 2}$$

$$NO_2^- + 1/2\ O_2 \rightarrow NO_3^-\ Rx\ 3$$

These added ions exceed soil capacity and end up runoff water, polluting our rivers and lakes.

- Denitrification: This anaerobic reaction pathway reproduces N_2 from nitrate Here is the net reaction:

$$2NO_3^- + 10e^- + 12H^+ \rightarrow 2N_2 + 6H_2O$$

- Anammox reaction: This more recently discovered bacterial anaerobic reaction pathway converts ammonium and nitrate to N_2. Here is the net reaction

$$NO_2^- + NH_3^+ \rightarrow N_2 + 2H_2O$$

- Ammonification (not to be confused with mummification) occurs when plants and animals decompose, which returns ammonium to the soil for reuse by plants and microbes.

These reactions are shown in the abbreviated Nitrogen Cycle shown in Figure 22.1.2.

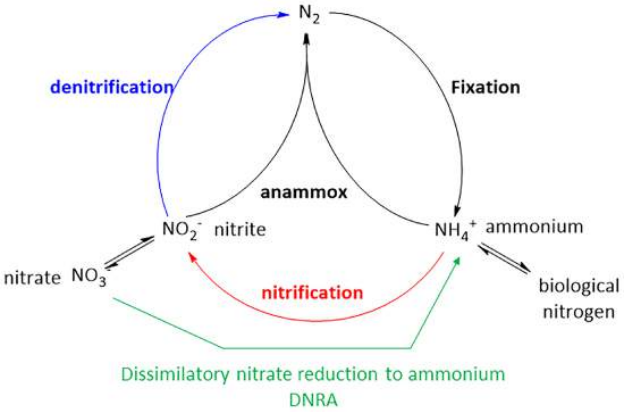


Figure 22.1.2: Abbreviated Nitrogen Cycle

Nitrogen metabolites are nutrients for plants and perhaps the most important nutrients in the regulation of plant growth (primary productivity) and in regulating life diversity in the biosphere. All living organisms require feedstocks to produce energy and as substrates for biosynthetic reactions. Which is used depends on the organism. Plants are primary producers so they use their synthesized carbohydrates for both energy production and biosynthesis. For carnivores, proteins and their derived amino acids are the source of energy (through oxidation) and serve as biosynthetic precursors. For omnivorous organisms, the source of energy depends on the "fed" state. With abundant food resources, carbohydrates, and lipids are the source of energy. Unlike carbohydrates and lipids, which can be stored as glycogen and triacylglycerols for future use, excess protein, and their associated amino acids can not be stored, so amino acids can be eliminated or used for oxidative energy.

In the fed state, carbohydrates are the main source, while in the unfed state, lipids take a predominant role. Under starving conditions, the organism's own proteins are broken down and used for oxidative energy production and for any biosynthesis that remains. In diseased states like diabetes, which can be likened to a starving state in the presence of abundant carbohydrates, both lipids and amino acids become the sources of energy.

How are amino acids in animals oxidatively metabolized? Many pathways could be used to do so but it would seem logical that NH_4^+ would be removed and the carbons in the remaining molecule would eventually enter glycolysis or the TCA cycle in the form of ketoacids. NH_4^+ is toxic in high concentrations. Ammonium is not oxidized to nitrite or nitrates in humans as occurs in the soil by microorganisms. It can be recycled back into nucleotides or amino acids, and excess amounts are eliminated from the organism. Both processes must be highly controlled. We will turn out attention to the oxidation of amino acids in the next section.

22.1.2: NITROGENASE: AN INTRODUCTION

Beauty is in the eye of the beholder.

As the domain of biochemistry covers the entire biological world, the extent of coverage of a given topic in textbooks can depend, in part, on the interest and experiences of the author(s) presenting the material. Is relevance a metric that should determine coverage? If so, books focused on human or medical biochemistry would surely omit photosynthesis. If topics are selected based on their importance for life, then photosynthesis must surely be covered. If so, then nitrogenase must also be included. If the degree of chemical difficulty for a chemical reaction and the amazing eloquence of the evolved biochemistry solution is considered, then both photosynthesis and nitrogen fixation must be presented. Even though nitrogen fixation is a reductive reaction, it shares strong similarities with the oxygen-evolving complex of photosynthesis. They catalyze enormously important redox reactions that involve an abundant atmospheric gas using a very complicated and unique inorganic metallic cofactor that evolution has selected as uniquely suited for the job.

Every first-year student of chemistry can draw the Lewis structure of dinitrogen, N_2, which contains a triple bond and a lone pair on each nitrogen. If Lewis structures speak to them, they should be able to state that the triple bond makes N_2 extraordinarily stable, thus explaining why we can breathe an atmosphere containing 80% N_2 and not die. If they have taken biology, they are also aware that very few biological organisms can utilize N_2 as a substrate, as this requires breaking bonds between the nitrogen atoms, a chemical process reserved for nitrogen "fixing" bacteria found in rhizomes of certain plants. Lastly, they probably memorized that high pressure and temperature are needed in the Haber-Bosch process used to convert N_2 and H_2 to ammonia, NH_3. As with any scientific advance, the Haber-Bosch process has brought both harm (it's used for explosive weapons) and good (fertilizers). This process now fixes enough N_2 in the form of fertilizers to support half of the world's population, with nitrogenase supporting the rest. Efforts are underway to genetically modify plants to make nitrogenase, eliminating the need for fertilizers but perhaps creating unforeseen problems of its own.

You might be surprised to find out that at room temperature the equilibrium constant favors ammonia formation, hence $\Delta G^0 < 0$. The reaction is favored enthalpically as it is exothermic at room temperature. It is disfavored entropically as should be evident from the balanced equation:

$$N_2(g) + 3\,H_2(g) \rightarrow 2\,NH_3(g).$$

The thermodynamic parameters for the reaction (per mol) are ΔH° = –46.2 kJ, ΔS° = –389 J K–1, and ΔG° = –16.4 kJ at 298 K

The entropy is negative since the reaction proceeds from 4 molecules to two molecules. From an enthalpy perspective, if you raise the temperature of an exothermic reaction, you drive it in a reverse direction. If you increase the pressure, you shift the equilibrium to the side that has the fewest number of molecules.

If the reaction is favored thermodynamically at room temperature, why doesn't it proceed? This story sounds familiar as this same descriptor applies to the oxidation of organic molecules with dioxygen. There we showed using MO theory that the reaction is kinetically slow. Same with NH_3 formation. A superficial way to see this is that we must break bonds in the stable N_2 to start the reaction, leading to a high activation energy, and making the reaction kinetics sluggish.

One could jump-start the reaction by raising the temperature, but that would slow an exothermic reaction. The Keq (or K_D) and ΔG^0 are functions of temperature and for this reaction, the reaction becomes disfavored at higher temperatures. The solution Haber found was high pressure, forcing the reaction to the side that has fewer molecules of gas, and high temperature to overcome the activation energy barrier and make the reaction kinetically feasible. A complex metal catalyst (magnetite - Fe_3O_4 -with metal oxides like CaO and Al_2O_3 which prevent reduction of the Fe with H_2) provides an absorptive surface to bring reagents together and facilitate bond breaking in H_2 and N_2.

In photosynthesis, the **oxygen-evolving complex** (OEC) with Mn, Fe, S, and Ca is used to oxidize another very stable and ubiquitous molecule, H_2O. Now we explore the amazing mechanisms behind the nitrogenase complex which fixes N_2 to form NH_3 in a reductive fashion.

What might be needed to drive this reaction biologically? You might surmise the list to include:

- a source of energy, most likely ATP, to facilitate this complex reaction;
- a source of electrons as the N atoms move from an oxidation state of 0 in elemental N to 3- in NH3; this source turns out to be a protein called flavodoxin or ferredoxin. Of course, these electrons also have interesting sources before they were in the electron carriers of these proteins;
- some pretty amazing metal centers to accept and donate electrons in a controlled way; these centers are mostly FeS clusters with an additional cluster containing molybdenum (Mo). The clusters are named F, P, and M
- a source of hydrogen; you might have guessed correctly that it's not H_2 gas (from where would that come?), but H^+ ions which are pretty ubiquitously available.
- a net reaction that is different that the Haber-Bausch process ($N_2 + 3H_2 \rightarrow 2\,NH_3$).

Here is the actual reaction catalyzed by nitrogenase:

$$N_2 + 8\,e^- + 16\,ATP + 8\,H^+ \rightarrow 2\,NH_3 + H_2 + 16\,ADP + 16\,Pi.$$

Let's think a bit about the reaction. As electrons are added the attraction between the nitrogen atoms must decrease. Eventually, bonds between them must be broken. Protons could be easily added to maintain charge neutrality. A basic mechanism might involve intermediates as shown in Figure 22.1.3.

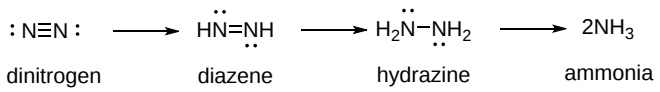


Figure 22.1.3: Possible intermediates in the conversion of dinitrogen to ammonia by nitrogenase

Nitrogenase can also other small molecules with triple bonds, including C≡O: and H-C≡C-H.

22.1.3: THE STRUCTURE OF NITROGENASE

Nitrogenase is a multiprotein complex in which the functional biological unit is built from **two** sets of the following dimeric structures:

- a homodimer of subunits, E and F, which have binding sites for the mobile carrier of electrons (the protein ferredoxin or flavodoxin), ATP, and an FeS cofactor (4Fe-4S, called the F cluster) which accept electrons. These subunits are hence called the nitrogenase reductase subunits
- a heterodimer of alpha and beta subunits. The a (alpha chain) binds the 8Fe-7S F cluster and the Fe-S-Mo M cluster. These subunits comprise the (di)nitrogenase catalytic subunits. The iron-molybdenum M cluster is in the α subunit and is where N_2 reduction occurs. The P-cluster is between the α and β subunits and facilitates electron flow between the Fe-protein (F cluster) and FeMo-cofactor (M cluster)

For clarity, one-half of the overall structure of the protein complex with bound ATP and metal centers is shown in Figure 22.1.4.

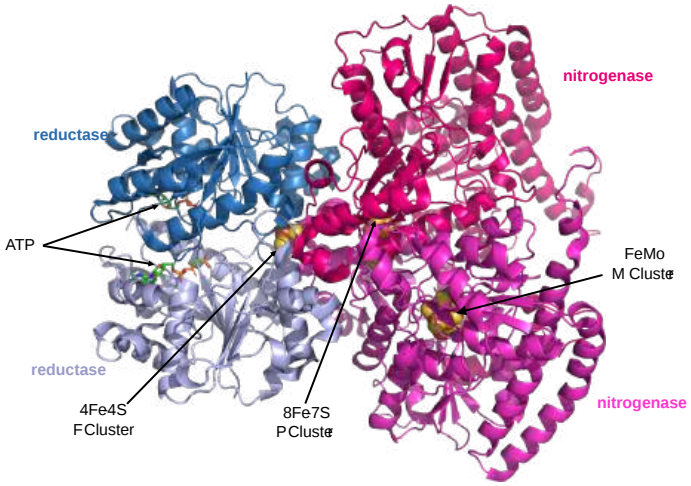
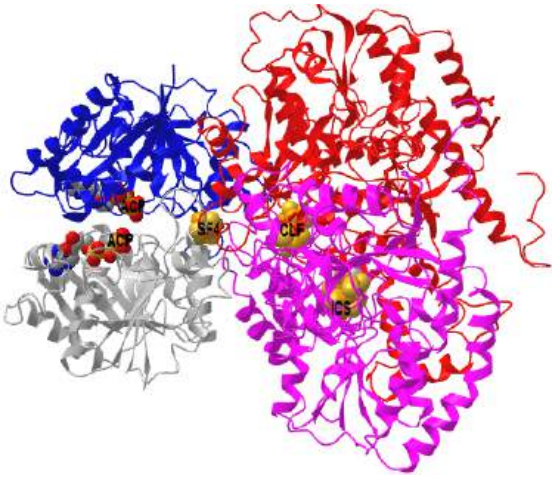


Figure 22.1.4: Nitrogenase structure (4wzb)

This half-structure consists of a homodimer of the reductase monomers and a heterodimer of nitrogenase subunits.

Figure 22.1.5 shows an interactive iCn3D model of the half structure of nitrogenase complex from Azotobacter vinelandii (4WZB) (long load).



NCBI iCn3D Figure 22.1.5: Nitrogenase complex from Azotobacter vinelandii (4WZB). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...zPNjfPRwq8MSeA (long load))

The structure is color-coded in a fashion similar to Figure 4. The F cluster is labeled as SF4, the P cluster as CLF (FE(8)-S(7) cluster), and the M cluster as ICS (iron-sulfur-molybdenum cluster with interstitial carbon).

The reductase subunits (called Av2 in *Azotobacter vinelandii*), accept electrons from ferredoxin and is where the ATP analog, ACP (phosphomethylphosphonic acid adenylate ester), and the 4Fe-4S F cluster is bound. The nitrogenase subunits (called Av1 in *Azotobacter vinelandii*) convert N_2 to NH_3 and is where the 8Fe-7S P cluster, FeMo (C Fe_7 Mo S_9) M Cluster is bound.

An enhanced view of the bound cofactors and ATP is shown in the same spatial orientation in Figure 22.1.6 below.

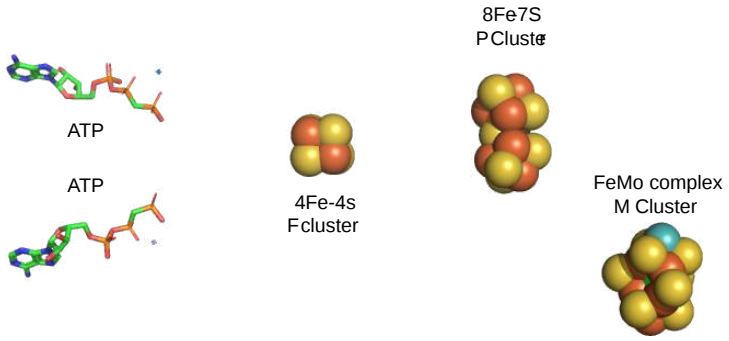


Figure 22.1.6: Bound cofactor metal clusters and ATP in nitrogenase

You can easily image the direction of the flow of electrons from the F cluster to the P cluster to the M cluster.

The metal centers are shown in Figure 22.1.7 in more detail in both line and space fill views.

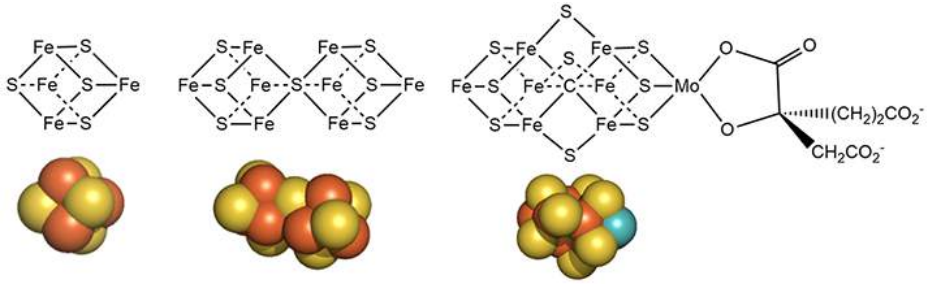


Figure 22.1.7: Detailed structures of the metal cofactors in nitrogenase

Mo is bound to 3 sulfur ions and oxygen from 3-hydroxy-3-carboxy-adipic acid as shown in Figure 7 above.

The M cluster has an interstitial carbide ion that derives from -CH_3 attached to the sulfur of S-adenosyl-methionine (SAM) allowing the carbide to be labeled with either ^{13}C or ^{14}C for mechanistic studies. These labeled carbides are not exchanged or used as a substrate when the enzyme undergoes catalytic turnover. Hence it seems that the carbide probably just stabilizes the M cluster. it won't be shown in the figures below showing more detailed mechanisms.

We'll focus on two features of the reaction mechanism, ATP hydrolysis and the flow of electrons and protons to N_2 as it is reduced to NH_3.

22.1.4: ATP HYDROLYSIS

ATP binds in the reductase subunit (AV2) where ferredoxin brings in electrons, and where the F metal cluster is bound. Much as GTP hydrolysis controls conformational change and subunit dissociation in the heterotrimeric $G_{\alpha\beta\gamma}$ in signal transductions, ATP hydrolysis in reductase (AV2) subunits, drives not only electron transfer but dissociation/reassociation of the reductase and nitrogenase catalytic subunits. It appears that 2 ATPs are hydrolyzed per electron transferred from the F cluster to the Mo M cluster. Since the oxidation numbers of N are 0 and -3 in N_2 and NH_3, respectively, sequential rounds of ATP hydrolysis and dissociation/reassociation occur. Note that the Fe-protein hydrolyzes ATP only when bound to the MoFe-protein.

The overall reaction involves the reduction of N_2 to two molecules of NH_3 at the FeMo cofactor. It involves a reductive elimination of hydrides that are bridged by Fe ions (Fe-H-Fe) in a reaction that also makes H_2 as a by-product. We'll describe this complex reaction next.

22.1.5: NITROGENASE REACTION: PART 1 - ADDITION OF ELECTRONS AND PROTONS

The sequential path of electrons from the reductase subunit containing the F cluster to the P and M clusters in the nitrogenase subunit should be apparent from the figures above. We will concentrate on the binding of N_2 and how it receives electrons from the M cluster. Figure 22.1.8 shows the FeMo-cofactor and some adjacent amino acid residues. Mo is labeled but not shown in spacefill. HCA is a bound molecule of 3-hydroxy-3-carboxy-adipic acid, which interacts with Mo. The carbide is shown in the middle as the green sphere and labeled CX.

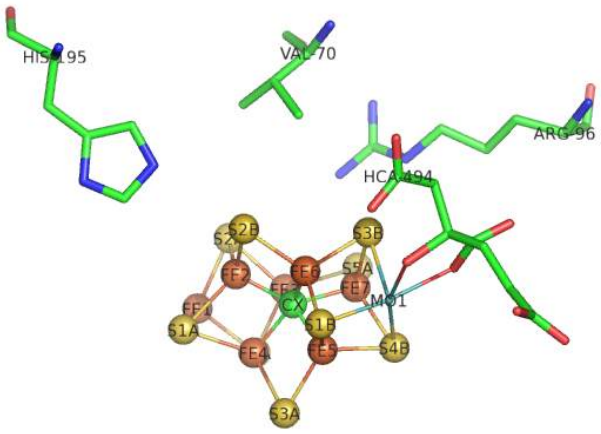


Figure 22.1.8: Side chains near the M cluster

If Val 70 is mutated to Ile, a substrate appears not to access the cluster suggesting that N_2 may interact with the top part of the structure with the residues shown acting as gatekeepers. His side chains are often found at enzyme active sites so you might expect His 195 to be a general acid/base. Mutations lead to drastic losses in the reduction of N_2. His 195 is involved in hydrogen bonds to sulfur S2B and bridges Fe2 and Fe3 in Figure 8, where reduction of N_2 likely occurs. If His 195 moves, it can form short H bonds between the imidazole N and an H bond to HFe2. If the ring is rotated 180^0, no proton transfer occurs from the surface. It appears that His195 might be involved in the first N_2 protonation event.

The Lowe and Thorneley (LT) model has been proposed as a mechanism for dinitrogen reduction. In this model, an electron and proton are added to the oxidized form of the enzyme (E_o) to produce E_1. This is repeated 3 more times to form sequentially, E_2, E_3, and E_4. Only then does N_2 bind and the reduction of N_2 occurs. Two of the added electrons are accepted by H^+ ions which form H_2, which is liberated on N_2 binding. Hence only 6 electrons are added to the actual N_2 molecule, in agreement with the change in oxidation numbers discussed above. The Lowe and Throneley model is shown in Figure 22.1.9.

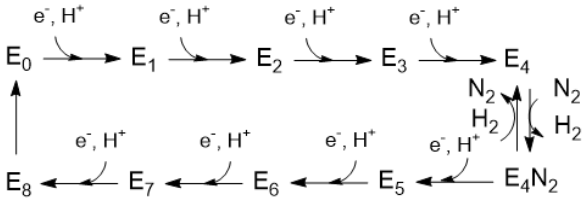


Figure 22.1.9: Lowe and Throneley model for electron and proton additions in nitrogenase

The crystal structure shows 2 ATP analogs bound to the reductase subunit. The stoichiometry of the reaction shows 16 ATP used. Simple math suggests that 2 ATP are cleaved to support the entry of one electron into the complex, assuming 8 transferred electrons (6 to N_2 and 2 to 2 protons to form H_2).

Part 1 - E1-E4: A potential structure for the E4 intermediate is shown in Figure 22.1.10.

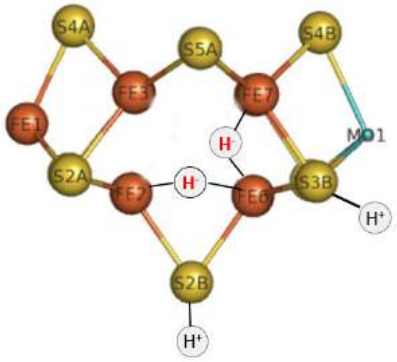


Figure 22.1.10: The E4 Janus intermediate in the reduction of N_2.

Note the central carbide is not shown. This is often called the **Janus intermediate** as it is halfway through the catalytic cycle. It is named for Janus, the Roman god of beginnings and transitions, and has been ascribed to gates, doors, doorways, and passages. Janus is typically

shown with two faces, one looking to the future and one to the past (image below: DOI:10.1590/2177-6709.21.1.018-023.oin. License CC BY 4.0 Creative Commons Attribution 4.0 International)

The hydrides bridge 2 Fe ions so these are examples of three-center, two-electron bonds. The H^+ ions in Figure 10 balance the charge from the hydrides.

How does this reaction occur? We must look to organometallic chemistry to help us understand the mechanism of this and subsequent steps. Hydride equivalents have been added to the metals, associated with the oxidation of metal ions in the center. This particular reaction is called an **oxidative addition**. Presumably, the sulfur ions act as Lewis bases as they gain protons from a Lewis acid, probably His 195.

Oxidative addition reactions

Figure 22.1.11 shows oxidative additions to metal centers for three different types of reactants.

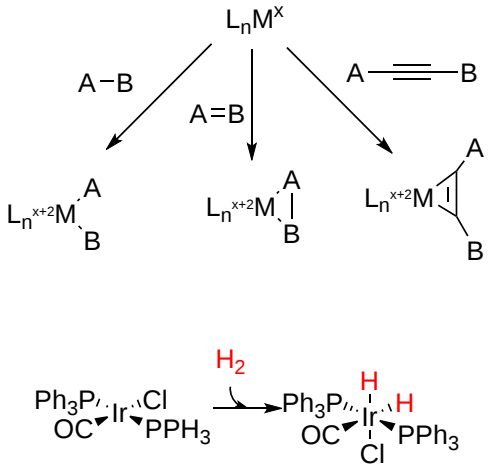


Figure 22.1.11: Three different types of oxidative addition reactions (after Schaller, http://employees.csbsju.edu/cschaller/ROBI1.htm)

In oxidative insertion, the oxidation state of the metal ion increases, hence the name oxidative. The hydrogens are now hydrides. Note the example for insertion of H_2, an example similar to the proposed hydride additions to the M cluster. Oxidative reduction occurs most readily when the two oxidation states of the metal ion are stable. It is likewise favored for metal centers that are not sterically hindered (makes sense if A-B is to be added) and if A-B has a low bond dissociation energy.

One way to study reaction intermediates is to trap them. If N_2 can't access the binding site and the temperature is reduced, the accumulated hydrides (and for charge balance the H+s) in E4 might leave in the opposite reaction, reductive elimination, which we will discuss below, as the reaction goes back to E1. In the elimination, they could form H_2 as metal gains back electrons in a reduction. The Val70Ile discussed above would allow an intermediate to be trapped.

22.1.6: NITROGENASE REACTION: PART 2 - REDUCTION OF N_2

Step E4 to E5 seems a bit bizarre as H_2 gas is released. This would seem to waste ATP but we should trust evolution has led to this mechanism for a reason. This mechanism, the reverse of oxidative addition, is another classic organometallic reaction, **reductive elimination**.

Reductive elimination reactions

In this reaction, a molecule is eliminated or expelled from the complex as the metal ion is reduced and adds two electrons. Figure 22.1.12 shows reductive elimination. Reductive elimination occurs most readily in higher oxidation state metal centers which can be stabilized on reduction. It occurs most readily from electron-rich ligands and if the other surrounding ligands are bulky. The dissociating species must also be cis to each other in the transition metal complex so they can form a bond with each other when they leave.

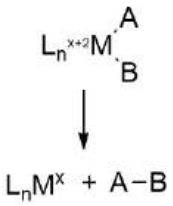


Figure 22.1.12: Reductive Elimination reaction (after Schaller, ibid)

Oxidative addition and reductive elimination (OA/RE) reactions at metal centers are often coupled together in organometallic **catalytic cycles**, in the same way as a histidine can act as a general acid and then accept a proton back as a general acid to complete the catalytic cycle. In the OA/RE reactions at metal centers, some rearrangements or other modifications can also occur. Think about it. The FeS clusters must return to their original oxidation state after the complete LT cycle. We will encounter another organometallic reaction after the addition of N_2, **migratory insertion**, in the second half of the reaction. Another advantage of coupling OA/RE is that the positive charge or oxidation state on the transition metal complex does not get too high, which is unstable. Making a cation with a positive charge more positive becomes more difficult, much as removing a second proton from a polyprotic acid is more difficult than removing the first (as reflected in the higher pKa for removal of the second proton).

Is H_2(g) really released? To study this, investigators have used alternative substrates like acetylene, HC≡CH (similar to N≡N), in the presence of D_2 and N_2 in an aqueous system. It helps to see Figure 22.1.9 again.

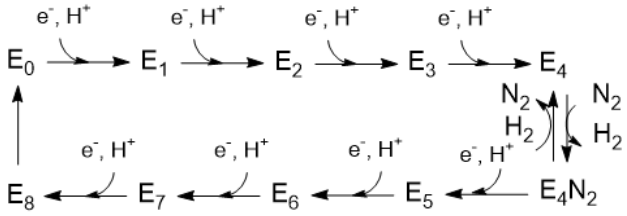


Figure 22.1.9: Lowe and Throneley model for electron and proton additions in nitrogenase

The acetylene was reduced and formed $C_2H_2D_2$ and C_2H_3D. Hence E4 must have had 2 Ds in it, and E2 probably 1. These results support the **reversible** reductive elimination mechanism for the E4 to E4:N_2 reaction above. Previously it had been shown that H^+ are reduced by D_2 in the presence of D_2 and N_2 in an aqueous system, so these results are consistent. In additional support, deuterium from D_2 is not incorporated into products ($C_2H_2D_2$, C_2H_3D, or HD) in the absence of N_2.

Let's return for a moment to the bridging hydrides as shown again Figure 22.1.10.

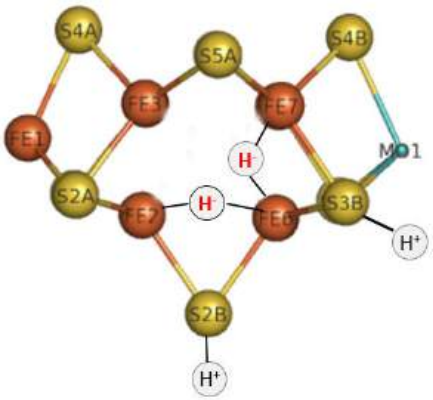


Figure 22.1.10: The E4 Janus intermediate in the reduction of N_2

To summarize, it appears likely that the reductive elimination of the two proximal bridging hydrides is the mechanism for the formation of H_2. The bridging hydrides, which are strong bases, are much less likely to be protonated than if the hydrides were terminal. A simple and competing protonation reaction could form H_2 as well, and if that occurred the reducing equivalents of the bridging hydrides would be lost. Hence the bridging hydrides (share by two metal centers) are more stable and hence can "wait" for the incoming N_2 reactants before their reducing equivalents are lost. They may convert to terminal hydrides eventually to facilitate substrate reduction (hydrogenation).

The Janus intermediate E4 is now in a position to bind N_2 and release H_2. For each H^- that binds to the M cluster, two H^+s bind to the M cluster sulfides for electrostatic stabilization.

Migratory Insertions

We need to consider one last common type of reaction at a metal center, **migratory insertions (MI)**. In a MI reaction, a group attached to a metal ion center is transferred to another group attached to the same metal. Figure 22.1.13 shows four examples of MI reactions.

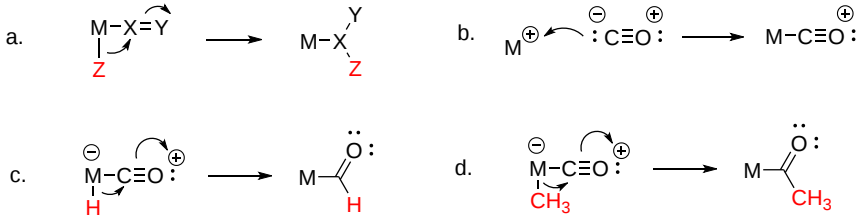


Figure 22.1.13: Migratory insertion reactions at metal centers

Panel (a) shows a generic MI reaction. Panel (b) shows the interaction of carbon monoxide, :C≡O: (isoelectronic and analogous to :N≡N:), with a metal center. Panel (c) shows how a hydride could engage in 1:1 insertion as it shifts and covalently bonds to the first atom of another ligand bound to the metal center. Note that metal does not have to have a negative charge. This could theoretically be important for the reduction of N_2 bound as a ligand through a coordinate covalent (dative) bond to the metal. Panel (d) shows the migration of an alkyl group.

We will see that the MI reaction is involved in adding Hs to N_2 starting not with N_2 but with the N_2H_2 stage as the intermediates insert into an Fe-H bond. N_2 is not reactive to the insertion of a hydride as is carbon monoxide, CO, which provides a positive oxygen to facilitate electron flow during the insertion. In addition, the oxygen become neutral after the reaction.

This offers a great explanation for how Nature chose the FeMo cluster for nitrogen fixation. The interaction of two hydrides requires a 4 Fe face (coordinated with the carbon) that allows for the storage of reducing equivalents for the initial reduction of N_2. The large M cluster is more stable and effectively held together by a central carbide anion. This also allows the metal centers to never change their oxidation state by more than 1 charge unit.

The source of the protons to form N_2H_2 comes directly from the two H_2 attached to the two sulfurs as they can't come from the hydrides which are released as H_2. The seeming wasteful reductive elimination of H_2 and energy is required to allow the kinetically unreactive N_2 molecule to bind to the reduced and activate 4Fe face which is also electrostatically facilitated by the 2 bound protons in what has been called a push (reductive)-pull (protonation) reaction. This first step in N_2 reduction to N_2H_2 is the hardest.

22.1.7: OXIDATION STATES OF NITROGENASE FE CENTERS

First Half: It would be difficult to assign specific oxidation states to each Fe ion in the M complex. Instead, we can assign relative changes in the oxidation states as the reaction proceeds from E0 to E4. In each step of the LT model, 1 electron is added. We will first assign this to an average Fe ion, M^0, with an arbitrarily assigned oxidation state of 0. On the addition of 1 electron, the oxidation state would go from M^0 to M^{-1} as the metal is reduced. The M^{-1} state is then oxidized as an electron is transferred to H^+, and when 2 electrons are transferred, a single H^- is made. The diagram in Figure 22.1.14 shows the change in oxidation state in going from E0 to E4.

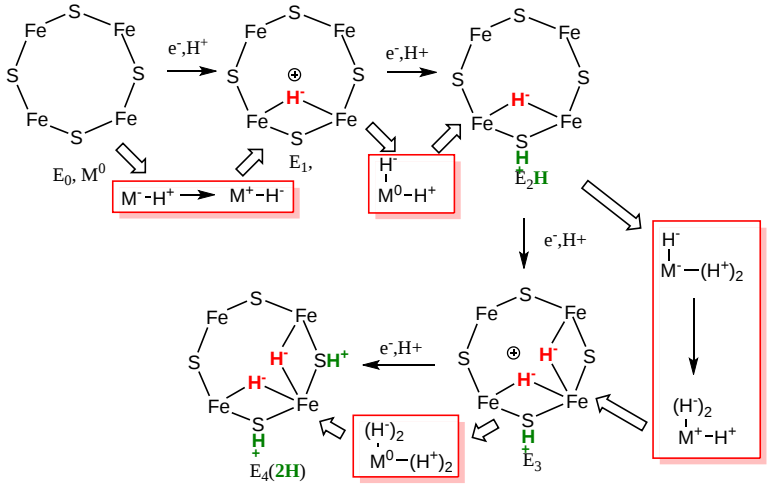


Figure 22.1.14: Change in oxidation state in going from E0 to E4

The red boxes highlight thermodynamic cycle-like steps which show how changes in the redox state of the Fe ions (M) could be visualized. Note that in going from E0 to E4, the actual oxidation state of M changes from 0 to +1 to 0 to +1 and back to 0. That is quite amazing given that 4 electrons have been added. Note also that in the red box going from step E0 to E1, M goes from -1 to +1 which corresponds to our

description of an oxidative addition when the metal center loses two electrons. This mechanism shows that nitrogenase could be considered a "hydride storage device".

Second Half (facing forward to production NH_3):

How does N_2 initially interact with E4? It must depend on how the hydrides are released as H_2, which evidence shows occurs by **reductive elimination** (re) and not hydride protonation (hp). On addition, the N_2 very quickly is converted to diazene, HN=NH, with the departing H_2 taking with it 2 H^+s and 2 electrons (or reducing equivalents). These events could occur as shown in Figure 22.1.15.

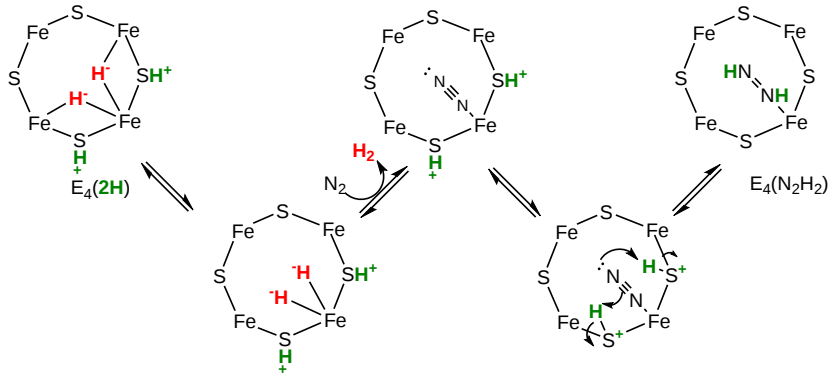


Figure 22.1.15: Reaction mechanism for the formation of N_2H_2

Now, with N_2 bound as diazene (N_2H_2) and H_2 released, the rest of the reaction could occur as shown below. One new step, a migratory insertion, is shown in Figure 22.1.16.

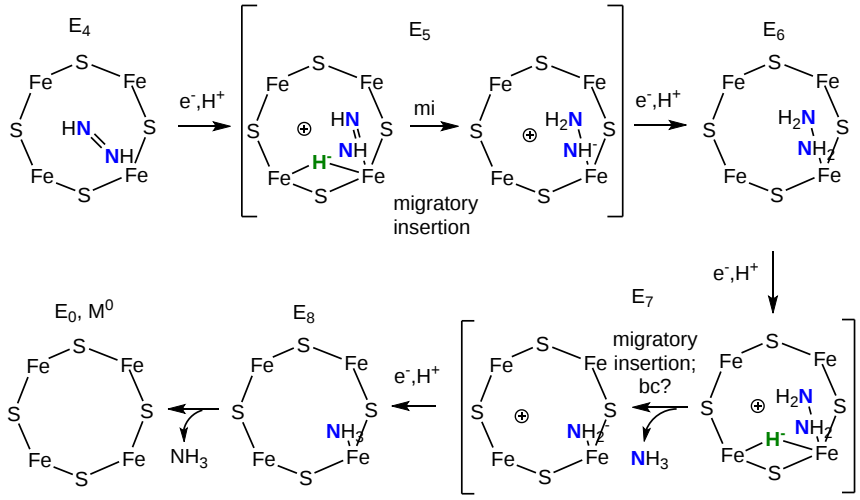


Figure 22.1.16: Reaction mechanism for the conversion N_2H_2 to NH_3.

The two halves of the reaction are similar with bridging hydrides utilized - the E4 Janus intermediate links the two halves together.

22.2: BIOSYNTHESIS OF AMINO ACIDS

22.2.1: INTRODUCTION

By the time many students get to the study of amino acid biosynthesis, they have seen so many pathways that learning new pathways for the amino acids seems daunting, even though they can be clustered into subpathways. Most know that from a nutrition perspective, amino acids can be divided into nonessential and essential (need external dietary supplementation) amino acids. These are shown for humans below.

- Nonessential amino acids: Alanine, Asparagine, Aspartate, Cysteine, Glutamate, Glutamine, Glycine, Proline, Serine, Tyrosine
- Essential amino acids: Arginine*, Histidine, Isoleucine, Leucine, Lysine, Methionine*, Phenylalanine*, Threonine, Tryptophan, Valine

Three of the essential amino acids can be made in humans but need significant supplementation. Arginine is depleted in processing through the urea cycle. When cysteine is low, methionine is used to replace it so its levels fall. If tyrosine is low, phenylalanine is used to replace it.

The amino acids can be synthesized from glycolytic and citric acid cycle intermediates as shown in Figure 22.2.1

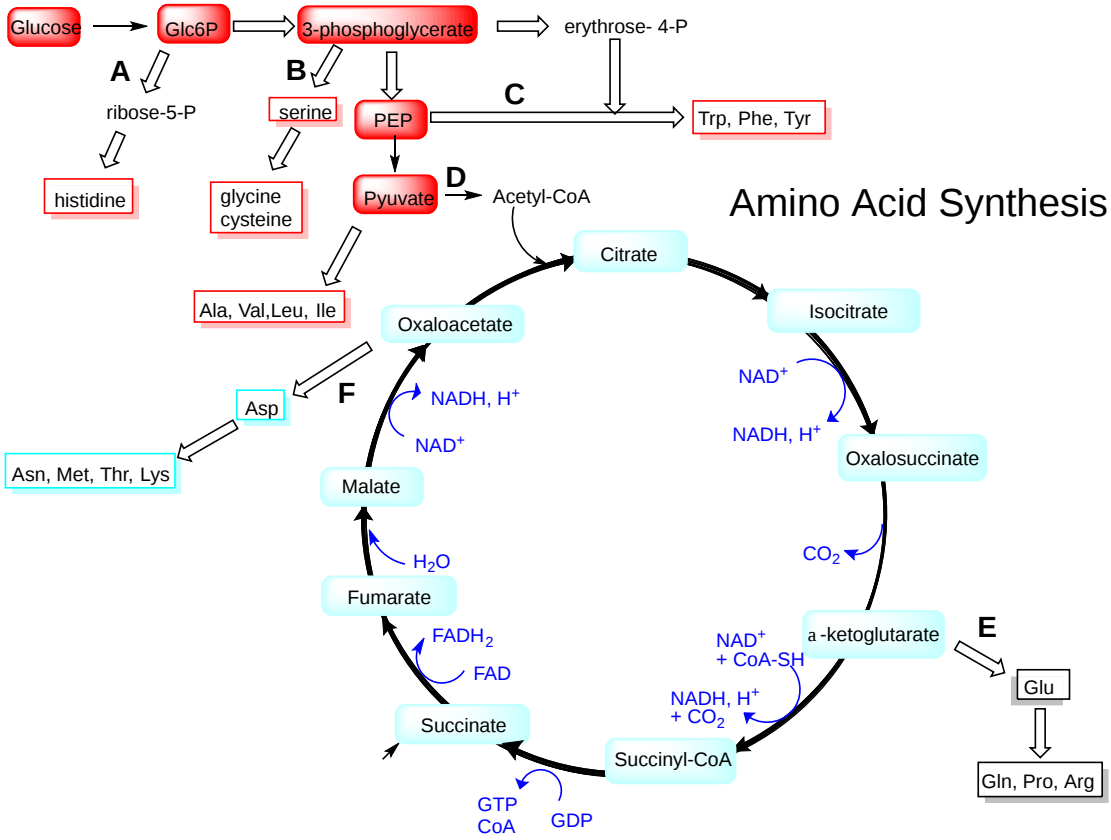


Figure 22.2.1: Summary amino acid synthesis from glycolytic and TCA intermediates

For this chapter subsection, we will provide only the basic synthetic pathways in abbreviated form without going into mechanistic or structural details

22.2.2: AMINO ACID SYNTHESIS FROM GLYCOLYTIC INTERMEDIATES

22.2.2.1: FROM GLUCOSE-6-PHOSPHATE: HISTIDINE

The synthesis of histidine from a phosphorylated form of ribose (derived from glucose-6-phosphate) is shown in Figure 22.2.2.

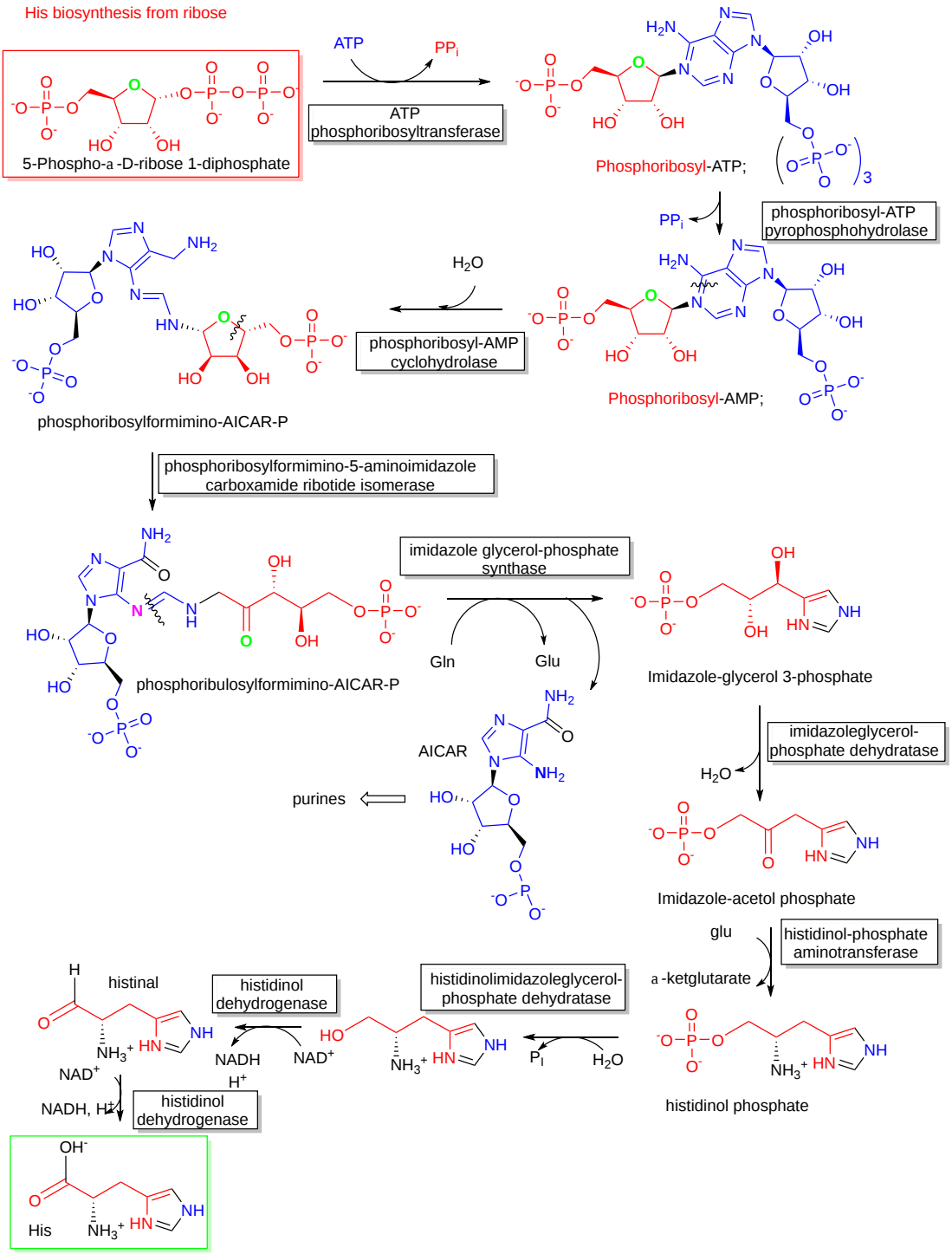


Figure 22.2.2: Synthesis of histidine from a phosphorylated form of ribose

22.2.2.2: FROM 3-PHOSPHOGLYCERATE: SERINE, GLYCINE, AND CYSTEINE

The synthesis of serine, glycine, and cysteine from 3-phosphoglycerate is shown in Figure 22.2.3.

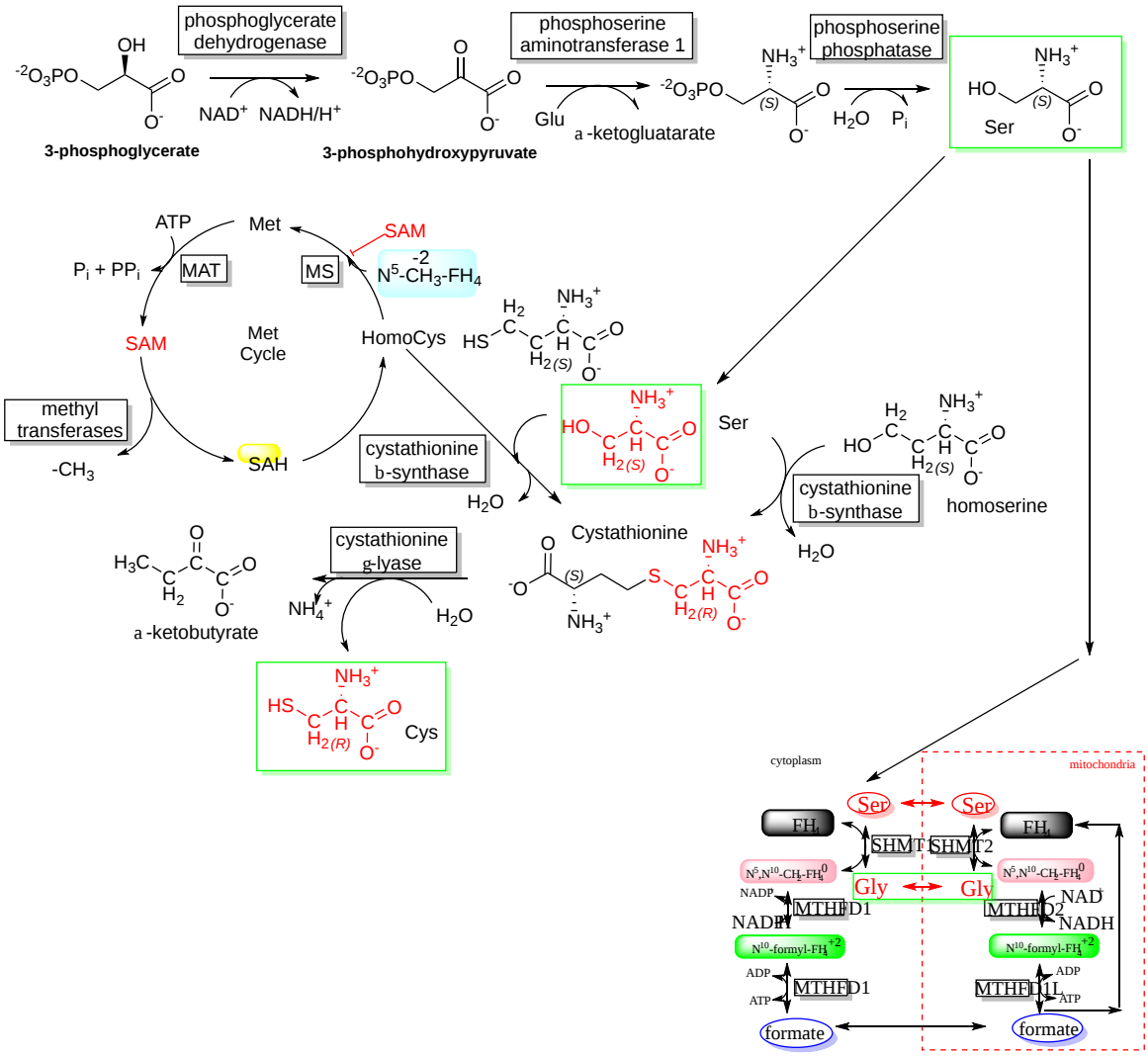


Figure 22.2.3: The synthesis of serine, glycine, and cysteine from 3-phosphoglycerate

22.2.2.3: FROM PHOSPHENOL PYRUVATE: THE AROMATICS - TRP, PHE, AND TYR

The synthesis of the first of the biosynthetic pathways for the aromatic amino acids phenylalanine, tryptophan, and tyrosine from phosphoenolpyruvate up to chorismate is shown in Figure 22.2.4.

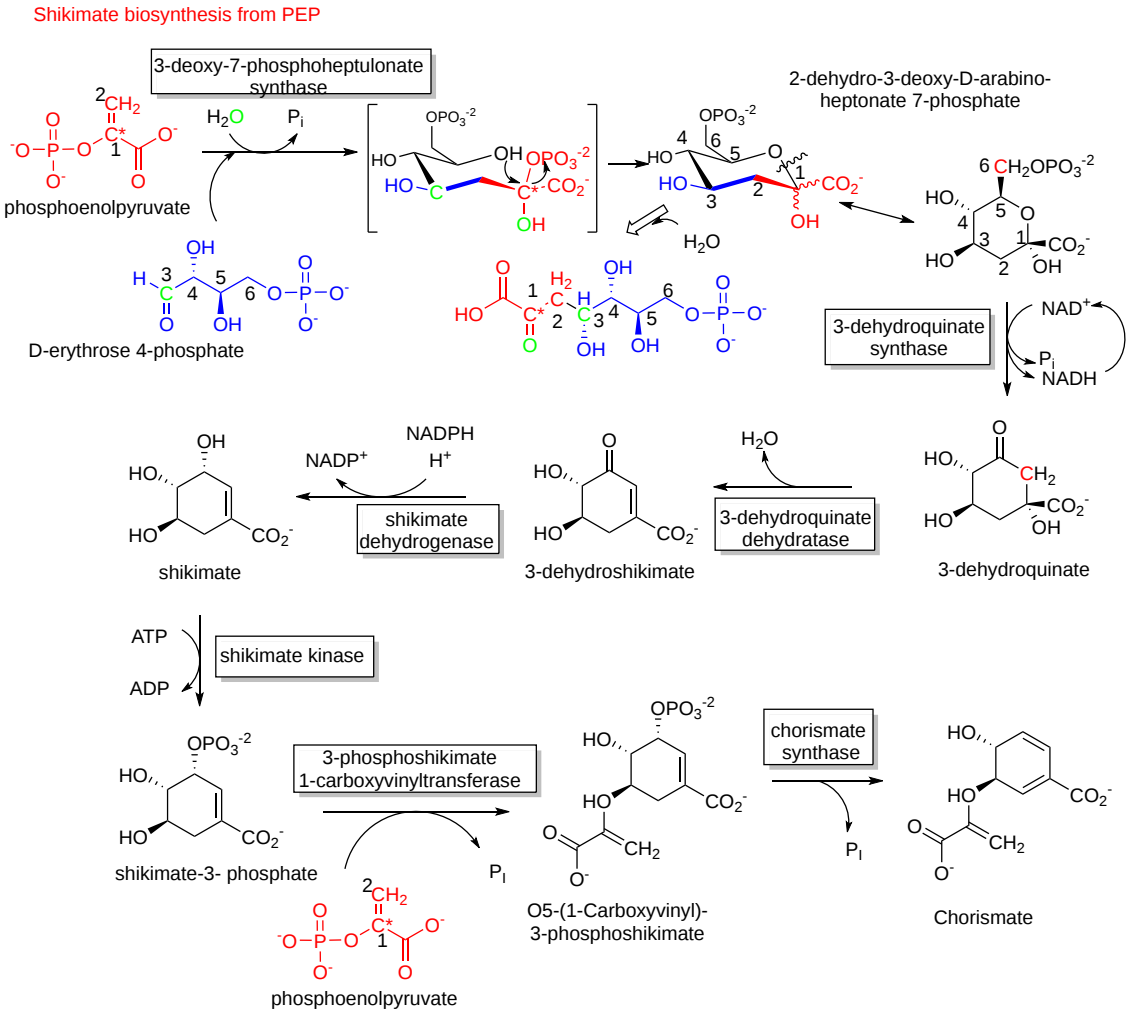


Figure 22.2.4: Synthesis of the first of the biosynthetic pathways for the aromatic amino acids phenylalanine, tryptophan, and tyrosine from phosphoenolpyruvate up to chorismate

Chorismate to tryptophan

The synthesis of the second half of the biosynthetic pathway for tryptophan from chorismate is shown in Figure 22.2.5

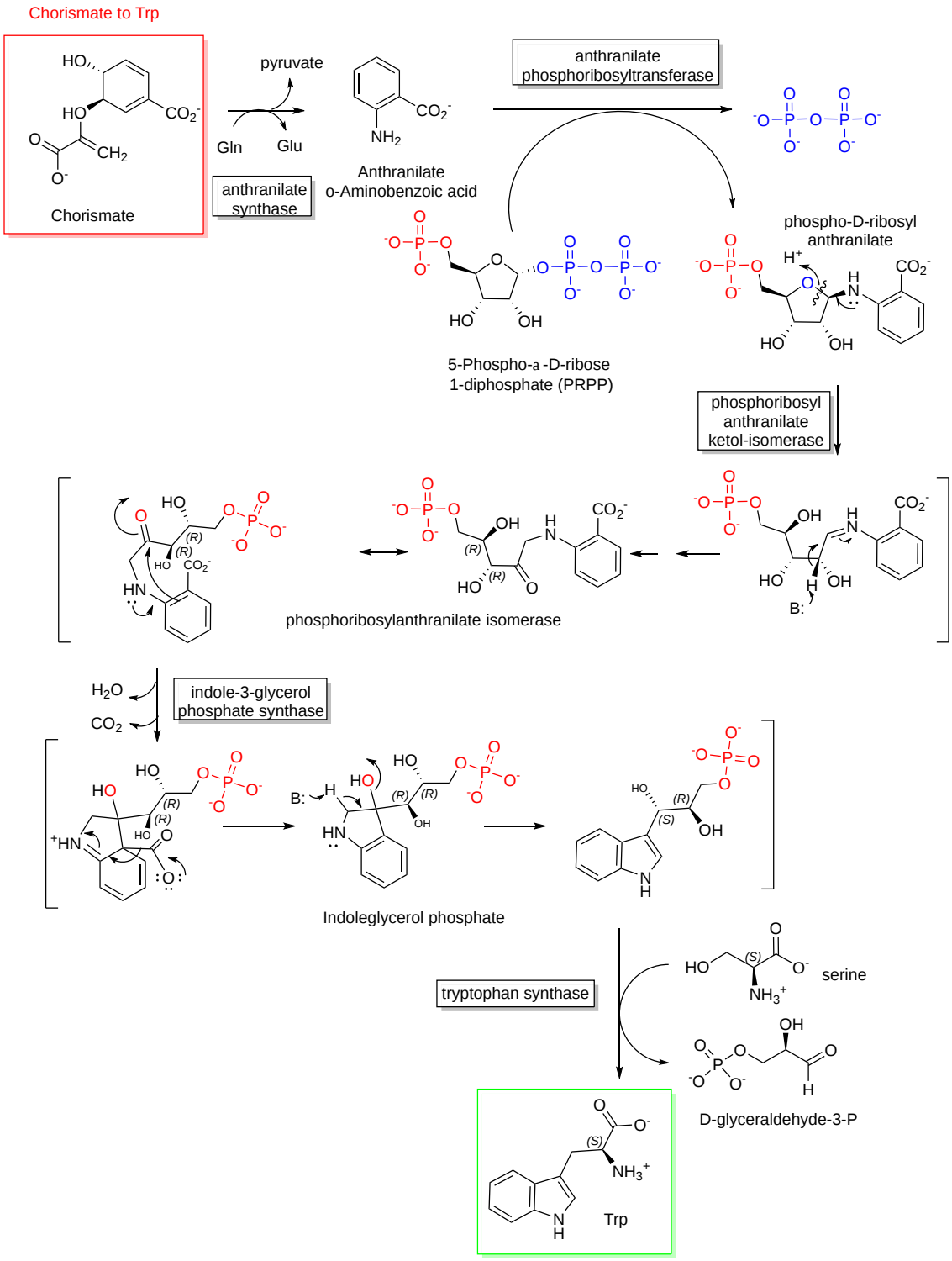


Figure (\PageIndex{5}\): Synthesis of the second half of the biosynthetic pathways for the aromatic amino acid tryptophan from chorismate

Chorismate to Phe and Tyr

The synthesis of the second half of the biosynthetic pathway for phenylalanine and tyrosine from chorismate is shown in Figure 22.2.6

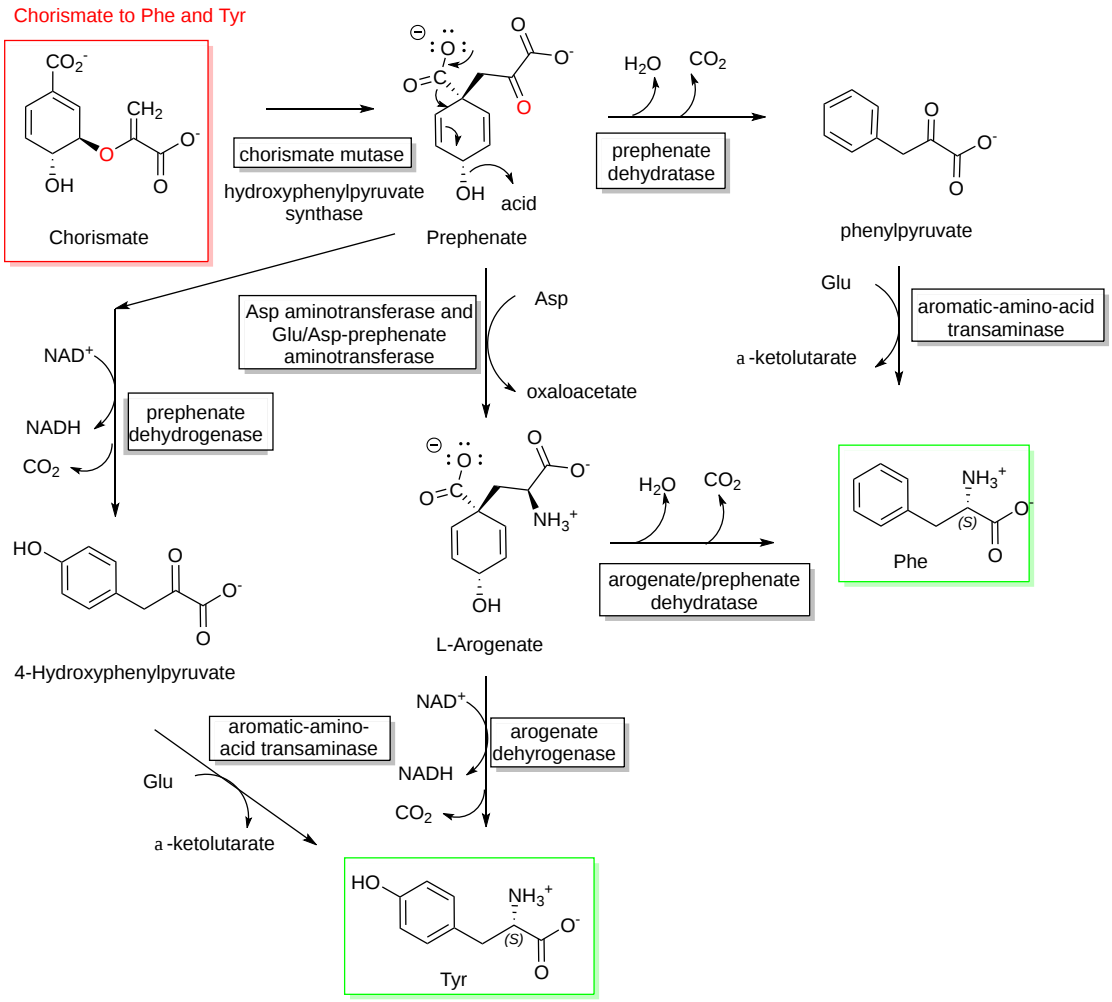


Figure 22.2.6: Synthesis of the second half of the biosynthetic pathway for phenylalanine and tyrosine from chorismate

22.2.2.4: FROM PYRUVATE: ALA, VAL, LEU, ILE

Ala can easily be synthesized from the alpha-keto acid pyruvate by a transamination reaction, so we will focus our attention on the others, the branched-chain nonpolar amino acids Val, Leu, and Ile.

The synthesis of valine, leucine, and isoleucine from pyruvate is shown in Figure 22.2.7.

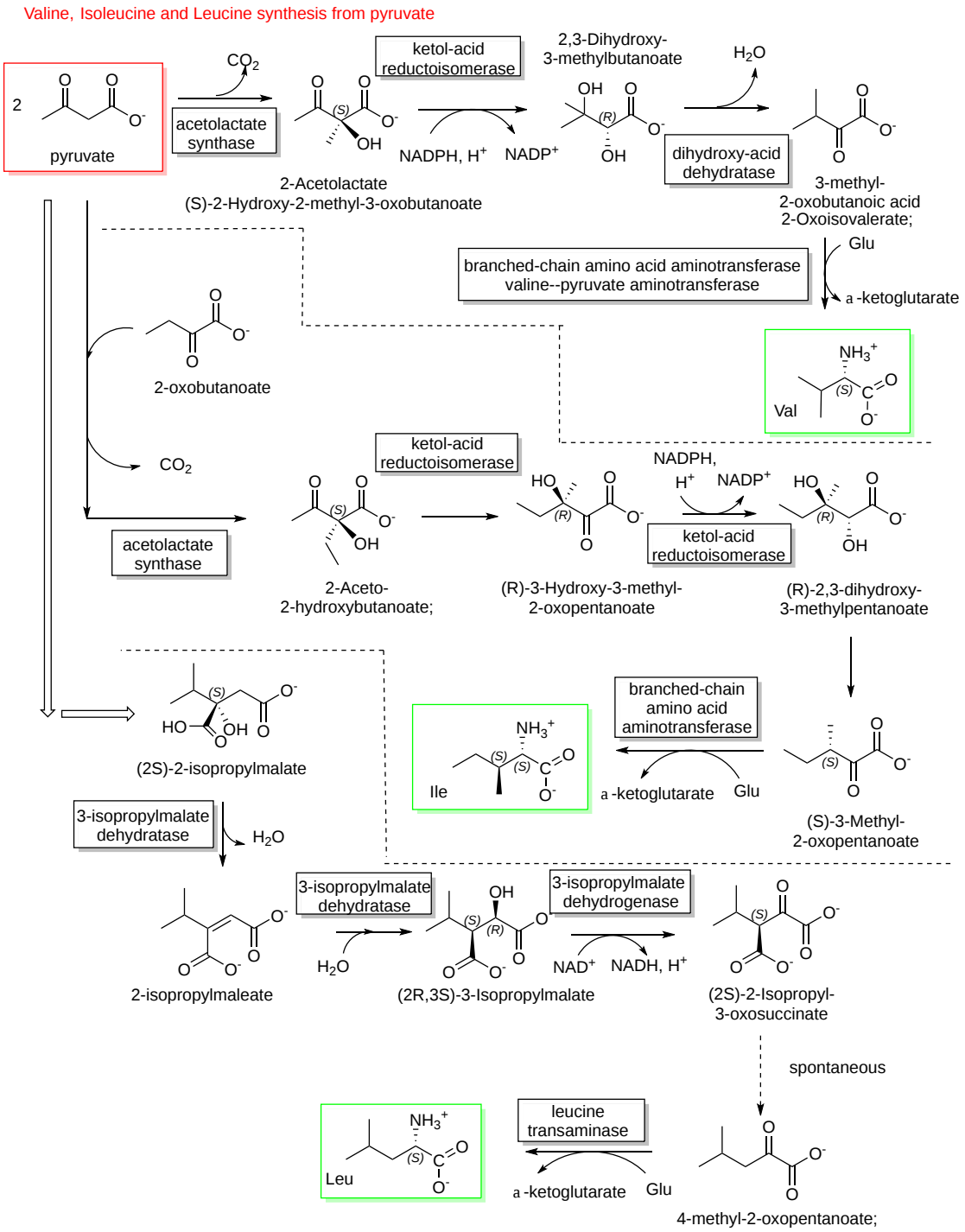


Figure 22.2.7: The synthesis of valine, leucine, and isoleucine from pyruvate

22.2.3: TCA INTERMEDIATES

22.2.3.1: FROM A-KETOGLUATARATE: GLU, GLN, PRO, ARG

Since amino acid metabolism is so complex, it's important to constantly review past learning. Figure 22.2.8 from section 18.2 shows the relationship among Glu, Gln, and keto acids.

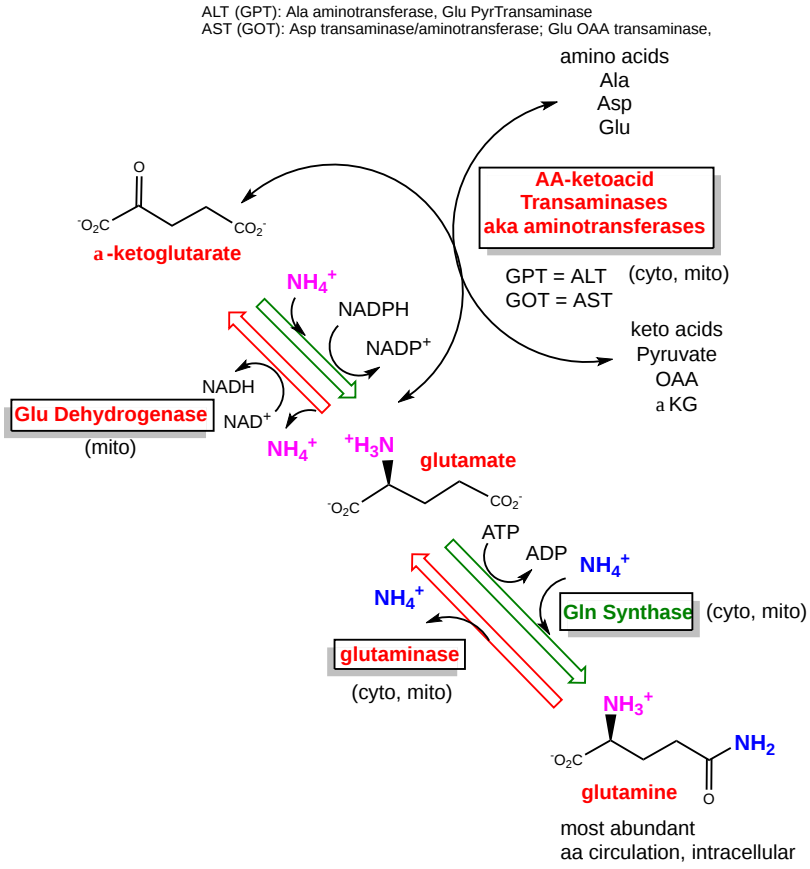


Figure 22.2.8: Glutamate and glutamine synthesis from α-ketoglutarate

As is evident from the figure, glutamic acid can be made directly through the transamination of α-ketoglutarate by an ammonia donor, while glutamine can be made by the action of glutamine synthase on glutamic acid.

Arginine is synthesized in the urea cycle as we have seen before. It can be made from α-ketoglutarate through the following sequential intermediates: N-acetylglutamate, N-acetylglutamate-phosphate, N-acetylglutamate-semialdehyde, N-acetylornithine to N-acetylcitruline. The is deacetylated and enters the urea cycle.

The pathway for conversion of α-ketoglutarate to proline is shown in Figure 22.2.9.

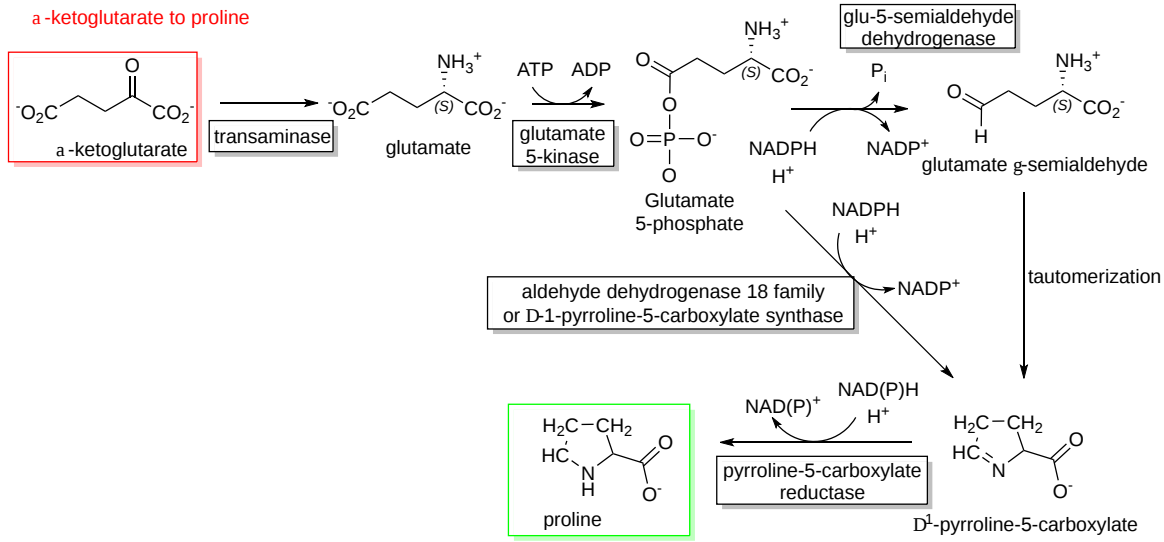


Figure 22.2.9: Conversion of α-ketoglutarate to proline

22.2.3.2: FROM OXALACETATE: ASP, ASN, MET, THR, LYS

22.2.3.2.1: OAA TO ASPARTATIC ACID

This is a simple transamination

22.2.3.2.2: ASPARTIC ACID TO ASPARAGINE

This is catalyzed by the enzyme Asparagine Synthase as shown in the reaction equation below:

$$\text{Aspartate} + \text{Glutamine} + \text{ATP} + H_2O \rightarrow \text{Asparagine} + \text{Glutamic Acids} + \text{AMP} + PP_i$$

22.2.3.2.3: ASPARTIC ACID TO LYSINE

There are two pathways.

- The diaminopimelic acid (DAP) pathway uses aspartate and pyruvate and forms diaminopimelic acid as an intermediate. It's found in bacteria, some fungi, and archaea and in plants.
- The aminoadipic acid (AAA) pathway uses α-ketoglutarate and acetyl-CoA and forms aminoadipic acid as an intermediate. It is used by fungi.,

Here we present just the synthesis of lysine from aspartate and pyruvate using the diaminopimelic acid DAP pathway. The pathway is shown in Figure 22.2.10.

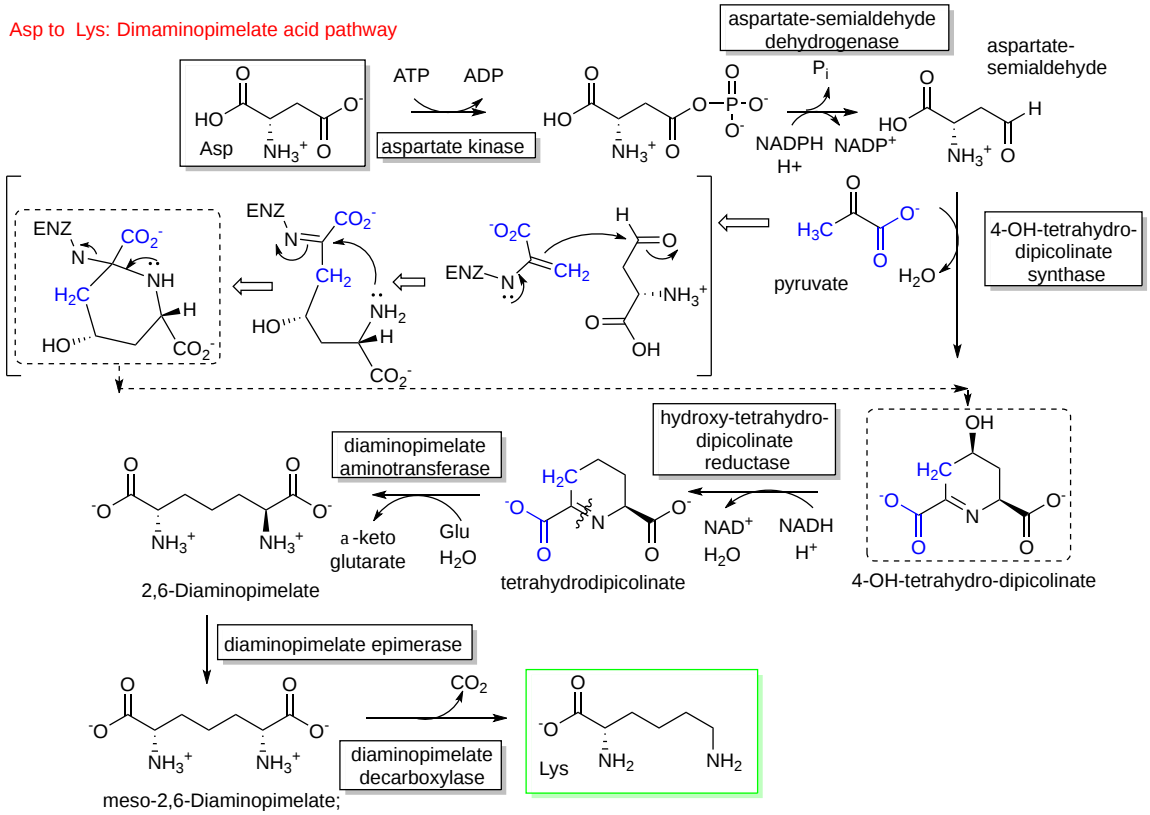


Figure 22.2.10: The synthesis of lysine from aspartic acid in the diaminopimelic acid DAP pathway

.

22.2.3.2.4: ASPARTIC ACID TO THREONINE

The conversion of aspartic acid to threonine is shown in Figure 22.2.11.

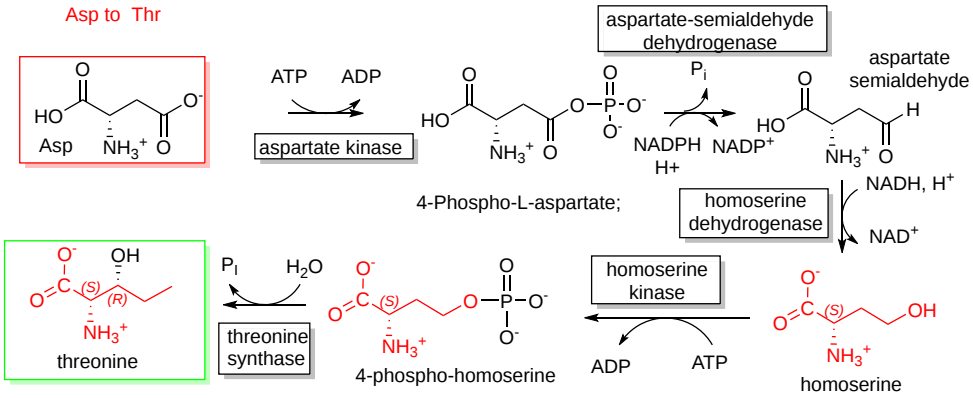


Figure 22.2.11: The conversion of aspartic acid to threonine

22.2.3.2.5: ASPARTIC ACID TO METHIONINE

The conversion of aspartic acid to methionine is shown in Figure 22.2.12.

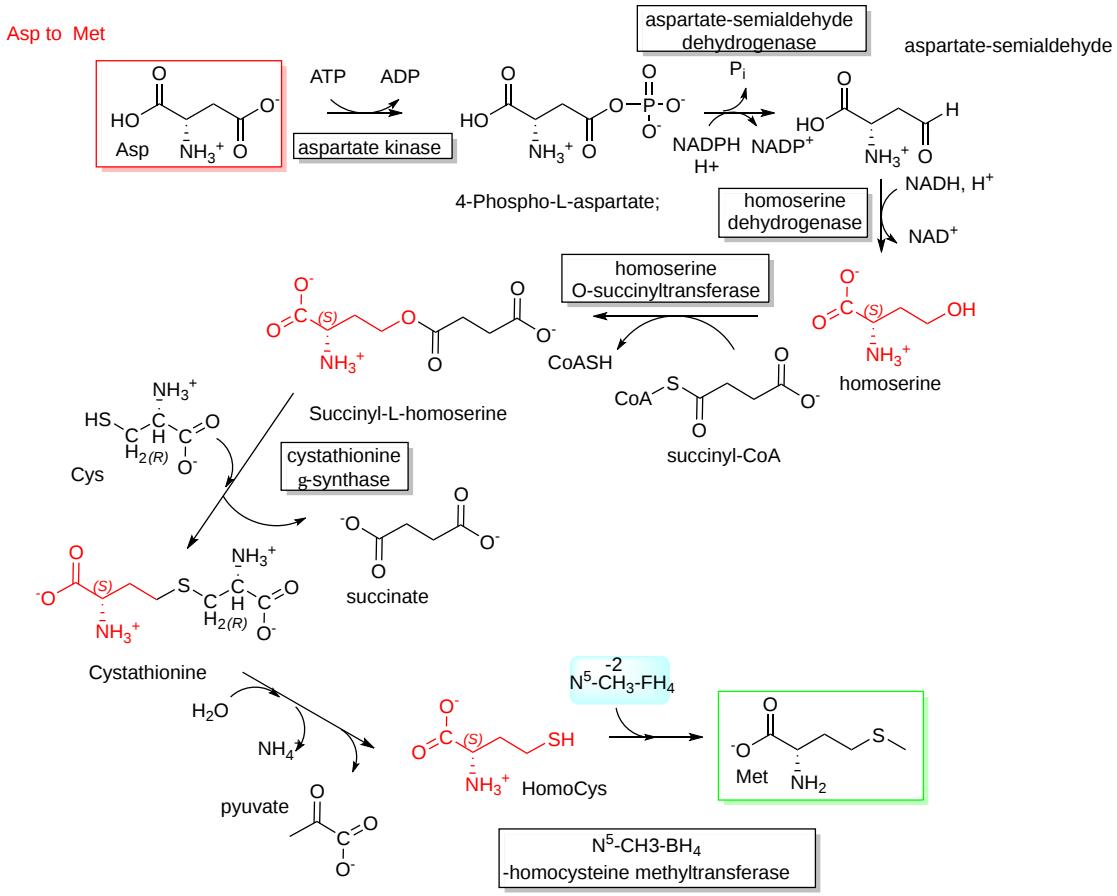


Figure 22.2.12: The conversion of aspartic acid to methionine

22.3: MOLECULES DERIVED FROM AMINO ACIDS

22.3.1: INTRODUCTION

Once made or ingested, amino acids have many metabolic fates. Of course, they are used for the synthesis of proteins. Aspartate and glutamate (and indirectly glutamine) can be converted to oxaloacetate and α-ketoglutarate, respectively, and used in the citric acid cycle for energy production. They can also be used for gluconeogenesis using mitochondrial and cytoplasmic enzymes. Branched-chain amino acids can be converted to acetyl-CoA and used in energy production or fat synthesis. A review summary of the use of amino acids in energy and biosynthetic metabolic pathways is shown in Figure 22.3.1.

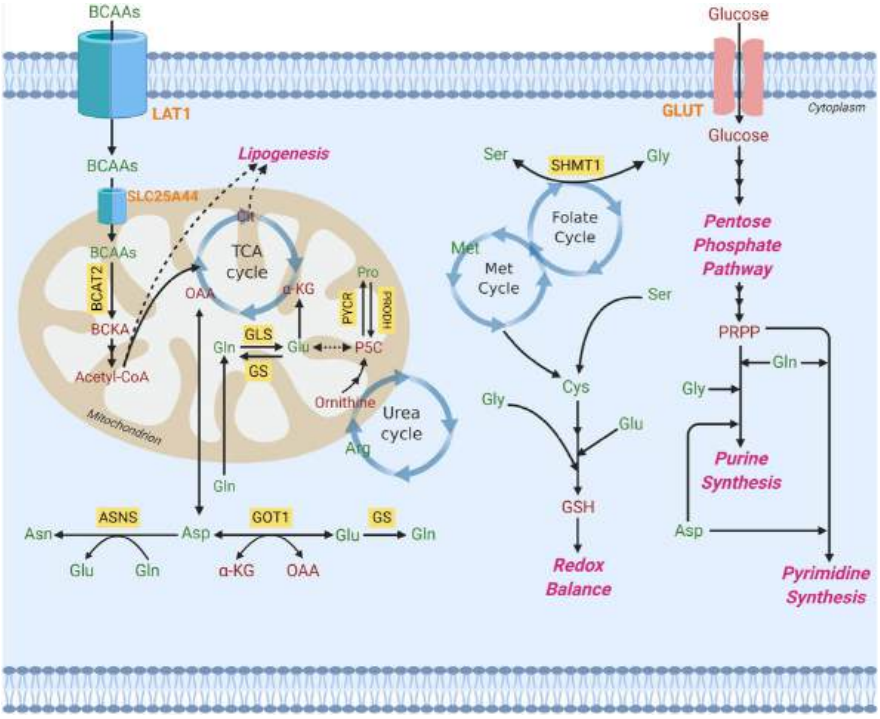


Figure 22.3.1: Review summary of the use of amino acid in energy and biosynthetic metabolic pathways. Lieu, E.L., Nguyen, T., Rhyne, S., *et al.* Amino acids in cancer. *Exp Mol Med* **52,** 15–30 (2020). https://doi.org/10.1038/s12276-020-0375-3. Creative Commons Attribution 4.0 International License, http://creativecommons.org/licenses/by/4.0/.

Amino acids are shown in green and other metabolites are in red. Orange represents transporters. Yellow boxes signify enzymes. Lesser known abbreviations for species include SHMT1 serine hydroxymethyltransferase, cytosolic, BCAT branched-chain amino acid transaminase, mitochondrial, BCAA branched-chain amino acid (valine, leucine, isoleucine), BCKA branched-chain ketoacid, GOT1 aspartate transaminase, cytosolic (AST), GLS glutaminase, GS glutamine synthetase (cytosolic and mitochondrial), ASNS asparagine synthetase, PRODH pyrroline-5-carboxylate dehydrogenase, PYCR pyrroline-5-carboxylate reductase, P5C pyrroline-5-carboxylate, GSH glutathione, PRPP phosphoribosyl pyrophosphate, LAT1 large-neutral amino acid transporter 1, SLC25A44 solute carrier family 25 member 44, GLUT glucose transporter,

Cancer cells from an increased need for fuels and biosynthetic intermediates. Both can come from amino acids as described previously. Glutamine is a key amino acid, especially if glucose is depleted as α-ketoglutarate (α-KG) and subsequently oxaloacetate (OAA generated from it powers the TCA cycle as fumarate, malate, and citrate are significantly increased. Hence it is both anaplerotic and a source of fuel. Similar increases in citrate occur in hypoxia. Aerobic glycolysis (Warburg effect) occurs in cancer cells, which show enhanced glucose uptake and conversion to lactate even in the presence of oxygen. This process can go so quickly that the amount of ATP produced in cancer cells from aerobic glycolysis can be similar to the from oxidative metabolism in the mitochondria, even though it is far less efficient. More information on cancer cell metabolism is found in Chapter 23.

In this chapter, we will discuss the conversion of amino acids to other molecules not directly involved in those metabolic pathways. We will focus on their use for the synthesis of polyamines, heme, and neurotransmitters in this chapter section. We won't discuss detailed mechanisms or structures for the proteins and enzymes involved in these pathways. In the next chapter section (22.4), we will present amino acids as substrates in the synthesis of pyrimidine and purine bases for nucleotides and nucleic acids.

22.3.2: POLYAMINE SYNTHESIS

If a non-quaternary amine has a single positive charge when protonated, a polyamine can have multiple positive charges. Hence they would be expected to bind to almost any negatively charged biomolecule, but especially those with multiple negative charges. These would include the polyanions RNA and DNA as well as proteins and lipid bilayers. They would then have the potential ability to regulate many features of cell life, including DNA replication and transcription, RNA translation, and a multitude of binding interactions. The question arises if these interactions are nonspecific, or specific, in which case they can be considered key regulators of cellular activity. Polyamine response elements have been found that regulate the transcription of genes including c-Myc and c-Jun. Polyamines have been shown tumor growth and aggressiveness.

The main biological polyamines include putrescine, spermine, and spermidine. Another is cadaverine. Given their names, you can surmise that they smell horrible. The synthesis of three polyamines from arginine and SAM is shown in Figure 22.3.2.

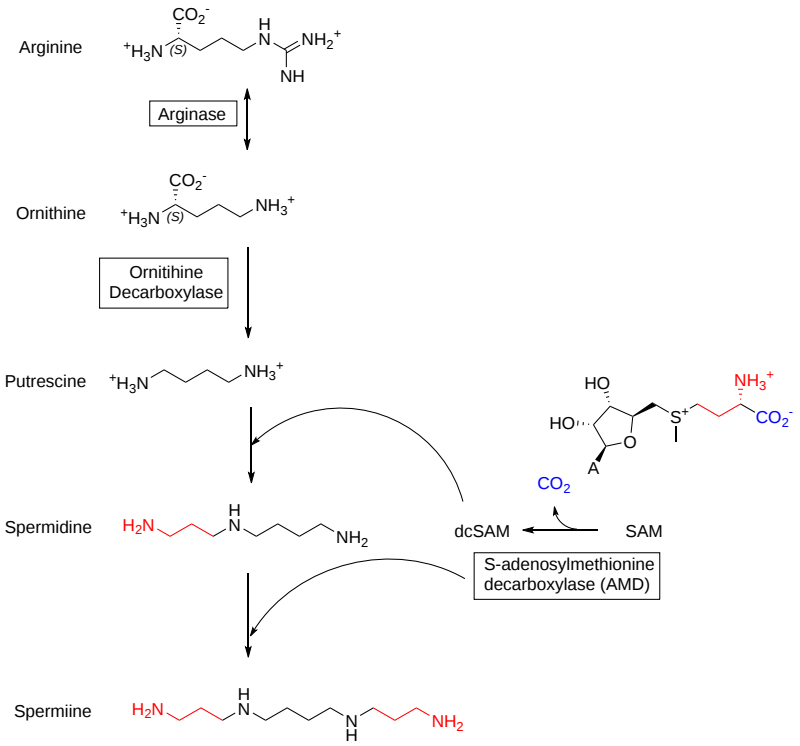


Figure 22.3.2: Polyamine synthesis form arginine and SAM

22.3.3: GLUTATHIONE SYNTHESIS AND REDOX BALANCE

Glutathione, γ-glutamylcysteinylglycine (GSH), is a chief regulator of the oxidation state of a cell. As a disulfide bond can be cleaved and hence reduced by the excess concentration of a thiol (sulfhydryl) like b-mercaptoethanol (which gets oxidized in the process), the free thiol on glutathione can act as a reducing agent in the cell. The production of reactive oxygen species (ROS) in normal but especially tumor cells, which have increased O_2 demand and use, is countered by the generation of an antioxidant defense state. This is characterized in part by increased levels of reductants such as NADPH but especially glutathione. It can react with H_2O_2 through the enzyme GSH peroxidase to form water and the oxidized disulfide form of GSH, GSSG. The GSSG is oxide back to GSH by glutathione reductase (GR) and NADPH. Figure 22.3.3 shows the synthesis of glutathione from glutamate, cysteine, and glycine.

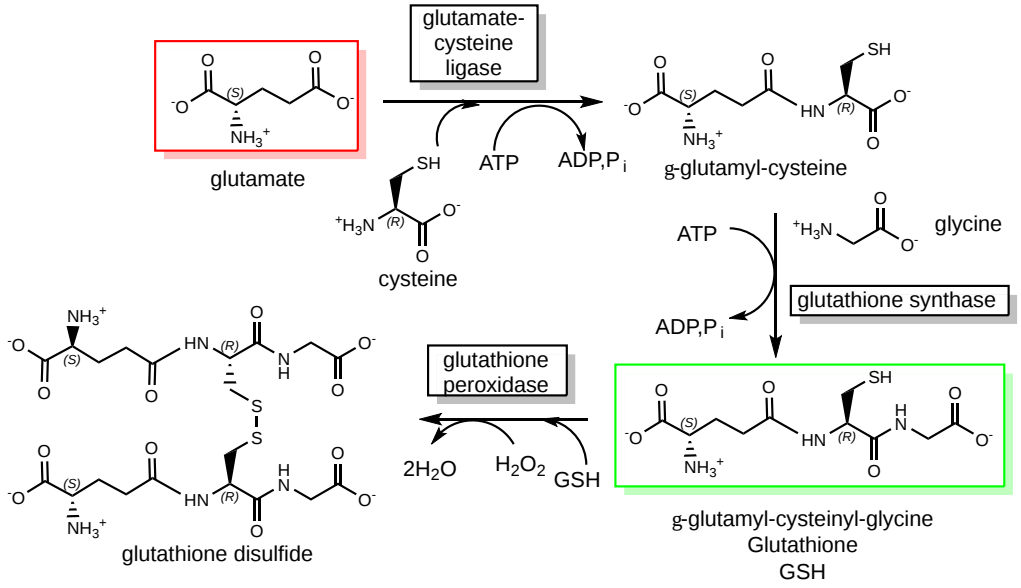


Figure 22.3.3: Synthesis of glutathione

NADPH is generated in the cell by the phosphopentose pathway metabolism of glucose and by malic enzyme. It can also be generated from the Ser-Glycine One Carbon Cycle (SGOT) that we saw in Chapter 18.4, which is shown again in Figure 22.3.4. Under appropriate conditions, this cycle can produce NADPH.

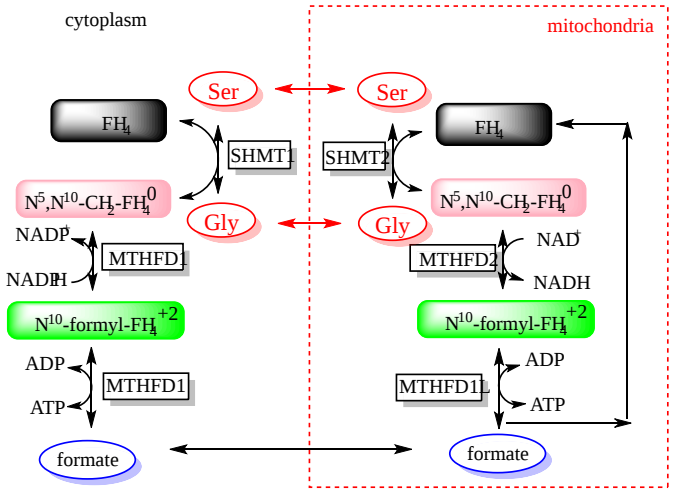


Figure 22.3.4: The Ser-Gly One Carbon (SGOC) Cycle

Serine Hydroxymethyltransferases (SHMTs) 2 is upregulated by HIF1α and helps maintain the NADPH/$NADP^+$ ratio. Given the connection between the SGOC and the methionine cycle through folate, a decrease in serine concentration leads to a decrease in GSH.

22.3.4: HEME BIOSYNTHESIS

This section is derived from Aminat S. Ogun; Neena V. Joy; Menogh Valentine. https://www.ncbi.nlm.nih.gov/books/NBK537329/. Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/)

Heme is a macrocyclic tetrapyrrole ring structure containing two nonpolar vinyl groups on one edge and two charge propionates on the other. It is extensively conjugated with 26 π electrons (4n+2 = 4(6)+2) so it is aromatic. The molecule without Fe^{2+} is called protoporphyrin IX and with a centrally-coordinate Fe^{2+}, it is called heme. The structures of both are shown in Figure 22.3.5.

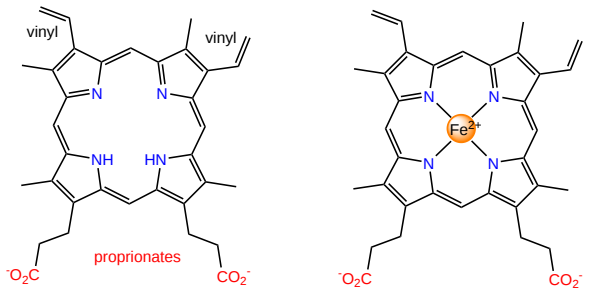


Figure 22.3.5: Structure of protoporphyrin IX and heme

It is found in oxygen-binding proteins and as substrates and cofactors for enzymes involved in electron transport. It is synthesized in the bone marrow and liver. Alternative forms of heme include heme b (in hemoglobin), heme a (cytochrome a), and heme c (cytochrome c).

Its synthesis, as expected given its macrocyclic structure, is complicated. The key enzyme in the pathway for regulation is 5'-Aminolevulinic acid synthase (ALA-S). Liver and bone express ALAS2 while ALAS1 is expressed in all tissues. The synthesis starts in the mitochondria and ends in the cytosol. The overall pathway for heme synthesis is shown in Figure 22.3.6.

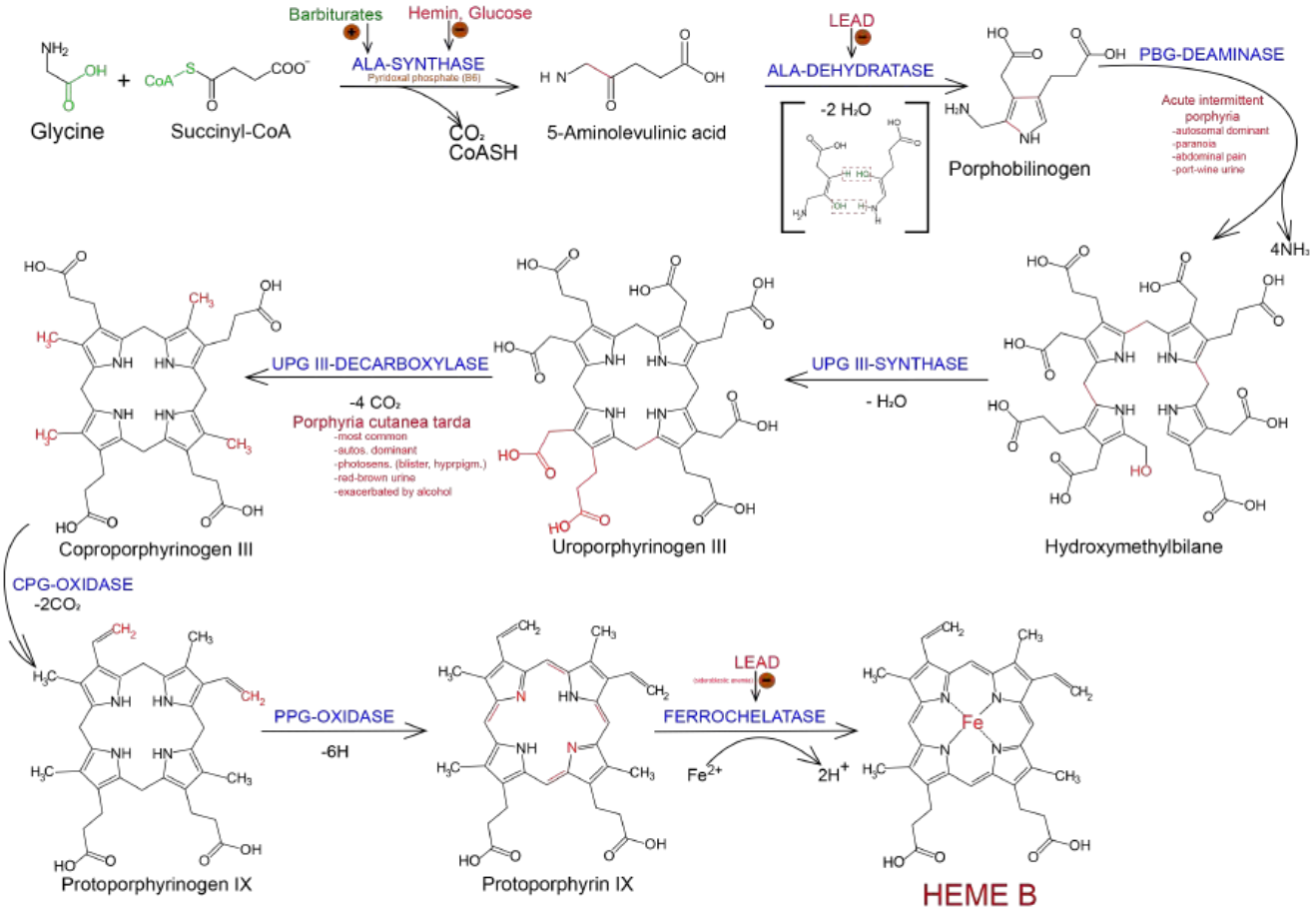


Figure 22.3.6: Heme biosynthetic pathway. Wikimedia Commonsile: Heme-Synthesis-Chemical-Details-Mirror.svg

5'-Aminolevulinic acid synthase (ALA-S), a pyridoxal phosphate-dependent enzyme, catalyzes the rate-limiting step in heme synthesis in the liver and erythroid cells. It is highly regulated There are two forms of ALA Synthase, ALAS1, and ALAS2. All cells express ALAS1 while only the liver and bone marrow expresses ALAS2. The gene for ALAS2 is on the X-chromosome. After the synthesis of ALA in the mitochondria, it moves into the cytoplasm for the remaining steps.

Figure 22.3.7 shows a likely mechanism for the first committed step, the production of ALA. This enzyme is used in the synthesis of all tetrapyrroles, including heme, chlorophyll, and cobalamin.

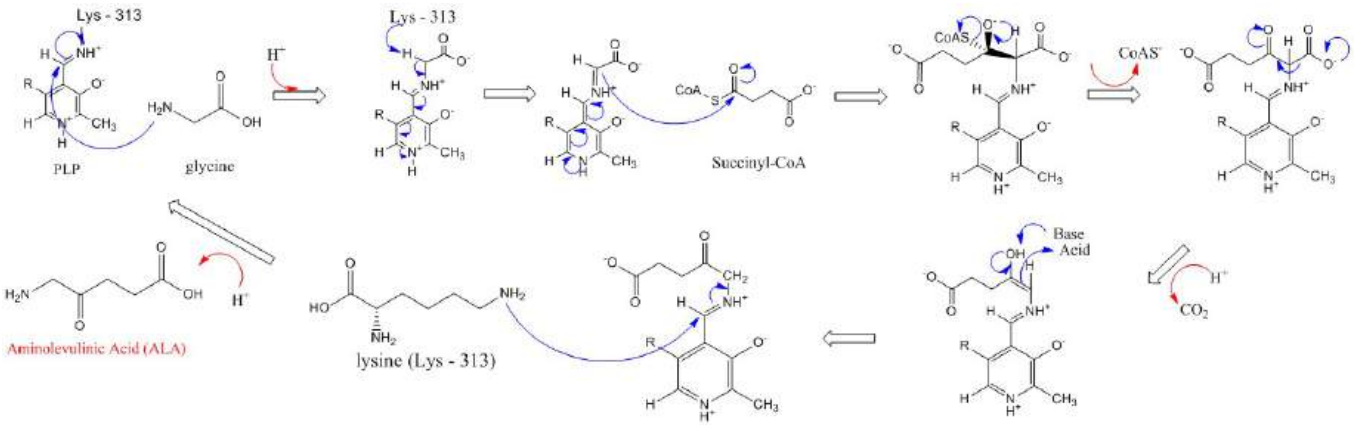


Figure 22.3.7: Mechanism for 5'-Aminolevulinic acid synthesis by ALAS (Wikipedia. https://en.Wikipedia.org/wiki/Aminol..._acid_synthase)

The pathway shown in Figure 6 above is called the C4 pathway and is found in mammals, fungi, and purple nonsulfur bacteria. A C5 pathway is found in most bacteria, all archaea, and plants. The biosynthetic pathway for heme synthesis in E. Coli is shown in Figure 22.3.8.

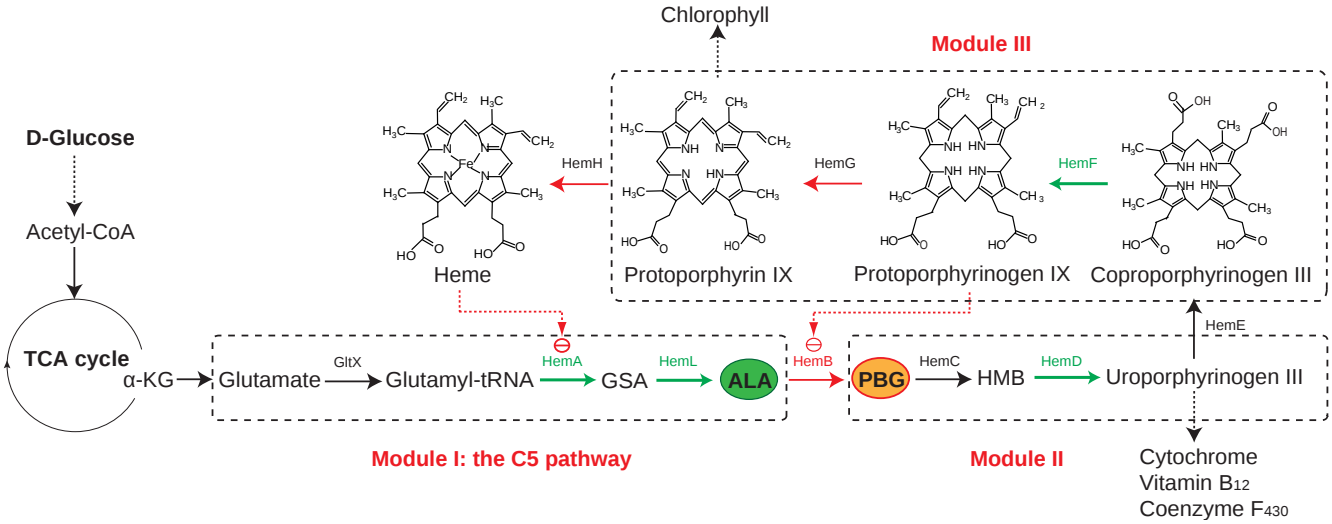


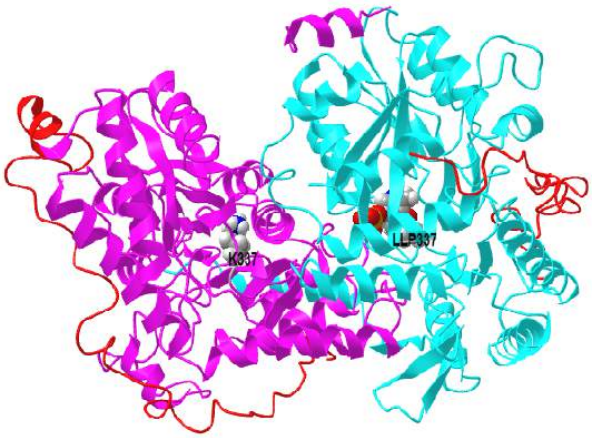
Figure 22.3.8: Heme pathway in *E. coli*. Zhang, J., Kang, Z., Chen, J. *et al.* Optimization of the heme biosynthesis pathway for the production of 5-aminolevulinic acid in *Escherichia coli*. *Sci Rep* **5,** 8584 (2015). https://doi.org/10.1038/srep08584. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/

The pathway is divided into three modules (module I, module II, and module III in the dotted box). The arrows in green and red represent the enzymes that are positive and negative to ALA accumulation, respectively. Dotted red arrows represent the feedback inhibition. α-KG: α-ketoglutarate, GSA: glutamate-1-semialdehyde, ALA: 5-aminolevulinic acid, PBG: porphobilinogen, HMB: hydroxymethylbilane, GltX: glutamyl-tRNA synthetase, HemA: glutamyl-tRNA reductase, HemL: glutamate-1-semialdehyde aminotransferase, HemB: 5-aminolevulinic acid dehydratase, HemC: porphobilinogen deaminase, HemD: uroporphyrinogen III synthase, HemE: uroporphyrinogen decarboxylase, HemF: coproporphyrinogen III oxidase, HemG: protoporphyrin oxidase, HemH: ferrochelatase.

In immature red blood cells (reticulocytes), heme increase globin protein synthesis. The hormone erythropoietin increases heme synthesis. In the liver, heme is part of cytochrome P450s. Increased concentration of drugs causes increases in ALAS1 to produce the cytochrome P450s to metabolize them. Also, low heme concentration increases ALAS1 transcription. Mutations in ALAS2 can lead to X-linked sideroblastic anemia from decreased heme production even as Fe^{2+} continues to enter the cell.

Yeast ALAS is a homodimer with PLP covalently attached through a Schiff base link to lysine 337 of just one of the monomers. The structures of a noncovalent complex of PLP with ALAS (pdb 5TXR) and the covalently bound one (5TXT) show large changes in the protein conformation. PLP when covalently attached reorders the active. A C-terminal extension not found in bacteria wraps around the dimer and binds near the active site and is important for activity. Mutations in the tail can result in human diseases.

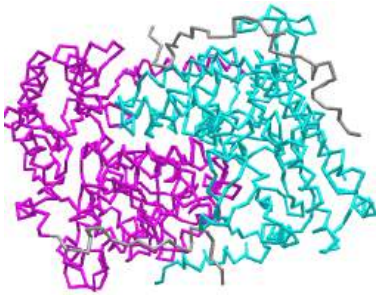
Figure 22.3.9 shows an interactive iCn3D model of the 5-aminolevulinic acid synthase with covalently attached PLP (5TXT).



NCBI iCn3D Figure 22.3.9: 5-aminolevulinic acid synthase with covalently attached PLP (5TXT). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?fTBbWuS3HPTP8uRm9

Lysine 337 (spacefill, CPK colors, labeled) in the A chain (magenta, no bound PLP) is shown. Lys 337 in the B chain (cyan) is covalently linked to PLP. The side chain of lysine 337 covalently attached to PLP is shown in spacefill, CPK colors, and labeled. The C-terminal extension (493–548) is shown as a red backbone chain. The very distal end of the extension is disordered and missing in the B chain (cyan).

Figure 22.3.10 shows an interactive iCn3D model of aligned 5-aminolevulinic acid synthase with free PLP (not covalently attached, 5TXR) and with covalently attached PLP (5TXT).



NCBI iCn3D Figure 22.3.11: Alignment of 5-aminolevulinic acid synthase with free PLP (5TXR) and with and with covalently attached PLP (5TXT). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...aqN4Qpi2pSLkD7

The 5TXT structure contains two molecules of a stabilizing molecule shown in stick form, which you can ignore. The A chain is shown in magenta and the B chain is in cyan. Press "a" to toggle back and forth between the structures. The C-terminal extension is missing from the figure.

Figure 22.3.11 shows another view of heme synthesis which emphasizes the role of mitochondrial and cytoplasmic enzymes.

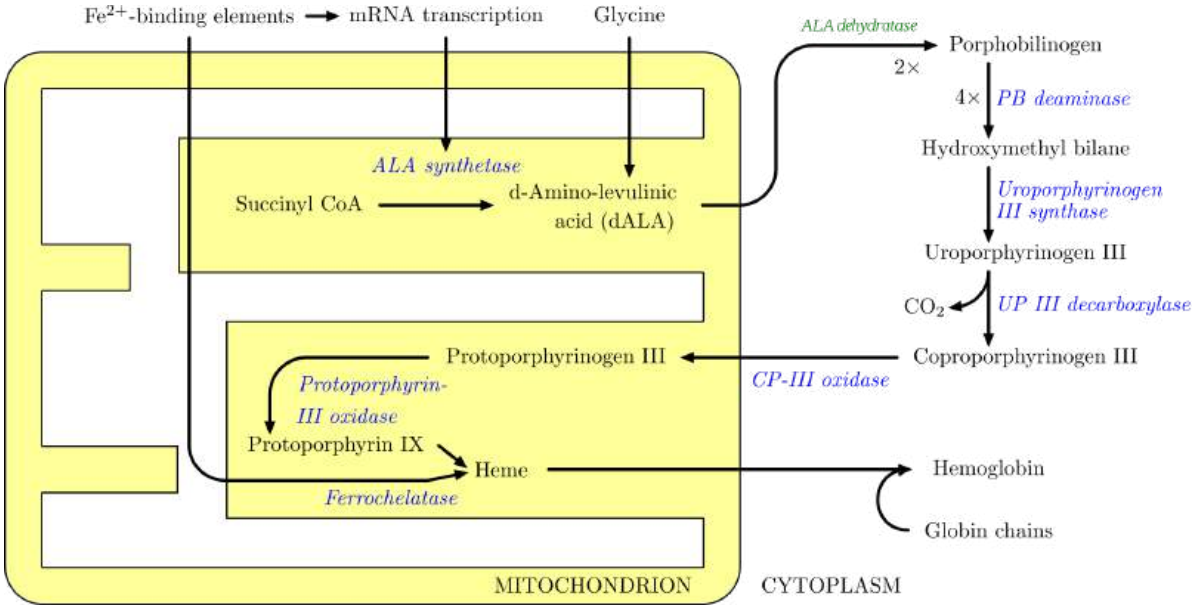


Figure 22.3.11: Mitochondrial and cytosolic contributions to heme synthesis. https://commons.wikimedia.org/wiki/F..._synthesis.png

22.3.5: NEUROTRANSMITTERS

The section below is modified from Manorama Patri. **Synaptic Transmission and Amino Acid Neurotransmitters.** DOI: 10.5772/intechopen.82121. https://www.intechopen.com/books/neu...rotransmitters. Creative Commons Attribution 3.0 License,

There are three major categories of amino acids and their derivatives act as neurotransmitters are:

1. Amino acids: The neurotransmitters of this group are involved in fast synaptic transmission and are inhibitory and excitatory in action (primarily glutamic acid, GABA, aspartic acid, and glycine).
2. Amines: Amines are modified amino acids such as biogenic amines, e.g., catecholamines. The neurotransmitters of this group involve in slow synaptic transmission and are inhibitory and excitatory in action (noradrenaline, adrenaline, dopamine, serotonin, and histamine).
3. Others: The ones which do not fit in any of these categories (acetylcholine and nitric oxide). Amino acids are among the most abundant of all neurotransmitters present within the central nervous system (CNS).

Amino acid transmitters provide the majority of excitatory and inhibitory neurotransmission in the nervous system. Amino acids used for synaptic transmission are compartmentalized (e.g., glutamate, compartmentalized from metabolic glutamate used for protein synthesis by packaging the transmitter into synaptic vesicles for subsequent Ca^{2+}-dependent release). Amino acid neurotransmitters are all products of intermediary metabolism except GABA. Unlike all the other amino acid neurotransmitters, GABA is not used in protein synthesis and is produced by an enzyme (glutamic acid decarboxylase; GAD) uniquely located in neurons.

Here is some more specific information:

- Glutamate: Glutamate is used at the great majority of fast excitatory synapses in the brain and spinal cord. Glutamate binds to glutamate receptors of which there are many subtypes based on other molecules (some amino acid derivatives) that can bind to them. These other molecules include N-methyl-D-aspartate (NMDA), α-amino-3-hydroxy-5-methyl-4-isoxazolepropionate (AMPA) kainate, and quisqualate.
- Aspartate: Aspartate is the most abundant excitatory neurotransmitter in the CNS. Like glycine, aspartate is primarily localized to the ventral spinal cord. Note that the two major excitatory neurotransmitters both have carboxylic acid side chains.
- Gamma-aminobutyric acid (GABA): GABA, which is not one of the canonical amino acids used in protein biosynthesis, is the most ubiquitous inhibitory neurotransmitter in the brain.
- Glycine: lycine receptors are ligand-gated ion channels that increase Cl^- influx and hence are generally inhibitory. Hydroxymethyl transferase converts the amino acid serine to glycine. Glycine has been found to play a role in the functional modulation of NMDA receptors

The pathways for the synthesis of amino acid-derived bioactive amines and neurotransmitters are shown in Figure 22.3.12.

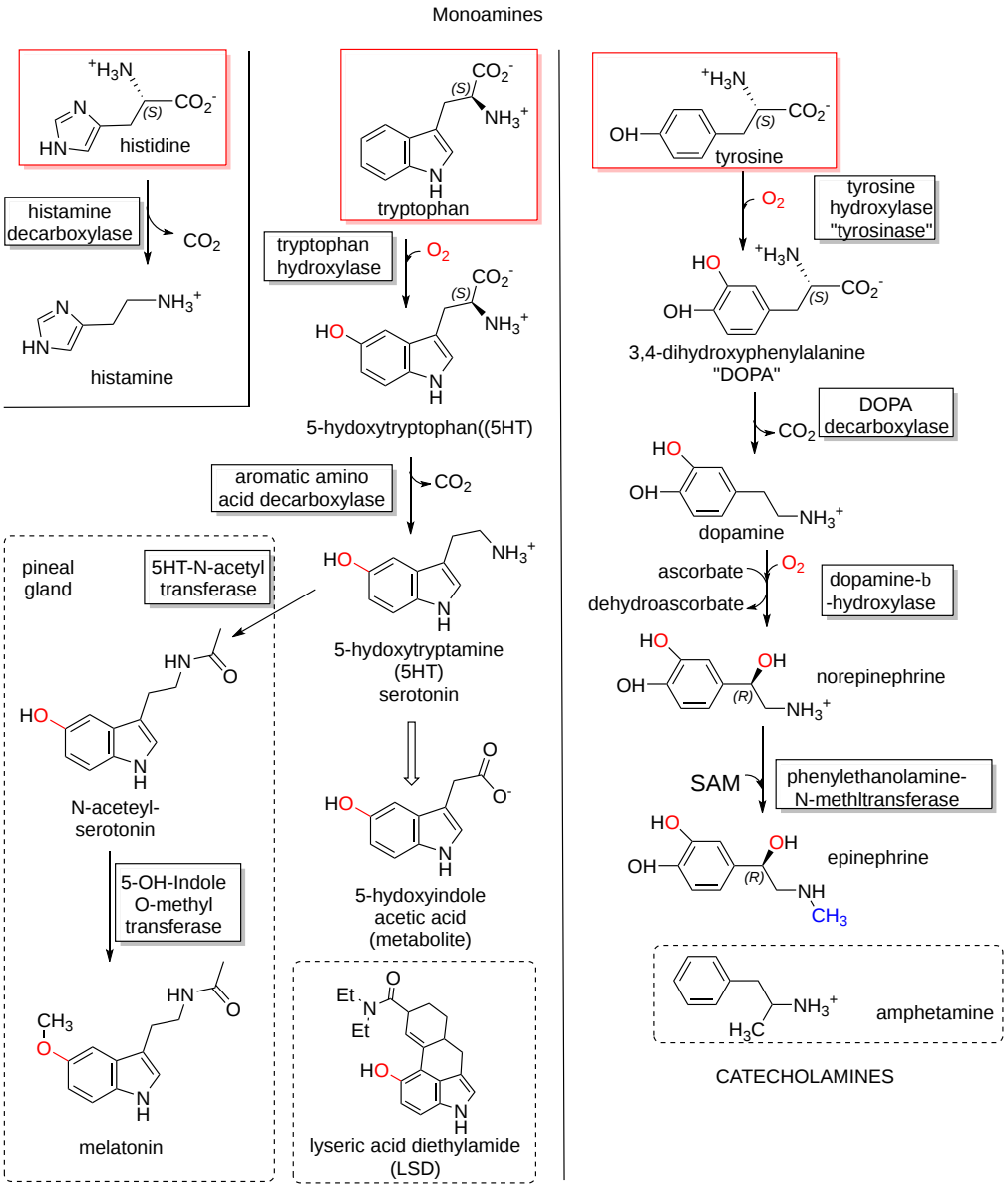


Figure 22.3.12: Pathways for the synthesis of amino acid-derived bioactive amines and neurotransmitters

Note the structural similarity of the psychotropic and hallucinogenic drug LSD to serotonin (5HT), amphetamines to norepinephrine and epinephrine, and melatonin (a substance some take as a sleeping aid and which forms in the dark at night in brains. The name **catecholamines** derive from the common name of the 1,2-dihydroxybenzene group (catechol).

The first and rate-limiting step in catecholamine synthesis is catalyzed by **tyrosine hydroxylase (TH).** It has no heme but it has an Fe^{2+} and tetrahydrobiopterin as a cofactor used in the synthesis of dihydroxyphenylalanine (DOPA). Tyrosine hydroxylase is rate-limiting for the synthesis of all three transmitters.

The enzyme is inhibited by catecholamines including dopamine, a downstream product, and is activated by phosphorylation on serine 40. The structures of TH in the absence of dopamine and the pSer40 state are known. The protein is a tetramer with a regulatory domain (dimer) and catalytic domain (also a dimer) separated by 15 Å.

The mammalian TH is a member of the aromatic amino acid hydroxylases (AAAHs) which are mainly found as homotetramers. Each subunit has 3 domains:

- The N-terminal **regulatory domain (RD)** that has an unstructured variable-length section followed by an ACT (aspartate kinase-chorismate mutase-TyrA) domain. The N-terminal tail contains serine 40 which on phosphorylation relieves the inhibition when dopamine is bound to the catalytic domain.
- a central **catalytic domain (CD)** that binds Fe^{2+}, aromatic amino acid substrates, and the tetrahydrobiopterin cofactor
- C-terminal oligomerization domain (OD) which leads to dimer and tetramer formation.

Figure 22.3.13 shows a potential model that illustrates dopamine (DA)-mediated feedback inhibition and its regulation by serine 40 phosphorylation through the interaction of the N-terminal tail of the regulatory domain (RD) with the catalytic domain. All forms containing bound dopamine (yellow star) are inactive.

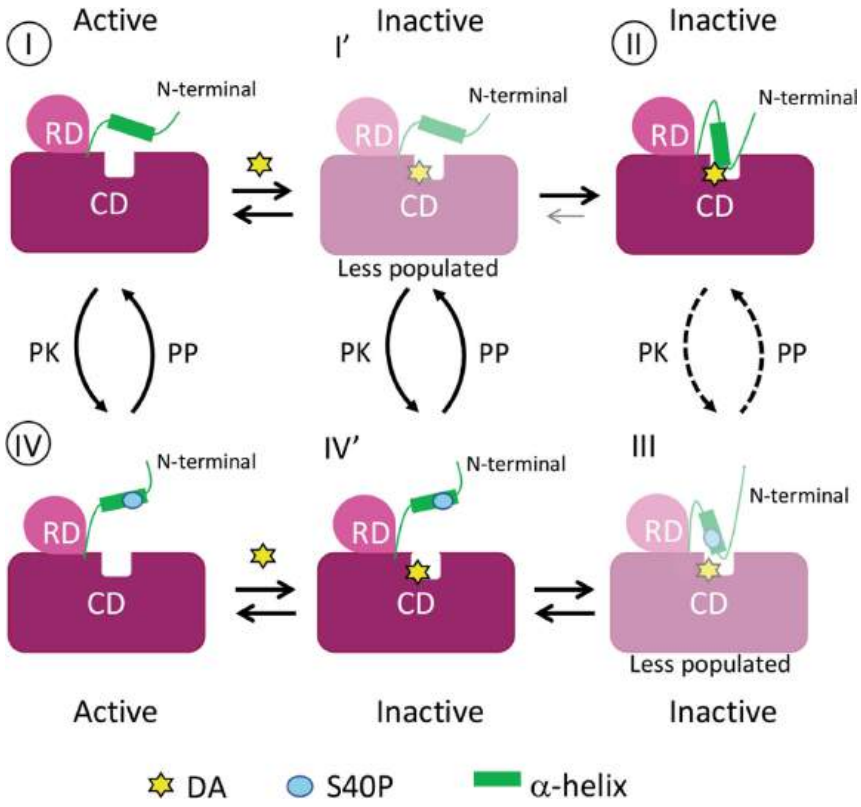


Figure 22.3.13: Cartoon model of DA-mediated feedback inhibition and its regulation by S40 phosphorylation. Bueno-Carrasco, M.T., Cuéllar, J., Flydal, M.I. *et al.* Structural mechanism for tyrosine hydroxylase inhibition by dopamine and reactivation by Ser40 phosphorylation. *Nat Commun* **13,** 74 (2022). https://doi.org/10.1038/s41467-021-27657-y. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

In the active, apo, and non-phospho states, the 39−58 α-helix of the N-terminal regulatory domain of TH is detached from the main structure (I, apo-TH). The feedback inhibitor DA binds to the TH active site, most likely in the open conformation (I′, TH(DA)). DA-binding favors the interaction of the N-terminal α-helix with the same binding site, which blocks DA exit and contributes to the high-affinity binding and strong inhibition of TH activity (II, TH(DA)). Protein Kinase (PK) phosphorylation of S40 in TH(DA), leads to state III (THS40p(DA)), prompting the detachment of the α-helix from the TH active site (IV′), which opens up for DA-dissociation and activation (IV, THS40p). PKs and protein phosphatase(s) (PP) control the transition between THS40p and unphosphorylated TH for both DA bound (I′ ↔ IV′ and II ↔ III) and apo-TH (I ↔ IV). States I′ and III are expected to be only transiently populated during DA binding as states II and IV' will be more stable. Hence states I′ and III states are faded. S40 is also expected to be less accessible in state II than in state I, which is indicated by stippled lines for phosphorylation of TH in state II. The case for dephosphorylation is not known, but it could be expected that state III is a poorer substrate for PP than the open states IV′ and IV. The dephosphorylation reaction III → II is therefore also stippled. The states where we provide structural details in this work (I, II, and IV) are marked with circles.

Figure 22.3.14 shows a model of the TH active site changes on phosphorylation of serine 40 using structural and molecular dynamics approaches that led to the cartoon model above.

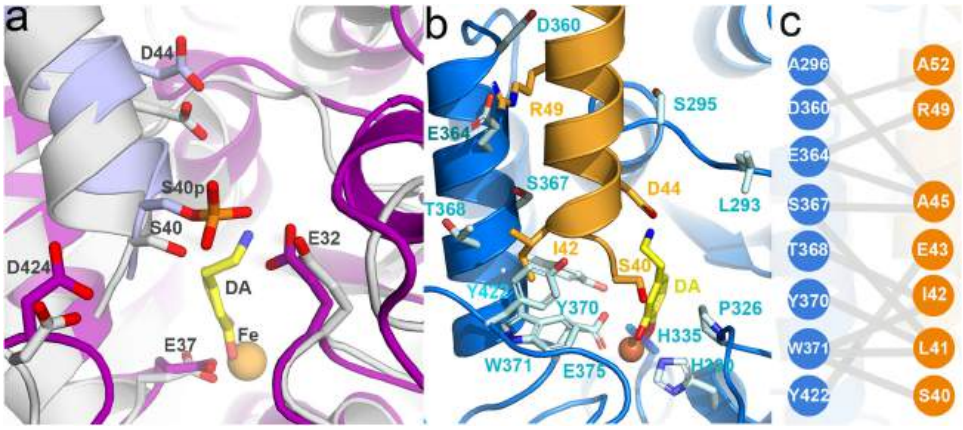


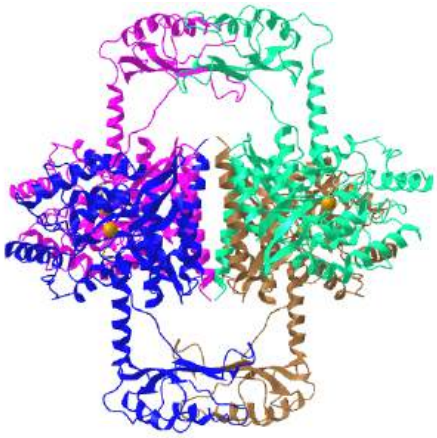
Figure 22.3.14: **Modeling of the TH active site.**

Panel (a) shows models demonstrating the effect of serine 40 phosphorylation on the interaction of the N-terminal α-helix with bound dopamine (DA). Representative conformations from the last 50 ns of a 500 ns MD simulations for TH(DA) (grey ribbon) and pS40-TH(DA) (light blue ribbon) are shown. The resulting structures show a slight shift of the N-terminal α-helix upon phosphorylation, most probably due to electrostatic repulsion between the phosphate and E325, E375, and D424.

Panel (b) shows a detailed view of the atomic model of the TH(DA) active site. (left) The N-terminal α-helix (orange), establishes connections with the adjacent helix D360-E375 and with residues of the 290–297 and 420–429 loops (blue, right).

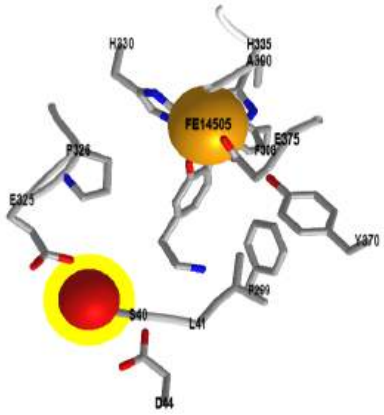
Pane (c) shows a cartoon depicting the interactions established between residues of the N-terminal α-helix that enters the active site, and residues of adjacent regions.

Figure \(\PageIndex{15}\) below shows an interactive iCn3D model of the full-length tyrosine hydroxylase in complex with dopamine (residues 40-497) in which the regulatory domain (residues 40-165) has been included only with the backbone atoms (6zvp)



NCBI iCn3D Figure 22.3.15: Tyrosine hydroxylase dopamine complex (6zvp). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...hu6vJn22eDkWw7

Finally, Figure 22.3.16 shows an interactive iCn3D model of the active site of tyrosine hydroxylase in complex with dopamine (6zvp)



NCBI iCn3D Figure 22.3.16: Active site of tyrosine hydroxylase dopamine complex (6zvp). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?sx5obQnzojmKXuLS9

The side chains binding the active site Fe^{2+} and the interaction of Fe^{2+} with dopamine (LDP) are shown in sticks and labeled. The oxygen of serine 40 is shown as a red sphere. Dopamine is the molecule containing the 1,2-dihydroxybenzene.

Figure 22.3.17 shows a final summary presentation of the conversion of phenylalanine, tyrosine, and tryptophan to neurotransmitters.

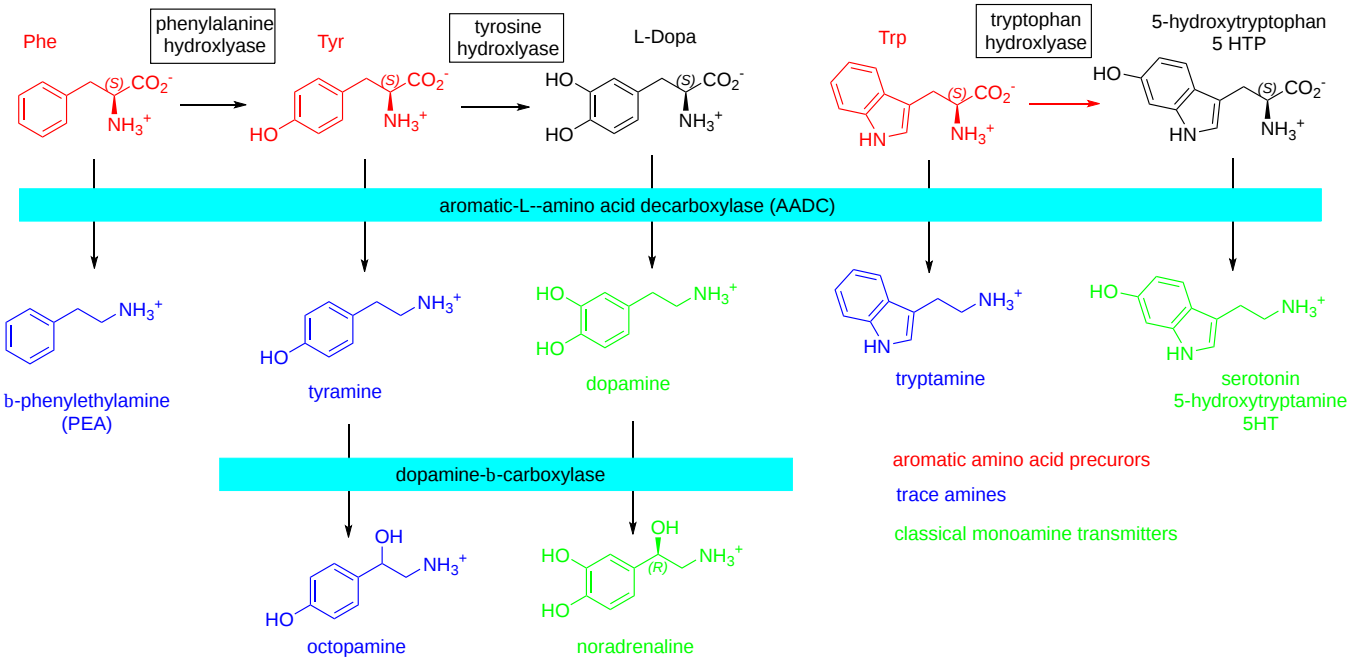


Figure 22.3.17: Comparison of monoamine synthesis pathways. Adapted from Hochman, Shawn. (2015). Neural Regeneration Research. 10. 10.4103/1673-5374.169625. Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International

22.4: BIOSYNTHESIS AND DEGRADATION OF NUCLEOTIDES

22.4.1: INTRODUCTION

We conclude our exploration of metabolic pathways with the biosynthesis and breakdown of nucleotides, the monomers that comprise nucleic acids. We can't also forget the important role of ATP as the universal carrier of biological free energy, as well as the nucleotides involved in signal transduction (GTP in heterotrimeric G proteins, small G proteins, and ATP as substrate in protein phosphorylations by kinases). As with the other later sections on metabolism, we won't focus much on detailed reaction mechanisms or enzyme structures, with one exception, the enzyme that converts nucleotides to deoxynucleotides.

Nucleotide synthesis is often included in chapters on amino acid metabolism as almost every atom in the purine and pyrimidine ring derives from them as shown in Figure 22.4.1.

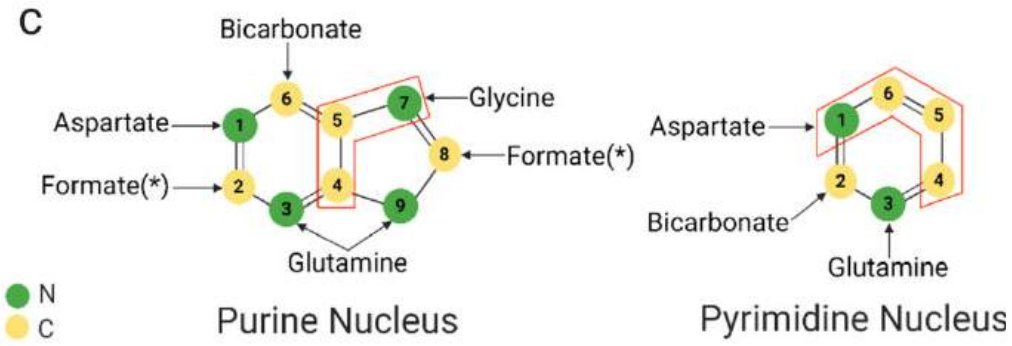


Figure 22.4.1: Source of atoms in nucleotide bases. Lieu, E.L., Nguyen, T., Rhyne, S., *et al.* Amino acids in cancer. *Exp Mol Med* **52,** 15–30 (2020). https://doi.org/10.1038/s12276-020-0375-3. Creative Commons Attribution 4.0 International License, http://creativecommons.org/licenses/by/4.0/.

For purines, glutamine and aspartate provide the nitrogen for the nucleotide's rings. They also provide the NH_3s for the ring substituents (glutamine for adenine and aspartate for guanine). Glycine and formate provide the carbon atoms for the rings. The one carbon molecule formate, which derives from glycine, is also added to purine rings. Glycine provides a one-carbon unit indirectly through the main carrier of activated one-carbon units, 5,10-meTHF, which is converted to formate through 10-formyl THF.

Pyrimidines are much smaller and their synthetic pathway reflects that. Instead of being synthesized as nucleobases as in the case of purines, they are made as ribonucleotides as they are linked to phosphoribosyl pyrophosphate (PRPP). Glutamine and aspartate again provide the ring C and N atoms. The one-carbon unit derives from serine-to-glycine conversion. A methyl group from the activated 1C donor, 5,10-meTHF, is added to dUMP to make dTMP.

22.4.2: PURINE SYNTHESIS

The material below derives from De Vitto, H.; Arachchige, D.B.; Richardson, B.C.; French, J.B. The Intersection of Purine and Mitochondrial Metabolism in Cancer. Cells 2021, 10, 2603. https://doi.org/10.3390/cells10102603. **Creative Commons Attribution License**

Mammals have two pathways for purine synthesis, a de novo pathway and a salvage pathway to recycle nucleotide bases. The salvage pathway is typically sufficient as purine bases come from nucleic acid breakdown. The resulting free bases (adenine, guanine, and hypoxanthine are connected to **phosphoribosyl pyrophosphate (PRPP)** to form nucleoside monophosphates (NMP) using either **adenine phosphoribosyltransferase (*APRT*)** to form AMP or hypoxanthine-guanine phosphoribosyltransferase (*HGRT*) to form IMP and GMP. PRPP is a substrate in both the salvage and de novo pathways. The overall de novo and salvage pathways for purine synthesis are described in detail in Figure 22.4.2.

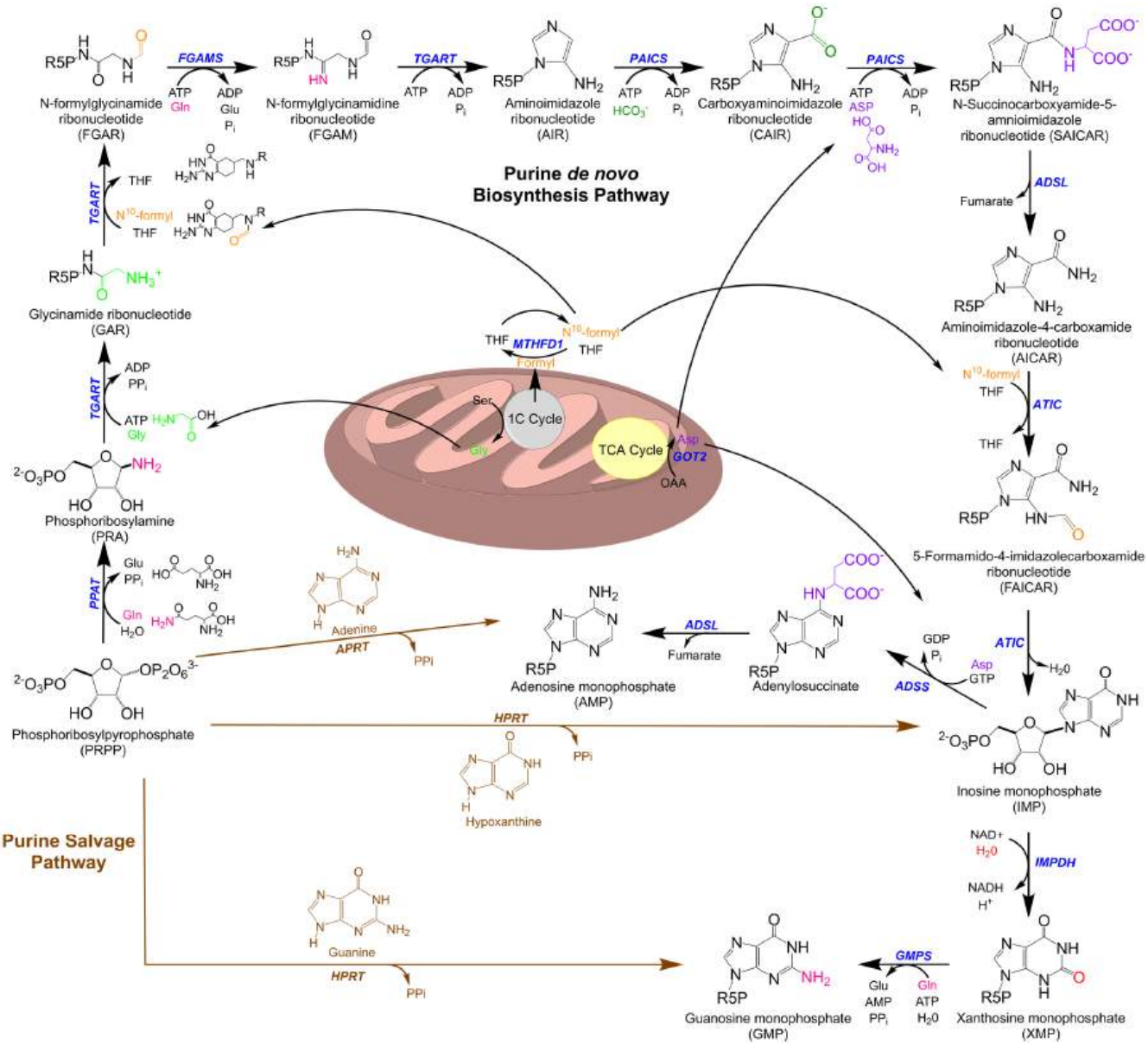


Figure 22.4.2: Purine metabolic pathways. De Vitto, H.; Arachchige, D.B.; Richardson, B.C.; French, J.B. The Intersection of Purine and Mitochondrial Metabolism in Cancer. Cells 2021, 10, 2603. https://doi.org/10.3390/cells10102603. **Creative Commons Attribution License**

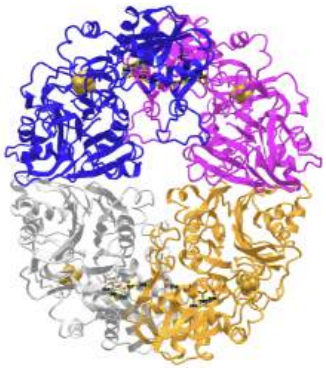
The conserved de novo biosynthesis pathway to generate IMP consists of 10 chemical steps catalyzed by 6 gene products in humans. These include the trifunctional enzyme *TGART*, composed of GAR synthetase (GARS), GAR transformylase (GARTfase), and AIR synthetase (AIRS) domains; the bifunctional enzymes *PAICS*, composed of CAIR synthetase/AIR carboxylase (CAIRS) and SAICAR synthetase (SAICARS), and *ATIC*, composed of AICAR transformylase (AICART) and IMP cyclohydrolase (*IMPCH*); and three monofunctional enzymes, phosphoribosyl amidotransferase (*PPAT*), formyl glycin amidine ribonucleotide synthetase (*FGAMS*), and adenylosuccinate lyase (*ADSL*). Downstream IMP is converted to (1) GMP through stepwise reactions of IMP dehydrogenase (*IMPDH*) followed by GMP synthetase (*GMPS*) and (2) AMP via adenylosuccinate synthetase (*ADSS*) followed by *ADSL*. The salvage pathway requires PRPP to generate IMP and GMP through one-step reactions mediated by hypoxanthine phosphoribosyltransferase (*HPRT*) utilizing hypoxanthine and guanine bases. AMP is generated by adenine phosphoribosyltransferase (*APRT*) utilizing adenine base and PRPP as substrates. Mitochondria supply precursors for purine de novo biosynthesis including glycine, N^{10}-formyl THF, and aspartic acid through their one-carbon cycle (1C cycle) and tricarboxylic acid cycle (TCA).

The de novo pathway kicks in when there is high demand for purines. Six enzymes are required for the 10-step pathway. Three of these are multifunctional enzymes catalyzing multiple steps in the pathway, comprising the two bifunctional enzymes phosphoribosylaminoimidazole carboxylase (*PAICS)* and AICAR transformylase/inosine monophosphate cyclohydrolase (*ATIC)* and the trifunctional enzyme glycinamide

ribonucleotide transformylase (*TGART*). When active, the pathway is limited both by substrate availability and by the reaction rate of its initial step, the conversion of PRPP to phosphoribosylamine (PRA) by **phosphoribosylpyrophosphate amidotransferase (*PPAT*)**. The final product of the de novo biosynthesis pathway, IMP is a substrate for the production of both AMP and GMP. 6 ATP are used to make 1 IMP from PRPP. None are required in the salvage pathway.

PPAT is also called Glutamine phosphoribosylpyrophosphate amidotransferase or amidophosphoribosyltransferase. It catalyzes the rate limiting step, is tightly regulated. *PPAT* possesses two nucleotide-binding sites near the active site, allowing for feedback control by downstream purine nucleotides via allosteric inhibition.

Figure 22.4.3 shows an interactive iCn3D model of the Glutamine Phosphoribosylpyrophosphate Amidotransferase from Arabidopsis thaliana (6LBP)



NCBI iCn3D Figure 22.4.3: Glutamine Phosphoribosylpyrophosphate Amidotransferase from Arabidopsis thaliana (6LBP). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...uWnf1bTq1odk88

The plant enzyme is a homotetramer, which each subunit having a Fe_4S_4 center (spacefill). The active site residues in each subunit (DDS 432-444 and DS 369-370) are shown as colored sticks and labeled.

Regulation of IMP production also occurs through enzyme phosphorylation. For example, Thr 397 on PPAT is phosphorylated by protein kinase B (PKB). PRPP concentrations also affects the rate. Another regulation of the flux through the de novo pathway is through the condensation of the enzymes into the purinosome, which contains *PPAT*, *TGART*, formylglycinamidine ribonucleotide synthetase (*FGAMS*), *PAICS*, adenylosuccinate lyase (*ADSL*), and *ATIC*. The purinosome also interacts with the mitochondria which would allow high local concentrations of ATP.

A cartoon view of the purinosome is shown in Figure 22.4.4.

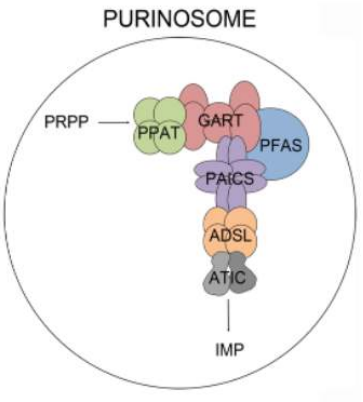


Figure 22.4.4: Cartoon showing the purinosome. Baresova et al (2018) PLoS ONE 13(7): e0201432. https://doi.org/10.1371/journal.pone.0201432. Creative Commons Attribution License,

Figure 22.4.5 shows another view of de novo IMP synthesis in which the origin of each atom in the purine ring is shown in color.

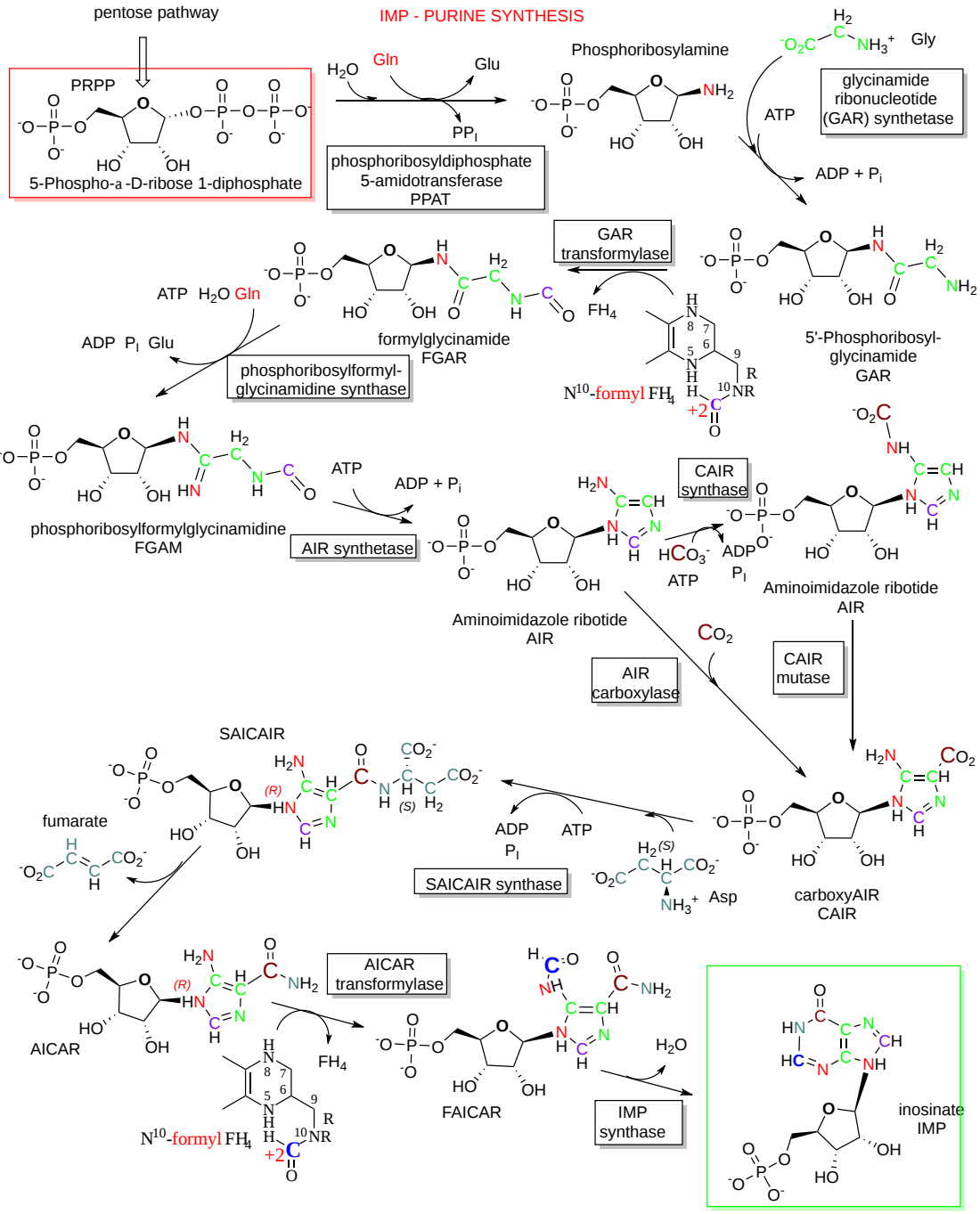


Figure 22.4.5: De novo IMP synthesis showing the origin of atoms in IMP.

Figure 22.4.6 shows an expanded view of the conversion of IMP to GTP and ATP.

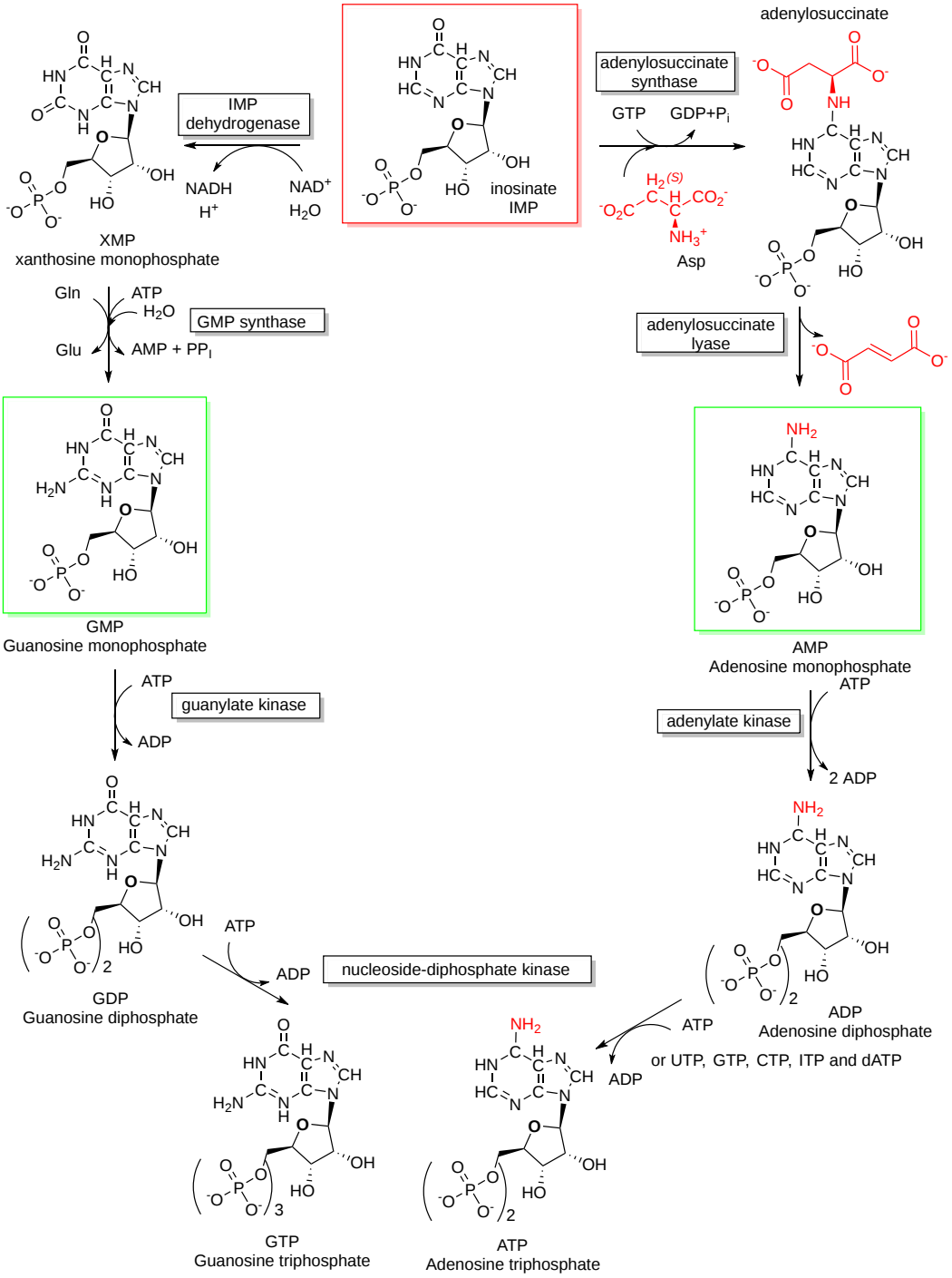


Figure 22.4.6: Conversion of IMP to GTP and ATP

22.4.3: PYRIMIDINE SYNTHESIS

As mentioned in the introduction, pyrimidines have a much simpler biosynthetic pathway. Instead of being synthesized as nucleobases as in the case for purines, they are made as ribonucleotides as they are linked to phosphoribosyl pyrophosphate (PRPP). Glutamine and aspartate again provide the ring C and N atoms. The one-carbon unit derives from serine-to-glycine conversion. A methyl group from the activated 1C donor, 5,10-meTHF, is added to dUMP to make dTMP.

The first step in the pathway for pyrimidine synthesis is the condensation of aspartate and carbamoyl phosphate. We have seen the synthesis of carbamoyl phosphate in the urea cycle by the enzyme carbamoylphosphate synthase I (CPSI) in Chapter 18.3 but present the reaction again in Figure 22.4.7.

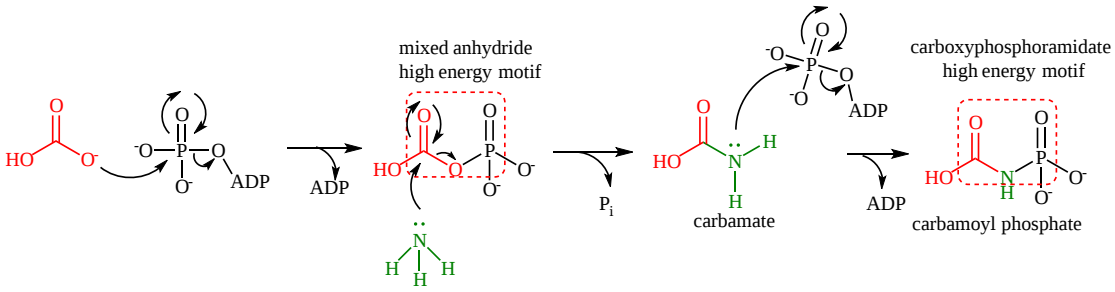


Figure 22.4.7: Synthesis of carbamoyl phosphate

A different cytosolic version of the enzyme, CPS II, is used to synthesize both arginine and pyrimidine nucleotides. It uses glutamine as a donor of NH_3.

The pathway for the synthesis of UTP and CTP are shown in Figure 22.4.7. It does not explicitly show the synthesis of carbamoylphosphate, which is an intergyral part of the pathway and one of the rate-limiting steps in pyrimidine synthesis.

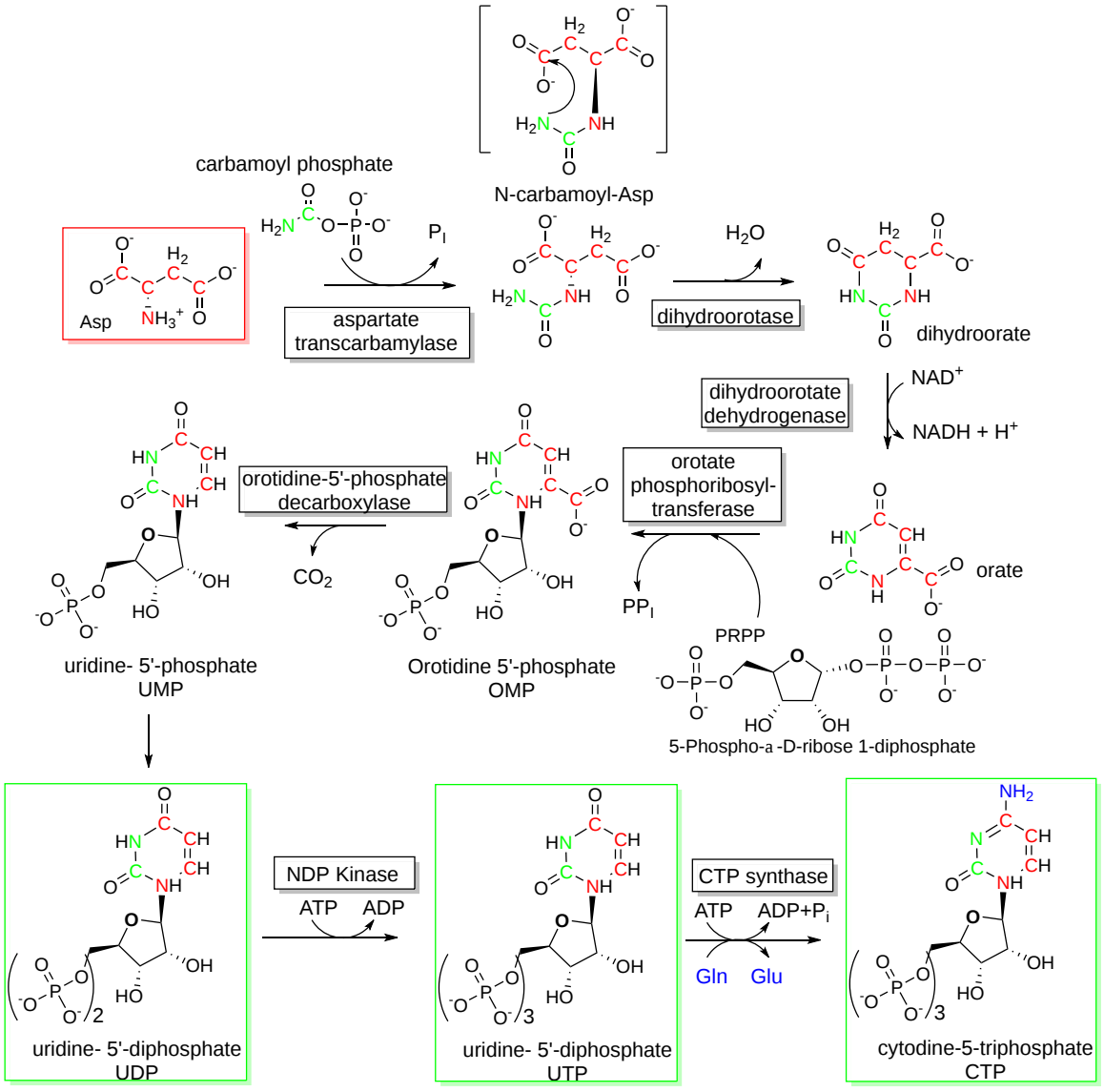


Figure 22.4.8: De novo synthesis of UDP, UTP and CTP

UDP and CDP can be converted to dCDP and dUDP, then on to dCPT and dUTP, and to dTMP as shown in Figure 22.4.9.

Deoxythymidine-5'-monophosphate Synthesis

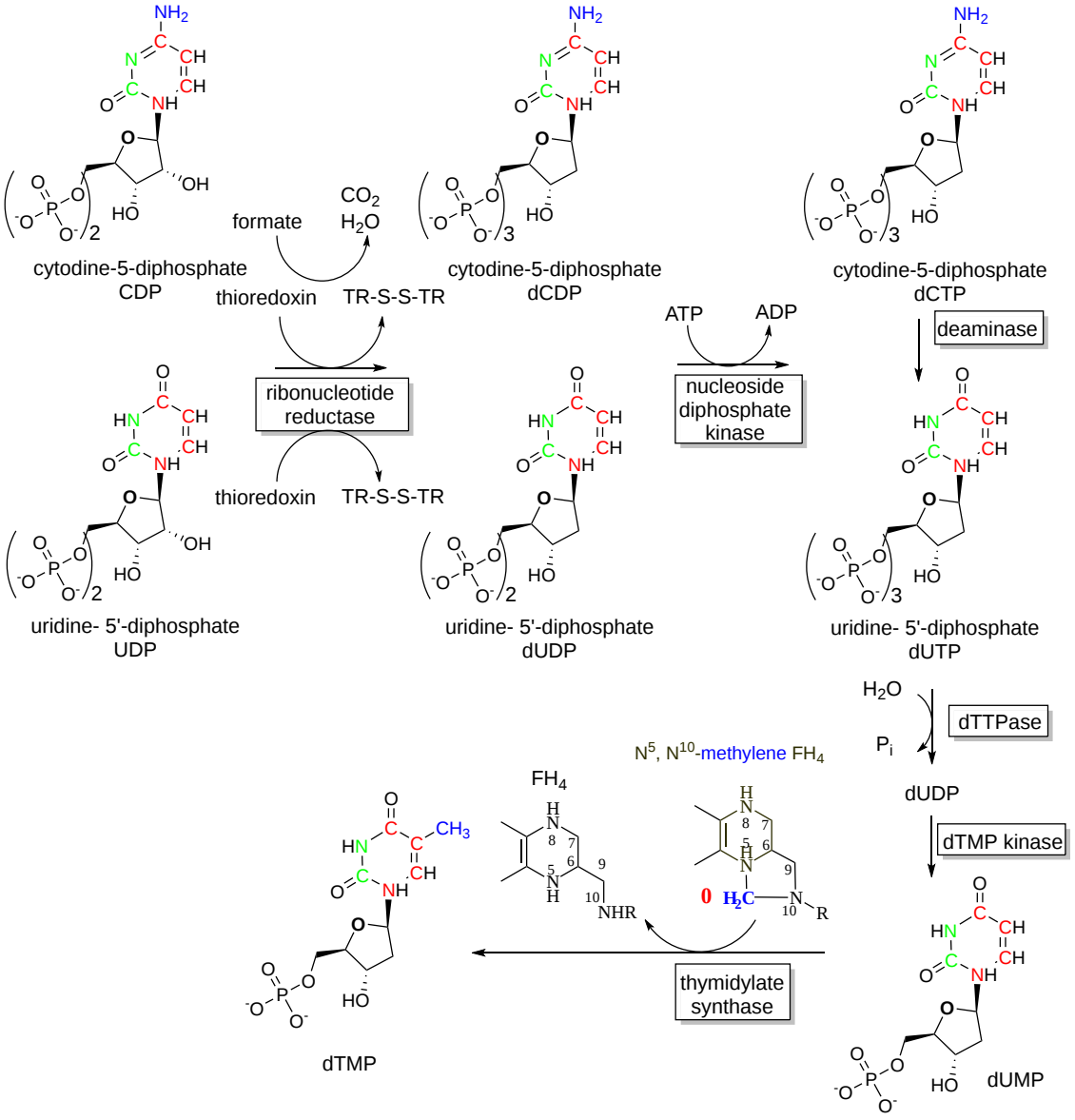


Figure 22.4.9: Synthesis of dTMP

Some of the material below derives from Li et al. *Int. J. Mol. Sci.* **2021**, *22*(19), 10253; **https://doi.org/10.3390/ijms221910253.** Creative Commons Attribution (CC BY) license (**https://creativecommons.org/licenses/by/4.0/**)

In purine synthesis, we saw multifunctional enzymes that catalyze several steps, as well as the assembly of the enzymes in the de novo pathway into the purinosome. In an analogous fashion, three different enzyme activities that catalyze the first three combined rate limiting steps of pyrimidine synthesis, **C**arbamoyl-phosphate synthetase, **A**spartate transcarbamoylase, and **D**ihydroorotase are found in a single, multifunction protein referred to as **CAD**. Its structure is a hexamer of a 243K monomer. It has 4 domains that include

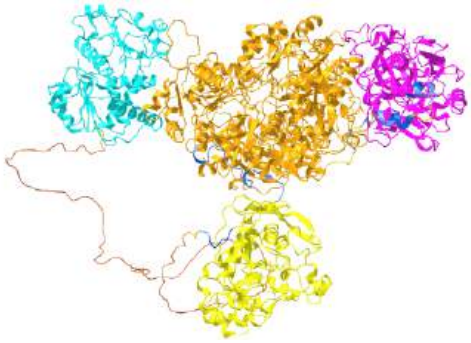
- glutamine amidotransferase (GATase) which "moves" HCO_3^-, glutamine, and ATP to the CPSIIase domain
- carbamoylphosphate synthetase II (CPSIIase): This has two parts, CPSaseA, and CPSase B. They combine functionally with GATase to form a glutamine-dependent carbamoylphosphate synthase (CPSase)
- aspartate transcarbamylase (ATCase) acts as a homotrimer
- dihydroorotase (DHOase) catalyzes the reversible cyclization reaction.

The CAD protein is a 'fusion' protein encoding these four enzymatic activities of the pyrimidine pathway. Figure 22.4.10 shows the domain structure of the CAD protein.

Figure 22.4.10: Domain structure of CAD

The red represents glutamine amidotransferase and the blue the carbamoyl phosphate synthase ATP binding domain. More specifically, the following amino acid stretches comprise the different domains: GATase (2-365), CPSase A (395-933), CPSlase B (934-1455), DHOase (1456-1788), and ATCase (1918-2225)

Figure 22.4.11 shows an interactive iCn3D model of the AlphaFold predicted model of the CAD protein (P27708)



NCBI iCn3D Figure 22.4.11: AlphaFold predicted model of the CAD protein (P27708). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...5V2XUmfsz1qjZ6

The GATase domain is magenta, the CPSase domains orange, the DHOase domain yellow, and the ATCase domain cyan. The structure of the disordered loops could not be modeled.

22.4.4: RIBONUCLEOTIDE REDUCTASES (RNRS)

Ribonucleotide reductases (RNRs), also called **ribonucleoside-diphosphate reductase,** catalyze the oxidation of the C2'-OH on the ribose ring to C2'-H through a free radical mechanism for the oxidation of for all NDPs including ADP, GDP, CDP and UDP (which converts to dUDP and in a different reaction to dTDP). The reaction is as follows:

[thioredoxin]-dithiol + ribonucleoside 5'-diphosphate ↔ [thioredoxin]-disulfide + 2'-deoxyribonucleoside 5'-diphosphate + H_2O

To refresh your mind, thioredoxin is a small protein (12K) that is part of a complex with thioredoxin reductase and thioredoxin-interacting protein. It has two key sulfhydryls at the active site which act as reducing agents as they get converted to a disulfide as shown in Figure 22.4.12.

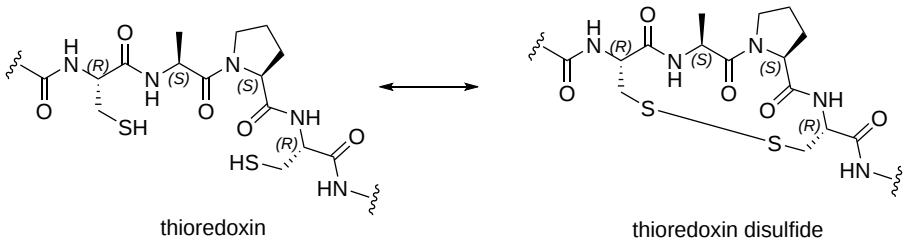


Figure 22.4.12: Reduced thioredoxin and its oxidized form

This single enzyme is so critical to cellular life that we will examine it more closely.

There are several classes of these enzymes (Ia-Ie, II, and III). The class I enzymes generally use a di-transition metal complex as a cofactor while class II uses adenosylcobalamin. We will focus on class I enzymes, which have two subunits, called α and β or M1 and M2, respectively. The NDP binds in an α/M1 subunit active site which is developed only in the dimer. The β/M2 subunit is often referred to as the radical-generating subunit as it contains the transition metal complex that generates a free tyrosyl radical cation critical to the reaction.

Almost all class I ribonucleotide reductases (RNRs) use transition metal ions located in the β/M2 subunit in the catalytic cycle for the dehydroxylation of the 2' OH on the ribose ring of the nucleotide. The metal ion complex a β/M2 subunit tyrosine to a tyrosine free radical cation which oxidizes an active site cysteine in the α/M1 subunit to form a thiol radical cation ($Cys•^+$), called a thiyl radical. This abstracts a

H• from the 3'C on the ribose of the substrate, forming a 3'C radical cation. This facilitates a dehydration reaction which leads to the dehydroxylation of the 2' OH, regenerating the thiyl radical. Reducing equivalents to restore the catalytic function of the enzyme come from the oxidation of a thioredoxin disulfide bound in the other subunit of the protein or a formate.

Given the importance of these enzymes, they must be highly regulated. There are two regulatory sites:

- a specificity site: determines nucleotide (NDP) specificity
- an activity site: regulates catalytic activity

The specificity and activity sites are in the α/M1 subunit where allosteric regulators dNTPs and ATP bind to different sites. When ATP is bound, the enzyme uses CDP and UDP as substrates. When dGTP is bound, ADP is the preferred substrate. Finally, when dTTP is bound, GDP is the preferred substrate. The enzyme is inhibited by dATP binding to the actual active site.

Figure 22.4.13s shows an abbreviated mechanism and cartoon showing the activities of the two subunits.

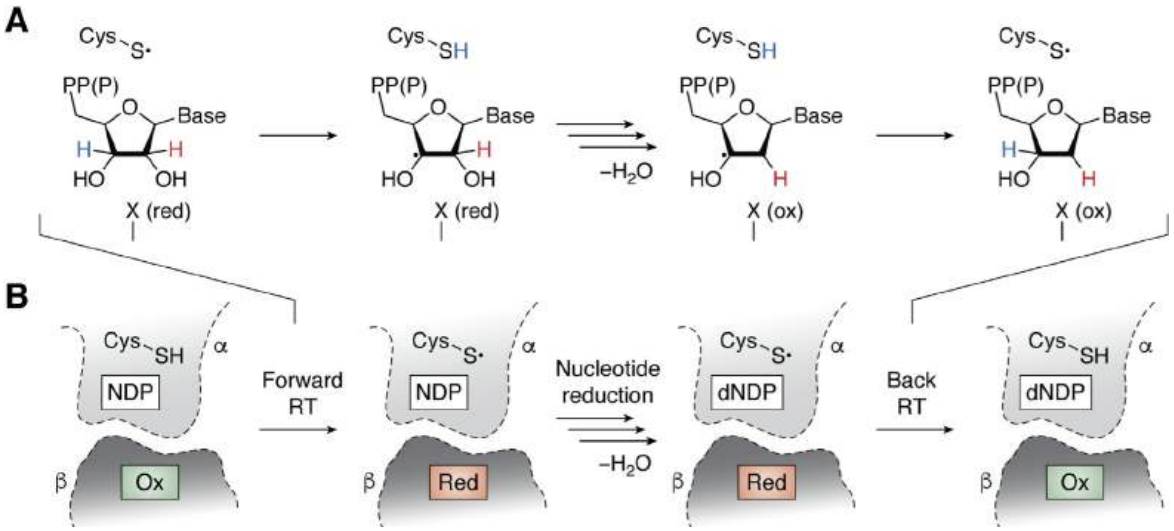


Figure 22.4.13: Abbreviated mechanism and a cartoon showing the activities of the two subunits. in class I RNR. *A*, universal mechanism for nucleotide reduction in RNRs. *B*, diagram of the steps involved in radical translocation in class I RNRs. Ruskoski and. Boal. J. Biol. Chem. (2021) 297(4) 101137. DOI:https://doi.org/10.1016/j.jbc.2021.101137. CC-BY license (http://creativecommons.org/licenses/by/4.0/)

Radical formation starts at tyrosine 122 in the metal center site in the β/M2 subunit. Electron transfer then occurs across the two subunits from a very distant active site Cys 439 in the α subunit. which enables the formation of the thiyl radical cation ($Cys^{\bullet+}$).

The structure of an E. Coli Type IA enzyme with bound ligands has been determined after much effort that involved trapping of a long-life intermediate. It required replacement of Tyr 122 in both β chains with 2,3,5-trifluorotyrosine, which allowed the structure to be determined by cyro-EM. Tyr 122 in the β chain forms starts the process of electron transfer as it becomes the Tyr $122^{\cdot+}$ radical cation.

A detailed mechanisms showing both electron transfer to $Y122^{\cdot+}$ and accompanying proton transfer is shown in Figure 22.4.14.

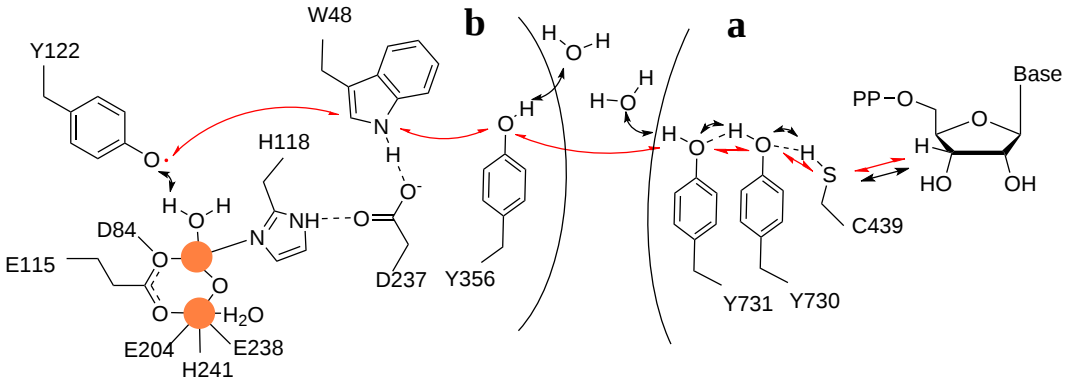
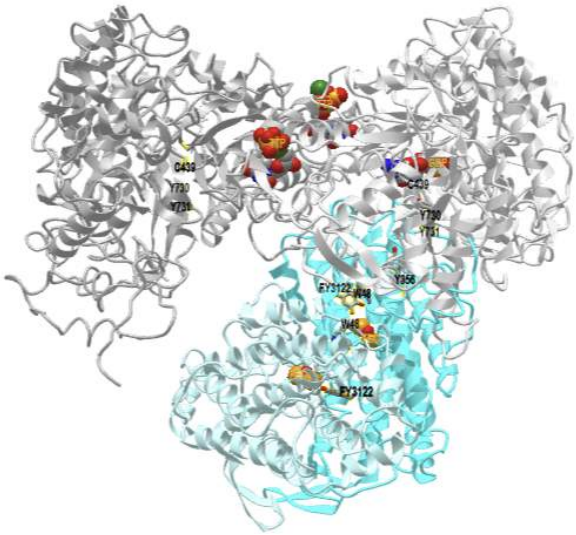


Figure 22.4.14: Pathway of electron and proton transfers for the formation of the thiyl radical cation ($Cys^{\bullet+}$). after Kang et al. Science (2020). DOI: 10.1126/science.aba6794)

The path for electron flow from the sulfur of C439 in the β subuit to regenerate Y122 is shown. That electron transfer occurs over a very long distance of 35 Å as it hops from the original sulfur donor to the acceptor.

Figure 22.4.15 shows an interactive iCn3D model of the holocomplex of E. coli class Ia ribonucleotide reductase with GDP and TTP (6W4X).



NCBI iCn3D Figure 22.4.15: Holocomplex of E. coli class Ia ribonucleotide reductase with GDP and TTP (6W4X). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...71ubUDRovXURG8

The structure is a tetramer of two α chains (different shades of gray) and two β chains (different shades of cyan). The GDP and TTP in the α subunit and the Fe cluster in the β subunit are shown in CPK spacefill and labeled. The key side chains in the α and β chains that participate in electron transfer over the 35 Å distance are shown in CPK-colored sticks and labeled. The Fe^{2+} ions are in the form of a μ-oxo-diron complex (2 Fe ions coordinated by 1 oxide).

Figure 22.4.16 shows specificity and catalytic sites of the biologically active tetramer form for Class 1 RNR as well as the transition metal cofactors (Fe, Mn, or both) in various Class 1 RNRs.

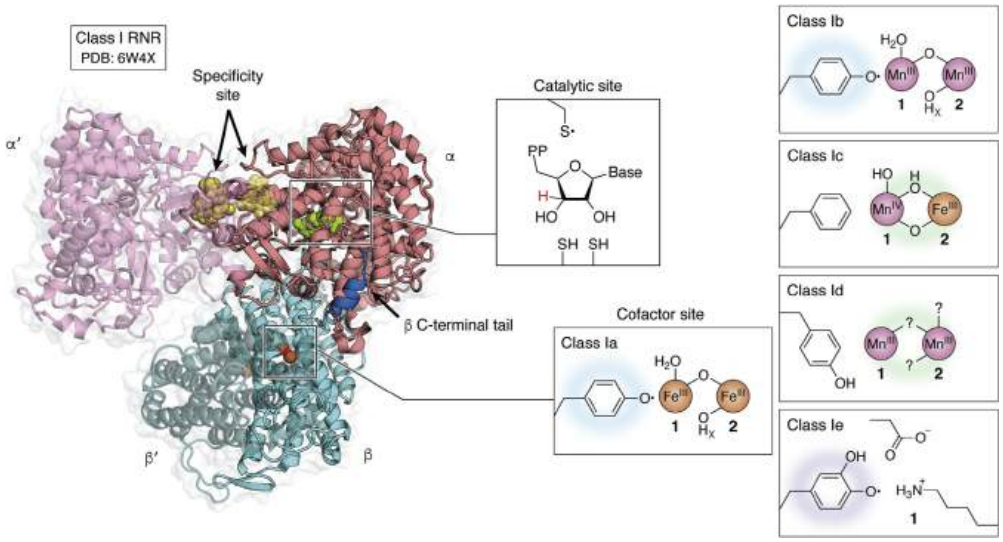
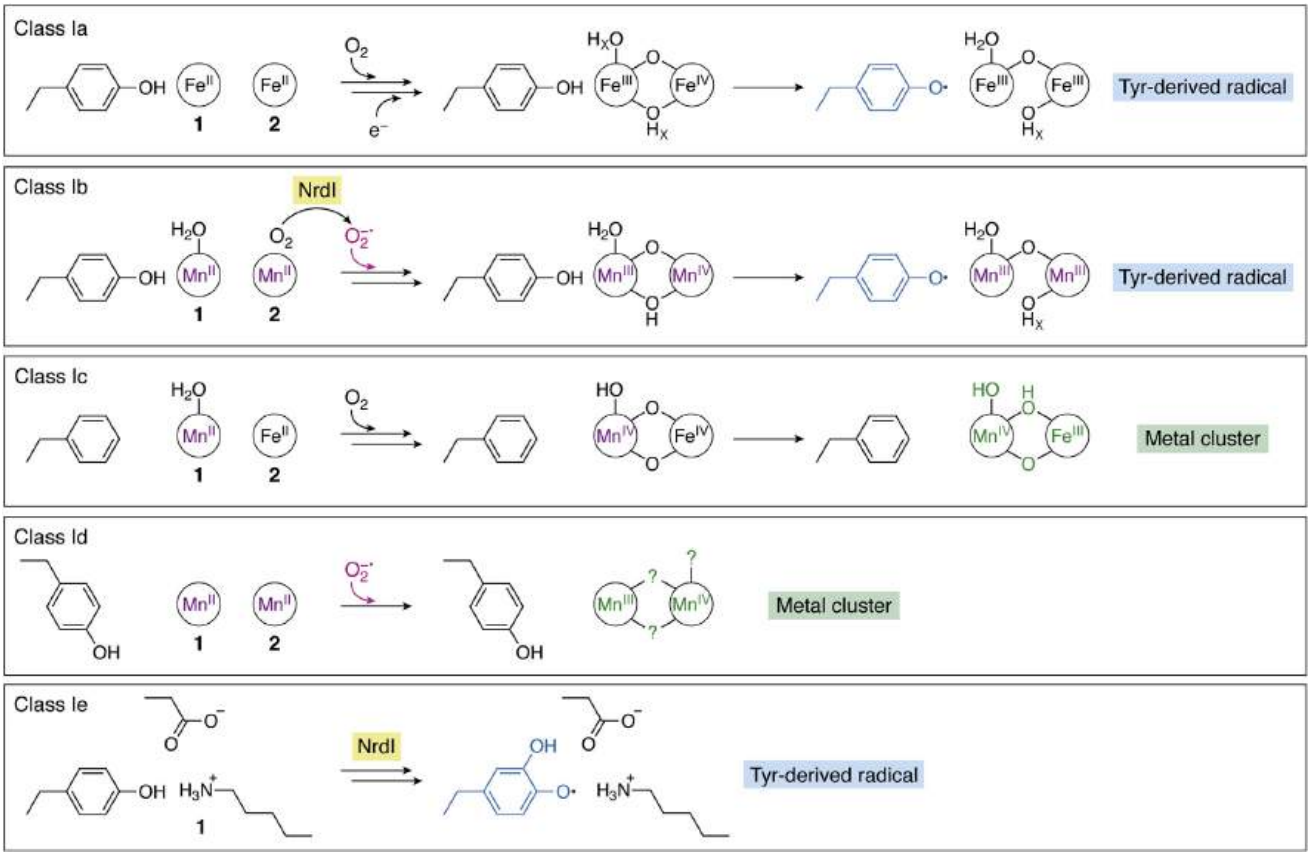


Figure 22.4.16: Quaternary structure of the active holoenzyme complex in class I RNR (PDB accession code 6W4X). *Insets* show the location of the active site in the catalytic α subunit (*middle top*) and the metallo- or radical cofactor (*middle bottom* and *far right*) in the β subunit

The tyrosyl free radical forms on binding of dioxygen to the transition metal ion center electron and subsequent loss of an electron from tyrosine, as shown in Figure 22.4.17:

Figure 22.4.17: Cofactor assembly mechanisms for class I RNRs. Manganese-dependent enzymes are highlighted in *purple*. Superoxide-dependent RNRs are highlighted in *hot pink*. Subclasses that require a NrdI activase are indicated with a *yellow box*. Metal-centered Cys oxidants shown in *green* and Tyr-derived radical Cys oxidants shown in *blue*.

The mechanisms of cofactor actions are not fully understood. O_2 initially adds to the Fe^{2+}/Fe^{2+} cluster which forms a peroxo-Fe^{3+}/Fe^{3+}



intermediate.

Structural features of the active site for different class I RNRs along with redox-active transition metal ions are shown in Figure 22.4.18.

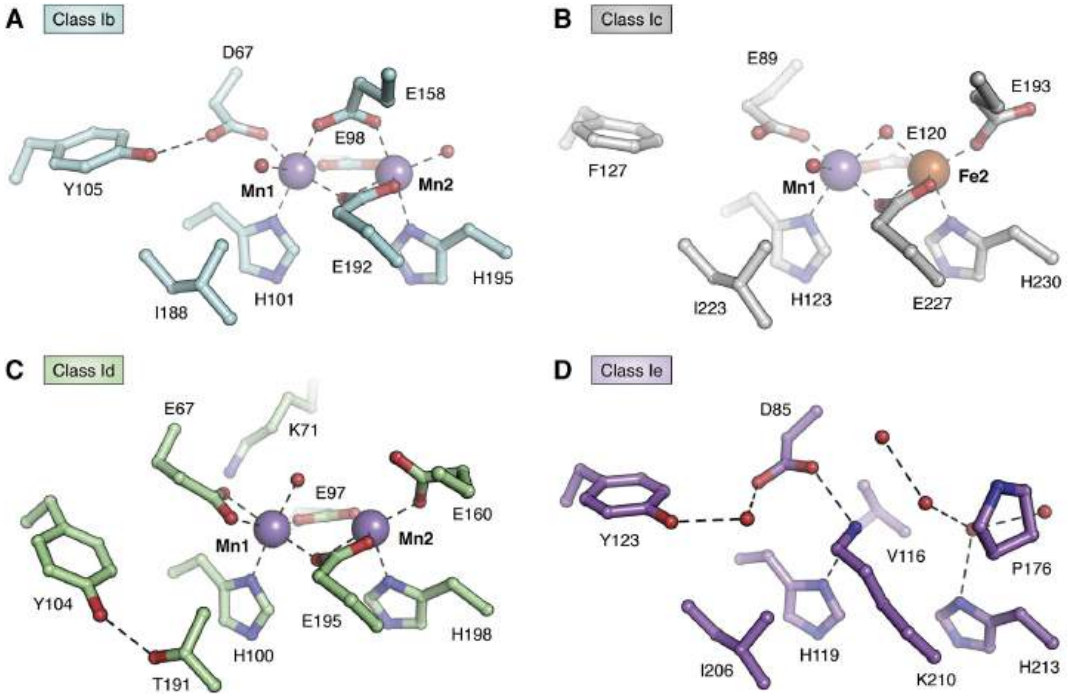


Figure 22.4.18: Comparison of metal-binding sites in class I RNRs

The PDB codes for each structure are as follows: A, (3N3A), B, 4M1I), C, (6CWP), and D (6EBO). Water molecules are shown as *red spheres*.

Now let's look at the regulation of class IA RNRs in more detail. Let's consider perhaps the most important allosteric regulators which bind 15 Å from the active site which affect RNR enzymatic activity:

- dATP: inhibits RNRs when it binds to the α subunit
- ATP: reverses the inhibition by dATP

dATP/ATP also affects RNR enzyme specificity as they tilt the preference of RNR towards pyrimidine substrates, where TTP and dGTP promote purine substrate binding. These same rules apply to human and E. coli RNRs, with the locations of the active sites also being the same.

The allosteric regulators appear to affect the quaternary structure of the enzyme.

In E. Coli, the binding of dADP converts the structure of the enzyme from an active $\alpha_2\beta_2$ form to an inhibited $\alpha_4\beta_4$ ring structure. When dATP is bound, a "**cone**" domain in α forms interactions with the β subunit, leading to the formation of a dimer of the $\alpha_2\beta_2$ tetramer. ATP reverses this effect by displacing dATP and pushing the equilibrium towards the active $\alpha_2\beta_2$ form. This is shown in Panel A of Figure 22.4.19.

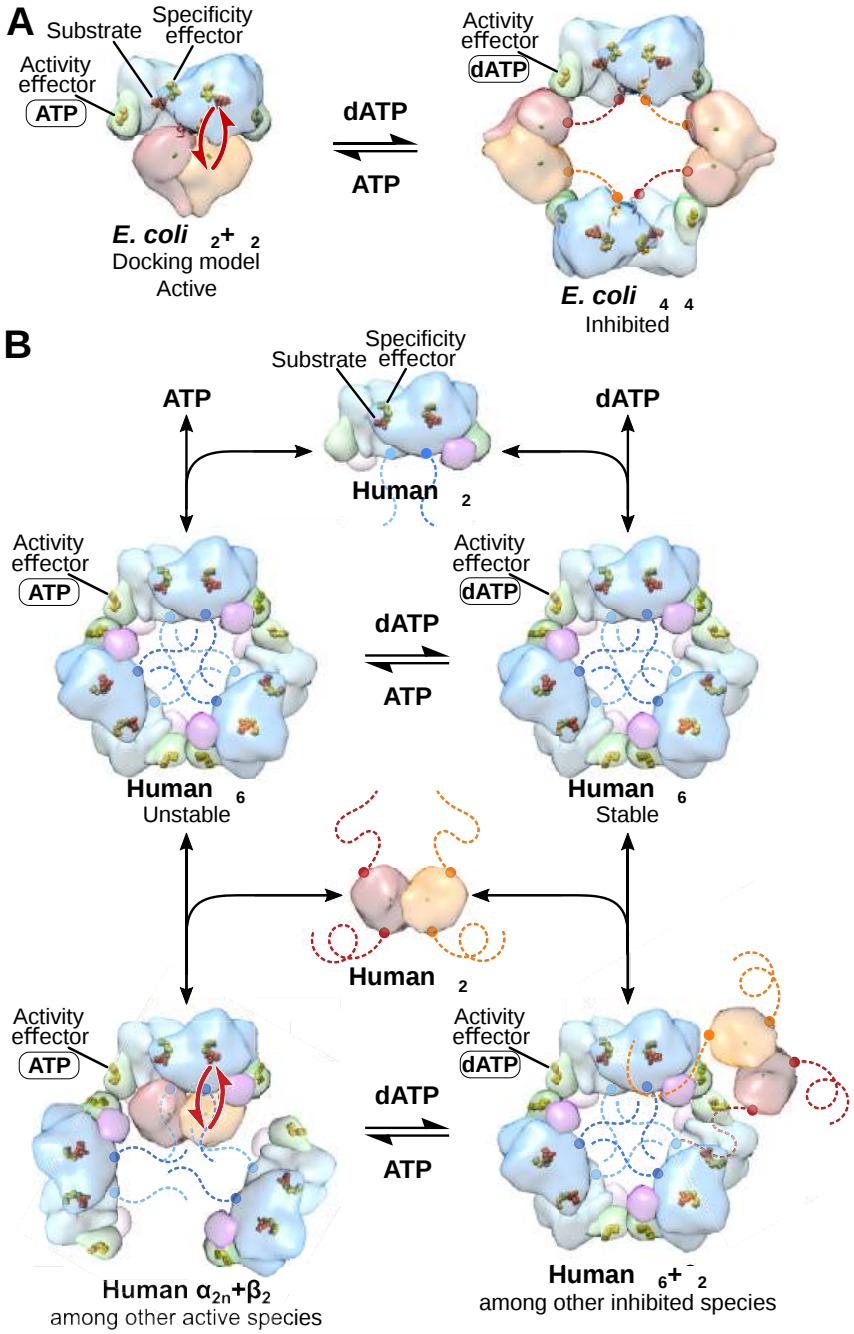


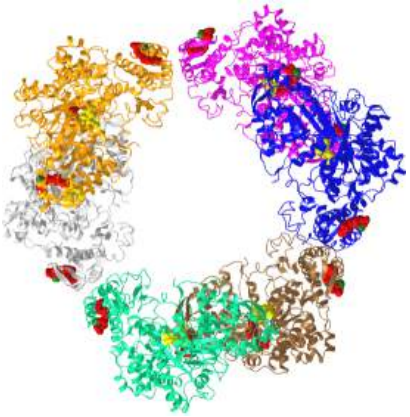
Figure 22.4.19: **Comparison of mechanisms of allosteric regulation of activity for *E. coli* and human RNRs.** Brignole et al. 3.3-Å resolution cryo-EM structure of human ribonucleotide reductase with substrate and allosteric regulators bound eLife 7:e31502. https://doi.org/10.7554/eLife.31502. Attribution 4.0 International (CC BY 4.0).

Panel B above illustrates the regulation of the human enzyme, which appears to be quite different. In the absence of either dATP or ATP, RNRs exist just as α_2 dimers. On binding either dATP or ATP, the α_2 dimer is converted to an α_6 (a trimer of dimers) structure. Both are inactive in the absence of β subunits (which provide the metal cofactor site)

- When β_2 is added to the dATP bound α_6 hexamer, the hexamer becomes stabilized, but the resulting complex is inhibited.
- When β_2 is added to the ATP bound α_6 hexamer, the hexamer becomes destabilized and into smaller structures which are active.

Hence the ratio of cellular dATP/ATP changes the aggregation state of the RNR and hence its activity in both E. Coli and humans.

Figure 22.4.20 shows an interactive iCn3D model of the Human ribonucleotide reductase large subunit (alpha) with dATP and CDP (6AUI).



NCBI iCn3D Figure 22.4.20: Human ribonucleotide reductase large subunit (alpha) with dATP and CDP (6AUI). (Copyright; author via source). Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...w1sFwwYSAD4nG6

Each of the α subunits in the hexamer is shown in a different color. dATPs (allosteric inhibitors) are shown in spacefill, red. CDPs (substrates) are shown in spacefill yellow.

The hole in the middle of the structure prevents β_2 from binding in a catalytically productive fashion, so the structure is inactive.

A specific loop, loop 2, assists in the determination of RNRs specificity. It is between the base of dATP and the base of the substrate CDP. The backbone of the loop interacts with the adenine base and orients Gln 288 in the active site to interact with the cytosine of the substrate CDP. In both systems, backbone atoms of this loop ‘read out’ the adenine base and position Gln288 into the active site to recognize the cytosine of CDP. Figure 22.4.20 gives further details on the origin of substrate specificity.

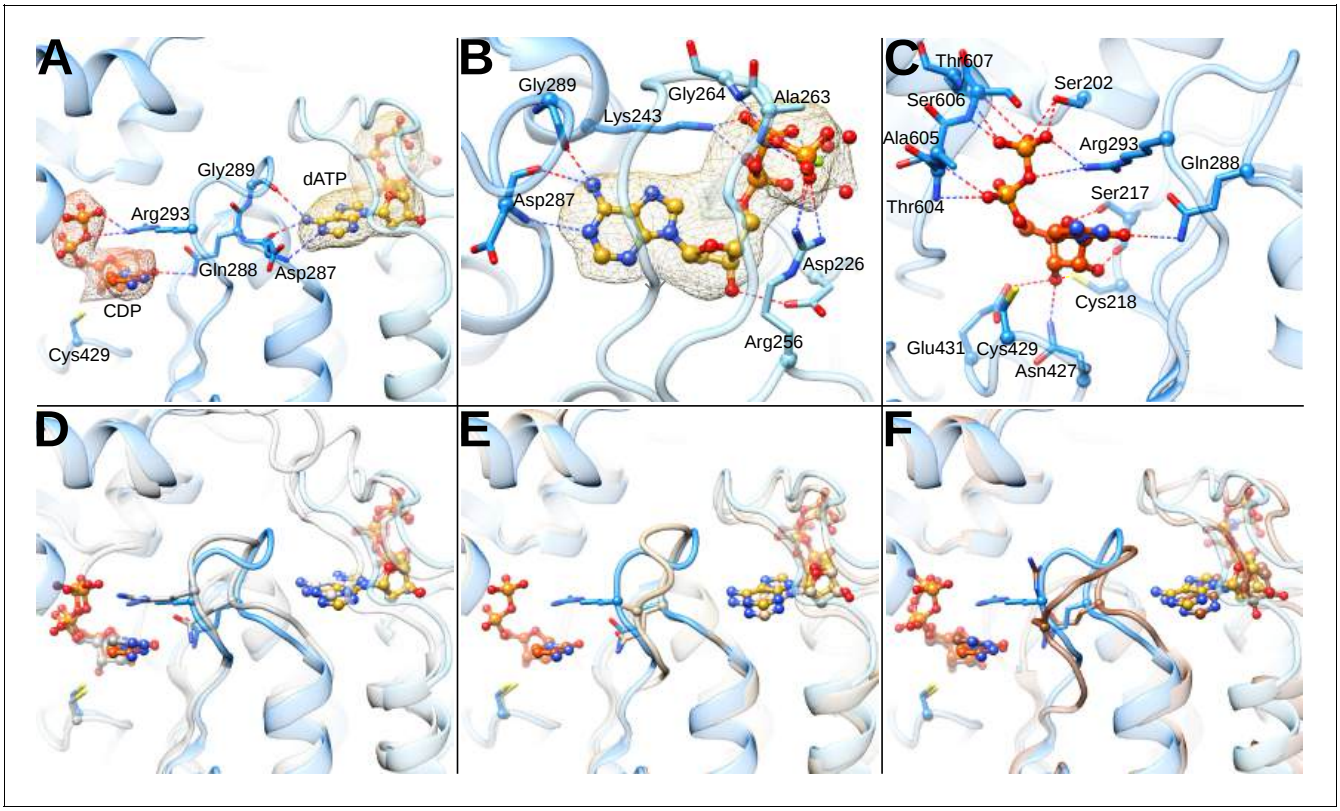


Figure 22.4.21: Determinants of substrate specificity are conserved from *E. coli* to humans**.**

Panel (**A**) shows residues of human α (blue) interacting with CDP (carbons in orange) in the active site and dATP (carbons in yellow) in the specificity site. Density for CDP in orange mesh and for dATP in the yellow mesh.

Panel (**B**) zooms in on dATP in the specificity site. Water molecules and oxygen atoms are in red, nitrogen in blue, magnesium in green, and phosphate in gold.

Panel (**C**) zooms in on CDP in the active site.

Panel (**D**) shows an overlay of human α from the α_6 EM structure in blue with *E. coli* α from the $\alpha_4\beta_4$-CDP-dATP cocrystal structure in gray (PDB: 5CNS) shows a nearly identical loop 2 conformation positioning Gln288 and Arg293 (Gln294 and Arg298 in *E. coli*).

Panel (**E**) shows an overlay of human α from the α_6 EM structure in blue with the crystal structure of human α with N- and C-termini truncated (residues 77–742) cocrystallized with dATP in tan (PDB: 2WGH) shows similar positioning of dATP but an altered conformation of loop 2 in the absence of bound CDP. The CDP shown is from the α_6 EM structure.

Panel (**F**) shows an overlay of human α from the α_6 EM structure in blue with equivalent residues of yeast α structure with CDP and AMPPNP in brown (PDB: 2CVU) and shows a conformation of loop 2 that is distinct from that seen in structures of *E. coli* and human α

22.4.4.1: TRANSCRIPTIONAL REGULATION OF RNRS

Given the importance of this key enzyme, it should come as no surprise that its levels are regulated at the transcription level. As the activity of RNR is determined by its polymeric quaternary state, the transcriptional activation of the genes for RNRs is controlled in bacteria by quaternary states of the RNR-specific transcriptional repressor NrdR. The transcription factors bind to a specific DNA sequence called an NrdR box, which precedes the start site of transcription for RNR, where RNA polymerase binds.

The NrdR protein acts to repress the synthesis of RNR genes. Its aggregation state is controlled by dATP/ATP ratios. When dATP is high, the NrdR binds to DNA and represses the synthesis of the RNR gene. In contrast when ATP is high, the protein does not bind to DNA and hence it can not repress transcription. The association /dissociation of the repressor is controlled by the aggregation state of NrdR. When abundant, NrdR exists as a 12-mer complex with two molecules of ATP bound per monomer. This ATP-bound 12-mer can't bind DNA, so transcription of the gene for RNR is not repressed. As dATP increases, one ATP is displaced so each monomer has 1 dATP and 1 ATP bound. This causes the NrdR to covet to an 8-mer. A 4-mer (tetrameric) version of this protein binds to the NrdR box sequence at the start of the RNR gene, repressing its synthesis.

Figure 22.4.22 shows the **dodecameric, octameric, and tetrameric structures of NrdR and their functions.**

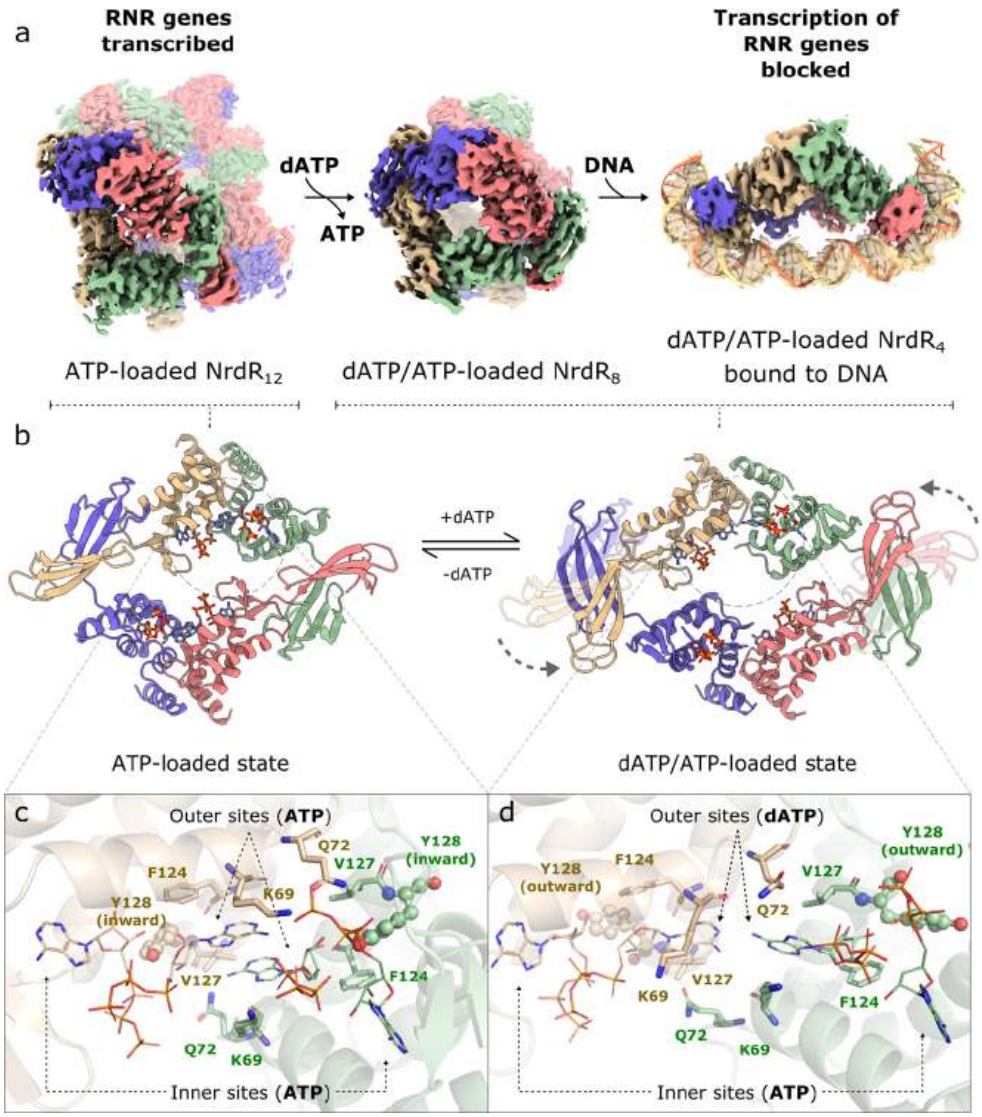


Figure 22.4.22: **Mechanism of NrdR function involving dodecameric, octameric, and tetrameric structures.** Rozman Grinberg, I., Martínez-Carranza, M., Bimai, O. *et al.* A nucleotide-sensing oligomerization mechanism that controls NrdR-dependent transcription of ribonucleotide reductases. *Nat Commun* **13,** 2700 (2022). https://doi.org/10.1038/s41467-022-30328-1. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Panel (a) shows a surface representation of the cryo-EM maps for the dodecameric, octameric, and DNA-bound tetrameric NrdR structures.

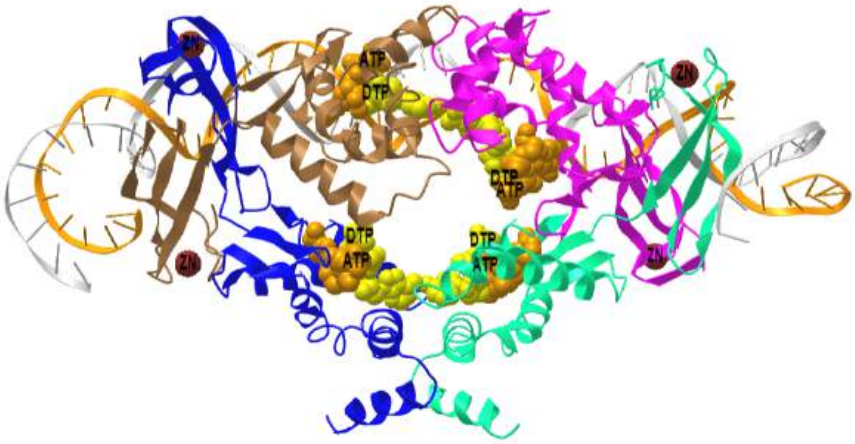
Panel (b) shows a cartoon representation of the ATP-loaded NrdR tetramer (left) and the dATP/ATP-loaded tetramer (right). Chains A, B, C, and D are colored in beige, green, pink, and blue, respectively.

Panel (c) shows the interface between the ATP cones in chain A (beige) and chain B (green) in the ATP-loaded dodecamer.

Panel (**d**) shows the dATP/ATP-loaded tetramer.

Panels **c**, and **d** were made from the same perspective, based on an alignment of the ATP-cones in chains A and B in both structures.

Figure 22.4.23 shows an interactive iCn3D model of the Streptomyces coelicolor dATP/ATP-loaded NrdR in complex with its cognate DNA (7P3F)**.**



NCBI iCn3D Figure 22.4.23: Streptomyces coelicolor dATP/ATP-loaded NrdR in complex with its cognate DNA (7P3F). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...2u3HMRPmJHe9X9

The monomers of the NrdR tetramer are shown in different colors. The yellow spheres are dATP and the orange ones are ATP.

22.4.5: PYRIMIDINE SALVAGE PATHWAY

There is also a salvage pathway as shown in Figure 22.4.24.

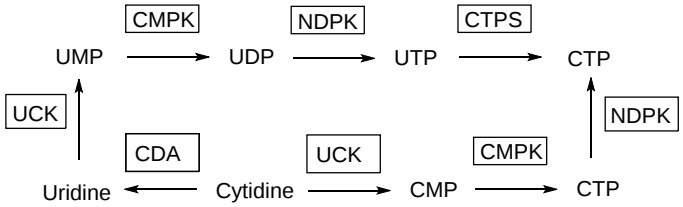


Figure 22.4.24: Pyrimidine salvage pathway (after Wang et al. Frontiers in Oncology, 11 (2021). https://www.frontiersin.org/article/...nc.2021.684961. DOI=10.3389/fonc.2021.684961

Cells at rest use the salvage pathway reactants derived from nucleic acid degradation generic nucleoside pools.

22.4.6: NUCLEOTIDE DEGRADATION

22.4.6.1: PURINES

The pathway for purine degradation is shown in Figure 22.4.25.

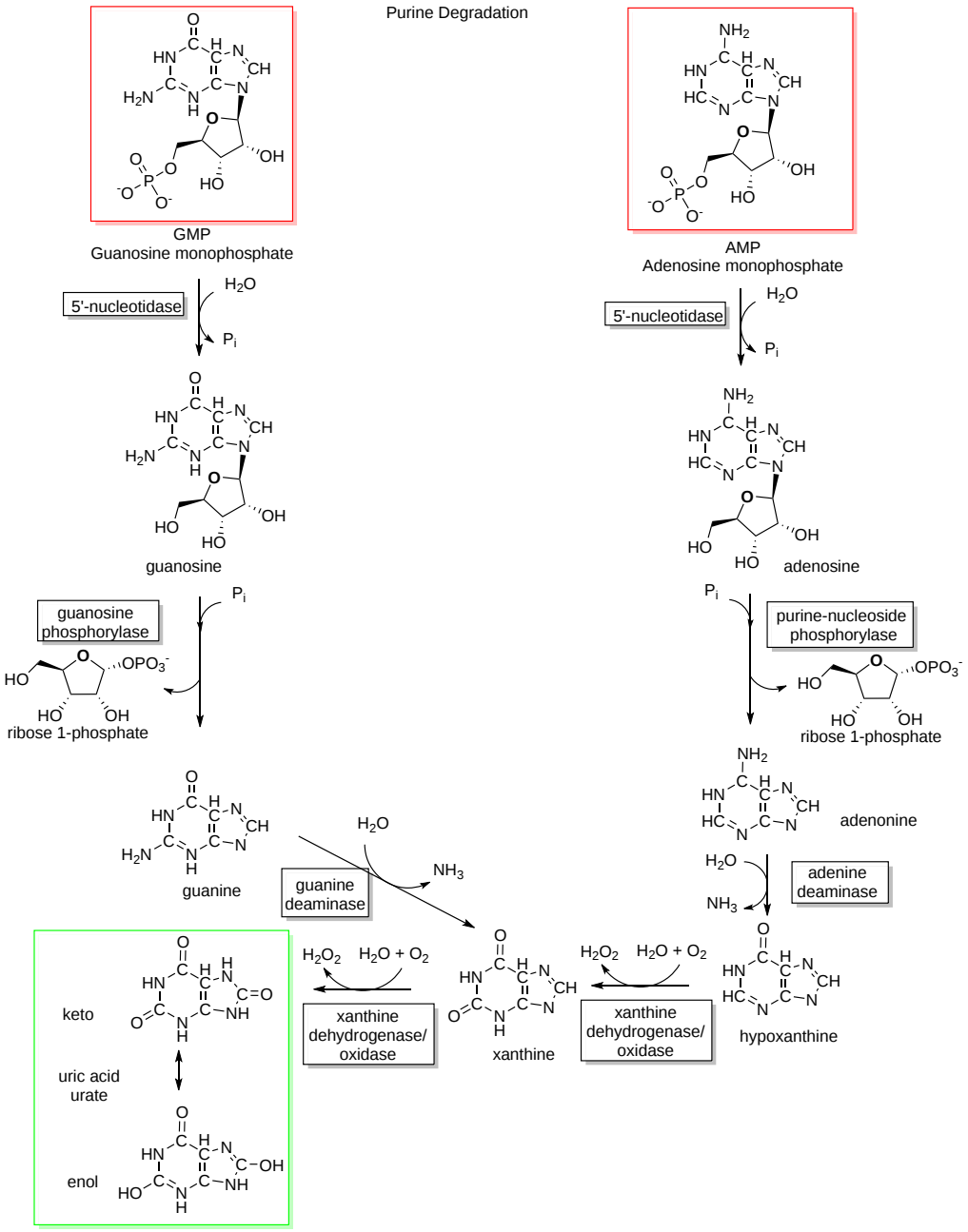


Figure 22.4.25: Purine degradation

22.4.6.2: PYRIMIDINE DEGRADATION

The pathway for pyrimidine degradation is shown in Figure 22.4.26.

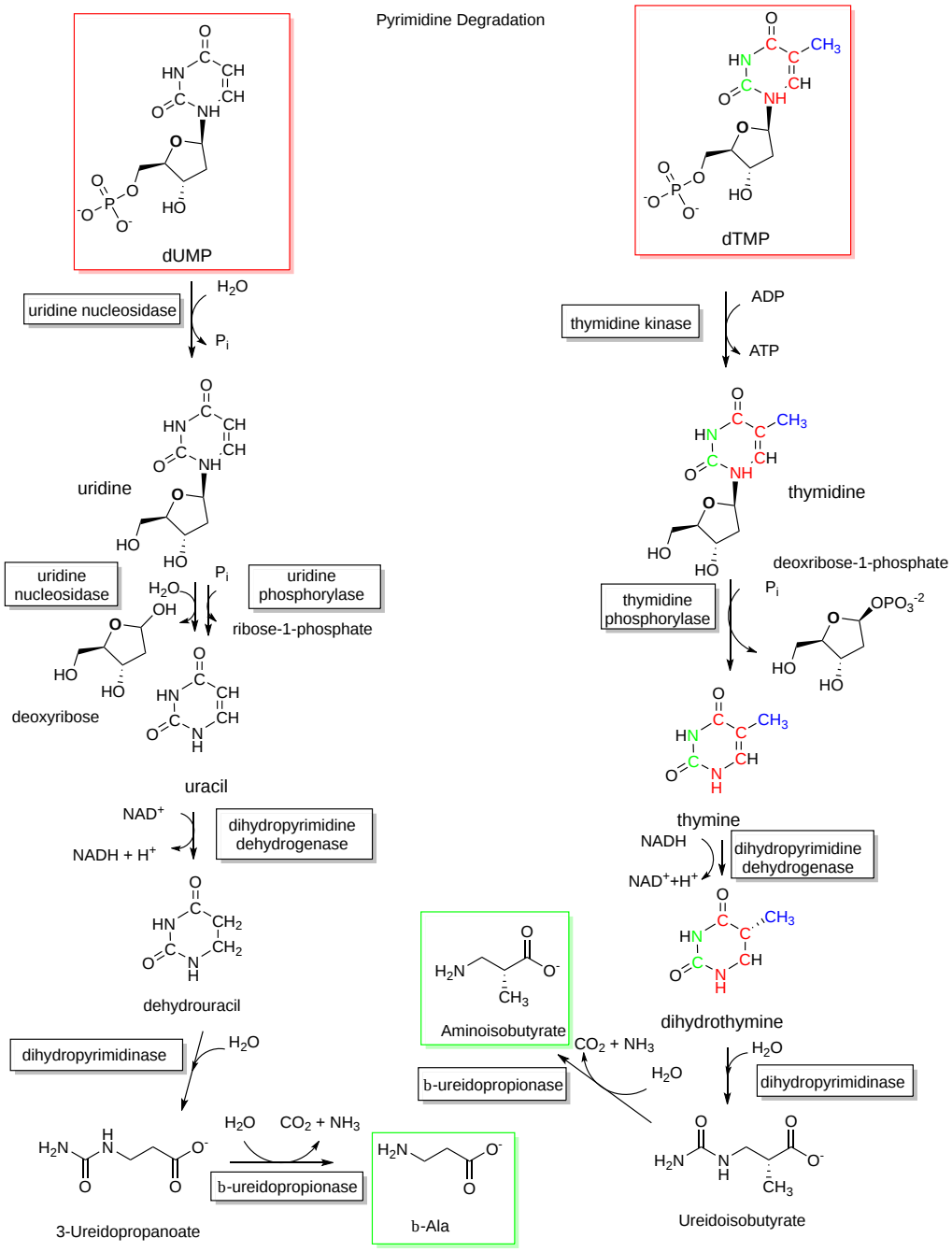


Figure 22.4.26: Pyrimidine degradation

CHAPTER OVERVIEW

23: Chromosome Structure

23.1: Gene Mapping and Chromosomal Karyotypes

23.2: DNA Transposable Elements

23.3: Chromosome Packaging

23.1: Gene Mapping and Chromosomal Karyotypes

23.1.1: Introduction

Genes provide instructions to build living organisms and each specific gene maps to the same chromosome in every cell. This physical gene location within the organism's chromosomes is called the gene loci. If two genes are found on the same chromosome, especially when they are near one another, they are said to be linked.

Genetic linkage is the tendency of DNA sequences that are close together on a chromosome to be inherited together during the meiosis phase of sexual reproduction. Two genetic markers that are physically near to each other are unlikely to be separated into different chromatids during chromosomal crossover and are therefore said to be more *linked* than markers that are far apart. In other words, the nearer two genes are on a chromosome, the lower the chance of recombination between them, and the more likely they are to be inherited together. Markers on different chromosomes are perfectly *unlinked*.

Genetic linkage is the most prominent exception to Gregor Mendel's Law of Independent Assortment. The first experiment to demonstrate linkage was carried out in 1905. At the time, the reason why certain traits tend to be inherited together was unknown. Later work revealed that genes are physical structures related by physical distance.

The typical unit of genetic linkage is the centimorgan (cM). A distance of 1 cM between two markers means that the markers are separated into different gametes on average once per 100 meiotic products, thus once per 50 meioses. A **linkage map** (also known as a **genetic map**) is a table for a species or experimental population that shows the position of its known genes or genetic markers relative to each other in terms of recombination frequency, rather than a specific physical distance along each chromosome. Linkage maps were first developed by Alfred Sturtevant, a student of Thomas Hunt Morgan. Figure 23.1.1 shows a gene linkage map with the relative positions of allelic characteristics on the second Drosophila chromosome.

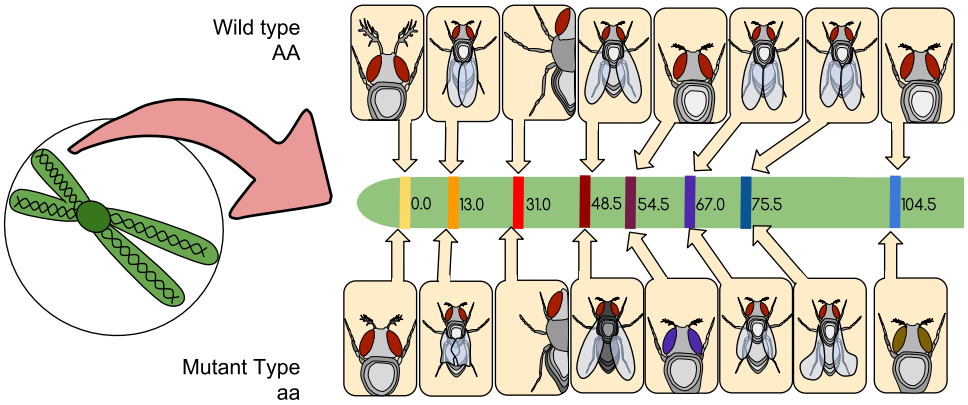


Figure 23.1.1: Gene linkage map showing the relative positions of allelic characteristics on the second Drosophila chromosome. The alleles on the chromosome form a linkage group due to their tendency to form together into gametes. The distance between the genes (map units) is equal to the percentage of crossing-over events that occurs between different alleles. This diagram is also based on the findings of Thomas Hunt Morgan in his Drosophila cross. (CC BY-SA 4.0 International; Twaanders17 via Wikipedia

A linkage map is a map based on the frequencies of recombination between markers during the crossover of homologous chromosomes. The greater the frequency of recombination (segregation) between two genetic markers, the further apart they are assumed to be. Conversely, the lower the frequency of recombination between the markers, the smaller the physical distance between them. Historically, the markers originally used were detectable phenotypes (enzyme production, eye color) derived from coding DNA sequences; eventually, confirmed or assumed noncoding DNA sequences such as microsatellites or those generating restriction fragment length polymorphisms (RFLPs) have been used.

Linkage maps help researchers to locate other markers, such as other genes by testing for genetic linkage of the already known markers. In the early stages of developing a linkage map, the data are used to assemble **linkage groups**, a set of genes that are known to be linked. As knowledge advances, more markers can be added to a group, until the group covers an entire chromosome. For well-studied organisms, the linkage groups correspond one-to-one with the chromosomes.

Traditional studies used to physically map genes onto specific chromosomes were painstaking and involved using restriction enzymes to fragment the genome of an organism and then clone the fragments into YACs or BACs creating a DNA library. The library could then be screened with specific genetic probes to determine which fragment contained a gene of interest. The

fragments would then need to be sequenced and reassembled using overlapping patterns. Today, the sequencing of entire genomes from nearly any organism is possible and relatively easy in comparison. Thus, a traditional genetic map can more readily be overlayed on the physical chromosomal map of an organism as shown in Figure 23.1.2 This was one of the overarching goals of the human genome project.

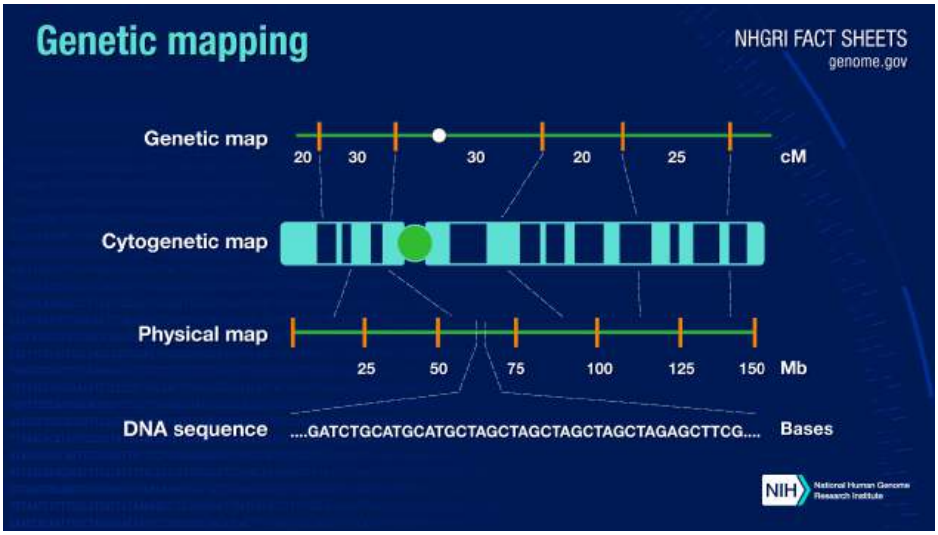


Figure 23.1.2: Traditional genetic map overlayed on the physical chromosomal map of an organism. The National Institutes of Health

23.1.2: Karyotypes

The entire chromosome set of a species is known as a **karyotype**, which can be thought of as a global map of the nuclear genome. **Karyotyping** is the process by which the condensed chromosomes of an organism are stained and photographed using light microscopy. Karyotyping can be used to determine the chromosome complement of an individual, including the number of chromosomes and any abnormalities.

Karyotypes describe the chromosome count of an organism and what these chromosomes look like under a light microscope. Attention is paid to their length, the position of the centromeres, banding pattern, any differences between the sex chromosomes, and any other physical characteristics. The preparation and study of karyotypes are part of the larger field of **cytogenetics**. The field of **cytogenetics** involves the study of inheritance in relation to the structure and function of chromosomes. Thus, karyotyping is a fundamental process within this field.

The study of whole sets of chromosomes is sometimes known as **karyology**. The chromosomes are depicted (by rearranging a photomicrograph) in a standard format known as a *karyogram* or *idiogram*: in pairs, ordered by size and position of centromere for chromosomes of the same size, as shown in Figure 23.1.3

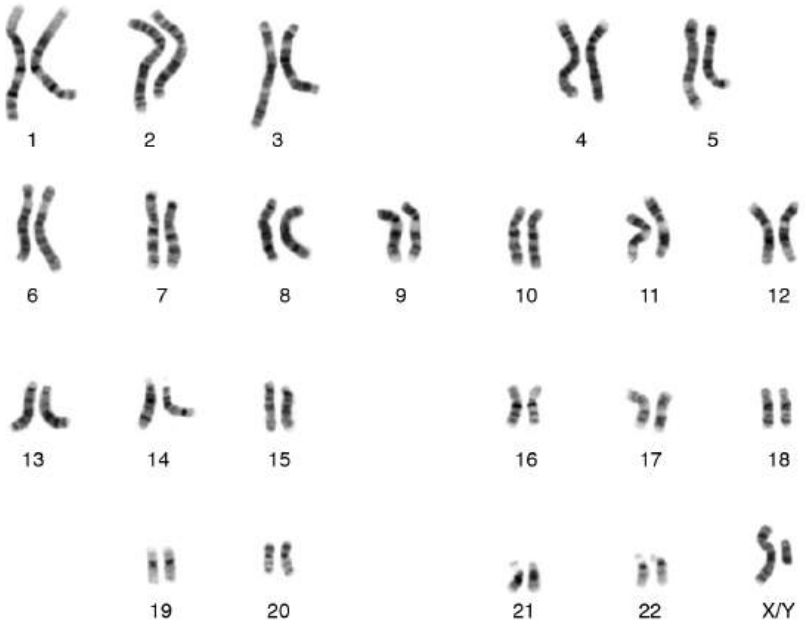


Figure 23.1.3: Human Karyogram. The National Cancer Institute

The basic number of chromosomes in the somatic cells of an individual or a species is called the **somatic number** and is designated ***2n*.** In the germ line (the sex cells) the chromosome number is ***n*** (humans: n = 23). Thus, in humans 2n = 46. In normal **diploid** organisms, autosomal chromosomes are present in two copies. There may, or may not, be sex chromosomes. **Polyploid** cells have multiple copies of chromosomes and **haploid** cells, usually, gametes, have single copies. Karyotypes can be used for many purposes; such as to study chromosomal aberrations, cellular function, taxonomic relationships, medicine and to gather information about past evolutionary events (***karyosystematics***).

During the chromosomal staining processes used to produce a karyotype, the staining intensity along the chromosome can vary due to localized sequence and structural differences. These banding patterns are an inherent characteristic of a chromosome and can be utilized as a diagnostic tool. Typically, karyotypes are prepared from cells that are actively undergoing mitosis. The mitotic progression is blocked in prometaphase or metaphase when chromosomes exist in their most condensed state. The cells are lysed, but the nuclei are retained intact and are subsequently treated with a chemical fixing agent. Once fixed, a number of different types of stains can be used to visualize the chromosomes.

One of the first types of chromosomal staining procedures developed was known as Q-banding, which was developed in 1970 by Torbjorn Caspersson. This technique uses the DNA-alkylating dye, quinacrine, which forms a covalent link with the DNA. Researchers noted that the staining patterns resulting from this technique were consistent and repeatable, demonstrating that banding patterns can be used to identify and characterize individual chromosomes. Giemsa dye, as shown in Figure 24.3, is more commonly used today, as it can be used with bright field microscopy and produces high detail banding patterns. A specific technique, called G-banding uses Giemsa staining following the treatment of mitotic chromosomes with the protease, trypsin. Pre-treating the sample with trypsin before staining causes the partial breakdown of chromosomal proteins leading to chromosomal relaxation. This allows more thorough staining of the chromosomes when treated with Giemsa dye. When the chromosomal region is more tightly packed into heterochromatin, it tends to stain more darkly with the Giemsa dye, than the more lightly packaged euchromatic regions. Heterochromatic regions tend to have higher A-T content and don't contain as many gene regions as euchromatic regions. Euchromatic regions stain more lightly with G-banding. Other types of staining with Giemsa, include R-banding or Reverse-banding, which involves heating the DNA before staining. This is thought to cause the melting of A-T-rich regions, reducing the Giemsa staining, when compared to G-C-rich, gene-containing regions of the chromosomes.

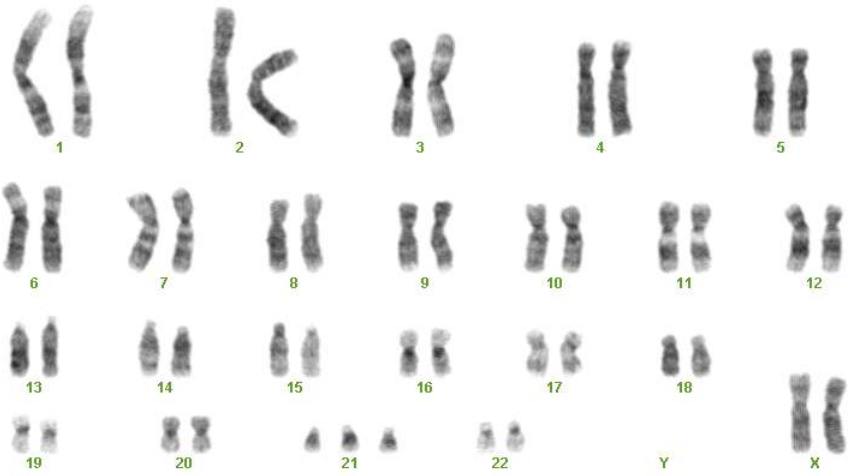


Figure 23.1.4: Female Down Syndrome Karyotype Demonstrating Trisomy 21. (Karyotype prepared by Fatma Soylemez). Image from: Soylemez, F. (2014) Chapter in Health Problems in Down Syndrome.

When visualizing a karyotype, the chromosomal images are aligned so that heterologous chromosomes are paired together and positioned such that the p-arm (short arm) is on top and the q-arm (long arm) points downward. Karyotypes can be used to quickly identify gross chromatic abnormalities that are larger than a few megabases in difference. This includes abnormalities such as ***aneuploidy*** (the addition or absence of an entire chromosome), or ***translocations*** (the transfer of part of a chromosome to a neighboring chromosome), as shown in Figure 23.1.4 and Figure 23.1.5.

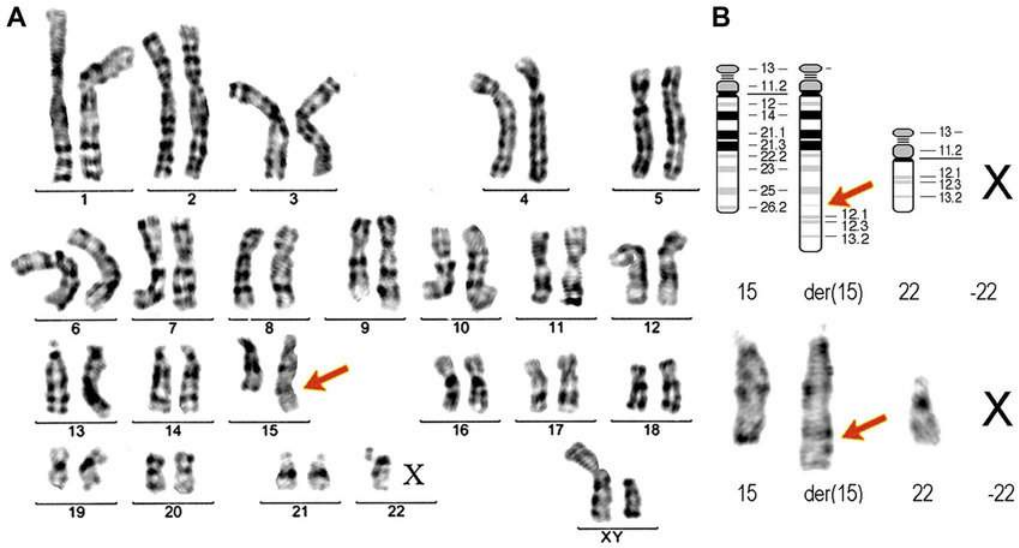


Figure 23.1.5: Karyotype of a Translocation (A) G-banding karyotype of a patient with t(15;22) translocation. Chromosome 22 monosomy with an unbalanced translocation is observed. The red arrows indicate the abnormal chromosomes. X indicated total monosomy. Original magnification, x110. (B) Ideogram and partial G-banding karyotype showing a more defined image of the translocation of the q-arm of chromosome 22 to chromosome 15. Image from: Gug, C., et al (2018) Exp & Ther Med 16(4) 6609

The telomeric regions of chromosomes can also be identified using fluorescent staining techniques, as shown in Figure 23.1.6. The structure of telomeric chromosomal regions is described in section 24.3.

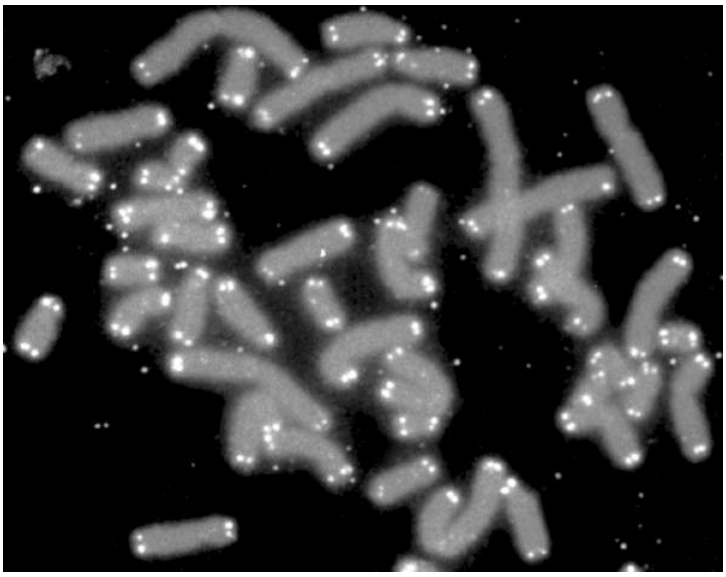


Figure 23.1.6: Telomere Staining. Image from: U.S. Department of Energy Human Genome Program

More recently, techniques such as chromosome painting, use fluorescently labeled probes to hybridize with specific chromosomes or even specific gene regions of a chromosome. Karyotypes originating from this technique are called spectral karyotypes. This technique can be especially useful in identifying translocations that have occurred in human cells, as shown in Figure 23.1.7.

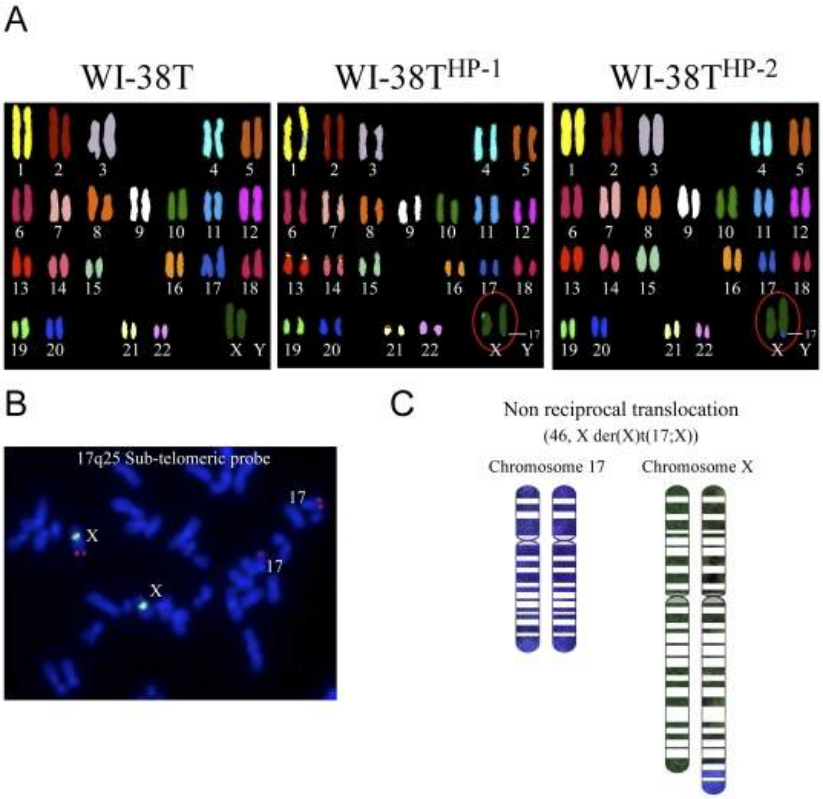


Figure 23.1.7: Spectral karyotyping (SKY) analysis of a fibroblast cell line derived from lung tissue (WI-38T) Different strains of WI-38T were developed and two showed high proliferation rates typical of the premalignant state. (A) shows SKY analysis for WI-38T and the two high proliferation strains. Both high proliferation strains show translocations of chromosome 17 to the X chromosome. (B) shows specific staining of WI-38T(HP-1) for the 17q25 region of chromosome 17 (pink) which is visible on both copies of chromosome 17 and also on one of the X chromosomes, indicating the translocation. (C) shows a graphic image of the translocation. Image from: Buganim, Y., et al (2010) PLoS ONE 5(3) e9657

23.1.3: References

Wikipedia contributors. (2021, May 23). Genetic linkage. In *Wikipedia, The Free Encyclopedia*. Retrieved 00:19, July 23, 2021, from https://en.Wikipedia.org/w/index.php?title=Genetic_linkage&oldid=1024743818

Wikipedia contributors. (2021, July 6). Karyotype. In *Wikipedia, The Free Encyclopedia*. Retrieved 01:16, July 23, 2021, from https://en.Wikipedia.org/w/index.php?title=Karyotype&oldid=1032263708

23.2: DNA Transposable Elements

23.2.1: Introduction

Eukaryotic genomes contain an abundance of repeated DNA, and some repeated sequences are mobile. **Transposable elements** (TEs) are defined as DNA sequences that can move from one location to another in the genome. TEs have been identified in all organisms, prokaryotic and eukaryotic, and can occupy a high proportion of a species' genome. For example, transposable elements comprise approximately 10% of several fish species, 12 % of the *C. elegans* genome, 37% of the mouse genome, 45% of the human genome, and up to >80% of the genome of some plants like maize. From bacteria to humans, transposable elements have accumulated over time and continue to shape genomes through their mobilization.

TEs were discovered by Barbara McClintock during experiments conducted in 1944 on maize. Since they appeared to influence phenotypic traits, she named them **controlling elements**. However, her discovery was met with less than enthusiastic reception by the genetic community. Her presentation at the 1951 Cold Spring Harbor Symposium was not understood and at least not very well received. She had no better luck with her follow-up publications and after several years of frustration decided not to publish on the subject for the next two decades. Not for the first time in the history of science, an unappreciated discovery was brought back to life after some other discovery has been made. In this case, it was the discovery of insertion sequences (IS) in bacteria by Szybalski group in the early 1970s. In the original paper, they wrote: "Genetic elements were found in higher organisms which appear to be readily transposed from one to another site in the genome. Such elements, identifiable by their controlling functions, were described by McClintock in maize. It is possible that they might be somehow analogous to the presently studied IS insertions". The importance of McClintock's original work was eventually appreciated by the genetic community with numerous awards, including 14 honorary doctoral degrees and a Nobel Prize in 1983 "for her discovery of mobile genetic elements". Her picture is shown in Figure 23.2.1.

Figure 23.2.1: Barbara McClintock (1902-1992). The photo was taken in her lab at the Department of Genetics, Carnegie Institution at Cold Spring Harbor, New York. This photo was taken when McClintock received the American Association of University Women Achievement Award in 1947 for her work on cytogenetics. She received the Nobel Prize in Physiology or Medicine in 1983 for her discovery of mobile genetic elements. Image from: Smithsonian Institution

The mobilization of TEs is termed **transposition or retrotransposition**, depending on the nature of the intermediate used for mobilization. There are several ways in which the activity of TEs can positively and negatively impact a genome; for example, TE mobilization can promote gene inactivation, modulate gene expression or induce illegitimate recombination. Thus, TEs have played a significant role in genome evolution. For example, DNA transposons can inactivate or alter the expression of genes by insertion within introns, exons, or regulatory regions. In addition, TEs can participate in the reorganization of a genome by the mobilization of non-transposon DNA or by acting as recombination substrates. This recombination would occur by homology between two sequences of a transposon located in the same or different chromosomes, which could be the origin of several types of chromosome alterations. Indeed, TEs can participate in the loss of genomic DNA by internal deletions or other mechanisms.

The reduction in fitness suffered by the host due to transposition ultimately affects the transposon, since host survival is critical to the perpetuation of the transposon. Therefore, strategies have been developed by host and transposable elements to minimize the

deleterious impact of transposition, and to reach equilibrium. For example, some transposons tend to insert in nonessential regions in the genome, such as heterochromatic regions, where insertions will likely have a minimal deleterious impact. In addition, they might be active in the germ line or embryonic stage, where most deleterious mutations can be selected against during fecundation or development, allowing only non-deleterious or mildly deleterious insertions to pass to successive generations. New insertions may also occur within an existing genomic insertion to generate an inactive transposon or can undergo self-regulation by **overproduction-inhibition**. On the other hand, host organisms have developed different mechanisms of defense against high rates of transposon activity, including DNA-methylation to reduce TE expression, several RNA interference-mediated mechanisms, mainly in the germ line, or through the inactivation of transposon activity by the action of specific proteins.

In some cases, transposable elements have been “domesticated” by the host to perform a specific function in the cell. A well-known example is RAG proteins, which participate in V(D)J recombination during antibody class switching, and exhibit a high similarity to DNA transposons, from which these proteins appear to be derived. Another example is the centromeric protein CENP-B, which seems to have originated from the pogo-like transposon. The analogous human mariner *Himar1* element has been incorporated into the SETMAR gene, which consists of the histone H3 methylase gene and the Himar1 transposase domain. This gene is involved in the non-homologous end-joining pathway of DNA repair and has been shown to confer resistance to ionizing radiation. From a genome-wide view, it has been estimated that ~25% of human promoter regions and ~4% of human exons contain sequences derived from TEs. Thus, we are likely underestimating the rate of domestication events in mammalian genomes.

The first TE classification system was proposed by Finnegan in 1989 and distinguished two classes of TEs characterized by their transposition intermediate: RNA (class I or retrotransposons) or DNA (class II or DNA transposons). The transposition mechanism of class I is commonly called “copy and paste” and that of class II, “cut and paste.” In 2007 Wicker et al. proposed a hierarchical classification based on TEs structural characteristics and mode of replication, as shown in Figure 23.2.2

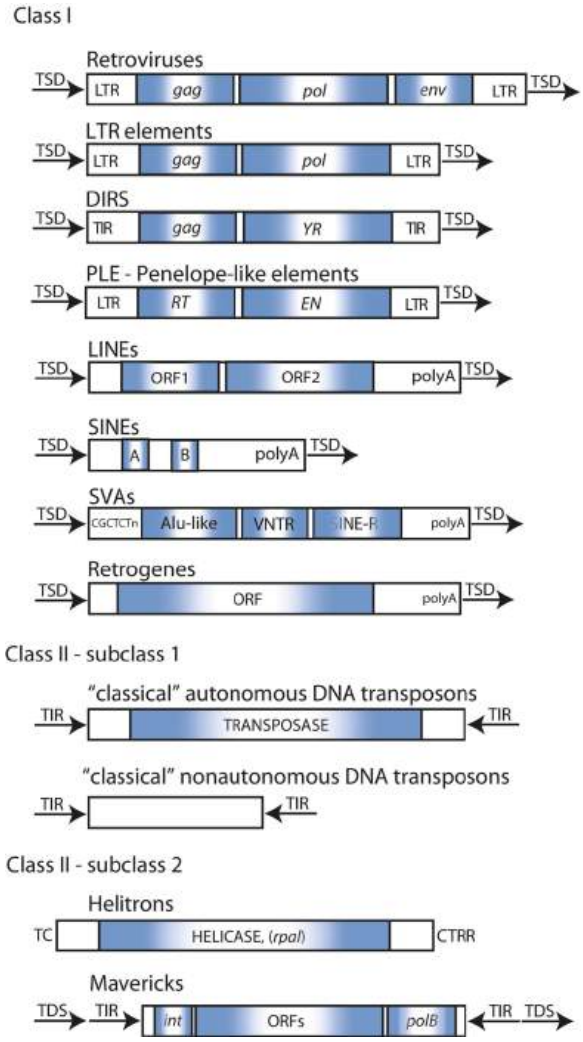


Figure 23.2.2: Structures of eukaryotic mobile elements. See text for detailed discussion Image from Makalowski, W, et. al. (2019)

23.2.2: Class I: Mobile Elements

As mentioned above, class I TEs transpose through an RNA intermediary. The RNA intermediate is transcribed from genomic DNA and then reverse-transcribed into DNA by a TE-encoded reverse transcriptase (RT), followed by reintegration into a genome. Each replication cycle produces one new copy, and as a result, class I elements are the major contributors to the repetitive fraction in large genomes. Retrotransposons are divided into five orders: LTR retrotransposons, DIRS-like elements, Penelope-like elements (PLEs), LINEs (*l*ong *in*terspersed *e*lements), and SINEs (*s*hort *in*terspersed *e*lements). This scheme is based on the mechanistic features, organization, and reverse transcriptase phylogeny of these retroelements. Accidentally, the retrotranscriptase coded by an autonomous TE can reverse-transcribe another RNA present in the cell, e.g., mRNA, and produce a retrocopy of it, which in most cases results in a pseudogene.

The LTR retrotransposons are characterized by the presence of *l*ong *t*erminal *r*epeats (LTRs) ranging from several hundred to several thousand base pairs. Both exogenous retroviruses and LTR retrotransposons contain a *gag* gene that encodes a viral particle coat and a *pol* gene that encodes a reverse transcriptase, ribonuclease H, and an integrase, which provide the enzymatic machinery for reverse transcription and integration into the host genome. Reverse transcription occurs within the viral or viral-like particle (GAG) in the cytoplasm, and it is a multi-step process. Unlike LTR retrotransposons, exogenous retroviruses contain an *env* gene, which encodes an envelope that facilitates their migration to other cells. Some LTR retrotransposons may contain remnants of an *env* gene, but their insertion capabilities are limited to the originating genome. This would rather suggest that they originated in exogenous retroviruses by losing the *env* gene. However, there is evidence that suggests the contrary, given that LTR retrotransposons can acquire the *env* gene and become infectious entities. Presently, most of the LTR sequences (85%) in the human genome are found only as isolated LTRs, with the internal sequence being lost most likely due to homologous recombination between flanking LTRs. Interestingly, LTR retrotransposons target their reinsertion to specific genomic sites, often around genes, with putative important functional implications for a host gene. It is estimated that 450,000 LTR copies make up about 8% of our genome. LTR retrotransposons inhabiting large genomes, such as maize, wheat, or barley, can contain thousands of families. However, despite the diversity, very few families comprise most of the repetitive fraction in these large genomes. Notable examples are Angela (wheat), BARE1 (barley), Opie (maize), and Retrosor6 (sorghum).

The DIRS order clusters structurally diverged groups of transposons that possess a tyrosine recombinase (YR) gene instead of an integrase (INT) and do not form target site duplications (TSDs). Their termini resemble either split direct repeats (SDR) or inverted repeats. Such features indicate a different integration mechanism than that of other class I mobile elements. DIRS were discovered in the slime mold (*Dictyostelium discoideum*) genome in the early 1980s, and they are resent in all major phylogenetic lineages including vertebrates. It has been shown that they are also common in hydrothermal vent organisms.

Another order, termed *Penelope*-like elements (PLE), has wide, though patchy distribution from amoebae and fungi to vertebrates with copy numbers up to thousands per genome. Interestingly, no PLE sequences have been found in mammalian genomes, and apparently, they were lost from the genome of *C. elegans*. Although PLEs with an intact ORF have been found in several genomes, including *Ciona* and *Danio*, the only transcriptionally active representative, *Penelope*, is known from *Drosophila virilis*. It causes the hybrid dysgenesis syndrome characterized by the simultaneous mobilization of several unrelated TE families in the progeny of dysgenic crosses. It seems that *Penelope* invaded *D. virilis* quite recently, and its invasive potential was demonstrated in *D. melanogaster*. PLEs harbor a single ORF that codes for a protein containing reverse transcriptase (RT) and endonuclease (EN) domains. The PLE RT domain more closely resembles telomerase than the RT from LTRs or LINEs. The EN domain is related to GIY-YIG intron-encoded endonucleases. Some PLE members also have LTR-like sequences, which can be in a direct or an inverse orientation, and have a functional intron.

LINEs do not have LTRs; however, they have a poly-A tail at the 3′ ends and are flanked by the TSDs. They comprise about 21% of the human genome and among them L1 with about 850,000 copies is the most abundant and best-described LINE family. L1 is the only LINE retroposon still active in the human genome. In the human genome, there are two other LINE-like repeats, L2 and L3, distantly related to L1. A contrasting situation has been noticed in the malaria mosquito *Anopheles gambiae*, where around 100 divergent LINE families compose only 3% of its genome. LINEs in plants, e.g., Cin4 in maize and Ta11 in *Arabidopsis thaliana*, seem rare as compared with LTR retrotransposons. A full copy of mammalian L1 is about 6 kb long and contains a PolII promoter and two ORFs. The ORF1 codes for a non-sequence-specific RNA binding protein that contains zinc finger, leucine zipper, and coiled-coil motifs. The ORF1p functions as a chaperone for the L1 mRNA. The second ORF encodes an endonuclease, which makes a single-stranded nick in the genomic DNA, and a reverse transcriptase, which uses the nicked DNA to prime reverse transcription of LINE RNA from the 3′ end. Reverse transcription is often unfinished, leaving behind fragmented copies of LINE elements; hence most of the L1-derived repeats are short, with an average size of 900 bp. LINEs are part of the CR1 clade, which

has members in various metazoan species, including fruit flies, mosquito, zebrafish, pufferfish, turtles, and chicken. Because they encode their own retrotransposition machinery, LINE elements are regarded as autonomous retrotransposons.

SINEs evolved from RNA genes, such as 7SL and tRNA genes. By definition, they are short, up to 1000 base pairs long. They do not encode their own retrotranscription machinery and are considered nonautonomous elements and in most cases are mobilized by the L1 machinery. The outstanding member of this class from the human genome is the Alu repeat, which contains a cleavage site for the *Alu*I restriction enzyme that gave its name. With over a million copies in the human genome, Alu is probably the most successful transposon in the history of life. Primate-specific Alu and its rodent relative B1 have limited phylogenetic distribution suggesting their relatively recent origins. The mammalian-wide interspersed repeats (MIRs), by contrast, spread before eutherian radiation, and their copies can be found in different mammalian groups including marsupials and monotremes. SVA elements are unique primate elements due to their composite structure. They are named after their main components: SINE, VNTR (a variable number of tandem repeats), and Alu. Usually, they contain the hallmarks of the retroposition, i.e., they are flanked by TSDs and terminated by a poly(A) tail. It seems that SVA elements are nonautonomous retrotransposons mobilized by L1 machinery, and they are thought to be transcribed by RNA polymerase II. SVAs are transpositionally active and are responsible for some human diseases. They originated less than 25 million years ago, and they form the youngest retrotransposon family with about 3000 copies in the human genome.

Retro(pseudo)genes are a special group of retroposed sequences, which are products of reverse transcription of a spliced (mature) mRNA. Hence, their characteristic features are an absence of promoter sequence and introns, the presence of flanking direct repeats, and a 3′-end polyadenosine tract. Processed pseudogenes, as sometimes retropseudogenes are called, have been generated in vitro at a low frequency in the human HeLa cells via mRNA from a reporter gene. The source of the reverse transcription machinery in humans and other vertebrates seems to be active L1 elements. However, not all retroposed messages have to end up as pseudogenes. About 20% of mammalian protein-encoding genes lack introns in their ORFs. It is conceivable that many genes lacking introns arose by retroposition. Some genes are known to be retroposed more often than others. For instance, in the human genome, there are over 2000 retropseudogenes of ribosomal proteins. A genome-wide study showed that the human genome harbors about 20,000 pseudogenes, 72% of which most likely arose through retroposition. Interestingly, the vast majority (92%) of them are quite recent transpositions that occurred after primate/rodent divergence. Some of the retroposed genes may undergo quite complicated evolutionary paths. An example could be the RNF13B retrogene, which replaced its own parental gene in the mammalian genomes. This retrocopy was duplicated in primates, and the evolution of this primate-specific copy was accompanied by the exaptation of two TEs, Alu and L1, and intron gain via changing a part of the coding sequence into an intron leading to the origin of a functional, primate-specific retrogene with two splicing variants.

23.2.3: Class II: Mobile Elements

Class II elements move by a conservative cut-and-paste mechanism; the excision of the donor element is followed by its reinsertion elsewhere in the genome. DNA transposons are abundant in bacteria, where they are called insertion sequences, but are also present in all phyla. Two subclasses of DNA transposons have been distinguished, based on the number of DNA strands that are cut during transposition.

Classical “cut-and-paste” transposons belong to subclass I, and they are classified as the TIR order. They are characterized by terminal inverted repeats (TIR) and encode a transposase that binds near the inverted repeats and mediates mobility. This process is not usually a replicative one, unless the gap caused by excision is repaired using the sister chromatid. When inserted at a new location, the transposon is flanked by small gaps, which, when filled by host enzymes, cause duplication of the sequence at the target site. The length of these TSDs is characteristic of particular transposons. Nine superfamilies belong to the TIR order, including *Tc1-Mariner*, *Merlin*, *Mutator*, and *PiggyBac*. The second-order Crypton consists of a single superfamily of the same name. Originally thought to be limited to fungi, now it is clear that they have a wide distribution, including animals and heterokonts. A heterogeneous, small, nonautonomous group of elements MITEs also belong to the TIR order, which in some genomes amplified to thousands of copies, e.g., *Stowaway* in the rice genome, *Tourist* in most bamboo genomes, or *Galluhop* in the chicken genome.

Subclass II includes two orders of TEs that, just as those from subclass I, do not form RNA intermediates. However, unlike “classical” DNA transposons, they replicate without double-strand cleavage. Helitrons replicate using a rolling-circle mechanism, and their insertion does not result in the target site duplication. They encode tyrosine recombinase along with some other proteins. Helitrons were first described in plants, but they are also present in other phyla, including fungi and mammals. Mavericks are large transposons that have been found in different eukaryotic lineages excluding plants. They encode various numbers of proteins that

include DNA polymerase B and an integrase. Kapitonov and Jurka suggested that their life cycle includes a single-strand excision, followed by extrachromosomal replication and reintegration to a new location.

23.2.4: TEs are not randomly distributed in the genome

As seen in the previous section, TEs are highly diverse and in principle, every TE sequence in a genome can be affiliated to a (sub)family, superfamily, subclass, and class. This is summarized in Figure 23.2.3 However, much like the taxonomy of species, the classification of TEs is in constant flux, perpetually subject to revision due to the discovery of completely novel TE types, the introduction of new levels of granularity in the classification, and the ongoing development of methods and criteria to detect and classify TEs.

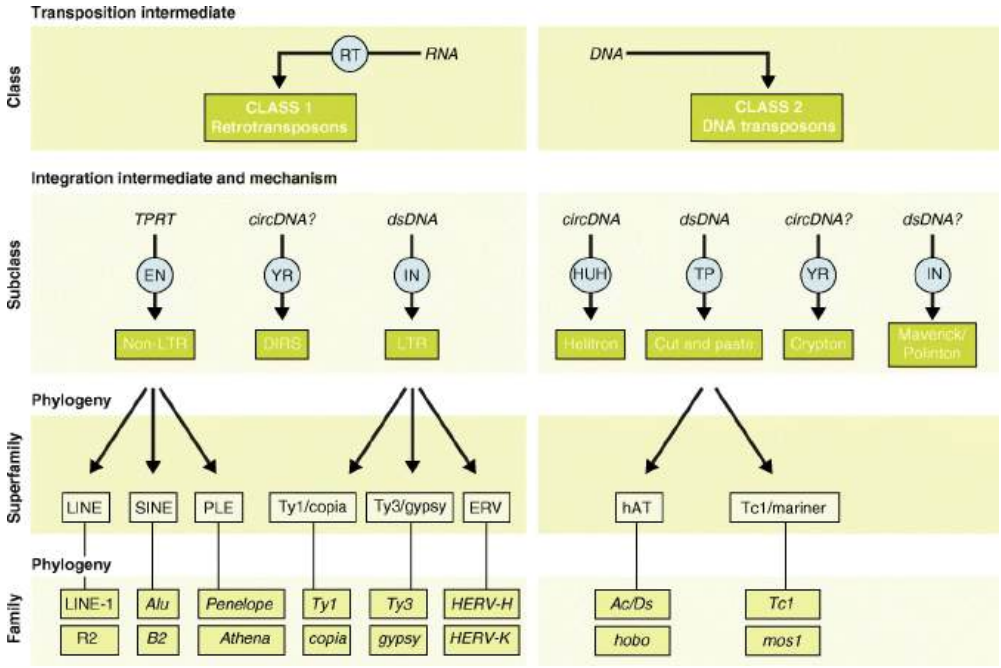


Figure 23.2.3: Classification of Eukaryotic Transposable Elements. This schematic shows the key features and relationships between TE classes, subclasses, superfamilies, and families. Blue circles represent TE-encoded enzymes. circDNA circular DNA intermediate, DIRS Dictyostelium repetitive sequence, dsDNA linear double-stranded DNA intermediate, EN endonuclease, IN integrase, PLEs Penelope-like elements, HUH, Rep/Helicase protein with HUH endonuclease activity, RT reverse transcriptase, TP transposase, TPRT target primed reverse transcription, YR tyrosine recombinase. Image from: Bourque, G., et. al. (2018) Genetic Biology 19:199

The genome may be viewed as an ecosystem inhabited by diverse communities of TEs, which seek to propagate and multiply through sophisticated interactions with each other and with other components of the cell. These interactions encompass processes familiar to ecologists, such as parasitism, cooperation, and competition. Thus, it is perhaps not surprising that TEs are rarely, if ever, randomly distributed in the genome, as shown in Figure 23.2.4 TEs exhibit various levels of preference for insertion within certain features or compartments of the genome. These are often guided by opposite selective forces, a balancing act of facilitating future propagation while mitigating deleterious effects on host cell function. At the most extreme end of the site-selection spectrum, many elements have evolved mechanisms to target specific loci where their insertions are less detrimental to the host but favorable for their propagation. For instance, several retrotransposons in species as diverse as slime mold and budding and fission yeast have evolved independently, but convergently, the capacity to target the upstream regions of genes transcribed by RNA polymerase III, where they do not appear to affect host gene expression but retain the ability to be transcribed themselves.

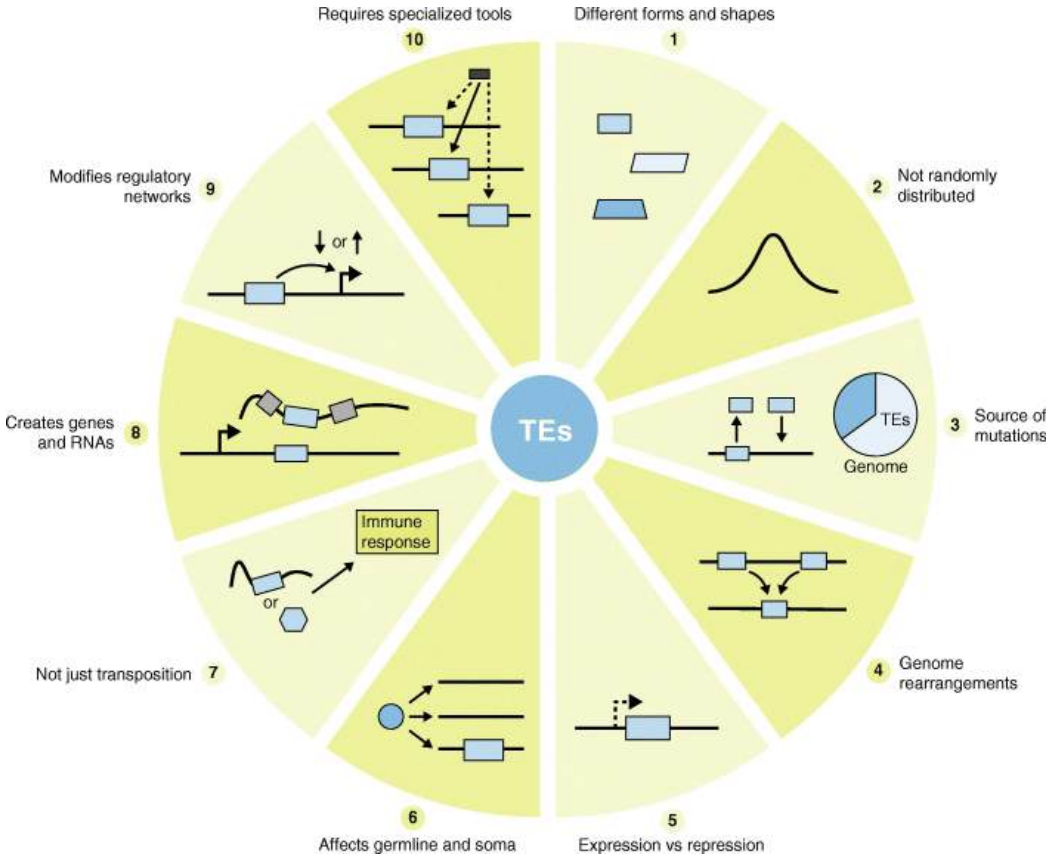


Figure 23.2.4: Ten things you should know about transposable elements (TEs). Examples of how TEs can impact genomes in direct and indirect ways. Blue boxes represent TEs, gray boxes represent canonical exons, and the black box represents a sequencing read. Right-angled arrows represent gene sor TE promoters. Image from: Bourque, G., et. al. (2018) Genetic Biology 19:199

23.2.5: TEs are an extensive source of mutations and genetic polymorphisms

TEs occupy a substantial portion of the genome of a species, including a large fraction of the DNA unique to that species. In maize, where Barbara McClintock did her seminal work, an astonishing 60 to 70% of the genome is comprised of LTR retrotransposons, many of which are unique to this species or its close wild relatives, but the less prevalent DNA transposons are currently the most active and mutagenic (Fig. 24.2.4). Similarly, the vast majority of TE insertions in *Drosophila melanogaster* are absent at the orthologous site in its closest relative *D. simulans* (and vice versa), and most are not fixed in the population. Many TE families are still actively transposing and the process is highly mutagenic; more than half of all known phenotypic mutants of *D. melanogaster* isolated in the laboratory are caused by spontaneous insertions of a wide variety of TEs. Transposition events are also common and mutagenic in laboratory mice, where the ongoing activity of several families of LTR elements are responsible for 10–15% of all inherited mutant phenotypes. This contribution of TEs to genetic diversity may be underestimated, as TEs can be more active when organisms are under stress, such as in their natural environment.

Because TE insertions rarely provide an immediate fitness advantage to their host, those reaching fixation in the population do so largely by genetic drift and are subsequently eroded by point mutations that accumulate neutrally. Over time, these mutations result in TEs that can no longer encode transposition enzymes and produce new integration events. For instance, our (haploid) genome contains ~ 500,000 L1 copies, but more than 99.9% of these L1 copies are fixed and no longer mobile due to various forms of mutations and truncations. It is estimated that each person carries a set of ~ 100 active L1 elements, and most of these are young insertions still segregating within the human population. Thus, as for any other organism, the ‘reference’ human genome sequence does not represent a comprehensive inventory of TEs in humans. Thousands of ‘non-reference’, unfixed TE insertions have been cataloged through whole genome sequencing and other targeted approaches. On average, any two human haploid genomes differ by approximately a thousand TE insertions, primarily from the L1 or Alu families. The number of TE insertion polymorphisms in a species with much higher TE activity such as maize dwarfs the number in humans.

If TEs bring no immediate benefit to their host and are largely decaying neutrally once inserted, how do they persist in evolution? One key to this conundrum is the ability of TEs not only to propagate vertically but also horizontally between individuals and

species. There is now a large body of evidence supporting the idea that horizontal transposon transfer is a common phenomenon that affects virtually every major type of TE and all branches of the tree of life. While the cellular mechanisms underlying horizontal transposon transfer remain murky, it is increasingly apparent that the intrinsic mobility of TEs and ecological interactions between their host species, including those with pathogens and parasites, facilitate the transmission of elements between widely diverged taxa.

TEs are associated with genome rearrangements and unique chromosome feature

Transposition represents a potent mechanism of genome expansion that over time is counteracted by the removal of DNA via deletion. The balance between the two processes is a major driver in the evolution of genome size in eukaryotes. Several studies have demonstrated the impact and range of this shuffling and cycling of genomic content on the evolution of plant and animal genomes. Because the insertion and removal of TEs are often imprecise, these processes can indirectly affect surrounding host sequences. Some of these events occur at high enough frequency to result in vast amounts of duplication and reshuffling of host sequences, including genes and regulatory sequences. For example, a single group of DNA transposons (MULEs) has been responsible for the capture and reshuffling of ~ 1,000 gene fragments in the rice genome. Such studies have led to the conclusion that the rate at which TEs transpose, which is in part under host control, is an important driver of genome evolution.

In addition to rearrangements induced as a byproduct of transposition, TEs can promote genomic structural variation long after they have lost the capacity to mobilize. In particular, recombination events can occur between the highly homologous regions dispersed by related TEs at distant genomic positions and result in large-scale deletions, duplications, and inversions (Fig. 24.2.4). TEs also provide regions of microhomology that predispose to template switching during repair of replication errors leading to another source of structural variants. These non-transposition-based mechanisms for TE-induced or TE-enabled structural variation have contributed substantially to genome evolution. These processes can also make the identification of actively transposing elements more difficult in population studies that infer the existence of active elements through the detection of non-reference insertions.

TEs also contribute to specialized chromosome features. An intriguing example is in *Drosophila*, where LINE-like retrotransposons form and maintain the telomeres in replacement of the telomerase enzyme which has been lost during dipteran evolution. This domestication event could be viewed as a replay of what might have happened much earlier in eukaryotic evolution to solve the ‘end problem’ created by the linearization of chromosomes. Indeed, the reverse transcriptase component of telomerase is thought to have originated from an ancient lineage of retroelements. TE sequences and domesticated transposase genes also play structural roles at centromeres.

There is an intricate balance between TE expression and repression

To persist in evolution, TEs must strike a delicate balance between expression and repression (Fig. 24.2.4). Expression should be sufficient to promote amplification, but not so vigorous as to lead to a fitness disadvantage for the host that would offset the benefit to the TE of increased copy numbers. This balancing act may explain why TE-encoded enzymes are naturally suboptimal for transposition and why some TEs have evolved self-regulatory mechanisms controlling their own copy numbers. A variety of host factors are also employed to control TE expression, which includes a variety of small RNA, chromatin, and DNA modification pathways, as well as sequence-specific repressors such as the recently profiled KRAB zinc-finger proteins. However, many of these silencing mechanisms must be at least partially released to permit the developmental regulation of host gene expression programs, particularly during early embryonic development. For example, genome-wide loss of DNA methylation is necessary to reset imprinted genes in primordial germ cells. This affords TEs an opportunity, as reduced DNA methylation often promotes TE expression. Robust expression of a TE in the germ lineage (but not necessarily in the gametes themselves) is often its own downfall. In one example of a clever trick employed by the host, TE repression is relieved in a companion cell derived from the same meiotic product as flowering plant sperm. However, this companion cell does not contribute genetic material to the next generation. Thus, although TEs transpose in a meiotic product, the events are not inherited. Instead, TE activity in the companion cell may further dampen TE activity in sperm via the import of TE-derived small RNAs.

Another important consequence of the intrinsic expression/repression balance is that the effects of TEs on a host can vary considerably among tissue types and stages of an organism’s life cycle. From the TE’s perspective, an ideal scenario is to be expressed and active in the germline, but not in the soma, where expression would gain the TE no advantage, only disadvantages. This is indeed observed among many species, with ciliates representing an extreme example of this division—TEs are actively deleted from the somatic macronucleus but retained in the micronucleus, or germline. Another example is the P-elements in *Drosophila*, which are differentially spliced in the germline versus soma. Many organisms, including plants, do not differentiate germ lineage cells early in development; rather, they are specified from somatic cells shortly before meiosis commences. Thus, TEs

that transpose in somatic cells in plants have the potential to be inherited, which suggests that the interest of TEs and hosts are in conflict across many more cells and tissues than in animals with a segregated germline.

TEs are insertional mutagens in both germline and soma

Like other species, humans contend with a contingent of currently active TEs where the intrinsic balance between expression and repression is still at play. For us, this includes L1 and other mobile elements that depend on L1-encoded proteins for retrotransposition. These elements are responsible for new germline insertions that can cause genetic disease. More than 120 independent TE insertions have been associated with human disease. The rate of de novo germline transposition in humans is approximately one in 21 births for *Alu* and one in 95 births for L1.

Historically, little attention has been given to transposition in somatic cells and its consequences, because somatic transposition may be viewed as an evolutionary dead-end for the TE with no long-term consequences for the host species. Yet, there is abundant evidence that TEs are active in somatic cells in many organisms (Fig. 24.2.4). In humans, L1 expression and transposition have been detected in a variety of somatic contexts, including early embryos and certain stem cells. There is also a great deal of interest in mobile element expression and activity in the mammalian brain, where L1 transposition has been proposed to diversify neuronal cell populations. One challenge for assessing somatic activity has rested with the development of reliable single-cell insertion site mapping strategies.

Somatic activity has also been observed in human cancers, where tumors can acquire hundreds of new L1 insertions. Just like for human polymorphisms, somatic activity in human cancers is caused by small numbers of so-called ‘hot’ L1 loci. The activities of these master copies vary depending on the individual, tumor type, and timeframe in the clonal evolution of the tumor. Some of these de novo L1 insertions disrupt critical tumor suppressors and oncogenes and thus drive cancer formation, although the vast majority appear to be ‘passenger’ mutations. Host cells have evolved several mechanisms to keep TEs in check. However, as the force of natural selection begins to diminish with age and completely drops in post-reproductive life, TEs may become more active.

TEs can be damaging in ways that do not involve transposition

TEs are best known for their mobility, and their ability to transpose to new locations. While the breakage and insertion of DNA associated with transposition represent an obvious source of cell damage, this is not the only or perhaps even the most common mechanism by which TEs can be harmful to their host. Reactivated transposons harm the host in multiple ways. First, de-repression of transposon loci, including their own transcription, may interfere with transcription or processing of host mRNAs through a myriad of mechanisms. Genome-wide transcriptional de-repression of TEs has been documented during replicative senescence of human cells and several mouse tissues, including the liver, muscle, and brain. De-repression of LTR and L1 promoters can also cause oncogene activation in cancer. Second, TE-encoded proteins such as the endonuclease activity of L1 ORF2p can induce DNA breaks and genomic instability. Third, accumulation of RNA transcripts and extrachromosomal DNA copies derived from TEs may trigger an innate immune response leading to autoimmune diseases and sterile inflammation (Fig. 24.2.4). Activation of interferon response is now a well-documented property of transcripts derived from endogenous retroviruses and may give immunotherapies a boost in identifying and attacking cancer cells. The relative contribution of all the above mechanisms in organismal pathologies remains to be determined.

Following transcription (and sometimes splicing) of TEs, the next step in the process involves the translation of the encoded proteins and, for retroelements, reverse transcription of the TEs into cDNA substrates suitable for transposition. Once engaged by a TE-encoded reverse transcriptase protein, the resulting cytosolic DNAs and RNA:DNA hybrids can alert inflammatory pathways. An example of this is seen in patients with Aicardi–Goutières syndrome, where the accumulation of TE-derived cytosolic DNA is due to mutations in pathways that normally block TE processing or degrade TE-derived DNA. Although not all TEs encode functional proteins, some do, including a few endogenous retroviruses capable of producing Gag, Pol, or envelope (Env) proteins. Overexpression of these Env proteins can be cytotoxic and has been linked to at least two neurodegenerative diseases, multiple sclerosis, and amytrophic lateral sclerosis. Small accessory proteins produced by the youngest human endogenous retrovirus (HERV) group, HERV-K (HML-2), may play a role in some cancers but the evidence remains circumstantial.

Key coding and non-coding RNAs are derived from TEs

Although usually detrimental, there is growing evidence that TE insertions can provide the raw material for the emergence of protein-coding genes and non-coding RNAs, which can take on important and, in some cases essential, cellular function (Fig. 24.2.4). The process of TE gene ‘domestication’ or exaptation over evolutionary time contributes to both deeply conserved functions and more recent, species-specific traits. Most often, the ancestral or a somewhat modified role of a TE-encoded gene is harnessed by the host and conserved, while the rest of the TE sequence, and hence its ability to autonomously transpose, has been

lost. Spectacular examples of deeply conserved TE-derived genes are *Rag1* and *Rag2*, that catalyze V(D)J somatic recombination in the vertebrate immune system. Both genes, and probably the DNA signals they recognize, were derived from an ancestral DNA transposon around 500 million years ago. Indeed, DNA transposases have been co-opted multiple times to form new cellular genes.

The *gag* and *env* genes of LTR retrotransposons or endogenous retroviruses (ERVs) have also been domesticated numerous times to perform functions in placental development, contribute to host defense against exogenous retroviruses, act in brain development, and play other diverse roles. One of the most intriguing examples of TE domestication is the repeated, independent capture of ERV *env* genes, termed *syncytins*, which appear to function in placentation by facilitating cell–cell fusion and syncytiotrophoblast formation. Notably, one or more *syncytin* genes have been found in virtually every placental mammalian lineage where they have been sought, strongly suggesting that ERVs have played essential roles in the evolution and extreme phenotypic variability of the mammalian placenta. Another example of a viral-like activity re-purposed for host cell function is provided by the neuronal *Arc* gene, which arose from the *gag* gene from a LTR retrotransposon domesticated in the common ancestor of tetrapod vertebrates. Genetic and biochemical studies of murine Arc show that it is involved in memory and synaptic plasticity and has preserved most of the ancestral activities of Gag, including the packaging and intercellular trafficking of its own RNA. Remarkably, flies appear to have independently evolved a similar system of trans-synaptic RNA delivery involving a *gag*-like protein derived from a similar yet distinct lineage of LTR retrotransposons. Thus, the biochemical activities of TE-derived proteins have been repeatedly co-opted during evolution to foster the emergence of convergent cellular innovations in different organisms.

TEs can donate their own genes to the host, but they can also add exons and rearrange and duplicate existing host genes. In humans, intronic *Alu* elements are particularly prone to be captured as alternative exons through cryptic splice sites residing within their sequences. L1 and SVA (SINE/VNTR/Alu) elements also contribute to exon shuffling through transduction events of adjacent host sequences during their mobilization. The reverse transcriptase activity of retroelements is also responsible for the trans-duplication of cellular mRNAs to create ‘processed’ retrogenes in a wide range of organisms. The L1 enzymatic machinery is thought to be involved in the generation of tens of thousands of retrogene copies in mammalian genomes, many of which remain transcribed and some of which have acquired new cellular functions. This is a process still actively shaping our genomes; it has been estimated that 1 in every 6000 humans carries a novel retrogene insertion.

TEs also make substantial contributions to the non-protein coding functions of the cell. They are major components of thousands of long non-coding RNAs in human and mouse genomes, often transcriptionally driven by retroviral LTRs. Some of these TE-driven lncRNAs appear to play important roles in the maintenance of stem cell pluripotency and other developmental processes. Many studies have demonstrated that TE sequences embedded within lncRNAs and mRNAs can directly modulate RNA stability, processing, or localization with important regulatory consequences. Furthermore, TE-derived microRNAs and other small RNAs processed from TEs can also adopt regulatory roles serving host cell functions. The myriad of mechanisms by which TEs contribute to coding and non-coding RNAs illustrate the multi-faceted interactions between these elements and their host.

TEs contribute cis-regulatory DNA elements and modify transcriptional networks

Cis-regulatory networks coordinate the transcription of multiple genes that function in concert to orchestrate entire pathways and complex biological processes. In line with Barbara McClintock’s insightful predictions, there is now mounting evidence that TEs have been a rich source of material for the modulation of eukaryotic gene expression (Fig. 24.2.4). Indeed, TEs can disperse vast amounts of promoters and enhancers, transcription factor binding sites, insulator sequences, and repressive elements. The varying coat colors of agouti mice provide a striking example of a host gene controlling coat color whose expression can be altered by the methylation levels of a TE upstream of its promoter. In the oil palm, the methylation level of a TE that sits within a gene important for flowering ultimately controls whether or not the plants bear oil-rich fruit.

As TE families typically populate a genome as a multitude of related copies, it has long been postulated that they have the potential to donate the same cis-regulatory module to ‘wire’ batteries of genes dispersed throughout the genome. An increasing number of studies support this model and suggest that TEs have provided the building blocks for the assembly and remodeling of cis-regulatory networks during evolution, including pathways underlying processes as diverse as pregnancy, stem cell pluripotency, neocortex development, innate immunity in mammals, or the response to abiotic stress in maize. Indeed, TE sequences harbor all the necessary features of a ‘classical’ gene regulatory network. They are bound by diverse sets of transcription factors that integrate multiple inputs (activation/repression), respond to signals in both *cis* and *trans*, and are capable of co-ordinately regulating gene expression. In this context, TEs are highly suitable agents to modify biological processes by creating novel cis-regulatory circuits and fine-tuning pre-existing networks.

Outlook

As potent insertional mutagens, TEs can have both positive and negative effects on host fitness, but it is likely that the majority of TE copies in any given species—and especially those such as humans with small effective population size—have reached fixation through genetic drift alone and are now largely neutral to their host. When can we say that TEs have been co-opted for cellular function? The publication of the initial ENCODE paper, which asserted 'function for 80% of the genome', was the subject of much debate and controversy. Technically speaking, ENCODE assigned only 'biochemical' activity to this large fraction of the genome. Yet critics objected to the grand proclamations in the popular press (The Washington Post Headline: "Junk DNA concept debunked by new analysis of the human genome") and to the ENCODE consortium's failure to prevent this misinterpretation. To these critics, ignoring evolutionary definitions of function was a major misstep.

This debate can be easily extended to include TEs. TEs make up the vast majority of what is often referred to as 'junk DNA'. Today, the term is mostly used (and abused) by the media, but it has deep roots in evolutionary biology. Regardless of the semantics, what evidence is needed to assign a TE with a function? Many TEs encode a wide range of biochemical activities that normally benefit their own propagation. For example, TEs often contain promoter or enhancer elements that highjack cellular RNA polymerases for transcription and autonomous elements encode proteins with various biochemical and enzymatic activities, all of which are necessary for the transposon to replicate. Do these activities make them functional?

The vast differences in TEs between species make standard approaches to establishing their regulatory roles particularly challenging. For example, intriguing studies on the impact of HERVs, in particular HERV-H, in stem cells and pluripotency must be interpreted using novel paradigms that do not invoke deep evolutionary conservation to imply function, as these particular ERVs are absent outside of great apes. Evolutionary constraints can be measured at shorter time scales, including the population level, but this remains a statistically challenging task, especially for non-coding sequences. Natural loss-of-function alleles may exist in the human population and their effect on fitness can be studied if their impact is apparent, but these are quite rare and do not allow systematic studies. It is possible to engineer genetic knockouts of a particular human TE locus to test its regulatory role but those are restricted to in-vitro systems, especially when the orthologous TE does not exist in the model species. In this context, studying the impact of TEs in model species with powerful genome engineering tools and vast collections of mutants and other genetic resources, such as plants, fungi, and insects, will also continue to be extremely valuable.

Finally, a growing consensus is urging more rigor when assigning cellular function to TEs, particularly for the fitness benefit of the host. Indeed, a TE displaying biochemical activity (such as those bound by transcription factors or lying within open chromatin regions) cannot be equated to a TE that shows evidence of purifying selection at the sequence level or, when genetically-altered, result in a deleterious or dysfunctional phenotype. Recent advances in editing and manipulating the genome and the epigenome en masse yet with precision, including repetitive elements, offer the promise for a systematic assessment of the functional significance of TEs.

23.2.6: References

1. Munoz-Lopez, M. and Garcia-Perez, J.L. (2010) DNA Transposons: Nature and Applications in Genomics. Curr Genomics 11(2):115-128. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2874221/

2. Makalowski, W., Gotea, V., Pande, A. and Makalowska, I. (2019) Transposable Elements: Classification, Identification, and Their Use As a Tool For Comparative Genomics. In: Anisimova M. (eds) Evolutionary Genomics. Methods in Molecular Biology, vol 1910. Humana, New York, NY. https://doi.org/10.1007/978-1-4939-9074-0_6

3. Bourque, G., Burns, K.H., Gehring, M., Borbunova, V., Seluanov, A., Hammell, M., Imbeault, M., Izvak, Z., Levin, H.L., Macfarlan, T.S., Mager, D.L., Feschotte, C. (2018) Ten things you should know about transposable elements. Genome Biology 19: 199. Available at: https://genomebiology.biomedcentral.com/articles/10.1186/s13059-018-1577-z#Fig1

23.3: Chromosome Packaging

Some of the material in this chapter section comes from Chapter 8.4, Chromosomes and Chromatin, as it was important to describe it earlier in the structure/function unit. In addition, some biochemistry courses might not get to the material in a late chapter in a text. Repetition of some of the material is easier in an online textbook as well.

23.3.1: Introduction

Recall from Chapter 8, that within eukaryotic cells, DNA is organized into long linear structures called ***chromosomes***, as shown in Figure 23.3.1. A ***chromosome*** is a deoxyribonucleic acid (DNA) molecule with part or all of the genetic material (genome) of an organism. Most eukaryotic chromosomes include packaging proteins which, aided by chaperone proteins, bind to and condense the DNA molecule to prevent it from becoming an unmanageable tangle. Before typical cell division, these chromosomes are duplicated in the process of DNA replication, providing a complete set of chromosomes for each daughter cell. The replicated arms of a chromosome are called ***chromatids***. Before being separated into the daughter cells during mitosis, replicated chromatids are held together by a chromosomal structure called the ***centromere***.

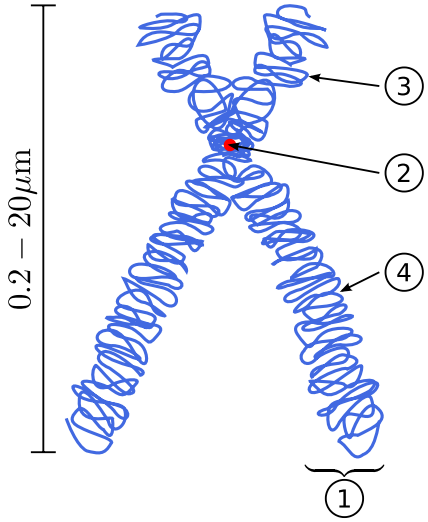


Figure 23.3.1: Diagram of Replicated and Condensed Eukaryotic Chromosome. (1) Chromatid – one of the two identical parts of the chromosome after the S phase. (2) Centromere – the point where the two chromatids are joined together. (3) Short arm is termed *p*; the Long arm is termed q. Image by: Magnus Manske, Dietzel65, and Tryphon

Eukaryotic organisms (animals, plants, fungi, and protists) store most of their DNA inside the cell nucleus as linear nuclear DNA, and some in the mitochondria as circular mitochondrial DNA or in chloroplasts as circular chloroplast DNA. In contrast, prokaryotes (bacteria and archaea) do not have organelle structures and thus, store their DNA only in a region of the cytoplasm known as the ***nucleoid region***. Prokaryotic chromosomes consist of ***double–stranded circular DNA***.

The genome of a cell is often significantly larger than the cell itself. For example, if the DNA from a human cell containing 46 chromosomes were stretched out in a line, it would extend more than 6 feet (2 meters)! How is it possible that the genetic information not only fits into the cell but fits into the cell nucleus? Eukaryota solves this problem by a combination of supercoiling and packaging DNA around the ***histone family of proteins*** (described below). Prokaryotes do not contain histones (with a few exceptions). Prokaryotes tend to compress their DNA using ***nucleoid-associated-proteins (NAPs)*** and ***supercoiling*** (Figure 24.3.2).

23.3.2: Supercoiling

DNA supercoiling refers to the over- or under-winding of a DNA strand, and is an expression of the strain on that strand, as shown in Figure 23.3.2 Supercoiling is important in many biological processes, such as compacting DNA, and regulating access to the genetic code. DNA supercoiling strongly affects DNA metabolism and possibly gene expression. Additionally, certain enzymes such as ***topoisomerases*** can change DNA topology to facilitate functions such as DNA replication or transcription.

In a “relaxed” double-helical segment of B-DNA, the two strands twist around the helical axis once every 10.4–10.5 base pairs of sequence. Adding or subtracting twists, as some enzymes can do, impose strain. If a DNA segment under twist strain were closed into a circle by joining its two ends and then allowed to move freely, the circular DNA would contort into a new shape, such as a

simple figure-eight (Figure 23.3.2). Such a contortion is a ***supercoil***. The noun form "supercoil" is often used in the context of DNA topology.

Figure 23.3.2: DNA Supercoiling. The supercoiled structure of linear DNA molecules with constrained ends. The helical nature of the DNA duplex is omitted for clarity. Image by: Richard Wheeler

Positively supercoiled (overwound) DNA is transiently generated during DNA replication and transcription, and, if not promptly relaxed, inhibits (regulates) these processes. The simple figure eight is the simplest supercoil and is the shape a circular DNA assumes to accommodate one too many or one too few helical twists. The two lobes of the figure- eight will appear rotated either clockwise or counterclockwise with respect to one another, depending on whether the helix is over- or underwound. For each additional helical twist being accommodated, the lobes will show one more rotation about their axis. As a general rule, the DNA of most organisms is negatively supercoiled.

Lobal contortions of a circular DNA, such as the rotation of the figure-eight lobes above, are referred to as ***writhe.*** The above example illustrates that twist and writhes are interconvertible. Supercoiling can be represented mathematically by the sum of ***twist*** and ***writhe*** (Figure 23.3.2 The ***twist*** is the number of helical turns in the DNA and the writhe is the number of times the double helix crosses over on itself (these are the supercoils). Extra helical twists are positive and lead to positive supercoiling, while subtractive twisting causes negative supercoiling. Many topoisomerase enzymes sense supercoiling and either generate or dissipate it as they change DNA topology.

In addition to forming supercoiled structures, circular chromosomes from bacteria have been shown to undergo the processes of ***catenation*** and ***knotting*** upon the inhibition of topoisomerase enzymes. ***Catenation*** is the process by which two circular DNA strands are linked together like chain links, whereas ***DNA knotting*** is the interlooping structures occurring within a single circular DNA structure, as shown in Figure 23.3.3 *In vivo*, the action of topoisomerase enzymes is critical to keep knots and catenoids from tangling the DNA structure. Catenanes are effectively **topologically linked circular molecules**

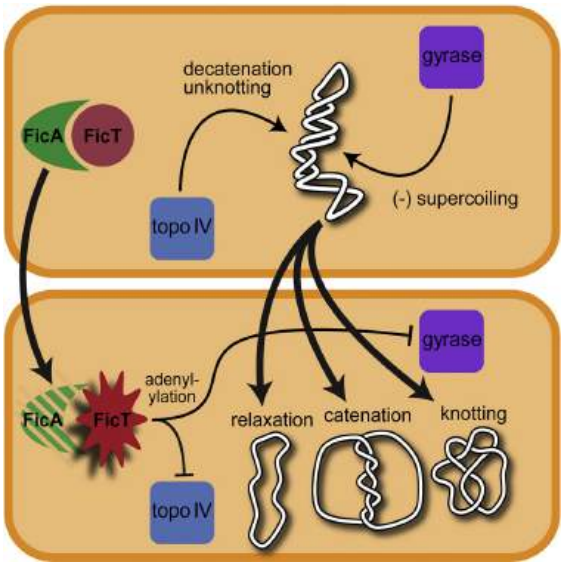


Figure 23.3.3: DNA Catenation and Knotting. The upper structure shows the negatively supercoiled form of bacterial DNA. The inhibition of topoisomerase enzyme activity leads to the relaxation, catenation, and knotting of the chromosomal structure. Image from: Harms, A. et al. (2015) Cell Reports 12(9):1497-1507.

In part, because chromosomes may be very large, segments in the middle may act as if their ends are anchored. As a result, they may be unable to distribute excess twist to the rest of the chromosome or to absorb twist to recover from underwinding—the segments may become *supercoiled*, in other words. In response to supercoiling, they will assume an amount of writhe, just as if their ends were joined.

Supercoiled circular DNA forms two major structures; a ***plectoneme*** or a ***toroid***, or a combination of both (Figure 24.3.2). A negatively supercoiled DNA molecule will produce either a one-start left-handed helix, ***the toroid***, or a two-start right-handed helix with terminal loops, the ***plectoneme. Plectonemes*** are typically more common in nature, and this is the shape most bacterial plasmids will take (Figure 4.10). For larger molecules, it is common for hybrid structures to form – a loop on a toroid can extend into a plectoneme, as shown in Figure 23.3.4 DNA supercoiling is important for DNA packaging within all cells, and seems to also play a role in gene expression.

Figure 23.3.4: Bacterial DNA Supercoiling. Atomic force microscopy (AFM) visualization of torsionally relaxed (A), and negatively supercoiled (B) bacterial plasmids pBR322. (C) Electron microscopy image of the E. coli chromosomal DNA displaying a hybrid toroidal-plectoneme structure. Image A and B from Witz, G. and Stasiak, A. (2009) Nucleic Acids Research 38(7):2119-2133. Image C from Prokaryotic Chromosomes

23.3.3: Topoisomerases

Topoisomerase can change the tension in supercoiled DNA. Think of how you untie a knot. It takes a lot of work sometimes, and if it's too hard, you simply cut the impediment to unknotting. Topoisomerases work by making transient breaks in the DNA before unwinding and religation. There are two main types of topoisomerases, topo I and topo II. It's very hard to describe their activities with just words and static diagrams. View the video below and you will get a great sense of what the enzymes do and how they are different.

With this background, we can now explore each enzyme in more detail. They are both targets of cancer drugs which makes them even more interesting.

Topo I enzyme relaxes DNA by nicking one stand. The dsDNA then rotates around the non-nicked strand. It unwinds new DNA and allows the condensation of chromosomes. When both DNA and RNA polymerase makes new DNA and RNA strands, respectively, they increase the supercoiling of the nucleic acid. Topoisomerases relax them. They also play a role in the regulation of gene expression by affecting gene promoters where RNA polymerase binds, with negative supercoiling enhancing transcription and positive supercoiling inhibiting it.

Figure 23.3.5 shows the topology of DNA and an overview of the mechanisms of Topo I and II. : DNA topology and DNA topoisomerase mechanism

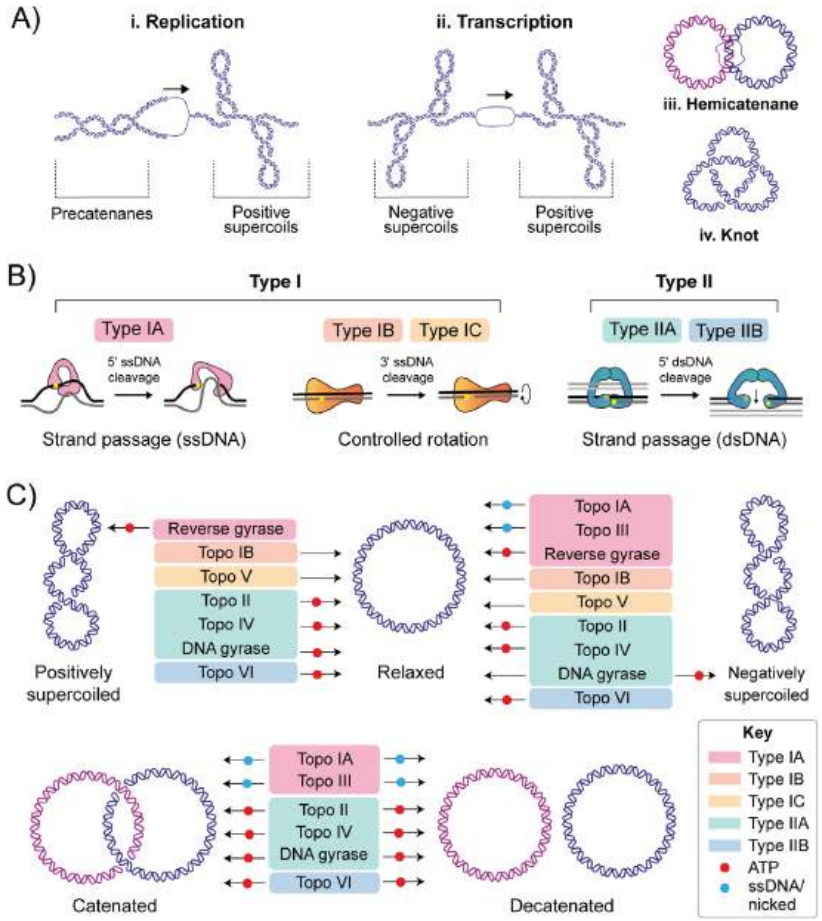


Figure 23.3.5: DNA topology and DNA topoisomerase mechanisms. Shannon J. McKie, Keir C. Neuman, and Anthony Maxwell. Bioessays (2021). https://doi.org/10.1002/bies.202000286. Attribution 4.0 International (CC BY 4.0)

(A) Topological consequences of DNA metabolism. i) During DNA replication, strand separation leads to positive supercoiling ahead of the advancing protein machinery, and precatenane formation behind. Precatenanes form as the newly-synthesized duplexes wrap around one another, and, if not removed before completion of replication, catenated DNA molecules are formed. ii) During transcription, strand separation leads to positive supercoiling ahead of the advancing protein machinery, and negative supercoil formation behind. iii) Hemicatenanes are a possible end result of replication, in which the parental strands of the replicated duplexes remain base-paired. iv: DNA knotting can also occur as a result of DNA replication in which a DNA molecule is intramolecularly linked.

(B) Summary of topo categories and mechanism. The topos are categorized based on whether they catalyze single- (type I) or double-stranded (type II) DNA breaks. The type I topos are further subdivided to type IA, IB, and IC. Type IA form a transient covalent bond to the 5′ DNA phosphate and function via a strand passage mechanism. Type IB and IC form a transient covalent bond to the 3′ DNA phosphate and function via a controlled-rotation mechanism. Type II topos are further subdivided into type IIA and IIB. Both form a transient covalent bond to the 5′ DNA phosphate of both strands of the duplex and function via a strand-passage mechanism.

(C) Summary of the topological manipulations performed by DNA topoisomerases, namely relaxation of positive and negative supercoils and decatenation. Type IA topos are color-coded pink, type IB are orange, type IC are yellow, type IIA are green, and type IIB are blue. The requirement of ATP or ssDNA for activity is denoted using a red or blue circle, respectively

Topoisomerase I (Topo I):

Class I topoisomerases wrap around the DNA and cut one strand. Keeping that spot in place, the helix can spin to reduce strain caused by either over- or underwinding. After these geometric contortions, the single-stranded DNA nick is repaired and the tension is relieved. Type IA topoisomerase from E. Coli is shown in Figure 23.3.6

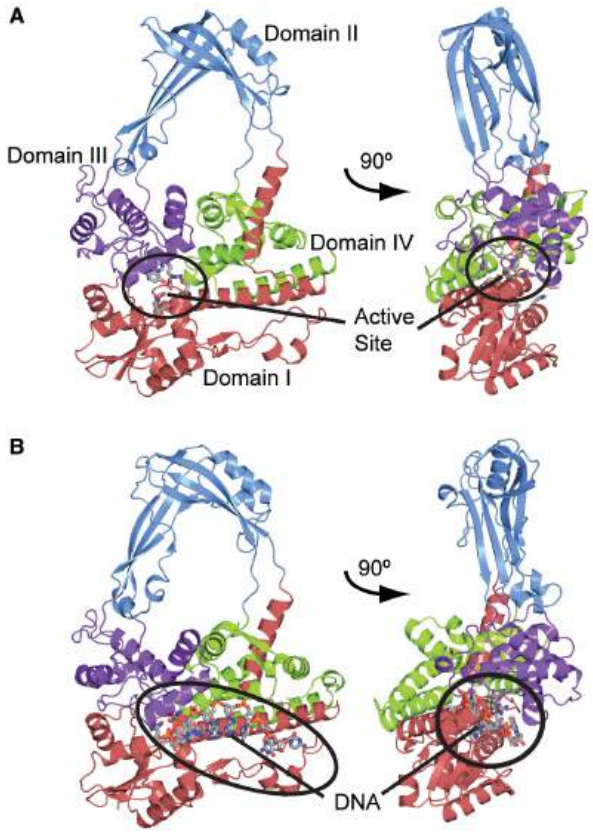


Figure 23.3.6 Structure of a type IA topoisomerase. E. coli topoisomerase III is shown to illustrate the overall structure of a type IA topoisomerase and the typical toroidal fold observed in all members of this type.

(A) Diagram showing the structure of the apo-enzyme [PDB 1D6M(7)]. In the absence of DNA, the active site, found at the intersection of domains I and III (encircled), is buried.

(B) Diagram showing the structure of a complex with single-stranded DNA [PDB 1I7D (23)]. Note the movement of domains that occurs to accommodate DNA. In both diagrams, the four major domains of the protein are colored red, blue, purple, and green for domains I, II, III, and IV, respectively. The single-stranded DNA binding groove, shown circled in black, extends from domain IV to the active site. The active site residues as well as the single-stranded DNA in the complex are shown in a ball and stick representation. Dasgupta T, Ferdous S, Tse-Dinh YC. Mechanism of Type IA Topoisomerases. Molecules. 2020 Oct 17;25(20):4769. doi: 10.3390/molecules25204769. PMID: 33080770; PMCID: PMC7587558. Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

A simplified cartoon mechanism for Topo I is shown in Figure 23.3.7.

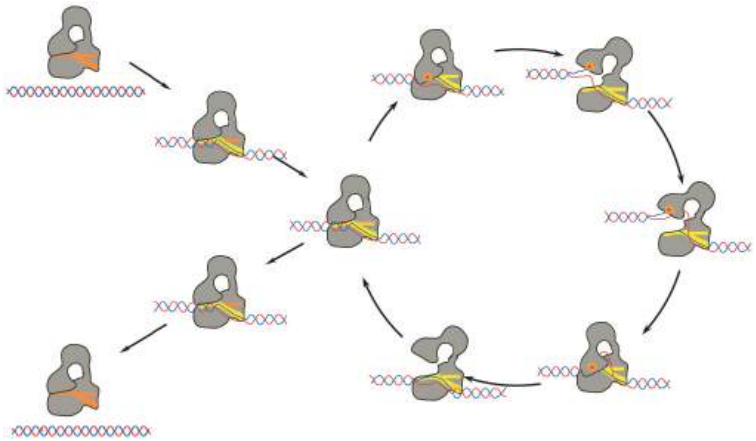


Figure 23.3.7: Diagram showing the proposed mechanism of DNA relaxation by type IA topoisomerases. The mechanism involves several transient
conformational intermediates both of the protein and the DNA. The sequence of the steps and the intermediates are hypothetical and more states are likely to be involved in the cycle. Processivity by the enzyme requires that after one relaxation event, the protein continues to another relaxation cycle without releasing the DNA. In the diagram,
the protein is shown in grey, and the DNA in red/blue. The orange dot represents the presence of the covalent protein/DNA complex. The single-stranded DNA binding groove is shown in red or yellow. Dasgupta, T et al. ibid.

A more detailed view of the domain structure and mechanism for Topo I is shown in Figure 23.3.8

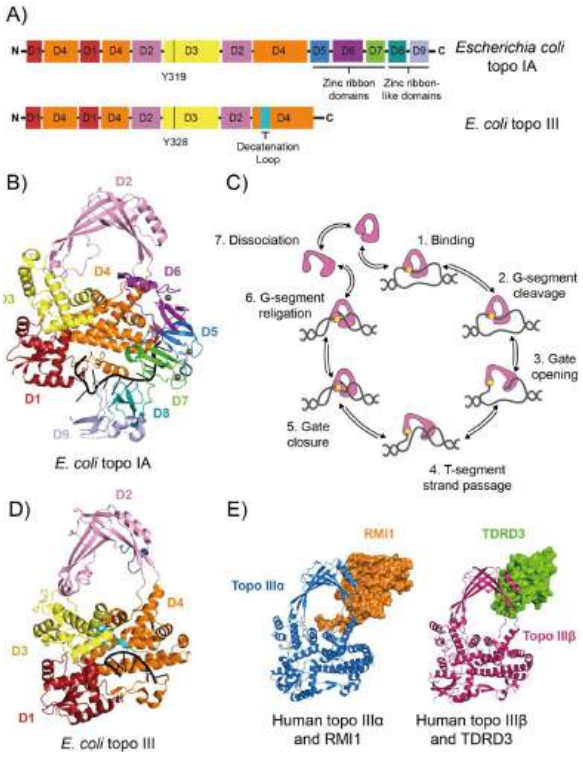


Figure 23.3.8: Type IA DNA topoisomerases. Dasgupta, T et al. ibid.

(A) Protein domain organization of *Escherichia coli* DNA topoisomerase IA (topo IA) and DNA topoisomerase III (topo III). Black vertical lines represent the active site tyrosines.

(B) Crystal structure of *E. coli* topo I bound to ssDNA (PDB: 4RUL).[20]

(C) Strand-passage mechanism for type IA topos. (1) topo binds the G-segment ssDNA region, (2) the G-segment is cleaved. (3) The topo DNA gate is opened, (4) which allows T-segment transfer through the cleaved G-strand. (5) The DNA gate is closed, (6) and the G-strand is re-ligated, changing the linking number by 1. (7) The topo can then go through another round of relaxation or dissociate from the DNA. Type IA topo (domains 1–4) is in pink, the active site tyrosine is yellow and the DNA is grey.

(D) Crystal structure of *E. coli* topo III bound to ssDNA (PDB: 2O54).[26]

(E) Crystal structures of human topo IIIα (blue) bound to RMI1(orange) (PDB: 4CGY),[39] and human topo IIIβ (magenta) bound to TDRD3 (green) (PDB: 5GVE).[60] For panels A, B, and C, the topo I and III domains are color-coded as follows: D1 is red, D2 is pink, D3 is yellow, D4 is orange, D5 is marine blue, D6 is purple, D7 is green, D8 is teal, and D9 is light blue

Figure 23.3.9 shows an interactive iCn3D model of E.Coli topoisomerase I in complex with ssDNA (4RUL).

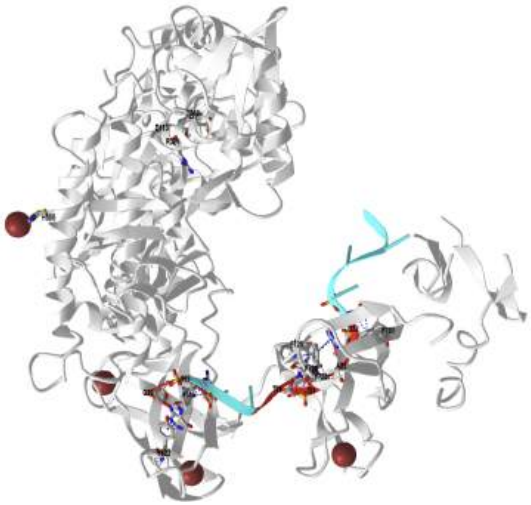


Figure 23.3.9: E.Coli topoisomerase I in complex with ssDNA (4RUL). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...JxcFvds2e8p2WA

The topoisomerase is shown in gray and the single-stranded DNA is in cyan. 4 Zn^{2+} ions are shown. Three form Zn-finger motifs. Two of these are shown near the protein binding site for DNA. A fourth Zn is bound to a single His 566 in the structure. The active site amino acids of the enzyme, which includes residues D111, D113, Y319, and R321 from D1 and D3, are shown and labeled. Also shown are active site residues, E115, F139 and Y312. There are 4 Cys-Zn ribbon domains. The ones that interact with the ssDNA involve π-stacking, some of which are illustrated in the iCn3D model.

Class II topoisomerases (Topo II)

A series of enzymes are included in this class including DNA gyrase and Topo (IV) from prokaryotes and Topo II from eukaryotes. In eukaryotes, they help sister chromosomes separate if they get tangled during cell division. This enzyme works by: this enzyme makes a double-stranded cut, moves on the helix through the cut, and reseals the cut.

- binding the gate segment (G -segment) ds-DNA at a DNA gate where a double-stranded break is made
- binding the transport segment (T-segment) ds-DNA at the N-gate where the nucleotide ATP binds
- The T-segment DNA moves through the break in the G-segment and released the C-gate
- The G- and T-segments are reconnected.

After this, the N gate reopens to allow the process to occur again.

The domain structure of Topo IIs and the general mechanism of action are shown in Figure 23.3.10

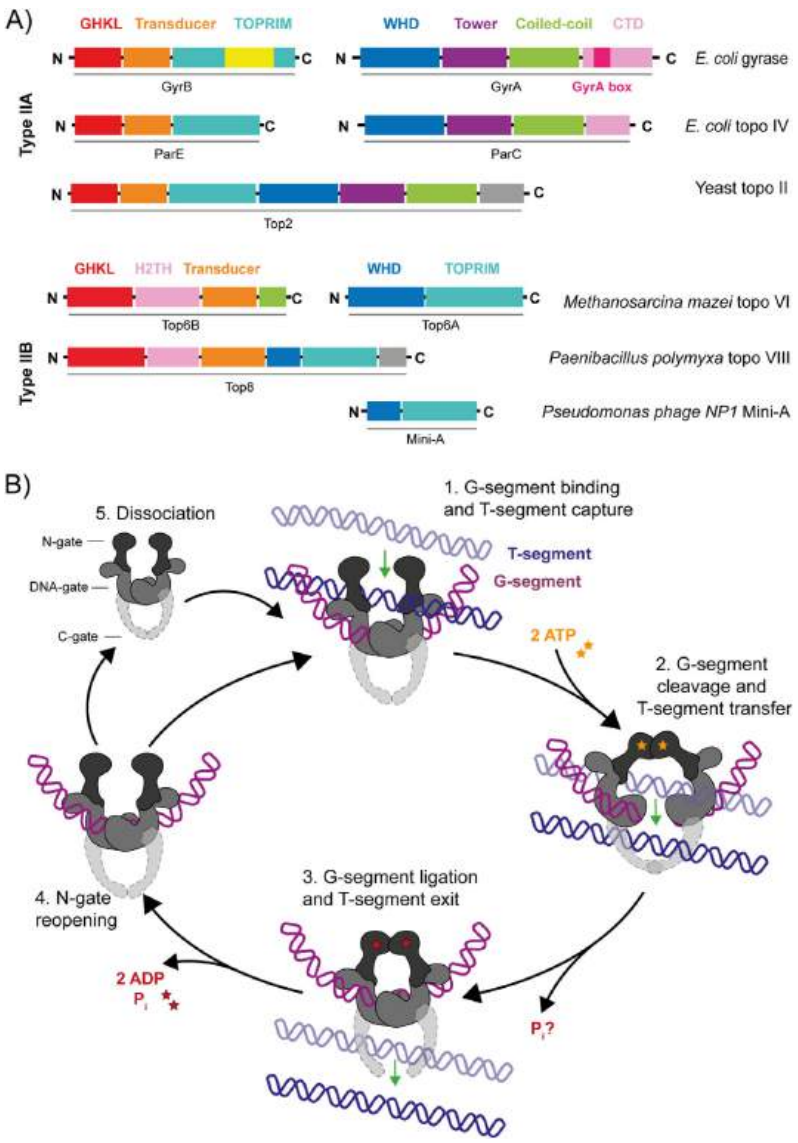


Figure 23.3.10 Type II DNA topoisomerases: domain organization and mechanism.

(A) Protein domain organization for the type IIA topos: *E. coli* DNA gyrase, *E. coli* DNA topoisomerase IV (topo IV), yeast DNA topoisomerase II (topo II), *Methanosarcina mazei* DNA topoisomerase VI (topo VI), *Paenibacillus polymyxa* DNA topoisomerase VIII (plasmid-borne), and *Pseudomonas phage NP1* Mini-A.

(B) type II topo strand passage mechanism. (1) G-segment is bound at the DNA gate and the T-segment is captured. (2) ATP binding stimulates dimerization of the N-gate, the G-segment is cleaved and the T-segment is passed through the break. (3) The G-segment is re-ligated and T-segment exits through the C-gate. For type IIB topos, there is no C-gate so once the T-segment passes through the G-segment, it is released from the enzyme. (4) Dissociation of ADP and P_i allows N-gate opening, a scenario where the enzyme either remains bound to the G-segment, ready to capture a consecutive T-segment, or (5) dissociates from the G-segment.

Figure 23.3.11shows an interactive iCn3D model of Yeast Topoisomerase II-DNA-AMPPNP complex (4GFH) .

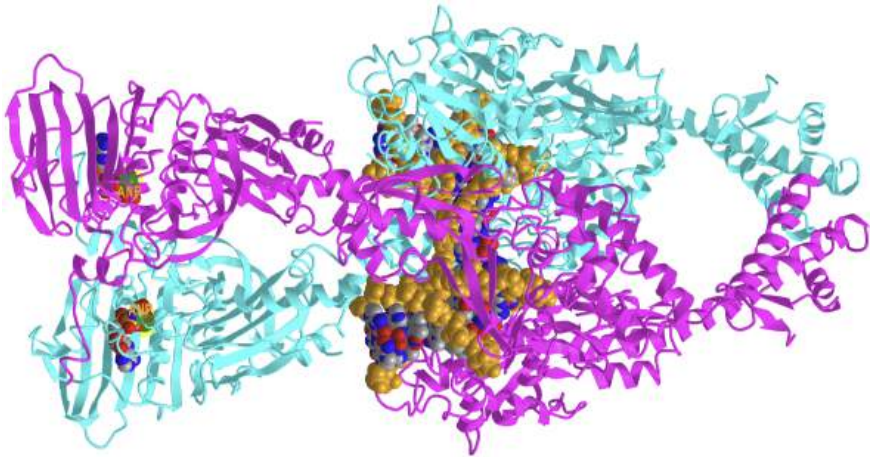


Figure 23.3.11: Yeast Topoisomerase II-DNA-AMPPNP complex (4GFH). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...b5EHLfMwde5Dw8

The two protein chains in the homodimer are shown in magenta and cyan. The double-stranded DNA is shown with a brown backbone and CPK-colored spheres. ANP (AMPPNP), a nonhydrolyzable ATP analog, and Mg^{2+} ions are shown in spacefill and labeled.

The ATP binding and ATPase domain of one of the monomers (cyan for example) is adjacent to the nuclease-cutting domain of the other monomer (magenta). This requires some conformational gymnastics as the ATP binding and cleavage domain move around each other to allow the DNA strand to pass in the right direction and to reset the enzyme.

Note the circular nature of chloroplast and mitochondrial DNA, suggesting a bacterial origin for both of these organelle structures. Sequence alignments further lend support for the ***endosymbiotic theory,*** which proposes that bacteria were engulfed by early eukaryotic organisms and subsequently became symbiotic to their eukaryotic counterpart, rather than being digested.

A reminder about mitochondrial DNA

In the cells of extant organisms, the vast majority of the proteins present in the mitochondria (numbering approximately 1500 different types in mammals) are coded for by nuclear DNA. However, sequencing of the human mitochondrial genome has revealed 16,569 base pairs encoding 13 proteins (Figure 24.3.5). Many of the mitochondrially produced proteins are required for electron transport during the production of ATP, as shown in Figure 23.3.12

Figure 23.3.12: Mitochondrial Genome. Mitochondria are organelle structures containing a double membrane, thought to have originated as an independent prokaryotic organism that was originally engulfed by a eukaryotic organism, where it became a symbiotic counterpart. Mitochondria contain circular chromosomal DNA that shares high sequence similarity with alpha protobacteria. The human mitochondrial genome contains 16,569 base pairs encoding 13 proteins and ribosomal RNA (rRNA) components. Images adapted from: The National Human Genome Research Institute and Shanel, Knopfkind, and JHC.

23.3.4: Histones and DNA packing

Within eukaryotic chromosomes, chromatin proteins, known as ***histones***, compact and organize DNA. These compacting structures guide the interactions between DNA and other proteins, helping control which parts of the DNA are transcribed.

Histones are highly alkaline proteins found in eukaryotic cell nuclei that package and order the DNA into structural units called nucleosomes. They are the chief protein components of chromatin, acting as spools around which DNA winds, and playing a role in gene regulation. Without histones, the unwound DNA in chromosomes would be very long (a length-to-width ratio of more than 10 million to 1 in human DNA). For example, each human diploid cell (containing 23 pairs of chromosomes) has about 1.8 meters of DNA; wound on the histones, the diploid cell has about 90 micrometers (0.09 mm) of chromatin.

Five major families of histones exist: H1/H5, H2A, H2B, H3, and H4. Histones H2A, H2B, H3, and H4 are known as the core histones, while histones H1/H5 are known as the linker histones.

The core histones all exist as dimers, which are similar in that they all possess the histone fold domain: three alpha helices linked by two loops. It is this helical structure that allows for interaction between distinct dimers, particularly in a head-tail fashion (also called the handshake motif). The resulting four distinct dimers then come together to form one octameric nucleosome core, approximately 63 Angstroms in diameter. Around 146 base pairs (bp) of DNA wrap around this core particle 1.65 times in a left-handed super-helical turn to give a particle of around 100 Angstroms across, called a nucleosome, as shown in Figure 23.3.13

Figure 23.3.13: Nucleosome Core Structure. Histones H2A and H2B dimerize, and Histones H3 and H4 dimerize. Two dimers of each join to form a histone core octamer. The DNA double helix winds 1.65 times around the octamer core forming the nucleosome structure. Image adapted from: Nucleosome Structure

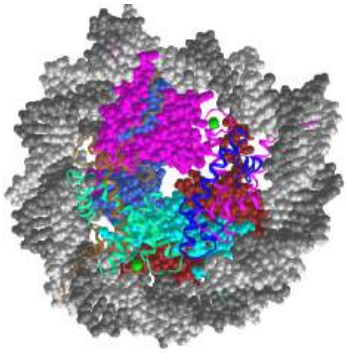
The linker histone H1 binds the ***nucleosome*** at the entry and exit sites of the DNA, thus locking the DNA into place and allowing the formation of a higher order structure, as shown in Figure 23.3.14 The most basic such formation is the 10 nm fiber or beads on a string conformation. This involves the wrapping of DNA around nucleosomes with approximately 50 base pairs of DNA separating each pair of nucleosomes (also referred to as linker DNA).

Figure 23.3.14: Overall Nucleosome Structure. (A) Side view diagram of the nucleosome structure with the histone octamer shown in blue, the DNA double helix in red, and the histone H1 linker in green. (B) Shows a top-view rendering of the histone octamer with the associated DNA helix. Note that the Histone tails from H3 and H2B protrude from the DNA. Image A from: Darekk2 Image B from EMW

The nucleosome contains over 120 direct protein-DNA interactions and several hundred water-mediated ones. Direct protein – DNA interactions are not spread evenly about the octamer surface but rather located at discrete sites. These are due to the formation of two types of DNA binding sites within the octamer; the α1α1 site, which uses the α1 helix from two adjacent histones, and the L1L2 site formed by the L1 and L2 loops. Salt links and hydrogen bonding between both side-chain basic and hydroxyl groups and main-chain amides with the DNA backbone phosphates form the bulk of interactions with the DNA. This is important, given that the ubiquitous distribution of nucleosomes along genomes requires it to be a non-sequence-specific DNA-binding factor. Although nucleosomes tend to prefer some DNA sequences over others, they are capable of binding practically to any sequence, which is thought to be due to the flexibility in the formation of these water-mediated interactions. In addition, non-polar interactions are made between protein side-chains and the deoxyribose groups, and an arginine side-chain intercalates into the DNA minor groove at all 14 sites where it faces the octamer surface. The distribution and strength of DNA-binding sites about the octamer surface distort the DNA within the nucleosome core. The DNA is non-uniformly bent and also contains twist defects. The twist of free B-form DNA in solution is 10.5 bp per turn. However, the overall twist of nucleosomal DNA is only 10.2 bp per turn, varying from a value of 9.4 to 10.9 bp per turn.

The histone tail extensions constitute up to 30% by mass of the histones but are not visible in the crystal structures of nucleosomes due to their high intrinsic flexibility, and have been thought to be largely unstructured (Figure 4.14). The N-terminal tails of histones H3 and H2B pass through a channel formed by the minor grooves of the two DNA strands, protruding from the DNA every 20 bp. The N-terminal tail of histone H4, on the other hand, has a region of highly basic amino acids (16-25), which, in the crystal structure, forms an interaction with the highly acidic surface region of a H2A-H2B dimer of another nucleosome, being potentially relevant for the higher-order structure of nucleosomes. This interaction is thought to occur under physiological conditions also, and suggests that acetylation of the H4 tail distorts the higher-order structure of chromatin.

Figure 23.3.15shows an interactive iCn3D model of the human nucleosome (3afa). One member of each pair of histones is shown in cartoon rendering, while the other member of the pair is shown in the same color but in spacefill rendering. The structure of a human nucleosome (3afa) is shown below (H2A is shown in cyan, H2B in blue, H3 in magenta, and H4 in purple). Each strand of DNA is shown in a different shade of gray.



NCBI iCn3D Figure 23.3.15 Human nucleosome (3afa). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...B2SwQHYDLj4BJ6

The formation of the DNA double helix represents the first-order packaging of the chromosome structure. The formation of nucleosomes represents the second level of packaging for eukaryotic chromosomes. *In vitro* data suggests that nucleosomes are then arranged into either a solenoid structure which consists of 6 nucleosomes linked together by the Histone H1 linker proteins or a zigzag structure that is similar to the solenoid construct, as shown in Figure 23.3.16 Both the solenoid and zigzag structures are approximately 30 nm in diameter. The solenoid and zigzag structures reported from *in vitro* data have not yet been confirmed to occur *in vivo.*

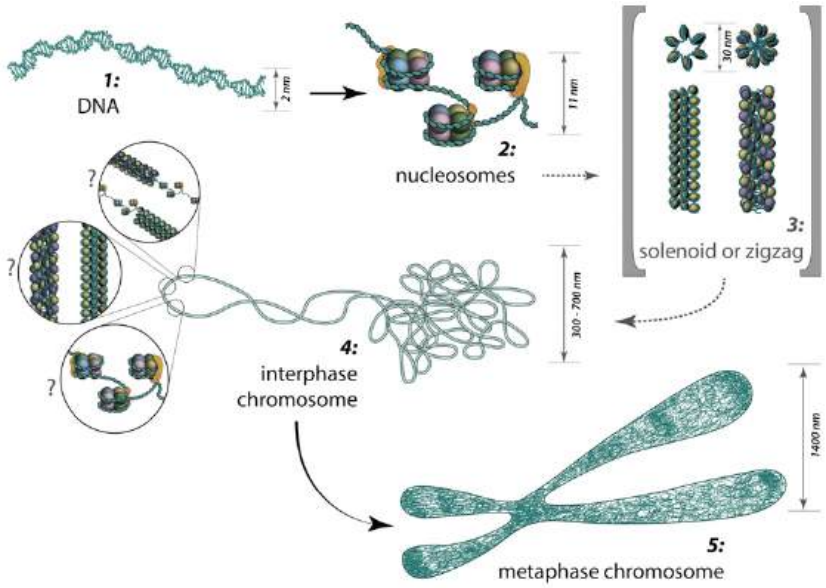


Figure 23.3.16: Chromosome Structure. (1) DNA double helix is approximately 2 nm in diameter. (2) The nucleosome core structure is approximately 11 nm in diameter. (3) The solenoid/zigzag structure is approximately 30 nm in diameter and is proposed to form chromosome loops (4) during cellular interphase and more condensed chromosome territories (5) during mitosis. Image by: MBInfo

During interphase, each chromosome occupies a spatially limited, roughly elliptical domain which is known as a chromosome territory (CT). Each chromosome territory is comprised of higher-order chromatin units of ~1 Mb each. These units are likely built up from smaller loop domains that contain the solenoid/zigzag structural motifs. On the other hand, 1Mb domains can themselves serve as smaller units in higher-order chromatin structures.

Chromosome territories are known to be arranged radially around the nucleus. This arrangement is both cell and tissue-type specific and is also evolutionarily conserved. The radial organization of chromosome territories was shown to correlate with their gene density and size. In this case, the gene-rich chromosomes occupy interior positions, whereas larger, gene-poor chromosomes, tend to be located around the periphery. Chromosome territories are also dynamic structures, with genes able to relocate from the periphery towards the interior once they have been ‘switched on’. In other cases, genes may move in the opposite direction, or simply maintain their position. The eviction of genes from their chromosome territories into the interchromatin compartment or a neighboring chromosome territory is often accompanied by the formation of large decondensed chromatin loops.

23.3.5: Models describing chromosome territory arrangement

With the development of high-throughput biochemical techniques, such as 3C (chromosome conformation capture) and 4C (chromosome conformation capture-on-chip and circular chromosome conformation capture), numerous spatial interactions between neighboring chromatin territories have been described, as shown in Figure 23.3.17. These descriptions have been supplemented with the construction of spatial proximity maps for the entire genome (e.g., for a human lymphoblastoid cell line). Together, these observations and physical simulations have led to the proposal of various models that aim to define the structural organization of chromosome territories:

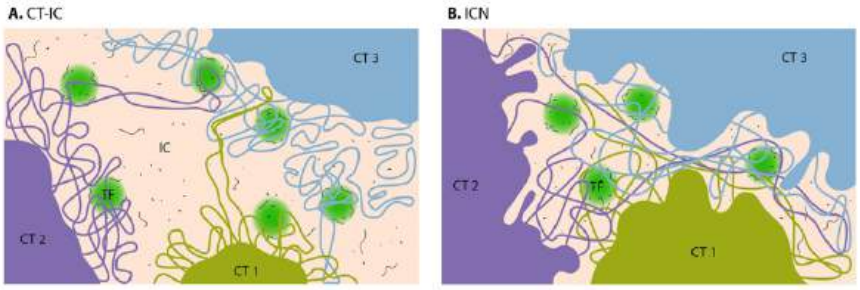


Figure 23.3.17: Computer Models of Chromosome Territory (CT) Structure. In the CT-IC model, the space between discrete CTs can be visualized in light and electron microscope and is called the interchromatin compartment (IC). Transcription factories (TF, green color) are localized predominantly in perichromatin region. In the ICN model, the interchromatin compartment is not apparent. Instead, the space between CTs is occupied by intermingling decondensed chromatin loops, which often share the same transcription factories. Image by: MBInfo

1. The chromosome territory-interchromatin compartment (CT-IC) model describes two principal compartments: chromosome territories (CTs) and an interchromatin compartment (IC). In this model, chromosome territories build up an interconnected chromatin network that is associated with an adjacent 3D space called the interchromatin compartment. The latter can be observed using both light and electron microscopy.

Within a single chromosome territory, the interphase chromosome is divided into defined regions based on the level of chromosome condensation. Here, the inner part of the interphase chromosome is comprised of more condensed chromatin domains or higher-order chromatin fibers, while a thin (<200 nm) layer of more decondensed chromatin, known as the perichromatin region, can be found around the chromosomal periphery. Functionally, the perichromatin region represents the major transcriptional compartment and is also the region where most co-transcriptional RNA splicing takes place. DNA replication and DNA repair are also predominately carried out within the perichromatin region. Finally, nascent RNA transcripts, referred to as perichromatin fibrils, are also generated in the perichromatin region. Perichromatin fibrils are then subjected to the splicing events by the factors, provided by the interchromatin compartment.

The lattice model, proposed by Dehgani et al. is based on reports that transcription also occurs within the inner, more condensed chromosome territories and not only at the interface between the interchromatin compartment and the perichromatin region. Using ESI (electron spectroscopic imaging), Dehgani et al. showed that chromatin was organized as an array of deoxyribonucleoprotein fibers of 10–30 nm in diameter. In this study, the interchromatin compartments, which are described in the CT-IC model as large channels between chromosome territories, were not apparent. Instead, chromatin fibers created a loose meshwork of chromatin throughout the nucleus that intermingled at the periphery of chromosome territories. Thus, inter- and intra-chromosomal spaces within this meshwork are essentially contiguous and together form the intra-nuclear space.

2. The interchromatin network (ICN) model predicts that intermingling chromatin fibers/loops can make both cis- (within the same chromosome) and trans- (between different chromosomes) contacts. This intermingling is uniform and makes a distinction between the chromosome territory and interchromatin compartment functionally meaningless. The advantage of the ICN model is that it permits high chromatin dynamics and diffusion-like movements. The authors propose that ongoing transcription influences the degree of intermingling between specific chromosomes by stabilizing associations between particular loci. Such interactions are likely to depend on the transcriptional activity of the loci and are therefore cell-type specific.

The cell type-specific organization of chromosome territories has been studied by measuring the volume and frequency of intermingling between heterologous chromosomes. By using 3C (chromosome conformation capture) and FISH (fluorescence in situ hybridization) to map the regions of chromosome intermingling, it was revealed that these regions contain a higher density of active genes and are enriched with markers of transcriptional activation and repression, such as activated RNAPII. By comparing the positions of the CTs in undifferentiated mouse embryonic stem (ES) cells, ES cells in early stages of differentiation, and terminally differentiated NIH3T3 cells, it was shown that fully differentiated cells had a higher enrichment of RNAPII, compared to undifferentiated or less-differentiated cells. The findings support the notion that the intermingling regions have functional significance in the nucleus and provide a basis for understanding how the radial and relative positions of chromosomal territories evolve during the process of differentiation, explaining their organization in a cell type-dependent manner.

3. The Fraser and Bickmore model emphasizes the functional importance of giant chromatin loops, which originate from chromosome territories and expand across the nuclear space to share transcription factories. In this case, both *cis*- and *trans*- oops of decondensed chromatin can be co-expressed and co-regulated by the same transcription factory.

4. The Chromatin polymer models assume a broad range of chromatin loop sizes and predict the observed distances between genomic loci and chromosome territories, as well as the probabilities of contacts being formed between given loci. These models apply physics-based approaches that highlight the importance of entropy for understanding nuclear organization. By proposing the existence of conformational chromatin ensembles with structures based on three possible homopolymer states, these models also provide alternative structures to the traditional 30 nm chromatin fiber, which has been brought into question following recent studies.

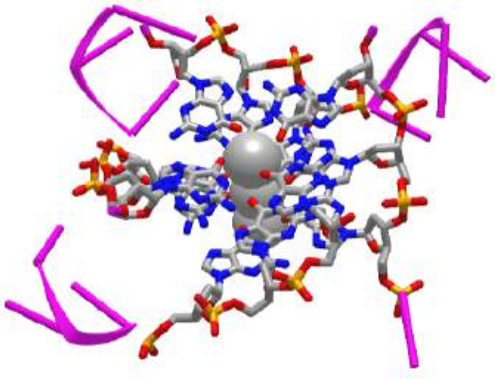
With a lack of experimental evidence to support these described models, it must be remembered that they serve only to hypothesize the structural and chemical properties of intermediate chromatin structures and to highlight unanswered questions. For example, the mechanisms that exist to control the rate and the extent of chromatin movement remain to be defined

23.3.6: Telomere Structures

At the ends of the linear chromosomes are specialized regions of DNA called telomeres, shown in Figure 23.3.18 The main function of these regions is to allow the cell to replicate chromosome ends using the enzyme telomerase, as the enzymes that normally replicate DNA cannot copy the extreme 3′ ends of chromosomes. These specialized chromosome caps also help protect the DNA ends, and stop the DNA repair systems in the cell from treating them as damage to be corrected. In human cells, telomeres are usually lengths of single-stranded DNA containing several thousand repeats of a simple TTAGGG sequence.

Figure 23.3.18: Telomere Structure. (A) Telomeres are located at the end of chromosomes, where they help protect against the loss of DNA during replication. (B) DNA quadruplex formed by telomere repeats. The looped conformation of the DNA backbone is very different from the typical DNA helix. The green spheres in the center represent potassium ions. Image (A) by MBInfo and Image (B) by: Thomas Splettstoesser

In human cells, telomeres contain 300-8000 repeats of a simple **TTAGGG** sequence. The repetitive **TTAGGG** sequences in telomeric DNA can form unique higher-order structures called **quadruplexes**. Figure 23.3.19shows an interactive iCn3D model of parallel quadruplexes from human telomeric DNA (1KF1). The structure contains a single DNA strand **(**5'-AGGGTTAGGGTTAGGGTTAGGG-3') which contains four TTAGGG repeats.



NCBI iCn3D Figure 23.3.19 A buried phenylalanine in low molecular weight protein tyrosyl phosphatase (1xww) (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...y5joFHDgWJQsQ6

Rotate the model to see 3 parallel layers of quadruplexes. In each layer, **4** noncontiguous guanine bases interact with a K^+ ion. Hover over the guanine bases in one layer and you will find that one layer consists of guanines 4, 10, 16, and 22, which derive from the last G in each of the repeats in the sequence of the oligomer used (5'-AGGGTTAGGGTTAGGGTTAGGG-3'). These quadruplexes certainly serve as recognition and the binding site for telomerase proteins. The guanine-rich telomere sequences which can form quadruplex may also function to stabilize chromosome ends

During DNA replication, the double-stranded DNA is unwound and DNA polymerase synthesizes new strands. However, as DNA polymerase moves in a unidirectional manner (from 5' to 3'), only the leading strand can be replicated continuously. In the case of the lagging strand, DNA replication is discontinuous. In humans, small RNA primers attach to the lagging strand DNA, and the DNA is synthesized in small stretches of about 100-200 nucleotides, which are termed Okazaki fragments. The RNA primers are removed, and replaced with DNA and the Okazaki fragments are ligated together. At the end of the lagging strand, it is impossible to attach an RNA primer, meaning that there will be a small amount of DNA lost each time the cell divides. This 'end replication problem' has serious consequences for the cell as it means the DNA sequence cannot be replicated correctly, with the loss of genetic information.

To prevent this, telomeres are repeated hundreds to thousands of times at the end of the chromosomes. Each time cell division occurs, a small section of telomeric sequences is lost to the end replication problem, thereby protecting the genetic information. At some point, the telomeres become critically short. This attrition leads to cell senescence, where the cell is unable to divide, or apoptotic cell death. Telomeres are the basis for the Hayflick limit, the number of times a cell can divide before reaching senescence.

Telomeres can be restored by the enzyme telomerase, which extends telomeres length (Figure 24.3.10). Telomerase activity is found in cells that undergo regular division, such as stem cells and lymphocyte cells of the immune system. Telomeres can also be extended through the Alternative Lengthening of Telomeres (ALT) pathway. In this case, rather than being extended, telomeres are switched between chromosomes by homologous recombination. As a result of the telomere swap, one set of daughter cells will have shorter telomeres, and the other set will have longer telomeres.

A downside to telomere extension is the potential for uncontrolled cell division and cancer. Abnormally high telomerase activity has been found in the majority of cancer cells, and non-telomerase tumors often exhibit ALT pathway activation. As well as the potential for losing genetic information, cells with short telomeres are at high risk for improper chromosome recombination, which can lead to genetic instability and aneuploidy (an abnormal number of chromosomes).

These guanine-rich telomere sequences may also stabilize chromosome ends by forming structures of stacked sets of four-base units, rather than the usual base pairs found in other DNA molecules (Figure 24.3.10). Here, four guanine bases form a flat plate and these flat four-base units then stack on top of each other, to form a stable G-quadruplex structure. These structures are stabilized by hydrogen bonding between the edges of the bases and the chelation of a metal ion in the center of each four-base unit. Other structures can also be formed, with the central set of four bases coming from either a single strand folded around the bases or several different parallel strands, each contributing one base to the central structure.

In addition to these stacked structures, telomeres also form large loop structures called telomere loops or T-loops. Here, the single-stranded DNA curls around in a long circle stabilized by telomere-binding proteins. At the very end of the T-loop, the single-stranded telomere DNA is held onto a region of double-stranded DNA by the telomere strand disrupting the double-helical DNA and base pairing to one of the two strands. This triple-stranded structure is called a displacement loop or D-loop.

23.3.7: References

Börner, R., Kowerko, D., Miserachs, H.G., Shaffer, M., and Sigel, R.K.O. (2016) Metal ion induced heterogeneity in RNA folding studied by smFRET.Coordination Chemistry Reviews327 DOI: 10.1016/j.ccr.2016.06.002 Available at:https://www.researchgate.net/publication/303846502_Metal_ion_induced_heterogeneity_in_RNA_folding_studied_by_smFRET

Hardison, R. (2019) B-Form, A-Form, and Z-Form of DNA. Chapter in: R. Hardison's Working with Molecular Genetics.Published by LibreTexts. Available at:https://bio.libretexts.org/Bookshelves/Genetics/Book%3A_Working_with_Molecular_Genetics_(Hardison)/Unit_I%3A_Genes%2C_Nucleic_Acids%2C_Genomes_and_Chromosomes/2%3A_Structures_of_Nucleic_Acids/2.5%3A_B-Form%2C_A-Form%2C_and_Z-Form_of_DNA

Lenglet, G., David-Cordonnier, M-H., (2010) DNA-destabilizing agents as an alternative approach for targeting DNA: Mechanisms of action and cellular consequences.Journal of Nucleic Acids2010, Article ID: 290935, DOI: 10.4061/2010/290935 Available at:https://www.hindawi.com/journals/jna/2010/290935/

Mechanobiology Institute (2018) What are chromosomes and chromosome territories? Produced by the National University of Singapore.Available at:https://www.mechanobio.info/genome-regulation/what-are-chromosomes-and-chromosome-territories/

National Human Genome Research Institute (2019) The Human Genome Project.National Institutes of Health. Available at:https://www.genome.gov/human-genome-project

Wikipedia contributors. (2019, July 8). DNA. In Wikipedia, The Free Encyclopedia. Retrieved 02:41, July 22, 2019, from https://en.Wikipedia.org/w/index.php?title=DNA&oldid=905364161

Wikipedia contributors. (2019, July 22). Chromosome. In Wikipedia, The Free Encyclopedia. Retrieved 15:18, July 23, 2019, from https://en.Wikipedia.org/w/index.php?title=Chromosome&oldid=907355235

Wikilectures. Prokaryotic Chromosomes (2017) In MediaWiki, Available at:https://www.wikilectures.eu/w/Prokaryotic_Chromosomes

Wikipedia contributors. (2019, May 15). DNA supercoil. In Wikipedia, The Free Encyclopedia. Retrieved 19:40, July 25, 2019, from https://en.Wikipedia.org/w/index.php?title=DNA_supercoil&oldid=897160342

Wikipedia contributors. (2019, July 23). Histone. In Wikipedia, The Free Encyclopedia. Retrieved 16:19, July 26, 2019, from https://en.Wikipedia.org/w/index.php?title=Histone&oldid=907472227

Wikipedia contributors. (2019, July 17). Nucleosome. In Wikipedia, The Free Encyclopedia. Retrieved 17:17, July 26, 2019, from https://en.Wikipedia.org/w/index.php?title=Nucleosome&oldid=906654745

Wikipedia contributors. (2019, July 26). Human genome. In Wikipedia, The Free Encyclopedia. Retrieved 06:12, July 27, 2019, from https://en.Wikipedia.org/w/index.php?title=Human_genome&oldid=908031878

Wikipedia contributors. (2019, July 19). Gene structure. In Wikipedia, The Free Encyclopedia. Retrieved 06:16, July 27, 2019, from https://en.Wikipedia.org/w/index.php?title=Gene_structure&oldid=906938498

CHAPTER OVERVIEW

24: DNA Metabolism

24.1: DNA Replication

24.2: DNA Mutations, Damage, and Repair

24.3: DNA Recombination

24.1: DNA Replication

24.1.1: Introduction

The elucidation of the structure of the double helix by James Watson and Francis Crick in 1953 provided a hint as to how DNA is copied during the process of ***DNA replication***. Separating the strands of the double helix would provide two templates for the synthesis of new complementary strands, but exactly how new DNA molecules were constructed was still unclear. In one model, ***semiconservative replication***, the two strands of the double helix separate during DNA replication, and each strand serves as a template from which the new complementary strand is copied. After replication in this model, each double-stranded DNA includes one parental or "old" strand and one daughter or "new" strand. There were two competing models also suggested: ***conservative*** and ***dispersive***, which are shown in Figure 24.1.1.

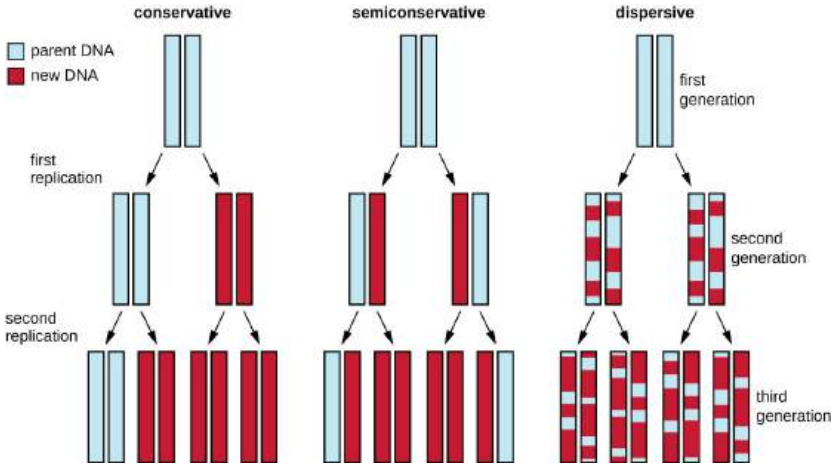


Figure 24.1.1: Three Models of DNA replication. In the ***conservative model,*** parental DNA strands (blue) remained associated in one DNA molecule while new daughter strands (red) remained associated in newly formed DNA molecules. In the ***semiconservative model,*** parental strands separated and directed the synthesis of a daughter strand, with each resulting DNA molecule being a hybrid of a parental strand and a daughter strand. In the ***dispersive model***, all resulting DNA strands have regions of double-stranded parental DNA and regions of double-stranded daughter DNA. Figure by Parker, N., et.al. (2019) Openstax

Matthew Meselson and Franklin Stahl devised an experiment in 1958 to test which of these models correctly represents DNA replication, as shown in Figure 24.1.2 They grew the bacterium, *Escherichia coli* for several generations in a medium containing a "heavy" isotope of nitrogen (^{15}N) that was incorporated into nitrogenous bases and, eventually, into the DNA. This labeled the parental DNA. The *E. coli* culture was then shifted into a medium containing ^{14}N and allowed to grow for one generation. The cells were harvested and the DNA was isolated. The DNA was separated by ultracentrifugation, during which the DNA formed bands according to its density. DNA grown in ^{15}N would be expected to form a band at a higher density position than that grown in ^{14}N. Meselson and Stahl noted that after one generation of growth in ^{14}N, the single band observed was intermediate in position in between DNA of cells grown exclusively in ^{15}N or ^{14}N. This suggested either a semiconservative or dispersive mode of replication. Some cells were allowed to grow for one more generation in ^{14}N and spun again. The DNA harvested from cells grown for two generations in ^{14}N formed two bands: one DNA band was at the intermediate position between ^{15}N and ^{14}N, and the other corresponded to the band of ^{14}N DNA.

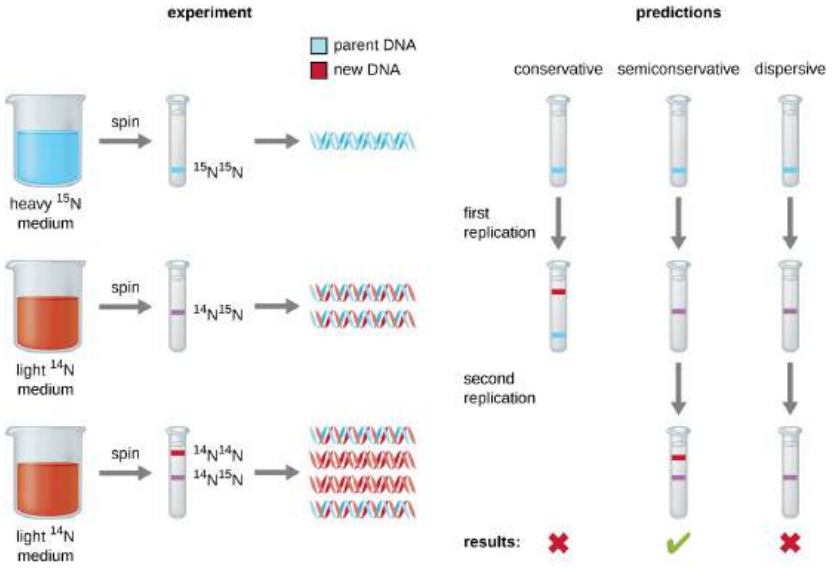


Figure 24.1.2: Meselson and Stahl experimented with *E. coli* grown first in heavy nitrogen (^{15}N) then in ^{14}N. DNA grown in ^{15}N (blue band) was heavier than DNA grown in ^{14}N (red band), and sedimented to a lower level on ultracentrifugation. After one round of replication, the DNA sedimented halfway between the ^{15}N and ^{14}N levels (purple band), ruling out the conservative model of replication. After a second round of replication, the dispersive model of replication was ruled out. These data supported the semiconservative replication model. Figure by Parker, N., et.al. (2019) *Openstax*

These results could only be explained if DNA replicates in a semiconservative manner. Therefore, the other two models were ruled out. As a result of this experiment, we now know that during DNA replication, each of the two strands that make up the double helix serves as a template from which new strands are copied. The new strand will be complementary to the parental or “old” strand. The resulting DNA molecules have the same sequence and are divided equally into the two daughter cells.

Think about It: What would have been the conclusion of the Meselson-Stahl experiment if, after the first generation, they had found two bands of DNA?

To synthesize double-stranded DNA, the parental strands must separate so DNA polymerases can copy both strands. As all DNA polymerases synthesize new DNA in a 5' to 3' direction from a 3' to 5' template, different mechanisms are used to faithfully synthesize both parental strands. The general mechanism is shown in Figure 24.1.3

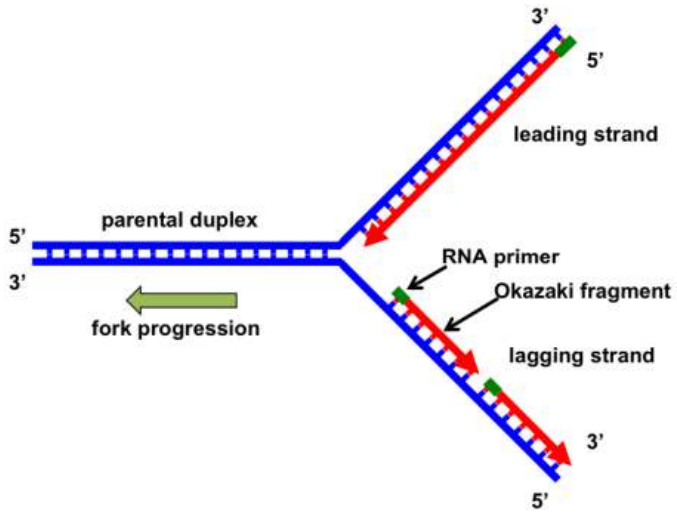


Figure 24.1.3 The replication fork. Leading-strand synthesis proceeds continuously in the 5' to 3' direction. Lagging-strand synthesis also occurs in the 5' to 3' direction, but in a discontinuous manner. An RNA/DNA primer (labeled in green) initiates leading-strand synthesis and every Okazaki fragment on the lagging strand.

Small RNA primers are needed for the new strands. Short (1000-2000 NT) DNA (Okazaki) fragments are made on the 3'-5' parental strand. Ultimately the RNA primers are degraded and filled, and the Okazaki fragments ligated. We will discuss replication in detail

for E. Coli, a model prokaryote, followed by replication in eukaryotes.

Figure 24.1.4 shows a general overview of a DNA "replication fork" from where DNA strand synthesis proceeds..

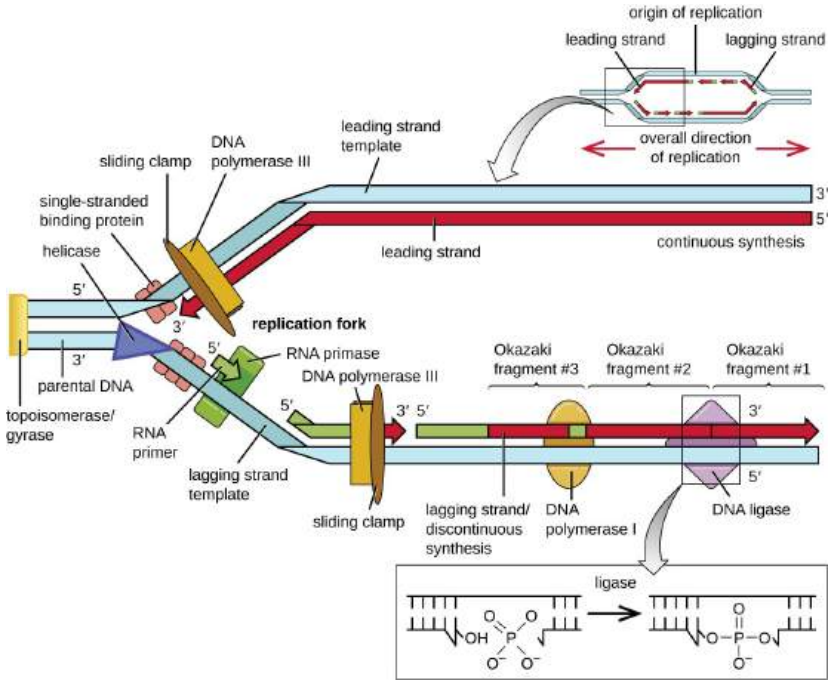


Figure 24.1.4: General Overview of a DNA Replication Fork. At the origin of replication, topoisomerase II relaxes the supercoiled chromosome. Two replication forks are formed by the opening of the double-stranded DNA at the origin, and helicase separates the DNA strands, which are coated by single-stranded binding proteins to keep the strands separated. DNA replication occurs in both directions. An RNA primer complementary to the parental strand is synthesized by RNA primase and is elongated by DNA polymerase III through the addition of nucleotides to the 3′-OH end. On the leading strand, DNA is synthesized continuously, whereas on the lagging strand, DNA is synthesized in short stretches called Okazaki fragments. RNA primers within the lagging strand are removed by the exonuclease activity of DNA polymerase I, and the Okazaki fragments are joined by DNA ligase. Figure by Parker, N., et.al. (2019) Openstax

24.1.2: DNA Replication in E. Coli

DNA replication has been well studied in bacteria primarily because of the small size of the genome and the mutants that are available. *E. coli* has 4.6 million base pairs (Mbp) in a single circular chromosome and all of it is replicated in approximately 42 minutes, starting from a single ***origin of replication*** and proceeding around the circle bidirectionally (i.e., in both directions), as shown in Figure 24.1.5. This means that approximately 1000 nucleotides are added per second. The process is quite rapid and occurs with few errors. *E. coli* has a single ***origin of replication***, called *oriC*, on its one chromosome. The origin of replication is approximately 245 base pairs long and is rich in adenine-thymine (AT) sequences.

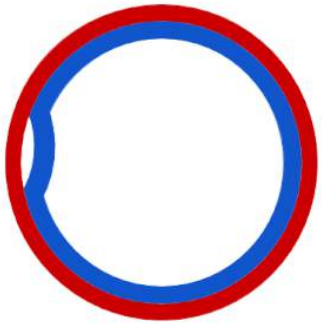


Figure 24.1.5: Prokaryotic DNA Replication. Replication of DNA in prokaryotes begins at a single origin of replication, shown in the figure to the left, and proceeds in a bidirectional manner around the circular chromosome until replication is complete. The bidirectional nature of replication creates two replication forks that are actively mediating the replication process. The right-hand figure shows a dynamic model of this process. The red and blue dots represent the incorporation of daughter strand nucleotides during the process of replication. Figures from: Daniel Yuen at David Tribe Derivatives and Catherinea228

24.1.2.1: Replication Overview - E. Coli

The open regions of DNA that are actively undergoing replication are called **replication forks**. All the proteins involved in DNA replication aggregate at the ***replication forks*** to form a replication complex called a **replisome.** The initial assembly of the complex that initiates primer synthesis is called the **primosome**. Table 24.1.1 below show the components that assemble at the replication fork to form the E. Coli replisome.

E. coli Gene	Enzyme/Protein Function	Description
dnaA	Initiator Protein	Melts DNA at oriC, exposing two template ssDNA strands
dnaB	Helicase	Unwinds the DNA helix at the front end of each replication fork during replication
dnaC	Helicase Loader	Loads the DnaB Helicase onto the ssDNA template strands
dnaG	Primase	Synthesizes RNA primers used to initiate DNA synthesis
dnaE	α-Catalytic Subunit of DNA Polymerase III	Catalytic subunit of the main replicative polymerase during DNA replication
dnaQ	ε-Editing Subunit of DNA Polymerase III	Editing subunit of the main replicative polymerase during DNA replication
dnaN	β-clamp subunit of DNA Polymerase III	Clamping subunit of the main replicative polymerase during DNA replication
polA	DNA Polymerase I	Processes Okazaki fragments and also fills in gaps during DNA repair processes
polB	DNA Polymerase II	Proofreading and editing, especially on lagging strand synthesis and some involvement in DNA repair
ssb	Single Stranded Binding Proteins (SSB)	Bind with single-stranded regions of DNA in the replication fork and prevent the strands from rejoining
A dimer encoded by *gyrA* and *gyrB*	DNA Gyrase	Type II Topoisomerase involved in releiving positive supercoiling tension caused by the action of Helicase
A dimer encoded by *parC* and *parE*	Topoisomerase IV	Type II Topoisomerase involved in decatenation of daughter chromosomes during DNA replication
ligA	DNA Ligase	Fixes nicks in the DNA backbone during DNA replication, DNA damage, and DNA repair processes
Note: Only the genes involved in the formation of the catalytic domain of DNA polymerase III are listed		

Table 24.1.1: **Enzymes involved in DNA Replication in the prokaryote, E. coli**

In *E. coli*, DNA replication is initiated at the single origin of replication, *oriC.* Binding of the initiator protein, **DnaA**, locally unfolds the DNA to form two template ssDNA, which bind **DnaB helicase.** A DnaB hexamer adds to each strand in a process promoted by **DnaC**, a helicase loader. The **single-stranded DNA binding protein B (SSPB)**binds to and protects the rest of the ssDNA, preventing further binding by DnaB. The **primase, DnaG**, is recruited to the site by the DnaB hexamer and synthesizes the RNA primers. DnaB also recruits DNA polymer III holoenzyme (PolIII HE) which binds through a β clamp. All of the bound proteins collectively form the replisome. An overview of E. Coli **replisome** is shown in Figure 24.1.6

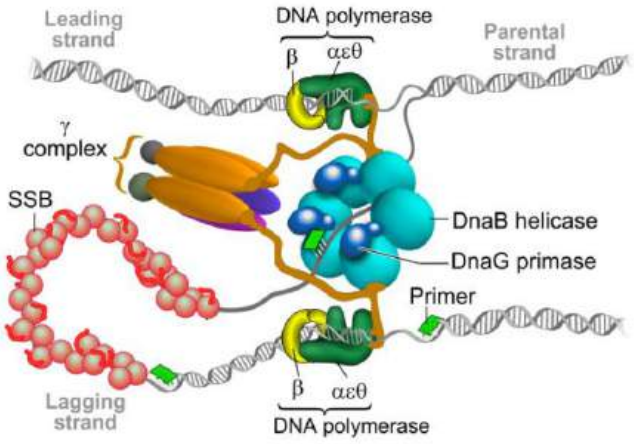


Figure 24.1.6: The bacterial replisome. Ilic, S.; Cohen, S.; Singh, M.; Tam, B.; Dayan, A.; Akabayov, B. DnaG Primase—A Target for the Development of Novel Antibacterial Agents. *Antibiotics* **2018**, *7*, 72. https://doi.org/10.3390/antibiotics7030072 Creative Commons Attribution (CC BY) license (**http://creativecommons.org/licenses/by/4.0/**)

Once assembled, replisomes move in opposite directions from the single oriC in the E. Coli chromosome. They meet at the opposite ends at a **termination site (ter)** to which ***Tus*** proteins are bound that create 'replication fork traps'. After completion of DNA replication, the newly synthesized genomes are separated and segregated to daughter cells.

An alternative term, the **primosomes**, is used to describe a subcomplex of the replisome which starts replication of the E. Coli chromosome, as well as some phages and plasmids. It contains 6 proteins including helicases and primases, and catalyzes the movement of the replication form by unwinding and primer synthesis. The motor protein **helicases**, use ATP to move along the ds-DNA backbone, unraveling it as it proceeds. The human genome has genes for 64 RNA and 31 DNA helicases (about 1% of eukaryotic genes).

24.1.2.2: Primase and Polymerase activities

The synthesis of both RNA strands by the DnaG primase, and DNA strands by DNA polymerase III holoenzyme (pol III) occurs at each start site for an Okazaki fragment. Both enzymes bind to the conserved carboxy-terminal tail of the single-stranded DNA-binding protein (SSB). It turns out that they can both be bound simultaneously.

The primase (DnaG) has three domains:

- N-terminus that binds the template
- RNA polymerase domain
- C-terminus that binds helicase and the C-terminus of SSB.

Primase is displaced by polIII after about 10 nucleotides have been added to the RNA primer so DNA synthesis can now occur at the 3' end of the primer.

Figure 24.1.7 shows how the primase to polymerase switch is made.

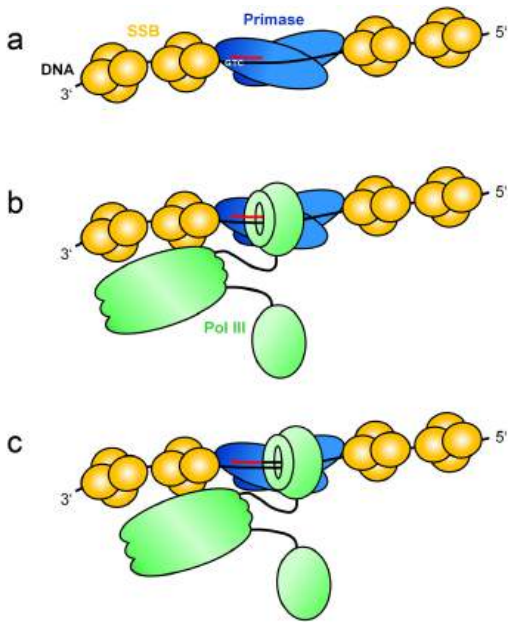


Figure 24.1.7: Schematic representation of the primase-to-polymerase switch during DNA replication in *E. coli*. Bogutzki, A., Naue, N., Litz, L. *et al. Sci Rep* **9,** 14460 (2019). https://doi.org/10.1038/s41598-019-51031-0. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/

Panel (**a**): Two primase molecules cooperate in the synthesis of the RNA primer (red).

Panel (**b**): For elongation of the primer, pol III enters the complex, whereupon primase and pol III are concurrently bound to the primed site, possibly via interactions with the C-termini of an adjacent SSB tetramer.

Panel (**c**): During pol III-mediated elongation of the primer by several nucleotides, both enzymes stay bound to the template. Only after the primer has been elongated by more than 10 nucleotides, one of the primases is released in the G4ori system. It is presumably the displacement of SSB by pol III that causes the consequent dissociation of primase. Whereas in the G4ori system, the two primases are positioned by hairpin structures that prevent SSB from binding to this part of the origin. At the *E. coli* replication fork, primases are brought into contact via their interaction with the replicative helicase DnaB.

What happens if the replication fork does not move to the termination site? If the DNA is damaged or if the replisome falls off of the chromosome, it can rebind and restart using **Pri** proteins. PriA is a DNA helicase that can bind to replication forks through DNA motifs and through interactions with SSBs. Other proteins involved include PriB, PriC, DnaT, DnaC, DnaB helicase, and DnaG primase as illustrated in Figure 24.1.8

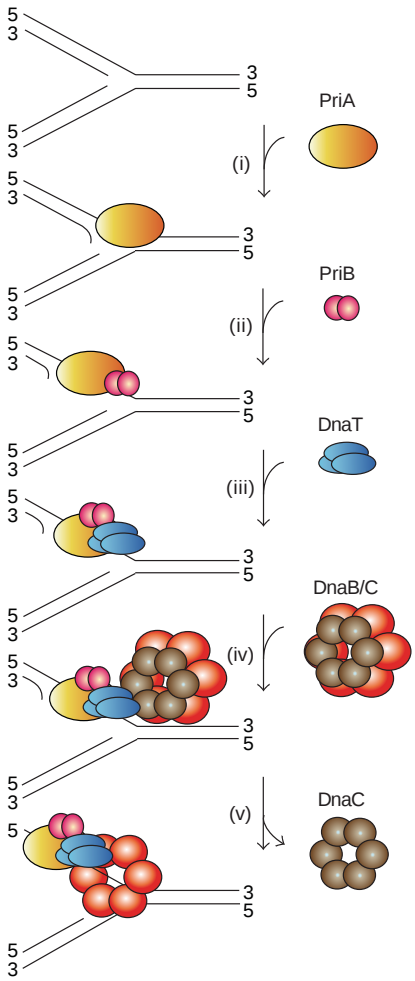


Figure 24.1.8 Primosome restart assembly. The proposed assembly mechanism is as follows. (i) PriA recognizes and binds to a replication fork, (ii) PriB joins PriA to form a PriA-PriB-DNA ternary complex, (iii) DnaT participates in this nucleocomplex to form a triprotein complex, in which PriB is released from ssDNA due to recruitment of DnaT, (iv) the PriA-PriB-DnaT-DNA quaternary complex loads the DnaB/C complex, and (v) DnaB is loaded on the lagging strand template. Yen-Hua Huang and **Cheng-Yang Huang.** BioMed Research International (2014). https://doi.org/10.1155/2014/195162. Huang and Cheng-Yang Huang. This is an open access article distributed under the Creative Commons Attribution License,

The primosome and replisome are complicated in structure and in their functional activity. Words go only so far in painting an image of how it works. To help we show a few different images of the replisome of E. Coli below.

The first is shown in Figure 24.1.9

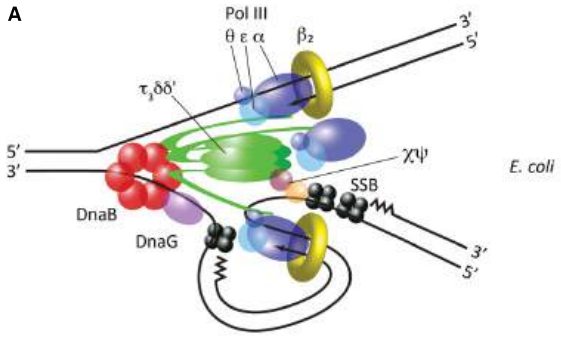


Figure 24.1.9: **Replisome architecture in bacteria. (A)** Architecture of the *E. coli* replisome, derived from *in vitro* studies and direct observation *in vivo*. Beattie TR, Reyes-Lamothe R. A. Front Microbiol. 2015 doi: 10.3389/fmicb.2015.00562. Creative Commons Attribution License (CC BY)

In this diagram, the leading strand is shown in the upper right end of the diagram. The central bottom loop shows the lagging strand. The $\tau_3\delta\delta'\psi\chi$ is the clamp loader and the DnaB (red) is hexameric. The ssDNA in the lagging strand loop is bound by ssDNA binding proteins (SSB).

Figure 24.1.10shows the rebind primosome which is mostly similar to the regular one.

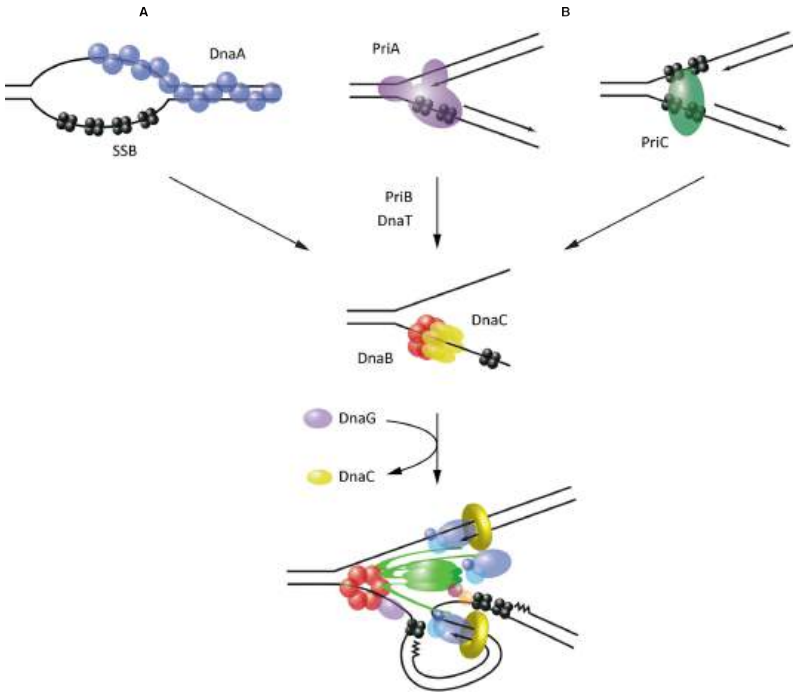


Figure 24.1.10 **Mechanisms of helicase loading leading to replisome assembly in *E. coli*. (A)**Recognition and melting of the *oriC* locus during initiation by DnaA. **(B)**Recognition of abandoned fork structures during replisome reloading by PriA and PriC. All pathways converge on the loading of the replicative helicase DnaB, which acts as an assembly platform for the remaining replisome components.

Finally, Figure 24.1.11shows models of DNA polymerase for lagging strand synthesis.

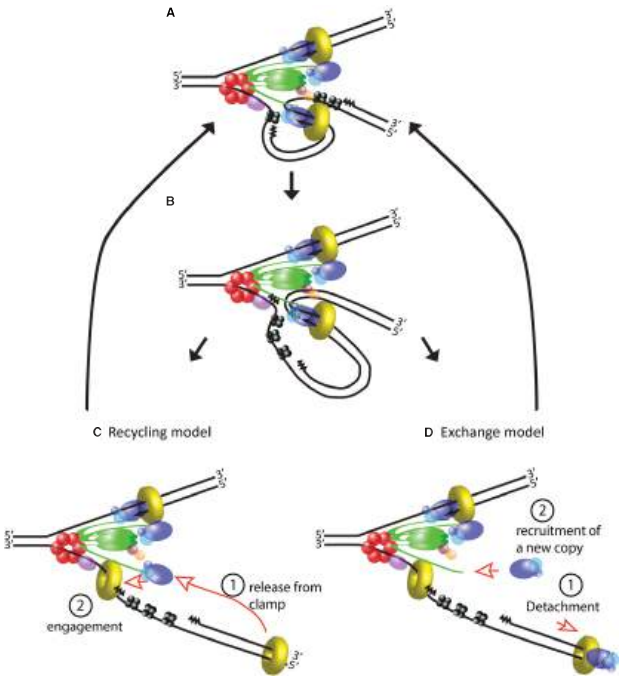


Figure 24.1.11: **Usage of DNA polymerase during lagging strand synthesis. (A)**S chematic of the *E. coli* replisome during the elongation step of an Okazaki fragment. **(B)**Lagging strand polymerase meets the RNA primer of the previous Okazaki fragment and stops synthesis. **(C)** Current model of events following completion of an Okazaki fragment. DNA polymerase is released from the β clamp (step 1) and the same molecule rebinds to a new β clamp to start the next Okazaki fragment (step 2). **(D)**An alternative model based on evidence from T4 and T7 replisomes. After completing the Okazaki fragment, the DNA polymerase detaches from the rest of the replisome (step 1). A new molecule of DNA polymerase is recruited to the replisome (step 2) and engages in the synthesis of a new Okazaki fragment. In this tentative model, a local pool of "spare" polymerases may facilitate their exchange and additional components may exchange along with the polymerase (not depicted)

24.1.3: E. Coli DNA Polymerases

E. Coli has 5 DNA polymerases. **DNA polymerase I** aids in lagging strand synthesis as it removes the RNA primers and incorporates DNA in its place. **DNA polymerase II,** may play an editing role following lagging strand synthesis by DNA polymerase I. DNA polymerases I and II also play a role in DNA repair, as do **DNA polymerases IV and V**.

DNA polymerases are shaped like a right hand in overall shape with three domains named palm, fingers, and thumb. The bottom of the cleft formed by the three domains forms the polymerase active site in the Palm domain, The monomeric nucleotides to be added bind through the finger domain, while the thumb domains facilitates the dissociation of the newly synthesized DNA. These features are illustrated for a polymerase that requires host thioredoxin for a bacteriophage T7 DNA polymerase in Figure 24.1.12

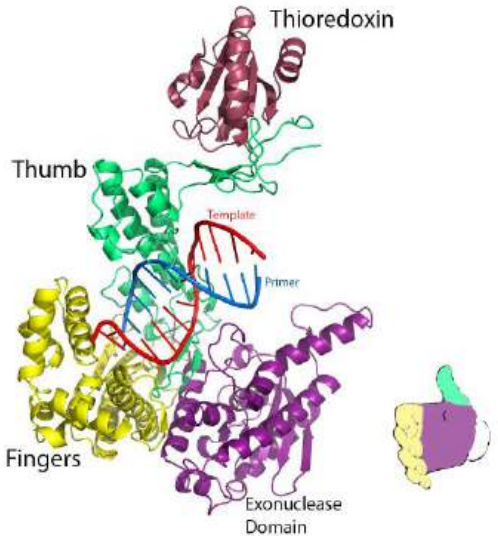


Figure 24.1.12 Structure of T7 DNA replication complex. Melum 103 - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/inde...curid=38408627

24.1.3.1: DNA Polymerase III

Pol III is a fascinating enzyme. It consists of an αεθ core with both 5'-3' polymerase and 3′−5′ proofreading activities, a β_2 ring-shaped "sliding clamp" that keeps the enzyme on the DNA track (processive) without iteratively jumping off and rebinding (distributive), and a $(\tau/\gamma)_3\delta\delta'\psi\chi$ clamp loader. The SSB protein has a conserved amphiphilic C-terminus that binds both DnaG (primase) and the χ subunit of the clamp loader. A

After primer addition by DnaG, the $\beta_2$clamp of polIII is brought to the end of the primer terminus by the clamp loader, after which α and ε subunits bind the clamp. The holoenzyme can add ∼1000 Nt/s and over 150 kb without falling off. Hence it is a very processive enzyme.

Figure 24.1.13 shows an interactive iCn3D model of the the E. coli replicative DNA polymerase III (alpha, $beta_2$, epsilon, tau complex) bound to DNA (5FKV)

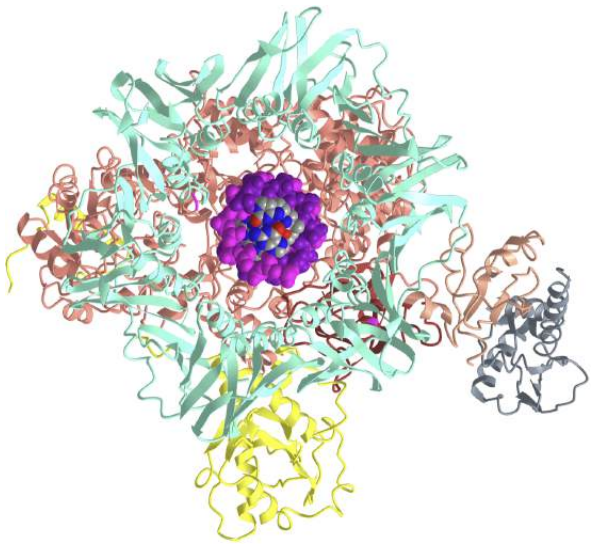


Figure 24.1.13: E. coli replicative DNA polymerase III (alpha, $beta_2$, epsilon, tau complex) bound to DNA (5FKV). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...vf4pCugQms7tw8

- N-terminal domains of α (αNTD, residues 1–963, are colored in salmon
- OB (964–1072) on the C-term domain of α (αCTD) colored brown,
- τ-binding domains (TBD, 1173–1160) on the C-term domain of α (αCTD) colored dark salmon,

- ε in yellow
- θ in orange (?)
- β_2 in aquamarine
- τ_C in slate gray
- DNA in spacefill, backbone magenta and purple, bases in CPK colors

Overall, there are significant conformational changes in the DNA Polymerase III complex upon binding to the DNA that cause the tail of the polymerase to move from interacting with the clamp in the DNA-bound state to a position 35 Å away from the clamp in the DNA-free state. It has been hypothesized that this large conformational change may help the polymerase act as a switch to facilitate the lagging strand synthesis. On the lagging strand, the polymerase repositions to a newly primed site every ∼1000 bp. To do so, the polymerase needs to release both clamp and DNA. The switch-like movement of the polymerase tail may play a part in the release and consequent repositioning of the polymerase at the end of the Okazaki fragment.

Video 25.1.1: **DNA Binding Induces Large Conformational Changes in the DNA Polymerase III Complex**(click link to view). The video shows the linear morphing of the DNA-free to the DNA-bound state showing the large conformation change between the two states. The green subunit is the β-clamp, The α-subunit is shown in orange with the active-site residues in magenta, the α-C-terminal domain (α-CTD shown in brown, the ε-subunit in yellow, and the τ-tail shown in blue. Video from: Fernandez-Liero, R., et al. (2015) eLife 4:e11134

The complex can also proofread the newly synthesized DNA. This requires some conformational changes in the polIII complex, including a rotation/tilt of the dsDNA against the β_2 ring-shaped "sliding clamp". The thumb domain moves between the two DNA strands containing a mismatch and produced a distorted DNA. The epsilon subunit, a nuclease, can reach the mismatched nucleotide and clip it off.

Figure 24.1.14shows an interactive iCn3D model of the E. coli replicative DNA polymerase III-clamp-exonuclease-theta complex bound to DNA in the editing mode (5M1S)

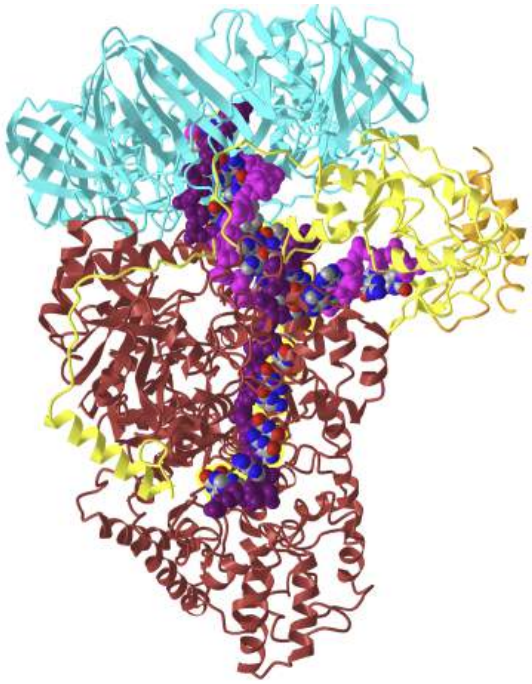


Figure 24.1.14: E. coli replicative DNA polymerase III-clamp-exonuclease-theta complex bound to DNA in the editing mode (5M1S). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...gedpgkHdqG1PGA

- PolIII α brown,
- PolIII ε in yellow
- PolIII θ in orange
- PolIII β_2 in cyan
- DNA in spacefill, backbone primer in magenta, the template in purple, and bases in CPK colors

24.1.3.2: DNA Polymerase I

DNA polymerase I, as does polIII, has a 5' to 3' polymerase activity. Also, both have a 3' to 5' exonuclease activity for proofreading as well as a 5'-3' exonuclease to remove RNA primers. It contains three domains, a 5'-3' exonuclease followed by a 3'-5' exonuclease, then the polymerase domain. Selective proteolysis between the first two domains produces the **Klenow fragment.** In contrast, the 5'-3' exonuclease of polIII is in the separate epsilon subunit.

Figure 24.1.15shows an interactive iCn3D model of the predicted AlphaFold structure of E. Coli DNA Polymerase I (P00582)

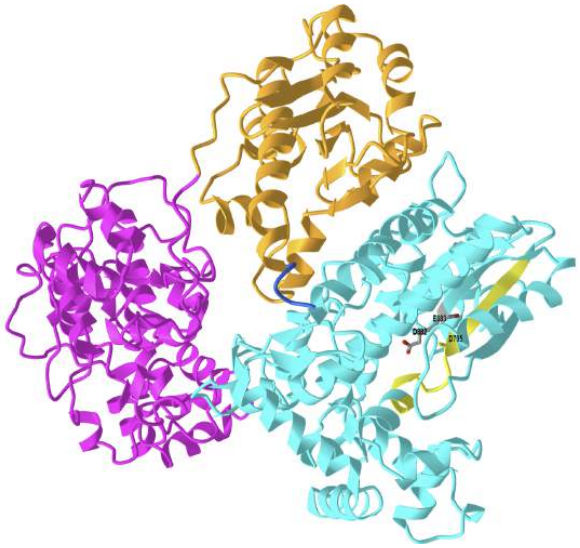


Figure 24.1.15: Predicted AlphaFold structure of E. Coli DNA Polymerase I (P00582). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...YhbSsUtpnkTfz8

https://structure.ncbi.nlm.nih.gov/i...oNSC2GJAsMEUk6

- 5' to 3' exonuclease, 1-323, magenta
- 3' to 5' exonuclease, 324-517, orange
- 5' to 3' polymerase, 521-928, cyan
- Val700-Arg713, Motif A, yellow
- Klenow fragment: 324-928

Key aspartate and glutamates involved in the polymerase active site are shown in sicks and labeled. Motif A is conserved in prokaryotic DNA polymerases. Essential roles of motif A in catalysis include interaction with the incoming dNTP and coordination with two divalent metal ions that are required for the polymerization reaction. Note the distance between the 3' to 5' exonuclease and the 5'-3' polymerase.

24.1.4: Other enzyme activities

24.1.4.1: DNA Ligase

DNA Ligase enzymes seal the breaks in the backbone of DNA that are caused during DNA replication, DNA damage, or during the DNA repair process. The biochemical activity of DNA ligases results in the sealing of breaks between 5′-phosphate and 3′-hydroxyl termini within a strand of DNA. DNA ligases have been differentiated as being ATP-dependent or NAD^+-dependent depending on the co-factor (or co-substrate) that is used during their reaction. Typically, more than one type of DNA ligase is found within an organism.

Figure 24.1.16shows the structure of *E. coli* LigA in complex with nicked adenylated DNA from PDB 2OWO, visualized by UCSF Chimera. The various domains are indicated by different colors and relate to Pfam domains indicated.

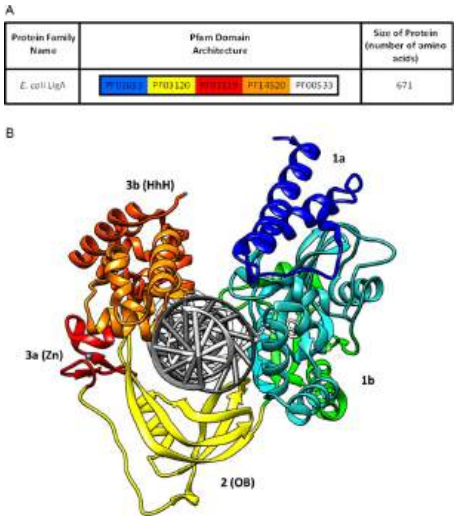


Figure 24.1.16 Structure of DNA ligase. Pergolizzi, G., Wagner, G.K, and Bowater, R.P. (2016) Biosci Rep 36(5) e00391

DNA ligase enzyme is covalently modified by the addition of the AMP moiety to a Lysine residue on the enzyme. The AMP derives from the ATP or NADH cofactor. The downstream 5'-phosphate at the site of the DNA nick is able to mediate a nucleophilic attack on the AMP-enzyme complex, causing the AMP to transfer to the 5'-phosphate position of the DNA. The AMP serves as a good leaving group for the nucleophilic attack of the upstream 3'-OH with the 5'-phosphate to seal the DNA backbone, and release the AMP. DNA ligase can use either adenosine triphosphate (ATP) or nicotinamide adenine dinucleotide (NAD^+) as a cofactor. Figure 24.1.17shows a mechanism of the ligation reaction, which is powered by ATP hydrolysis.

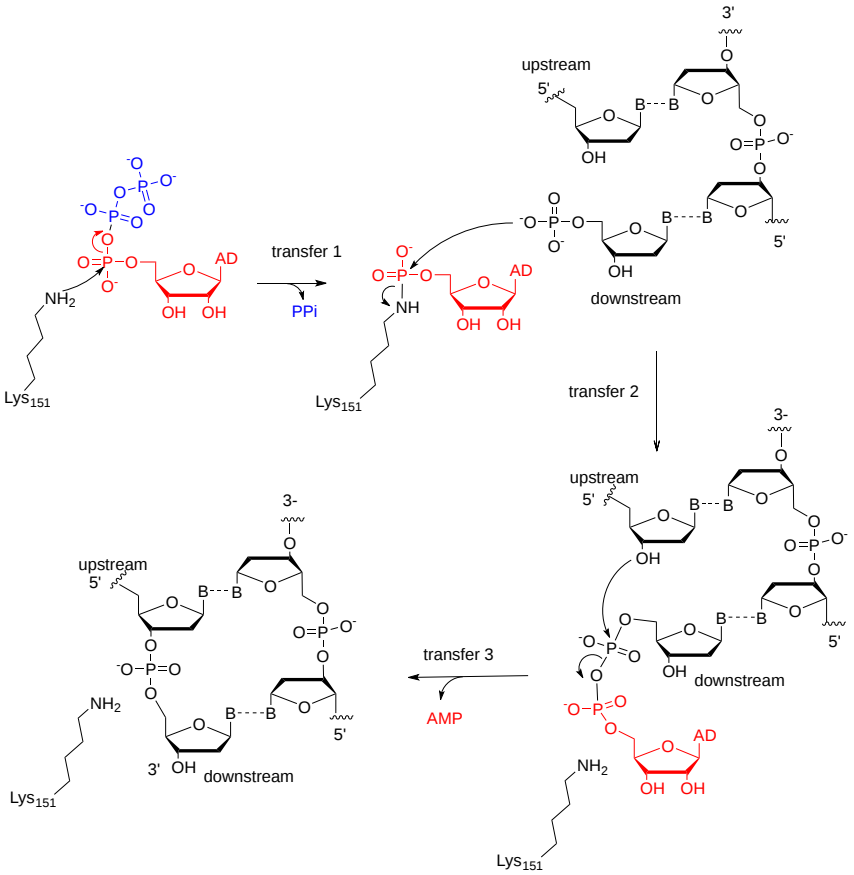


Figure 24.1.17: DNA Ligase Reaction. After Showalter, A. (2002). .

Figure 24.1.18shows an interactive iCn3D model of Human DNA Ligase I bound to 5'-adenylated, nicked DNA (1X9N)

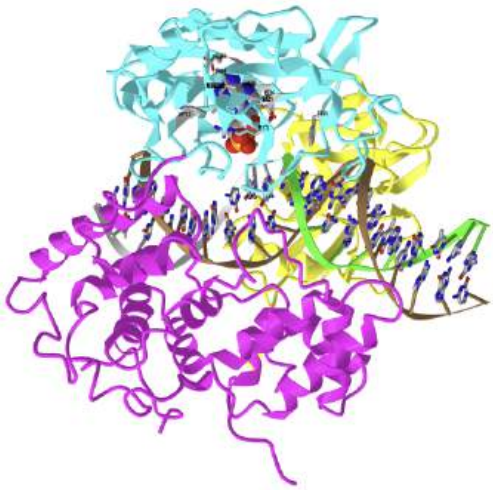


Figure 24.1.18: Human DNA Ligase I bound to 5'-adenylated, nicked DNA (1X9N). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...VNgobsRuLTsXZA

The protein has three domains:

- DNA binding domain (DBD), shown in magenta. In contrast to most DNA binding proteins, Ligase I binds to the *minor* grove around the area of DNA damage.
- Adenylation domain (AdD), shown in cyan, has covalently attached cofactor AMP and key catalytic residues shown as sticks and labeled. It ligates the broken DNA and forms a phosphodiester bond.
- The OB-fold domain (OBD), shown in yellow, facilitates catalysis as it binds and unwinds over the short region.

The DNA strands are as follows:

- ss DNA terminated with dideoxy is shown in green
- ss template DNA is shown in brown
- ss 5'-phosphorylated DNA is shown in gray

The AdD and OBD domains are similar in structure to other covalent nucleotidyltransferases involved in DNA and RNA ligation and capping of messenger RNA. Glu 566, Glu 621, and Arg 573 interact with AMP and probably help determine the specificity of the AMP cofactor (over GTP). Divalent cations are required for catalysis but are not present in the above structure. A E566K mutation leads to severe immunodeficiency. Lys 568 forms the covalent AMP adduct. The dideoxynucleoside in the structure is not optimally positioned for reaction with 5'P of AppDNA. Glu 720 and Glu 621 are highly conserved and presumably involved in metal ion binding.

Zoom into the structure to observe the changes at the 3' end of the dideoxy DNA and 5' end of the phosphorylated DNA. These nucleotides are given the abbreviation X in iCn3D as they are modified.

24.1.4.2: Topoisomerases

We just studied these in a previous section but here is a review for this section. The unwinding of the double-stranded helix at the replication fork generates winding tension in the DNA in the form of positive supercoils upstream of the replication fork. Enzymes called **topoisomerases** counteract this by introducing negative supercoils into the DNA in order to relieve this stress in the helical molecule during replication. There are four known topoisomerase enzymes found in *E. coli* that fall into two major classes, Type I Topoisomerases and Type II Topoisomerases, as shown in Figure 24.1.19 Topoisomerase I and III are Type I topoisomerases, whereas DNA gyrase and Topoisomerase IV are Type II topoisomerases.

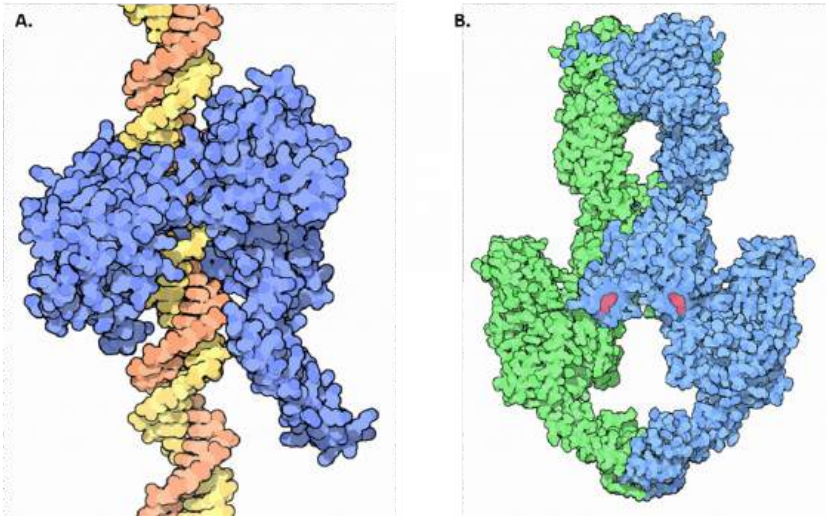


Figure 24.1.19: Structures of Type I and Type II Topoisomerase. (A) Crystal structure of a Type I Topoisomerase, shown in blue, bound to DNA, shown in orange and yellow. (B) Crystal structure of a Type II Topoisomerase. Topoisomerase II forms a tetrameric structure, shown in green and blue. ATP cofactors (pink) are shown bound to the enzyme.

Goodsell, D.S. (2015) RCSD PDB-101 Molecule of the Month

Type I Topoisimerase

Type I Topoisomerases relieve tension caused during the winding and unwinding of DNA. One way that they can do this is by making a cut or nick in one strand of the DNA double helix as shown in Figure 24.1.19 The 5'-phosphoryl side of the nicked DNA strand remains covalently bound to the enzyme at a tyrosine residue, while the 3'-end is held noncovalently by the enzyme. The Type I topoisomerases rotate or spin the 3'-end of the DNA around the intact DNA strand. This releases the overwinding in the DNA and effectively releases tension. The enzyme completes the reaction by resealing the phosphodiester backbone or ***ligating*** the broken strand back together. Overall, only one strand of the DNA is broken during the reaction mechanism and there is ***NO*** requirement of ATP during the reaction. The *E. coli* Topo I enzyme can only remove negative DNA supercoils, but not positive ones. Thus, this enzyme is not involved in relieving the positive supercoiling caused by the DNA helicase during replication. This is in contrast to eukaryotic Topo I that can relieve both positive and negative supercoiling. Although *E. coli* Topoisomerase I is not directly involved in relieving the tension caused by DNA replication, it is essential for *E. coli* viability. It is thought to help balance the actions of the Type II topoisomerases and help maintain optimal supercoiling density within the chromosomal DNA. Thus, Topo I is thought to help maintain the homeostatic balance of chromosome supercoiling within *E. coli.* Topo III, which is also a Type I Topoisomerase, appears to play a role in the decatenation of the daughter chromosomes during DNA replication, but does not play a role in the relaxation of supercoiling.

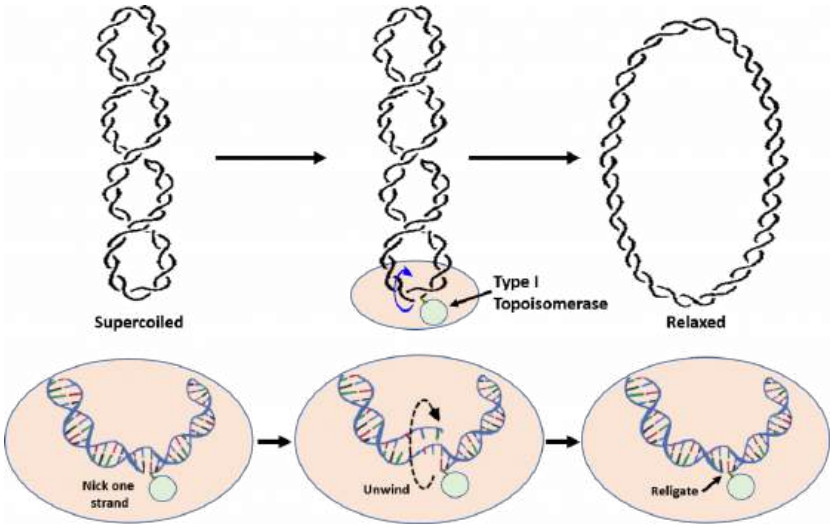


Figure 24.1.19: Reaction of Type I Topoisomerases. During the reaction, Type I Topoisomerases nick one strand of the DNA. One end of the nicked DNA is covalently attached to the enzyme, shown in light green in the lower diagrams. The other end is held non-covalently and rotated around the double helix to unwind the supercoiling and relax the DNA. Once supercoiling tension is released, the backbone is resealed and the Topoisomerase I enzyme is released. emixed from: Notahelix and JoKalliauer, and the National Human Genome Research Institute

Type II Topoisomerases have multiple functions within the cell. They can increase or decrease winding tension within the DNA or they can unknot or decatanate DNA that has become tangled with another strand as shown in Figure 24.1.20 It does so by a more dangerous method than their Type I counterparts, by breaking both strands of the DNA during their reaction mechanism. The enzyme is covalently attached to both broken sides while the other DNA helix is passed through the break. The double-stranded break is then resealed.

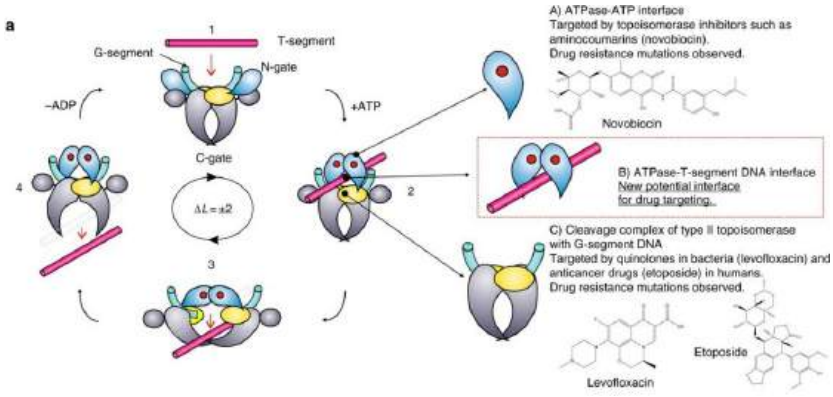


Figure 24.1.20: Reaction of Type II Topoisomerases. Figure modified fromLaponogov, I., et al. (2018) *Nature Communications* 9:2579. .

The proposed type II topoisomerase reaction cycle is exemplified by topoisomerase IV. Topoisomerase IV subunits are denoted in grey, cyan, and yellow. The gate or G-DNA is in green and the transported or T-DNA is in mauve. ATP bound to the ATPase domains is denoted by a red dot. In step 1, the G-DNA binds with the enzyme. ATP and the T-DNA segment associated with the enzyme in step 2. In step 3, the G-DNA is cleaved and the T-DNA is passed through the break. Drug-targetable domains within the type II topoisomerase complex are highlighted in subsections A, B, and C with examples on the right-hand side of the figure.

Type II Topoisomerase - D

DNA gyrase is the type II topoisomerase enzyme that is primarily involved in relieving positive supercoiling tension that results due to the helicase unwinding at the replication fork. Type II Topoisomerases, especially Topo IV, also address a key mechanistic challenge that faces the bacterial replisome during the termination of DNA replication. The circular nature of the bacterial chromosome dictates that a pair of replisomes that initiate from a single origin of replication will eventually converge on each other in a head-to-head orientation. Positive supercoiling accumulates between the the two replisomes as they converge, but the activity of DNA gyrase, which normally removes positive supercoils, becomes limited by the decreasing amount of template DNA available. Instead, supercoils may diffuse behind the replisomes, forming precatenanes between newly replicated DNA; in *E.coli* these must be resolved by Topo IV for chromosome segregation to occur.

24.1.5: Termination of Replication

If starting replication is critically important and obviously highly controlled as illustrated above, then termination of replication must be equally critical, otherwise genome instability would arise. There is one discrete origin of replication in E. Coli, oriC, with a defined sequence. In contrast, there are 10, 23 base-pair, nonpalindromic **termination sites (Ter)**of slightly different sequences. These bind the **termination protein, Tus**. The affinity of Tus for the Ter site depends on the Ter sequence and, in general,is tight with a K_D in the picomolar range. There are two types of Ter-Tus complexes, one an **open**"**permissive**" conformation that allows replication to continue, and a **locked**, "nonpermissive" form that stops it. In the nonpermissive conformation, a key and conserved cytosine on the leading strand at a conserved GC base pair is flipped out into a cytosine binding pocket, which you can think of as a "stop sign" for replication.

If you think of the E. Coli circular chromosome as a clock with the oriC at 12 o'clock, there are 5 Ter sequences as the replication fork move counter-clockwise at about 7 o'clock and another 5 as the fork move clockwise at around 5 o'clock. The sequences run in opposite polarity to prevent the left-side replication fork from entering the right-hand side as it moves around the chromosome and vice versa. The replisome displaces the Tus proteins at the permissive Ter sites but stops at the nonpermissive site, where DnaB helicase unwinds the DNA, flipping out the cytosine as the locked conformation forms. These processes are illustrated in Figure 24.1.21

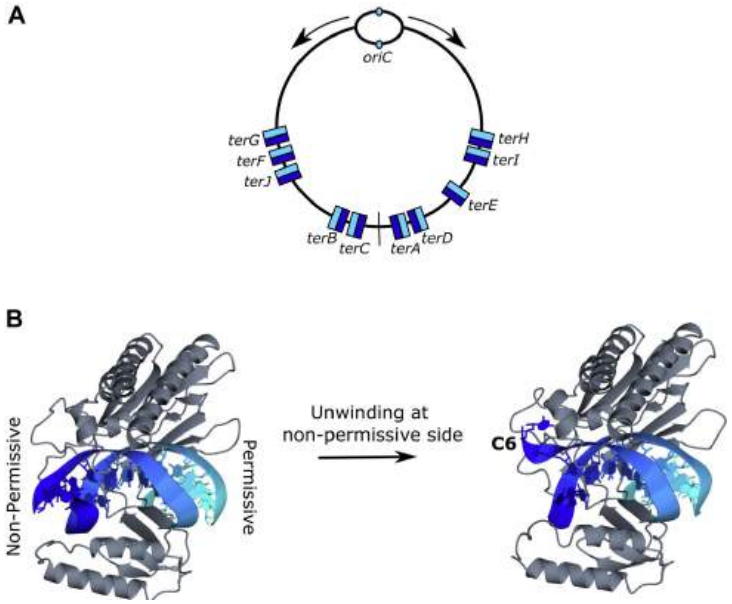


Figure 24.1.21: DNA replication and Tus-*ter* termination trap in *E. coli*. Katie H. Jameson et al., JBC, 297, (2021). DOI:https://doi.org/10.1016/j.jbc.2021.101409. Creative Commons Attribution (CC BY 4.0)

Panel A, *E. coli* contains a single circular chromosome, which replicates bidirectionally from a single origin (*small oval*). The direction of replisome travel from the origin is depicted by *arrows*. Chromosomal midpoint indicated by a *straight line*. The location of the *ter* sites on the *E. coli* chromosome is shown relative to *oriC*. Permissive orientation is displayed in *light blue*, nonpermissive orientation is displayed in *dark blue*.

Panel B, the structure of Tus-*ter* (PDB ID: 2I06) illustrating the nonpermissive and permissive faces (*left*) and the “locked” conformation formed by DNA unwinding at the nonpermissive face (*right*). The cytosine base at position 6 of *ter* (C6), which flips into a specific binding site on the nonpermissive face of Tus to form the “lock,” is indicated.

Replication can proceed at the Ter site if the replisome is moving from the light blue to dark blue sequences of the TER site. The TER site hence exhibits polarity.

Panel A, *E. coli* contains a single circular chromosome, which replicates bidirectionally from a single origin (*small oval*). The direction of replisome travel from the origin is depicted by *arrows*. Chromosomal midpoint indicated by a *straight line*. The location of *ter* sites on the *E. coli* chromosome is shown relative to *oriC*. Permissive orientation is displayed in *light blue*, nonpermissive orientation is displayed in *dark blue*.

Panel B, structure of Tus-*ter* (PDB ID: 2I06) illustrating the nonpermissive and permissive faces (*left*) and the "locked" conformation formed by DNA unwinding at the nonpermissive face (*right*). The cytosine base at position 6 of *ter* (C6), which flips into a specific binding site on the nonpermissive face of Tus to form the "lock," is indicated.

Figure 24.1.22Crystal structure (PDB code: 2I06) of the locked Tus–*Ter* complex shows the flipped C(6) base at the non-permissive face (5)

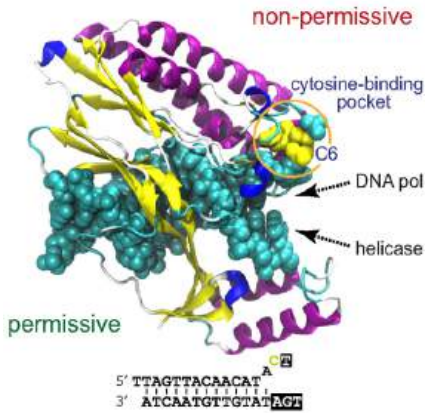
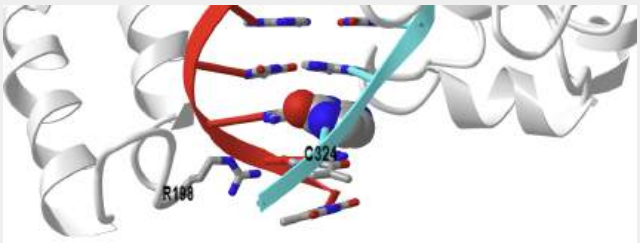
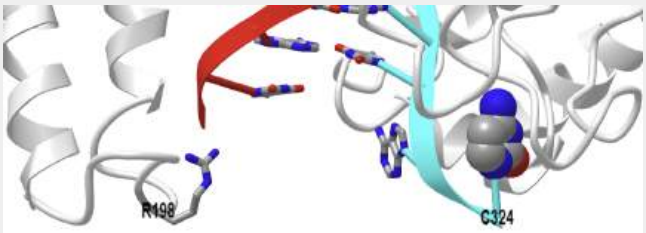


Figure 24.1.22 Crystal structure (PDB code: 2I06) of the locked Tus–*Ter* complex shows the flipped C(6) base at the non-permissive face (5). Pandey et al. Nucleic Acids Res. 2015 Jul 13; 43(12): 5924–5935. doi: 10.1093/nar/gkv527. Creative Commons Attribution License (http://creativecommons.org/licenses/by-nc/4.0/),

Figure 24.1.23shows interactive iCn3D models of the Escherichia Coli Replication Terminator Protein (Tus) Complexed With TerA DNA in open (left) or locked form (right).

Escherichia Coli Replication Terminator Protein (Tus) Complexed With TerA DNA (2I05)	Escherichia Coli Replication Terminator Protein (Tus) Complexed With DNA- Locked form (2I06)
	
(Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?7mc7UTqLBbgRZP8p8	(Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?bScTvhNGAMD456vt9

In the figures, the C6 that is flipped out in the locked nonpermissive form (right) is shown in spacefill and labeled as C324. The Arg 198 alters its orientation to allow the C6 to flip out and form the locked conformational form.

A summary of the process of DNA replication is shown in Video 25.1.2

24.1.5.0.1: Click Here to View Video

Video 9.2 Overview of the DNA Replication Process

Video from: Yourgenome, animated by Polymime Animation Company, Ltd

24.1.6: DNA Replication of Extrachromosomal Elements: Plasmids and Viruses

To copy their nucleic acids, plasmids and viruses frequently use variations on the pattern of DNA replication described for prokaryote genomes. We will focus here on one style known as the ***rolling circle method.***

Whereas many bacterial ***plasmids*** replicate by a process similar to that used to copy the bacterial chromosome, other plasmids, several ***bacteriophages***, and some ***viruses*** of eukaryotes use ***rolling circle*** replication as shown in Figure 24.1.24

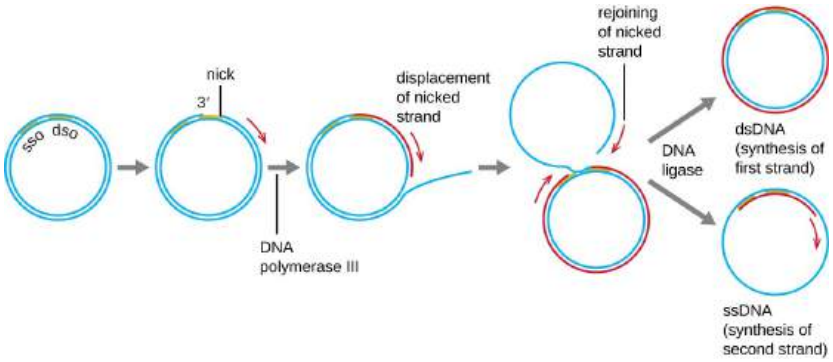


Figure 24.1.24 **Rolling Circle Replication.** The process of rolling circle replication is initiated by a single stranded nick in the DNA. Within prokaryotes, DNA polymerase III is utilized to generate the daughter strand. DNA ligase rejoins nicks in the backbone and enables the initiation of DNA synthesis of the second daughter strand. Figure by Parker, N., et.al. (2019) *Openstax*.

The circular nature of plasmids and the circularization of some viral genomes on infection make this possible. Rolling circle replication begins with the enzymatic nicking of one strand of the double-stranded circular molecule at the ***double-stranded origin (dso) site***. In bacteria, DNA polymerase III binds to the 3′-OH group of the nicked strand and begins to unidirectionally replicate the DNA using the un-nicked strand as a template, displacing the nicked strand as it does so. Completion of DNA replication at the site of the original nick results in the full displacement of the nicked strand, which may then recircularize into a single-stranded DNA molecule. RNA primase then synthesizes a primer to initiate DNA replication at the ***single-stranded origin (sso) site*** of the single-stranded DNA (ssDNA) molecule, resulting in a double-stranded DNA (dsDNA) molecule identical to the other circular DNA molecule.

24.1.7: DNA Replication in Eukaryotes

24.1.7.1: The Cell Cycle

The ***cell cycle*** is an ordered series of events involving cell growth and cell division that produces two new daughter cells. Cells on the path to cell division proceed through a series of precisely timed and carefully regulated stages of growth, DNA replication, and division that produce two genetically identical cells. The cell cycle has two major phases, ***interphase,*** and the ***mitotic phase***, as shown

in Figure 24.1.25 During interphase, the cell grows and DNA is replicated. During the mitotic phase, the replicated DNA and cytoplasmic contents are separated and the cell divides. Watch this video about the cell cycle: http://openstax.org/l/biocellcyc

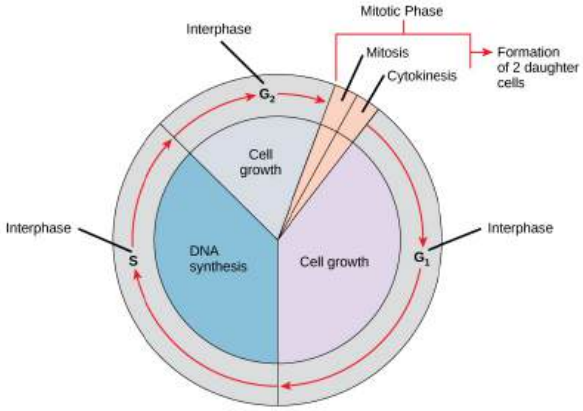
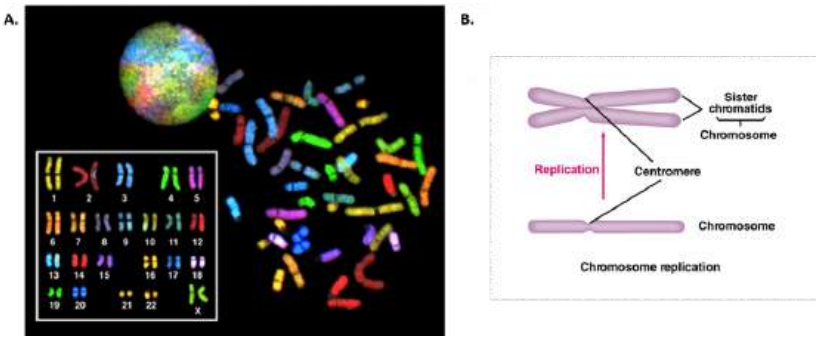


Figure 24.1.25 **Diagram of the Cell Cycle.** Fowler, S., et.al. (2013) *Openstax*

A cell moves through a series of phases in an orderly manner. During interphase, G_1 involves cell growth and protein synthesis, the S phase involves DNA replication and the replication of the centrosome, and G_2 involves further growth and protein synthesis. The mitotic phase follows interphase. Mitosis is nuclear division during which duplicated chromosomes are segregated and distributed into daughter nuclei. Usually, the cell will divide after mitosis in a process called cytokinesis in which the cytoplasm is divided and two daughter cells are formed.

During interphase, the cell undergoes normal processes while also preparing for cell division. For a cell to move from interphase to the mitotic phase, many internal and external conditions must be met. The three stages of interphase are called G_1, S, and G_2. The first stage of interphase is called the G_1 phase, or first gap, because little change is visible. However, during the G_1 stage, the cell is quite active at the biochemical level. The cell is accumulating the building blocks of chromosomal DNA and the associated proteins, as well as accumulating enough energy reserves to complete the task of replicating each chromosome in the nucleus. Throughout interphase, nuclear DNA remains in a semi-condensed chromatin configuration. In the S phase (synthesis phase), DNA replication results in the formation of two identical copies of each chromosome—***sister chromatids***—that are firmly attached at the centromere region,as shown in Figure 24.1.26 At this stage, each chromosome is made of ***two sister chromatids*** and is a duplicated chromosome. The centrosome is duplicated during the S phase. The two centrosomes will give rise to ***the mitotic spindle,*** the apparatus that orchestrates the movement of chromosomes during mitosis. In mammals, the centrosome consists of a pair of rod-like ***centrioles*** at right angles to each other. Centrioles help organize cell division. Centrioles are not present in the centrosomes of many eukaryotic species, such as plants and most fungi.



24.1.7.0.1:

Figure 24.1.26 **Human Chromosome Structure.** Figure A from: The National Human Genome Research Institute, and Figure B from: The School of Biomedical Sciences Wiki

(A) Shows a spectral karyogram of a normal human female. Humans have a total of 23 pairs of chromosomes for a total of 46. Each pair of chromosomes are referred to as ***homologous chromosomes*** as they contain copies of the same gene regions. Each of the homologous pairs of chromosomes is stained the same color. Chromosomes are shown in their condensed, unreplicated state. (B) Shows a schematic diagram of a single chromosome before (lower diagram) and after (upper diagram) replication. Upon replication, the identical copies of the chromosome are called ***sister chromatids*** and are linked together at the ***centromere*** structure.

Figure A from: The National Human Genome Research Institute, and Figure B from: The School of Biomedical Sciences Wiki

In the G_2 phase, or second gap, the cell replenishes its energy stores and synthesizes the proteins necessary for chromosome manipulation. Some cell organelles are duplicated, and the cytoskeleton is dismantled to provide resources for the mitotic spindle. There may be additional cell growth during G_2. The final preparations for the mitotic phase must be completed before the cell is able to enter the first stage of mitosis. To make two daughter cells, the contents of the nucleus and the cytoplasm must be divided. The mitotic phase is a multistep process during which the duplicated chromosomes are aligned, separated, and moved to opposite poles of the cell, and then the cell is divided into two new identical daughter cells. The first portion of the mitotic phase, ***mitosis,*** is composed of five stages, which accomplish nuclear division. The second portion of the mitotic phase, called ***cytokinesis***, is the physical separation of the cytoplasmic components into two daughter cells.

If cells are not traversing through one of the phases of interphase or mitosis, they are said to be in G_0 or a resting state. If cells enter G_0 permanently, they are said to have entered a stage of ***replicative senescence*** and will no longer be maintained for long-term viability within the organism.

The progression of cells through the cell cycle requires the coordinated actions of specific protein kinases, known as ***cyclin-dependent kinases***. Cyclin-dependent kinases are usually abbreviated as CDK or CDC proteins. CDK/CDC proteins require the binding of a regulatory cyclin protein to become activated, as shown in Figure 24.1.27. The major cyclin proteins that drive the cell cycle in the forward direction, are expressed only at discrete times during the cell cycle. When activated by a cyclin counterpart, CDK/CDC enzymes phosphorylate downstream targets involved with cell cycle progression. For example, the primary cyclin-CDK complex involved in the initiation of DNA replication during S-phase is the CyclinE-CDK2 complex. CDK2 is activated by the expression and binding of Cyclin E during late G1 phase. This causes CDK2 to phosphorylate downstream targets, including the retinoblastoma tumor suppressor protein, pRb. pRB normally binds and inhibits the activity of transcription factors from the E2F family. Following the release of E2F transcription factors from pRb, E2Fs activate the transcription of genes involved in DNA replication and the progression of cells into S-phase.

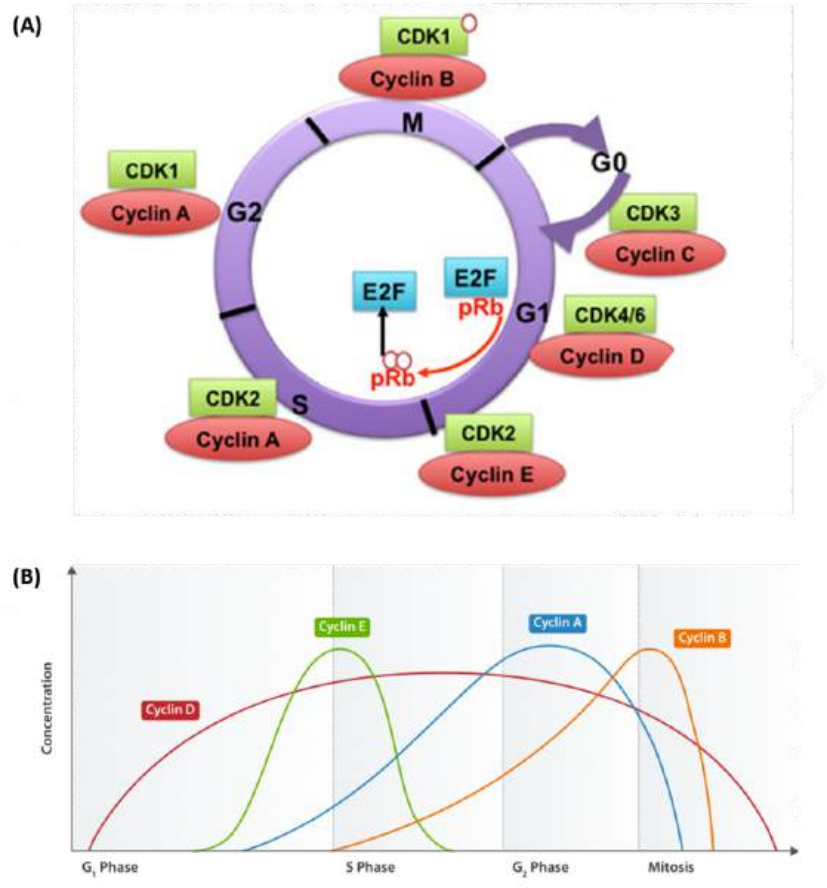


Figure 24.1.26: Cyclin Dependent Kinases (CDKs) and their Cyclin Regulatory Subunits. Figure A from: Aleem, E., and Arceci, R.J. (2015) *Front. Cell and Dev* 3 (16) and Figure B from: Cyclinexpression_waehrend-Zellzyklus

Panel (A) shows CDK-cyclin complexes with direct functions in regulating the cell cycle are shown. CDK3/cyclin C drives cell cycle entry from G_0. CDK4/6/cyclin D complexes initiate phosphorylation of the retinoblastoma protein (pRb) and promote the activation of CDK2/cyclin E complex. In late G1, CDK2/cyclin E complex completes phosphorylation and inactivation of pRb, which releases the E2F transcription factors and G1/S transition takes place. DNA replication takes place in S phase. CDK2/cyclin A complex regulates progression through S phase and CDK1/cyclin A complex through G2 phase in preparation for mitosis (M). Mitosis is initiated by CDK1/cyclin B complex.

Panel (B) Shows the cyclical nature of cyclin expression during cell cycle progression. Cyclin abundance is regulated by transcriptional expression and rapid protein degradation. Thus, their biological activity is targeted at very specific time points during the cell cycle progression.

24.1.7.1: Replication Initiation

Origin organization, specification, and activation in eukaryotes are more complex than in bacterial or archaeal kingdoms and significantly deviate from the paradigm established for prokaryotic replication initiation. The large genome sizes of eukaryotic cells, which range from 12 Mbp in *S. cerevisiae t*o 3 Gbp in humans, necessitates that DNA replication starts at several hundred (in budding yeast) to tens of thousands (in humans) origins to complete DNA replication of all chromosomes during each cell cycle, as shown in Figure 24.1.27.

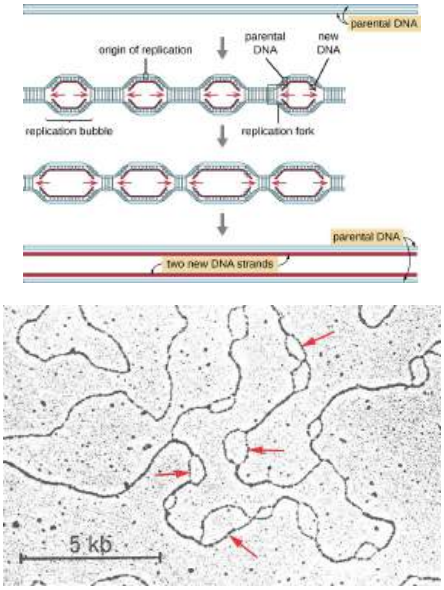


Figure 24.1.27: **Eukaryotic chromosomes are typically linear, and each contains multiple origins of replication**. The top figure is a graphic representation of the eukaryotic origins of replication, while the bottom image is a Cryo-electron micrograph image. The figure on the top is from *Parker, N. et al.* and the figure on the bottom is from *Fritensky, B. and Brien, N*

With the exception of *S.cerevisiae* and related *Saccharomycotina* species, eukaryotic origins do not contain consensus DNA sequence elements but their location is influenced by contextual cues such as local DNA topology, DNA structural features, and chromatin environment. Nonetheless, eukaryotic origin function still relies on a conserved initiator protein complex to load replicative helicases onto DNA during the late M and G1 phases of the cell cycle, a step known as ***origin licensing***. In contrast to their bacterial counterparts, replicative helicases in eukaryotes are loaded onto origin duplex DNA in an inactive, double-hexameric form and only a subset of them (10–20% in mammalian cells) is activated during any given S phase, events that are referred to as ***origin firing***. The location of active eukaryotic origins is therefore determined on at least two different levels, origin licensing to mark all potential origins, and origin firing to select a subset that permits assembly of the replication machinery and initiation of DNA synthesis. The extra licensed origins serve as backup and are activated only upon slowing or stalling of nearby replication forks, ensuring that DNA replication can be completed when cells encounter replication stress. Together, the excess of licensed origins and the tight cell cycle control of origin licensing and firing embody two important strategies to prevent under- and overreplication and to maintain the integrity of eukaryotic genomes.

24.1.7.2: Human Primosome

In humans,, the primosome contains primase and DNA polymerase α (Polα), and makes RNA-DNA primers to which deoxynucleotides are added by polymerases δ and ϵ. Hence there are two catalytic sites for addition or ribo- and deoxyribonucleotides. The structure of the human primosome and the C-terminal domain of the primase large subunit ($p58_C$) with bound DNA/RNA duplex is presented below. $p58_C$ coordinates the catalytic activities.

As with other polymerases, primase synthesis of RNA primers has the following steps:

- initiation (rate limiting): primase binds to DNA and makes a dinucleotide RNA;
- elongation, which is not as fast as DNA replication since it is less processive, adding only around 10 nucleotides. These short fragments are moved to Polα where deoxynucleotides are added with inactivation of the primase
- termination.

The structures of the enzymes are as follows:

Human Polα consists of a :

- large catalytic subunit (p180) with a C-terminal $p180_C$ domain with two Zn^{2+} binding modules.
- smaller accessory subunit (p70) with an N-terminal ($p70_N$), a phosphodiesterase, and oligonucleotide/oligosaccharide-binding (OB) domains.

Human primase consist of

- catalytic (p49)
- regulatory (p58) subunits with two domains, the N-terminal ($p58_N$), which interacts with p49 and which connects primase and Polα, and a C-terminal ($p58_C$) which contains an iron-sulfur cluster involved in substrate binding and primase activity.

The structures are shown in Figure 24.1.28

Figure 24.1.28 Structure of the human primosome hetero-tetramer complex. Baranovskiy et al. JBC, 291, 10006-10020 (2016). DOI: https://doi.org/10.1074/jbc.M116.717405. Creative Commons Attribution (CC BY 4.0)

Panel A shows a schematic representation of the domain organization. The flexibly tethered domains are shown as separate parts. $p58_C$ coordinates the iron-sulfur cluster. *Exo** is an exonuclease domain with no associated activity due to the evolutionary substitution of the catalytic amino acid residues; *PDE*, phosphodiesterase.

Pane B shows the crystal structure of the primosome. Subunits are shown as schematics and colored as in *A*. The α-carbons of catalytic aspartates are shown as *purple spheres*.

Figure 24.1.29shows an interactive iCn3D model of the Human primosome without nucleic acids (5EXR)

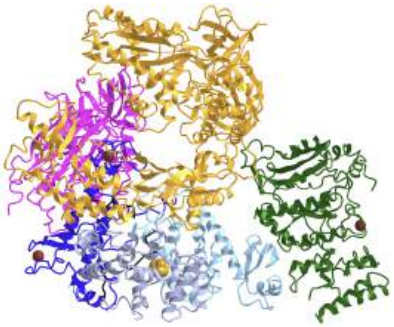


Figure 24.1.29: Human primosome without nucleic acids (5EXR). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...12dKTjz7hkcD97

DNA primase small: catalytic (p49) - dark green

DNA primase large: regulatory (p58) subunits. p58 has two distinct domains, N-terminal ($p58_N$ light blue) and C-terminal ($p58_C$ gray/purple), connected with an 18-residue linker (253–270) . $p58_N$ interacts with p49 and connects primase with Polα), and an iron-sulfur cluster containing $p58_C$ plays an important role in substrate binding and primase activity

DNA polymerase alpha catalytic subunit: large catalytic subunit (p180). has p180core (orange) and linker 1251-1265 then the C-terminal domain ($p180_C$ - blue) connects to small subunit p70. ($p180_C$) contains Zn1 and Zn2 bind site

DNA polymerase alpha subunit B: smaller accessory subunit (p70) with 3 domains: $p70_N$ (light green) then linker 79-156 BOTH NOT SHOWN IN STRUCTURE), the P70 phosphodiesterase, and oligonucleotide/oligosaccharide-binding (OB) domains (combined magenta).

Figure 24.1.30shows an interactive iCn3D model of the C-terminal domain of the **human DNA primase large subunit** with bound DNA template/RNA primer (5F0Q)

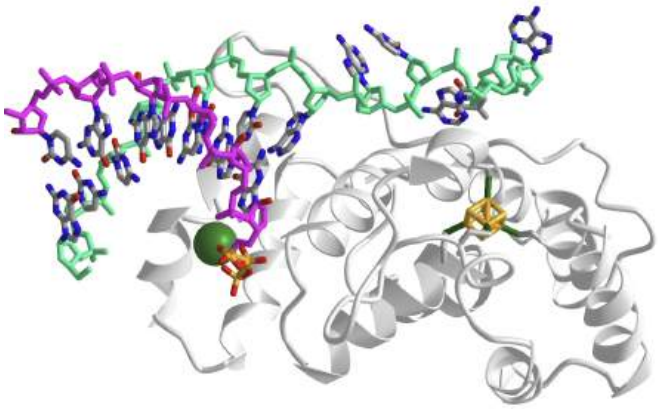


Figure 24.1.30: C-terminal domain of the human DNA primase large subunit with bound DNA template/RNA primer (5F0Q). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...jqV6UxcgeYNbM7

The ss-DNA is shown with a pale green backbone while the RNA backbone is shown in magenta. The FeS cluster and a Mg^{2+} ion are shown in the catalytic subunit. The Mg^{2+} is shown interacting with a terminal GTP of the RNA.

Figure 24.1.31shows an interactive iCn3D model of the catalytic core of human DNA polymerase alpha in a ternary complex with an RNA-primed DNA template and dCTP (4QCL)

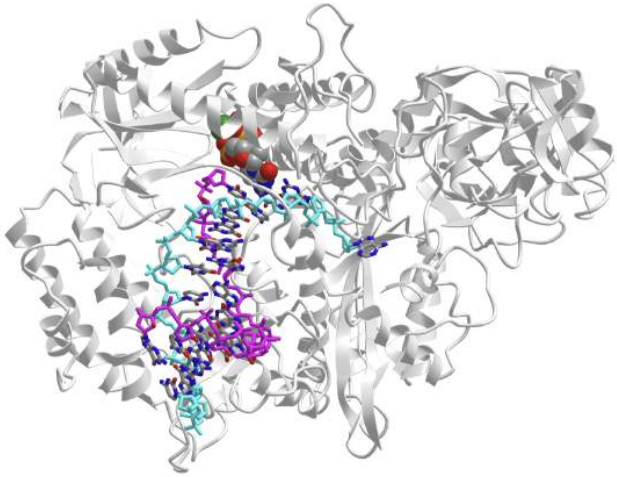


Figure 24.1.31: Catalytic core of human DNA polymerase alpha in ternary complex with an RNA-primed DNA template and dCTP (4QCL). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...gYFMnVSVJZiYA6

The ss-DNA is shown with a cyan backbone while the RNA primer backbone is shown in magenta. dCTP is shown in spacefill.

24.1.7.3: Eukaryotic DNA polymerases

Similar to DNA replication in prokaryotes, DNA replication in eukaryotes occurs in opposite directions between the two new strands at the replication fork. Within eukaryotes, two replicative polymerases synthesize DNA in opposing orientations, as shown in Figure 24.1.32 ***Polymerase ε (epsilon)*** synthesizes DNA in a continuous fashion, as it is "pointed" in the same direction as DNA unwinding. Similar to bacterial replication, this strand is known as the ***leading strand.*** In contrast, ***polymerase δ (delta)*** synthesizes DNA on the opposite template strand in a fragmented, or discontinuous, manner and this strand is termed the ***lagging strand***. The discontinuous stretches of DNA replication products on the lagging strand are known as Okazaki fragments and are about 100 to 200 bases in length at eukaryotic replication forks. Owing to the "lagging" nature, the lagging strand generally contains a longer stretch of ssDNA that is coated by single-stranded binding proteins, which stabilizes ssDNA templates by preventing secondary structure formation or other transactions at the exposed ssDNA. In eukaryotes, ssDNA stabilization is maintained by the heterotrimeric complex known as ***replication protein A (RPA)*** (Figure 9.19). Each Okazaki fragment is preceded by an RNA primer, which is displaced by the procession of the next Okazaki fragment during synthesis. In eukaryotic cells, a small amount of the DNA segment immediately upstream of the RNA primer is also displaced, creating a flap structure. This flap is then cleaved by endonucleases (such as Fen1, discussed later). At the replication fork, the gap in DNA after removal of the flap is sealed by ***DNA ligase I***. Owing to the relatively short nature of the eukaryotic Okazaki fragment, DNA replication synthesis occurring discontinuously on the lagging strand is less efficient and more time consuming than leading-strand synthesis.

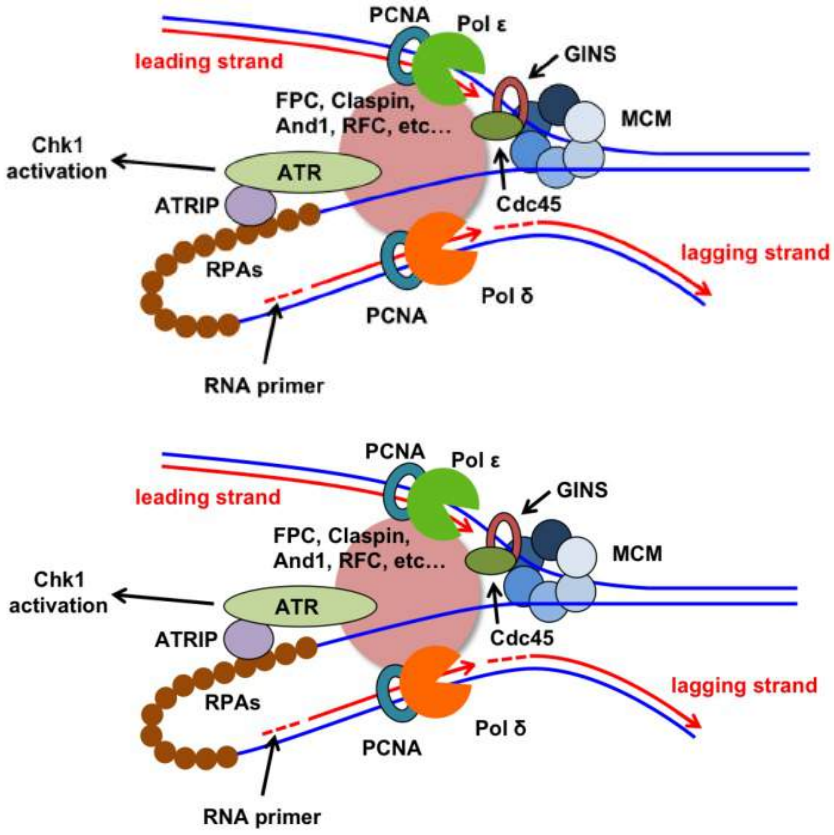


Figure 24.1.32: The Eukaryotic Replisome Complex Coordinates DNA Replication. *Lemanm A.R. and Noguchi, E. (2013) Genes 4(1):1-32*.

Replication on the leading and lagging strands is performed by Pol ε and Pol δ, respectively. Many replisome factors (including the FPC [fork protection complex], Claspin, And1, and RFC [the replication factor C clamp loader]) are charged with regulating polymerase functions and coordinating DNA synthesis with the unwinding of the template strand by Cdc45-MCM [mini-chromosome maintenance]-GINS [go-ichi-ni-san]. The replisome also associates with checkpoint proteins as DNA replication and genome integrity surveillance mechanisms.

Figure 24.1.33shows an interactive iCn3D model of the Core human replisome (7PFO). (long load time)

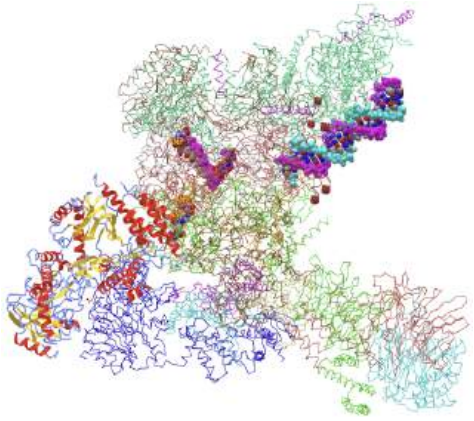


Figure 24.1.33: Core human replisome (7PFO). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...XQQijVnottcwa7 (long load time)

The leading DNA strand backbone is shown in spacefill magenta while the lagging strand backbone is shown in cyan. The DNA bases are shown as CPK spheres. The ATP analog, phosphaminophosphonic acid-adenlate ester, is shown in spacefill with CPK colors and

labeled. The C-alpha traces of the different protein subunits are all shown in different colored alpha-C traces, except the DNA polymerase epsilon catalytic subunit A which is shown in cartoon form and colored by secondary structure. (long load time)

At the eukaryotic replication fork, three distinct replicative polymerase complexes contribute to canonical DNA replication: α, δ, and ε. These three polymerases are essential for the viability of the cell. Because DNA polymerases require a primer on which to begin DNA synthesis, first, polymerase α (Pol α) acts as a replicative primase. Pol α is associated with an RNA primase and this complex accomplishes the priming task by synthesizing a primer that contains a short ~10-nucleotide RNA stretch followed by 10 to 20 DNA bases. Importantly, this priming action occurs at replication initiation at origins to begin leading-strand synthesis and also at the 5' end of each Okazaki fragment on the lagging strand.

However, Pol α is not able to continue DNA replication. From *in vitro* studies, it was observed that DNA replication must be "handed off" to another polymerase to continue synthesis. The polymerase switching requires clamp loaders. Initially, it was thought that Pol δ performed leading-strand replication and that Pol α completed each Okazaki fragment on the lagging strand. Using mutator polymerase variants and mapping nucleotide misincorporation events, Kunkel and colleagues found that Pol ε and Pol δ mutations lead to mismatched nucleotide incorporation only on the leading and lagging strands, respectively. Thus, normal DNA replication requires the coordinated actions of three DNA polymerases: Pol α for priming synthesis, Pol ε for leading-strand replication, and Pol δ for generating Okazaki fragments during lagging-strand synthesis.

In eukaryotes, DNA polymerases are grouped into seven families (A, B, C, D, X, Y, and RT). Crystal structures of the three nuclear replicative DNA polymerases demonstrate that they belong to the B family (Figure 25.1.17). All three replicative DNA polymerases are multi-subunit enzymes as shown in Table 24.1.2below.

Table 25.1.2 Subunits of the Major Eukaryotic Replicative DNA Polymerases

Polymerase	**Species**		**Function**
Polymerase α	***H. sapiens***	***S. cerevisiae***	
Catalytic or A-subunit	POLA1 (p180)	POL1	Catalytic subunit; polymerase activity; inactivated exonuclease
B-subunit	POLA2 (p70)	POL12	Regulatory subunit
Primase small subunit	PRIM1 (p49)	PRI1	Primase
Primase large subunit	PRIM2 (p58)	PRI2	Primase
Polymerase δ	***H. sapiens***	***S. cerevisiae***	
Catalytic or A-subunit	POLD1 (p125)	POL3	Catalytic subunit; has both polymerase and exonuclease activity
B-subunit	POLD2 (p50)	POL31	Accessory subunit
C-subunit	POLD3 (p66 or p68)	POL32	Accessory subunit
D-subunit	POLD4 (p12)	–	Accessory subunit
Polymerase ε	***H. sapiens***	***S. cerevisiae***	
Catalytic or A-subunit	POLE or POLE1	POL2	Catalytic subunit; has both polymerase and exonuclease activity
B-subunit	POLE2	DPB2	Accessory subunit
C-subunit	POLE3 (p17; CHRAC17)	DPB3	Accessory subunit
D-subunit	POLE4 (p12)	DPB4	Accessory subunit

Table 24.1.2 Subunits of the Major Eukaryotic Replicative DNA Polymerases. *Doublié, S. and Zahn, K.E. (2014) Front. Microbiol 5:444*

All B family polymerases are composed of five subdomains, the fingers, thumb, and palm which constitute the core of the enzyme, as well as an exonuclease domain and an N-terminal domain (NTD). The palm, a highly conserved fold composed of four antiparallel β

strands and two helices, harbors two strictly conserved catalytic aspartates, located in motif A, **D**XXLYPS and motif C, DT**D**S , as shown in Figure 24.1.34

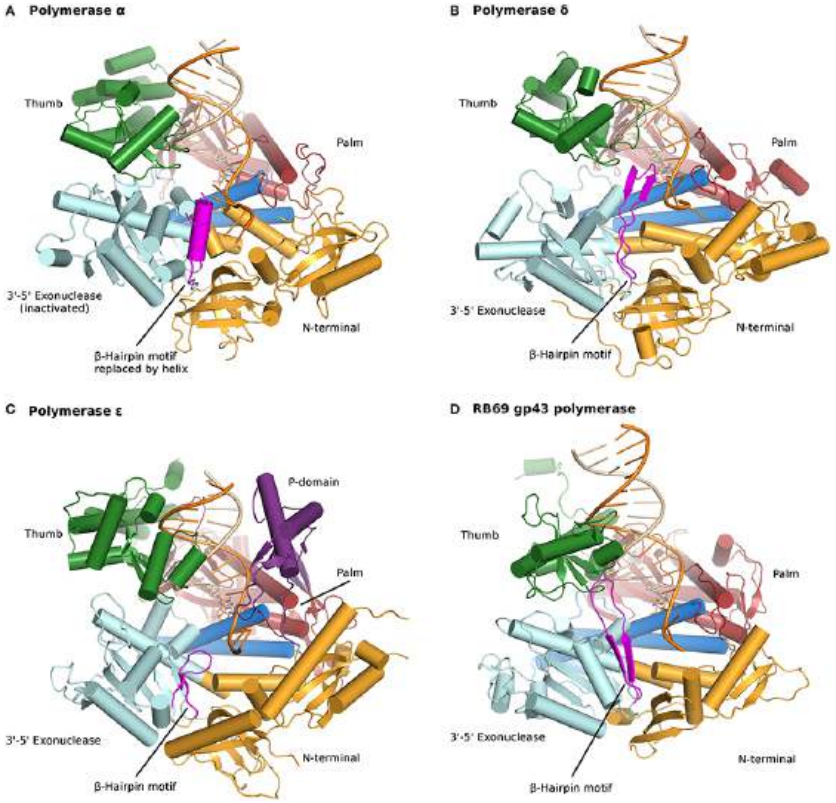


Figure 24.1.34: Structure of polymerases. *Doublié, S. and Zahn, K.E. (2014) Front. Microbiol 5:444*

This fold is shared by a very large group of enzymes, including DNA and RNA polymerases, reverse transcriptases, CRISPR polymerase, and even reverse (3′–5′) transferases. In contrast, the thumb and the fingers subdomains exhibit substantially more structural diversity. The fingers undergo a conformational change upon binding DNA and the correct incoming nucleotide. This movement allows residues in the finger subdomain to come in contact with the nucleotide in the nascent base pair. The thumb holds the DNA duplex during replication and plays a part in processivity. The exonuclease domain carries a 3′–5′ proofreading activity, which removes misincorporated nucleotides. The NTD seems to be devoid of catalytic activity. In pol δ the NTD comprises three motifs, one has a topology resembling an OB fold, one a single-stranded DNA binding motif, and the another has a RNA-binding motif (RNA Recognition Motif or RRM). The NTD likely plays a role in polymerase stability and fidelity through its interactions with other domains.

DNA polymerases require additional factors to support DNA replication *in vivo*. DNA polymerases have a semi-closed hand structure, which allows them to load onto DNA and translocate. This structure permits DNA polymerase to hold the single-stranded template, incorporate dNTPs at the active site, and release the newly formed double strand. However, the conformation of DNA polymerases does not allow for their stable interaction with the template DNA. To strengthen the interaction between template and polymerase, DNA sliding clamps have evolved, promoting the processivity of replicative polymerases. In eukaryotes, this sliding clamp is a homotrimer known as ***proliferating cell nuclear antigen (PCNA),*** which forms a ring structure. The PCNA ring has polarity with a surface that interacts with DNA polymerases and tethers them securely to DNA. PCNA-dependent stabilization of DNA polymerases has a significant effect on DNA replication because it enhances polymerase processivity up to 1,000-fold (Figure 25.1.19).

The ***DNA helicases (MCM proteins)*** and polymerases must also remain in close contact at the replication fork (Figure 25.1.19). If unwinding occurs too far in advance of synthesis, large tracts of ssDNA are exposed. This can activate DNA damage signaling or induce aberrant DNA repair processes. To thwart these problems, the eukaryotic replisome contains specialized proteins that are designed to regulate the helicase activity ahead of the replication fork. These proteins also provide docking sites for physical interaction between helicases and polymerases, thereby ensuring that duplex unwinding is coupled with DNA synthesis.

24.1.7.4: Control of Origin Firing

Origin usage in eukaryotes can be dynamic, with origin firing at different sites depending on cell type and developmental stage. Nevertheless, the mechanism of replisome assembly and origin firing is highly conserved. During late mitosis and G_1 phase, cell cycle proteins, such as Cdc6, associate with *Ori* sites throughout the genome and recruit the helicase enzymes, ***MCMs 2-7*** as shown in Figure 24.1.35 At this time, double hexamers of the MCM2-7 complex are loaded at replication origins. This generates a ***pre-replication complex (pre-RC)***. Origins with an associated pre-RC are considered licensed for replication. Licensed replication origins can then be "fired," when replication actually initiates at the *Ori*. Origin firing is brought about by multiple phosphorylation events carried out by the cyclin E-CDK2 complex at the onset of S phase and by other ***cyclin-dependent kinases (CDKs)*** prior to individual origin firing (Figure 24.1.35). ***Cyclin-dependent kinases (CDKs)*** are the families of protein kinases first discovered for their role in regulating the cell cycle. They are also involved in regulating transcription, mRNA processing, and the differentiation of nerve cells. CDKs are activated through the binding of an associated cyclin regulatory protein. Without a cyclin, CDKs exhibit little kinase activity. Following the phosphorylation of the pre-RC, origin melting occurs and DNA unwinding by the helicase generates ssDNA, exposing a template for replication (Figure 24.1.35). The replisome then begins to form with the localization of replisome factors such as Cdc45. DNA synthesis begins on the melted template, and the replication machinery translocates away from the origin in a bidirectional manner.

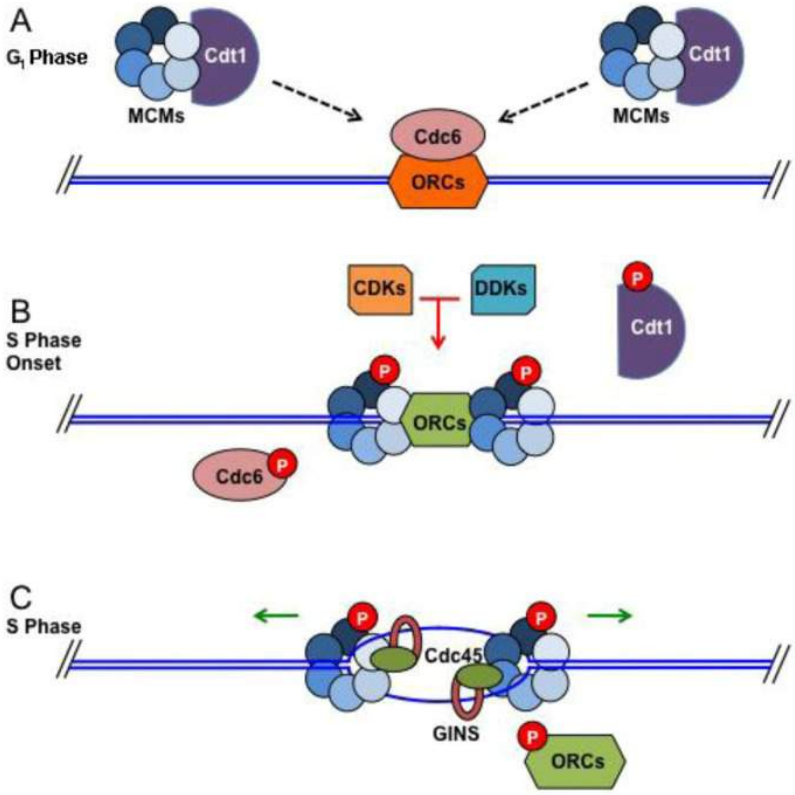


Figure 24.1.35: MCM2-7 loads onto DNA at replication origins during G1 and unwinds DNA ahead of replicative polymerases. Figure from: *Lemanm A.R. and Noguchi, E. (2013) Genes 4(1):1-32*

Pane (**A**) shows the combined activities of Cdc6 and Cdt1 bring MCM complexes (shown as blue circles of varying shades) to replication origins.

Panel (**B**) shows CDK/DDK-dependent phosphorylation of pre-RC components leads to replisome assembly and origin firing. Cdc6 and Cdt1 are no longer required and are removed from the nucleus or degraded

Panel (**C**) shows MCMs and associated proteins (GINS and Cdc45 are shown) unwinding DNA to expose template DNA. At this point, replisome assembly can be completed and replication initiated. "P" indicates phosphorylation.

24.1.7.5: Replication through Nucleosomes

Eukaryotic genomes are substantially more complicated than the smaller and unadorned prokaryotic genomes. Eukaryotic cells have multiple noncontiguous chromosomes, each of which must be compacted to allow packaging within the confined space of a nucleus. As seen in chapter 4, chromosomes are packaged by wrapping ~147 nucleotides (at intervals averaging 200 nucleotides) around an

octamer of histone proteins, forming the ***nucleosome***. The histone octamer includes two copies each of histone H2A, H2B, H3, and H4. In chapter 8, it was highlighted that histone proteins are subject to a variety of post-translational modifications, including phosphorylation, acetylation, methylation, and ubiquitination that represent vital epigenetic marks. The tight association of histone proteins with DNA in nucleosomes suggests that eukaryotic cells possess proteins that are designed to remodel histones ahead of the replication fork, in order to allow smooth progression of the replisome. It is also essential to reassemble histones behind the fork to reestablish the nucleosome conformation. Furthermore, it is important to transmit the epigenetic information found on the parental nucleosomes to the daughter nucleosomes, in order to preserve the same chromatin state. In other words, the same histone modifications should be present on the daughter nucleosomes as on the parental nucleosomes. This must all be done while doubling the amount of chromatin, which requires the incorporation of newly synthesized histone proteins. This process is accomplished by ***histone chaperones*** and ***histone remodelers***, which are discussed below and shown in Figure 24.1.36

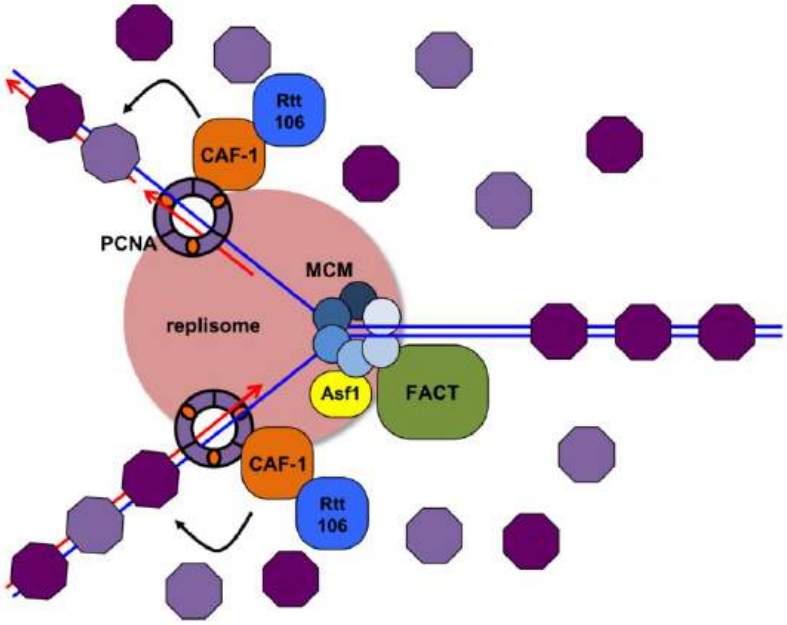


Figure 24.1.36: Nucleosome displacement and deposition during DNA replication. *Lemanm A.R. and Noguchi, E. (2013) Genes 4(1):1-32*

Histones are removed from chromatin ahead of the replication fork. FACT may facilitate this process. Asf1 recruits histone H3-H4 dimers to the replication fork. CAF-1 and Rtt106 load newly synthesized (light purple) histones to establish chromatin behind the fork. Previously loaded histones (dark purple) are also deposited on both daughter DNA strands. The histone chaperones involved in these processes are associated with replisome proteins: CAF-1/Rtt106 with PCNA and FACT/Asf1 with MCMs.

Several histone chaperones are known to be involved in replication-coupled nucleosome assembly, including the ***FACT complex***. The ***FACT complex*** components were originally identified as proteins that greatly stimulate transcription by RNA polymerase II. In budding yeast, FACT was found to interact with DNA Pol α-primase complex, and the FACT subunits were found to interact genetically with replication factors. More recently, studies showed that FACT facilitates DNA replication *in vivo* and is associated with the replisome in budding yeast and human cells. The FACT complex is a heterodimer that does not hydrolyze ATP, but facilitates the “loosening” of histones in nucleosomes

24.1.7.6: Replication Fork Barriers and the Termination of Replication

In prokaryotes, such as the *E. coli*, bidirectional replication initiates at a single replication origin on the circular chromosome and terminates at a site approximately opposed from the origin. This replication terminator region contains DNA sequences known as *Ter* sites, polar replication terminators that are bound by the Tus protein. The *Ter*-Tus complex counteracts helicase activity, resulting in replication termination. In this way, prokaryotic replication forks pause and terminate in a predictable manner during each round of DNA replication.

In eukaryotes, the situation differs. Replication termination typically occurs by the collision of two replication forks anywhere between two active replication origins. The location of the collision can vary based on the replication rate of each of the forks and the timing of origin firing. Often, if a replication fork is stalled or collapsed at a specific site, replication of the site can be rescued when a replisome traveling in the opposite direction completes copying the region. However, there are numerous programmed ***replication***

fork barriers (RFBs) and replication “challenges” throughout the genome. To efficiently terminate or pause replication forks, some fork barriers are bound by RFB proteins in a manner analogous to *E. coli* Tus. In these circumstances, the replisome and the RFB proteins must specifically interact to stop replication fork progression.

24.1.8: Telomeres and Replicative Senescence

24.1.8.1: The End Replication Problem

We have discussed the structure of telomers in the previous section. Let's look now at their activity/function. In humans, ***telomeres*** consist of hundreds to thousands of repetitive sequences of TTAGGG at chromosomal ends for maintaining genomic integrity. Because the DNA replication is asymmetric along double strands, RNA primer sequence at the 3′-hydroxyl end cannot be replaced by DNA polymerase I, as there is no 3'-OH primer group present for the polymerase to extend the DNA chain. This causes the loss of 30–200 nucleotides with each DNA replication and cell division and is known as ***the end replication problem***. ***Telomeres*** provide a repetitive noncoding sequence of DNA at their 3′ en, to prevent the loss of critical genetically encoded information during replication. Moreover, telomeres are coated with a complex of six capping proteins, also known as ***shelterin proteins***, which are packed into a compact ***T-loop structure*** that hides the ends of the chromosomes. This prevents the DNA repair machinery from mistaking chromosomal ends for double-stranded DNA breaks, as shown in Figure 24.1.37. Therefore, telomeres have been proposed as a ***mitotic clock*** that measures how many times a cell has divided and in essence, gives a cell a defined lifetime.

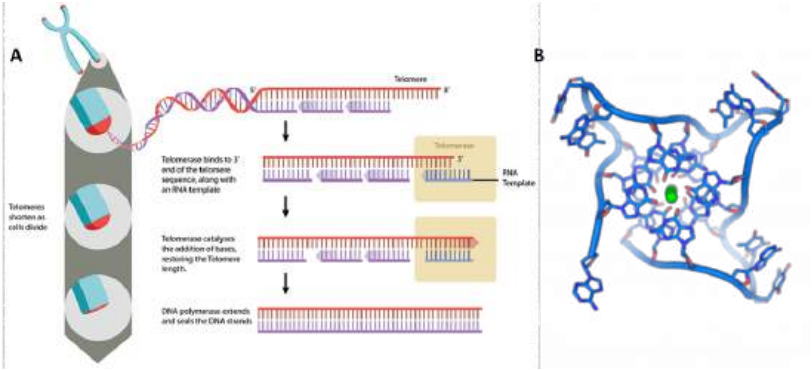


Figure 24.1.37: Telomere Structure. *Image (A) by: MBInfo and Image (B) by:Thomas Splettstoesser*

Pane (A): shows telomeres located at the end of chromosomes, where they help protect against the loss of DNA during replication.

Panel (B) shows DNA quadruplex formed by telomere repeats. The looped conformation of the DNA backbone is very different from the typical DNA helix, this is known as ***T-loop formation***. The green spheres in the center represent potassium ions.

The human ***telomerase enzyme*** is responsible for maintaining and elongating telomeres and consists of an ***RNA component (TERC)*** and a ***reverse transcriptase (TERT)***, that serves as the catalytic component, as shown in Figure 24.1.38 The ***TERT*** uses the ***TERC*** as a template to synthesize new telomeric DNA repeats at a single-stranded overhang to maintain telomere length (Figure 25.1.26). Some cells such as germ cells, stem cells, hematopoietic progenitor cells, activated lymphocytes, and most cancer cells constitutively express telomerase and maintain telomerase activity to overcome telomere shortening and cellular senescence. However, most other somatic cells generally have a low or undetectable level of telomerase activity and concomitantly limited longevity. Interestingly, overall telomerase activity decreases with age, but increases markedly in response to injury, suggesting a role for telomerase in cellular regeneration during wound healing. The telomere length and integrity are regulated through the interplay between the ***telomerase*** and ***shelterin proteins.***

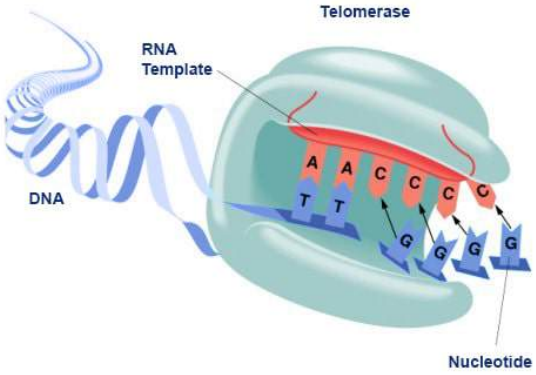


Figure 24.1.37: Conceptual Model of Telomerase Activity. Abbexa Ltd.

The active site of the telomerase enzyme contains the RNA template, TERC (shown in red) and aligns with the last few telomeric bases at the end of the chromosome (shown in blue). This creates a single-stranded overhang that can be used as a template by the TERT reverse transcriptase to extend the telomere sequence.

In vivo, shortened telomeres and damaged telomeres generally caused by reactive oxygen species (ROS) are usually assumed to be the main markers of cellular aging and are thought to be the main cause of ***replicative senescence***. *In vitro*, telomeres lose approximately 50–200 bp at each division due to the end-replication problem. Approximately 100 mitoses are thought to be sufficient to reach the ***Hayflick limit***, or the maximum number of mitotic events allowed prior to entering replicative senescence. Cells in continual renewal, such as blood cells, compensate for telomere erosion by expressing telomerase, the only enzyme able to polymerize telomeric sequences *de novo* at the extremity of telomeres. Knocking out telomerase components, such as the catalytic subunit (TERT) or the RNA template (TERC), induces several features of aging in mice. In humans, germline mutations in telomerase subunits are responsible for progeroïd syndromes, such as Dyskeratosis congenita, a rare genetic form of bone marrow failure. Furthermore, healthy lifespan in humans is positively correlated with longer telomere length and patients suffering from age-related diseases and premature aging have shorter telomeres compared with healthy individuals. An accumulation of unrepaired damage within telomeric regions has also been shown to accumulate in aging mice and non-human primates, suggesting that damage of telomeres with age may also be contributing to age-driven disease states and reduced health span.

Thus, one could argue that the activation and expression of telomerase may be a way of reducing age-related diseases and increasing overall longevity. However, the constitutive expression of telomerase, unfortunately, is a characteristic of almost all cancer cells. It is therefore, no surprise that transgenic animals over-expressing the catalytic subunit of telomerase (mTERT), develop cancers earlier in life. However, over-expression of telomerase in mice that are highly resistant to cancers has shown large increases in median lifespan and significantly reduced age-associated disorders. Since humans are not highly resistant to cancer, this is not a feasible option for humans. However, additional studies in mice, where constitutive expression of telomerase is only introduced into a small percentage of host cells using adenovirus gene therapy techniques has yielded more promising results. Adenoviruses are a group of viruses that form an icosahedral protein capsid that houses a linear double-stranded DNA genome. Infections in humans typically cause symptoms of the common cold and are usually mild in nature. These are a good target for gene therapy, as the DNA that they carry can be mutated, so that they are deficient in their ability to replicate once they have infected the host. They can also be transformed to carry a gene of interest into the host, where that gene can then integrate into the host genome. Experiments in mice that were infected with an adenovirus carrying the mTERT gene showed that mTERT was delivered to a wide range of tissues within the body, and increased telomere length within those tissues. Furthermore, the mTERT expressing mice were healthier than their litter mates and displayed a reduction in disabling conditions associated with physiological aging such as osteoporosis and insulin resistance, as shown in Figure 24.1.38 Cognitive skills and metabolic functions were also improved. Noticeably, mice treated with gene therapy did not have increased incidence in cancer rates, suggesting that in at least for the short-lived mouse species, gene therapy approach to increased telomerase activity is safe. Within these animals, the median lifespan was increased by 24% when animals were treated at 1 year of age, and by 13% if treated at 2 years of age.

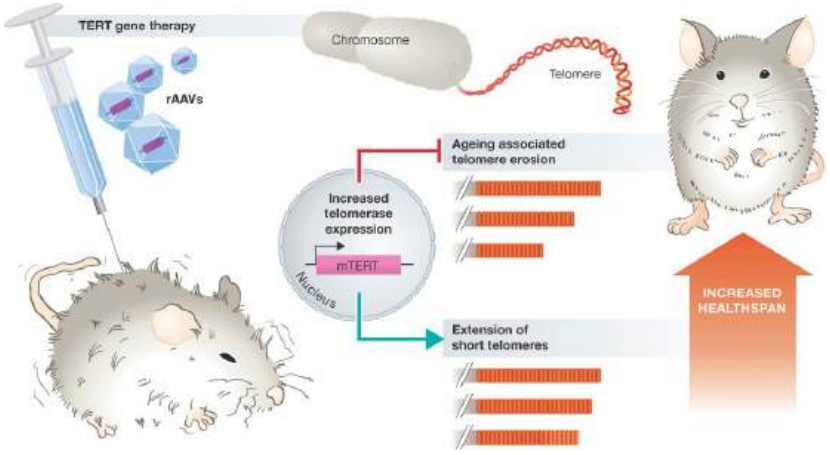


Figure 24.1.38: Promoting healthspan in Mice using a Telomerase Gene Therapy. Delivery of the catalytic subunit of telomerase (TERT) using a modified adenovirus vector (rAAV) suppresses aging-associated telomere erosion and extends short telomeres in a variety of tissues. Consequently, animals display improved healthspans and extended lifespans. Boccardi, V. and Herbig, U. (2012) EMBO Mol Med 4:685-687.

24.1.8.2: Replication and Repair of Telomere Sequences

In addition to the end replication problem, telomeric DNA (telDNA) replication and repair is a real challenge due to the different structural features of telomeres. First, the nucleotide sequence itself consists of a hexanucleotide motif (TTAGGG) repeated over kilobases, with the 5′-3′ strand named the “G-strand” due to its high content in guanine. During the progression of the replication fork, the lagging strand, corresponding to the G-strand, forms G-quadruplex (G4) structures, which have to be resolved to allow fork progression and to complete replication, as shown in Figure 24.1.39 Secondly, R-loops corresponding to highly stable RNA:DNA hybrids, involving the long non-coding telomeric transcript TERRA (telomeric repeat-containing RNA) also have to be dissociated. Thirdly, the extremity of telomeres adopts a specific loop structure, the T-loop, which has to be unraveled. This is the loop that hides the double-stranded end from the DNA damage sensors, and is locked by the hybridization of the 3′ single strand overhang extremity with the above 3′-5′ strand, thereby displacing the corresponding 5′-3′ strand to form a D-loop (displacement loop) structure (Figure 24.1.38). Lastly, replication also has to deal with barriers encountered elsewhere in the genome, such as torsions and a condensed heterochromatic environment.

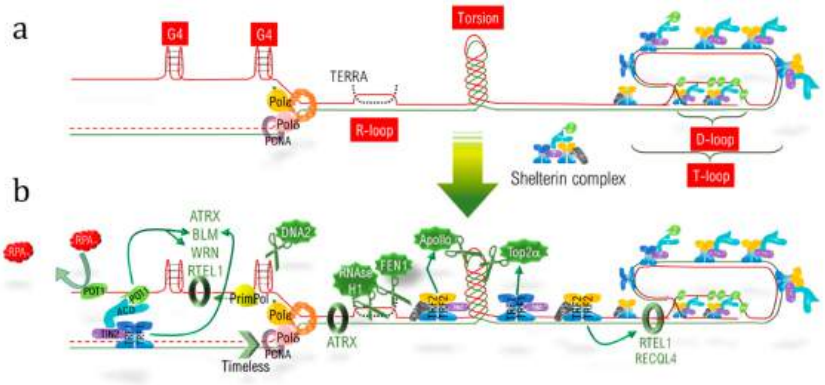


Figure 24.1.39: Obstacles and solutions to replicate telomeres. Billiard, P. and Poncet, D.A. (2019) Int J. Mol. Sci. 20(19) 4959

Panel (a) shows the Telomeric sequence, with the G-strand in a solid red line and the C-strand in a solid green line, is depicted. The terminal D-loop structuring the much larger T-loop is stabilized by the shelterin complex. The replisome (PCNA, pol ε, etc) polymerizes a new G-strand (depicted in a dotted red line) and frees the parental G-strand, enabling the formation of G4 secondary structure. R-loops corresponding to TERRA hybridization (in dotted black lines) with the 3'-5' strand and torsions due to the fork progression are also shown.

Panel (b) shows replication helpers, such as helicases, either helping in G4 unwinding or in D-loop unlocking are depicted. The DNAses (Top2a, DNA2) and RNAses (RNAse H1 and FEN1) help in resolving torsions and RNA:DNA heteroduplexes, while Timeless stimulates the replisome and POT1 competes with RPA1 for binding of the single-strand and helps in G4 dissolution. The shelterin components, POT1, TRF1, and TRF2 help in loading the helper proteins (fine green arrows)

Since telomeres face a host of obstacles to completing the replication process, as discussed in Figure 25.1.28, the cell possesses a set of specialized machinery to fully achieve their replication, such as the RTEL1, TRF1, and TRF2 proteins, DNAses, RNAsses, and Timeless. The recruitment of these factors is orchestrated by the shelterin complex.

At the molecular level, the GGG telomeric repeats are particularly sensitive to ROS, which produce stretches of 8-oxoguanine that are especially difficult to repair. Coupled with inefficient telomere repair, these ROS-induced lesions produce single and double-strand breaks, and/or generate replicative stress, ultimately resulting in telomere shortening. The presence of unrepaired single or tandem 8-oxoguanine drastically inhibits the binding of TRF1 and TRF2, and impairs the recruitment of telomerase, especially when ROS damage is localized in the 3′ overhang. This type of damage contributes to telomere deprotection and shortening. Strikingly, ROS (and other metabolic stresses) also induce the relocation of TERT to mitochondria, as observed (i) in primary neurons after oxidative stress; (ii) in neurons exposed to the tau protein; (iii) in Purkinje neurons subjected to excitotoxicity; and (iv) in cancer cell lines treated with a G4 ligand. Mitochondrial TERT increases the inner membrane potential, as well as the mtDNA copy number, and decreases ROS production with a protective effect on mtDNA. Mitochondria are also critical sensors of cellular damage and contribute to the processes of autophagy and apoptosis (programmed cell death). The relocalization of TERT following chromosomal damage in the nucleus, may indicate one mechanism the mitochondria utilizes to monitor cellular stress and damage.

24.1.9: Replication of Mitochondrial DNA

Mammalian mitochondria contain multiple copies of a circular, double-stranded DNA genome approximately 16.6 kb in length, as shown in Figure 24.1.40 The two strands of mtDNA differ in their base composition, with one being rich in guanines, making it possible to separate a heavy (H) and a light (L) strand by density centrifugation. The mtDNA contains one longer ***noncoding region (NCR)*** also referred to as the control region. In the NCR, there are promoters for polycistronic transcription, one for each mtDNA strand; the light strand promoter (LSP) and the heavy strand promoter (HSP). The NCR also harbors the origin for H-strand DNA replication (O_H). A second origin for L-strand DNA replication (O_L) is located outside the NCR, within a tRNA cluster.

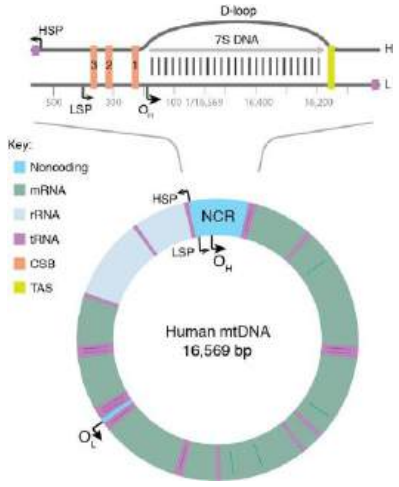


Figure 24.1.40: Map of human mtDNA. Falkenberg, M. (2018) *Essays Biochem* 62(3):287-296

Falkenberg, M. (2018) *Essays Biochem* 62(3):287-296

As shown in Figure 24.1.40 the genome encodes for 13 mRNA (green), 22 tRNA (violet), and 2 rRNA (pale blue) molecules. There is also a major noncoding region (NCR), which is shown enlarged at the top in blue. The major NCR contains the heavy strand promoter (HSP), the light strand promoter (LSP), three conserved sequence boxes (CSB1-3, orange), the H-strand origin of replication (O_H), and the termination-associated sequence (TAS, yellow). The triple-stranded displacement-loop (D-loop) structure is formed by a premature termination of nascent H-strand DNA synthesis at TAS. The short H-strand replication product formed in this manner is termed 7S DNA. A minor NCR, located approximately 11,000 bp downstream of O_H, contains the L-strand origin of replication (O_L).

A dedicated DNA replication machinery is required for its maintenance. Mammalian mtDNA is replicated by proteins distinct from those used for nuclear DNA replication and many are related to replication factors identified in bacteriophages. DNA polymerase γ (POLγ) is the replicative polymerase in mitochondria. In human cells, POLγ is a heterotrimer with one catalytic subunit (POLγA) and two accessory subunits (POLγB). POLγA belongs to the A family of DNA polymerases and contains a 3′–5′ exonuclease domain that acts to proofread the newly synthesized DNA strand. POLγ is a highly accurate DNA polymerase with a frequency of misincorporation lower than 1×10^{-6}. The accessory POLγB subunit enhances interactions with the DNA template and increases both the catalytic activity and the processivity of POLγA. The DNA helicase TWINKLE travels in front of POLγ, unwinding the double-

stranded DNA template. TWINKLE forms a hexamer and requires a fork structure (a single-stranded 5′-DNA loading site and a short 3′-tail) to load and initiate unwinding. Mitochondrial single-stranded DNA-binding protein (mtSSB) binds to the formed ssDNA, protects it against nucleases, and prevents secondary structure formation

The most accepted model of DNA replication in the mitochondria is the ***strand displacement model,*** as shown in Figure 24.1.41 Within this model, DNA synthesis is continuous on both the H- and L-strand. There is a dedicated origin for each strand, O_H and O_L. First, replication is initiated at O_H and DNA synthesis then proceeds to produce a new H-strand. During the initial phase, there is no simultaneous L-strand synthesis and mtSSB covers the displaced, parental H-strand. By binding to single-stranded DNA, mtSSB prevents the mitochondrial RNA polymerase (POLRMT) from initiating random RNA synthesis on the displaced strand. When the replication fork has progressed about two-thirds of the genome, it passes the second origin of replication, O_L. When exposed in its single-stranded conformation, the parental H-strand at O_L folds into a stem–loop structure. The stem efficiently blocks mtSSB from binding and a short stretch of single-stranded DNA in the loop region, therefore remains accessible, allowing POLRMT to initiate RNA synthesis. POLRMT is not processive on single-stranded DNA templates. After adding approximately 25 nucleotides, it is replaced by POLγ and L-strand DNA synthesis is initiated. From this point, H- and L-strand synthesis proceeds continuously until the two strands have reached full circle. Replication of the two strands is linked, since H-strand synthesis is required for the initiation of L-strand synthesis. DNA Ligase III is used to complete the ligation of the newly formed DNA strands.

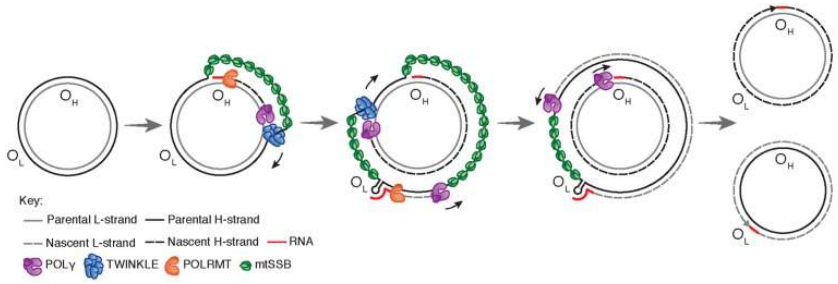


Figure 25.1.24 Replication of the human mitochondrial genome.
Mitochondrial DNA replication is i
nitiated at O
H
and proceeds unidirectionally to produce the full-length nascent H-strand.
mtSSB
binds and protects the exposed, parental H-strand. When the replisome passes O
L
, a stem–loop structure is formed that blocks mtSSB binding, presenting a single-stranded loop-region from which POLRMT can initiate primer synthesis. The transition to L-strand DNA synthesis takes place after about 25 nt, when POLγ replaces POLRMT at the 3′-end of the primer. Synthesis of the two strands proceeds in a continuous manner until two full, double-stranded DNA molecules have been formed.

Figure from:
Falkenberg, M. (2018) *Essays Biochem* 62(3):287-296

During DNA replication, the parental molecule remains intact, which poses a steric problem for the moving replication machinery. Topoisomerases belonging to the type 1 family can relieve torsional strain formed in this way, by allowing one of the strands to pass through the other. In mammalian mitochondria, TOP1MT a type IB enzyme can act as a DNA “swivel”, working together with the mitochondrial replisome. Furthermore, replication of circular DNA often causes the formation of catenanes, or interlocked circles that need to be separated from one another. The type 1A topoisomerase, topoisomerase 3α (Top3α), is required to resolve the hemicatenane structure that can form during mtDNA replication.

Curiously, not all replication events initiated at O_H continue to full circle. Instead, 95% are terminated after about the first 650 nucleotides at a sequence known as the ***termination associated sequences (TAS)*** (Figure 25.1.23). This creates a short DNA fragment known as the 7S DNA, that remains bound to the parental L-strand, while the parental H-strand is displaced (Figure 25.1.23). As a result, a triple-stranded ***displacement loop structure, a D-loop,*** is formed. The functional importance of the D-loop structure is unclear and how replication is terminated at TAS is also not known.

24.1.9.0.1: Mastering the Content

Which of the following is the enzyme that replaces the RNA nucleotides in a primer with DNA nucleotides?

1. DNA polymerase III
2. DNA polymerase I
3. primase

4. helicase

[reveal-answer q="628075"]Show Answer[/reveal-answer]
[hidden-answer a="628075"]Answer b. DNA polymerase I is the enzyme that replaces the RNA nucleotides in a primer with DNA nucleotides.[/hidden-answer]

Which of the following is not involved in the initiation of replication?

1. ligase
2. DNA gyrase
3. single-stranded binding protein
4. primase

[reveal-answer q="820951"]Show Answer[/reveal-answer]
[hidden-answer a="820951"]Answer a. Ligase is not involved in the initiation of replication.[/hidden-answer]

Which of the following enzymes involved in DNA replication is unique to eukaryotes?

1. helicase
2. DNA polymerase
3. ligase
4. telomerase

[reveal-answer q="650146"]Show Answer[/reveal-answer]
[hidden-answer a="650146"]Answer d. Telomerase is unique to eukaryotes.[/hidden-answer]

Which of the following would be synthesized using 5′-CAGTTCGGA-3′ as a template?

1. 3′-AGGCTTGAC-4′
2. 3′-TCCGAACTG-5′
3. 3′-GTCAAGCCT-5′
4. 3′-CAGTTCGGA-5′

[reveal-answer q="429167"]Show Answer[/reveal-answer]
[hidden-answer a="429167"]Answer c. 3′-GTCAAGCCT-5′[/hidden-answer]

The enzyme responsible for relaxing supercoiled DNA to allow for the initiation of replication is called _____.
[reveal-answer q="855893"]Show Answer[/reveal-answer]
[hidden-answer a="855893"]The enzyme responsible for relaxing supercoiled DNA to allow for the initiation of replication is called **DNA gyrase or topoisomerase II**.[/hidden-answer]

Unidirectional replication of a circular DNA molecule like a plasmid that involves nicking one DNA strand and displacing it while synthesizing a new strand is called _____.
[reveal-answer q="378861"]Show Answer[/reveal-answer]
[hidden-answer a="378861"]Unidirectional replication of a circular DNA molecule like a plasmid that involves nicking one DNA strand and displacing it while synthesizing a new strand is called**rolling circle replication**.[/hidden-answer]

More primers are used in lagging strand synthesis than in leading strand synthesis.
[reveal-answer q="25479"]Show Answer[/reveal-answer]
[hidden-answer a="25479"]True[/hidden-answer]

1. Why is primase required for DNA replication?
2. What is the role of single-stranded binding protein in DNA replication?
3. Below is a DNA sequence. Envision that this is a section of a DNA molecule that has separated in preparation for replication, so you are only seeing one DNA strand. Construct the complementary DNA sequence (indicating 5′ and 3′ ends).DNA sequence: 3′-T A C T G A C T G A C G A T C-5′
4. Review Figure 1 and Figure 2. Why was it important that Meselson and Stahl continue their experiment to at least two rounds of replication after isotopic labeling of the starting DNA with ^{15}N, instead of stopping the experiment after only one round of replication?
5. If deoxyribonucleotides that lack the 3′-OH groups are added during the replication process, what do you expect will occur?

24.1.9.1: 25.1.7 References

1. Parker, N., Schneegurt, M., Thi Tu, A-H., Lister, P., Forster, B.M. (2019) Microbiology. *Openstax*. Available at: https://opentextbc.ca/microbiologyopenstax/
2. Principles of Biochemistry/Cell Metabolism I: DNA replication. (2017, August 6). *Wikibooks, The Free Textbook Project*. Retrieved 19:07, October 31, 2019 from en.wikibooks.org/w/index.php?title=Principles_of_Biochemistry/Cell_Metabolism_I:_DNA_replication&oldid=3259729.
3. Kaiser, G.E. (2015) Prokaryotic Cell Anatomy. *Community College of Baltimore County.* Available at: http://faculty.ccbcmd.edu/~gkaiser/SoftChalk%20BIOL%20230/Prokaryotic%20Cell%20Anatomy/nucleoid/nucleoid/nucleoid3.html
4. The RCSB PDB "Molecule of the Month": Inspiring a Molecular View of Biology D.S. Goodsell, S. Dutta, C. Zardecki, M. Voigt, H.M. Berman, S.K. Burley (2015) *PLoS Biol* **13**(5): e1002140. doi: 10.1371/journal.pbio.1002140
5. Wikipedia contributors. (2020, May 7). Helicase. In *Wikipedia, The Free Encyclopedia*. Retrieved 13:38, June 9, 2020, from en.Wikipedia.org/w/index.php?title=Helicase&oldid=955303097
6. Windgassen, T.A., Wessel, S.R., Bhattacharyya, B., and Keck, J.L. (2017) Mechanisms of bacterial DNA replication restart. *Nuc Acids Res* 46(2):504-519. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5778457/
7. Xu, Z-Q., Dixon, N.E. (2018) Bacterial Replisomes. *Curr Opin Struct Biol* 53:159-168. Available at: https://www.sciencedirect.com/science/article/pii/S0959440X18300952
8. Liu, B., Eliason, W.K., and Steitz, T.A. (2013) Structure of a helicase-helicase loader complex reveals insights into the mechanism of bacterial primosome assembly. Nature Comm 4:2495. Available at: https://www.researchgate.net/publication/256764134_Structure_of_a_helicase-helicase_loader_complex_reveals_insights_into_the_mechanism_of_bacterial_primosome_assembly
9. Xu, Z-Q., and Dixon, N.E. (2018) Bacterial Replisomes. *Curr Op Struc Biol* 53:159-168. Available at: https://www.sciencedirect.com/science/article/pii/S0959440X18300952
10. Fernandez-Leiro, R., Conrad, J., Scheres, S.HW., and Lamers, M.H. (2015) cryo-EM structures of the E. coli replicative DNA polymerase reveal its dynamic interactions with the DNA sliding clamp, exonuclease and τ. *eLife* 4:e11134. Available at: https://elifesciences.org/articles/11134
11. Ekundayo, B. and Bleichert, F. (2019) Origins of DNA replication. *PLOS* 15(12): e1008556. Available at: https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1008320
12. Leman, A.R., and Noguchi, E. (2013) The replication fork: understanding the eukaryotic replication machinery and the challenges to genome duplication. *Genes* 4(1): 1-32. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3627427/
13. Doublié, S. and Zahn, K.E. (2014) Structural insights into eukaryotic DNA replication. *Front. Microbiol.* 5:444. Available at: https://www.frontiersin.org/articles/10.3389/fmicb.2014.00444/full
14. Billard, P., and Poncet D.A. (2019) Replication stress at telomeric and mitochondrial DNA: Common origins and consequences on ageing. *Int J. Mol Sci* 20(19):4959. Available at: https://www.mdpi.com/1422-0067/20/19/4959/htm
15. Yeh, J-K., and Wang, C-Y. (2016) Telomeres and telomerase in cardiovascular diseases. *Genes* 7(9)58. Available at: https://www.mdpi.com/2073-4425/7/9/58/htm
16. Boccardi, V. and Herbig, U. (2012) Telomerse gene therapy: a novel approach to combat aging. *EMBO Mol Med* 4:685-687. Available at: https://www.embopress.org/doi/epdf/10.1002/emmm.201200246
17. Wikipedia contributors. (2020, April 26). Cyclin-dependent kinase. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:52, June 30, 2020, from https://en.Wikipedia.org/w/index.php?title=Cyclin-dependent_kinase&oldid=953307433
18. Falkenberg, M. (2018) Mitochondrial DNA replication in mammalian cells: overview of the pathway. Essays Biochem 62(3):287-296. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6056714/
19. Folwer, S., et. al. (2013) Concepts of Biology. *Openstax.* Available at: https://openstax.org/details/books/concepts-biology?Book%20details
20. Aleem, E. and Arceci, R.J. (2015) Targeting cell cycle regulators in hematologic malignancies. Frontiers in Cell and Developmental Biology 3(16). Available at: https://www.researchgate.net/publication/275354547_Targeting_cell_cycle_regulators_in_hematologic_malignancies

24.2: DNA Mutations, Damage, and Repair

The integrity of the DNA structure for cell viability is underscored by the vast amounts of cellular machinery dedicated to ensuring its accurate replication, repair, and storage. Even still, mutations within the DNA are a fairly common event.

24.2.1: DNA Mutations

Mutations are random changes that occur within the sequence of bases in DNA. They can be large-scale, altering the structure of the chromosomes, or small scale where they only alter a few or even a single base or nucleotide. Mutations can occur for many reasons. For example, DNA mutations can be caused by mistakes made by the DNA polymerase during replication. As noted in chapter 9, DNA polymerases are highly processive enzymes that contain proofreading and editing functions. With these safeguards, their error rates are typically very low and range from one in a million bases to one in a billion bases. Even with such high fidelity, this error rate will lead to between 3 and 3,000 errors within the human genome for each cell undergoing DNA replication. DNA mutations can also result through the replication of DNA that has been damaged by endogenous or exogenous agents. The next section will highlight common types of DNA damage and their effects. If a DNA polymerase encounters a damaged DNA base in the template DNA during replication it may place a random nucleotide base across from the lesion. For example, an adenine-containing nucleotide will often be added across a lesion, regardless of what the correct match should be. This can lead to the formation of ***transition*** or ***transversion*** mutations.

A ***transition*** mutation is a point mutation that changes a purine nucleotide to another purine (A ↔ G) or a pyrimidine nucleotide to another pyrimidine (C ↔ T). ***Transversion*** refers to the substitution of a purine for a pyrimidine or vice versa. Sometimes lesions may cause bases to be skipped during replication or cause extra nucleotides to be inserted into the backbone. DNA polymerases can also slip during the replication of regions of the DNA that have repeated sequences or large stretches repeating a single base. Larger lesions or cross-links in the DNA during replication can lead to more catastrophic DNA damage including DNA strand breaks. Mutations may also occur during the processes of mitosis and meiosis when sister chromatids and/or homologous chromosomes are separated from one another.

In nature, ***mutagenesis,*** or the process of generating DNA mutations, can lead to changes that are harmful, beneficial, or have no effect. Harmful mutations can lead to cancer and various heritable diseases, but beneficial mutations are the driving force of evolution. In 1927, Hermann Muller first demonstrated the effects of mutations with observable changes in chromosomes. He induced mutagenesis by irradiating fruit flies with X-rays.

When a mutation is caused by an environmental factor or a chemical agent, that agent is called a ***mutagen.*** Typical mutagens include chemicals, like those inhaled while smoking, and radiation, such as X-rays, ultraviolet light, and nuclear radiation. Different mutagens have different modes of damaging DNA and are discussed further in the next section. It is important to note that DNA damage, in and of itself, does not necessarily lead to the formation of a mutation in the DNA. There are elaborate DNA repair processes designed to recognize and repair different types of DNA lesions. Fewer than 1 in 1,000 DNA lesions will result in a DNA mutation. The processes of DNA damage recognition and repair are the focus of later sections within this chapter.

24.2.1.1: Types of Mutations

There are a variety of types of mutations. Two major categories of mutations are germline mutations and somatic mutations.

- ***Germline mutations*** occur in gametes, the sex cells, such as eggs and sperm. These mutations are especially significant because they can be transmitted to offspring and every cell in the offspring will have the mutations.
- ***Somatic mutations*** occur in other cells of the body. These mutations may have little effect on the organism because they are confined to just one cell and its daughter cells. Somatic mutations also cannot be passed on to offspring.

Mutations also differ in the way that the genetic material is changed. Mutations may change an entire chromosome or just one or a few nucleotides.

Chromosomal alterations are mutations that change chromosome structure or number. They occur when a section of a chromosome breaks off and rejoins incorrectly or does not rejoin at all. Possible ways these mutations can occur are illustrated in the figure below. Chromosomal alterations are very serious. They often result in the death of the cell or organism in which they occur. If the organism survives, it may be affected in multiple ways. An example of a human chromosomal alteration is the mutation that causes Down Syndrome. It is a duplication mutation that leads to developmental delays and other abnormalities. It occurs when the individual inherits an extra copy of chromosome 21. It is also called trisomy ("three-chromosome") 21. Thus, large-scale mutations in the chromosomal structure include ***(1) Amplifications (including gene duplications)*** where repetition of a

chromosomal segment or presence of an extra piece of a chromosome broken piece of a chromosome may become attached to a homologous or non-homologous chromosome so that some of the genes are present in more than two doses leading to multiple copies of all chromosomal regions, increasing the dosage of the genes located within them, ***(2) Deletions*** of large chromosomal regions, leading to loss of the genes within those regions, and ***(3) Chromosomal Rearrangements*** such as ***translocations*** (which interchange of genetic parts from nonhomologous chromosomes), ***insertions*** (which insert segments of one chromosome into another nonhomologous chromosome), and ***inversions*** (which invert or flip a section of a chromosome into the opposite orientation), as shown in Figure 24.2.1.

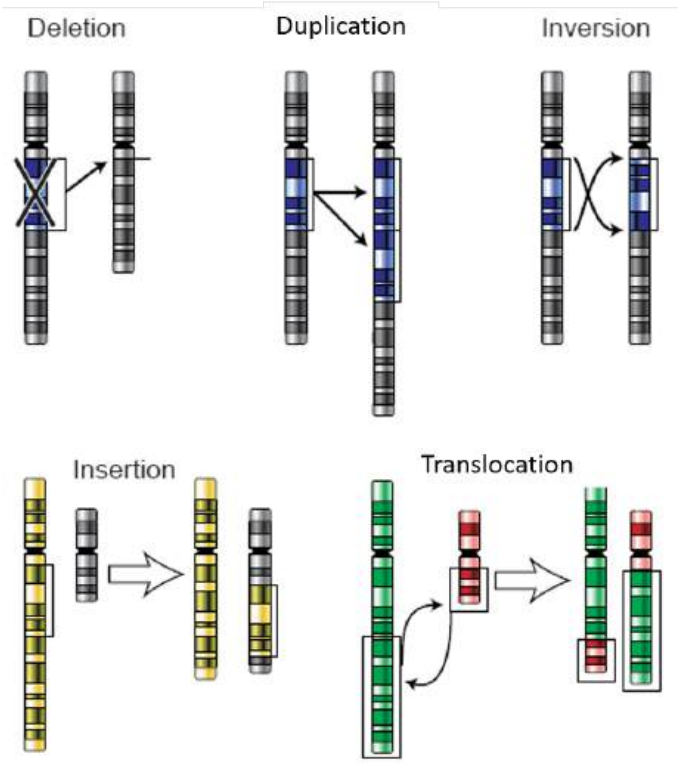


Figure 24.2.1: Chromosomal Alterations. Chromosomal alterations are major changes in the genetic material.*Image modified from Dietzel65*

There are also smaller mutations that can occur that only alter a single nucleotide or a small number of nucleotides within a localized region of the DNA. These are classified according to how the DNA molecule is altered. One type, a ***point mutation***, affects a single base and most commonly occurs when one base is substituted or replaced by another. Mutations also result from the addition of one or more bases, known as an ***insertion***, or the removal of one or more bases, known as a ***deletion***.

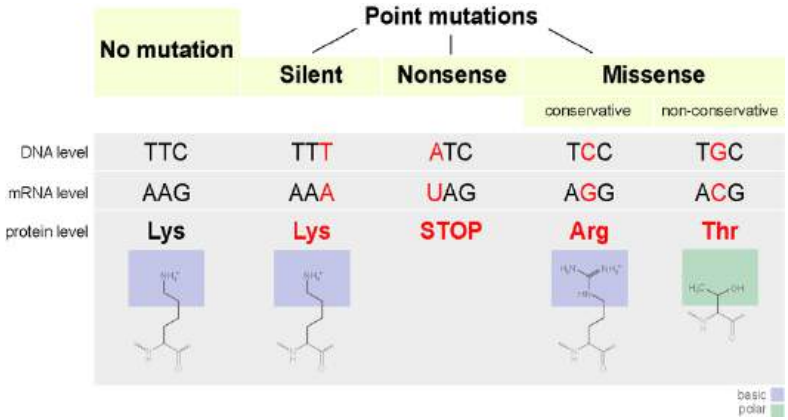


Figure 24.2.2: **The Potential Effects of Point Mutations on Protein Coding Regions.** The image shows various types of point mutations (silent, missense, and nonsense), which may lead to changes in the protein structure. Figure from: Jonsta247

Point mutations (Table 24.2.1 and Figure 24.2.2) may have a wide range of effects on protein function. As a consequence of the degeneracy of the genetic code, a ***point mutation*** will commonly result in the same amino acid being incorporated into the resulting

polypeptide despite the sequence change. This change would not affect the protein's structure and is thus called a ***silent mutation***. A ***missense mutation*** results in a different amino acid being incorporated into the resulting polypeptide. The effect of a missense mutation depends on how chemically different the new amino acid is from the wild-type amino acid. The location of the changed amino acid within the protein also is important. For example, if the changed amino acid is part of the enzyme's active site or greatly affects the shape of the enzyme, then the effect of the missense mutation may be significant. Many missense mutations result in proteins that are still functional, at least to some degree. Sometimes the effects of missense mutations may be only apparent under certain environmental conditions; such missense mutations are called ***conditional mutations***. Rarely, a missense mutation may be beneficial. Under the right environmental conditions, this type of mutation may give the organism that harbors it a selective advantage. Yet another type of point mutation called a ***nonsense mutation***, converts a codon encoding an amino acid (a sense codon) into a stop codon (a nonsense codon). Nonsense mutations result in the synthesis of proteins that are shorter than the wild type and typically not functional.

Table 24.2.1: Types of Point Mutations

Type	Description	Example	Effect
Silent	mutated codon codes for the same amino acid	CAA (glutamine) → CAG (glutamine)	none
Missense	mutated codon codes for a different amino acid	CAA (glutamine) → CCA (proline)	variable
Nonsense	a mutated codon is a premature stop codon	CAA (glutamine) → UAA (stop) usually	serious

Smaller scale ***deletions*** and ***insertions*** also cause various effects. Because codons are triplets of nucleotides, insertions or deletions in groups of three nucleotides may lead to the insertion or deletion of one or more amino acids and may not cause significant effects on the resulting protein's functionality. However, ***frameshift mutations***, caused by insertions or deletions of a number of nucleotides that are not a multiple of three are extremely problematic because a shift in the reading frame results, as shown in Figure 24.2.3. Because ribosomes read the mRNA in triplet codons, frameshift mutations can change every amino acid after the point of the mutation. The new reading frame may also include a stop codon before the end of the coding sequence. Consequently, proteins made from genes containing frameshift mutations are nearly always nonfunctional.

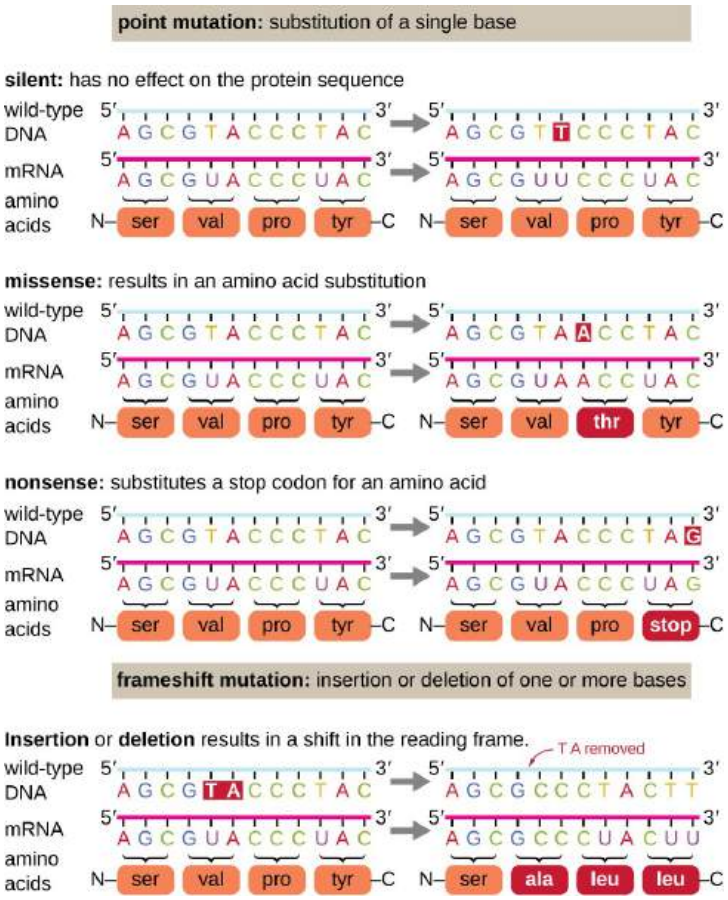


Figure 24.2.3: Summary of Small Scale Mutations on Protein Coding Regions. DNA alterations that lead to changes in the protein sequence encoded by the DNA can include point mutations, DNA insertions, and DNA deletions. Figure from: Parker, et al (2019) Microbiology from Openstax

The majority of mutations have neither negative nor positive effects on the organism in which they occur. These mutations are called neutral mutations. Examples include silent point mutations, which are neutral because they do not change the amino acids in the proteins they encode.

Some mutations have a positive effect on the organism in which they occur. They are referred to as beneficial mutations. If they occur in germline cells (eggs or sperm) these traits can be heritable and passed from one generation to the next. Beneficial mutations generally code for new versions of proteins that help organisms adapt to their environment. If they increase an organism's chances of surviving or reproducing, the mutations are likely to become more common within a population over time. There are several well-known examples of beneficial mutations. Here are just two:

1. Mutations have occurred in bacteria that allow the bacteria to survive in the presence of antibiotic drugs. The mutations have led to the evolution of antibiotic-resistant strains of bacteria.
2. A unique mutation is found in people in a small town in Italy. The mutation protects them from developing atherosclerosis, which is the dangerous buildup of fatty materials in blood vessels. The individual in which the mutation first appeared has even been identified.

Harmful mutations can also occur. Imagine making a random change in a complicated machine such as a car engine. The chance that the random change would improve the functioning of the car is very small. The change is far more likely to result in a car that does not run well or perhaps does not run at all. By the same token, any random change in a gene's DNA is more likely to result in the production of a protein that does not function normally or may not function at all, than in a mutation that improves the function. Such mutations are likely to be harmful. Harmful mutations may cause genetic disorders or cancer.

- A genetic disorder is a disease, syndrome, or other abnormal condition caused by a mutation in one or more genes or by a chromosomal alteration. An example of a genetic disorder is cystic fibrosis. A mutation in a single gene causes the body to produce thick, sticky mucus that clogs the lungs and blocks ducts in digestive organs. Genetic disorders are usually caused by gene mutations that occur within germline cells and are heritable.

- Illnesses caused by mutations that occur within an individual, but are not passed on to their offspring, are mutations that occur in somatic cells. Cancer is a disease caused by an accumulation of mutations within somatic cells. It results in cells that grow out of control and form abnormal masses of cells called tumors. It is generally caused by mutations in genes that regulate the cell cycle, DNA repair, angiogenesis, and other genes that favor cell growth and survival. Because of the mutations, cells with the mutated DNA have evolved to divide without restrictions, hide from the immune system, and develop drug resistance.

24.2.2: Types of DNA Damage

DNA damage, due to environmental factors and normal metabolic processes inside the cell, occurs at a rate of 1,000 to 1,000,000 molecular lesions per cell per day. While this constitutes only 0.000165% of the human genome's approximately 6 billion bases (3 billion base pairs), if left unrepaired can cause mutations in critical genes (such as tumor suppressor genes) can impede a cell's ability to carry out their function and appreciably increase the likelihood of tumor formation and disease states such as cancer.

The vast majority of DNA damage affects the primary structure of the double helix; that is, the bases themselves are chemically modified. These modifications can, in turn, disrupt the molecules' regular helical structure by introducing non-native chemical bonds or bulky adducts that do not fit in the standard double helix. Unlike proteins and RNA, DNA usually lacks tertiary structure, and therefore damage or disturbance does not occur at that level. DNA is, however, supercoiled and wound around "packaging" proteins called histones (in eukaryotes), and both superstructures are vulnerable to the effects of DNA damage.

Several types of DNA damage can occur due either to normal cellular processes or due to the environmental exposure of cells to DNA-damaging agents. DNA bases can be damaged by: (1) oxidative processes, (2) alkylation of bases, (3) base loss caused by the hydrolysis of bases, (4) bulky adduct formation, (5) DNA crosslinking, and (6) DNA strand breaks, including single and double-stranded breaks. An overview of these types of damage is described below.

24.2.2.1: Oxidative Damage

Reactive oxygen species (ROS) can cause significant cellular stress and damage including oxidative DNA damage. Hydroxyl radicals ($^{\bullet}OH$) are one of the most reactive and electrophilic of the ROS and can be produced by ultraviolet and ionizing radiations or from other radicals arising from enzymatic reactions. The $^{\bullet}OH$ can cause the formation of 8-oxo-7,8-dihydroguanine (8-oxoG) from guanine residues, among other oxidative products, as shown in Figure 24.2.4 Guanine is the most easily oxidized of the nucleic acid bases because it has the lowest ionization potential among the DNA bases. The 8-oxo-dG is one of the most abundant DNA lesions, and it is considered as a biomarker of oxidative stress. It has been estimated that up to 100,000 8-oxo-dG lesions can occur daily in DNA per cell. The reduction potential of 8-oxo-dG is even lower (0.74 V vs. NHE) than that of guanosine (1.29 V vs NHE). Therefore, it can be further oxidized creating a variety of secondary oxidation products.

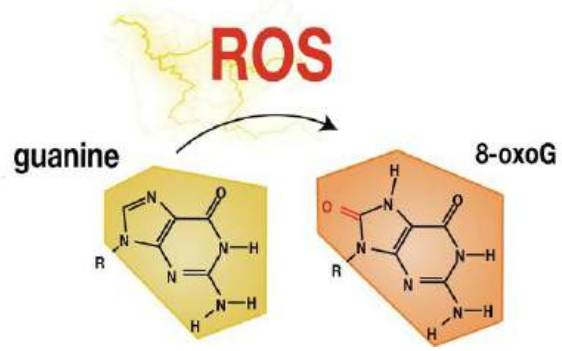


Figure 24.2.4: Reactive Oxygen Species and DNA Damage. Under conditions of oxidative stress, 8-oxoG is the result of reactive oxygen species (ROS) modifying guanine. Figure from: Poetsch, A.R. (2020) Comp & Struct Biotech J. 18:207-219.

As mentioned previously, increased levels of 8-oxo-dG in tissue can serve as a biomarker of oxidative stress. Furthermore, increased levels of 8-oxo-dG are frequently found associated with carcinogenesis and other disease states, as shown in Figure 24.2.5. During the replication of DNA that contains 8-oxo-dG, adenine is most often incorporated across from the lesion. Following replication, the 8-oxo-dG is excised during the repair process and thymine is incorporated in its place. Thus, 8-oxo-dG mutations typically result in a G to T transversion.

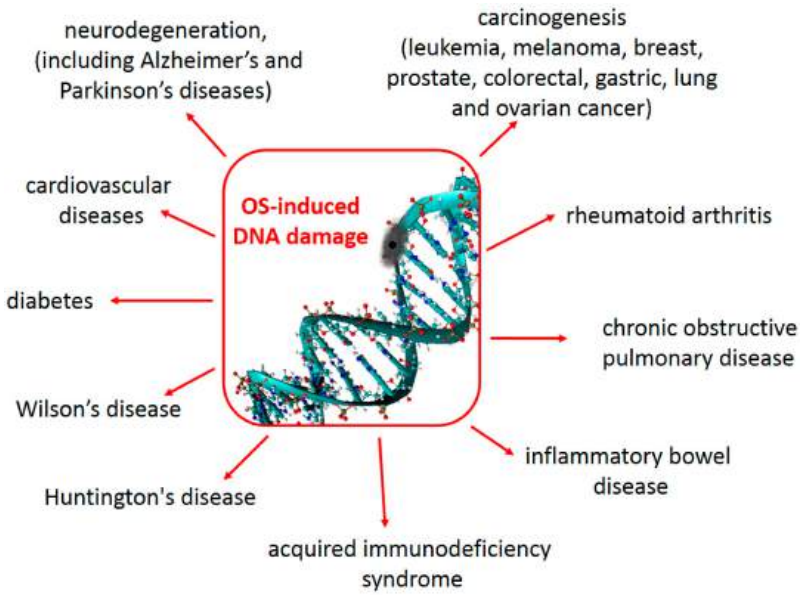


Figure 24.2.5: Oxidative Stress and Human Health. DNA damage caused by oxidative stress has been associated with many human disease states. Figure from: Galano, A., Tan, D-X., and Reiter, R.J. (2018) Molecules 23(3)530

24.2.2.2: Alkylation of Bases

Alkylating agents are widespread in the environment and are also produced endogenously, as by-products of cellular metabolism. They introduce lesions into DNA or RNA bases that can be cytotoxic, mutagenic, or neutral to the cell. Figure 24.2.6 depicts the major reactive sites on the DNA bases that are susceptible to alkylation. Cytotoxic lesions block replication, interrupt transcription, or signal the activation of apoptosis, whereas mutagenic ones are miscoding and cause mutations in newly synthesized DNA. The most common type of alkylation is methylation with the major products including N7-methylguanine (7meG), N3-methyladenine (3meA), and O6-methylguanine (O6meG). Smaller amounts of methylation also occurs on other DNA bases, and include the formation of N1-methyladenine (1meA), N3-methylcytosine (3meC), O4-methylthymine (O4meT), and methyl phosphotriesters (MPT).

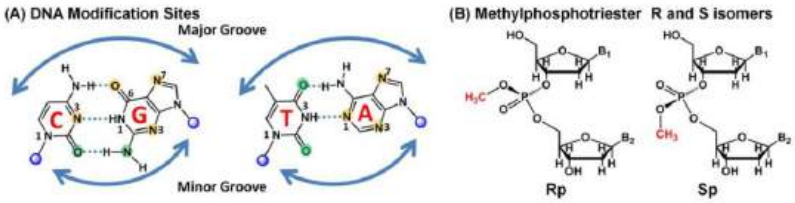


Figure 24.2.6: Major Sites of Alkylation in DNA. The diagram shows DNA Watson-Crick base pairs with principal damage sites modified by small alkylating agents (methyl and ethyl). (A) Sites of base modification by SN1 and SN2 alkylating agents. The orange indicates major damage sites and the green minor damage sites. (B) Phosphotriester formation is indicated by the presence of methyl groups, along with the Rp and Sp isomers. Figure from: Ahmad, A., Nay, S.L. and O'Conner, T.R. (2015) Chapter 4 Direct Reversal Repair in Mammalian Cells

Alkylating agents can cause damage to all exocyclic nitrogens and oxygens in DNA and RNA, as well as at ring nitrogens (Figure 25.2.6A). However, the percentage of each base site modified depends on the alkylating agent, the position in DNA or RNA, and whether nucleic acids are single- or double-stranded. Interestingly, O-alkylations are more mutagenic and harmful than N-alkylations, which may be more cytotoxic, but not as mutagenic.

As we will explore in Chapter 13, methylation of DNA also serves as an important mechanism regulating gene expression.

24.2.2.3: Base Loss

An ***AP site (apurinic/apyrimidinic site)***, also known as an ***abasic site***, is a location in DNA (also in RNA but much less likely) that has neither a purine nor a pyrimidine base, either spontaneously or due to DNA damage, as shown in Figure 24.2.7. It has been estimated that under physiological conditions 10,000 apurinic sites and 500 apyrimidinic may be generated in a cell daily.

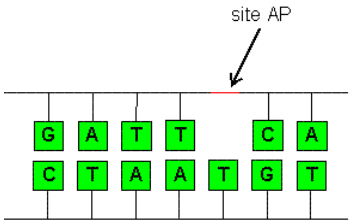


Figure 24.2.7: Abasic Sites. Apurinic and Apyrimidimic (AP) sites occur due to unstable hydrolysis. Figure from Mayhew

AP sites can be formed by spontaneous depurination, but also occur as intermediates in base excision repair, the repair process described in section 25.2.5. If left unrepaired, AP sites can lead to mutation during semiconservative replication. They can cause replication fork stalling and are often bypassed by ***translesion synthesis***, which is discussed in greater detail in section 12.8. In *E. coli*, adenine is preferentially inserted across from AP sites, known as the "A rule". The situation is more complex in higher eukaryotes, with different nucleotides showing a preference depending on the organism and environmental conditions.

24.2.2.4: Bulky Adduct Formation

Some chemicals are biologically reactive and will form covalent linkages with biological molecules such as DNA and proteins creating large ***bulky adducts***, or appendages, that branch off from the main molecule. We will use the mutagen/carcinogen, **benzo[*a*]pyrene,** as an example for this process.

Benzo[*a*]pyrene is a polycyclic aromatic hydrocarbon that forms during the incomplete combustion of organic matter at temperatures between 300°C (572°F) and 600°C (1,112°F). The ubiquitous compound can be found in coal tar, tobacco smoke, and many foods, especially grilled meats. Benzo[*a*]pyrene is a ***procarcinogen*** that needs to be biologically activated by metabolism before it forms a reactive metabolite, as in Figure 24.2.8 Normally, when the body is exposed to foreign molecules, it will start a metabolic process that makes the molecule more hydrophilic and easier to remove as a waste product. Unfortunately, in the case of benzo[*a*]pyrene, the resulting metabolite is a highly reactive epoxide that forms a bulky adduct preferentially with guanine residues in DNA. If left unrepaired, during DNA replication an adenine will usually be placed across from the lesion in the daughter molecule. Subsequent repair of the adduct will result in the replacement of the damaged guanine base with thymine, causing a G --> T transversion mutation.

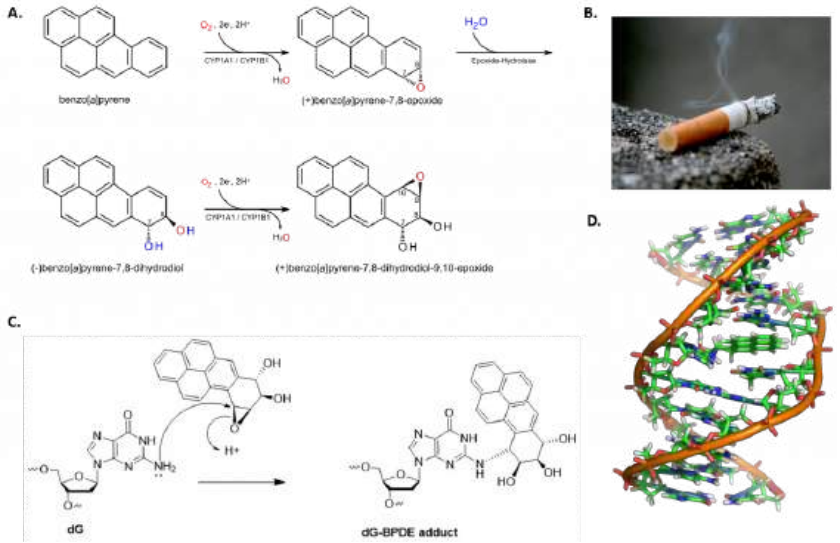


Figure 24.2.8: Benzo[*a*]pyrene Activation and Adduct Formation. (A) Benzo[*a*]pyrene is biologically activated to a dihydrodial epoxide during normal metabolic processes. (B) Cigarette smoke is a source of benzo[*a*]pyrene. (C) The activated (+)benzo[*a*]pyrene-7,8-dihydrodiol-9,10-epoxide can form a DNA adduct with guanine residues. (D) The benzo[*a*]pyrene adduct forms a bulky lesion that distorts the DNA backbone. Figures from: (A) Eleska, (B) Elizabeth Aguilera (C) DI93biochemist, and (D) Zephryis

24.2.2.5: DNA Crosslinking

Crosslinking of DNA occurs when various exogenous or endogenous agents react with two nucleotides of DNA, forming a covalent linkage between them. This crosslink can occur within the same strand ***(intrastrand)*** or between opposite strands of

double-stranded DNA ***(interstrand)***, as shown in Figure 24.2.9. These adducts interfere with cellular metabolism, such as DNA replication and transcription, triggering cell death.

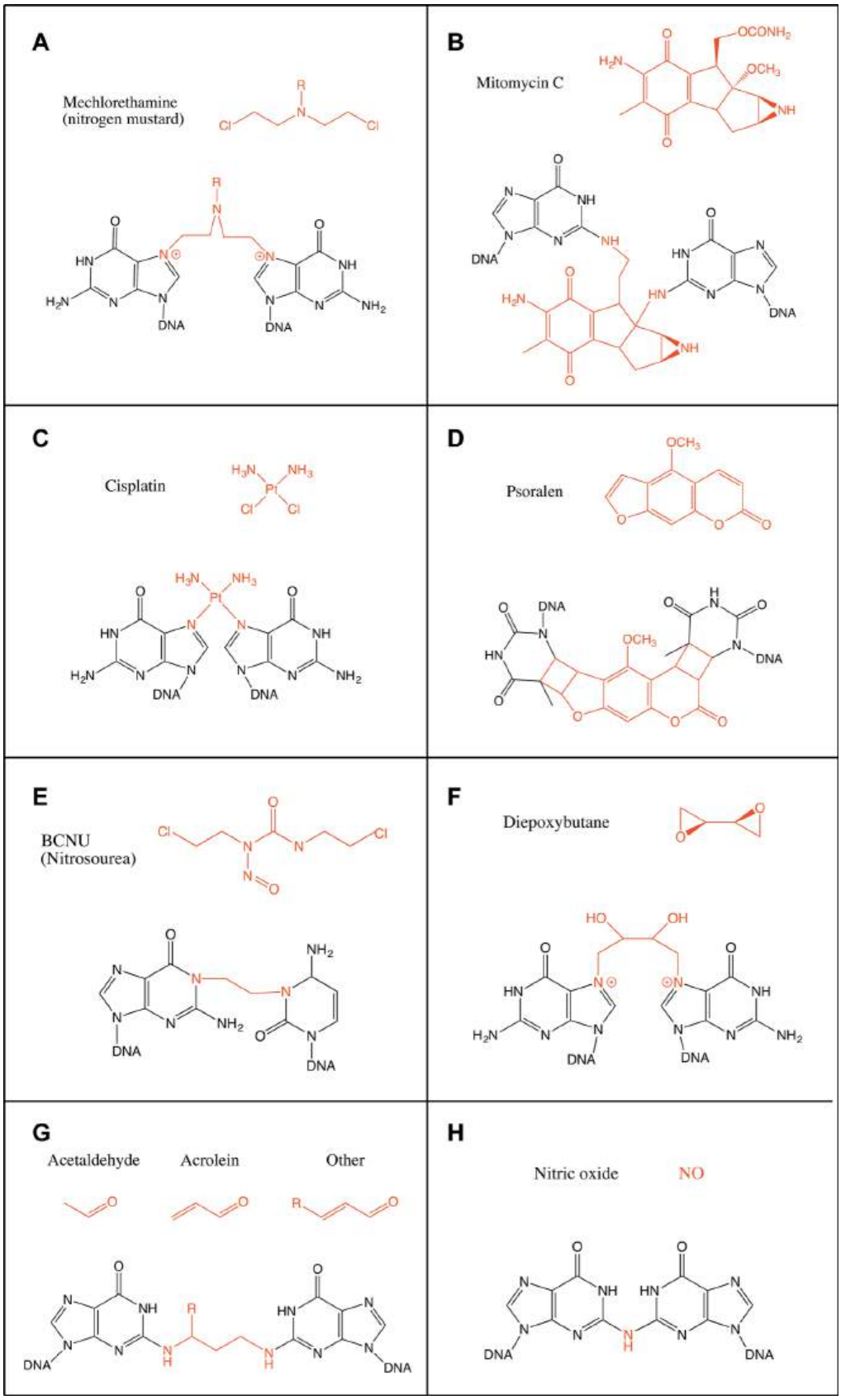

A
Mechlorethamine
(nitrogen mustard)
B
Mitomycin C
C
Cisplatin
D
Psoralen
E
BCNU
(Nitrosourea)
F
Diepoxybutane
G
Acetaldehyde
Acrolein
Other
H
Nitric oxide
NO

Figure 24.2.9: f Intrastrand and Interstrand DNA Crosslinks. Lopez and Martinez, Cell. Mol. Life Sci. (2016) 73:3097–3114 DOI 10.1007/s00018-016-2218-x. Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/),

UV light can cause molecular crosslinks to form between two pyrimidine residues, commonly two thymine residues, that are positioned consecutively within a strand of DNA, as shown in Figure 24.2.10 Two common UV products are cyclobutane pyrimidine dimers (CPDs) and 6–4 photoproducts. These premutagenic lesions alter the structure and possibly the base pairing. Up to 50–100 such reactions per second might occur in a skin cell during exposure to sunlight, but are usually corrected within seconds by photolyase reactivation or nucleotide excision repair. Uncorrected lesions can inhibit polymerases, cause misreading during transcription or replication, or lead to the arrest of replication. Pyrimidine dimers are the primary cause of melanomas in humans.

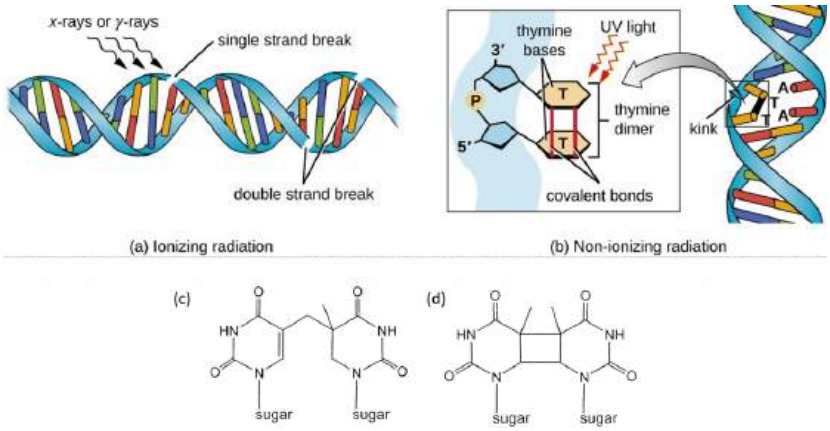


Figure 24.2.10: **Double-Stranded DNA breaks and Thymine Dimer Formation** (a) Ionizing radiation, such as X-rays and γ-rays contain enough energy to cause single and double-stranded breaks in the DNA backbone. (b) Energy from non-ionizing radiation such as UV light is directly absorbed by the DNA causing thymine residues that are adjacent within a single DNA strand to become cross-linked. (c) a 6,4-dimer that forms a single covalent bond and (d) thymine-thymine cyclobutane dimer that forms two covalent bonds between the thymine residues. Figures from: (a and b) Parker, et al (2019) Microbiology from Openstax and (c and d) Smokefoot

24.2.2.6: DNA Strand Breaks

Ionizing radiation such as that created by radioactive decay or in cosmic rays causes breaks in DNA strands (see Figure above). Low-level ionizing radiation may induce irreparable DNA damage (leading to replication and transcription errors needed for neoplasia or may trigger viral interactions) leading to premature aging and cancer. Chemical agents that form crosslinks within the DNA, especially interstrand crosslinks, can also lead to DNA strand breaks if the damaged DNA undergoes DNA replication. Crosslinked DNA can cause topoisomerase enzymes to stall in the transition state when the DNA backbone is cleaved. Instead of relieving supercoiling and resealing the backbone, the stalled topoisomerase remains covalently linked to the DNA in a process called ***abortive catalysis***. This leads to the formation of a single-stranded break in the case of Top1 enzymes or double-stranded breaks in the case of Top2 enzymes. DNA double-strand breaks due to topoisomerase stalling can also occur during the transcription of DNA, as shown in Figure 24.2.11 Abortive ***catalysis*** and the formation of DNA strand breaks during transcriptional events may serve as a damage sensor within the cell and help to instigate DNA damage response signaling pathways that initiate DNA repair processes.

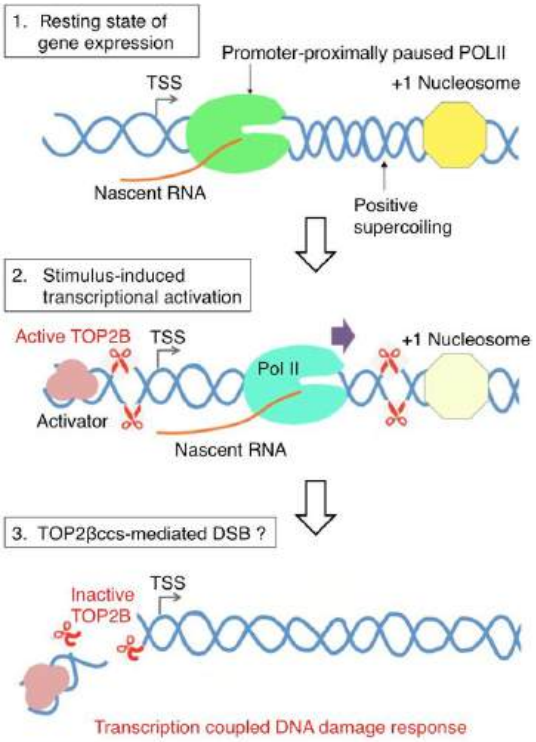


Figure 24.2.11: Model of Topoisomerase IIB (**TOP2B)-mediated DNA Double Strand Breaks During Transcription.** Morimoto, S., et al (2019) Genes 10:868

Panel (1) shows that in the uninduced state of transcription, Pol II is paused between +25 and +100 from the transcription start site. The pausing is attributed to different elements including pausing-stabilizing transcription factors, the +1 nucleosome, and DNA structure and torsion. Positive supercoiling ahead of Pol II may require the function of TOP2B.

Panel (2) shows transcription activation induced by various stimuli activates TOP2B to resolve DNA torsion in the promoter and gene body.

Panel (3) shows that in this process, double-strand breaks could be formed from abortive catalysis of TOP2B, which occurs frequently in some genes. This may be responsible for DNA damage response signaling that has been observed in a number of stimulus-inducible genes in humans. Figure from:

24.2.3: Cellular Stress and DNA Damage Response

Genetic damage produced by either exogenous or endogenous mechanisms represents an ongoing threat to the cell. To preserve genome integrity, eukaryotic cells have evolved repair mechanisms specific for different types of DNA Damage. However, regardless of the type of damage a sophisticated surveillance mechanism, that elicits ***DNA damage checkpoints***, detects and signals its presence to the DNA repair machinery. ***DNA damage checkpoints*** have been functionally conserved throughout eukaryotic evolution, with most of the relevant players in the checkpoint response highly conserved from yeast to humans. Checkpoints are induced to delay cell cycle progression and to allow cells time to repair damaged DNA before DNA replication, as shown in Figure 24.2.12 Once the damaged DNA is repaired, the checkpoint machinery triggers signals that will resume cell cycle progression. Within cells, multiple pathways contribute to DNA repair, but independent of the specific repair pathway involved, three phases of checkpoint activation are traditionally identified: (1) Sensing of damage, (2) Activating the signaling cascade, and (3) Switching on downstream effectors. The sensor phase recognizes the damage and activates the signal transduction phase to block cell cycle progression and select the appropriate repair pathway.

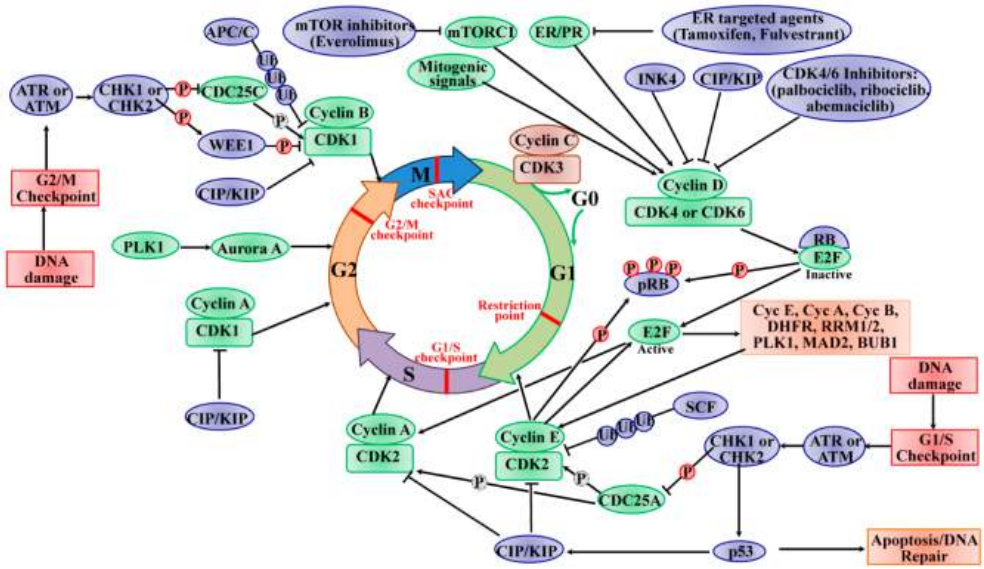


Figure 24.2.12: Cell Fates Following DNA Damage. Cell cycle checkpoints are induced by DNA damage, shown in red. Cyclin-Dependent Kinase Inhibitors (CIP/KIP), shown in purple, block cell cycle progression in all phases of the cell cycle (G1, S, G2, or M) following DNA damage by inhibiting cyclin-dependent kinase complexes (shown in green). Signaling cascades activated in response to DNA damage also elicit DNA repair pathways, or if the DNA damage is severe, programmed cell death (Apoptosis) will be activated. Image from: Ding, L., et al. (2020) Int. J. Mol. Sci. 21(6):1960

In addition to blocking cell cycle progression, DNA damage sensors also activate DNA repair mechanisms that are specific for the type of damage present. For example, single-stranded DNA breaks are repaired primarily by Base Excision Repair, bulky DNA adducts, and crosslinks are repaired by Nucleotide Excision Repair, and smaller nucleotide mutations, such as alkylation are repaired by Mismatch Repair. Cells also have two major mechanisms for repairing Double-Strand-Breaks (DSBs). They include Non-Homologous End-Joining (NHEJ) and Homologous Recombination (HR). If damage is too extensive to be repaired, apoptotic pathways will be elicited. In the following sections, details about the major DNA repair pathways will be given.

In multicellular organisms, the response to DNA damage can result in two major physiological consequences: (1) Cells can undergo cell cycle arrest, repair the damage, and re-enter the cell cycle, or (2) cells can be targeted for cell death (apoptosis) and removed from the population. The cell cycle process is highly conserved and precisely controlled to govern the genome duplication and separation into the daughter cells. The cell cycle consists of four distinct and ordered phases, termed G0/G1 (gap 1), S (DNA synthesis), G2 (gap 2), and M (mitosis). Multiple checkpoints exist within each stage of the cell cycle to ensure the faithful replication of DNA in the S phase and the precise separation of the chromosomes into daughter cells. The G1 and G2 phases are critical regulatory checkpoints, whereby the restriction point between the G1 and S phase determines whether the cells enter the S phase or exit the cell cycle to halt at the G0 phase. The cell cycle progression requires the activity of cyclin-dependent kinases (CDKs), a group of serine/threonine kinases. CDKs are activated when they form complexes with cyclin regulatory proteins that are expressed specifically at different stages of the cell cycle. Cyclins bind to and stabilize CDKs in their active conformation. The formation of cyclin/CDKs controls the cell-cycle progression via phosphorylation of the target genes, such as tumor suppressor protein retinoblastoma (Rb).

During DNA damage, the cell cycle is arrested or blocked by the action of cyclin-dependent kinase inhibitors. As noted in Figure 12.12, this is a complicated signal transduction cascade that has many downstream effects. A primary function of cell cycle arrest is that CDK inhibition allows time for DNA repair before cell-cycle progression into the S-phase or mitosis. As shown in Figure 25.2.12, two major cell-cycle checkpoints respond to DNA damage; they occur pre- and post-DNA synthesis in the G1 and G2 phases, respectively, and impinge on the activity of specific CDK complexes. The checkpoint kinases phosphatidylinositol 3-kinase (PI3K)-like protein kinases (PI3KKs) ataxia telangiectasia and Rad3-related (ATR) or ataxia telangiectasia mutated (ATM) protein, and the transducer checkpoint kinases CHK1 (encoded by the CHEK1 gene) and CHK2 (encoded by the CHEK2 gene) are key regulators of DNA damage signaling. The DNA damage signaling is detected by ATM/ATR, which then phosphorylates and activates CHK2/CHK1, respectively. The activated CHK2 is involved in the activation of p53, leading to p53-dependent early phase G1 arrest to allow time for DNA repair. The activation of p53 induces the expression of the Cyclin-Dependent Kinase Inhibitor (CKI) $p21^{CIP1}$ gene, leading to the inhibition of cyclin E/CDK2 complexes and subsequent upregulation of DNA repair machinery.

If the DNA repair cannot be completed successfully or the cells cannot program to respond to the stresses of viable cell-cycle arrest, the cells face the fate of apoptosis induced by p53. The activated CHK1 mediates temporary S phase arrest through phosphorylation to inactivate CDC25A, causing ubiquitination and proteolysis. Moreover, the activated CHK1 phosphorylates and inactivates CDC25C, leading to cell-cycle arrest in the G2 phase. The active CHK1 also directly stimulates the phosphorylation of WEE1, resulting in enhancing the inhibitory Tyr15 phosphorylation of CDK2 and CDK1 and subsequent cell-cycle blocking in the G2 phase. The activity of WEE1 can also be stimulated by the low levels of CDK activity in the G2 cell-cycle phase. The SAC, also known as the mitotic checkpoint, functions as the monitor of the correct attachment of the chromosomes to the mitotic spindle in metaphase, which is regulated by the TTK protein kinase (TTK, also known as monopolar spindle 1 (MPS1)). The activation of SAC transiently induces cell-cycle arrest by inhibiting the activation of APC/C. To establish and maintain the mitotic checkpoint, the TTK recruits many checkpoint proteins to kinetochores during mitosis via phosphorylating its substrates to ensure adequate chromosome segregation and genomic integrity. In this way, the genomic instability from chromosome segregation defects is protected by SAC. Once the SAC is passed, the APC/C E3 ligase complex stimulates and tags cyclin B and securin for ubiquitin-mediated degradation, leading to the initiation of mitosis. In a word, the checkpoints offer a failsafe mechanism to ensure the genomic integrity from the parental cell to the daughter cell. The signal transduction cascade of checkpoint activation eventually converges to CDK inhibition, which indicates the CDK function as a key driver of cell-cycle progression.

24.2.3.1: Mismatch Repair

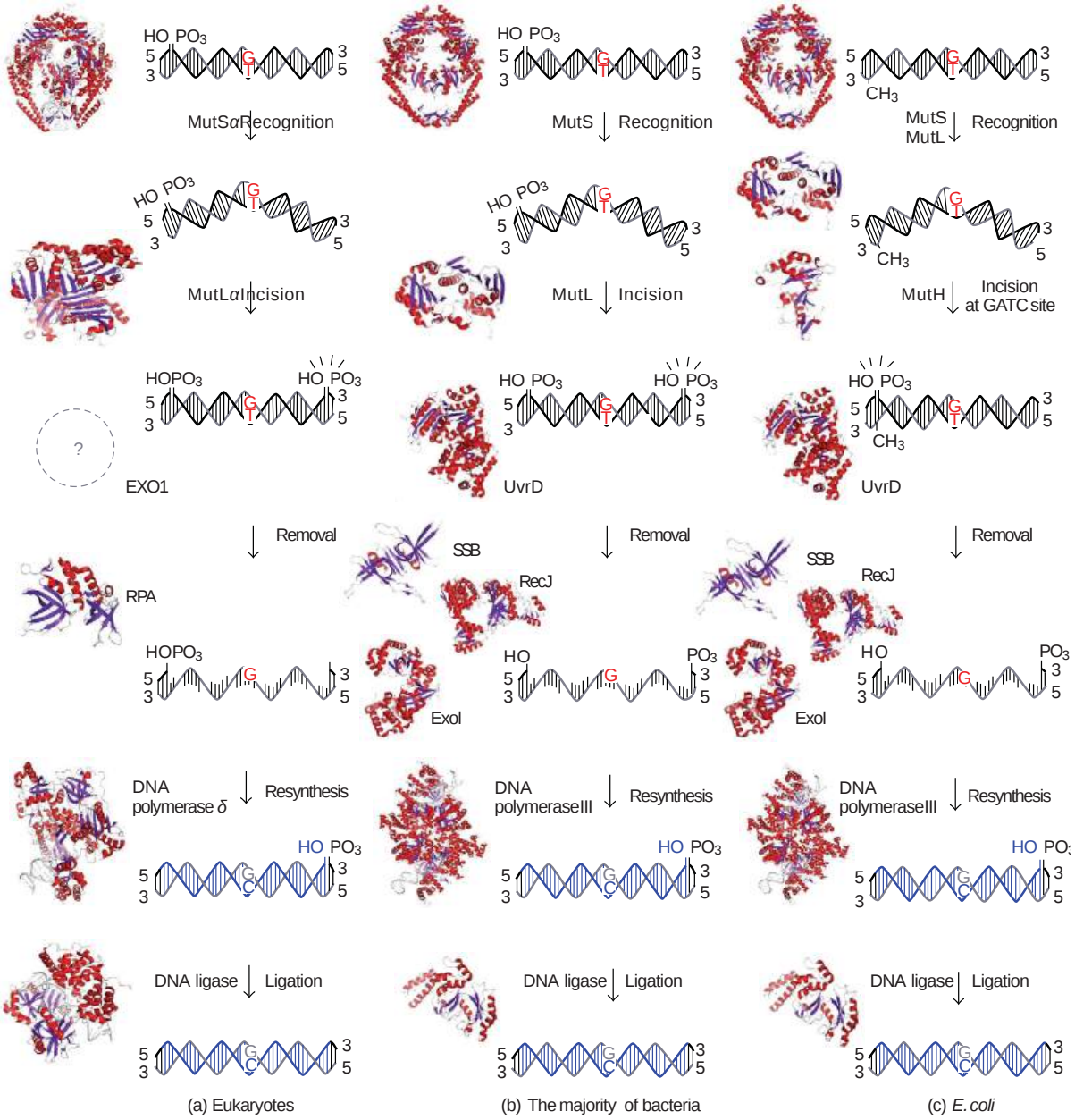
DNA **mismatch repair (MMR)** is a highly conserved DNA repair system that greatly contributes to maintaining genome stability through the correction of mismatched base pairs and small modifications of DNA bases, such as alkylation. The major source of mismatched base pairs is replication error, although it can arise also from other biological processes. Thus, the MMR machinery must have a mechanism for determining which strand of the DNA is the template strand and which strand has been newly synthesized. In *E. coli*, methylation of the DNA is a common post-replicative modification that occurs. Thus, in newly synthesized DNA, the unmethylated strand is recognized as the new strand, and the methylated strand is used as the template to repair mismatches. In *E.coli*, MMR increases the accuracy of DNA replication by 20–400-fold. Mutations and epigenetic silencing in MMR genes have been implicated in up to 90% of human hereditary nonpolyposis colon cancers, indicating the significance of this repair system in maintaining genomic stability. Post-replicative MMR is performed by the long-patch MMR mechanism in which multiple proteins are involved and a relatively long tract of the oligonucleotide is excised during the repair reaction. In contrast, particular kinds of mismatched base pairs are repaired through very short-patch MMR in which a short oligonucleotide tract is excised to remove the lesion. Table$24.2.2$below shows mismatch repair enzymes in bacteria, yeast, and humans.

Molecular function	*Thermus thermophilus*	*Escherichia coli*	*Saccharomyces cerevisiae*	*Homo sapiens*
Mismatch recognition	MutS	MutS	MutS α (MSH2/MSH6) MutS β (MSH2/MSH3)	MutS α (MSH2/MSH6) MutS β (MSH2/MSH3)
Strand incision	β-Clamp [*1]	—	PCNA	PCNA
	clamp-loader [*1]		RFC	RFC
	MutL		MutL α (MLH1/PMS1) MutL γ [*2] (MLH1/MLH3)	MutL α (MLH1/PMS2) MutL γ [*2] (MLH1/MLH3)
Strand incision	—	MutH	—	—
Match making	MutL	MutL	MutL α (MLH1/PMS1) MutL β (MLH1/MLH2) MutL γ (MLH1/MLH3)	MutL α (MLH1/PMS2) MutL β (MLH1/PMS1) MutL γ (MLH1/MLH3)
Strand excision (single-stranded DNA-binding)	SSB	SSB	RPA	RPA
Strand excision (exonuclease)	RecJ	RecJ	EXO1 [*3]	EXO1 [*3]
	ExoI	ExoI		
		ExoVII		
		ExoX		
Strand excision (helicase)	UvrD	UvrD	—	—
Repair synthesis	DNA polymerase III	DNA polymerase III	DNA polymerase δ	DNA polymerase δ

[*1] The involvement of bacterial clamp and clamp-loader in the strand incision reaction has not yet been confirmed. [*2] It is demonstrated that the endonuclease motif in MLH3 is responsible for in vivo MMR; however, the endonuclease activity of MutL γ has not yet been confirmed biochemically. [*3] In yeast and human, EXO1 has the 5' -flap endonuclease activity in addition to 5' - 3' exonuclease activity.

Table24.2.2: Mismatch repair enzymes in bacteria, yeast, and human

MMR in eukaryotes and most bacteria directs the repair to the error-containing strand of the mismatched duplex by recognizing the strand discontinuities. On the other hand, E. coli MMR reads the absence of methylation as a strand discrimination signal. The **MutS** protein recognizes mismatches, In both MMR systems, strand discrimination is conducted by **nicking endonucleases. MutL** homologs from eukaryotes and most bacteria incise the discontinuous strand to introduce the entry or termination point for the excision reaction. In E. coli, **MutH** nicks the unmethylated strand of the duplex to generate the entry point of excision. Figure 24.2.13shows different MMR pathway models.



24.2.4:

Figure 24.2.13 A schematic representation of MMR pathway models. Fukui, K. (2010) J. Nuc. Acids 260512. Creative Commons Attribution License

Vertical panel (a): Eukaryotic MMR. A DNA mismatch is generated by the misincorporation of a base during DNA replication. MutSα recognizes base-base mismatches and MutLα nicks the 3-or5-side of the mismatched base on the discontinuous strand. The resulting DNA segment is excised by the EXO1 exonuclease, in cooperation with the single-stranded DNA-binding protein RPA. The DNA strand is resynthesized by DNA polymerase δ and DNA ligase 1.

Vertical panel (b): MMR in mutH-less bacteria. Mismatched bases are recognized by MutS. After the incision of the discontinuous strand by MutL, the error-containing DNA strand is removed by the cooperative functions of DNA helicases, such as UvrD, the exonucleases RecJ and ExoI, and the single-stranded DNA-binding protein SSB. DNA polymerase III and DNA ligase fill the gap to complete the repair.

Vertical panel (c): E. coli MMR. MutS recognizes mismatched bases, and MutL interacts with and stabilizes the complex. Then, MutH endonuclease is activated to incise the unmethylated GATC site to create an entry point for the excision reaction. DNA

helicase, a single-stranded DNA-binding protein, and several exonucleases are involved in the excision reaction. PDB IDs of crystal structures in this figure are 2O8B (human MutSα), 1H7S (human MutLα), 1L1O (human RPA), 3IAY (human DNA polymerase δ), 1X9N (human DNA ligase 1), 1E3M (bacterial MutS), 1B63 (bacterial MutL), 2AZO (E. coli MutH), 2ISI (bacterial UvrD), 2ZXO (bacterial RecJ), 3C95 (bacterial ExoI), 2CWA (bacterial SSB), 2HQA (bacterial DNA polymerase III), and 2OWO (bacterial DNA ligase).

Figure 24.2.14shows an interactive iCn3D model of the E. Coli DNA Mismatch Repair Protein Muts Binding to a G-T Mismatch (1E3M).

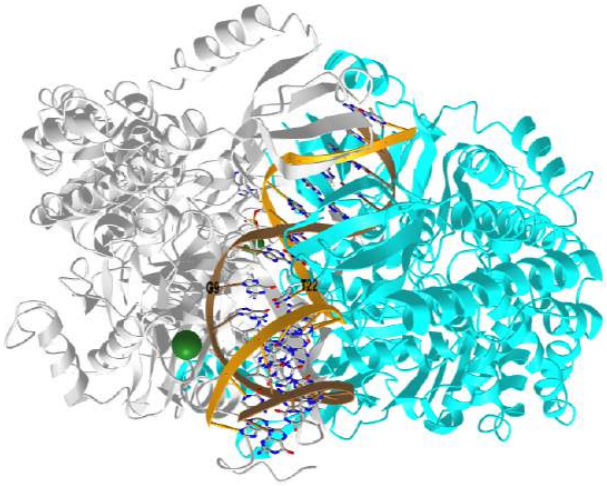


Figure 24.2.14: E. Coli DNA Mismatch Repair Protein MutS Binding to a G-T Mismatch (1E3M). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?YHB4ptjaS1mt7gp3A

The MutS monomers are colored gray and cyan. The mismatched G9-T22 base pair is labeled. ADP is shown in spacefill. Phe 36 from the gray monomer is shown in magenta.

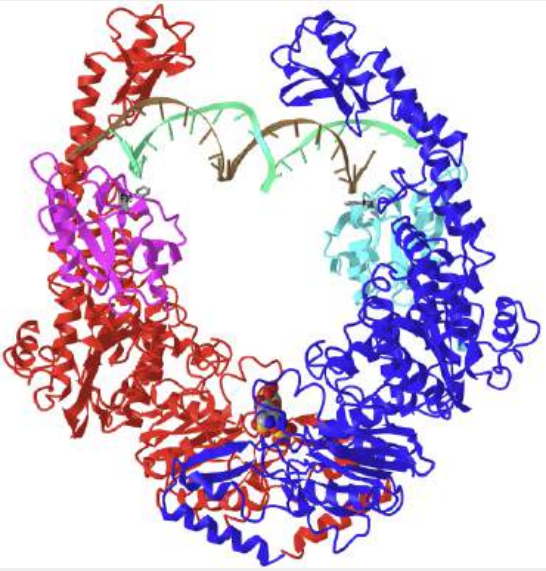
The conformation of the monomers is different, so the dimer displays pseudo symmetry. Both subunits contribute to DNA binding, but only one (gray) binds both ADP and the actual mismatched GT base pair, the llatterthrough minor grove interactions, which kinks the DNA. General major grove interaction clamps the DNA. Note how far away the ADP binds. Phenylalanine 36 in the gray subunit (which binds the mismatch) inserts adjacent to the mismatch.

ATP is bound and hydrolyzed to ADP by the MutS protein on binding the mismatch. Next, a dimer of MutL binds in a process that also requires ATP. MutH, a nuclease, also binds to MutL. The bound DNA is scanned until a "signal" is detected. In E. Coli, the signal is a GATC sequence that is methylated on just one strand and nicked by the MutH on the unmethylated GATC. Helicase II binds and unwinds the DNA in the region of the mismatch. Exonucleases (3' to 5' or 5' to 3') remove the sequence on the mismatched. PolII and DNA ligase then repair the DNA.

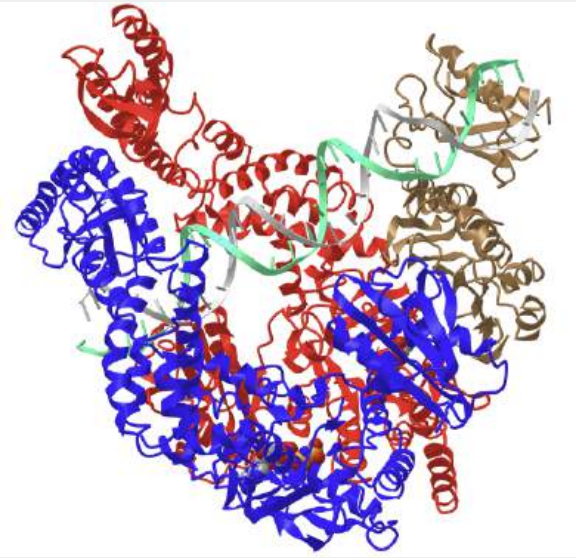
MutS is yet another fascinating enzyme as it must scan millions of DNA bases without initiating repair until it localizes a mismatch. A series of sequential conformation changes that lead to specific recognition of the mismatch must occur.

Fernandez-Leiro et al have determined the structure of MutS in a variety of stages along the repair pathway. Figure 24.2.15 shows interactive iCn3D models of the E. Coli MutS scanning form (EMD-11791, PDB 7AI5) and the more progressed MutS:MutL kink clamped form (EMD-11795, PDB 7AIC)

E. Coli MutS scanning form (EMD-11791, PDB 7AI5)	E. Coli MutS with MutL in kink clamped form (EMD-11795, PDB 7AIC)



Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...pdHkXxzXsZnqu6



Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...hQ6PeKbQeCw916

Figure 24.2.15: E. Coli MutS scanning form (EMD-11791, PDB 7AI5) (left) and the more progressed MutS:MutL kink clamped form (EMD-11795, PDB 7AIC) (right)

In the scanning form, the 2 monomers have been color coded as follows: monomer 1, N-terminal the interacts with DNA magenta, with the rest of the protein in red; monomer 2, the N-terminal part reacts with the DNA cyan, and the rest of the chain blue. In the kink-clamped state (right) the N-terminal magenta and N-terminal cyan sections were not present in the resolved structure. ATP is shown in spacefill.

These structures suggest that during scanning by the homoduplex of normal DNA, the conformational change necessary for MutS to morph to the kink-clamped state can not occur due to a steric block. Kinking of the DNA at the mismatch removes the steric block.

Click the links to download videos to get animations showing the role of MutS in mismatch repair. (Fernandez-Leiro, R., Bhairosing-Kok, D., Kunetsky, V. *et al.* The selection process of licensing a DNA mismatch for repair. *Nat Struct Mol Biol* **28**, 373–381 (2021). https://doi.org/10.1038/s41594-021-00577-7, with permission)

 Video 1: Molecular mechanism of DNA mismatch repair initiation. Front and side views of MutS passing through the first four stages of the repair cascade: DNA scanning, mismatch recognition, intermediate state, and MutL recruitment. The movements show a computational morphing between the four cryo-EM structures. MutS monomer A is shown in a pale-green color, monomer B in pale blue, DNA in dark gray, and $MutL^{LN40}$ in yellow.

 Video 2: Mismatch repair licensing at a mismatch. Top and side views of MutS as it transforms from the DNA-scanning state to the mismatch-bound state. The initial part of the movie represents the movement of monomer B relative to monomer A during the scanning state, as derived from the principal component analysis of the multibody refinement. Note that the MutS dimer explores multiple conformations, attempting to distort the DNA, without crossing over the opposite monomer. When a mismatch is present in the DNA, it allows MutS to deform and kink the DNA and the two MutS monomers to cross over in a clockwise manner. Movements show a computational morphing between the different states. MutS monomer A is shown in a pale-green color, monomer B in pale blue, and DNA in gray. The DNA mismatch is highlighted in pink

 Video 3: Multiple conformational changes of mismatch and connector domains tracking DNA. Front and side views of MutS as it goes from the mismatch-bound state to the $MutL^{LN40}$-bound clamp state via the intermediate state. MutS monomer A is shown in a pale green color, monomer B in pale blue, and DNA in dark gray. DNA mismatch is highlighted in pink. The mismatch domain is shown in dark green and the connector domain in light green. The ends of a central helix in the connector domain are colored in red and blue for clarity. Movements show a computational morphing between the different states.

24.2.5: Base Excision Repair

Most oxidized bases are removed from DNA by enzymes operating within the Base Excision Repair (BER) pathway. Single-stranded DNA breaks can also be repaired through this process. Removal of oxidized bases in DNA is fairly rapid. For example, 8-oxo-dG was increased 10-fold in the livers of mice subjected to ionizing radiation, but the excess 8-oxo-dG was removed with a half-life of 11 minutes. 8-oxoG is excised by 8-oxoguanine DNA glycosylase (OGG1) leaving an apurinic site (AP site), as shown in Figure 24.2.16. AP sites are then processed further into single-strand breaks via backbone incision of AP-endonuclease 1 (APE1). In long patch base excision repair, the base and some additional nucleotides are replaced dependent on the activity of polymerase delta (Polδ) and epsilon (Polε) together with proliferating cell nuclear antigen (PCNA). The old strand is removed by Flap-endonuclease 1 (FEN1), before ligase I (LigI) ligates the backbone back together. Short patch base excision repair constitutes of polymerase beta (Polβ) replacing the single missing base, ligase III (LigIII) ligating the DNA backbone back together, and X-ray repair cross-complementing protein 1 (XRCC1) aiding the process and serving as a scaffold for additional factors.

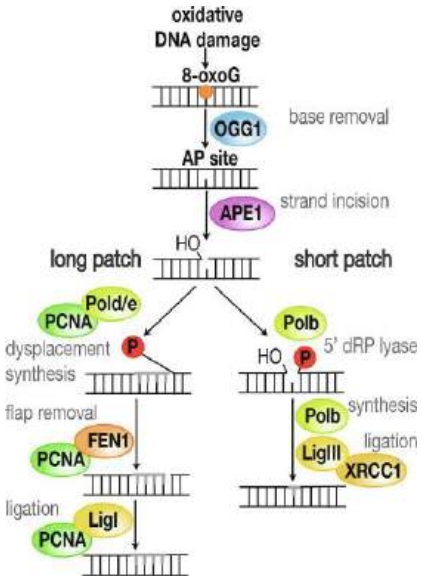


Figure (\PageIndex{16}\): Base Excision Repair. Poetsch, A.R. (2020) Comp & Struct Biotech J. 18:207-219.

Base excision repair (BER) of 8-oxo-7,8-dihydroguanine (8-oxoG). Oxidative DNA damage is repaired via several repair intermediates by base excision repair (BER). Through the emoval of the oxidized base, a reactive apurinic site (AP site) is formed. Incision of the strand creates a single-strand break, and the damaged site is then repaired through either short or long patch BER.

24.2.5.1: 25.2.6 Nucleotide Excision Repair

Bulky DNA adducts and DNA crosslinks, such as those caused by UV light are repaired using Nucleotide Excision Repair (NER) pathways. In higher eukaryotic cells, NER excises 24-32 nucleotide DNA fragments containing the damaged lesion with extreme accuracy. Reparative synthesis using the undamaged strand as a template, followed by ligation of the single-strand break that emerged as a result of the damage, is the final stage of DNA repair. The process involves the coordinated action of approximately 30 proteins that successively form complexes with variable compositions on the DNA. NER consists of two pathways distinct in terms of initial damage recognition. ***Global genome nucleotide excision repair (GG-NER)*** detects and eliminates bulky damages in the entire genome, including the untranscribed regions and silent chromatin, while ***transcription-coupled nucleotide excision repair (TC -NER)*** operates when damage to a transcribed DNA strand limits transcription activity. TC-NER is activated by the stalling of RNA polymerase II at the damaged sites of a transcribed strand, while GG-NER is controlled by the protein, XPC, a specialized protein factor that reveals the damage. A schematic GG-NER process is presented in Figure (\PageIndex{17}\) below.

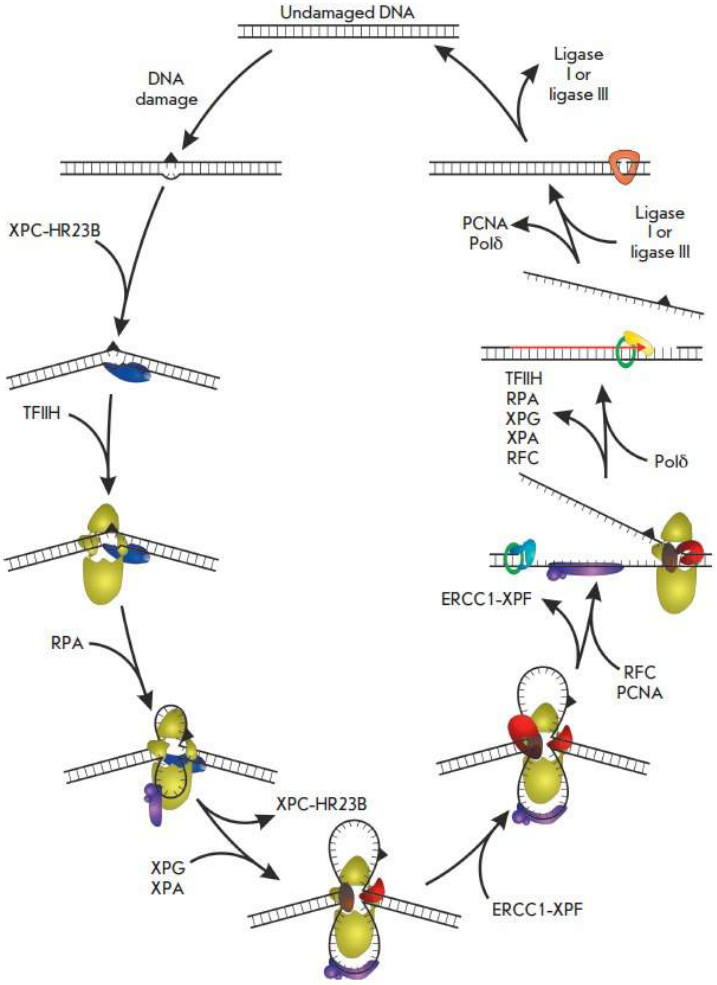


Figure (\PageIndex{17}\): Schematic of Global Genome Nucleotide Excision Repair. Figure from: Petruseva, I.O., et al (2014) 6(1):23-34.

Genetic mutations in NER pathway genes can result in UV-sensitive and high-carcinogenic pathologies, such as xeroderma pigmentosum (XP), Cockayne syndrome (CS), and trichothiodystrophy (TTD), as well as some neurodegenerative manifestations.

Xeroderma pigmentosum has provided the names of some of the genes involved in NER. Mutation of XP genes and loss of proper NER function causes the symptoms associated with the disease. People with XP have an impaired ability to repair bulky DNA adducts and crosslinks, such as thymine dimers that are caused by UV-light exposure. People suffering from XP have extreme photosensitivity, skin atrophy, hyperpigmentation, and a high rate of sunlight-induced skin cancer. The risk of internal tumors in XP patients is also 1,000-fold higher. Moreover, the disease is often associated with neurologic disorders. Currently, there is no effective treatment for this disorder.

The detection of bulky DNA lesions during NER is particularly challenging for a cell, which can be solved only through highly sensitive recognition that requires multiple protein components. In contrast to BER, where a damaged base is simultaneously recognized and eliminated by a single specialized glycosylase, specialized groups of proteins are responsible for the recognition of the lesion and the excision of the lesion in NER. In eukaryotic NER, universal sensor proteins perform the initial recognition of the total range of bulky damages. In the case of TC-NER, it occurs when the transcribing RNA polymerase II is stalled by damage; in GG-NER, these are complexes of the XPC factor and DDB1-DDB2 heterodimer (XPE factor) enhancing the repair of UV damage. In general, NER recognition of damage is a multistep process involving several proteins that form near damaged complexes of variable compositions. The process is completed by the formation of a preincision complex ready to eliminate a damaged DNA fragment by specialized NER endonucleases.

In a eukaryotic cell after stable XPC/DNA complex formation during the initial recognition of the damage, NER is performed by a ***repairasome***, which is a complex of variable composition and architecture consisting of a large number of subunits. Individual

subunits of the complex have no sufficient affinity and selectivity to the substrate (DNA containing bulky damage). The situation changes when specific protein complexes are established at the damage site. A total of 18 polypeptides must be accurately positioned within two or three DNA turns when a stable structure ready for damage removal is formed and excision starts. The structure of NER-associated proteins provides the possibility of contact with the DNA substrate and of dynamic specific protein-protein interactions. The changes in interactions performed by the same protein are one of the mechanisms that regulate the repair process and fine-tune the complexes, providing high-precision nucleotide excision repair.

24.2.5.2: 25.2.7 Repair of Double-Stranded DNA Breaks

Cells have evolved two main pathways to repair double-strand breaks within the DNA: ***the non-homologous end-joining (NHEJ) pathway***, which ensures direct resealing of DNA ends; and ***the homologous recombination (HR) pathway*** that relies on the presence of homologous DNA sequences for DSB repair, as shown in Figure (\PageIndex{18}\) below.

NHEJ repair is the simplest and most widely utilized mechanism to repair DSB that occur in DNA. Repair by NHEJ involves direct resealing of the two broken ends independently of sequence homology. Although being active throughout the cell cycle, NHEJ is relatively more important during the G1 phase. Proteins required for NHEJ include but are not restricted to, the highly conserved Ku70/Ku80 heterodimeric complex, DNA-dependent protein kinase catalytic subunit (DNA-PKcs), and DNA Ligase IV (LIG4) in complex with XRCC4. By directly binding DNA ends, Ku70/Ku80 ensures protection against exonucleases and, as such, acts as an inhibitor of HR. Very short sequence homologies are likely to help DNA end alignment before NHEJ-dependent repair, however, they are not strictly required. NHEJ protects genetic integrity by rejoining broken strands of DNA that may otherwise be lost during DNA replication and cell regeneration. However, during the process of NHEJ, insertions or deletions within the joined regions may occur (Fig 25.2.17).

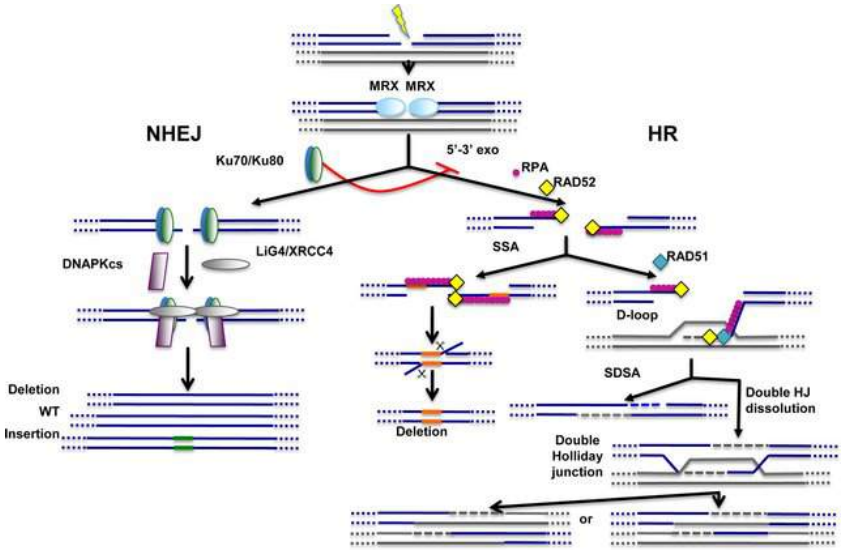


Figure (\PageIndex{18}\): Main Pathways of DNA Double-Stranded Break Repair. Diccotignies, A. (2013) Frontiers in Genetics 4(48):48

Non-homologous end-joining (NHEJ) and homologous recombination (HR) pathways act competitively to repair DNA double-strand breaks (DSBs). Key players of NHEJ and HR are depicted. The MRE11/RAD50/XRS2 (MRX) complex is recruited very early at DNA ends and appears to play important roles for both NHEJ and HR. Ku70/Ku80 heterodimer is required for NHEJ and, through inhibition of DNA end resection (5′–3′ exo), acts as a repressor of HR. The fidelity of NHEJ-dependent DSB repair is low and, most of the time, associated with nucleotide deletions and/or insertions at repair junctions. The common early step of HR-dependent mechanisms is the formation of ssDNA which is then coated by replication protein A (RPA). Single-strand annealing (SSA) mechanism requires the presence of direct repeats (shown in orange) on both sides of the break. SSA does not imply any strand invasion process and is therefore not dependent on RAD51 protein. Strand invasion and D-loop formation are however common steps of synthesis-dependent strand annealing (SDSA) and double Holliday junction (HJ) dissolution mechanisms. In the latter case, double Holliday junctions are resolved with or without crossing over.

In contrast to NHEJ, homologous recombination (HR) requires a homologous DNA sequence to serve as a template for DNA-synthesis-dependent repair and involves extensive DNA-end processing. As expected, HR is extremely accurate, as it leads to precise repair of the damaged locus using DNA sequences homologous to the broken ends. HR predominantly uses the sister

chromatid as a template for DSB repair, rather than the homologous chromosome. Correspondingly, HR is largely inhibited while cells are in the G1 phase of the cell cycle when the sister chromatid has not yet been replicated, as shown in panel (A) of Figure (\PageIndex{19}\) below. HR repair mechanisms play a bigger role in DSB repair that occurs after S-phase DNA replication (S-phase, G2, and M).

Repair through HR is not defined by a unique mechanism but operates through various mechanistically distinct DSB repair processes, including synthesis-dependent strand annealing (SDSA), double Holliday junction resolution, and single-strand annealing (SSA). The common step for HR-dependent DSB repair mechanisms is the initial formation of single-stranded DNA (ssDNA) for pairing with homologous DNA template sequences. For this to occur, the 5' DNA strand at the DSB is processed by multiple nucleases and accessory proteins to create a 3' ssDNA section that can be used as a template for recombination (see Figure 18 above).

Panel B of Figure (\PageIndex{19}\) below provides a more detailed look at the HR process. During the highly regulated process of HR, three main phases can be distinguished. Firstly, 3′-single-stranded DNA (ssDNA) ends are generated by nucleolytic degradation of the 5′-strands. This first step is catalyzed by endonucleases, including the MRN complex (consisting of Mre11, Rad50, and Nbs1). In the second step, the ssDNA-ends are coated by replication protein A (RPA) filaments. In the third step, RPA is replaced by Rad51 in a BRCA1- and BRCA2-dependent process, to ultimately perform the recombinase reaction using a homologous DNA template.

Importantly, HR is not only employed to repair DNA lesions induced by DNA-damaging agents but is also essential for proper chromosome segregation during meiosis. The relevance of HR in these physiological processes is illustrated by its strict requirement during development. Mice lacking key HR genes, such as *Brca1*, *Brca2*, or *Rad51*, display extensive genetic alterations which lead to early embryonic lethality. Whereas homozygous inactivation of HR genes is usually embryonic lethal, heterozygous inactivation of,*BRCA1* and *BRCA2*, does not interfere with cellular viability but rather predisposes individuals to cancer, including breast and ovarian cancer. The tumors that develop in individuals with heterozygous *BRCA1/2* mutations invariably lose their second *BRCA1/2* allele, indicating that in certain cancers, the absence of BRCA1/2 is compatible with cellular proliferation. How exactly such tumors cope with their HR defect is currently not fully understood.

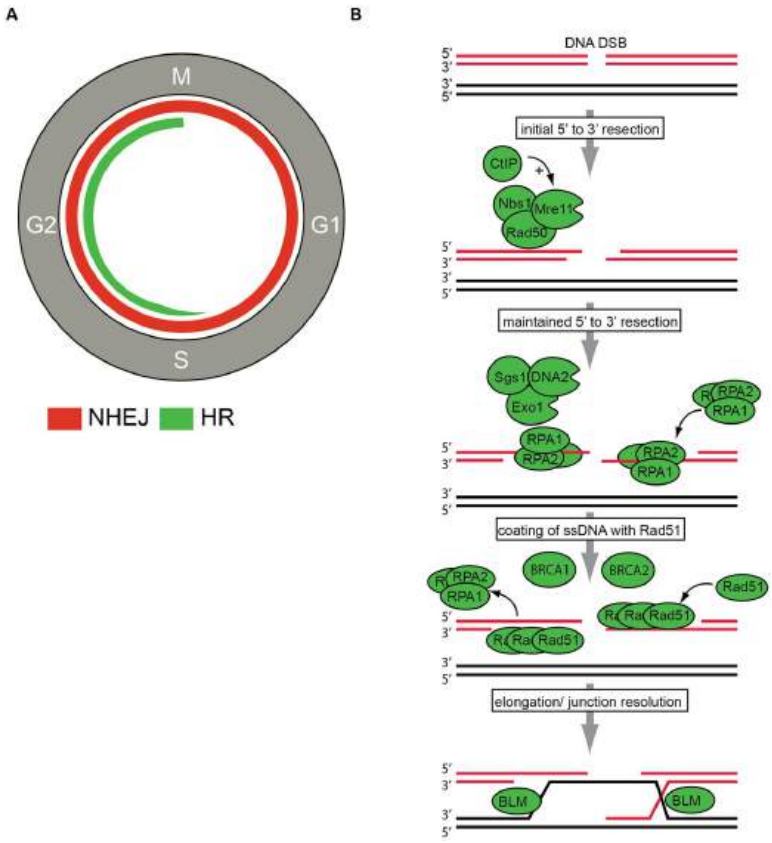


Figure (\PageIndex{19}\): **DNA double-strand break (DSBs) repair.** Krajewska, M., Fehrmann, R.S.N., de Vries, E.G.E., and van Vugt, A.A.T.M. (2015) Front. Genet. 6:96

Panel (A) shows the DNA DSBs repair pathways in the context of cell cycle regulation. Non-homologous end joining (NHEJ) can be performed throughout the cell cycle and is indicated with the red line. Homologous recombination (HR) can only be employed in S/G2 phases of the cell cycle and is indicated in green.

Pane (B) shows the key steps in the HR repair pathway are indicated. After DSB recognition, 5′–3′ end resection is initiated by the MRN (Mre11, Rad50, Nbs1) complex and CtIP. Subsequently, further resection by the Exo1, DNA2, and Sgs1 proteins is conducted to ensure 'maintained' resection. Then, resected DNA ends are bound by replication protein A (RPA). The actual recombination step within HR repair, termed strand exchange, is executed by the recombinase Rad51. Rad51 replaces RPA to eventually assemble helical nucleoprotein filaments called '*presynaptic filaments.*' This process is facilitated by other HR components, including BRCA1 and BRCA2. The final step of junction resolution is executed by helicases including Bloom syndrome, the RecQ helicase-like (BLM) helicase.

24.2.5.3: Error-Prone Bypass and Translesion Synthesis

If DNA is not repaired before DNA replication, the cell must employ another strategy to replicate the DNA, even in the presence of a DNA lesion. This is important to avoid causing double-stranded DNA breaks that can occur when a replisome stalls at the replication fork. Under these circumstances, another strategy that cells use to respond to DNA damage is to bypass lesions found during DNA replication and continue with the replicative process. DNA damage bypass can occur by recombination mechanisms or through a novel mechanism called ***translesion synthesis***. ***Translesion synthesis*** employs an alternate DNA polymerase that can substitute for a DNA polymerase that has stalled at the replication fork due to DNA damage. Specialized DNA polymerases, that are active in regions with DNA damage, have active sites that can accommodate fluctuations in DNA topography that enable them to bypass the lesions and continue with the replicative process.

The evolution of DNA polymerases that can tolerate the presence of distorted DNA lesions and continue with the replicative process can be seen at all levels of life, from prokaryotic, single-celled organisms through eukaryotic multicellular organisms, including humans. In fact, within vertebrates, there has been a large expansion of DNA polymerases that play a role in DNA

damage bypass mechanisms and highlight the importance of these processes in damage tolerance and cell survival, as shown in Table (\PageIndex{3}\) below.

Translesion DNA Polymerases		
E. coli	*S. cervisiae*	*H. sapiens*
DNA pol IV	REV1	REV1
DNA pol V	DNA polζ	DNA polζ
	DNA polη	DNA polη
		DNA polι
		DNA polκ
		DNA polμ
		DNA polλ
		DNA polθ
		DNA polυ
		DNA polσ

Table (\PageIndex{3}\): **DNA Polymerases involved in Error-Prone Bypass**

The activity of error-prone DNA polymerases is tightly regulated to avoid the rampant introduction of mutations within the DNA sequence. One of the main mechanisms that is employed within a replisome that is stalled at the replication fork due to DNA damage, involves the monoubiquitination of PCNA. Recall from Chapter 9, that PCNA is the sliding clamp that enables the DNA polymerase to bind tightly enough with the DNA during replication to mediate efficient DNA synthesis. Monoubiquitination of PCNA enables the recruitment of a translesion DNA polymerases and the bypass of the damaged lesions during DNA synthesis.

During translesion synthesis, the polymerase must insert a dNTP opposite of the lesion. None of the dNTP bases will likely be able to form stable hydrogen bond interactions with the damaged lesion. Thus, the nucleotide that causes the least distortion or repulsion will usually be added across from the lesion. This can cause transition or transversion mutations to occur at the lesion location. Alternatively, translesion polymerases can be prone to slippage, and either causes an insertion or deletion mutation in the vicinity of the DNA lesion. These slippages can lead to frameshift mutations if they occur within gene coding regions. Thus, over a lifetime, translesion synthesis in multicellular organisms can lead to an accumulation of mutations within somatic cells and cause the formation of tumors and the disease of cancer.

Evolution by natural selection is also possible due to random mutations that occur within germ cells. Occasionally, germline mutations may lead to a beneficial mutation that enhances the survival of an individual within a population. If this gene proves to enhance the survival of the population, it will be selected over time within the population and cause the evolution of that species. An example of a beneficial mutation is the case of a population of people that show resistance to HIV infection. Since the first case of infection with human immunodeficiency virus (HIV) was reported in 1981, nearly 40 million people have died from HIV infection, the virus that causes acquired immune deficiency syndrome (AIDS). The virus targets helper T cells that play a key role in bridging the innate and adaptive immune response, infecting and killing cells normally involved in the body's response to infection. There is no cure for HIV infection, but many drugs have been developed to slow or block the progression of the virus. Although individuals around the world may be infected, the highest prevalence among people 15–49 years old is in sub-Saharan Africa, where nearly one person in 20 is infected, accounting for greater than 70% of the infections worldwide, as shown in Figure (\PageIndex{20}\) below. Unfortunately, this is also a part of the world where prevention strategies and drugs to treat the infection are the most lacking.

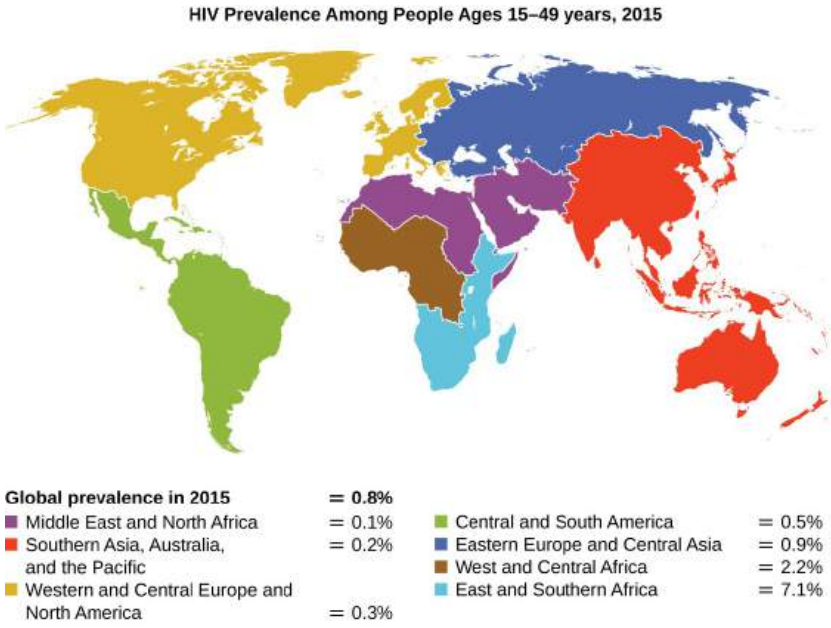


Figure (\PageIndex{20}\): HIV is highly prevalent in sub-Saharan Africa, but its prevalence is quite low in some other parts of the world. Figure from: Parker, et al (2019) Microbiology from Openstax

In recent years, scientific interest has been piqued by the discovery of a few individuals from northern Europe who are resistant to HIV infection. In 1998, American geneticist Stephen J. O'Brien at the National Institutes of Health (NIH) and colleagues published the results of their genetic analysis of more than 4,000 individuals. These indicated that many individuals of Eurasian descent (up to 14% in some ethnic groups) have a deletion mutation, called CCR5-delta 32, in the gene encoding CCR5. CCR5 is a coreceptor found on the surface of T-cells that is necessary for many strains of the virus to enter the host cell. The mutation leads to the production of a receptor to which HIV cannot effectively bind and thus blocks viral entry. People homozygous for this mutation have greatly reduced susceptibility to HIV infection, and those who are heterozygous have some protection from infection as well.

It is not clear why people of northern European descent, specifically, carry this mutation, but its prevalence seems to be highest in northern Europe and steadily decreases in populations as one moves south. Research indicates that the mutation has been present since before HIV appeared and may have been selected for in European populations as a result of exposure to the plague or smallpox. This mutation may protect individuals from the plague (caused by the bacterium *Yersinia pestis*) and smallpox (caused by the variola virus) because this receptor may also be involved in these diseases. The age of this mutation is a matter of debate, but estimates suggest it appeared between 1875 years to 225 years ago, and may have been spread from Northern Europe through Viking invasions.

This exciting finding has led to new avenues in HIV research, including looking for drugs to block CCR5 binding to HIV in individuals who lack the mutation. Although DNA testing to determine which individuals carry the CCR5-delta 32 mutation is possible, there are documented cases of individuals homozygous for the mutation contracting HIV. For this reason, DNA testing for the mutation is not widely recommended by public health officials so as not to encourage risky behavior in those who carry the mutation. Nevertheless, inhibiting the binding of HIV to CCR5 continues to be a valid strategy for the development of drug therapies for those infected with HIV.

24.2.6: Practice Problems

24.2.6.1: Multiple Choice

Which of the following is a change in the sequence that leads to the formation of a stop codon?

1. missense mutation
2. nonsense mutation
3. silent mutation
4. deletion mutation

[reveal-answer q="745512"]Show Answer[/reveal-answer]
[hidden-answer a="745512"]Answer b. A nonsense mutation is a change in the sequence that leads to formation of a stop codon.

[/hidden-answer]

The formation of pyrimidine dimers results from which of the following?

1. spontaneous errors by DNA polymerase
2. exposure to gamma radiation
3. exposure to ultraviolet radiation
4. exposure to intercalating agents

[reveal-answer q="709151"]Show Answer[/reveal-answer]
[hidden-answer a="709151"]Answer c. The formation of pyrimidine dimers results from exposure to ultraviolet radiation.[/hidden-answer]

Which of the following is an example of a frameshift mutation?

1. a deletion of a codon
2. missense mutation
3. silent mutation
4. deletion of one nucleotide

[reveal-answer q="688366"]Show Answer[/reveal-answer]
[hidden-answer a="688366"]Answer a. The deletion of one nucleotide is an example of a frameshift mutation.[/hidden-answer]

Which of the following is the type of DNA repair in which thymine dimers are directly broken down by the enzyme photolyase?

1. direct repair
2. nucleotide excision repair
3. mismatch repair
4. proofreading

[reveal-answer q="755583"]Show Answer[/reveal-answer]
[hidden-answer a="755583"]Answer a. In a direct repair, thymine dimers are directly broken down by the enzyme photolyase. [/hidden-answer]

Which of the following regarding the Ames test is true?

1. It is used to identify newly formed auxotrophic mutants.
2. It is used to identify mutants with restored biosynthetic activity.
3. It is used to identify spontaneous mutants.
4. It is used to identify mutants lacking photoreactivation activity.

[reveal-answer q="770537"]Show Answer[/reveal-answer]
[hidden-answer a="770537"]Answer b. It is used to identify mutants with restored biosynthetic activity.[/hidden-answer]

24.2.6.4: Fill in the Blank

A chemical mutagen that is structurally similar to a nucleotide but has different base-pairing rules is called a _____.

[reveal-answer q="702924"]Show Answer[/reveal-answer]
[hidden-answer a="702924"]A chemical mutagen that is structurally similar to a nucleotide but has different base-pairing rules is called a **nucleoside analog**.[/hidden-answer]

The enzyme used in light repair to split thymine dimers is called _____.

[reveal-answer q="939657"]Show Answer[/reveal-answer]
[hidden-answer a="939657"]The enzyme used in light repair to split thymine dimers is called **photolyase**.[/hidden-answer]

The phenotype of an organism that is most commonly observed in nature is called the _____.

[reveal-answer q="640686"]Show Answer[/reveal-answer]
[hidden-answer a="640686"]The phenotype of an organism that is most commonly observed in nature is called the **wild type**. [/hidden-answer]

24.2.6.5: True/False

Carcinogens are typically mutagenic.

[reveal-answer q="166576"]Show Answer[/reveal-answer]
[hidden-answer a="166576"]True[/hidden-answer]

24.2.6.6: Think about It

Why is it more likely that insertions or deletions will be more detrimental to a cell than point mutations?

24.2.6.7: Critical Thinking

Below are several DNA sequences that are mutated compared with the wild-type sequence: 3′-T A C T G A C T G A C G A T C-5′. Envision that each is a section of a DNA molecule that has separated in preparation for transcription, so you are only seeing the template strand. Construct the complementary DNA sequences (indicating 5′ and 3′ ends) for each mutated DNA sequence, then transcribe (indicating 5′ and 3′ ends) the template strands, and translate the mRNA molecules using the genetic code, recording the resulting amino acid sequence (indicating the N and C termini). What type of mutation is each?

Mutated DNA Template Strand #1: 3′-T A C T G T C T G A C G A T C-5′	
Complementary DNA sequence:	[practice-area rows="1"][/practice-area]
mRNA sequence transcribed from template:	[practice-area rows="1"][/practice-area]
Amino acid sequence of peptide:	[practice-area rows="1"][/practice-area]
Type of mutation:	[practice-area rows="1"][/practice-area]

Mutated DNA Template Strand #2: 3′-T A C G G A C T G A C G A T C-5′	
Complementary DNA sequence:	[practice-area rows="1"][/practice-area]
mRNA sequence transcribed from template:	[practice-area rows="1"][/practice-area]
Amino acid sequence of peptide:	[practice-area rows="1"][/practice-area]
Type of mutation:	[practice-area rows="1"][/practice-area]

Mutated DNA Template Strand #3: 3′-T A C T G A C T G A C T A T C-5′	
Complementary DNA sequence:	[practice-area rows="1"][/practice-area]
mRNA sequence transcribed from template:	[practice-area rows="1"][/practice-area]
Amino acid sequence of peptide:	[practice-area rows="1"][/practice-area]
Type of mutation:	[practice-area rows="1"][/practice-area]

Mutated DNA Template Strand #4: 3′-T A C G A C T G A C T A T C-5′	
Complementary DNA sequence:	[practice-area rows="1"][/practice-area]
mRNA sequence transcribed from template:	[practice-area rows="1"][/practice-area]
Amino acid sequence of peptide:	[practice-area rows="1"][/practice-area]
Type of mutation:	[practice-area rows="1"][/practice-area]

<h212references">25.2.10 References

1. <lifootnote-99-2">World Health Organization. " Global Health Observatory (GHO) Data, HIV/AIDS." http://www.who.int/gho/hiv/en/. Accessed August 5, 2016.↵
2. Parker, N., Schneegurt, M., Thi Tu, A-H., Lister, P., Forster, B.M. (2019) Microbiology. *Openstax*. Available at: https://opentextbc.ca/microbiologyopenstax/
3. Wikipedia contributors. (2020, July 15). DNA oxidation. In *Wikipedia, The Free Encyclopedia*. Retrieved 03:44, July 16, 2020, from https://en.Wikipedia.org/w/index.php?title=DNA_oxidation&oldid=967811859

4. Wakim, S. and Grewal, M. (2020) Human Biology. *Libretexts.* Available at: https://bio.libretexts.org/Bookshelves/Human_Biology/Book%3A_Human_Biology_(Wakim_and_Grewal)
5. Ahmad, A., Nay, S.L. and O'Conner, T.R. (2015) Chapter 4 Direct Reversal Repair in Mammalian Cells. Published through INTECH. Available at: https://cdn.intechopen.com/pdfs/48191.pdf
6. Wikipedia contributors. (2020, April 20). AP site. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:15, July 23, 2020, from https://en.Wikipedia.org/w/index.php?title=AP_site&oldid=952117602
7. Wikipedia contributors. (2020, July 4). Benzo(a)pyrene. In *Wikipedia, The Free Encyclopedia*. Retrieved 06:15, July 24, 2020, from https://en.Wikipedia.org/w/index.php?title=Benzo(a)pyrene&oldid=965990545
8. Wikipedia contributors. (2020, June 23). Pyrimidine dimer. In *Wikipedia, The Free Encyclopedia*. Retrieved 06:54, July 24, 2020, from https://en.Wikipedia.org/w/index.php?title=Pyrimidine_dimer&oldid=964108515
9. Morimoto, S., Tsuda, M., Bunch, H., Sasanuma, H., Ausin, C. and Takeda, S. (2019) Type II DNA Topoisomerases Cause Spontaneous Double-Strand Breaks in Genomic DNA. Genes 10:868. Available at: https://www.researchgate.net/publication/336916880_Type_II_DNA_topoisomerases_cause_spontaneous_double-strand_breaks_in_genomic_DNA/figures?lo=1
10. Ding, L., Cao, J., Lin, W., Chen, H., Xiong, X., Ao, H., Yu, M., Lin, J., Cui, Q. (2020) The Roles of the Cyclin-Dependent Kinases in Cell-Cycle Progression and Therapeutic Strategies in Human Breast Cancer. *Int. J. Mol Sci* 21(6):1960. Available at: https://www.mdpi.com/1422-0067/21/6/1960/htm
11. Verma, N., Franchitto, M., Zonfrilli, A., Cialfi, S., Palermo, R., and Talora, C. (2019) DNA Damage Stress: Cui Prodest? *Int. J. Mol. Sci.* 20(5):1073. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6429504/
12. Fukui, K. (2010) DNA Mismatch Repair in Eukaryotes and Bacteria. *J. Nuc. Acids* 260512. Available at: https://www.hindawi.com/journals/jna/2010/260512/#copyright
13. Petruseva, I.O., Evdokimov, A.N., and Lavrik, O.I. (2014) Molecular Mechanism of Global Genome Nucleotide Excision Repair. *Acta Naturae 6(1):23-34.* Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3999463/
14. Decottingnies, A. (2013) Alternative end-joining mechanisms: A historical perspective. *Frontiers in Genetics* 4(48):48. Available at: https://www.researchgate.net/publication/236129718_Alternative_end-joining_mechanisms_A_historical_perspective
15. Vitor, A.C., Huertas, P., Legube, G., and de Almeida, S.F. (2020) Studying DNA Double-Strand Break Repair: An Every-Growing Toolbox. *Front. Mol Biosci* 7:24. Available at: https://www.frontiersin.org/articles/10.3389/fmolb.2020.00024/full
16. Krajewska, M., Fehrmann, R.S.N., de Vries, E.G.E., and van Vugt, A.A.T.M. (2015) Regulators of homologous recombination repair as novel targets for cancer treatment. *Front. Genet.* 6:96. Available at: https://www.frontiersin.org/articles/10.3389/fgene.2015.00096/full

---

24.3: DNA Recombination

Search Fundamentals of Biochemistry

24.3.1: Homologous Recombination

Homologous recombination is a type of genetic recombination in which genetic information is exchanged between two similar or identical molecules of double-stranded or single-stranded nucleic acids (usually DNA as in cellular organisms but may be also RNA in viruses). As noted in section 25.2, This process is widely used by cells to accurately repair harmful breaks that occur on both strands of DNA, known as **double-strand breaks (DSB)**, in a process called **homologous recombinational repair (HRR)**. Homologous recombination also produces new combinations of DNA sequences during **meiosis**, the process by which eukaryotes make gamete cells, like sperm and egg cells in animals. These new combinations of DNA represent genetic variation in offspring, which in turn enables populations to adapt during evolution. Homologous recombination is also used in horizontal gene transfer to exchange genetic material between different strains and species of bacteria and viruses.

Although homologous recombination varies widely among different organisms and cell types, for double-stranded DNA (dsDNA) most forms involve the same basic steps. After a double-strand break occurs, sections of DNA around the 5' ends of the break are cut away in a process called *resection*. In the *strand invasion* step that follows, an overhanging 3' end of the broken DNA molecule then "invades" a similar or identical DNA molecule that is not broken. After strand invasion, the further sequence of events may follow either of two main pathways discussed below (see Models); the DSBR (double-strand break repair) pathway or the SDSA (synthesis-dependent strand annealing) pathway. Homologous recombination that occurs during DNA repair tends to result in non-crossover products, in effect restoring the damaged DNA molecule as it existed before the double-strand break.

There are several different ways to repair DSB as illustrated in Figure 24.3.1. The broken (or **resected**) DNA with a double-stranded break must find and come together (synapse), invade (or intertwine) with the DNA of a donor, that is homologous to it. The repair can then ensue. In somatic cells that undergo mitosis and not meiosis, the preferred donor is the sister chromatid (the copy of one chromosome made during cell division) and not the homologous chromosome (from the diploid cell). Variants include **synthesis-dependent strand annealing pathway (SDSA)**. Other variants include **nonhomologous end-joining (NHEJ), microhomology-mediated end-joining (MMEJ),** and **double Holiday junction (dHJ).**

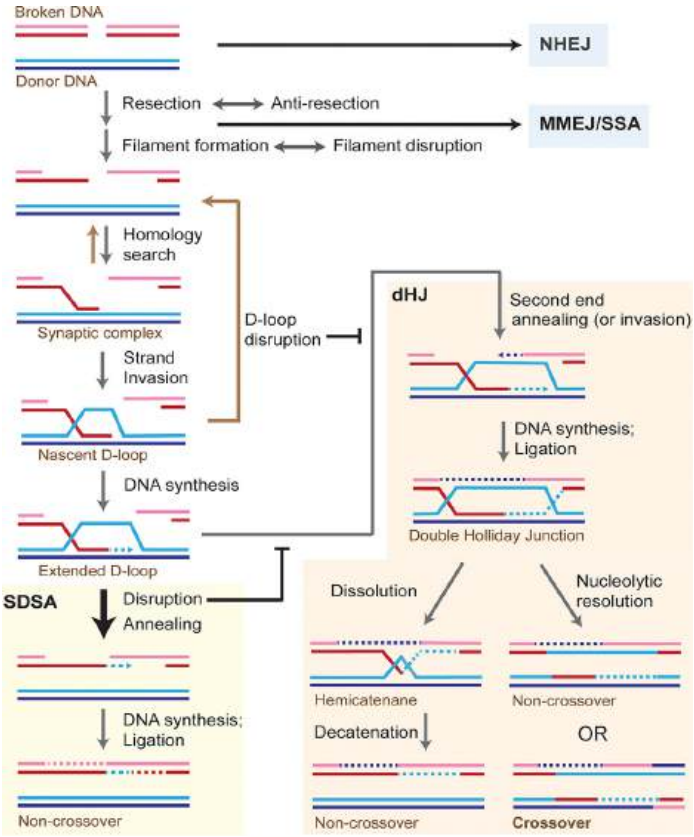
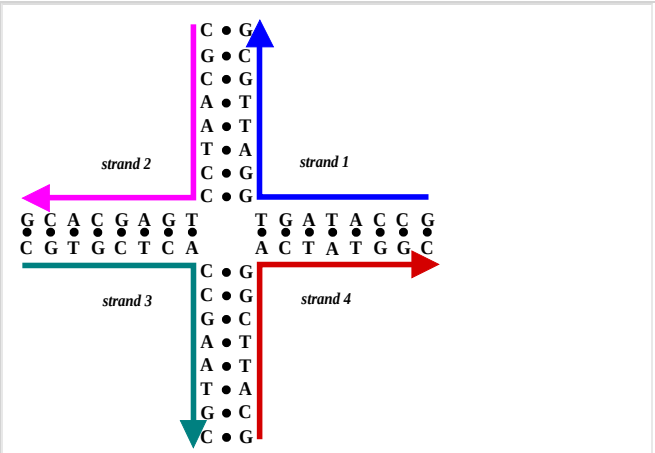


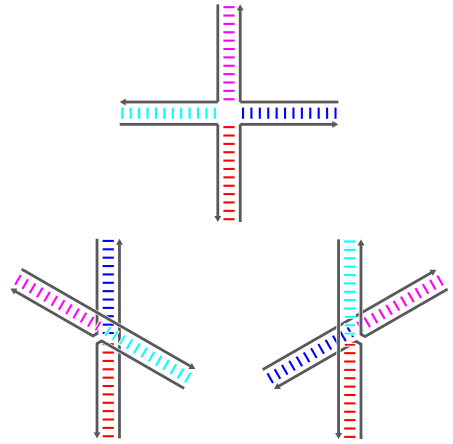
Figure 24.3.1: Model for the repair of DNA double-strand breaks by homologous recombination in somatic cells. Wright et al. J. Biol. Chem. (2018) 293(27) 10524 –10535. Creative Commons Attribution (CC BY 4.0)

When a DNA double-strand break (*DSB*) occurs in a DNA molecule, a repair can proceed by multiple pathways largely controlled by end resection. NHEJ is capable of repairing unresected or minimally resected DSBs in a template-independent fashion. MMEJ and single-strand annealing (SSA) rely on different extents of homology between the two DSB ends for repair independent of a donor molecule. Homologous recombination proceeds as shown in the figure using a homologous donor DNA. Most of the extended D-loops in somatic cells are disrupted and subsequently repaired by SDSA. The result of the repair by SDSA is always a noncrossover outcome, thus avoiding the loss of heterozygosity produced by somatic crossovers. SDSA occurs by disruption of the extended D-loop and annealing the newly synthesized DNA with the second end of the broken molecule. Alternatively, the newly synthesized strand may invade the second end. The extended D-loop can also undergo second-end capture or invasion to form a double Holliday junction (*dHJ*). This may either lead to a crossover or a noncrossover outcome. Invasion by the second break end makes dHJ formation and hence crossover outcome more likely for another model for crossover generation. dHJs can be dissolved into noncrossovers by the concerted action of the Sgs1–Top3–Rmi1 complex to migrate the two junctions toward each other and then decatenate the strands of the hemicatenane by the Top3 topoisomerase activity. Each *colored line* indicates a strand of DNA, and *dotted lines* represent DNA synthesis.

In the process of homologous recombination, a key intermediate is the **Holiday junction,** named after Robin Holiday who discovered it. It consists of branched nucleic acid with four double-stranded arms joined. A holiday junction is seen as the crossing of **red** and **blue** strands in the middle of Figure 1 labeled Nascent D-loop. Also, one is seen in the Extended D-loop just below it. A double Holiday junction is seen in the middle of the right-hand section. Two views of Holiday junctions are shown in Figure 24.3.2



By Донор - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/inde...curid=48470765



By Antony-22 - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/inde...curid=38557614

Figure 24.3.2 Two views of holiday junctions. The left-hand panel shows the primary and secondary sequences and some tertiary (3D) aspects of base-stacking conformational isomers of the Holliday junction. The bases nearest to the junction point determine which stacked isomer dominates.

Figure 24.3.3 shows an interactive iCn3D model of the structure of the Holliday junction intermediate in Cre-loxP site-specific recombination (3CRX).

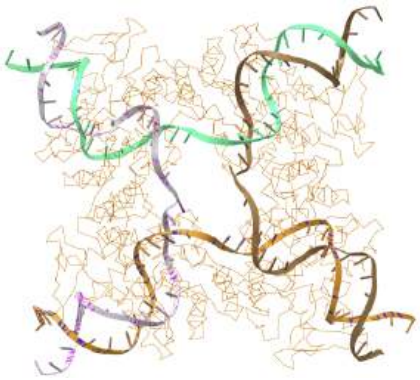


Figure 24.3.3: Structure of the Holliday junction intermediate in Cre-loxP site-specific recombination (3CRX). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...CQeGXUXiQa4AY8

The alpha carbon backbone of the four Cre recombinase monomers in the tetramer is shown in red. The DNA is nearly planar with a twofold-symmetric DNA intermediate that is similar to a square and stacked Holiday junction for the DNA in the unbound state.

Homologous recombination is conserved across all three domains of life as well as DNA and RNA viruses, suggesting that it is a nearly universal biological mechanism. The discovery of genes for homologous recombination in protists—a diverse group of eukaryotic microorganisms—has been interpreted as evidence that meiosis emerged early in the evolution of eukaryotes. Since their dysfunction has been strongly associated with increased susceptibility to several types of cancer, the proteins that facilitate homologous recombination are topics of active research. Homologous recombination is also used in gene targeting, a technique for introducing genetic changes into target organisms. For their development of this technique, Mario Capecchi, Martin Evans, and Oliver Smithies were awarded the 2007 Nobel Prize for Physiology or Medicine; Capecchi[3] and Smithies[4] independently discovered applications to mouse embryonic stem cells, however, the highly conserved mechanisms underlying the DSB repair model, including uniform homologous integration of transformed DNA (gene therapy), were first shown in plasmid experiments by Orr-Weaver, Szostack, and Rothstein.

Before the beginning of meiosis, the replication of the DNA is required to form sister chromatids, as shown in Step 1 in Figure 24.3.4 Once replicated, the DNA will condense and begin the process of meiosis. As the cells enter metaphase of meiosis, heterologous chromosomes are paired together (Step 2). When the homologous chromosomes are paired, they can under genetic mixing or DNA recombination form new chromosomal arrangements that are unique from the parental chromosomes (Step 3). Once the process of recombination is finished, the heterologous chromosomes are separated into two different daughter cells (Step 4). This is called ***Meiosis I***. At this stage the chromosomes have been reduced from the diploid to the haploid state, however, each chromosome set is still paired with its sister chromatid and needs to undergo a second round of cell division to produce the final set of four gametes (Step 5). This is called ***Meiosis II*** and results in the formation of four haploid gametes that are genetically unique.

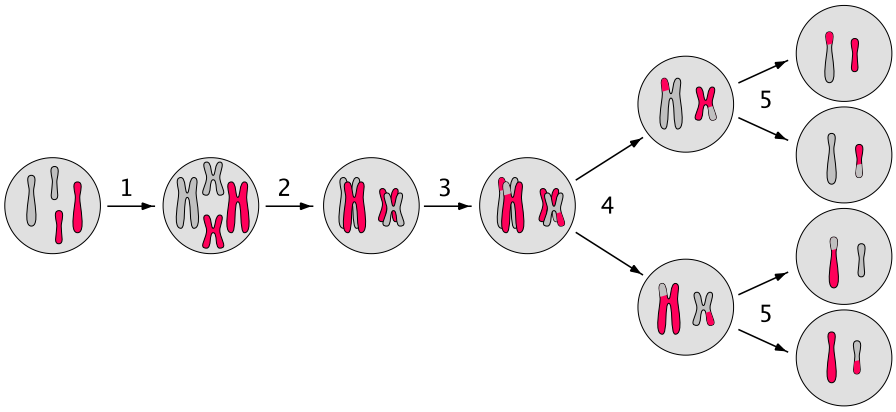


Figure 24.3.4: The Process of Meiosis. Peter Coxhead

During the process of ***meiosis,*** cell division is used to create the gametes or reproductive cells of an organism (the egg and the sperm cells). Meiosis results in the reduction of the genome from the 2n or diploid state to the 1n or haploid state. As you can see in Figure 25.3.1, the process of meiotic division results in the generation of four genetically unique haploid cells and involves the pairing of heterologous chromosomes during metaphase of meiosis. In humans, the meiotic process results in four viable sperm cells in the male and a single viable egg in the female. The other three cells produced in the female during the meiotic division are termed polar bodies and are very small and do not contain enough cytoplasmic components to survive. They get reabsorbed into the body. In either case, the resulting egg or sperm cell each carries a single copy of the genome and is in the haploid state.

It is during the process of meiosis that homologous recombination occurs in a controlled manner to introduce genetic variation into the resulting gametes. As a result, each egg and sperm cell has a unique genetic makeup that is a mixture of both parental copies of the genome.

Proper segregation during meiosis requires that homologs be connected by the combination of crossovers and sister chromatid cohesion. To generate crossovers, numerous double-strand breaks (DSBs) are introduced throughout the genome by the conserved Spo11 endonuclease. DSB formation and its repair are then highly regulated to ensure that homologous chromosomes contain at least one crossover and that no DSBs remain before meiosis I segregation. The synaptonemal complex (SC) is a meiosis-specific structure formed between homologous chromosomes during prophase that promotes DSB formation and biases the repair of DSBs to homologous chromosomes rather than back to the sister chromatids, ensuring that genetic recombination occurs. ***Synapsis,*** the pairing of homologous chromosomes, occurs when a particular recombination pathway is successful in establishing stable interhomolog connections.

24.3.2: Formation of the Synaptonemal Complex

In the 1950s, electron microscopists discovered an evolutionarily conserved, meiosis-specific structure formed between homologous chromosomes, unique to prophase I, called the SC, as shown in Figure 24.3.5. The SC physically connects homologs during prophase I and is removed before metaphase I, when homologs are connected instead by the combination of crossovers and sister chromatid cohesion. What is the function of the SC? Decades of research have shown that this elaborate chromosomal structure is critical for the regulation of recombination, the process by which crossovers are generated.

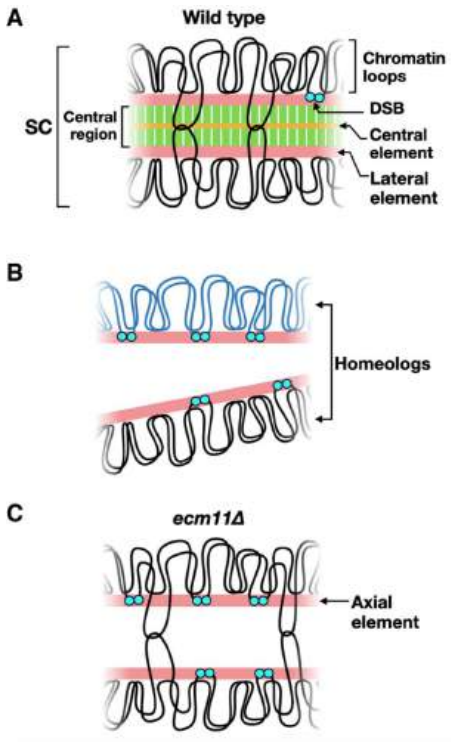


Figure 24.3.5: The relationship of DSB formation to synapsis in different situations. Hollingsworth, N.M. (2020) Genes and Dev 34:1562-1564.

Panel (*A*) shows the synaptonemal complex. When chromosomes form synapses, recombination intermediates contain double Holliday junctions (shown by intersecting loops). When cells exit the ***pachytene stage,*** the stage of meiotic prophase when chromosomes fully synapse, Holliday junctions are resolved to form crossovers, and the SC is disassembled. DSB formation is greatly reduced by synapsis but is not completely abolished until cells exit ***pachynema***.

Panel (*B*) shows the sequence diversity between homologous chromosomes largely inhibits recombination and synapsis, resulting in persistent DSB formation.

Panel (*C*) shows that in the absence of the central element, the transverse filament is not assembled, resulting in chromosomes that lack the central region. Recombination intermediates containing double Holliday junctions are still formed, but DSB formation is not down-regulated. Image from:

SC formation begins with the condensation of sister chromatids along meiosis-specific protein cores to make axial elements. Axial elements from homologous chromosomes are “zippered” together by the insertion of the central region. (Note that after synapsis, axial elements are called lateral elements (Panel A above). The central region is comprised of (1) transverse filaments located perpendicular to the lateral elements, and (2) the central element, which runs parallel to the lateral elements midway through the central region.

Assembly of the SC is initiated in the early stage of meiotic prophase I, which is commonly divided into five substages (leptotene, zygotene, pachytene, diplotene, and diakinesis). For proper assembly of the SC followed by the correct pairing of the homologous chromosome, lateral elements (LEs), which are composed of two main proteins (SYCP2 and SYCP3) should be formed along each chromosome at the initial stage, during leptotene. Later the two LEs associate with the linker part, known as the transverse filaments (TFs). TFs are primarily composed of the protein SYCP1. The central element (CE), which is composed of SYCE1, SYCE2, SYCE3, and TEX12, then connects to the LEs through the TFs, as shown in Figure 24.3.6

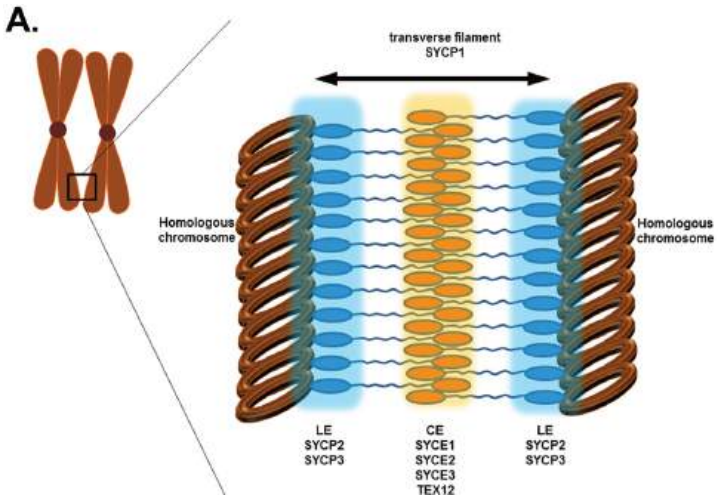


Figure 24.3.6: Organization of the Synaptonemal Complex. Image from: Seo, E.K., et al. (2016) PLOS one: DOI 10.1371

The lateral elements complete their pairing during the zygotene stage leading to the formation of the tripartite SC structure seen during the pachytene stage of the first meiotic prophase as shown in Figure 24.3.7 and Figure 24.3.8. This occurs in both males and females during gametogenesis.

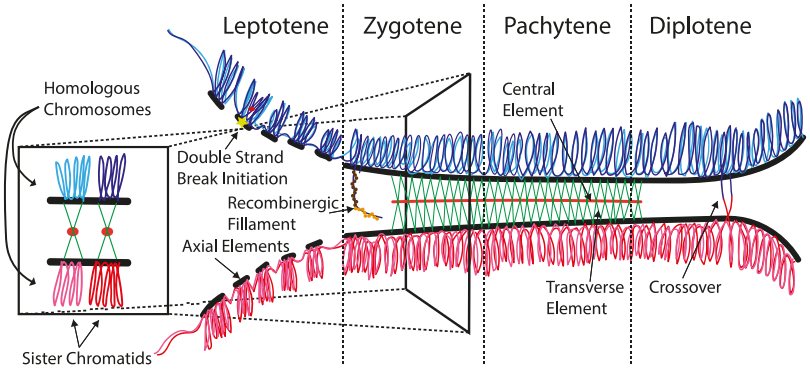


Figure 24.3.7: Schematic of the Synaptonemal Complex at Different Stages During Prophase I. Image from: Daniel Wells

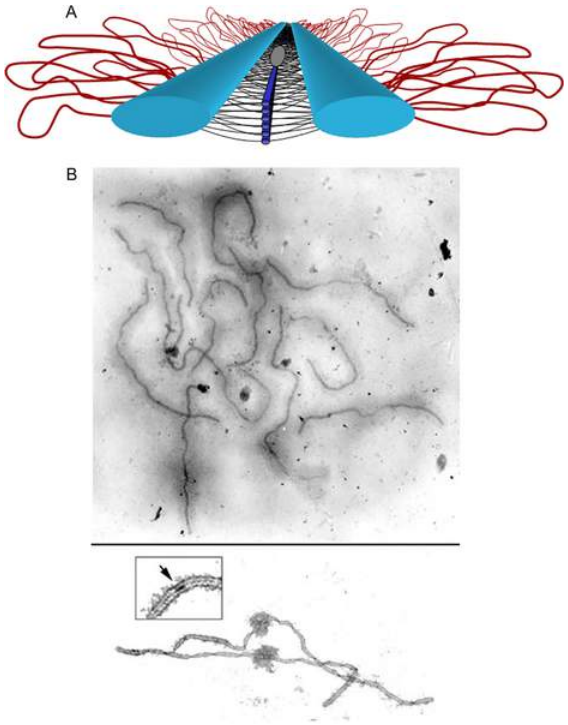


Figure 24.3.8: **The Synaptonemal Complex.** The Public Library of Science

Panel (A) shows a model of the SC. Lateral elements (light blue rods) of homologous chromosomes align and synapse together via a meshwork of transverse filaments (black lines) and longitudinal filaments (dark blue rods). The longitudinal filaments are collectively referred to as the “central element” of the SC. Ellipsoidal structures called recombination nodules (gray ellipsoid) are constructed on the central region of the SC. As their name implies, recombination nodules are believed to be involved in facilitating meiotic recombination (crossing over). The chromatin (red loops) of each homolog is attached to its corresponding lateral element. Because there are two “sister chromatids” in each homolog, two loops are shown extending laterally from each point along a lateral element.

Panel (B) Top shows a set of tomato SCs. Chromatin “sheaths” are visible around each SC showing two tomato SCs. The chromatin has been stripped from the SCs, allowing the details of the SC to be observed. Each SC has a kinetochore (“ball-like” structure) at its centromere. Recombination nodules, ellipsoidal structures found on the central regions of SCs, mark the sites of crossover events (see inset).

Zygotene is the sub-stage where synapsis between homologous chromosomes begins. It is also known as zygonema. This synapse can form up and down the chromosomes allowing numerous points of contact called 'synaptonemal complex', this can be compared to a zipper structure, due to the coils of chromatin. The SC facilitates synapsis by holding the aligned chromosomes together. After the homologous pairs synapse they are either called tetrads or bivalents. Bivalent is more commonly used at an advanced level as it is a better choice due to similar names for similar states (a single homolog is a 'univalent', and three homologs are a 'trivalent').

Once the synapse is formed it is called a bivalent (where a chromatid of one pair is synapsed/attached to the chromatid in a homologous chromosome and crossing over can occur. Subsequently, the synapses snap completing the crossing over of the genetic information. As a result, the variation in genetic material has increased significantly, because up and down the chromosome there has been an exchange of the mother and father's genetic material. The two sister chromatids separate from each other, but the homologous chromosomes remain attached. This makes the complex look much thicker. The SC is complete, allowing chiasma to form. This is what allows the crossing over alleles to occur as this is a process that only happens over a small region of the chromosomes.

The chiasma is a structure that forms between a pair of homologous chromosomes by crossover recombination and physically links the homologous chromosomes during meiosis as shown in Figure 24.3.9. Chiasmata are essential for the attachment of the homologous chromosomes to opposite spindle poles (bipolar attachment) and their subsequent segregation to the opposite poles during meiosis I.

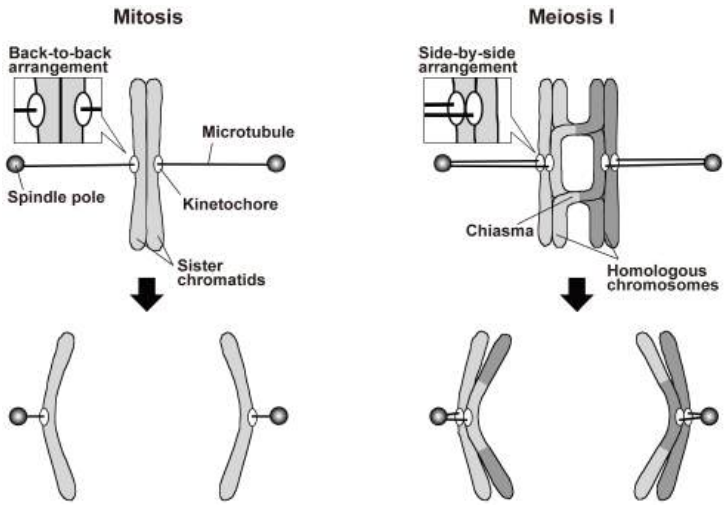


Figure 24.3.9: Chiasmata Promote Monopolar Attachment of Sister Chromatids and Their Co-Segregations Towards the Proper Pole during Meiosis I. Image from Hirose, Y., et al (2011) PLoS Genet. 7(3):e1001329

24.3.3: Mechanism of Homologous Recombination

Meiotic recombination is a tightly regulated process that is triggered by the programmed induction of DNA double-strand breaks (DSBs). Once formed, the ends of the DSBs are nucleolytically processed to generate 3′ single-stranded DNA (ssDNA) tails. Meiotic recombination factors then engage these ssDNA tails to form a nucleoprotein ensemble capable of locating DNA homology in the chromosome homolog and mediating invasion of the homolog to form a DNA joint called a displacement loop or D-loop. The 3′ end of the invading strand is extended by DNA synthesis, followed by the pairing of the non-invading 3′ single-stranded tail with the displaced ssDNA strand in the enlarged D-loop (second-end capture). After DNA synthesis and DNA ligation, a double Holliday Junction (dHJ) intermediate is formed. Resolution of the dHJ intermediate can result in crossover recombinants that harbor a reciprocal exchange of the arms of the homologous chromosomes.

Genetic studies have revealed that meiotic DSBs arise via the action of a protein ensemble that harbors the Spo11 protein, which bears homologous to archaeal Topo VIA, the catalytic subunit of a type II topoisomerase. Indeed, studies in *S. cerevisiae*, *S. pombe*, and *M. musculus* have shown that Spo11 becomes covalently conjugated to the 5′ ends of DNA through a tyrosine residue proposed to be the catalytic center of topoisomerase function. Thus, mutations in the putative catalytic tyrosine residue of Spo11 engender the same phenotype as *spo11* deletion in *S. cerevisiae*, *S. pombe*, *A. thaliana*, and *M. musculus*. All these observations suggest that Spo11 is directly involved in catalyzing DSB formation to trigger meiotic recombination. Figure 24.3.10provides an overview of this process.

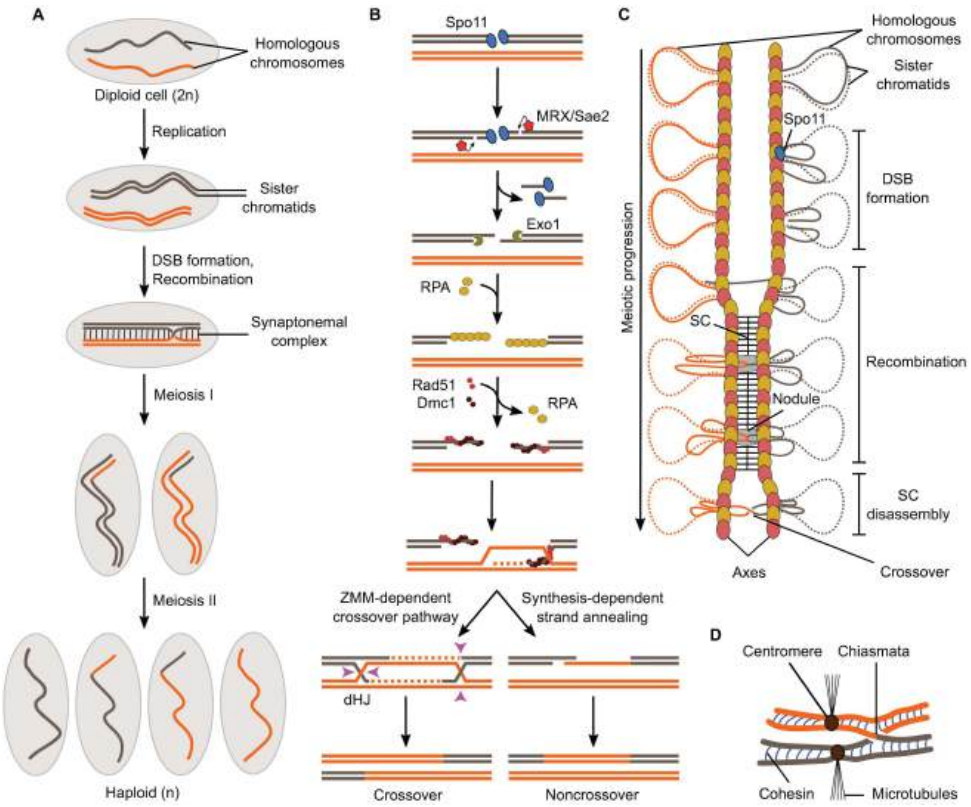


Figure 24.3.10: Overview of meiosis and meiotic recombination.Yadav, V.K. and Bouuaert, C.C. (2021) Front. Cell Dev. Biol 9:642737

Panel (A) shows a schematic of the formation of haploid gametes from a diploid cell with a single pair of homologous chromosomes. DSB formation and recombination promote homolog pairing and lead to the exchange of chromosomal fragments (crossovers) in the context of synapsed chromosomes.

Panel (B) shows meiotic recombination is initiated by Spo11-mediated DSB formation and leads to the formation of crossovers via a ZMM-dependent double Holliday Junction (dHJ) resolution pathway or non-crossovers by synthesis-dependent strand annealing.

Panel (C) shows the relationships between meiotic recombination and higher-order chromosome structure. DSB formation happens in the context of the loop-axis structure. As recombination progresses, the SC polymerizes between the axes and is disassembled before chromosome segregation. Axis proteins Red1 (red ovals) and Hop1 (yellow ovals) are shown.

Panel (D)shows that in metaphase I, homologs are held together through chiasmata and sister chromatid cohesion. Image from:

Following break formation, Spo11 remains covalently attached to the 5′-strands at both DNA ends and is released by an endonucleolytic cleavage reaction mediated by MRX (Mre11, Rad50, and Xrs2) and Sae2, which liberates Spo11 attached to a short oligonucleotide (Fig. 25.3.7B). The 5′-strands are further resected by 5′-3′ exonucleases to produce long single-stranded tails, which are coated with the ssDNA-binding protein, RPA. RPA is then replaced by recombinases Rad51 and Dmc1 that form a nucleoprotein filament and search for sequence similarity preferentially located on the homologous chromosome, producing D-loop structures. Following DNA synthesis using the homolog as a repair template, the recombination structures experience one of two main outcomes (Fig. 25.3.7B). The invading strand can be ejected from the donor by the action of helicases, which provides an opportunity for the DNA ends to re-anneal. This process is referred to as synthesis-dependent strand annealing (SDSA) and produces non-crossovers, that is, products not associated with reciprocal exchanges of chromosome fragments, but with the local transfer of genetic information from the repair template to the broken molecule (gene conversion). Alternatively, recombination structures are stabilized by the “ZMM” family of proteins and channeled through a pathway that produces mostly crossovers. Here, both ends of the break engage the donor to form a double Holliday Junction intermediate, which is resolved through a crossover-specific pathway that involves MutLγ and Exo1.

Every aspect of meiotic recombination is tied to the structural organization of the chromosomes (Fig. 25.3.7C). Early in the meiotic prophase, chromosomes organize as a series of DNA loops that are anchored along a nucleoprotein axis. DSB formation happens in the context of this loop-axis structure. As recombination progresses, polymerization of a proteinaceous structure called the

synaptonemal complex (SC) initiates between the two axes and elongates along their entire length. Recombination proceeds within the SC, inside a nodule embedded between the axes. After recombination is completed, the SC disassembles and crossovers, now cytologically visible as chiasmata, provide physical connections between the homologs until their segregation at anaphase (Fig. 25.3.7D).

24.3.4: References

1. Hollingsworth, N.M (2020) A new role for the synaptonemal complex in the regulation of meiotic recombination. *Genes and Dev*. 34: 1562-1564. Available at: http://genesdev.cshlp.org/content/34/23-24/1562.full

2. Seo, E.K., Choi, J.Y., Jeong, J-H., Kim Y-G, Park, H.H. (2016) Crystal structure of C-terminal coiled-coil domain of XYCP1 reveals the non-canonical anti-parallel dimeric structure of transverse filament at the synaptonemal complex. *PLOS one:* DOI 10.1371. Available at: https://www.researchgate.net/publication/306394048_Crystal_Structure_of_C-Terminal_Coiled-Coil_Domain_of_SYCP1_Reveals_Non-Canonical_Anti-Parallel_Dimeric_Structure_of_Transverse_Filament_at_the_Synaptonemal_Complex

3. Wikipedia contributors. (2021, April 18). Synaptonemal complex. In *Wikipedia, The Free Encyclopedia*. Retrieved 02:07, August 7, 2021, from https://en.Wikipedia.org/w/index.php?title=Synaptonemal_complex&oldid=1018542057

4. Wikipedia contributors. (2021, July 21). Homologous recombination. In *Wikipedia, The Free Encyclopedia*. Retrieved 02:17, August 7, 2021, from https://en.Wikipedia.org/w/index.php?title=Homologous_recombination&oldid=1034703880

5. School of Biomedical Wiki (Accessed Aug 2021) Meiosis Prophase I. Available at: https://teaching.ncl.ac.uk/bms/wiki/index.php/Meiosis_prophase_1

6. Hirose, Y., Suzuki, R., Ohba, T., Hinohara, Y., Matsuhara, H., Yoshida, M., Itabashi, Y., Murakami, H., and Yamamoto, A. (2011) Chaismata Promote Monopolar Attachment of Sister Chromatids and Their Co-Segregation Toward the Proper Pole during Meiosis I. *PLoS Genet.* 7(3):e1001329. Available at: https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1001329

7. Yeh, H-Y., Lin, S-W., Wu, Y-C., Chan, N-L., and Chi P. (2017) Functional Characterization of the Meiosis-Specific DNA Double-Strand Break Inducing Factor, SPO-11 from *C. elegans*. *Scientific Reports* 7:2370. Available at: https://www.nature.com/articles/s41598-017-02641-z#rightslink

8. Yadav, V.K., and Bouuaert, C.C. (2021) Mechanism and Control of Meiotic DNA Double-Strand Break Formation in S. cerevisiae. Front. Cell Dev. Biol. 9:642737. Available at: https://www.frontiersin.org/articles/10.3389/fcell.2021.642737/full

---

CHAPTER OVERVIEW

25: RNA Metabolism

25.1: DNA-Dependent Synthesis of RNA
25.2: RNA Processing
25.3: RNA-Dependent Synthesis of RNA and DNA
25.4: 26.4 References

25.1: DNA-Dependent Synthesis of RNA

25.1.1: Types of RNA

Structurally speaking, ***ribonucleic acid (RNA)***, is quite similar to DNA. However, whereas DNA molecules are typically long and double-stranded, RNA molecules are much shorter and are typically single-stranded. A ribonucleotide within the RNA chain contains ribose (the pentose sugar), one of the four nitrogenous bases (A, U, G, and C), and a phosphate group. The subtle structural difference between the sugars gives DNA added stability, making DNA more suitable for the storage of genetic information, whereas the relative instability of RNA makes it more suitable for its more short-term functions. The RNA-specific pyrimidine ***uracil*** forms a complementary base pair with adenine and is used instead of the thymine that is found in DNA. Even though RNA is single-stranded, most types of RNA molecules show extensive intramolecular base pairing between complementary sequences within the RNA strand, creating a predictable three-dimensional structure essential for their function, as shown in Figure 25.1.1and Figure 25.1.2

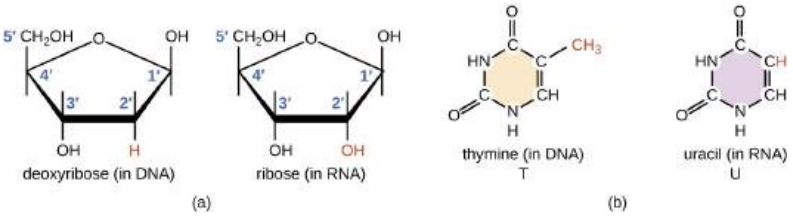


Figure 25.1.1: **RNA Structural Elements** (a) Ribonucleotides contain the pentose sugar ribose instead of the deoxyribose found in deoxyribonucleotides. (b) RNA contains the pyrimidine uracil in place of thymine found in DNA. Figure from: Parker, et al (2019) Microbiology from Openstax

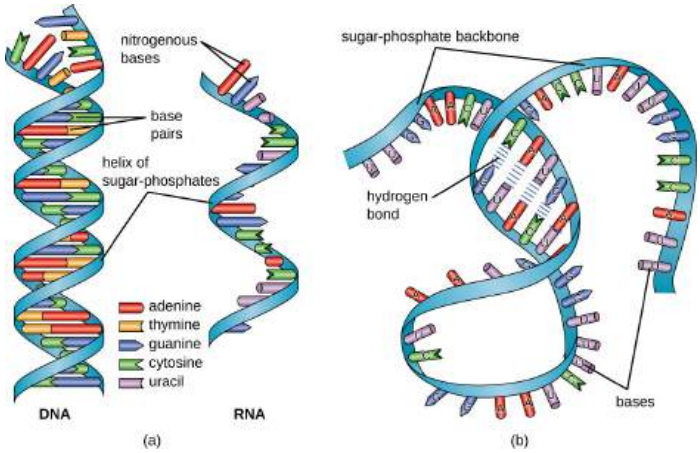


Figure 25.1.2: Parker, et al (2019) Microbiology from Openstax

RNA can largely be divided into two types, one that carries the code for making proteins or coding RNA, which is also called messenger RNA (mRNA), and non-coding RNA (ncRNA). The ncRNA can be subdivided into several different types, depending either on the length of the RNA or on the function. Size classification begins with ***the short ncRNAs*** (~20–30 nt), which include microRNAs (miRs), and small interfering (siRNAs); ***the small ncRNAs*** up to 200 nt, which include transfer RNA (tRNA), small nuclear RNA (snRNA), and small nucleolar RNA (snoRNA); and ***long ncRNAs*** (> 200 nt), which include ribosomal RNA (rRNA), enhancer RNA (eRNA) and long intergenic ncRNAs (lincRNAs), among others.

Cells access the information stored in DNA by creating RNA, through the process of ***transcription***, which then directs the synthesis of proteins through the process of ***translation***. The three main types of RNA directly involved in protein synthesis are messenger RNA (mRNA), ribosomal RNA (rRNA), and transfer RNA (tRNA). The mRNA carries the message from the DNA, which controls all of the cellular activities in a cell. If a cell requires a certain protein to be synthesized, the gene for this product is “turned on” and the mRNA is synthesized through the process of ***transcription***. The mRNA then interacts with ribosomes and other cellular machinery to direct the synthesis of the protein it encodes during the process of ***translation***. mRNA is relatively unstable and short-lived in the cell, especially in prokaryotic cells, ensuring that proteins are only made when needed.

rRNA and tRNA are stable types of RNA. In prokaryotes and eukaryotes, tRNA and rRNA are encoded by the DNA, where they are transcribed into long RNA molecules that are subsequently cut to release smaller fragments containing the individual mature

RNA species. In eukaryotes, synthesis, cutting, and assembly of rRNA into ribosomes takes place in the nucleolus region of the nucleus, but these activities occur in the cytoplasm of prokaryotes. Within the nucleolus region, ribosome assembly requires the activity of numerous snoRNAs.

Ribosomes are composed of rRNA and protein. As its name suggests, rRNA is a major constituent of ribosomes, composing up to about 60% of the ribosome by mass and providing the location where the mRNA binds. The rRNA ensures the proper alignment of the mRNA, tRNA, and the ribosomes; the rRNA of the ribosome also has an enzymatic activity (peptidyl transferase) and catalyzes the formation of the peptide bonds between two aligned amino acids during protein synthesis (Figure 26.1.3). Although rRNA had long been thought to serve primarily a structural role, its catalytic role within the ribosome was shown in 2000. Scientists in the laboratories of Thomas **Steitz** (1940–) and Peter **Moore** (1939–) at Yale University were able to crystallize the ribosome structure from *Haloarcula marismortui*, a halophilic archaeon isolated from the Dead Sea. Because of the importance of this work, Steitz shared the 2009 Nobel Prize in Chemistry with other scientists who made significant contributions to the understanding of ribosome structure. The structure and function of ribosomes will be discussed in further detail in Chapter 27.

Transfer RNA (tRNA) is the third prominent type of RNA involved in protein translation. tRNAs are usually only 70–90 nucleotides long. They carry the correct amino acid to the site of protein synthesis in the ribosome. It is the base pairing between the tRNA and mRNA that allows for the correct amino acid to be inserted in the polypeptide chain being synthesized, as shown in Figure 25.1.3 Any mutations in the tRNA or rRNA can result in global problems for the cell because both are necessary for proper protein synthesis.

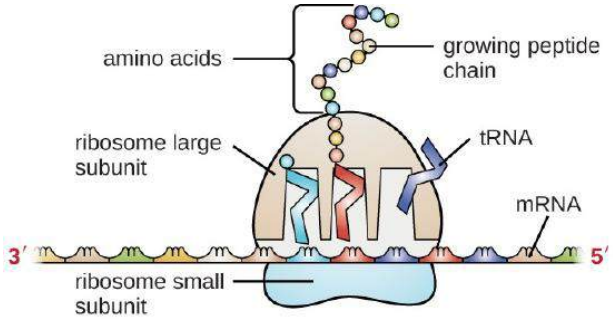


Figure 25.1.3: A generalized illustration of how mRNA and tRNA are used in protein synthesis within a cell. Figure from: Parker, et al (2019) Microbiology from Openstax

As described previously, some RNA molecules have enzymatic properties and serve as ribozymes. Within this chapter, the activity of snRNAs during the process of intron removal from mRNA sequences function as ribozymes and will be described. Furthermore, a detailed description of the enzymatic features of the ribosome structure will be provided in Chapter 27.

Other small ncRNAs and lncRNA molecules play a role in the regulation of transcriptional and translational processes. For example, the post-transcriptional expression levels of many genes can be controlled by RNA interference, in which miRNAs, specific short RNA molecules, pair with mRNA regions and target them for degradation, as shown in Figure 25.1.4 This process is aided by protein chaperones called argonautes. This antisense-based process involves steps that first process the miRNA so that it can base-pair with a region of its target mRNAs. Once the base pairing occurs, other proteins direct the mRNA to be destroyed by nucleases. Fire and Mello were awarded the 2006 Nobel Prize in Physiology or Medicine for this discovery.

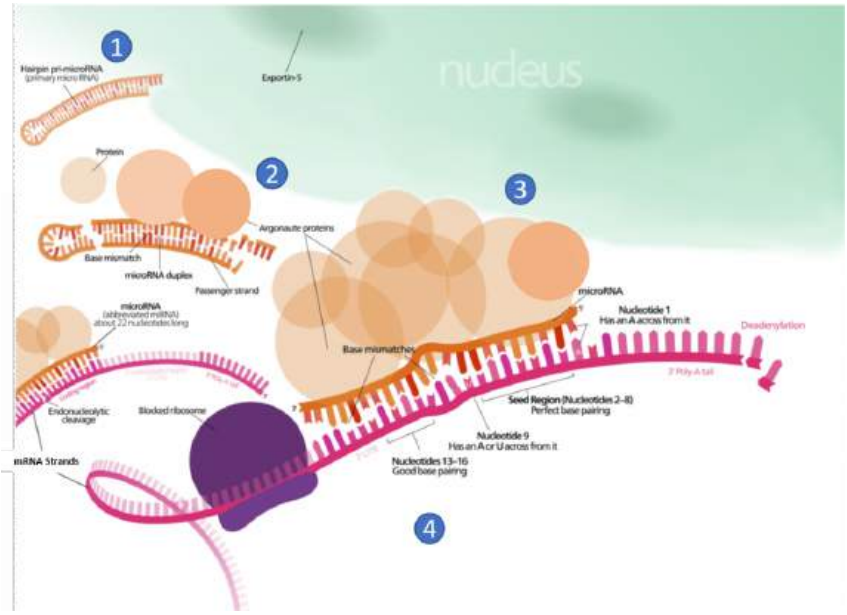


Figure 25.1.4: Role of Micro RNA (miRNA) in the Inhibition of Eukaryotic mRNA Translation. (1) A protein called Exportin-5 transports a hairpin primary micro RNA (pri-miRNA) out of the nucleus and into the cytoplasm. (2) An enzyme called Dicer (not shown), trims the pri-miRNA and removes the hairpin loop. A group of proteins, known as Argonautes, form a miRNA/protein complex. (3) miRNA/protein complex hydrogen bonds with mRNA based on complimentary sequence homology, and blocks translation. (4,5) The miRNA/protein complex binding speeds up the breakdown of the polyA tail of the mRNA, causing the mRNA to be degraded sooner. Figure modified from: Wikimedia Commons

At steady state, the vast majority of human cellular RNA consists of rRNA (~90% of total RNA for most cells) as shown in Figure 25.1.5. Although there is less tRNA by mass, their small size results in their molar level being higher than rRNA (Figure 26.1.5). Other abundant RNAs, such as mRNA, snRNA, and snoRNAs are present in aggregate at levels that are about 1–2 orders of magnitude lower than rRNA and tRNA (Figure 26.1.5). Certain small RNAs, such as miRNA and piRNAs can be present at very high levels; however, this appears to be cell type dependent. lncRNAs are present at levels that are two orders of magnitude less than total mRNA. Although the estimated number of different types of human lncRNAs may have a very restricted expression pattern and thus, accumulate to higher levels within specific cell types. For example, sequencing of mammalian transcriptomes has revealed more than 100,000 different lncRNA molecules can be produced, compared with the approximate 20,000 protein-coding genes. The diversity and functions of the transcriptome within biological processes are currently a highly active area of research.

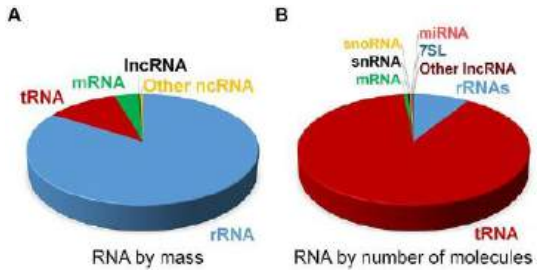


Figure 25.1.5: Estimate of RNA levels in a typical mammalian cell. The proportion of the various classes of RNA in mammalian somatic cells by total mass (A) and by the absolute number of molecules (B). The total number of RNA molecules is estimated at roughly 10^7 per cell. Other ncRNAs in (A) include snRNA, snoRNA, and miRNA. Note that due to their relatively large sizes, rRNA, mRNA, and lncRNA make up a larger proportion of the mass as compared to the overall number of molecules. Figure from: Palazzo, A. and Lee, E.S. (2015) Frontiers in Genetics 6:2

25.1.2: RNA Polymerases

This chapter will focus on the synthesis of RNA by DNA-dependent RNA Polymerase Enzymes (RNAPs). These enzymes are required to carry out the process of transcription and are found in all cells ranging from bacteria to humans. All RNAPs are multi-subunit assemblies, with bacteria having five core subunits, $\alpha_2\beta\beta'\omega$, that have **homologs** in archaeal and eukaryotic RNAPs. Bacterial RNAPs are the simplest form of RNA polymerases and provide an excellent system to study how they control transcription.

25.1.2.1: Prokaryotic RNA Polymerase Enzymes

The RNAP catalytic core within bacteria contains five major subunits ($\alpha_2\beta\beta'\omega$) (see Figure \(\PageIndex{7}\)) below. To position this catalytic core onto the correct promoter requires the association of a sixth subunit called the **sigma factor (σ)**. Within bacteria, there are multiple different sigma factors that can associate with the catalytic core of RNAP that help to direct the catalytic core to the correct DNA locations, where RNAP can then initiate transcription. For example, within *E. coli* σ^{70} is the housekeeping sigma factor that is responsible for transcribing most genes in growing cells. It keeps essential genes and pathways operating. Other sigma factors are activated during certain environmental situations, such as σ^{38} which is activated during starvation or when cells reach the stationary phase. When the sigma subunit associates with the RNAP catalytic core, the RNAP has then formed the holoenzyme. When bound to DNA, the holoenzyme conformation of RNAP can initiate transcription.

Transcription takes place in several stages. To start with, the RNA polymerase holoenzyme locates and binds to promoter DNA. At this stage the RNAP holoenzyme is it the closed conformation (RP_c), as shown in Figure 25.1.6. Initial specific binding to the promoter by sigma factors of the holoenzyme sets in motion conformational changes in which the RNAP molecular machine bends and wraps the DNA with mobile regions of RNAP playing key roles, as shown in Figure 25.1.6. Next, RNAP separates the two strands of DNA and exposes a portion of the template strand. At this point, the DNA and the holoenzyme are said to be in an 'open promoter complex' (RP_o), and the section of promoter DNA that is within it is known as a '**transcription bubble**'. Intermediates($I_{1\text{-}3}$) between RP_c and RP_o occur.

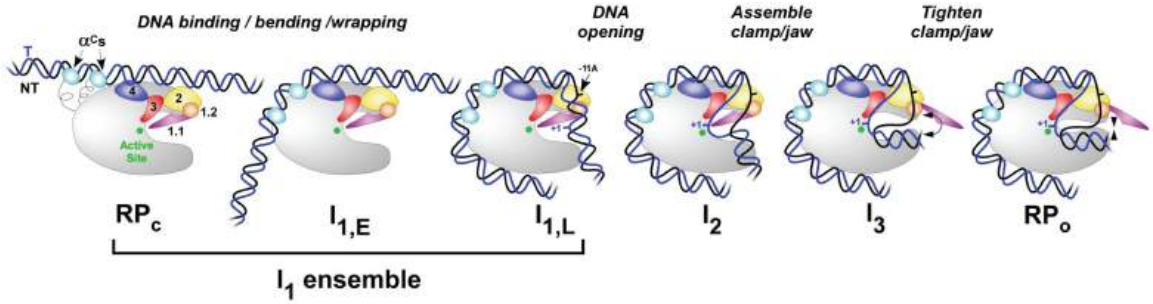


Figure 25.1.6: **Schematic Representation of *E.coli* Transcriptional Initiation.** Closed complexes like RPC, I1,E (early), or I1,L (late) can be significant members of the rapidly equilibrating I1 ensemble. RPo signifies the end of the initiation stage and entry into the elongation phase of RNA synthesis. The α domains are shown in light blue; σ domains are indicated by numbers 1.1, 1.2, 2, 3, and 4. Figure from: Ruff, E.F., et. al. (2015) Biomolecules 5(2):1035-1062.

In bacterial systems, the sigma factor locates the transcriptional start site using key DNA sequence elements located at -35 nucleotides and -10 nucleotides from the transcriptional initiation site, as shown in Panel A of Figure 25.1.7. This region is called the Pribnow box. For RNAP from *Thermus aquaticus*, the −35 element interacts exclusively with σ_4. The duplex DNA just upstream of the −10 element (−17 to −13) interacts with β′, σ_2, and σ_3 (Panel B). Flipping of the A_{-11}(nt) base from the duplex DNA into its recognition pocket in σ_2 is thought to be the key event in the initiation of promoter melting and the formation of the transcription bubble (Panel C). Once the transcription bubble has formed and transcription initiates, the sigma subunits dissociate from the complex and the RNAP catalytic subunit continues elongation on its own.

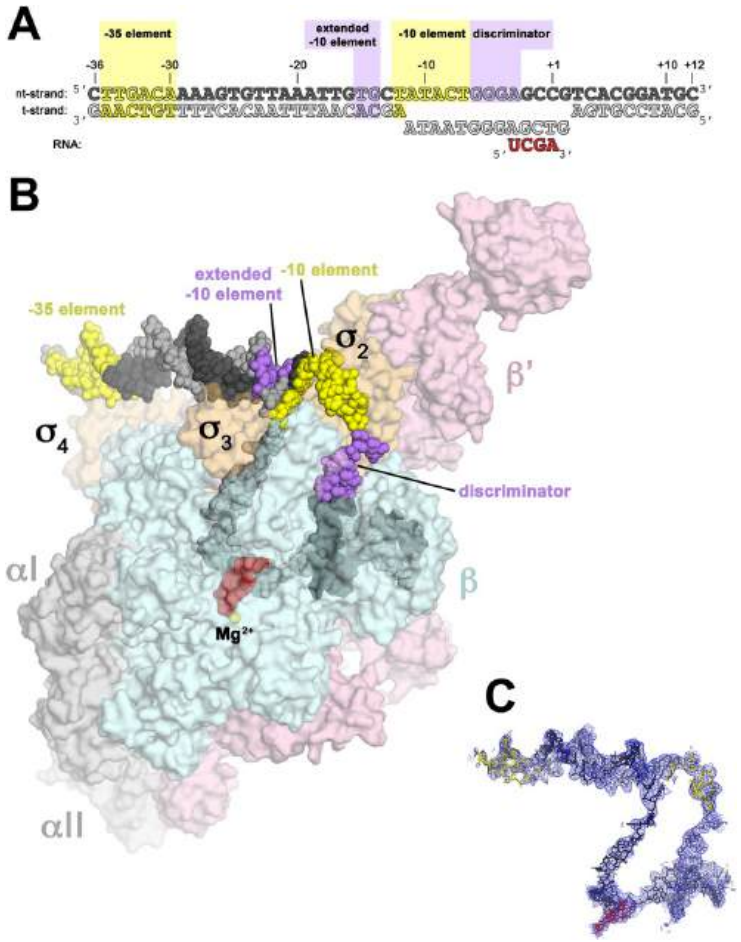


Figure 25.1.7: Structure of the RNAP Holoenzyme in *Thermus aquaticus.* (Figure modified from: Bae, B., et.al. (2015) eLife 4:e08504

Panel (A) shows oligonucleotides used for the crystallization of the RNAP holoenzyme in the open conformation. The numbers above denote the DNA position with respect to the transcription start site (+1). The −35 and −10 (Pribnow box) elements are shaded yellow, and the extended −10 and discriminator elements are purple. The nontemplate-strand DNA (top strand) is colored dark grey; template-strand DNA (bottom strand), light grey; RNA transcript, red.

Panel (**B**) shows the overall structure of RNAP holoenzyme in the open conformation bound with the DNA nucleotides. The nucleic acids are shown as CPK spheres and color-coded as in diagram A. Within RNAP, the αI, αII, ω, are shown in grey; β in light cyan; β′ in light pink; $\Delta 1.1\sigma^A$ in light orange. The *Taq* $E\Delta 1.1\sigma^A$ (Taq derives from *Thermus aquaticus*) is shown as a molecular surface and the forward portion of the RNAP holoenzyme is transparent to reveal the RNAP active site Mg^{2+} (yellow sphere) and the nucleic acids held inside the RNAP active site channel.

Panel (**C**) Electron density and model for RNAP holoenzyme nucleic acids in the open conformation. Color coding matches diagram A.

Figure 25.1.8 shows an interactive iCn3D model of the T. aquaticus transcription initiation complex containing bubble promoter and RNA (4XLN).

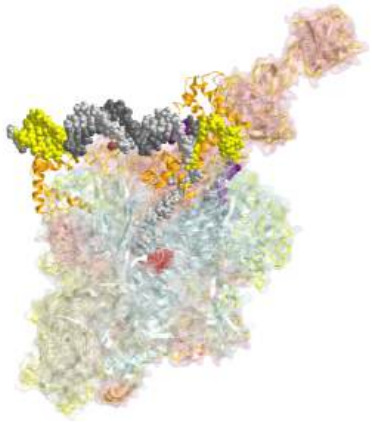


Figure 25.1.8: T. aquaticus transcription initiation complex containing bubble promoter and RNA (4XLN). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...pSXsY6HdMDdELA (long load time but worth the wait)

It is colored coded in fashion similar to that shown in Figure 25.1.7.

The rendering of the DNA is as follows:

- The non-template (nt-strand) DNA is colored dark grey; spheres
- template (t-strand) DNA, light grey; spheres
- Pribnow box yellow
- discriminator purple

The proteins are shown as surfaces with transparent secondary structures underneath. The color is as follows:

- (αI, αII, ω, light yellow;
- β, light cyan;
- β′, light pink;
- Taq EΔ1.1σ^{A}, light orange),

The RNA polymerase active site is located at the Mg^{2+} (black sphere) binding site. The nucleic acids are inside the RNAP active site channel.

25.1.2.2: Eukaryotic RNA Polymerase Enzymes

In eukaryotic cells, three RNAPs (I, II, and III) share the task of transcription, the first step in gene expression. RNA Polymerase I (Pol I) is responsible for the synthesis of the majority of rRNA transcripts, whereas RNA Polymerase III (Pol III) produces short, structured RNAs such as tRNAs and 5S rRNA. RNA Polymerase II (Pol II) produces all mRNAs and most regulatory and untranslated RNAs.

The three eukaryotic RNA polymerases contain homologs to the five core subunits found in prokaryotic RNAPs. In addition, the eukaryotic Pol I, Pol II and Pol III have five additional subunits forming a catalytic core that contains 10 subunits, as shown in Figure 25.1.9. The core has a characteristic crab-claw shape, which encloses a central cleft that harbors the DNA, and has two channels, one for the substrate NTPs and the other for the RNA product. Two ‘pinchers’, called the ‘clamp’ and ‘jaw’ stabilize the DNA at the downstream end and allow the opening and closing of the cleft. For transcription to occur, the enzyme has to maintain a

transcription bubble with separated DNA strands, facilitate the addition of nucleotides, translocate along the template, stabilize the DNA:RNA hybrid, and finally allow the DNA strands to reanneal. This is achieved by a number of conserved elements in the active site, which include the fork loop(s), rudder, wall, trigger loop, and bridge helix.

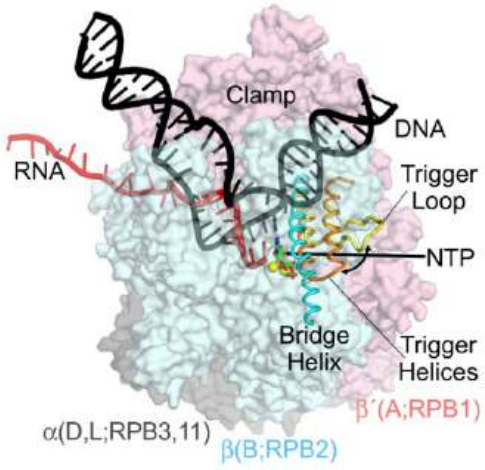


Figure 25.1.9: Structural Diagram of RNA Polymerase from *Thermus thermophilus* (PDB205j) with archaea adn eukaryotic homologs. Hein, P.P. and Landick, R. (2010) BMC Biology 8:141

DNA (black) is melting into a transcription bubble that allows template-strand pairing with RNA (red) in a 9-10 base pair RNA-DNA hybrid. The bridge helix (cyan) and trigger loop/helices (yellow/orange) lie on the downstream side of the active site. The presumed path of the NTP entry is indicated by the straight arrow. Interconversion of the trigger loop and trigger helices is indicated by the curved arrow. The RNA polymerase subunits are shown as semi-transparent surfaces with the identities of orthologous subunits in bacteria (α, β, and β', gray, blue, and pink, respectively), archaea (D, L, B, and A), and eukaryotic RNA polymerase II (RPB3, 11, RPB2, RPB1) indicated. The active site Mg^{2+} ions are shown as yellow spheres, and α,β-methylene-ATP in green and red.

Table 25.1.1 shows RNA polymerase (RNA pol) subunit composition in bacteria, archaea, and, yeast and plants (both eukaryotes).

Bacteria	Archaea	RNA pol I	RNA pol II	RNA pol III	RNA pol IV (plants)	RNA pol V (plants)
β	Rpo1 (RpoA)	RPA190	RPB1	RPC160	NRPD1	NRPE1
β'	Rpo2 (RpoB)	RPBA135	RPB2	RPC128	NRPD/E2	NRPD/E2
α	Rpo3 (RpoD)	RPAC40	RPB3	RPAC40	RPB3 [1]	RPB3 [1]
α	Rpo11 (RpoL)	RPAC19	RPB11	RPAC19	RPB11	RPB11
ω	Rpo6 (RpoK)	RPB6	RPB6	RPB6	RPB6 [1]	RPB6
	Rpo5 (RpoH)	RPB5	RPB5	RPB5	RPB5 [3]	NRPES5
	Rpb8 (RpoG)*	RPB8	RPB8	RPB8	RPB8 [1]	RPB8 [1]
	Rpo10 (RpoN)	RPB10	RPB10	RPB10	RPB10	RPB10
	Rpo12 (RpoP)	RPB12	RPB12	RPB12	RPB12	RPB12
	Rpo4 (RpoF)	RPA14	RPB4	RPC17	NRPD/E4	NRPD/E4
	Rpo7(RpoE)	RPA43	RPB7	RPC25	NRPD7 [1]	NRPE7

Bacteria	Archaea	RNA pol I	RNA pol II	RNA pol III	RNA pol IV (plants)	RNA pol V (plants)
		RPA12	RPB9	RPC11	NRPD9b	RPB9
	Rpo13*					
		RPA49		RPC53		
		RPA34.5		RPC37		
				RPC82		
				RPC34		
				RPC31		

Table 25.1.1: RNA polymerase (RNA pol) subunit composition. Abel, C., Verónica, M., I., G. A. , & Francisco, N. (2017). Subunits Common to RNA Polymerases. In (Ed.), The Yeast Role in Medical Applications. IntechOpen. https://doi.org/10.5772/intechopen.70936. Creative Commons Attribution 3.0 License

Schematic representations of the structure of the eukaryotic RNA pols I, II and III are shown in Figure 25.1.10

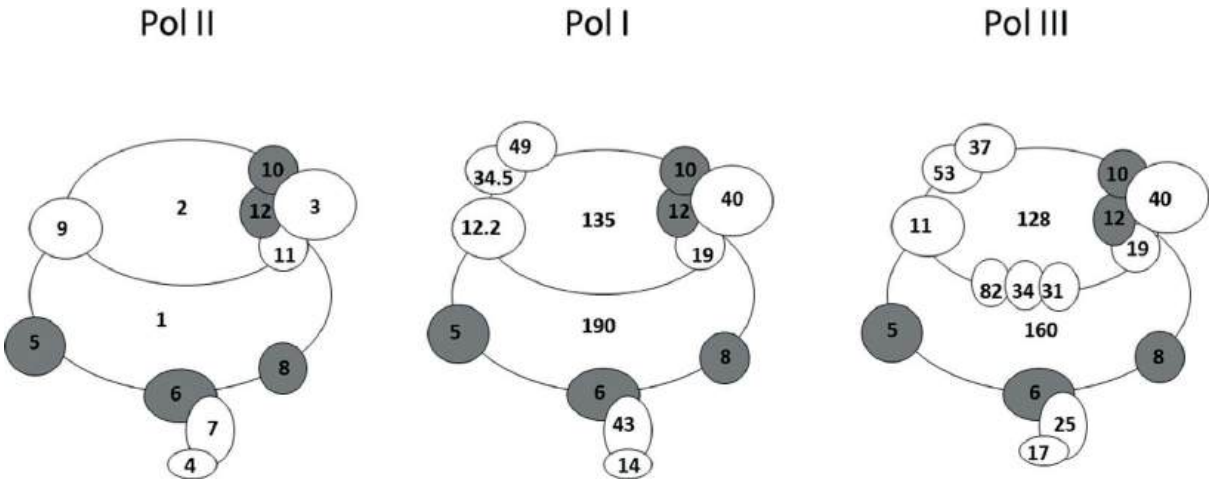


Figure 25.1.10 Schematic representation of the structure of the RNA pols I, II and III. Each RNA pol common subunit is indicated in grey. The numbers correspond to each subunit are indicated in Subunits Common to RNA Polymerases. Abel et al, ibid.

RNA polymerases must bind to DNA, and to host of **transcription factors (TF)** necessary for specific and regulatable transcription. (Note: RNAP is not considered a transcription factor.) The comparative structures of RNAP I-III are shown in Figure 25.1.11 The "stalk" is a structural feature found in eukaryotes but not in prokaryotes. The figure focuses mostly on a comparison of RNAP I and RNAP III.

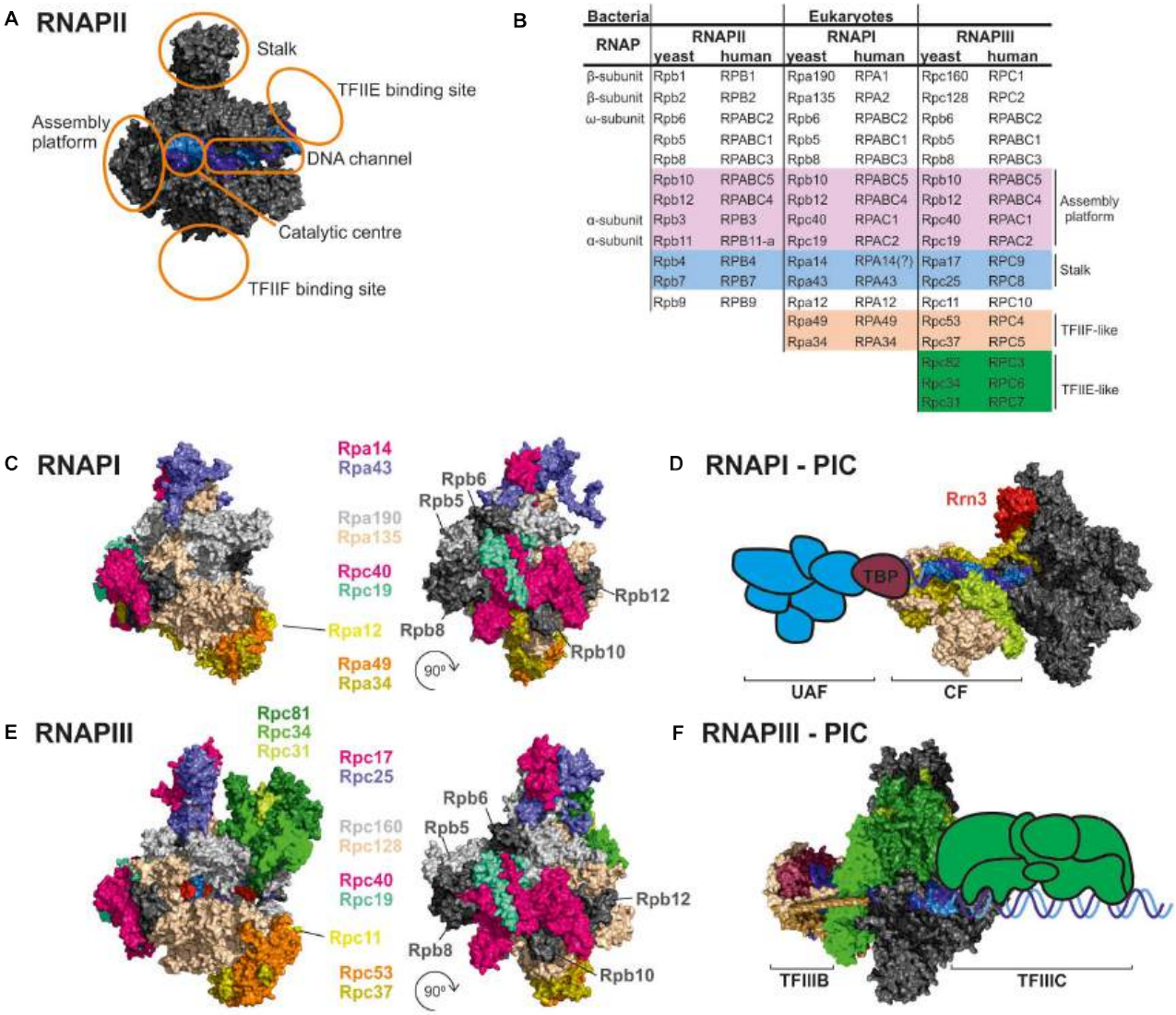


Figure 25.1.11: Comparison of RNAPI, II and III structures and transcription factors. Turowski TW and Boguta M (2021) Specific Features of RNA Polymerases I and III: Structure and Assembly. Front. Mol. Biosci. 8:680090. doi: 10.3389/fmolb.2021.680090. Creative Commons Attribution License (CC BY).

Panel (A) shows the general architecture of RNAPII, consisting of the catalytic core and stalk. RNAPII core consists of a DNA binding channel, catalytic center, and assembly platform. RNAPII binds multiple transcription factors (TFs). Some TFs are homologous to additional subunits of specialized RNAPs (i.e., TFIIF).

Panel (B) shows the subunit composition of eukaryotic RNAPs. Human nomenclature is shown for comparison. Please note that the C-terminal region of Rpa49 subunit harbors a “tandem winged helix” which is predicted in TFIIE and that human RNAPIII RPC7 subunit is coded by two isoforms α and β. The question mark indicates the name is unconfirmed.

Panel (C) shows the subunit composition of yeast RNAPI.

Panel (D) shows a Model of the RNAPI pre-initiation complex, showing an early intermediate with visible Rrn3 and core factor (CF). TATA-binding protein (TBP) and upstream-associated factor (UAF) are added schematically.

Panel (E) shows the subunit composition of yeast RNAPIII.

Panle (F) shows atomic model of RNAPIII pre-initiation complex with TFIIIB. The Rpc82/34/31 heterotrimer is involved in initiation and marked in green as in E. TFIIIC is added schematically. PDB: 5C4X, 5FJ8, 4C3J, 6EU0, and 6TPS

The stalks have two proteins that are not as homologous as the core subunits. These are highlighted in blue in Panel B of Figure 25.1.11 RNAP I and III have additional subunits compared to RNAP II. Overall RNAP I-III have 14, 12, and 17 subunits, respectively. RNAP I and III appear to have integrated transcription-like factors into their core enzyme. In contrast, RNAP II (which transcribes DNA to form messenger RNA), binds to discrete and separate transcription factors to form a **preinitiation complex (PIC)**which we will discuss below.

25.1.3: Transcription Factors and the Preinitiation Complex (PIC)

Unlike prokaryotic systems which can initiate the recruitment of RNAP holoenzymes directly onto the DNA promoter regions and mediate the conversion of RNAP to the open conformation, eukaryotic RNA polymerases require a host of additional general transcription factors (GTFs), to enable this process. Here we will focus on the activation of RNA Polymerase II as an example of the complexity of eukaryotic transcription initiation.

Class II gene transcription in eukaryotes is a tightly regulated, essential process controlled by a highly complex multicomponent machinery. A plethora of proteins, more than a hundred in humans, are organized in very large multiprotein assemblies that include a core of General Transcription Factors (GTFs). The GTFs include the factors TFIIA, TFIIB, TFIID, TFIIE, TFIIF, TFIIH, RNA polymerase (RNA pol II), as well as a large number of diverse complexes that act as co-activators, co-repressors, chromatin modifiers, and remodelers, as shown in Figure 25.1.12 Class II gene transcription is regulated at various levels: while assembling on chromatin, before and during transcription initiation, throughout elongation and mRNA processing, and termination. A host of activators and repressors has been reported to regulate transcription, including a central multisubunit complex called the ***Mediator*** that helps in the recruitment of GTFs and the activation of RNA Pol II. Here we will focus on the formation of the GTFs that make up the core ***preinitiation complex (PIC)*** during transcriptional activation.

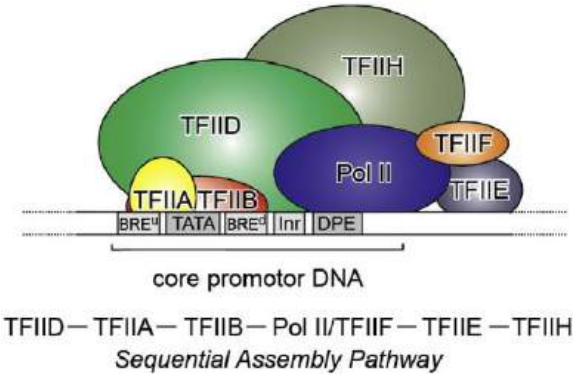


Figure 25.1.12: Transcription Preinitiation Complex (PIC). Gupta, K., et.al. (2016) J Mol Bio 428(12):2581-91.

Class II gene transcription in humans is brought about by over a hundred polypeptides assembling on the core promoter of protein-encoding genes, which then give rise to mRNA. A PIC on a core promoter is shown in a schematic representation. PIC contains, in addition to promoter DNA, the GTFs (TFIIA, B, D, E, F, and H), and RNA Pol II. PIC assembly is thought to occur in a highly regulated, stepwise fashion, as indicated. TFIID is among the first GTFs to bind the core promoter via its TATA-box Binding Protein (TBP) subunit. Nucleosomes at transcription start sites contribute to PIC assembly, mediated by signaling through epigenetic marks on histone tails. The Mediator (not shown) is a further central multiprotein complex identified as a global transcriptional regulator. TATA = TATA-box DNA; BRE^u = B recognition element upstream; BRE^d = B recognition element downstream; Inr = Initiator; DPE = Down-stream promoter element. Figure from:

Transcription of RNA pol II-dependent genes is triggered by the regulated assembly of the **Preinitiation Complex (PIC)**. PIC formation starts with the binding of TFIID to the core promoter. TFIID is a large megadalton-sized multiprotein complex with around 20 subunits made up of 14 different polypeptides: the TATA-box binding protein (TBP) and the **TBP-associated factors (TAFs)** (numbered 1–13), as shown in Figure 25.1.13 Some of the TAF subunits are present in two copies. A key feature in TAFs is the **histone fold domain (HFD),** which is present in 9 out of 13 TAFs in TFIID. The HFD is a strong protein–protein interaction motif that mediates specific dimerization. The HFD-containing TAFs are organized in discrete heterodimers, with the exception of TAF10, which is capable of forming dimers with two different TFIID components, TAF3 and TAF8. HFDs and several other structural features of TBP and the TAFs are well conserved between species.

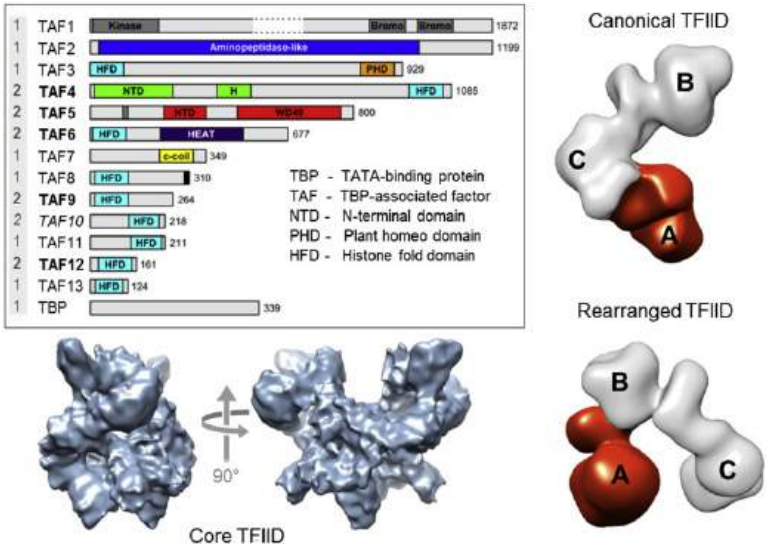


Figure 25.1.12: Human TFIID. Gupta, K., et.al. (2016) J Mol Bio 428(12):2581-91.

TFIID is a large megadalton-sized multiprotein complex comprising about 20 subunits made up of 14 different polypeptides. The constituent proteins of TFIID, TBP and the TAFs, are shown in a schematic representation depicted as bars (inset, left). Structured domains are marked and annotated. The presumed stoichiometry of TAFs and TBP in the TFIID holo-complex is given (far left, gray underlaid). TAF10 (in italics) makes histone fold pair separately with both TAF3 and TAF8. TAFs present in a physiological TFIID core complex extracted from eukaryotic nuclei are labeled in bold. The architecture of the TFIID core complex (EMD-2230) determined by cryo-EM is shown (bottom left) in two views related by a 90° rotation (arrows). The holo–TFIID complex is characterized by remarkable structural plasticity. Two conformations, based on cryo-EM data (EMD-2284 and EMD-2287), are shown on the right, a canonical form (top) and a more recently observed rearranged form (bottom). In the rearranged conformation, lobe A (colored in red) migrates from one extreme end of the TFIID complex (attached to lobe C) all the way to the other extremity (attached to lobe B).

TFIID was shown to adopt an asymmetric, horse-shoe shape with three almost equal-sized lobes (A, B, and C), exhibiting a considerable degree of conformational flexibility with at least two distinct conformations (open and closed), as shown in Figure 25.1.12 The TBP component of TFIID binds with a specific DNA sequence called the TATA box. This DNA sequence is found around 30 base pairs upstream of the transcription start site in many eukaryotic gene promoters. When TBP binds to a TATA box within the DNA, it distorts the DNA by inserting amino acid side chains between base pairs, partially unwinding the helix, and doubly kinking it. The distortion is accomplished through a great amount of surface contact between the protein and DNA. TBP binds with the negatively charged phosphates in the DNA backbone through positively charged lysine and arginine amino acid residues. The sharp bend in the DNA is produced through the projection of four bulky phenylalanine residues into the minor groove. As the DNA bends, its contact with TBP increases, thus enhancing the DNA-protein interaction. The strain imposed on the DNA through this interaction initiates the melting, or separation, of the strands. Because this region of DNA is rich in adenine and thymine residues, which base-pair through only two hydrogen bonds, the DNA strands are more easily separated.

The role of TAFs is complicated. Take for example TAF11 and TAF13. These act as competitive inhibitors of TBP to the TATA (Pribnow) box as well as TAF1 which somewhat mimics the structural features of the Pribnow box. TAF11/TAF13 binds to the DNA surface where TBP binds. suggesting a novel regulation of TFIID. These interactions are illustrated in Figure 25.1.14

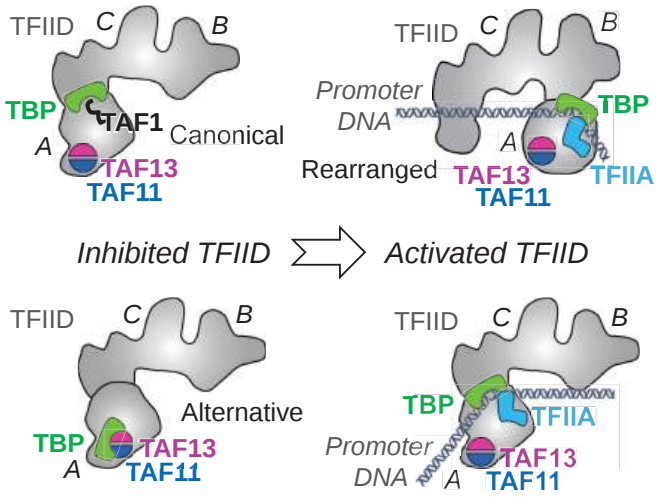


Figure 25.1.14 Novel TFIID regulatory state comprising TAF11/TAF13/TBP. Kapil Gupta,et al. (2017) Architecture of TAF11/TAF13/TBP complex suggests novel regulation properties of general transcription factor TFIID eLife 6:e30395. https://doi.org/10.7554/eLife.30395. Creative Commons Attribution License

Given its role in transcribing the DNA for thousands of messenger RNAs, let's focus on the preinitation complex for RNAP II. Structures are known for the closed and open promoters. A key component is TFH (see Figure 25.1.12). TFIIH opens the DNA for transcription. As if we need to complicate the structure of the preinitiation complex even more, it turns out the TFIIH is not a single dark green cartoon as shown in Figure 12, but a rather large complex itself. Its structure is shown in Figure 25.1.15

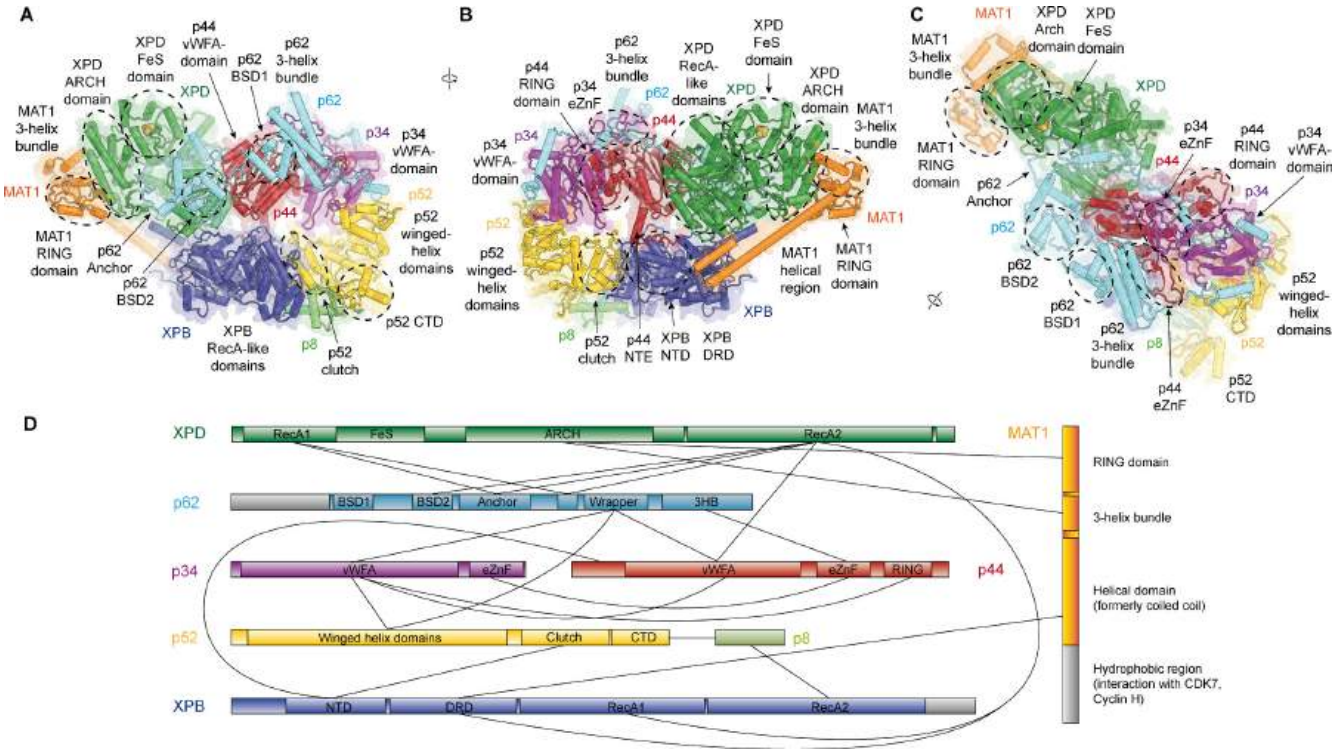


Figure 25.1.15 Structure of the TFIIH core complex. Greber et al. (2019). The complete structure of the human TFIIH core complex. *eLife* **8**:e44771. https://doi.org/10.7554/eLife.44771. Creative Commons Attribution License

Panes (**A, B, C**) shows three views of the structure of the TFIIH core complex and MAT1. Subunits are color-coded and labeled (in color); individual domains are labeled (in black) and circled if needed for clarity.

Panel (**D**) shows the domain-level protein-protein interaction network between the components of the TFIIH core complex and MAT1 derived from the interactions observed in our structure. Proteins are shown with the same colors as in A and major unmodeled regions are shown in grey. Abbreviations: CTD: C-terminal domain; DRD: DNA damage recognition domain; FeS: iron sulfur cluster domain; NTD: N-terminal domain; vWFA: von Willebrand Factor A.

The largest subunits are DNA-dependent helicases/translocases/ATPases XPB and XPD, which are bridged by MAT1. It appears that XPB starts and propagates a twist in the DNA which propagates ot open the DNA 30 BP downstream of the TATA box. This mostly likely is followed by the dissociation of TFIIH and a stoppage in DNA twisting, which allows RNA transcription to start.

Figure 25.1.16 shows an interactive iCn3D model of the human/mouse/mastadenovirus C RNA polymerase II core pre-initiation complex with open promoter DNA (7NVU).

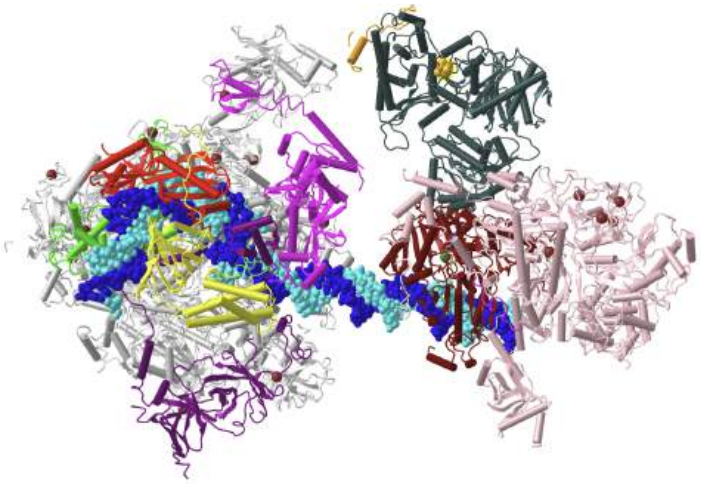


Figure 25.1.16: Human/mouse/a mastadenovirus C RNA polymerase II core pre-initiation complex with open promoter DNA (7NVU). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...Bc8itsYA7qctWA (long load time but worth the wait)

The following color schemes are used:

- MAT1 - orange
- TF2A - yellow
- TF2B - green
- TF2E - magenta
- TF2F- purple
- ATP-dependent translocase (helicase) subunit XPD - dark slate gray
- TBP (TATA box binding protein) - red
- TF2H - Pink
- 2H-XPB helicase - maroon
- NT-DNA - cyan
- template (T)-DNA - blue (Note: part of it is shown in a yellow cartoon in the interactive model)
- the SF_4 cofactor is shown in spacefill with CPK colors

In summary, the binding of TFIID to the core promoter is followed by the recruitment of further GTFs and RNA pol II. Several lines of evidence suggest that this process occurs in a defined, stepwise order and undergoes significant restructuring. First, PIC adopts an inactive state, the “closed” complex, which is incompetent to initiate transcription. In addition to TFIID, TFIIH is also critical for the shift of RNA Pol II from the closed to the open conformation. TFIIH has an ATP-dependent translocase activity within one of its subunits, that opens up about 11 to 15 base pairs around the transcription start site by moving along one DNA strand inducing torsional strain, leading to conformational rearrangements and the positioning of single-stranded DNA to the active site of RNA pol II. In this “open” complex, RNA pol II can enter elongation to transcribe throughout a gene in a highly processive manner without dissociating from the DNA template or losing the nascent RNA.

In most eukaryotes, after synthesizing about 20–100 bases, RNA pol II can pause (***Promoter proximal pause***) and then disconnect from promoter elements and other components of the transcription machinery, giving rise to a fully functional elongation complex

in a process called ***promoter escape***. The promoter-bound components of the PIC, in contrast, remain in place, and thus only TFIIB, TFIIF, and RNA pol II need to be recruited for re-initiation, significantly increasing the transcription rate in subsequent rounds of transcription. ***Promoter escape*** is preceded by an ***abortive transcription*** in many systems, where multiple short RNA products of 3 to 10 bases in length are synthesized.

In addition to promoter elements within the DNA, ***enhancer elements*** are also important for the initiation of transcription. ***Promoters*** are defined as DNA elements that recruit transcription complexes for the synthesis of coding and non-coding RNA. ***Enhancers*** are defined as DNA elements that positively regulate transcription at promoters over long distances in a position- and orientation-independent manner. However, studies have revealed that many enhancers can recruit Pol II and initiate transcription of enhancer RNA (eRNA), thus blurring the functional distinction between enhancers and promoters (Figure 10.13).

Enhancer transcription produces relatively short ncRNA. Furthermore, transcription at enhancers is unstable and often leads to the termination of elongation. In contrast, transcription initiation at most Pol II promoters is stable and produces long mRNAs. Topological studies revealed that enhancers come in close proximity to target gene promoters during transcription activation. According to current gene activation models, the ***Mediator complex*** forms a physical bridge between distant regulatory regions and promoters, thereby promoting looping. Transcription of at least a subset of genes regulated by enhancers occurs in bursts indicating a discontinuous process of transcription complex recruitment, assembly, and/or conversion to elongation-competent forms. The bursting phenomenon suggests that enhancer/promoter contacts may be transient and infrequent, as shown in Figure 25.1.17.

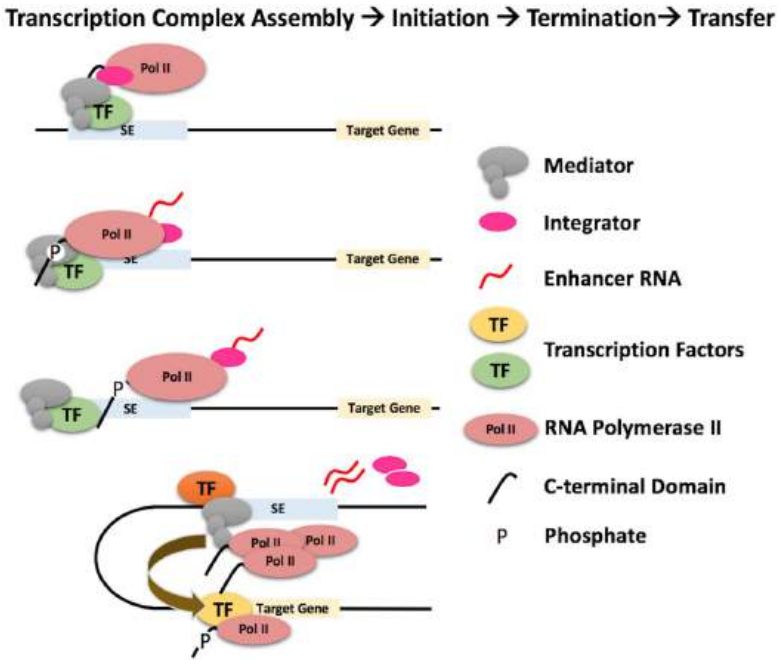


Figure 25.1.17: RNA polymerase II Enhancer Transfer Model. Gurumurthy, A., et.al. (2018) BioEssays 1800164

Depicted are the steps involved in the recruitment of Pol II to SEs, assembly into elongation-competent transcription complexes, transcription initiation, and elongation, abortion and termination, and transfer to target genes. Transcription factors recruit Mediator and other co-regulators to SEs. Mediator recruits Pol II and assembles a fraction into elongation competent transcription complexes. Transcription is initiated by phosphorylation of the CTD. Early abortion and transcription termination conferred by Integrator releases Pol II, which is dephosphorylated and transferred to target gene promoters. Super Enhancer Element (SE).

25.1.4: Transcriptional Elongation and Termination

25.1.4.1: Prokaryotic Transcriptional Elongation

The rate of transcription elongation by *E. coli* RNAP is not uniform. RNA synthesis is characterized by pauses, some of which may be brief and resolved spontaneously, whereas others may lead to the ***transcription elongation complex (TEC)*** backtracking.

Elongation rate and pausing are determined by template sequence and RNA structure (e.g., stem-loops) and involve at least two components of the RNAP catalytic center, the ***bridge helix (BH)*** and ***trigger loop (TL).*** Elongation is proposed to occur in three

steps, as shown in Figure 25.1.18 First, the TL folds in response to NTP binding. Mutational analyses indicate that this conformational change in the TL can be rate-limiting, and reflects the ability of the incoming NTP to bind to TEC. The second step is the incorporation of the NTP and the release of pyrophosphate. The third step involves the translocation of the RNAP down the DNA Template such that the next RNA nucleotide can be added to the nascent transcript.

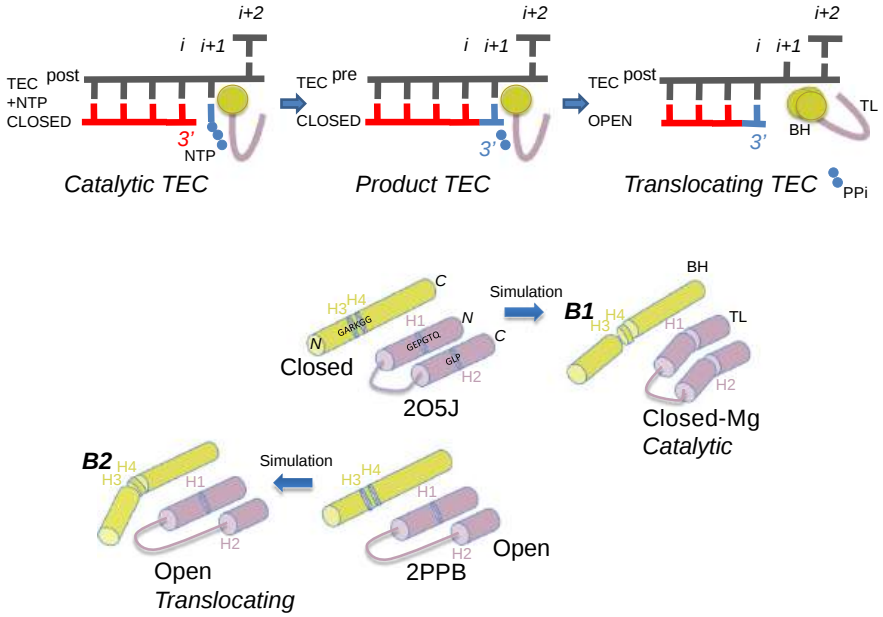


Figure 25.1.18: A model for Transcriptional Elongation. Kireeva, M., et. al. (2012) BMC Biophysics 5(1):11

The trigger loop hinges, bridge helix hinges, and bridge helix bending models are based on molecular dynamics simulations. At the top of the figure, diagrams of the closed TEC, the closed product TEC (after chemistry), and the translocating TEC are shown. DNA is grey; RNA is red; the NTP substrate (or incorporated NMP and pyrophosphate) is blue; the trigger loop (TL) is purple; the bridge helix (BH) is yellow. Interpretations of simulations are shown schematically below. Simulations indicate trigger loop hinges H1 and H2, bridge helix hinges H3 and H4 and bridge helix bend modes B1 (straighter) and B2 (more sharply bent).

Backtracking of TEC may take place after a brief pause in transcription, caused by the thermodynamic properties of nucleic acids sequences surrounding the elongation complex. In addition, misincorporation events render elongation complexes prone to backtracking by at least one bp. In this case, the rescue from backtracking through the cleavage of the 3' end of the erroneous transcript also may be seen as a proofreading reaction. Any backtracking event causes a pause or arrest of transcription elongation, which may limit its overall rate (the average speed of RNAP along the template) or the processivity (the fraction of RNAP molecules reaching the end of the gene).

While the general structure of the elongation complex (the transcription bubble, the RNA-DNA hybrid) remains unchanged during backtracking, the extension of RNA becomes impossible in this conformation. However, such complexes can be resolved by the hydrolytic activity of RNAP, which cleaves the phosphodiester bond in the active center of the backtracked complex, producing a new RNA 3' end in the active center. For single base backups, the hydrolytic reaction is catalyzed by a flexible domain of RNAP located in the secondary channel called the Trigger Loop (TL) and the two metal ions of the active center.

Longer sequences of backtracked TEC can restart when acted upon by GreA/B factors, which restore the 3'-end of the nascent transcript to the active center. GreA and GreB are transcript cleavage factors that act on backtracked elongation complexes. When Gre factors are bound in the secondary channel, Gre factors displace the TL from the active center, as shown in Figure 25.1.19 The displacement switches off the relatively slow TL-dependent intrinsic transcript hydrolysis, and imposes the highly efficient Gre-assisted hydrolysis. This efficiency is thought to be due to the stabilization of the second catalytic Mg^{2+} ion and an attacking water molecule by the Gre factors.

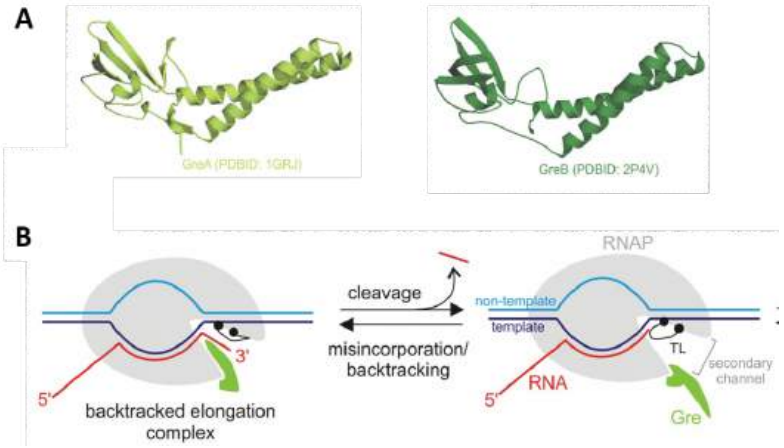


Figure 25.1.19: The Role of Gre Factors in Relieving Transcription Elongation Complex Backtracking. Zenkin, N., and Yuzenkova, Y. (2015) Biomolecules 5(3):1195-1209.

Panel (**A**) shows a ribbon diagram of the GreA and GreB proteins.

Panel (**B**) shows the mode of functioning of Gre factors. The Gre factor is bound to the active elongation complex but does not impose hydrolytic activity on it. Upon backtracking or misincorporation, the Gre factor protrudes its coiled-coil domain through the secondary channel of RNAP (shown in the lefthand diagram), where it substitutes for the catalytic domain Trigger Loop (TL). This substitution switches off the slow TL-dependent phosphodiester bond hydrolysis and, and instead, facilitates highly efficient Gre-dependent hydrolysis. After the resolution of the backtracked complex through RNA cleavage, the elongation complex returns to the active conformation, and the Gre factor gives way to the TL, which can now continue the catalysis of RNA synthesis (shown in the right-hand diagram). The controlled switching between Gre and the TL eliminates possible interference of Gre with the RNA synthesis.

25.1.4.2: Prokaryotic Transcriptional Termination

Transcription termination determines the ends of transcriptional units by disassembling the transcription elongation complex (TEC), thereby releasing RNA polymerases and nascent transcripts from DNA templates. Failure in termination causes transcription readthrough, which yields wasteful and possibly harmful intergenic transcripts. It can also perturb the expression of downstream genes when the unterminated TEC sweeps transcription initiation complexes off their promoters or collides with RNA polymerases that transcribe opposite strands.

Transcriptional termination in prokaryotes can be template-encoded and factor-independent ***(intrinsic termination)***, or require accessory factors, such as **Rho**, Mfd, and DksA. ***Intrinsic termination*** occurs at specific template sequences - an inverted repeat followed by a run of A residues. Termination is driven by the formation of a short stem-loop structure in the nascent RNA chain, as shown in Figure 25.1.20 RNA synthesis arrests and TEC dissociates at the 7th and 8th U of the run. Formation of the stem-loop dissociates the weak rU:dA hybrid. Stem-loop formation is hindered by upstream complementary RNA sequences that compete with the downstream portion of the stem, as well as by RNA: protein interactions in the RNA exit channel. Intrinsic termination depends critically upon timing. Hairpin folding and transcription of the termination point must be coordinated, so that the complete hairpin is formed by the time RNAP transcribes the termination point. The size of the stem, the sequence of the stem, and the length of the loop all affect termination efficiency.

The bridge α-helix in the β' subunit borders the active site and may have roles in both catalysis and translocation. Mutations in the YFI motif (β' 772-YFI-774) affect intrinsic termination as well as pausing, fidelity, and translocation of RNAP. One mutation, F773V, abolishes the activity of the λ tR2 intrinsic terminator, although neighboring mutations have little effect on termination. Modeling suggests that this unique phenotype reflects the ability of F773 to interact with the fork domain in the β subunit.

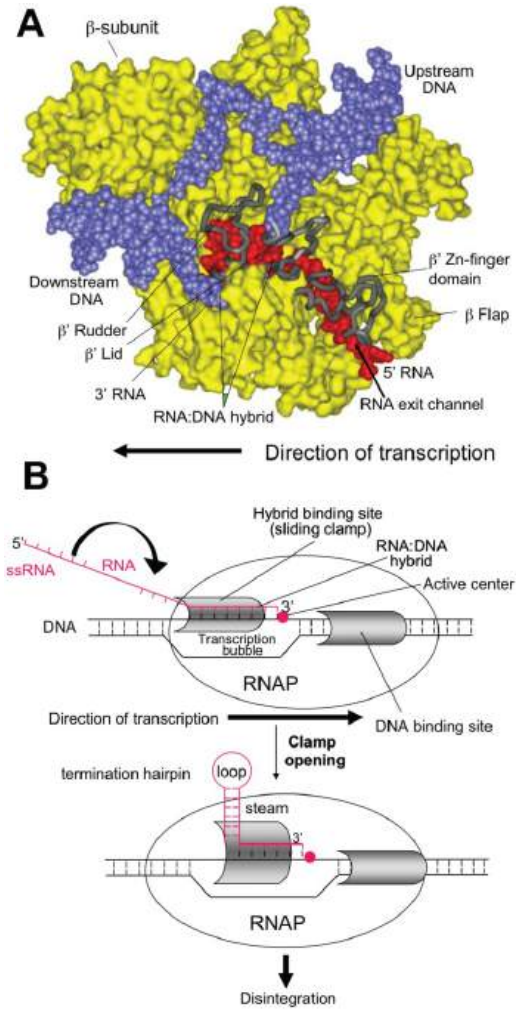


Figure 25.1.20: Model of Intrinsic Termination. Washburn, R.S., and Gottesman, M.E. (2015) Biomolecules 5(2)1063-1078.

Panel (A) shows the open conformation of the RNAP during transcriptional elongation. RNAP is shown in yellow, the DNA template in blue, and the nascent RNA in red. Key elements of the RNAP RNA exit channel are shown in grey and labeled as indicated.

Panel (B) shows the extension of the nascent RNA through the RNAP exit channel and the potential for forming the RNA hairpin structure when enough length has been achieved.

Panel (C) shows the clamp opening and disintegration of the TEC when the RNA hairpin structure is encountered at the transcriptional bubble.

Figure 25.1.21 shows an interactive iCn3D model of the T. thermophilus RNAP polymerase elongation complex with the NTP substrate analog (2O5J). (long load time)

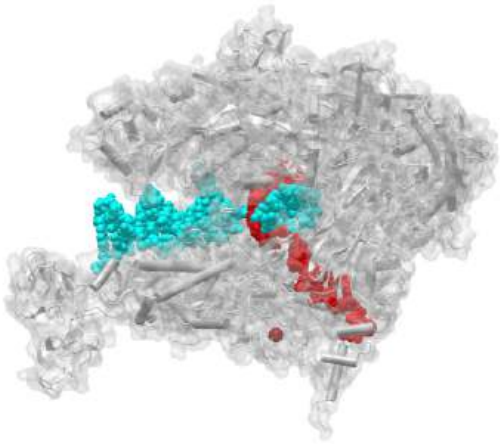


Figure 25.1.21: T. thermophilus RNAP polymerase elongation complex with the NTP substrate analog (2O5J). (Copyright; author via source). Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...g4RAeYA3FiGDLA (long load time)

Transcriptional termination can also be dependent upon accessory factors, such as the **Rho** protein. Transcription termination factor Rho is an essential protein in *E. coli* first identified for its role in transcription termination at Rho-dependent terminators, and is estimated to terminate ~20% of *E. coli* transcripts. The *rho* gene is highly conserved and nearly ubiquitous in bacteria. Rho is an RNA-dependent ATPase with RNA:DNA helicase activity, and consists of a hexamer of six identical monomers arranged in an open circle, as shown in Figure 25.1.22

Rho binds to single-stranded RNA in a complex multi-step pathway that involves two distinct sites on the hexamer. The primary binding site (PBS), distributed on the N-terminal domains around the hexamer (cyan), ensures initial anchoring of Rho to the transcript at a Rut (Rho utilization) site, a∼70 nucleotides (nt) long, cytidine-rich and poorly-structured RNA sequence. Each Rho monomer contains a subsite capable of binding specifically the base residues of a 5′-YC dimer (Y being a pyrimidine). Biochemical and structural data suggest that Rho initially binds to RNA in an open, 'lock-washer' conformation that closes into a planar ring as RNA transfers to the central cavity. There, the ssRNA contacts an asymmetric secondary binding site (SBS) (green), and this step, which presumably is rate-limiting for the overall reaction, leads to motor activation. Upon hydrolysis of ATP, the ssRNA is pulled upon conformational changes of the conserved Q and R loops of the SBS, leading to Rho translocation, and ultimately promoting RNA polymerase (RNAP) dissociation. The molecular mechanism of Rho translocation based on single-molecule fluorescence methods appears to be tethered tracking. The tethered tracking model postulates that Rho maintains its contacts between the PBS and the loading (Rut) site upon translocation (Panel B). This mechanism would allow Rho to maintain its high-affinity interaction with Rut, and implies the growth of an RNA loop between the PBS and the SBS upon translocation.

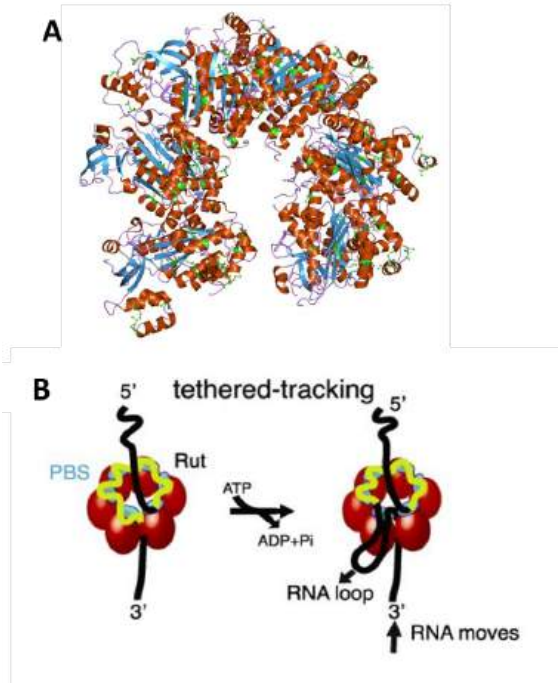


Figure 25.1.22: Schematic of Rho factor structure and mechanisms. (A) Jawahar Swaminathan and (B) Gocheva, V., et. al. (2015) Nuc Acids Res 43(1)10.1093.

Panel (A) shows the molecular structure of the Rho protein (PDB 1pv4)

Panel (B)shows how Rho assembles as a homo-hexameric ring (red spheres or tetragons), with RNA (black/yellow curve) binding to the primary binding sites (PBS, cyan) and the secondary binding sites inside the ring (SBS, green), where ATP-coupled translocation takes place. The Rut-specific binding site is depicted in yellow. The tethered-tracking model proposed that Rho translocates RNA while maintaining interactions between PBS and Rut. This model requires the formation of a loop that would shorten the extension of RNA upon translocation. Figure modified from:

Figure 25.1.23 shows an interactive iCn3D model of the E. Coli Rho transcription termination factor in complex with ssRNA substrate and ANPPNP (1PVO).

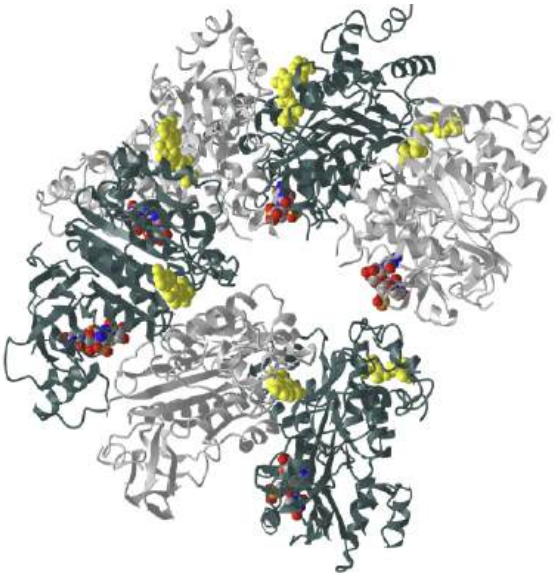


Figure 25.1.23: E. Coli Rho transcription termination factor in complex with ssRNA substrate and ANPPNP (1PVO). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...boj4YBbvevcsEA

The six subunits of the hexamer are shown in alternating slate gray and light gray. The di-ribonucleotides (5'-R(P*UP*C)-3') are shown in spacefill and colored CPK. ANPPNP is shown in spacefill yellow.

Figure 25.1.24shows an interactive iCn3D model of the closed ring structure of the E. Coli Rho transcription termination factor in a complex with nucleic acid in the motor domains (2HT1).

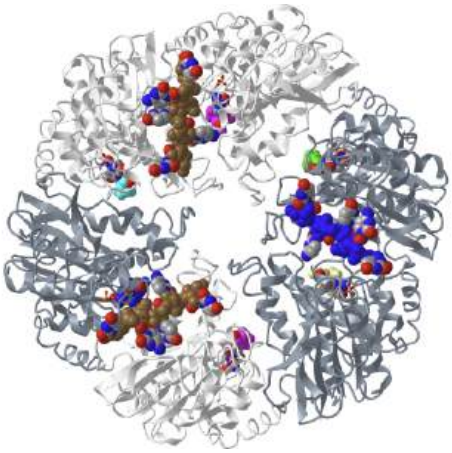


Figure 25.1.24: Closed ring structure of the E. Coli Rho transcription termination factor in complex with nucleic acid in the motor domains (2HT1). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...GVkVhPknbmiYj7

The six subunits of the hexamer are again shown in alternating slate gray and light gray. The ssRNAs are shown in spacefill with the backbones in one color and the bases in CPK colors.

Figure 25.1.25 shows an interactive iCn3D model of the E. coli Rho-dependent Transcription Pre-termination Complex (6XAS) (long load).

Figure 25.1.25: E. coli Rho-dependent Transcription Pre-termination Complex (6XAS). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...kUxfqN2TFPcSL9 (long load)

- The RNA polymerase subunits ($\alpha_2\beta\beta\omega$) are shown in light cyan
- The six subunits of the Rho hexamer are again shown in alternating slate gray and light gray
- NusA is shown in magenta
- Both DNA strands are shown in spacefill with the backbone blue and the bases red
- Only part of the RNA is shown (spacefill, backbone yellow, bases CPK), so it appears discontinuous

25.1.4.3:

25.1.4.4: Eukaryotic Transcriptional Termination

In eukaryotes, termination of protein-coding gene transcription by RNA polymerase II (Pol II) usually requires a functional polyadenylation (pA) signal, typically a variation of the AAUAAA hexamer. Nascent pre-mRNA is cleaved and the 5′ fragment is

polyadenylated at the pA site shortly downstream from the hexamer by cleavage and pA factors (CPFs). Two mechanisms have been suggested for pA-dependent transcription termination. In ***the allosteric model***, the pA signal and/or other termination signals bind with the pA signal downstream region (PDR) and induce reorganization of the Pol II complex. This includes the association or dissociation of endonuclease components such as the CPFs. This causes conformational changes in Pol II and TEC disassembly ensues. In ***the kinetic model,*** also known as ***the "torpedo" model***, cleavage at the pA site separates the pre-mRNA from the TEC, which continues synthesizing a downstream nascent transcript. This new transcript is a substrate of XRN2/Rat1p, a processive 5′-to-3′ exoribonuclease that catches up with, and disassembles, the TEC by an unknown mechanism.

The two pA-dependent models are not mutually exclusive, and unified models have been proposed. Loosely conserved pA signal sequences downstream of protein-coding genes bind to components of the polyadenylation factor (CF1) complex leading to the assembly of the cleavage and polyadenylation machinery. Termination is coupled to cleavage in a manner that has not yet been completely resolved, however, one of the major factors involved in yeast pA termination is the endonuclease, Ysh1. For example, the depletion of Ysh1 blocks TEC dissociation, but does not cause substantial readthrough at the termination site (Fig. 26.1.18 A&B). These results suggest that Ysh1 does not directly cause the pausing that occurs in the allosteric termination pathway, but rather plays a role in the dissociation of the Pol II complex from the DNA template, as shown in Figure 25.1.26 It should be noted that not all pA-dependent termination is dependent on Ysh1 and that other mechanisms of pA-mediated termination still remain to be elucidated.

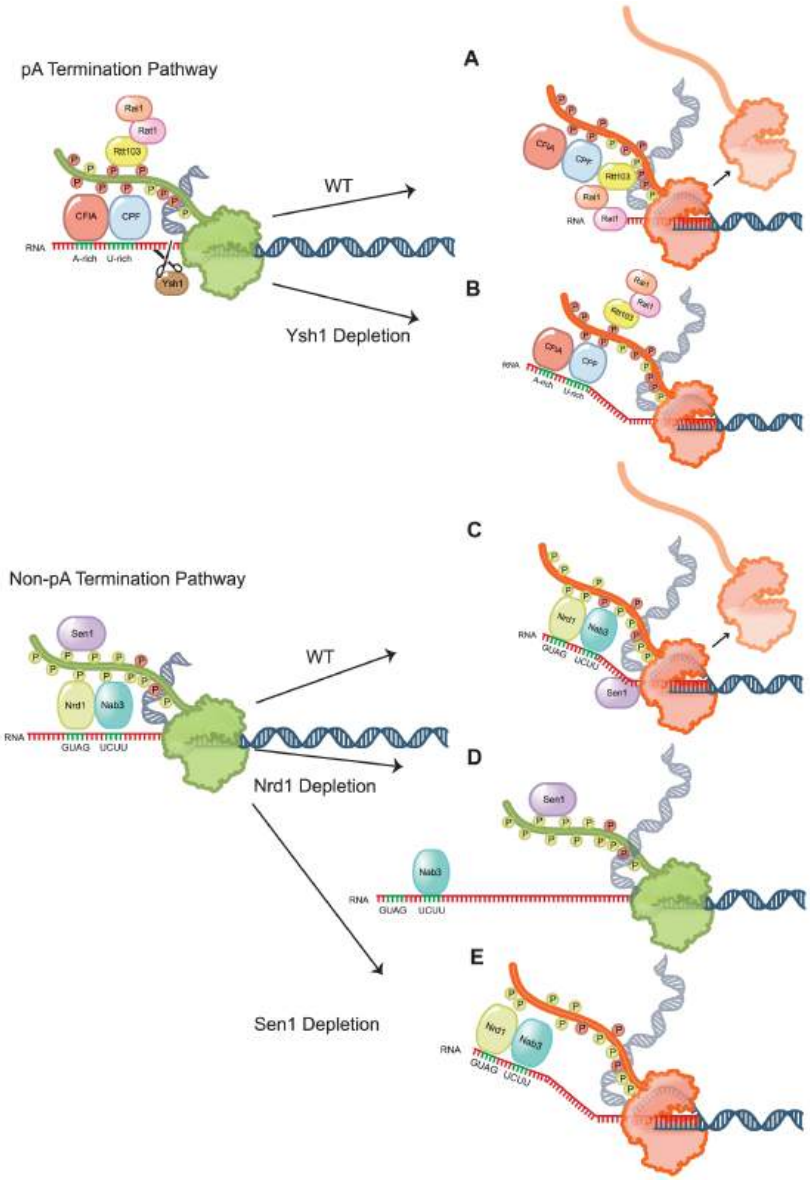


Figure 25.1.26: Schematic representation of Pol II termination after removal of non-pA and pA termination factors. *Schaughency, P., Merran, J., and Corden J.L. (2014) PLOS Genetics 10(10):e1004632*

Panel (A) shows the elongating Pol II (green) terminates pA transcripts (A) after an allosteric change (red) that reduces processivity.

Panel (B) shows the depletion of Ysh1 leads to minimally extended readthrough transcripts but does not block the allosteric change in Pol II.

Panel (C) shows how Nrd1 and Nab3 binding recruits Sen1 for termination of non-pA transcripts.

Panel (D) shows Pol II elongation complex lacking Nrd1 does not recognize termination sequences in the nascent transcript and thus does not facilitate the allosteric transition in Pol II. This leads to a processive readthrough.

Panel(E) shows how Nrd1 and Nab3 recognize terminator sequences allowing the allosteric change in Pol II but depletion of Sen1 blocks the removal of Pol II from the template. *Figure from:*

The mechanisms of termination of Pol II-mediated transcription differ for coding and non-coding transcripts. Coding transcripts and possibly some stable uncharacterized transcripts (SUTs) are nearly always processed at the 3′-end by the cleavage and polyadenylation (pA) machinery and are processed by the pA-dependent termination mechanisms described above. In contrast, ncRNAs are terminated and processed by an alternative pathway that, in yeast, requires the RNA-binding proteins Nrd1 and Nab3, as well as, the RNA helicase Sen1. Nrd1 and Nab3 recognize RNA sequence elements downstream of snoRNAs and CUTs and this

leads to the association of a complex that contains the DNA/RNA helicase Sen1 and the ***nuclear exosome***. The ***nuclear exosome*** is a complex of ribonucleases with 3' to 5' exonuclease and endonuclease activity. It functions to degrade unstable or incorrect RNA transcripts.

Both Nrd1 and Sen1 depletion lead to readthrough transcription of ncRNAs, suggesting their importance in non-pA-dependent transcription termination (Fig 26.1.18 C & D). Furthermore, depletion of Nrd1 also causes the accumulation of longer readthrough ncRNAs, suggesting its role in trafficking ncRNAs to the nuclear exosome following termination.

1. Parker, N., Schneegurt, M., Thi Tu, A-H., Lister, P., Forster, B.M. (2019) Microbiology. *Openstax*. Available at: https://opentextbc.ca/microbiologyopenstax/
2. Palazzo, A., and Lee, E.S. (2015) Non-coding RNA: what is function and what is junk? *Frontiers in Genetics* 6:2 Available at: file:///C:/Users/flatt/AppData/Local/Temp/fgene-06-00002.pdf
3. Wikipedia contributors. (2020, July 9). RNA. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:30, August 6, 2020, from https://en.Wikipedia.org/w/index.php?title=RNA&oldid=966784317
4. Burenina, O.Y., Oretskaya, T.S., and Kubareva, E.A. (2017) Non-Coding RNAs As Transcriptional Regulators in Eukaryotes. *Acta Naturae* 9(4):13-25. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5762824/
5. Khatter, H., Vorlander, M.K., and Muller C.W. (2017) RNA polymerase I and III: similar yet unique. Current Opinion in Structural Biology 47:88-94. Available at: https://www.sciencedirect.com/science/article/pii/S0959440X17300313
6. Wikipedia contributors. (2020, May 8). Sigma factor. In *Wikipedia, The Free Encyclopedia*. Retrieved 17:50, August 7, 2020, from https://en.Wikipedia.org/w/index.php?title=Sigma_factor&oldid=955570499
7. Bae, B., Felkistov, A., Lass-Napiokowska, A., Landick, R., and Darst, S.A. (2015) Structure of a bacterial RNA polymerase holoenzyme open protomer complex. *eLife* 4:e08504. Available at: https://elifesciences.org/articles/08504
8. Petrenko, N., Jin, Y., Dong, L., Wong, K.H., and Struhl, K. (2019) Requirements for RNA polymerase II preinitiation complex formation in vivo. *eLife* 8:e43654. Available at: https://elifesciences.org/articles/43654
9. Gupta, K., Sari-Ak, D., Haffke, M., Trowitzsch, S., and Berger, I. (2016) Zooming in on transcription preinitiation. *J Mol Biol*. 428(12):2581-2591. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4906157/
10. Wikipedia contributors. (2020, April 17). TATA-binding protein. In *Wikipedia, The Free Encyclopedia*. Retrieved 14:54, August 8, 2020, from https://en.Wikipedia.org/w/index.php?title=TATA-binding_protein&oldid=951583592
11. Patel, A.B., Greber, B.J., and Nogales, E. (2020) Recent insights into the structure of TFIID, its assembly, and its binding to core promoter. *Curr Op Struct Bio* 61:17-24. Available at: https://www.sciencedirect.com/science/article/pii/S0959440X19301113#fig0010
12. Ruff, E.F., Record, Jr., M.T., Artsimovitch, I., (2015) Initial events in bacterial transcription initiation. *Biomolecules* 5(2):1035-1062. Available at: https://www.mdpi.com/2218-273X/5/2/1035/htm
13. Kireeva, M., Opron, K., Seibold, S., Domecq, C., Cukier, R.I., Coulombe, B., Kashlev, M., and Burton, Z. (2102) Molecular dynamics and mutational analysis of the catalytic and translocation cycle of RNA polymerase. *BMC Biophysics* 5(1):11. Available at: https://www.researchgate.net/publication/225281979_Molecular_dynamics_and_mutational_analysis_of_the_catalytic_and_translocation_cycle_of_RNA_polymerase/figures?lo=1
14. Washburn, R.S., and Gottesman, M.E. (2015) Regulation of transcription elongation and termination. *Biomolecules* 5(2):1063-1078. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4496710/pdf/biomolecules-05-01063.pdf
15. Zenkin, N., and Yuzenkova, Y. (2015) New insights into the functions of transcription factors that bind the RNA polymerase secondary channel. Biomolecules 5(3):1195-1209. Available at: https://www.mdpi.com/2218-273X/5/3/1195/htm
16. Gocheva, V., LeGall, A., Boudvillain, M., Margeat, E., and Nollmann, M. (2015) Direct observation of the translocation mechanism of transcription termination factor Rho. *Nuc Acids Res* 43(1):10.1093. Available at: https://www.researchgate.net/publication/272162172_Direct_observation_of_the_translocation_mechanism_of_transcription_termination_factor_Rho
17. Miki, T.S., Carl, S.H., and Groβhans, H. (2017) Two disctinct transcription termination modes dictated by promoters. Genes & Dev 31:1-10. Available at: https://www.researchgate.net/publication/320350041_Two_distinct_transcription_termination_modes_dictated_by_promoters
18. Gurumurthy, A., Shen, Y., Gunn, E.M., Bungert, J. (2018) Phase separation and transcription regulation: Are Super-Enhancers and Locus Control Regions primary sites of transcription complex assembly? *BioEssays* 1800164. Available at: https://www.researchgate.net/publication/329331157_Phase_Separation_and_Transcription_Regulation_Are_Super-Enhancers_and_Locus_Control_Regions_Primary_Sites_of_Transcription_Complex_Assembly

19. Suñé-Pou, M., Prieto-Sánchez, Boyero-Corral, S., Moreno-Castro, C., El Yousfi, Y., Suñé-Negre, J.M., Hernández-Munain, C., and Suñé, C. (2017) Targeting splicing in the treatment of human disease. Genes 8(3):87. Available at: https://www.mdpi.com/2073-4425/8/3/87/htm
20. Schaughency, P., Merran, J., and Corden J.L. (2014) Genome-wide mapping of yeast RNA polymerase II termination. *PLOS Genetics* 10(10):e1004632 Available at: https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1004632
21. Nourse, J., Spada, S., and Danckwardt, S. (2020) Emerging roles of RNA 3'-end cleavage and polyadenylation in pathogenesis, diagnosis, and therapy of human disorders. *Biomolecules* 10(6):915. Available at: https://www.mdpi.com/2218-273X/10/6/915/htm
22. Wikipedia contributors. (2020, July 30). Five-prime cap. In *Wikipedia, The Free Encyclopedia*. Retrieved 05:53, August 11, 2020, from https://en.Wikipedia.org/w/index.php?title=Five-prime_cap&oldid=970240533
23. Cortes, T. and Cox, R.A. (2015) Transcription and translation of the rpsJ, rplN and rRNA operons of the tubercle bacillus. Microbiology (2015) 161:719-728. Available at: https://www.microbiologyresearch.org/docserver/fulltext/micro/161/4/719_mic000037.pdf?expires=1597159574&id=id&accname=guest&checksum=6FFC9C066EF41C7799FAE843CE94C49F
24. Hein, P.P. and Landick, R. (2010) The bridge helix coordinates movements of modules in RNA polymerase. *BMC Biology* 8:141. Available at: https://bmcbiol.biomedcentral.com/articles/10.1186/1741-7007-8-141
25. Gonatopoulos-Pournatzis, T., and Cowling, V.H. (2014) Cap-binding complex (CBC). *Biochem. J.* 457:231-242. Available at: https://www.researchgate.net/publication/259392894_Cap-binding_complex_CBC

25.2: RNA Processing

Post-transcriptional modifications of rRNA and tRNA will be topics of Chapter 27 as their structure and function in protein synthesis will be a focal point. Thus, this section will focus on post-transcriptional modifications of mRNA. We'll spend most of our time on eukaryotic RNA processing.

25.2.1: Prokaryotic RNA Processing

First, let's take a brief look at a fascinating feature of transcription in bacterial cells. Bacterial cells do not have extensive post-transcriptional modifications of mRNA primarily because transcription and translation are coupled processes. Bacterial cells lack the physical barrier of a nucleus, which allows transcription and translation machinery to function at the same time, enabling the concurrent translation of an mRNA while it is being transcribed (Fig 26.2.1). As mRNA is synthesized in prokaryotes, a ribosome binding motif called the Shine-Dalgarno sequence, located in the 5' untranslated region of the mRNA emerges early, allowing the ribosome to bind and translation to occur. In addition, the protein N-utilzation Substance, better known as **NusG,** plays a critical role. NusG has three separate domains and the functions of two of them are known. The NusG N-terminal domain (NusG-NTD) can bind to RNAP, whereas the C-terminal domain (NusG-CTD) can combine with the NusE (RpsJ) component of ribosomes. These two functions of NusG enable transcription to be coupled with translation. NusG CTD can also bind to Rho to terminate transcription, as shown in Figure 25.2.1.

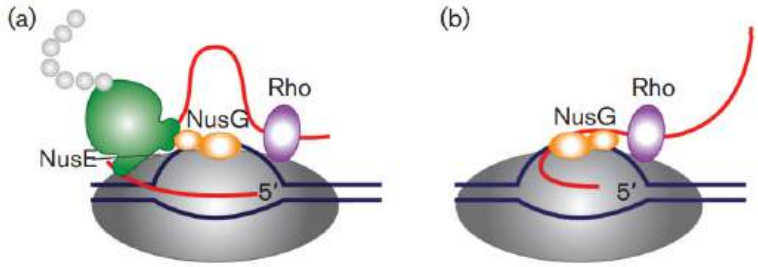


Figure 25.2.1: **The roles of NusG in transcription/translation coupling.** (a) Composition of an active RNAP complex. RNAP is shown in dark grey, DNA in blue and nascent RNA in red. The ribosome is shown in green with the nascent polypeptide chain in light grey; the bulge in the small subunit denotes the location of NusE (RpsJ). NusG is shown in orange: its shape denotes two functional sections. The larger section denotes the N-terminal domain, which binds to RNAP. The smaller section denotes the C-terminal domain, which interacts with NusE in situ. Rho is shown in purple. (b) After the translation is completed, NusG remains bound to RNAP and may also bind to Rho through the C-terminal domain leading to the termination of transcription. *Figure from: Cortes, T., and Cox, R.A. (2015) Microbiology 161:719-728.*

Another protein, NusA, slows RNAP, in contrast to NusG which increases its processivity, as reflected by the length of the RNA made before RNAP falls off.

25.2.2: Eukaryotic RNA Processing

In multicellular organisms, almost every cell contains the same genome, yet complex spatial and temporal diversity is observed in gene transcripts. This is achieved through multiple levels of processing leading from gene to protein, of which RNA processing is an essential stage. Following the transcription of a gene by RNA polymerases to produce a primary mRNA transcript, further processing is required to produce a stable and functional mature RNA product. This involves various processing steps including RNA cleavage at specific sites, intron removal, called ***splicing***, which substantially increases the transcript repertoire, and the addition of a 5'CAP. Another crucial feature of the RNA processing of most genes is the generation of 3′ ends through an initial endonucleolytic cleavage, followed in most cases by the addition of a poly(A) tail, a process termed 3′ end cleavage and polyadenylation (CPA).

25.2.3: Cleavage and 3'-Polyadenylation (CPA)

Polyadenylation is a required step for the correct termination of nearly all mRNA transcripts. Except for replication-dependent histone genes, metazoan protein-encoding mRNAs contain a uniform 3' end consisting of a stretch of adenosines. In addition to determining the correct transcript length at transcription termination, the poly(A) tail helps to ensure the translocation of the

nascent RNA molecule from the nucleus to the cytoplasm, enhances translation efficiency, acts as a signal feature for RNA degradation, and thereby contributes to the production efficiency of a protein.

Cleavage and polyadenylation (CPA) are carried out by the **cleavage/polyadenylation apparatus (CPA)**, a multi-subunit 3′ end processing complex, which involves over 80 proteins, comprised of four core protein subcomplexes, as shown in Figure 25.2.2 These consist of

1. **cleavage and polyadenylation specificity factor (CPSF)**, comprised of proteins CPSF1-4, factor interacting with PAPOLA and CPSF1 (FIP1L1), and WD repeat domain 33 (WDR33) (shown in green below);
2. cleavage stimulation factor (CstF), a trimer of CSTF1-3 (shown in red below;
3. cleavage factor I (CFI), a tetramer of two small nudix hydrolase 21 (NUDT21) subunits, and two large subunits of CPSF7 and/or CPSF6 (shown in orange in Figure 26.2.2 A); Note: Nudix are named for **nu**cleoside **di**phosphates hydroxylases.
4. cleavage factor II (CFII), composed of cleavage factor polyribonucleotide kinase subunit 1 (CLP1) and PCF11 cleavage and polyadenylation factor subunit (PCF11) (shown in yellow in Figure 26.2.2 A). Additional factors include symplekin, the **poly(A) polymerase (PAP)**, and the nuclear poly(A) binding proteins such as poly(A) binding protein nuclear 1 (PABPN1).

CPA is initiated by this complex recognizing specific sequences within the nascent pre-mRNA transcripts termed **polyadenylation signals (PAS)**. The PAS sequence normally consists of either a canonical 6 base sequence, the **AATAAA hexamer,** or a close variant usually differing by a single nucleotide (e.g., ATTAAA, TATAAA). It is located 10 to 35 nucleotides upstream of the cleavage site (CS) usually consisting of a CA dinucleotide. The PAS is also determined by surrounding auxiliary elements, such as upstream U-rich elements (USE), or downstream U-rich and GU-rich elements and G-rich sequences (DSE).

As soon as the nascent RNA molecule emerges from RNA polymerase II (RNA Pol II), the CPSF complex is recruited to the PAS **AATAAA hexamer,** through numerous interactions. Upon successful assembly of this macromolecular machinery, CPSF3 performs the endonucleolytic cleavage followed by a non-templated addition of approximately 50-100 A residues.

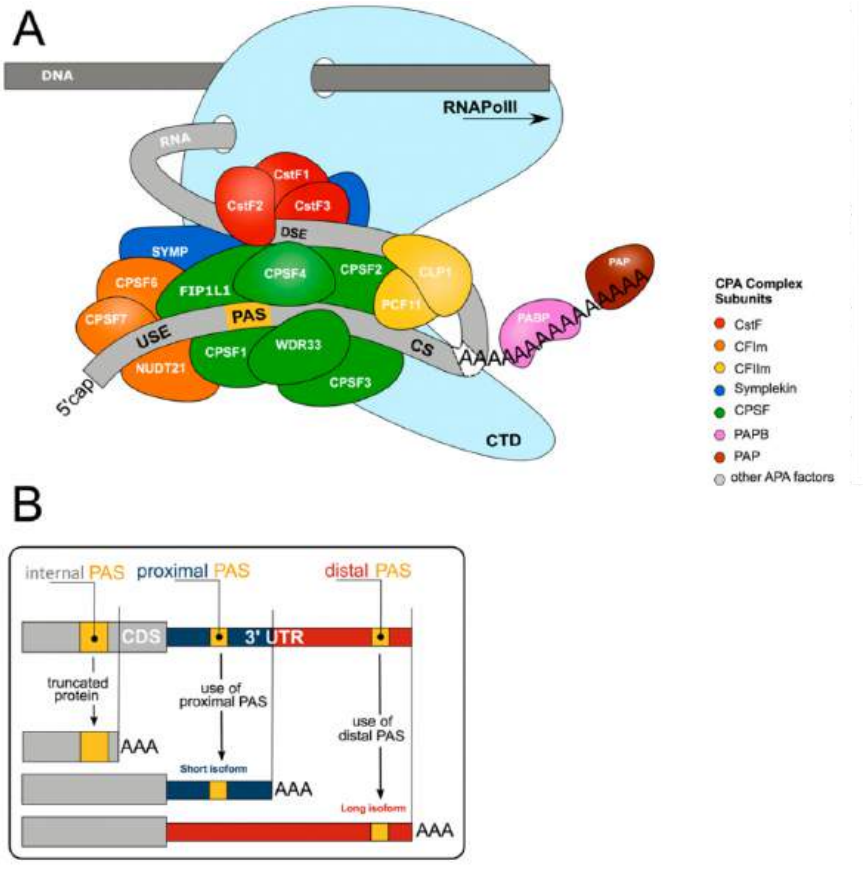


Figure 25.2.2: The human core 3′ end RNA processing machinery and impact on alternative polyadenylation. *Nourse, J., et. al. (2020) Biomolecules 10(6):915*

Panel (**A**) shows the core 3′ end processing machinery consists of complexes composed of multiple trans-acting proteins interacting with RNA via multiple cis-elements (USE = upstream sequence element; PAS = poly(A) signal; CS = cleavage site; DSE = downstream sequence element; CTD = C-terminal domain). Upon co-transcriptional assembly of these complexes, RNA cleavage and polyadenylation occur to form the 3′ end of the nascent RNA molecule.

Panel (**B**) shows more than 70% of all genes harbor more than one polyadenylation signal (PAS). This gives rise to transcript isoforms differing at the mRNA 3′ end. While alternative polyadenylation (APA) in 3′UTR changes the properties of the mRNA (stability, localization, translation), internal PAS usage (in introns or the coding sequence (CDS)) changes the C-termini of the encoded protein, resulting in different functional or regulatory properties.

Alternative polyadenylation (APA) occurs when more than one PAS is present within a pre-mRNA and provides an additional level of complexity in CPA-mediated RNA processing (Figure 26.2.2 B). Early studies revealed a significant portion of genes undergo APA, and with the advent of next-generation RNA sequencing technologies, the large-scale regulation of genes has become apparent, with approximately 70% of the transcriptome exhibiting APA regulation. As APA determines 3′UTR content and thus the regulatory features available to the mRNA, changes in the APA profile of a gene can have enormous impacts on expression.

For those trying to understand the structure and mechanism of the **cleavage/polyadenylation apparatus (CPA),** it is especially frustrating that different names are given to the constituents that comprise it, especially when comparing the proteins from different organisms. Hence it is useful to see multiple representations of the complex. Figure 25.2.3 shows a different cartoon representation of the cleavage and polyadenylation reactions. Note again the number of colored subcomplexes within the CPA as well as the different abbreviations shown for the individual proteins. This cartoon diagram is useful in visualizing the different steps involved.

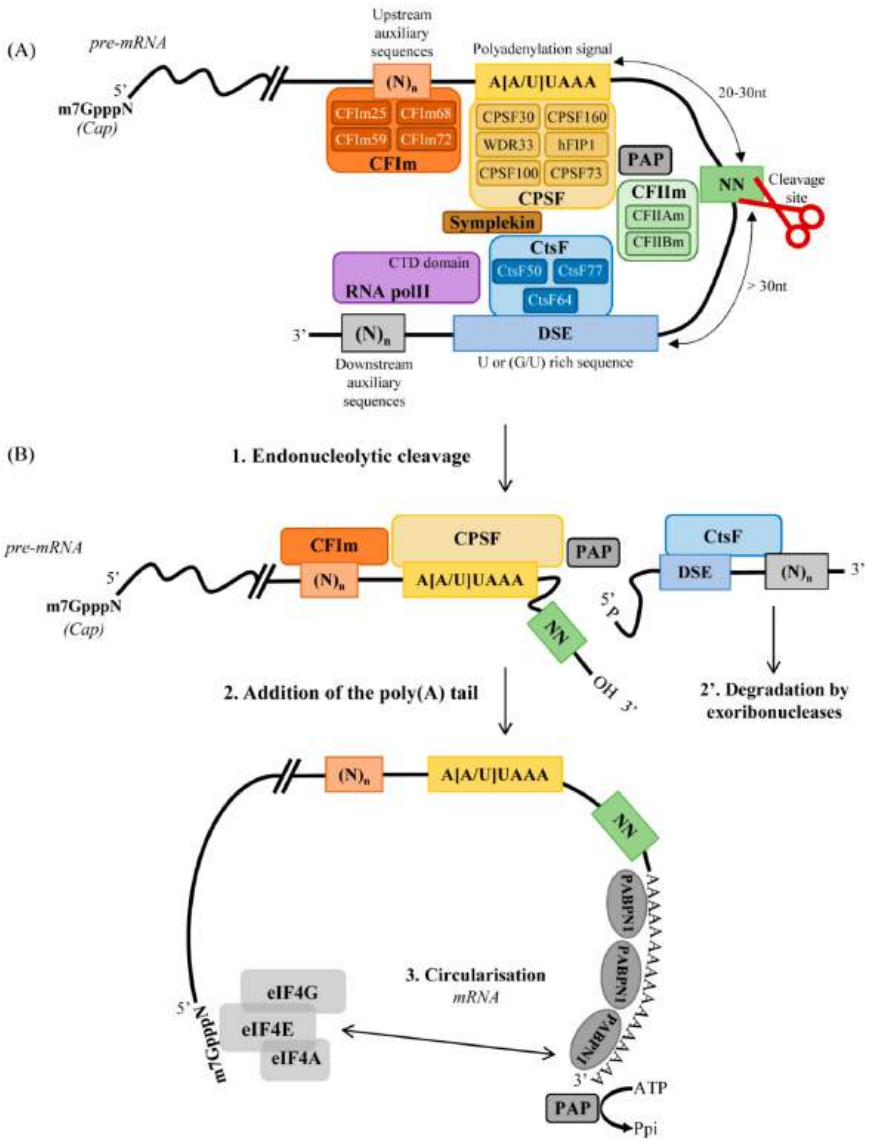


Figure 25.2.3: *Cis*-regulatory sequence elements and protein factors involved in cleavage and polyadenylation. Marsollier, A.-C.; Joubert, R.; Mariot, V.; Dumonceaux, J. Targeting the Polyadenylation Signal of Pre-mRNA: A New Gene Silencing Approach for Facioscapulohumeral Dystrophy. *Int. J. Mol. Sci.* **2018**, *19*, 1347. https://doi.org/10.3390/ijms19051347. **Creative Commons Attribution License**

Panel (**A**) shows that the specificity and efficiency of 3′end processing are determined by the binding of more than 80 RNA-binding proteins to regulatory *cis*-acting RNA sequence elements including the polyadenylation signal (PAS) A[A/U]UAAA; the cleavage site (represented by NN) and the downstream sequence element (DSE). Auxiliary sequences can be found near the polyadenylation signal or the DSE. The core processing complex, which is sufficient for the cleavage and polyadenylation, is composed of approximately 20 proteins, distributed in 8 complexes: the cleavage and polyadenylation specificity factor (CPSF), the cleavage stimulation factor (CstF); the mammalian cleavage factors I (CFIm) and the mammalian cleavage factors II (CFIIm); the single protein poly(A) polymerase (PAP); the single protein poly(A)-binding protein nuclear 1 (PABPN1); the single protein RNA polymerase II large subunit (Pol II); and the symplekin. Subunits of the different factors are indicated.

Panel (**B**) shows how CPSF and CstF are co-transcriptionally recruited to the poly(A) signal and the DSE respectively, causing an endonucleolytic cleavage of the pre-mRNA between the PAS and the DSE at the cleavage site. Two fragments are generated: one fragment with a free 5′phosphate group which is rapidly degraded by exoribonucleases and one fragment with a free 3′hydroxyl group on which 250 adenines will be added by PAP. The newly-synthetized poly(A) tail is covered by PAPBN1, allowing mRNA circularization and stabilization.

Now let's look at the structure of some of the complexes of the **cleavage/polyadenylation apparatus (CPA)**.

In yeast, the 3' processing is carried out by the cleavage and polyadenylation factor (CPF) which is called the CPSF in humans. On endonuclease cleavage, the RNA bound to RNA polymerase II is in two pieces, as shown in the figure above. The main mRNA now has a 3'-OH which is the site of polyadenylation. The minor cleavage fragment has a 5'-phosphate which gets degraded by the exonuclease Rat1, which as it cleaves the minor product helps displaces RNA polymerase II and helps to stop transcription.

In yeast, the CPF has 14 subunits with polymerase, nuclease, and phosphatase subcomplexes or "modules". The polymerase module, as the name implies, has the poly(A) polymerase, Pap1. Table 25.2.1 below shows some components of the polymerase module of both yeast CPF and human CPSF.

yeast Polymerase Module of CPF	Human mammalian polyadenylation specificity factor (mPSF) or CPSF
Cft1	CPSF160
Pfs2	WDR33
Yth1 - RNA binding subunit	CPSF30 - RNA binding subunit
Fip1 - Pap1 binding subunit	FIP1 - Pap1 binding subunit

Table 25.2.1 Some components of the polymerase module of both yeast CPF and human CPSF.

The **nuclease module** has an endonuclease (Ysh1) and a Mpe1 protein, which facilitate the cleavage site selection and polyadenylation. Table 25.2.2 below shows some components of the polymerase module in yeast and humans

yeast	human
endonuclease Ysh1	endonuclease CPSF73
pseudo-nuclease Cft2	pseudo-nuclease CPSF100
multidomain protein Mpe1	multidomain protein RBBP6

Table 25.2.2 Some components of the polymerase module in yeast and humans

Part of the Cft2 called the **yeast polymerase module interacting motif (yPIM)**, as its name implies, interacts with the polymerase module, in part through the interaction of key and conserved aromatic residues in it (F537, Y549, and F558) with a hydrophobic binding site in Cft1 and Pfs1. These interactions are key in activating and regulating the endonuclease and polyadenylation activities and hence controlling the termination of transcription.

Figure 25.2.4 shows an interactive iCn3D model of the yeast cleavage and polyadenylation specificity factor (CPF) polymerase module in complex with Mpe1, the yPIM of Cft2 and the pre-cleaved CYC1 RNA (7ZGR)

Figure 25.2.4: Yeast cleavage and polyadenylation specificity factor (CPF) polymerase module in complex with Mpe1, the yPIM of Cft2 and the pre-cleaved CYC1 RNA (7ZGR). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...ePGK2jiiuum7A7

The different parts of the complex are colored as shown below.

- Yth1 (mRNA 3' processing protein): magenta
- CFT1: green
- MPE1: orange
- Polyadenylation subunit 2 (Pfs2): yellow
- cleavage factor 2 protein (CF2P): cyan
- precleaved RNA CYC1: gray spacefill (backbone) with CPK-colored bases

25.2.4: 5'-CAP Formation

In eukaryotes, the 5′ cap, found on the 5′ end of the eventual mRNA molecule, consists of a guanine nucleotide connected to the mRNA via an unusual 5′ to 5′ triphosphate linkage, as shown in Figure 25.2.5. This guanosine is methylated on the 7 position directly after capping *in vivo* by a methyltransferase. It is referred to as the 7-methylguanylate cap, abbreviated m^7G.

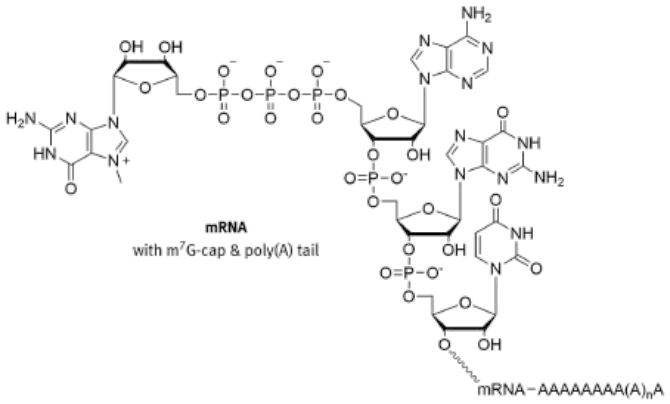


Figure 25.2.5: Structure of the 7-methylguanylate CAP. *Brisbane*

In multicellular eukaryotes and some viruses, further modifications exist, including the methylation of the 2′ hydroxy-groups of the first 2 ribose sugars of the 5′ end of the mRNA. Cap-1 has a methylated 2′-hydroxy group on the first ribose sugar, while cap-2 has methylated 2′-hydroxy groups on the first two ribose sugars. The 5′ cap is chemically similar to the 3′ end of an RNA molecule (the 5′ carbon of the cap ribose is bonded, and the 3′-OH unbonded). This provides significant resistance to 5′ exonucleases.

Small nuclear RNAs (snRNAs) contain unique 5′-caps. Sm-class snRNAs are found with 5′-trimethylguanosine caps, while Lsm-class snRNAs are found with 5′-monomethylphosphate caps. In bacteria, and potentially also in higher organisms, some RNAs are capped with NAD^+, NADH, or 3′-dephospho-coenzyme A. In all organisms, mRNA molecules can be decapped in a process known as messenger ***RNA decapping***.

For capping with 7-methylguanylate, the **capping enzyme complex (CEC)** binds to RNA polymerase II before transcription starts. As soon as the 5′ end of the new transcript emerges from RNA polymerase II, the CEC carries out the capping process (this kind of mechanism ensures capping, as with polyadenylation). The enzymes for capping can only bind to RNA polymerase II that is engaging in mRNA transcription, ensuring the specificity of the m^7G cap almost entirely to mRNA.

The 5′ cap has four main functions:

1. Regulation of nuclear export
2. Prevention of degradation by exonucleases
3. Promotion of translation (see ribosome and translation)
4. Promotion of 5′ proximal intron excision

In addition to the polyA tail, the nuclear export of RNA is regulated by the cap-binding complex (CBC), which binds to 7-methylguanylate-capped RNA, as shown in Figure 25.2.6. The CBC is then recognized by the nuclear pore complex and the mRNA is exported. Once in the cytoplasm after the pioneer round of translation, the CBC is replaced by the translation factors eIF4E and eIF4G of the eIF4F complex. This complex is then recognized by other translation initiation machinery including the ribosome, aiding in translation efficiency.

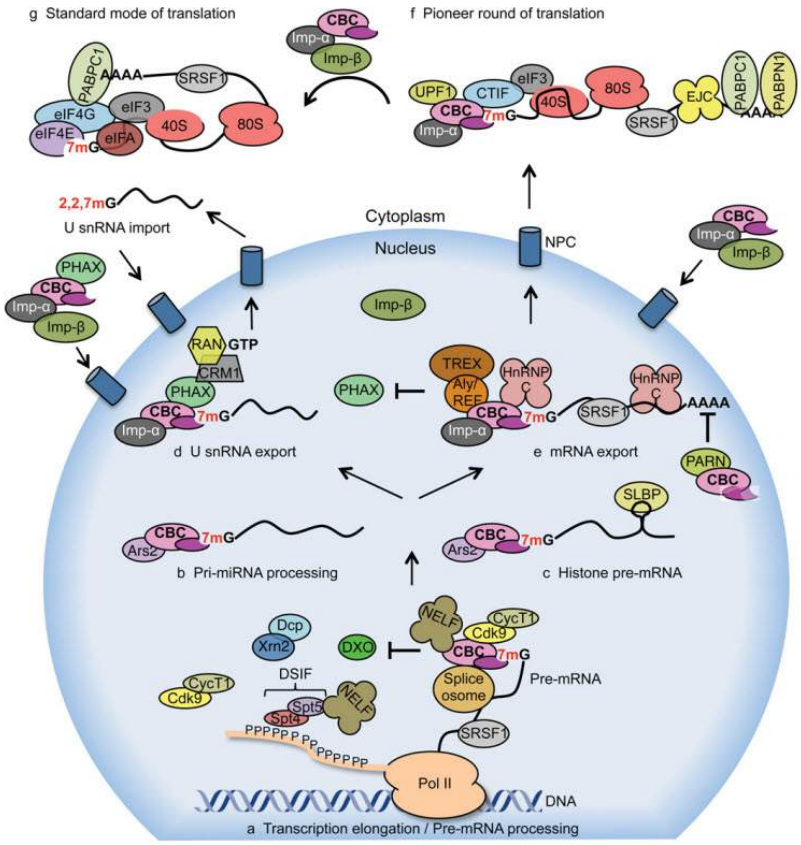


Figure 25.2.6: Importance of the 5-CAP during the lifespan of a mRNA Transcript. *Gonatopoulos-Pournatzis, T. and Crowing V. (2014) Biochemical Journal 457(2):231-42.*

Panel (a) CBC is required for pre-mRNA processing. The co-transcriptional binding of CBC to 7mG prevents the decapping activities of pre-mRNA degradation complexes [DXO (decapping exoribonuclease) and Dcp (decapping mRNA) Xrn2 (5′–3′ exoribonuclease 2)] and promotes pre-mRNA processing. CBC recruits P-TEFb [Cdk9/Cyclin T1 (CycT1)] to transcription initiation sites of specific genes promoting phosphorylation of the RNA pol II CTD at Ser2 residues. This results in the recruitment of splicing factors including SRSF1, which regulates both constitutive and alternative splicing events. Furthermore, CBC interacts with splicing machinery components that result in the spliceosomal assembly. CBC interacts with NELF and promotes pre-mRNA processing of replication-dependent histone transcripts.

Panel (b) CBC forms a complex with Ars2 and promotes miRNA biogenesis by mediating pri-miRNA processing.

Panel (c) CBC/Ars2 promotes pre-mRNA processing of replication-dependent histone transcripts.

Panel (d) CBC promotes the export of U snRNA. CBC interacts with PHAX, which recruits export factors including CRM1 and RAN·GTP.

Panel (e) CBC promotes the export of mRNA. For the export of transcripts over 300 nucleotides, hnRNP C interacts with CBC and inhibits the interaction between CBC and PHAX, allowing the CBC to interact with TREX and the transcript to be translocated to the cytoplasm. CBC interacts with the PARN deadenylase and inhibits its activity, protecting mRNAs from degradation.

Panel (f) CBC mediates the pioneer round of translation. Cbp80 interacts with CTIF, which recruits the 40S ribosomal subunit via eIF3 to the 5′ end of the mRNA for translation initiation. Upon binding of importin-β (Imp-β) to importin-α (Imp-α), mRNA is released from CBC and binds to eIF4E for the initiation of the standard mode of translation. CBC-bound mRNP components not found in eIF4E-bound mRNPs are CTIF, exon junction complex (EJC), and PABPN1.

Panel (g) The standard mode of translation is mediated by eIF4E cap-binding protein. eIF4E is a component of the eIF4F complex which promotes translation initiation.

Capping with 7-methylguanylate prevents 5′ degradation in two ways. First, degradation of the mRNA by 5′ exonucleases is prevented by functionally looking like a 3′ end. Second, the CBC and eIF4E/eIF4G block the access of decapping enzymes to the

cap. This increases the half-life of the mRNA, essential in eukaryotes as the export and translation processes take significant time.

The mechanism that promotes the 5′ proximal intron excision during splicing is not well understood, but the 7-methylguanylate cap appears to loop around and interact with the spliceosome, potentially playing a role in the splicing process.

The decapping of a 7-methylguanylate-capped mRNA is catalyzed by the decapping complex made up of at least Dcp1 and Dcp2, which must compete with eIF4E to bind the cap. Thus the 7-methylguanylate cap is a marker of an actively translating mRNA and is used by cells to regulate mRNA half-lives in response to new stimuli. During the decay process, mRNAs may be sent to P-bodies. P-bodies are granular foci within the cytoplasm that contain high levels of exonuclease activity.

25.2.4.1: **Triphosphatase and guanylyltransferase**

In capping the new mRNA, three different enzymes act sequentially:

- A phosphatase cleaves a terminal phosphate from the 5' end which has 3 phosphates at the start leaving 2 phosphates (the yeast triphosphatase is called Cet1);
- a GMP is added to the remaining diphosphate on the 5'-end to form a triphosphate in a reverse direction as shown in Figure 5 above (in yeast the RNA guanylyltransferase is called Ceg1);
- a methyl group is added to the N7 of the guanine base by a methylase also called a methyl transferase

These three enzymes are localized to a part of RNA polymerase that is highly phosphorylated, positioning them at the correct location for Cap formation. In yeast, the

Figure 25.2.7 shows an interactive iCn3D model of the Saccharomyces cerevisiae Cet1 (the triphosphatase)-Ceg1 (the guanylyltransferase) mRNA Capping Apparatus (3KYH)

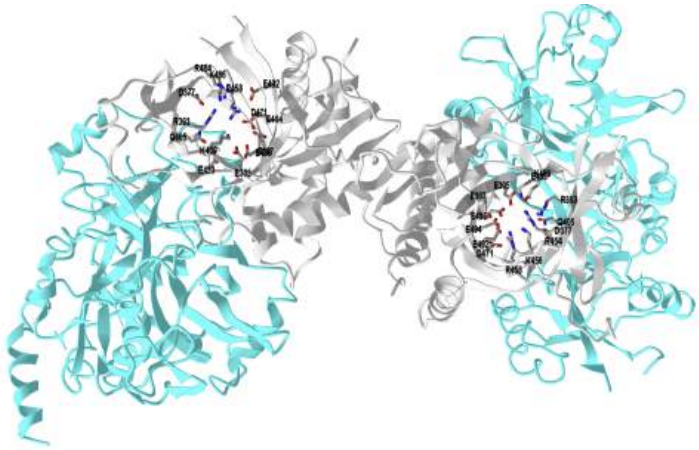


Figure 25.2.7: Saccharomyces cerevisiae Cet1-Ceg1 mRNA Capping Apparatus (3KYH) . (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...UWQ32mQrcF1eN6

The two enzymes exist as a heterotetramer of two homodimers. Two beta (Cet1-triphosphatase) subunits are shown in gray and the two alpha (Ceg1-guanylyltransferase) subunits are shown in cyan. The active site residues are shown in CPK-colored sticks and labeled. The two enzymes interact with yeast RNAP II mostly through Ceg1. Specifically, the Ceg1 oligonucleotide domain interacts with a motif, WAQKW (247-251), on Cet1. A conformational change in a flexible linker after that motif allows capping to ensue.

25.2.4.2: Methyl transferase

The next step is the methyl transfer to the N7 of guanine in the 5' end of the cap. Figure 25.2.8 shows an interactive iCn3D model of the Structure of a bacterial mRNA cap (Guanine-n7) methyltransferase (1RI1)

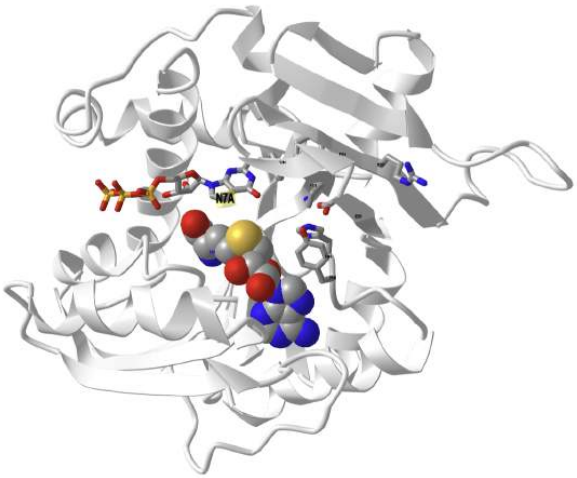


Figure 25.2.8: Structure of a bacterial mRNA cap (Guanine-n7) methyltransferase (1RI1) with bound S-adenoysl-homocysteine. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...YfaJ6W93zSDaX7

7-Methyl-guanosine-5'-triphosphate-5'-guanosine (GTG) is shown in sticks. The N7 nitrogen of guanine is labeled. S-Adenoslyl-L-homocysteine, the leaving group after methylation by S-adenosyl-L-methionine (SAM), is shown in spacefill. Key amino acid side chains in the active site are labeled (in small letters). The structure is most consistent with an in-line methyl transfer from SAM to the attacking nucleophile, the N7 of guanine. Specificity in most capping methylases occurs through noncovalent interactions (mostly base stacking, with two amino acid aromatic groups (an example of π-π stacking), or one aromatic and one nonpolar side chain. In the bacterial example shown above, the interactions include hydrophobic and hydrogen bonding interactions using Y284, F24, P175, E225, H144, and Y145.

25.2.4.3: Decapping

"What goes up must come down!" If the mRNA is capped during synthesis, there must be an enzyme to decap it. Figure 25.2.9 shows an interactive iCn3D model of the yeast mRNA decapping enzyme Dcp1-Dcp2 complex in the ATP bound closed conformation (2QKM)

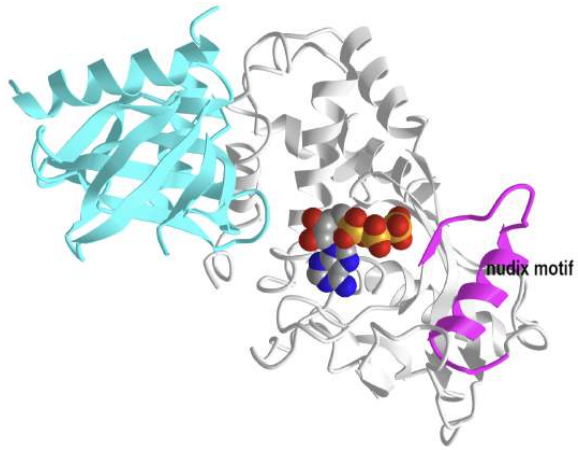


Figure 25.2.9: Yeast mRNA decapping enzyme Dcp1-Dcp2 complex in the ATP bound closed conformation (2QKM). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...hBPokeNeMfcGUA

The rendering is colored as shown below:

- a chain in cyan
- b chain in gray
- nudix motif in magenta
- ATP in spacefill

Dcp2p, as with many other enzymes, has an open (inactive) and closed (active) conformation suggesting that a conformational change between the two states regulates decapping. The ATP binding site demarcates the active site. the Dcp1 protein probably functions to stabilize the closed state.

25.2.5: mRNA Splicing

Eukaryotic genes that encode polypeptides are composed of coding sequences called ***exons*** (*ex*-on signifies that they are *ex*pressed) and intervening sequences called ***introns*** (*int*-ron denotes their *int*ervening role). Transcribed RNA sequences corresponding to introns do not encode regions of the functional polypeptide and are removed from the pre-mRNA during processing. All of the intron-encoded RNA sequences must be completely and precisely removed from a pre-mRNA before protein synthesis so that the exon-encoded RNA sequences are properly joined together to code for a functional polypeptide. If the process errs by even a single nucleotide, the sequences of the rejoined exons would be shifted, and the resulting polypeptide would be nonfunctional. The process of removing intron-encoded RNA sequences and reconnecting those encoded by exons is called ***RNA splicing.*** Intron-encoded RNA sequences are removed from the pre-RNA while it is still in the nucleus. Although they are not translated, introns appear to have various functions, including gene regulation and mRNA transport. On completion of these modifications, the ***mature transcript***, the mRNA that encodes a polypeptide, is transported out of the nucleus, destined for the cytoplasm for translation. Introns can be spliced out differently, resulting in various exons being included or excluded from the final mRNA product. This process is known as ***alternative splicing***. The advantage of alternative splicing is that different types of mRNA transcripts can be generated, all derived from the same DNA sequence. In recent years, it has been shown that some archaea also can splice their pre-mRNA.

The splicing reaction is catalyzed by the ***spliceosome***, a macromolecular complex formed by five small nuclear ribonucleoproteins (snRNPs), termed U1, U2, U4, U5, and U6, and approximately 200 proteins, as shown in Figure 25.2.10 Each of these snRNPs contains snRNAs that can interact with each other through intrastrand hydrogen bonding and hence which help localize the snRNPs to the complexes. The assembly of the spliceosome on pre-mRNA includes the binding of U1 snRNP, U2 snRNP, the pre-formed U4/U6-U5 triple snRNP, and the Prp19 complex. This assembly occurs through the recognition of several sequence elements on the pre-mRNA that define the exon/intron boundaries, which include the 5′ and 3′ splice sites (SS), the associated 3′ sequences for intron excision, the polypyrimidine (Py) tract, and the branch point sequence (BPS). The assembly of the spliceosome during the process is depicted in Figure 25.2.10

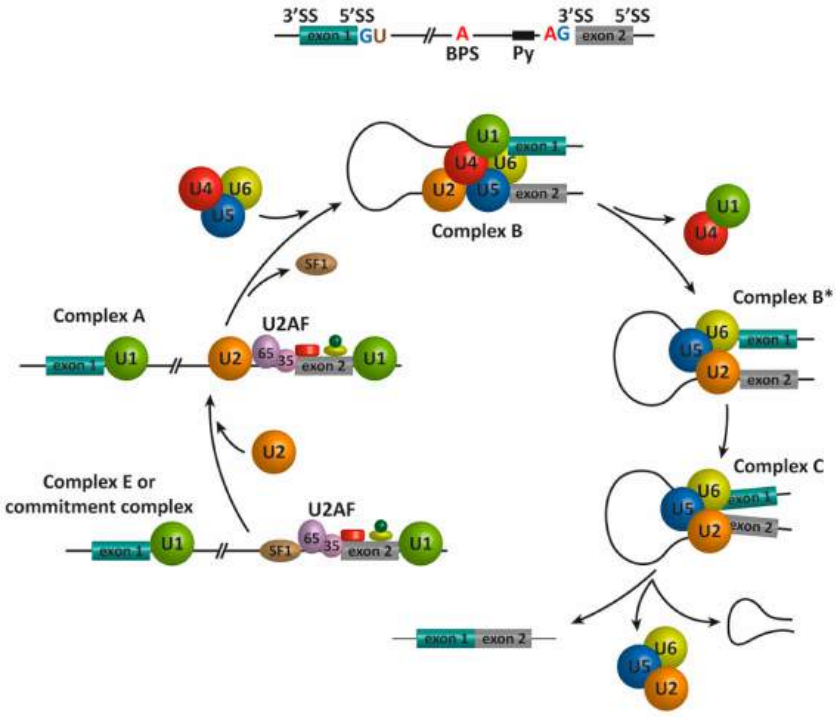


Figure 25.2.10: **Schematic representation of the spliceosome assembly and pre-mRNA splicing.** *Suñé-Pou, M., et. al. (2017) Genes 8(3):87* .

In the first step of the splicing process, the 5′ splice site (GU, 5′ SS) is bound by the U1 snRNP, and the splicing factors SF1/BBP and U2AF cooperatively recognize the branch point sequence (BPS), the polypyrimidine (Py) tract, and the 3′ splice site (AG, 3′ SS) to assemble complex E. The binding of the U2 snRNP to the BPS results in the pre-spliceosomal complex A. Subsequent steps lead to the binding of the U4/U5–U6 tri-snRNP and the formation of complex B. Complex C is assembled after rearrangements that detach the U1 and U4 snRNPs to generate complex B*. Complex C is responsible for the two transesterification reactions at the SS. Additional rearrangements result in the excision of the intron, which is removed as a lariat RNA, and ligation of the exons. The U2, U5, and U6 snRNPs are then released from the complex and recycled for subsequent rounds of splicing

There are different ways that pre-spliceosomes convert to full spliceosomes. The interactions of the small ribonucleoproteins U1 and U2, which bind at the 5' and 3' end of the introns respectively, are key. When multiple introns exist, a few different pathways occur. The interactions can be upstream (cross-introns) and downstream (cross-exons), as shown in Figure 25.2.11.

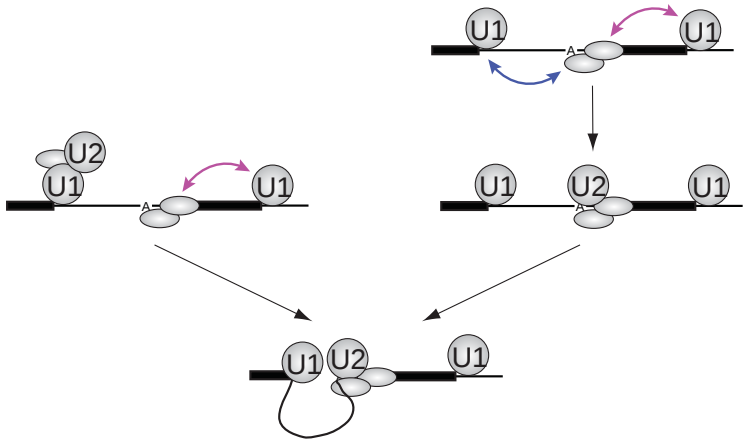


Figure 25.2.11: Implications of synergistic U2 recruitment for the mechanism of exon and intron recognition. Braun et al. (2018) . Synergistic assembly of human pre-spliceosomes across introns and exons. *eLife* **7**:e37751. https://doi.org/10.7554/eLife.37751. Creative Commons Attribution License

The left-hand side of the figure shows how the interaction of U1 and U2 can occur across a single intron, to form the mature version of the spliceosome showing the U1-U2 interaction. The colored arrow shows possible interactions that either speed up or slow down the U1-U2 mature interactions.

The right-hand side shows the formation of an intermediate cross-exon U2-U1 pair in which the exons that flank two introns are first delineated, which then leads to the physical U1-U2 interaction in the mature spliceosome when cross-intron interactions occur.

Now let's move to show the actual structures of two of the spliceosome structures shown in Figure 10 above. Figure 25.2.12shows an interactive iCn3D model of the human fully-assembled precatalytic spliceosome (pre-B complex) (6QX9)

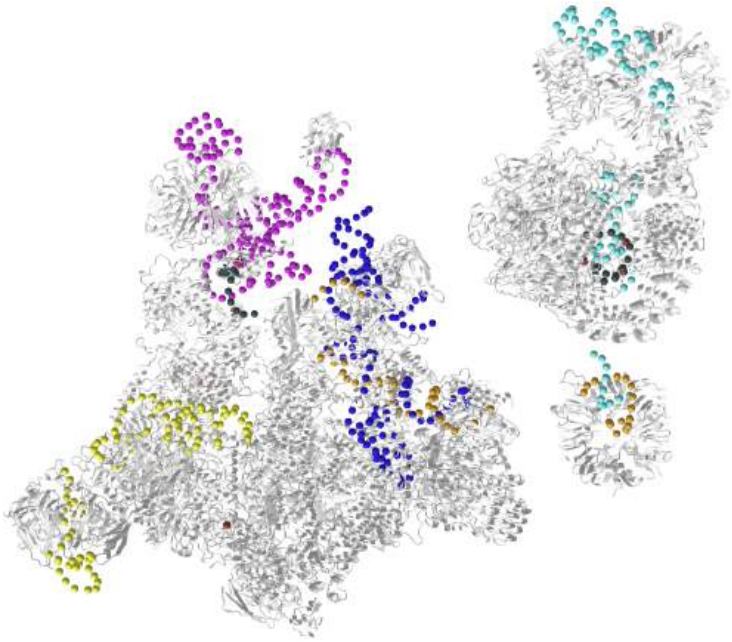


Figure 25.2.12: Human fully-assembled precatalytic spliceosome (pre-B complex) (6QX9). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...zLLvMh82m5u9M8

The full structure of the spliceosome is almost too complicated to understand even if displayed in a molecular model. It's better grasped in a way by the cartoon figure shown previously. To avoid unnecessary visual complexity, the protein subunits in the iCn3D model of the human spliceosome are shown in a gray cartoon, and only the RNA molecules in the U_n small ribonucleoproteins are shown in color, as described below. The resolution is such that each nucleotide in the RNA polymers is shown just as a single-colored sphere not connected to the next nucleotide in the RNA.

- U1 snRNA - magenta
- U6 snRNA - orange
- U5 snRNA - yellow
- U2 snRNA - cyan
- U4 snRNA - blue
- AdML pre-mRNA - black

Since it has U1, U2, U4, U5, and U6, it most closely resembles Complex B (or an immediate precursor pre-B of it) as shown in Figure 10. This form occurs before U1 snRNP dissociates. A helix from the 5'-single-stranded U1 snRNA inserts into a helicase in the complex between two RecA domains which bind ATP and unfold the nucleic acids. This allows the 5' single-stranded RNA to form interactions with an ACAGAGA sequence (mobile loop) in U6. These conformational changes allow the eventual separation of the U4 and U6 snRNA, freeing the snRNA in U6 is form the catalytic site. The B complex itself does not have a functioning active site. After the dissociation of U1 and U4 snRNPs, and the binding of another 20 or so proteins, the active spliceosome is formed.

Now let's look at the final catalytic structure, Complex C, which contains only three of the snRibonucleoproteins, U2, U5, and U6. Figure 25.2.13shows an interactive iCn3D model of the Human C Complex Spliceosome (7A5P)

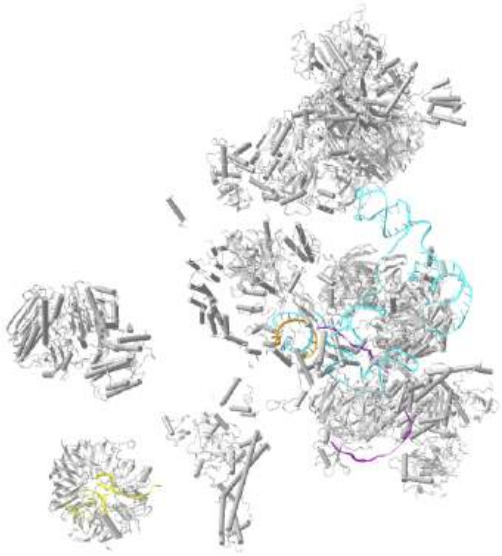


Figure 25.2.13: Human C Complex Spliceosome (7A5P). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...5co1nofAVEnyR6

The color coding is as follows:

- protein - gray
- U6 snRNA - orange
- U5 snRNA - yellow
- U2 snRNA - cyan
- AdML pre-mRNA - magenta

The interactions of the snRNAs are key for structure and catalysis in the spliceosome, and in the figures above the interactions are had to discern. Figure 25.2.14shows the interactions among U2, U5, and U6 snRNA in the human and yeast spliceosomes.

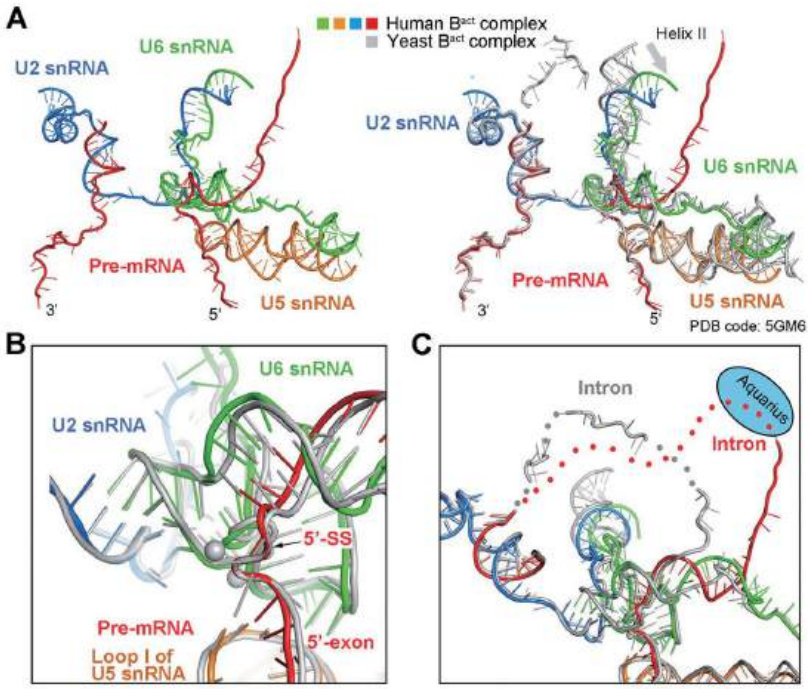


Figure 25.2.14: The RNA elements and the splicing active site of the human mature B^{act} complex. Zhang, X., Yan, C., Zhan, X. *et al.* Structure of the human-activated spliceosome in three conformational states. *Cell Res* **28**, 307–322 (2018). https://doi.org/10.1038/cr.2018.14. Creative Commons Attribution 4.0 Unported License. http://creativecommons.org/licenses/by/4.0/

Panel (A) shows the structure of the RNA elements in the core of the human mature B^{act} complex. The color-coded RNA elements of the human B^{act} complex are shown in the left panel, and their superposition with those of the *S. cerevisiae* B^{act} complex is displayed in the right panel. All yeast RNA elements are colored grey. The helix II of the U2/U6 duplex in the human B^{act} complex is bent relative to that in the yeast complex.

Panel (B) shows a structural overlay of the active site RNA elements between the human and *S. cerevisiae* B^{act} complexes.

Panel (C) shows that the U6/intron duplex in the human B^{act} complex is considerably longer than that in the *S. cerevisiae* B^{act} complex.

In mammals, the first catalytic step of the splicing reaction begins when the U1 snRNP binds the 5′ SS of the intron (defined by the consensus sequence AGGURAGU), and the splicing factors SF1 and U2AF cooperatively recognize the BPS, Py, and 3′ SS to assembled complex E or the commitment complex. Subsequently, U2 snRNP and additional proteins are recruited to the pre-mRNA BPS to form the pre-spliceosome or complex A. The binding of the U4/U6-U5 tri-snRNP forms the pre-catalytic spliceosome or complex B. After RNA-RNA and RNA-protein rearrangements at the heart of the spliceosome, U1 and U4 are released to form the activated complex B or complex B* This complex is responsible for executing the first catalytic step, through which the phosphodiester bond at the 5′ SS of the intron is modified by the 2′-hydroxyl of adenosine of the BPS to form a free 5′ exon and a branched intron, as shown in Figure 25.2.15. The reaction of the 2'-hydroxyl from the branch point adenosine nucleotide is known as a ***transesterification reaction***. During this process, additional rearrangements occur to generate the catalytic spliceosome or complex C (see Figure 10 above), which is responsible for catalyzing the second transesterification reaction leading to intron excision and exon–exon ligation. The resulting intron structure is referred to as a ***lariat structure***. After the second catalytic step, the U2, U5, and U6 snRNPs are released from the post-spliceosomal complex and recycled for additional rounds of splicing.

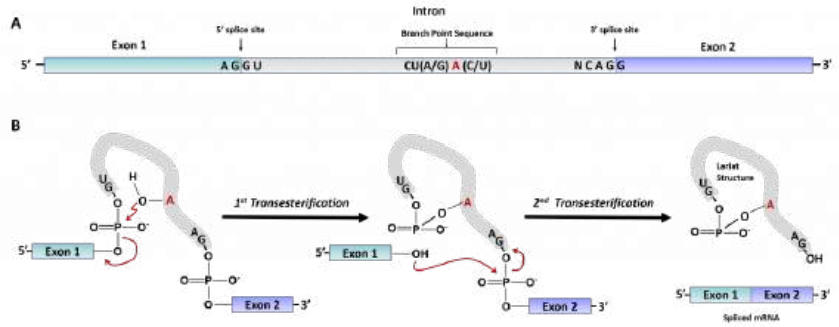


Figure 25.2.15: Transesterification Reactions Involved in mRNA Splicing.

Panel (A) shows a schematic diagram of the pre-mRNA with exons and introns indicated. Key sequences are required for splicing at the 5' and 3' intron locations, and for the recognition and positioning of the branch point Adenosine residue for the first transesterification reaction. (

Panel (B) shows a schematic of the two transesterification reactions required for intron removal. The branch point 2'-OH residue mediates attack on the 5'-phosphate of the intron guanosine residue located at the 5'-splice site. This releases the 3' hydroxyl of Exon 1 which subsequently mediates the attack of the 5' phosphate of the first guanosine residue in Exon 2. The 3' hydroxyl of the intron guanine residue is released forming the Lariat structure and Exon 1 is ligated to Exon 2.

Alternative Splicing (AS) offers an additional mechanism for regulating protein production and function. AS options are determined by the expression of or exposure to *trans* elements present within unique cellular locations and environments. Additional sequence elements within the mRNA, known as ***exonic*** and ***intronic*** splicing silencers or enhancers (ESS, ISS, ESE, and ISE, respectively), participate in the regulation of AS. Specific RNA-binding proteins, including heterogeneous nuclear ribonucleoproteins (hnRNPs) and serine/arginine-rich (SR) proteins, recognize these sequences to positively or negatively regulate AS, as shown in Figure 25.2.16. These regulators, together with an ever-increasing number of additional auxiliary factors, provide the basis for the specificity of this pre-mRNA processing event in different cellular locations within the body.

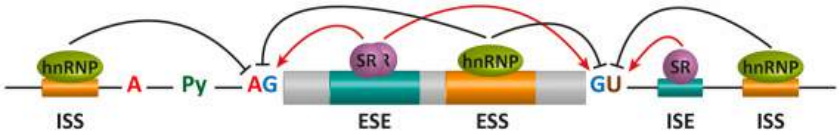


Figure 25.2.16: Alternative splicing (AS) regulation by *cis* mRNA elements and *trans-acting* factors. *Suñé-Pou, M., et. al. (2017) Genes 8(3):87*

The core *cis* sequence elements that define the exon/intron boundaries (5′ and 3′ splice sites (SS), GU-AG, polypyrimidine (Py) tract, and branch point sequence (BPS)) are poorly conserved. Additional enhancer and silencer elements in exons and introns (ESE: exonic splicing enhancers; ESI: exonic splicing silencers; ISE: intronic splicing enhancers; ISI: intronic splicing silencers) contribute to the specificity of AS regulation. *Trans-acting* splicing factors, such as SR family proteins and heterogeneous nuclear ribonucleoprotein particles (hnRNPs), bind to enhancers and silencers and interact with spliceosomal components. In general, SR proteins bound to enhancers facilitate exon definition, and hnRNPs inhibit this process. These *trans-acting* elements are expressed differentially within different locations or under different environmental stimuli to regulate AS.

There are several different types of AS events, which can be classified into four main subgroups. The first type is exon skipping, which is the major AS event in higher eukaryotes. In this type of event, a cassette exon is removed from the pre-mRNA as shown in Figure 25.2.17(panel A). The second and third types are alternative 3′ and 5′ SS selections (panels b and c). These types of AS events occur when the spliceosome recognizes two or more splice sites at one end of an exon. The fourth type is intron retention (panel d), in which an intron remains in the mature mRNA transcript. This AS event is much more common in plants, fungi, and protozoa than in vertebrates. Other events that affect the transcript isoform outcome include mutually exclusive exons (panel e), alternative promoter usage (panel f), and alternative polyadenylation (panel g).

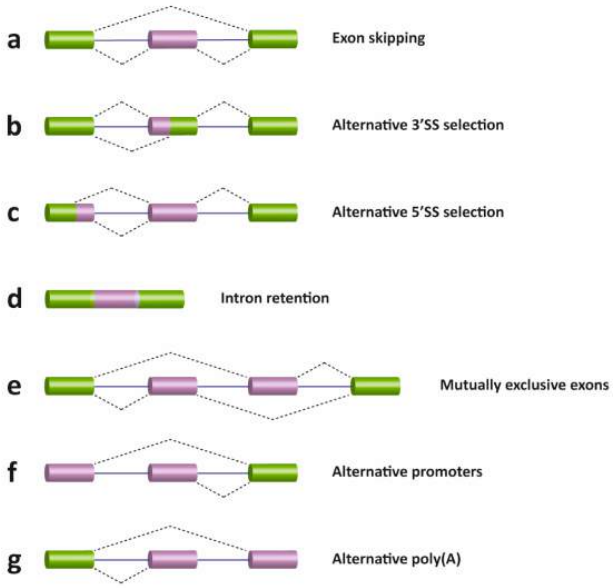


Figure 25.2.17: Schematic representation of different types of alternative transcriptional or splicing events, with exons (boxes) and introns (lines). Constitutive exons are shown in green and alternatively spliced exons are in purple. Dashed lines indicate the AS event. Exon skipping (**a**); alternative 3′ (**b**) and 5′ SS selection (**c**); intron retention (**d**); mutually exclusive exons (**e**); alternative promoter usage (**f**); and alternative polyadenylation (**g**) events are shown. Like alternative splicing (AS), the usage of alternative promoter and polyadenylation sites allows a single gene to encode multiple mRNA transcripts. *Suñé-Pou, M., et. al. (2017) Genes 8(3):87*

Here is a detailed and incredible narrated animation of splicing and the spliceosome.

25.2.6:

25.2.7: *References*

1. Parker, N., Schneegurt, M., Thi Tu, A-H., Lister, P., Forster, B.M. (2019) Microbiology. *Openstax*. Available at: https://opentextbc.ca/microbiologyopenstax/
2. Palazzo, A., and Lee, E.S. (2015) Non-coding RNA: what is function and what is junk? *Frontiers in Genetics* 6:2 Available at: file:///C:/Users/flatt/AppData/Local/Temp/fgene-06-00002.pdf
3. Wikipedia contributors. (2020, July 9). RNA. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:30, August 6, 2020, from https://en.Wikipedia.org/w/index.php?title=RNA&oldid=966784317
4. Burenina, O.Y., Oretskaya, T.S., and Kubareva, E.A. (2017) Non-Coding RNAs As Transcriptional Regulators in Eukaryotes. *Acta Naturae* 9(4):13-25. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5762824/
5. Khatter, H., Vorlander, M.K., and Muller C.W. (2017) RNA polymerase I and III: similar yet unique. Current Opinion in Structural Biology 47:88-94. Available at: https://www.sciencedirect.com/science/article/pii/S0959440X17300313
6. Wikipedia contributors. (2020, May 8). Sigma factor. In *Wikipedia, The Free Encyclopedia*. Retrieved 17:50, August 7, 2020, from https://en.Wikipedia.org/w/index.php?title=Sigma_factor&oldid=955570499
7. Bae, B., Felkistov, A., Lass-Napiokowska, A., Landick, R., and Darst, S.A. (2015) Structure of a bacterial RNA polymerase holoenzyme open protomer complex. *eLife* 4:e08504. Available at: https://elifesciences.org/articles/08504
8. Petrenko, N., Jin, Y., Dong, L., Wong, K.H., and Struhl, K. (2019) Requirements for RNA polymerase II preinitiation complex formation in vivo. *eLife* 8:e43654. Available at: https://elifesciences.org/articles/43654
9. Gupta, K., Sari-Ak, D., Haffke, M., Trowitzsch, S., and Berger, I. (2016) Zooming in on transcription preinitiation. *J Mol Biol*. 428(12):2581-2591. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4906157/
10. Wikipedia contributors. (2020, April 17). TATA-binding protein. In *Wikipedia, The Free Encyclopedia*. Retrieved 14:54, August 8, 2020, from https://en.Wikipedia.org/w/index.php?title=TATA-binding_protein&oldid=951583592
11. Patel, A.B., Greber, B.J., and Nogales, E. (2020) Recent insights into the structure of TFIID, its assembly, and its binding to core promoter. *Curr Op Struct Bio* 61:17-24. Available at: https://www.sciencedirect.com/science/article/pii/S0959440X19301113#fig0010
12. Ruff, E.F., Record, Jr., M.T., Artsimovitch, I., (2015) Initial events in bacterial transcription initiation. *Biomolecules* 5(2):1035-1062. Available at: https://www.mdpi.com/2218-273X/5/2/1035/htm
13. Kireeva, M., Opron, K., Seibold, S., Domecq, C., Cukier, R.I., Coulombe, B., Kashlev, M., and Burton, Z. (2102) Molecular dynamics and mutational analysis of the catalytic and translocation cycle of RNA polymerase. *BMC Biophysics* 5(1):11. Available at: https://www.researchgate.net/publication/225281979_Molecular_dynamics_and_mutational_analysis_of_the_catalytic_and_translocation_cycle_of_RNA_polymerase/figures?lo=1
14. Washburn, R.S., and Gottesman, M.E. (2015) Regulation of transcription elongation and termination. *Biomolecules* 5(2):1063-1078. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4496710/pdf/biomolecules-05-01063.pdf
15. Zenkin, N., and Yuzenkova, Y. (2015) New insights into the functions of transcription factors that bind the RNA polymerase secondary channel. Biomolecules 5(3):1195-1209. Available at: https://www.mdpi.com/2218-273X/5/3/1195/htm

16. Gocheva, V., LeGall, A., Boudvillain, M., Margeat, E., and Nollmann, M. (2015) Direct observation of the translocation mechanism of transcription termination factor Rho. *Nuc Acids Res* 43(1):10.1093. Available at: https://www.researchgate.net/publication/272162172_Direct_observation_of_the_translocation_mechanism_of_transcription_termination_factor_Rho
17. Miki, T.S., Carl, S.H., and Groβhans, H. (2017) Two distinct transcription termination modes dictated by promoters. Genes & Dev 31:1-10. Available at: https://www.researchgate.net/publication/320350041_Two_distinct_transcription_termination_modes_dictated_by_promoters
18. Gurumurthy, A., Shen, Y., Gunn, E.M., Bungert, J. (2018) Phase separation and transcription regulation: Are Super-Enhancers and Locus Control Regions primary sites of transcription complex assembly? *BioEssays* 1800164. Available at: https://www.researchgate.net/publication/329331157_Phase_Separation_and_Transcription_Regulation_Are_Super-Enhancers_and_Locus_Control_Regions_Primary_Sites_of_Transcription_Complex_Assembly
19. Suñé-Pou, M., Prieto-Sánchez, Boyero-Corral, S., Moreno-Castro, C., El Yousfi, Y., Suñé-Negre, J.M., Hernández-Munain, C., and Suñé, C. (2017) Targeting splicing in the treatment of human disease. Genes 8(3):87. Available at: https://www.mdpi.com/2073-4425/8/3/87/htm
20. Schaughency, P., Merran, J., and Corden J.L. (2014) Genome-wide mapping of yeast RNA polymerase II termination. *PLOS Genetics* 10(10):e1004632 Available at: https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1004632
21. Nourse, J., Spada, S., and Danckwardt, S. (2020) Emerging roles of RNA 3'-end cleavage and polyadenylation in pathogenesis, diagnosis, and therapy of human disorders. *Biomolecules* 10(6):915. Available at: https://www.mdpi.com/2218-273X/10/6/915/htm
22. Wikipedia contributors. (2020, July 30). Five-prime cap. In *Wikipedia, The Free Encyclopedia*. Retrieved 05:53, August 11, 2020, from https://en.Wikipedia.org/w/index.php?title=Five-prime_cap&oldid=970240533
23. Cortes, T. and Cox, R.A. (2015) Transcription and translation of the rpsJ, rplN and rRNA operons of the tubercle bacillus. Microbiology (2015) 161:719-728. Available at: https://www.microbiologyresearch.org/docserver/fulltext/micro/161/4/719_mic000037.pdf?expires=1597159574&id=id&accname=guest&checksum=6FFC9C066EF41C7799FAE843CE94C49F
24. Hein, P.P. and Landick, R. (2010) The bridge helix coordinates movements of modules in RNA polymerase. *BMC Biology* 8:141. Available at: https://bmcbiol.biomedcentral.com/articles/10.1186/1741-7007-8-141
25. Gonatopoulos-Pournatzis, T., and Cowling, V.H. (2014) Cap-binding complex (CBC). *Biochem. J.* 457:231-242. Available at: https://www.researchgate.net/publication/259392894_Cap-binding_complex_CBC

25.3: RNA-Dependent Synthesis of RNA and DNA

25.3.1: RNA Viruses

Infections with RNA viruses place a constant burden on our healthcare systems and economy. Over the past century, this has been particularly true for infections with the *Human immunodeficiency virus 1* (HIV-1), *Influenza A virus* (IAV), *Rotavirus* (RotaV), *West Nile virus* (WNV), *Dengue virus* (DV), *Measles virus* (MV), and the *Porcine reproductive and respiratory syndrome virus* (PRRSV). But also emergent RNA viruses can have considerable consequences, such as the *Severe acute respiratory syndrome-related coronavirus* (SARS-CoV) in 2003, the *Middle East Respiratory Syndrome coronavirus* (MERS-CoV), and most recently the *Severe acute respiratory syndrome-related coronavirus-2* (SARS-CoV-2) in December of 2019.

Eukaryotes and bacteria can be infected with a wide variety of RNA viruses. On average, these pathogens share little sequence similarity and use different replication and transcription strategies. RNA virus genomes can consist of double-stranded RNA (dsRNA) or single-stranded (ssRNA) as shown in Figure 25.3.1 (Panel a). In turn, the ssRNA viruses can be classified into positive sense (+RNA) and negative sense (−RNA) viruses, depending on the translatability of their genetic material. As summarized for four model RNA pathogens in Panel b, all RNA viruses use dedicated replication and transcription strategies to amplify their genetic material. The common denominator of these strategies is a conserved RNA-dependent polymerase domain. Figure 25.3.1 describes variations of the replication process in the different types of RNA viruses.

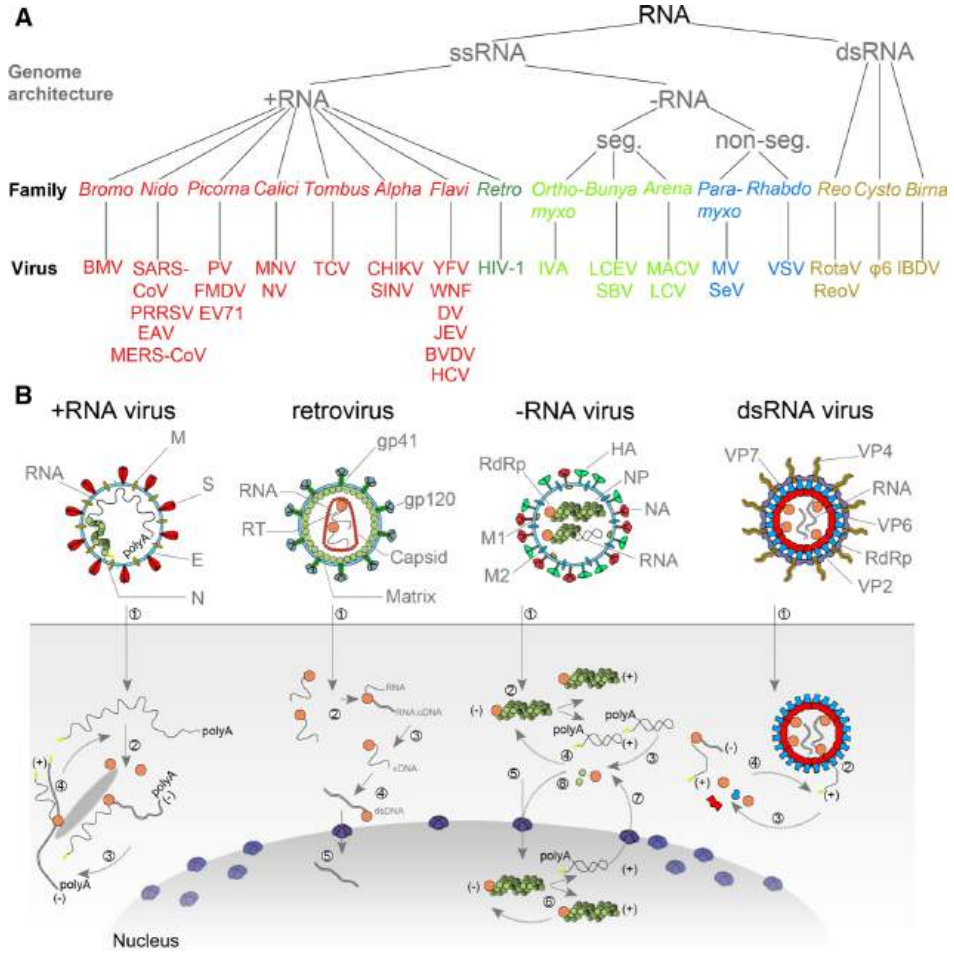


Figure 25.3.1: Taxonomy and Replication Strategies of RNA Viruses. te Velthuis, A.J.W. (2014) *Cell & Mol Life Sci 71:4403-4420*

Panel a shows a simplified taxonomy of the genome architecture of representative RNA viruses.

Panel b (*+RNA virus*)shows infection with a +RNA virus—as exemplified here with a CoV-like virion—releases a single-stranded RNA genome into the cytoplasm (*1*). (*2*) Translation of the 5′-terminal open-reading frame of the genome produces the viral replicase. (*3*) This multi-enzyme complex includes **RNA-dependent RNA polymerase (RdRp)** activity (*orange*) and associates

with intracellular membranes before −RNA synthesis commences. Newly synthesized −RNAs are subsequently used to produce new +RNAs (*4*), which are typically capped (*yellow*) and polyadenylated (polyA). **(*Retrovirus*)** HIV-1 genomes are packaged as ssRNA in virions. When the ssRNA is released (*1*) a cDNA copy is synthesized by the reverse transcriptase enzyme (RT) (*2*). The RNA is next degraded by the intrinsic RNase H activity in the RT (*3*) and the single-stranded cDNA is converted to dsDNA (*4*). The dsDNA is imported into the nucleus (*5*) for integration into the host's genetic material. **(−*RNA virus*)** (*1*) As illustrated here with an IAV-like particle, infection with an −RNA virus releases a viral RNA genome that is associated with a viral polymerase (*orange*) and nucleoprotein (*green*). (*2*) In the case of non-segmented −RNA viruses, these complexes support transcription to produce viral mRNAs or cRNAs. (*3*) Viral mRNAs are next translated and new viral proteins complex with cRNAs to synthesize new vRNAs. (*5*) The vRNA-containing complexes of some segmented −RNA viruses are imported into the nucleus of the host cell, where (*6*) the RdRp produces mRNAs or cRNAs. (*7*) mRNAs are transported to the cytoplasm, while cRNAs are bound by new viral proteins to form cRNPs for −RNA synthesis. **(*dsRNA virus*)** Fully duplexed RNA genomes lack cap and polyA elements. (*1*) The RdRp (*orange*), therefore, transcribes the viral genome inside the capsid of the virion (*blue* and *red*), so viral mRNAs can be (*2*) released into the cytoplasm as illustrated here with a rotavirus-like virion. In the cytoplasm, the mRNA is translated (*3*) or replicated by newly synthesized viral RdRps.

Viral RNA-dependent RNA Polymerases (RdRps) were discovered in the early 1960s from studies on mengovirus and polio viruses when it was observed that these viruses were not sensitive to actinomycin D, a drug that inhibits cellular DNA-directed RNA synthesis. This lack of sensitivity suggested that there is a virus-specific enzyme that could copy RNA from an RNA template and not from a DNA template. RdRps show some structural similarity to telomerase enzymes suggesting a potential ancestral relationship of telomerase with RdRps.

The most famous example of RdRp is of the polio virus. The viral genome is composed of RNA, which enters the cell through receptor-mediated endocytosis. From there, the RNA can act as a template for complementary RNA synthesis, immediately. The complementary strand is then, itself, able to act as a template for the production of new viral genomes that are further packaged and released from the cell ready to infect more host cells. The advantage of this method of replication is that there is no DNA stage; replication is quick and easy. The disadvantage is that there is no 'backup' DNA copy (Fig 26.3.1b; + RNA virus).

Reverse Transcriptase (RT) enzymes, on the other hand, that are common to retroviruses, convert the + RNA strand from the virus into a cDNA copy (Fig 26.3.1b; retroviruses). The RT enzyme then degrades the RNA and the single-stranded cDNA is converted to dsDNA. The dsDNA is then integrated into the host's genome where it can remain in a dormant state. Upon reactivation, + RNA will be manufactured, along with viral proteins used in the assembly of the infectious viral particles.

25.3.2: Structure and Mechanism of RNA Viral Polymerases

The RNA-dependent polymerase domain is a member of the superfamily of template-dependent nucleic acid polymerases and typically <400 amino acids in length. Its sequence is highly variable on average, with some regions showing less than ~10% conservation. Strong amino acid conservation can be observed, however, in regions that are directly involved in nucleotide selection or catalysis, such as the strictly conserved glycine and aspartates in the center of the domain. The prototypic RNA virus RNA polymerase domain harbors seven of such regions or motifs, which are arranged in the order G, F1–3, A, B, C, D, and E from amino- to carboxy-terminus. Each of the seven motifs in the RNA polymerase domain adopts a specific and conserved fold, as shown in Figure 25.3.2

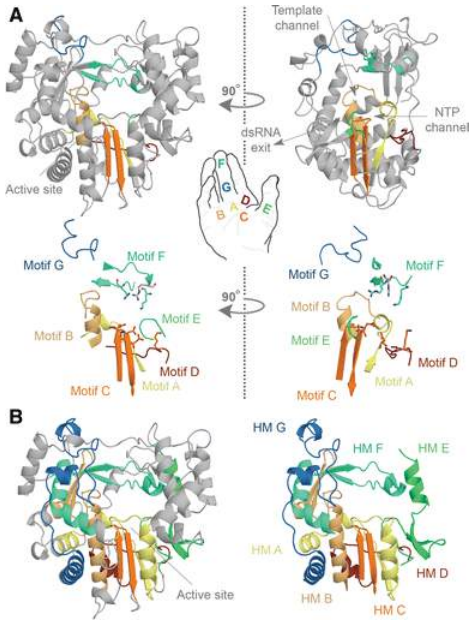


Figure 25.3.2: Conserved structural elements in the RNA virus polymerases. te Velthuis, A.J.W. (2014) *Cell & Mol Life Sci 71:4403-4420.*

Panel a shows the structure of the FMDV RdRp. The motifs A, B, C, D, E, F, and G are color-coded *yellow*, *gold*, *orange*, *red*, *light green*, *aquamarine*, and *blue*. Overall the polymerase structure adopts a shape that resembles a cupped right hand. Herein, motifs A–E lie on the palm, while motifs F and G are part of the fingers. In the side-view of the enzyme, the location of the template and NTP channels is indicated.

Panel b shows conserved structural elements of the FMDV RdRp. Homomorphic A–G were mapped and color-coded *yellow*, *gold*, *orange*, *red*, *light green*, *aquamarine*, and *blue*, respectively. Images A and B are based on PDB accession 2E9R.

Figure 25.3.3 shows an interactive iCn3D model of the Foot-and-mouth disease virus RNA-polymerase in complex with a template- primer RNA, PPi, and UTP (2E9Z)

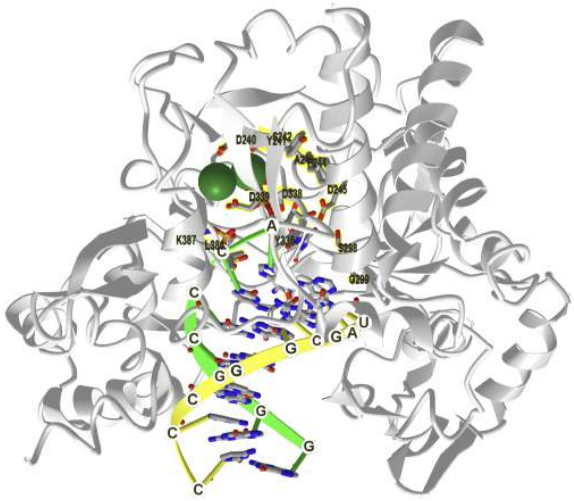


Figure 25.3.3: Foot-and-mouth disease virus RNA-polymerase in complex with a template- primer RNA, PPi and UTP (2E9Z). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...rZJwYWKHqcAd76

Active site region side chains are shown as sticks and labeled. The template stand is shown in yellow and the primer is in green. Note the conserved amino acids Tyr-336, the catalytic Asp-338, and Lys-387 in motifs C and E, respectively. The primer 3'OH forms an H bond with active site Asp 338 (on motif C). In some structures, Asp 338 is bound to a metal ion that interacts with the PPP of the incoming NTP. Two metal ions are involved in catalysis. Tyr 336 interacts with the primer nucleotide while Lys 387 and

Arg 388 interact with the primer backbone. These three residues are highly conserved. The acceptor base of the template strands is adjacent to the NTP binding site. There is no proofreading activity in viral RNA replication.

RNA-dependent RNA polymerases generally have a groove on top of the enzyme where RNA enters. It exits in the front. Nucleotides enter at a rear channel. This is illustrated in Figure 25.3.4 (Panel b). Following the convention for cellular DNA-dependent DNA polymerases (DdDp), the seven motifs and homomorphs are grouped into three subdomains. These subdomains are called fingers, palm, and thumb in reference to the polymerase domain's likeliness to a cupped right hand, as shown in Figure 25.3.4 (Panel A). The finger subdomain loops that interconnect the fingers with the thumb in the RNA-dependent RNA polymerases (RdRps) of +RNA and dsRNA viruses create an overall "closed-hand" conformation that is unique to RNA-dependent RNA polymerases (RdRps) and generally not seen in crystal structures of DNA-dependent DNA polymerases (DdDps) or reverse transcriptases (RTs).

The three subdomains of the RNA-dependent polymerases work together to facilitate the binding of RNA and nucleotides (NTPs). The thumb subdomain contains various residues that are involved in RNA binding and these generally pack into the minor groove of the template RNA. In some polymerases, the thumb also contains sequences that protrude into the template channel to help stabilize the initiating NTPs on the ssRNA template. Crucially, these protrusions are also able to undergo relatively large conformational rearrangements to facilitate the translocation of the template following the first condensation reaction. The other residues of the thumb subdomain contribute to the formation of the NTP tunnel, which sits at an ~110° angle with the template channel as shown in Figure 25.3.4, Panel A. The cavity is lined with positively charged amino acids, though charge interactions are likely not sufficient to guide NTPs into the interior. Indeed, molecular dynamics (MD) simulations have shown that the residues of the NTP channel can also explore a relatively large amount of space, which may allow the RdRp to actively "pump" NTPs down the channel.

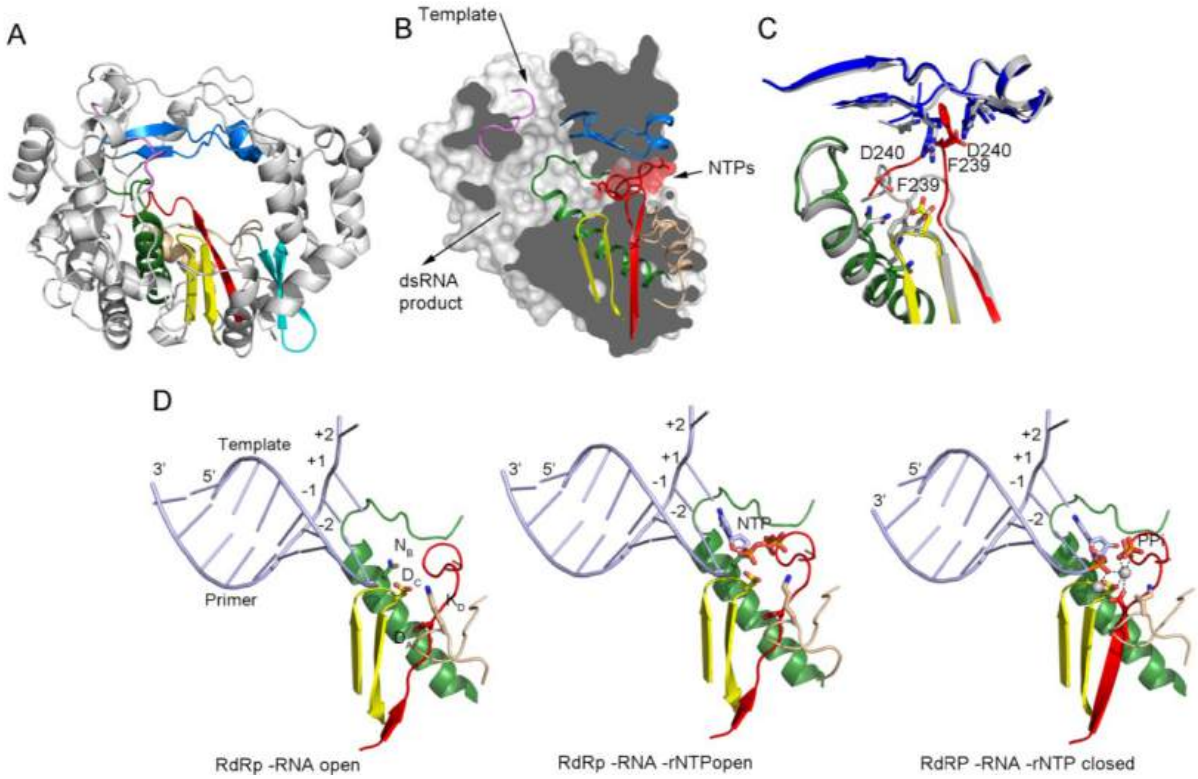


Figure 25.3.4: The Picornavirus (PV) Replication-Elongation Complexes for RNA Dependent RNA Polymerase. Ferrer-Orta, C., et al. (2015) Viruses 7(8):4438-4460

Panel (**A**) shows a ribbon representation of a typical picornaviral RdRP (model from the cardiovirus EMCV $3D^{pol}$, PDB id. 4NZ0). The seven conserved motifs are indicated in different colors: motif A, red; motif B, green; motif C, yellow; motif D, sand; motif E, cyan; motif F, blue; motif G, pink;

Panel (**B**) shows a lateral view of a surface representation of the enzyme (grey) that has been cut to expose the three channels that are the entry and exit sites of the different substrates and reaction products. The structural elements that support motifs A–G are also shown as ribbons. This panel also shows the organization of the palm sub-domain with motif A shown in two alternative

conformations: the standard conformation (PDB id. 4NZ0) found in the apo-form of most crystallized $3D^{pol}$ proteins and the altered conformation found in the tetragonal crystal form of the EMCV enzyme (PDB id. 4NYZ). The alterations affect mainly Asp240, the amino acid in charge of incoming ribonucleotide triphosphate (rNTP) selection, and the neighboring Phe239 that move ~10 Å away from its position in the enzyme catalytic cavity directed towards the entrance of the nucleotide channel, approaching to motif F;

Panel (**C**) shows a close-up of the structural superimposition of the two alternative conformations of the EMCV motif A;

Panel (**D**) shows the PV replication-elongation complexes. Sequential structures illustrating the movement of the different palm residues from a binary PV $3D^{pol}$-RNA open complex (left) to an open $3D^{pol}$-RNA-rNTP ternary complex (middle) where the incoming rNTP is positioned in the active site for catalysis and, a closed ternary complex (right) after nucleotide incorporation and pyrophosphate (PPi) release. The residues D_A (involved in rNTP selection through an interaction with the 2′ hydroxyl group), D_C (the catalytic aspartate of motif C), K_D (the general acid residue of motif D that can coordinate the export of the PPi group), and N_B (a conserved Asn of motif B, interacting with D_A) have been highlighted as sticks. The different structures correspond to the $3D^{pol}$-RNA (PDB id. 3OL6), $3D^{pol}$-RNA-CTP open complex (PDB id. 3OLB) and $3D^{pol}$-RNA-CTP closed complex (PDB id. 3OL7) structures of PV elongation complexes, respectively

The finger subdomain residues mainly pack into the major groove of the RNA template. Furthermore, upon entry of the template, they can unstack the single strand at position +3, as shown in Figure 25.3.5 (Panel A). The non-conserved structural elements of the fingers subdomain play a role in RNA binding as well. In particular, the fingertips of dsRNA and some +RNA virus RdRps allow the polymerase to “hold” the template without extensive conformational changes. The variations and extensions in the fingers subdomain has also been shown to play roles in protein–protein interactions, phosphorylation, oligomerization, and nuclear import. In contrast, the HIV-1 RT lacks such extensions and adopts a more “open-hand” or U-shaped binding cleft. As a consequence, the RT structure must rearrange its subdomains to coordinate the binding of the template, nascent strand, and incoming dNTPs.

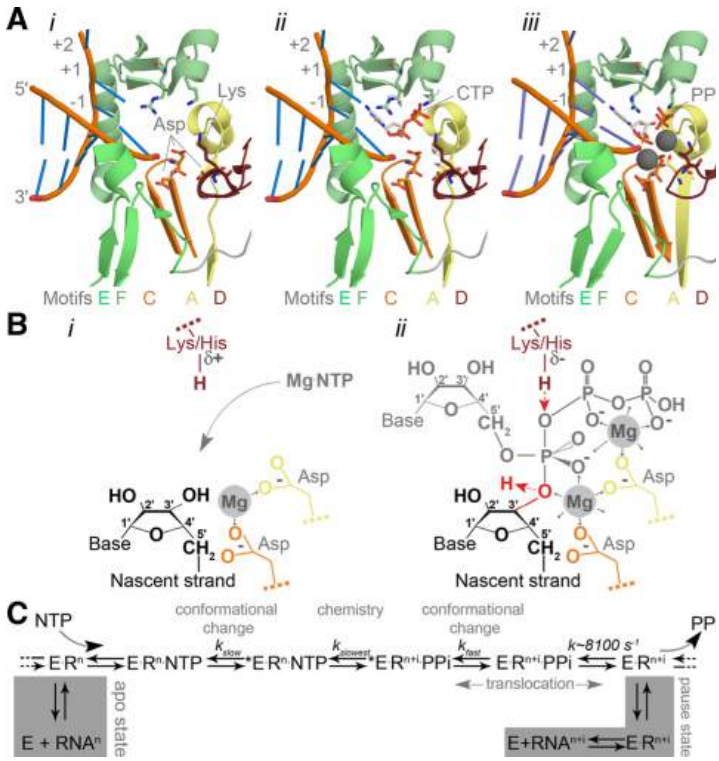


Figure 25.3.5: Catalysis in the RNA virus polymerase active site. te Velthuis, A.J.W. (2014) *Cell & Mol Life Sci 71:4403-4420.*

Panel a shows the structure of the PV active site as it moves from a native state or elongation complex (*i*) to an open complex (*ii*), and a closed complex (*iii*). The closed complex depicted here shows the active site after catalysis. Highlighted are the metal-binding aspartates of motifs A and C, and the lysine of motif D that acts as a general acid. Color coding by motif as in Fig. 26.3.2. Image based on PDB accessions 3Ol6, 3OLB, and 3OL7.

Panel b shows a schematic presentation of the RdRp active site. The aspartates (Asp) of motif A (*yellow*) and C (*orange*) bind divalent metal ions (marked Mg and *shaded grey*), which are used to coordinate the formation of a new phosphodiester bond at the 3′-OH (*red* in panel ii) of the nascent strand (*yellow*). The general acid (*red* Lys/His in panel ii) is positioned near the β-phosphate of the incoming NTP to protonate the PP_i leaving group.

Panel c shows a simplified schematic of the kinetic steps of RNA polymerases. *Asterisk* indicates a closed complex

The NTP and template entry channels meet at the palm subdomain - Figure 25.3.5, Panel A. This is a structure that is comprised of a central, partially formed three-stranded β-sheet, which is also present in RNA-recognition motifs (RRMs). The relative movement of these strands within the RRM is vital to catalysis and dependent on NTP binding. Only when a correct NTP binds can motif A and motif C align and the RRM fold be completed. This catalytically competent conformation of the active site is often referred to as the closed complex (not to be confused with the "closed-hand", which refers to the overall structure of the RdRp) - Figure 25.3.5, Panel d.

The polymerase reaction creates new phosphodiester bonds between NTPs using RNA as a template. The NTP substrates involved in this reaction are coordinated by two metal ions, which are bound by the conserved aspartates of motifs A and C - Figure 25.3.5 - panel a, i. They also position the NTP's triphosphate optimally for attack by the sugar moiety of the nascent strand once its 3′ carbon has lost a proton- Figure 25.3.5, panel b. The N-terminal aspartate of motif C thus uses a metal ion to fix the α-phosphate of the incoming NTP and reduce the pK_a of the nascent RNA's 3′-OH to facilitate the attack - Figure 25.3.5, panel b, ii. The amino-terminal carboxylate of motif A, on the other hand, stabilizes the β- and γ-phosphates with its divalent metal as well as the pentavalent (phosphorane) intermediate that forms during catalysis - Figure 25.3.5, panel b, ii. Structural and biochemical analyses have shown that the formation of this transient structure is dependent on the attack of the NTP's α-phosphate by the 3′-OH, which is often the rate-limiting catalytic step in NTP condensation - Figure 25.3.5, panel b.

Motif D's lysine or histidine assists the N-terminal aspartate of motif A in coordinating the β-phosphate of the incoming NTP, analogous to the trigger loop in DdDps. However, when the positively charged side chain of motif D approaches the β-phosphate, it can protonate the PP_i leaving group as well - Figure 25.3.5, panel b, ii. This second protonation step in the active site is not essential for the polymerase reaction, since catalysis can still take place when motif D's lysine has been mutated to a residue with a different pK_a. The polymerase reaction rate will nevertheless be 50- to 1,000-fold higher when a lysine or histidine is present. Recent data even suggests that motif D can coordinate the export of the PP_i group from the active site once catalysis has taken place, thereby triggering the translocation of the RNA.

Overall, the RNA replication process can be summarized with this four-step mechanism:

1. Nucleoside triphosphate (NTP) binding - initially, the RdRp presents with a vacant active site in which an NTP binds, complementary to the corresponding nucleotide on the template strand. Correct NTP binding causes the RdRp to undergo a conformational change.
2. Active site closure - the conformational change, initiated by the correct NTP binding, results in the restriction of active site access and produces a catalytically competent state.
3. Phosphodiester bond formation - two Mg^{2+} ions are present in the catalytically active state and arrange themselves in such a way around the newly synthesized RNA chain that the substrate NTP can undergo a phosphatidyl transfer and form a phosphodiester bond with the newly synthesized chain. With the use of these Mg^{2+} ions, the active site is no longer catalytically stable, and the RdRp complex changes to an open conformation.
4. Translocation - once the active site is open, the RNA template strand can move by one position through the RdRp protein complex and continue chain elongation by binding a new NTP, unless otherwise specified by the template.

Figure 25.3.6 shows a more detailed representation of the elongation catalytic cycle.

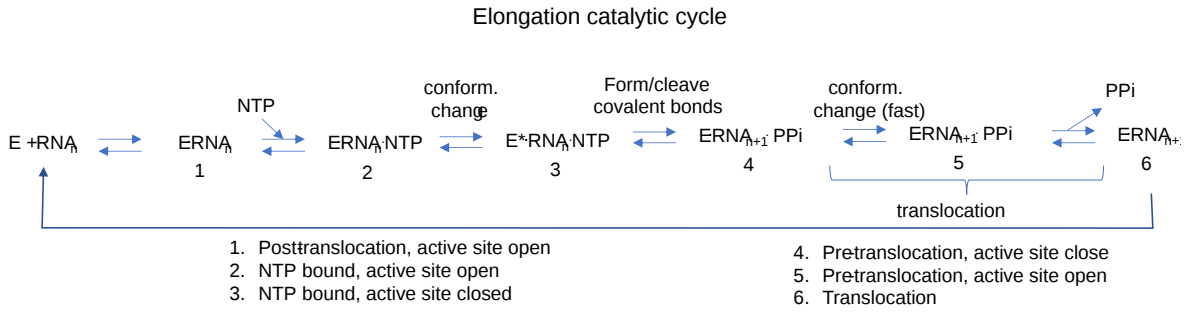


Figure 25.3.6: Elongation catalytic cycle of RNA-dependent RNA polymerases

RNA synthesis can be performed using a primer-independent (*de novo*) or a primer-dependent mechanism that utilizes a viral protein genome-linked (VPg) primer. The *de novo* initiation consists of the addition of a nucleoside triphosphate (NTP) to the 3'-OH of the first initiating NTP. During the following so-called elongation phase, this nucleotidyl transfer reaction is repeated with subsequent NTPs to generate the complementary RNA product. The termination of the nascent RNA chain produced by RdRp is not completely known, however, it has been shown that RdRp termination is sequence-independent.

One feature of RNA-dependent RNA polymerase replication is the immense error rate during transcription. RdRps and RTs are known to have a lack of fidelity on the order of 10^4 nucleotides, which is thought to be a direct result of their insufficient proofreading abilities. This high rate of variation is favored in viral genomes as it allows for the pathogen to overcome defenses developed by hosts trying to avoid infection allowing for evolutionary growth.

Let's look at another RdRp from the poliovirus. The virus has a single-stranded, positive-sense RNA genome, which makes it infectious in itself but much more infectious when replicated. It is translated into a polyprotein which is cleaved into about 12 separate proteins. One protein, 3Dpol, is an RNA-dependent RNA polymerase that transcribes the infecting +RNA strand into a -RNA strand, which then serves as a template for more +RNA strands.

Figure 25.3.7 shows an interactive iCn3D model of the Poliovirus polymerase elongation complex with 2',3'-dideoxy-CTP (3OLB). The model shows the palm, finger, and thumb domains.

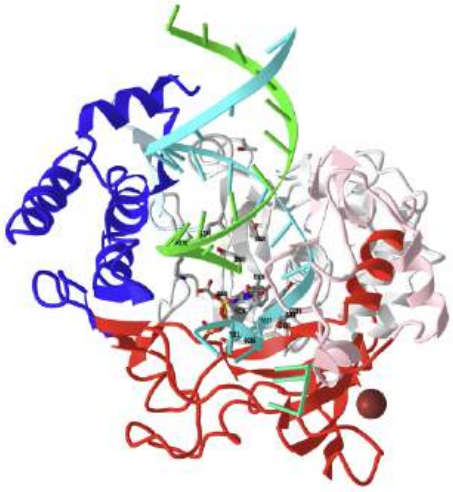


Figure 25.3.7: Poliovirus polymerase elongation complex with 2',3'-dideoxy-CTP (3OLB). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...f5xp9f7B6Qo4SA

Color coding is as follows:

- red: finger domain
- pink: pinky finger part of the finger domain
- blue: thumb domain
- light blue: primer grip at the beginning of the thumb domain
- cyan: RNA template

- green: RNA product
- gray: palm domain

Structural studies suggest that the pinky finger is involved in the initiation, while nucleotide binding and catalysis used the palm domain. The thumb domain appears to affect the translocation step. In most nucleases, the finger binds and positions NTPs in the active site. After catalysis, the reverse motion opens the active site which allows translocation by effectively ratcheting the template by one base pair which is driven by PPi release. These conformational changes can't be made so easily in +RNA stranded viral RdRps.

25.3.3: References

1. te Velthuis, A.J.W. (2014) Common and unique features of viral RNA-dependent polymerases. *Cell. & Mol Life Sci* 71:4403-4420. Available at: https://link.springer.com/article/10.1007/s00018-014-1695-z#Sec1

2. Wikipedia contributors. (2021, August 2). RNA-dependent RNA polymerase. In *Wikipedia, The Free Encyclopedia*. Retrieved 02:47, August 8, 2021, from https://en.Wikipedia.org/w/index.php?title=RNA-dependent_RNA_polymerase&oldid=1036813791

3. Ferrer-Orta, C., Ferrero, D., and Verdaguer, N. (2015) RNA-Dependent RNA Polymerases of Picornaviruses: From the Structure to Regulatory Mechanisms (2015 Viruses 7(8):4438-4460. Available at: https://www.mdpi.com/1999-4915/7/8/2829/htm

---

25.4: 26.4 References

<h210.6refs">26.4.6 References

1. Parker, N., Schneegurt, M., Thi Tu, A-H., Lister, P., Forster, B.M. (2019) Microbiology. *Openstax*. Available at: https://opentextbc.ca/microbiologyopenstax/
2. Palazzo, A., and Lee, E.S. (2015) Non-coding RNA: what is function and what is junk? *Frontiers in Genetics* 6:2 Available at: file:///C:/Users/flatt/AppData/Local/Temp/fgene-06-00002.pdf
3. Wikipedia contributors. (2020, July 9). RNA. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:30, August 6, 2020, from https://en.Wikipedia.org/w/index.php?title=RNA&oldid=966784317
4. Burenina, O.Y., Oretskaya, T.S., and Kubareva, E.A. (2017) Non-Coding RNAs As Transcriptional Regulators in Eukaryotes. *Acta Naturae* 9(4):13-25. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5762824/
5. Khatter, H., Vorlander, M.K., and Muller C.W. (2017) RNA polymerase I and III: similar yet unique. Current Opinion in Structural Biology 47:88-94. Available at: https://www.sciencedirect.com/science/article/pii/S0959440X17300313
6. Wikipedia contributors. (2020, May 8). Sigma factor. In *Wikipedia, The Free Encyclopedia*. Retrieved 17:50, August 7, 2020, from https://en.Wikipedia.org/w/index.php?title=Sigma_factor&oldid=955570499
7. Bae, B., Felkistov, A., Lass-Napiokowska, A., Landick, R., and Darst, S.A. (2015) Structure of a bacterial RNA polymerase holoenzyme open protomer complex. *eLife* 4:e08504. Available at: https://elifesciences.org/articles/08504
8. Petrenko, N., Jin, Y., Dong, L., Wong, K.H., and Struhl, K. (2019) Requirements for RNA polymerase II preinitiation complex formation in vivo. *eLife* 8:e43654. Available at: https://elifesciences.org/articles/43654
9. Gupta, K., Sari-Ak, D., Haffke, M., Trowitzsch, S., and Berger, I. (2016) Zooming in on transcription preinitiation. *J Mol Biol*. 428(12):2581-2591. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4906157/
10. Wikipedia contributors. (2020, April 17). TATA-binding protein. In *Wikipedia, The Free Encyclopedia*. Retrieved 14:54, August 8, 2020, from https://en.Wikipedia.org/w/index.php?title=TATA-binding_protein&oldid=951583592
11. Patel, A.B., Greber, B.J., and Nogales, E. (2020) Recent insights into the structure of TFIID, its assembly, and its binding to core promoter. *Curr Op Struct Bio* 61:17-24. Available at: https://www.sciencedirect.com/science/article/pii/S0959440X19301113#fig0010
12. Ruff, E.F., Record, Jr., M.T., Artsimovitch, I., (2015) Initial events in bacterial transcription initiation. *Biomolecules* 5(2):1035-1062. Available at: https://www.mdpi.com/2218-273X/5/2/1035/htm
13. Kireeva, M., Opron, K., Seibold, S., Domecq, C., Cukier, R.I., Coulombe, B., Kashlev, M., and Burton, Z. (2102) Molecular dynamics and mutational analysis of the catalytic and translocation cycle of RNA polymerase. *BMC Biophysics* 5(1):11. Available at: https://www.researchgate.net/publication/225281979_Molecular_dynamics_and_mutational_analysis_of_the_catalytic_and_translocation_cycle_of_RNA_polymerase/figures?lo=1
14. Washburn, R.S., and Gottesman, M.E. (2015) Regulation of transcription elongation and termination. *Biomolecules* 5(2):1063-1078. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4496710/pdf/biomolecules-05-01063.pdf
15. Zenkin, N., and Yuzenkova, Y. (2015) New insights into the functions of transcription factors that bind the RNA polymerase secondary channel. Biomolecules 5(3):1195-1209. Available at: https://www.mdpi.com/2218-273X/5/3/1195/htm
16. Gocheva, V., LeGall, A., Boudvillain, M., Margeat, E., and Nollmann, M. (2015) Direct observation of the translocation mechanism of transcription termination factor Rho. *Nuc Acids Res* 43(1):10.1093. Available at: https://www.researchgate.net/publication/272162172_Direct_observation_of_the_translocation_mechanism_of_transcription_termination_factor_Rho
17. Miki, T.S., Carl, S.H., and Groβhans, H. (2017) Two disctinct transcription termination modes dictated by promoters. Genes & Dev 31:1-10. Available at: https://www.researchgate.net/publication/320350041_Two_distinct_transcription_termination_modes_dictated_by_promoters
18. Gurumurthy, A., Shen, Y., Gunn, E.M., Bungert, J. (2018) Phase separation and transcription regulation: Are Super-Enhancers and Locus Control Regions primary sites of transcription complex assembly? *BioEssays* 1800164. Available at: https://www.researchgate.net/publication/329331157_Phase_Separation_and_Transcription_Regulation_Are_Super-Enhancers_and_Locus_Control_Regions_Primary_Sites_of_Transcription_Complex_Assembly
19. Suñé-Pou, M., Prieto-Sánchez, Boyero-Corral, S., Moreno-Castro, C., El Yousfi, Y., Suñé-Negre, J.M., Hernández-Munain, C., and Suñé, C. (2017) Targeting splicing in the treatment of human disease. Genes 8(3):87. Available at: https://www.mdpi.com/2073-4425/8/3/87/htm

20. Schaughency, P., Merran, J., and Corden J.L. (2014) Genome-wide mapping of yeast RNA polymerase II termination. *PLOS Genetics* 10(10):e1004632 Available at: https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1004632
21. Nourse, J., Spada, S., and Danckwardt, S. (2020) Emerging roles of RNA 3'-end cleavage and polyadenylation in pathogenesis, diagnosis, and therapy of human disorders. *Biomolecules* 10(6):915. Available at: https://www.mdpi.com/2218-273X/10/6/915/htm
22. Wikipedia contributors. (2020, July 30). Five-prime cap. In *Wikipedia, The Free Encyclopedia*. Retrieved 05:53, August 11, 2020, from https://en.Wikipedia.org/w/index.php?title=Five-prime_cap&oldid=970240533
23. Cortes, T. and Cox, R.A. (2015) Transcription and translation of the rpsJ, rplN and rRNA operons of the tubercle bacillus. Microbiology (2015) 161:719-728. Available at: https://www.microbiologyresearch.org/docserver/fulltext/micro/161/4/719_mic000037.pdf?expires=1597159574&id=id&accname=guest&checksum=6FFC9C066EF41C7799FAE843CE94C49F
24. Hein, P.P. and Landick, R. (2010) The bridge helix coordinates movements of modules in RNA polymerase. *BMC Biology* 8:141. Available at: https://bmcbiol.biomedcentral.com/articles/10.1186/1741-7007-8-141
25. Gonatopoulos-Pournatzis, T., and Cowling, V.H. (2014) Cap-binding complex (CBC). *Biochem. J.* 457:231-242. Available at: https://www.researchgate.net/publication/259392894_Cap-binding_complex_CBC

CHAPTER OVERVIEW

26: Protein Metabolism

26.1: The Genetic Code

26.2: Protein Synthesis

26.3: Translational Regulation and Protein Degradation

26.1: The Genetic Code

26.1.1: Overview of Translation

Within this chapter, we will cover the details of prokaryotic and eukaryotic translation. *Translation* is the process of converting the information housed in mRNA into the protein sequence. Essentially, you are translating the language of nucleotides into the language of amino acids. Recall that prokaryotic and eukaryotic transcription and translation systems differ in large part due to the compartmentalization of larger eukaryotic cells. Due to this compartmentalization, transcription and translation are separated spatially and temporally within the cell. Transcription occurs within the nucleus of eukaryotes and translation occurs within the cytoplasm. Prokaryotes do not have compartmentalization and have, thus, evolved a coupled transcription/translation system where both processes occur simultaneously. Both are illustrated in Figure 26.1.1.

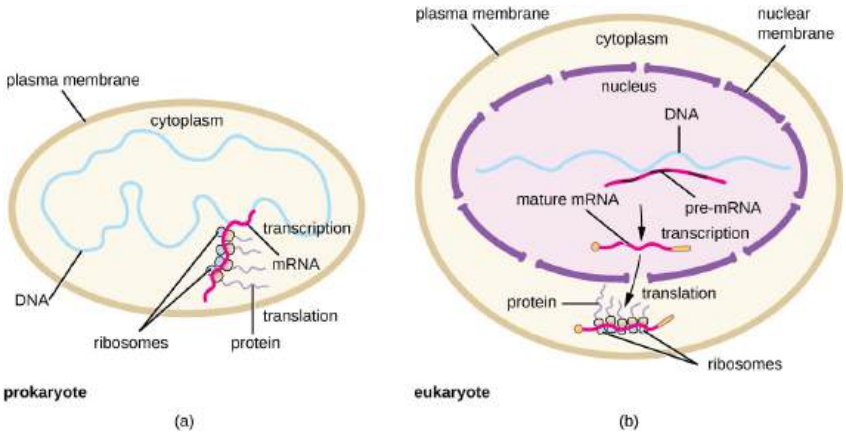


Figure 26.1.1: Cellular Location of Transcription and Translation in Prokaryotes and Eukaryotes. Baccei, A., and Rice, M. Lumen Learning

Panel (a) shows that prokaryotes lack cellular compartmentalization and show coupled transcription-translation processing;

Panel (b) shows that eukaryotes have a high degree of compartmentalization and separate the processes of transcription, which is in the nucleus of the cell, from the processes of translation, which is localized in the cytoplasm.

Recall that peptide formation is a dehydration reaction that combines the carboxylic acid of the upstream amino acid with the amine functional group of the downstream amino acid to form an amide linkage as shown in Figure 26.1.2. Water is the by-product. The ribosome (a large complex of peptides and rRNA molecules) serves as the enzyme that mediates this reaction. It requires a mature mRNA to serve as the template and directionally performs peptide bond synthesis from the N to the C-terminal of the growing peptide/protein. This is known as *N- to C-synthesis*. Note that the overall protonation state shown is very unlikely since under conditions when the carboxyl groups are protonated, so would the amines. This representation makes it easier to highlight the departing water

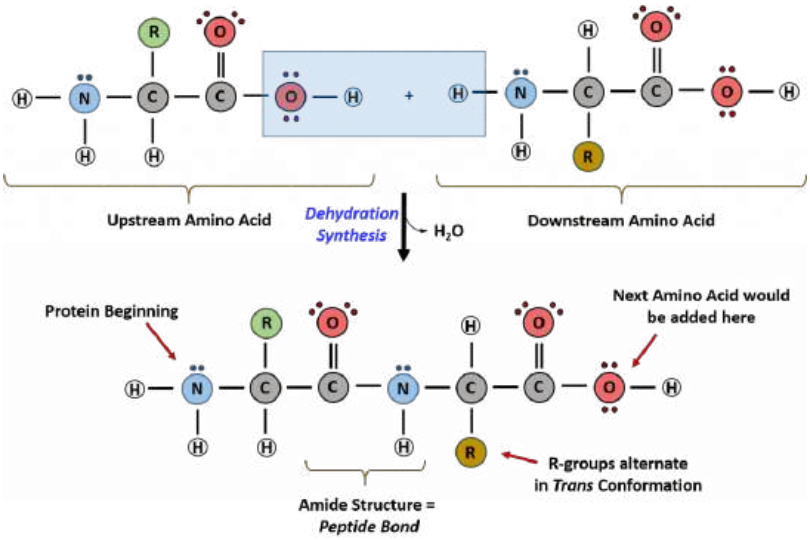


Figure 26.1.2: Formation of the Peptide Bond. The addition of two amino acids to form a peptide requires dehydration synthesis. The carboxylic acid of the upstream amino acid is joined with the amine functional group of the downstream amino acid to form the amide linkage. Within the ribosome, this reaction is highly directional and only occurs in the N to C orientation.

To maintain proper protein function, the error rate of translation is approximately 10^{-4} or 1 error in every 10,000 amino acids encoded. The fidelity of protein synthesis is maintained by the ribosome's ability to match the code from the template mRNA stand with the appropriate amino acid. Template mRNA is read by the ribosome in groups of three nucleotides, called a *codon*, as shown in Figure 26.1.3

Figure 26.1.3: Reading the mRNA Template. modified from: (A) Yikrazuul and (C) Zedalis, J., et al. Openstax

The template is non-overlapping and reads in discrete groups of three. This is known as the *reading frame* of the mRNA, and it is always read from the 5′ to 3′ direction. Thus, for each mRNA, there are three potential reading frames (panel B). Only one reading frame will be the correct one for protein synthesis. The ribosome must recognize and align the correct reading frame of the mRNA such that the correct codon sequences can be read. Small distinct tRNA molecules are tethered with specific amino acids and contain specific anticodons that complement mRNA codon sequences. The tRNA molecules can cycle on and off of the ribosome structure to hybridize with the correct codon sequences and chaperone the correct amino acid for peptide bond formation. The

ribosome then serves as a ribozyme and mediates the peptidyl transferase activity to form the peptide bond. The mRNA is then shifted to reveal the next codon within the sequence and the process is repeated until the entire protein has formed. Panel C shows the codon chart for all of the possible combinations of three nucleotides. 64 possible codon combinations are possible using the 4 nucleotide possibilities, but only 20 amino acids are encoded during protein synthesis. Each codon is specific for a single amino acid. There is *very little ambiguity* within the code.

However, there is *redundancy* within the code; i.e. many amino acids have more than one codon that encodes for that specific amino acid. To account for this redundancy, many tRNA molecules can recognize more than one codon using a single anticodon. This is known as *degeneracy. Degeneracy* usually occurs at the third position of the codon and is known as the *wobble base* position. *Degeneracy* helps to minimize the effects of mutations within the coding sequence, as mutations in the *wobble base* position will often lead to *silent mutation*– ie the mutation will still encode for the same amino acid.

In addition, if comparing the polarity of amino acids encoded by the different codons, neighboring codons typically encode for amino acids with similar polarity, as shown in Figure 26.1.4 This also helps to minimize the effects of mutations, by converting one amino acid within the sequence to one that has similar polarity. This type of mutation is more likely to cause less disturbance to the 3-dimensional structure of the resulting protein and retain biological function.

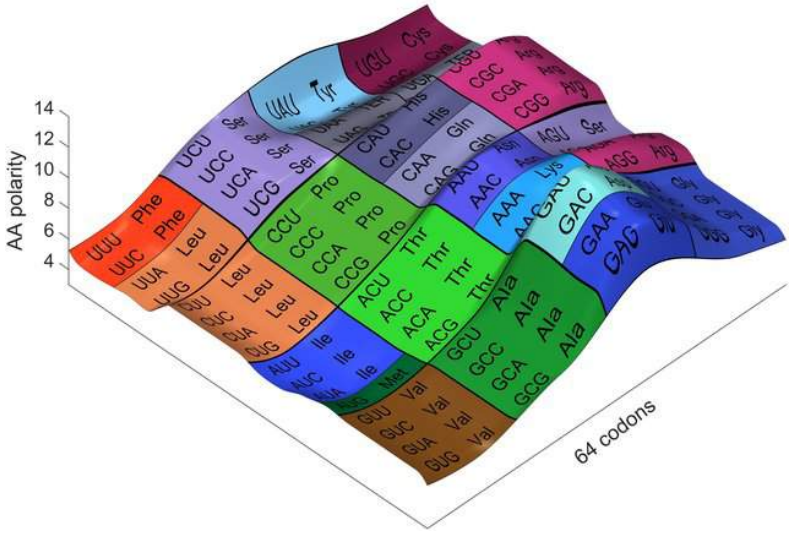


Figure 26.1.4: The Relationship of Codons and the Polarity of the Encoded Amino Acids. A representation of the genetic code, as a function of the measured polarity of each codon. The smoothness of the landscape shows that moving from one AA to its neighbor does not change the polarity too abruptly. Eckmann, J-P., Rougemont, J., and Tlusty, T. (2019) ResearchGate

Degeneracy within the genetic code also allows for differential A/T & G/C concentrations within species. For example, the G/C content of bacteria can range from as low as 30% to as high as 70%. Organisms living at high temperatures or extreme environments often have higher G/C content. This effectively increases the hydrogen bond strength between the strands of the DNA (G/C pairs have 3 H-bonds, whereas A/T pairs only have 2) and causes an increase in the melting temperature of the chromosome. Thus, the DNA is stable at a higher temperature or under more extreme ionic conditions, such as high salt. A more detailed discussion of how a single tRNA can function to recognize more than one codon is the topic of the next section.

The Genetic Code is universal for almost all species alive on the planet, providing support for a single origin of life. Most deviations in the code occur within the mitochondria of eukaryotic species, as shown in Figure 26.1.5

Figure 26.1.5: Variations in the Genetic Code. Universal genetic code (inside the box) and variations in animal mitochondrial genetic code (outside the box). Term = Termination Codon. Figure from: Watanabe, K., and Yokobori, S. (2011) J. Nuc. Acids Article ID: 623095, 12 pages

26.1.2: Transfer RNA (tRNA) Structure

Transfer RNAs (tRNAs) are central players in translation, functioning as adapter molecules between the informational level of nucleic acids and the functional level of proteins. Typically, tRNA molecules are between 76 – 90 nucleotides long and show a highly conserved secondary and tertiary structure. They also show the highest amount of nucleotide modification of all types of RNA with modifications concentrated in two hotspots—the anticodon loop and the tRNA core region, where the D- and T-loop interact with each other, stabilizing the overall structure of the molecule, as shown in Figure 26.1.6 These modifications can cause large rearrangements as well as local fine-tuning in the 3D structure of a tRNA.

The life of a transfer RNA (tRNA) molecule starts with a series of important maturation steps that can vary in their sequential order from case to case. Leader and trailer sequences are removed by a set of endo- and exonucleases, and in several tRNA precursors, splicing reactions excise intronic sequences. Furthermore, in many organisms, the sequence CCA, which represents the site of amino acid attachment, is not encoded but has to be added post-transcriptionally by CCA-adding enzymes. While all primary tRNA transcripts are composed of the four standard RNA bases A, C, G, and U, many of these nucleotides are modified, altering their properties in very different ways. Currently, 93 post-transcriptional modifications are known, and the variety of their functions is at least similarly diverse and not fully understood. The complexity of such modifications ranges from simple methylations at the bases or the ribose to rather complex and large base hypermodifications, whose synthesis often requires a whole cascade of enzymatic reactions. Modifications can alter a tRNA's shape in subtle ways, but can also lead to massive structural rearrangements. In addition, they ensure efficient translation by maintaining the anticodon loop structure and promoting correct codon-anticodon interactions.

After maturation, tRNAs have multiple interaction partners in their life cycle, ranging from aminoacyl-tRNA-synthetases that are responsible for amino acid attachment, to translation factors, ribosomes, and mRNAs. Apart from synthetases, these interaction partners do not specifically act on one individual tRNA transcript or isoacceptor, but on all tRNAs, similar to the above-mentioned CCA-adding enzyme. Thus, despite a high sequence variation, a cell's tRNAs show a well-conserved cloverleaf-like secondary structure that was originally discovered in 1965. The similar structure of all tRNA molecules allows them to bind to common protein synthesis machinery, such as the ribosome and CCA-adding enzymes. The cloverleaf consists of five parts: the acceptor stem (containing the tRNA's 5′- and 3′-ends), the D-arm, the anticodon arm, the variable loop, and the TΨC-arm (T-arm). At the 3′-terminus, the tRNA carries the CCA sequence, required for aminoacylation, tRNA positioning in the ribosome, and translation termination. In a conserved network of tertiary interactions, mostly between D- and T-loop, tRNAs fold into an L-shaped three-dimensional structure, which was first solved by Kim et al. in 1974, as shown in Figure 26.1.6 (Panel B). The anticodon and the amino acid-accepting CCA-ends are separated by the longest possible distance from each other. This conserved structure of a tRNA is essential for its recognition by the ribosome, other RNAs, and proteins and, consequently, for its functionality. For example, the CCA-adding enzyme uses the acceptor domain for substrate recognition, whereas aminoacyl-tRNA-synthetases use several recognition elements like anticodon, acceptor stem, or the discriminator position.

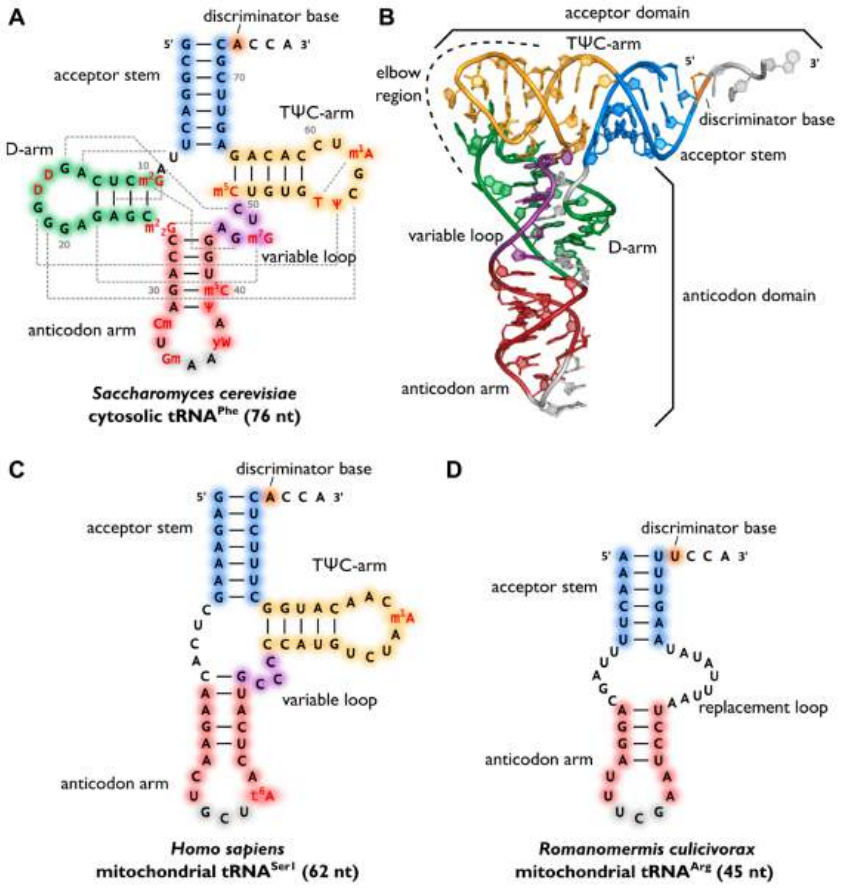


Figure 26.1.6: **Variability of transfer RNA (tRNA) structures.** Lorenz, C., et. al. (2017) Biomolecules 7(2):35

Panel (A) shows the canonical cloverleaf secondary structure of cytosolic $tRNA^{Phe}$ from S. cerevisiae is shown with acceptor stem (blue), D-arm (green), anticodon arm (red), variable loop (purple) and TΨC-arm (yellow). The anticodon is labeled in grey, the discriminator base in orange and post-transcriptional modifications in red. Grey dashed lines indicate tertiary interactions based on structural data and the length of the RNA is indicated in parenthesis;

Panel (B) shows the L-shaped tertiary structure of the cytosolic $tRNA^{Phe}$ from S. cerevisiae. Protein Data Bank entry (PDB): 1EHZ. The acceptor domain is composed of a stacked T-arm and acceptor stem, whereas D- and anticodon arm form the anticodon domain. The region where both domains come together and interact with each other via tertiary base pairing is also called the elbow region;

Panel(C) shows the secondary structure of human mitochondrial $tRNA^{Ser1}$, which lacks the whole D-arm;

Panel (D) shows the secondary structure of the mitochondrial $tRNA^{Arg}$ from the nematode *Romanomermis culicivorax*, which lacks both D- and T-arm. Instead, we find a so-called replacement loop. It represents the shortest tRNA found *in vivo*.

Surprisingly, not all tRNAs fold into the canonical cloverleaf structure. Especially many mitochondrial tRNAs are reduced in length and sometimes completely lack the D- or T-arm as shown in Figure 26.1.6, Panel C. In the mitochondria of nematodes, this situation is carried to an extreme, as tRNAs lacking one or even both arms seem to be the rule (Panel D).

Figure 26.1.7 shows an interactive iCn3D model of the yeast phenylalanine tRNA (1EHZ).

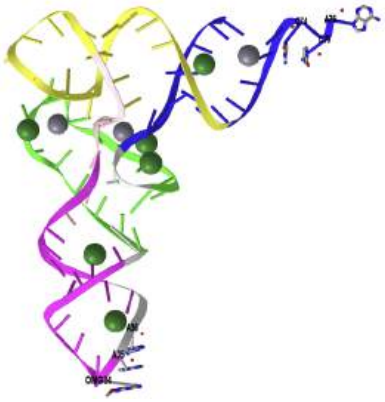


Figure 26.1.7: Yeast phenylalanine tRNA (1EHZ). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...yq7wdue7MNuT56

The coloring, which matches those in Figure 6 above, is shown below:

- Acceptor Stem - blue
- D Arm - green
- Anticodon= magenta
- Variable Loop - pink
- Tω C - yellow

Post-transcriptional enzyme-catalyzed modification of tRNA occurs at many base and sugar positions and influence specific anticodon–codon interactions and regulates translation, its efficiency, and fidelity. This phenomenon of nucleoside modification is most remarkable and results in a rich structural diversity of tRNA of which over 93 modifications have been characterized.

The variety of post-transcriptional modifications can be classified into two groups according to their complexity. The first group comprises the majority of modified bases, which have simple methylations at the ribose or base moiety that are usually introduced by a single enzymatic reaction. Simple modifications can be found at almost every position of the tRNA molecule with a high density in the tRNA core region, where tertiary interactions between D- and T-arm stabilize the three-dimensional fold, as shown in Figure 26.1.8. The second group includes complex modifications, whose synthesis requires the sequential activity of several enzymes. Most often these hypermodified nucleosides are found in the anticodon of tRNAs, where they play a direct role in codon recognition and create what is known as the *wobble base* or *wobble position*.

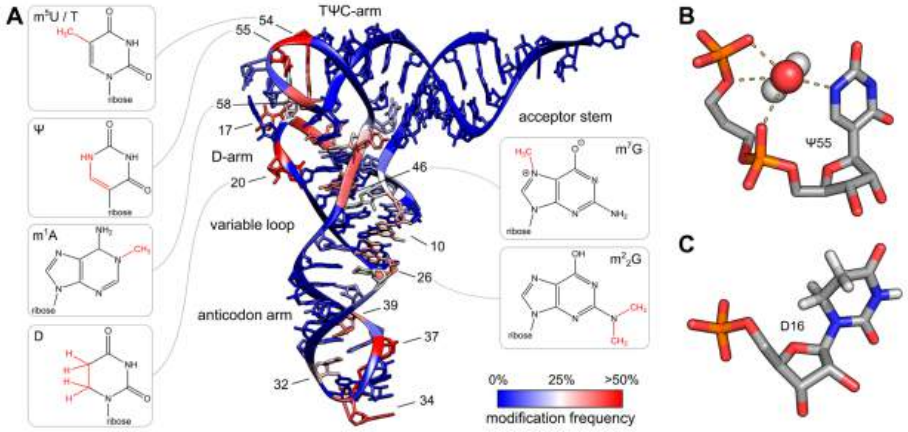


Figure 26.1.8: Post-transcriptional modifications in tRNA. Lorenz, C., et. al. (2017) Biomolecules 7(2):35

Panel (A) shows the colored tRNA structure shows the modification frequency of each base. The modification data were taken from the tRNAmodviz database and plotted on the crystal structure of $tRNA^{Phe}$ from S. cerevisiae. Blue-colored bases are rarely modified; red-colored bases are modification hotspots. tRNAs possess two regions with high modification levels—the anticodon loop (especially positions 34 and 37) and the core or elbow region, where D- and T-loop bases interact with each other and stabilize the tertiary fold. For some important positions, the chemical structure of the most frequent modification at this position is shown;

Panel (B) shows the three-dimensional structure of pseudouridine at position 55 of $tRNA^{Phe}$ from S. cerevisiae. The additional H-bond donor at N1 interacts with the 5′-adjacent phosphates via a coordinated water molecule. The hydrogen bound to N1 was not resolved in the crystal structure. The ribose shows a stabilizing C3′-endo conformation. PDB: 1EHZ-

Panel (C) shows the three-dimensional structure of D16 in the D-arm of $tRNA_i^{Met}$ from Schizosaccharomyces pombe. The C5-C6 bond of dihydrouridine is reduced, which leads to a non-planar structure of the base. The ribose takes the less stable C2′ -endo conformation. PDB: 2MN0.

A *wobble base pair* is a pairing between two nucleotides in RNA molecules that does not follow Watson-Crick base pair rules. The four main *wobble base pairs* are guanine-uracil (G-U), hypoxanthine-uracil (I-U), hypoxanthine-adenine (I-A), and hypoxanthine-cytosine (I-C), as shown in Figure 26.1.9 To maintain the consistency of nucleic acid nomenclature, “I” is used for hypoxanthine because hypoxanthine is the nucleobase of the inosine nucleotide; nomenclature otherwise follows the names of nucleobases and their corresponding nucleosides (e.g., “G” for both guanine and guanosine – as well as for deoxyguanosine). The thermodynamic stability of a wobble base pair is comparable to that of a Watson-Crick base pair. Wobble base pairs are fundamental in RNA secondary structure and are critical for the proper translation of the genetic code.

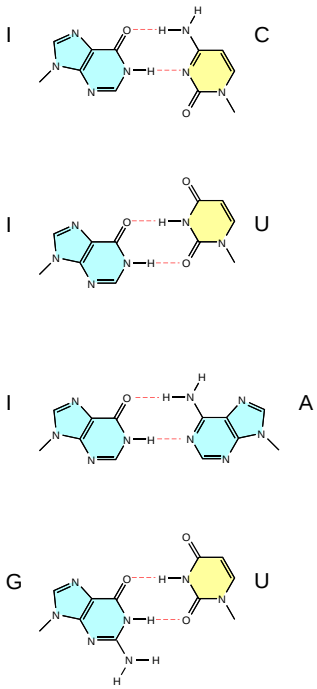


Figure 26.1.9: The Four Main Wobble Base Pairs. The top three diagrams represent the flexibility of the inosine nucleotide to form hydrogen bond interactions with cytosine, uracil, and adenine. tRNA and mRNA wobble base pairing also includes guanine hydrogen binding with uracil. Figure from Fdardel

The *wobble base* position is usually the first position of the anticodon (read in the 5′ – 3′ direction), which aligns with the 3rd position of the mRNA codon. This helps to explain the *degeneracy* found within the genetic code as shown in Figure 3 above and Figure 26.1.10 *Degeneracy* means that a single tRNA can recognize multiple different codons within mRNA.

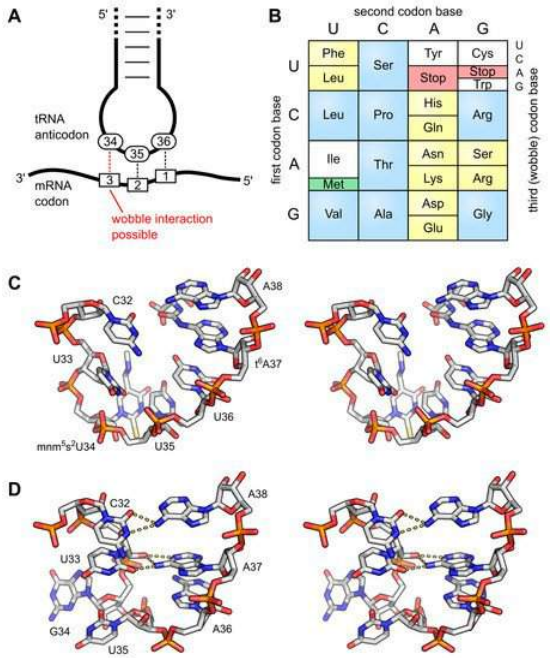


Figure 26.1.10: Anticodon Loop Structure and Codon Degeneracy. Lorenz, C., et. al. (2017) Biomolecules 7(2):35

Panel (A) shows the interaction of the anticodon bases (34–36) of a tRNA with the corresponding bases of the mRNA codons (3, 2, 1). A wobble interaction is possible between codon base 3 and anticodon base 34. The latter is frequently modified and directs the wobble interactions with the third codon base;

Panel (B) shows the standard genetic code is illustrated as a simple decoding table, 2-fold degenerate codon boxes are colored yellow, and 4-fold degenerate boxes are blue. Start and stop codons are colored green and red, respectively;

Panel (C) shows a stereo image of the well-structured anticodon loop of $tRNA^{Lys}$ from E. coli. Modifications mnm^5s^2U34 and t^6A37 prevent wrong base pairing inside the 7-nucleotide loop and promote the formation of the conserved U-turn motif. The stacked anticodon bases are located on the same side of the loop. PDB: 1FL8;

Panel (D) shows a stereo image of a collapsed and unmodified anticodon loop of $tRNA^{Tyr}$ from Bacillus subtilis. Here, bases 32 and 38 as well as 33 and 37 interact with each other and the U-turn motif is missing. The anticodon bases are not ordered and are on opposite sides of the loop. PDB: 2LAC.

A prominent example is $tRNA^{Ile}$ carrying the anticodon UAU. In principle, this anticodon can read codons AUA (for isoleucine) and AUG (for methionine). Yet, it was shown in some instances that $tRNA^{Ile}$ with unmodified UAU anticodon exists, but has a strong preference for its cognate AUA codon, while it rarely misreads AUG. In most organisms, however, $tRNA^{Ile}$ carries the anticodon CUA. To avoid misreading of the methionine codon by this tRNA, C34 (position 1 of the anticodon) is modified to lysidine (k^2C34, with the chemical structure shown in Figure 27.1.10, which restricts codon recognition to only AUA and thereby changes the amino acid identity of the tRNA from methionine to isoleucine. In the archaeal species, *Haloarcula marismortui*, *Methanococcus maripaludis, and Sulfolobus solfataricus*, this $tRNA^{Ile}$ carries a different modification at C34, fulfilling the same purpose of restricting the interaction to AUA codons. Here, the original cytosine is modified at the C2-oxo position, which is replaced by agmatine (decarboxy-arginine), resulting in agmatidine (C+ or agm^2C), as shown in Figure 26.1.11. A complimentary modification is that of N^4-acetylcytosine (ac^4C34, whose chemical structure is shown in Figure 27.1.8) in the elongator-$tRNA^{Met}$ of *E. coli*, which prevents the recognition of the AUA isoleucine codon. In non-plant mitochondria, however, both AUG and AUA codons are read as methionine. Hence, mitochondrial $tRNA^{Met}$ (carrying the anticodon CAU) has to recognize both codon forms. This is achieved by the introduction of 5-formylcytidine (f^5C, Figure 11, at position 34, a modification that pairs with both A and U residues at the corresponding codon position 3.

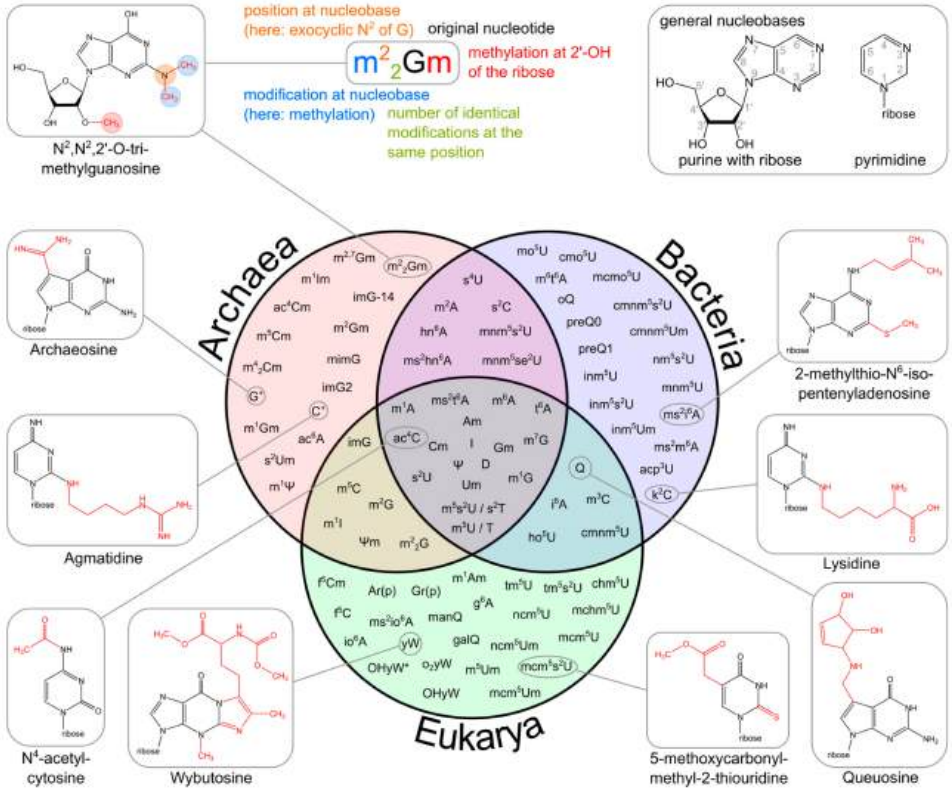


Figure 26.1.11: Variability of tRNA Modifications. Lorenz, C., et. al. (2017) Biomolecules 7(2):35

The upper part of the image illustrates the systematic abbreviation of RNA modifications with N^2,N^2,2′-O-trimethylguanosine ($m^2{}_2Gm$) as an example and also shows the atom numbering in the purine and pyrimidine rings as well as in the ribose. An abbreviation in front of the base letter describes a base modification, whereas letters after the base stand for ribose alterations. Superscripted numbers specify the position at the base and subscripted numbers indicate the frequency of identical modification at

the same position. Abbreviations are as follows: ac—acetyl, acp—aminocarboxypropyl, chm—carboxyhydroxymethyl, cmo—oxyacetic acid, cmnm—carboxymethylaminomethyl, f—formyl, g—glycinyl, gal—galactosyl, hn—hydroxynorvalylcarbamoyl, ho—hydroxy, i—isopentenyl, inm—isopentenylaminomethyl, io—cis-hydroxyisopentenyl, m—methyl, man—mannosyl, mchm—carboxyhydroxymethyl methyl ester, mcm—methoxycarbonylmethyl, mcmo—oxyacetic acid methyl ester, mnm—methylaminomethyl, mo—methoxy, ncm—carbamoylmethyl, nm—aminomethyl, r(p) —5-O-phosphono-b-d-ribofuranosyl, s—thio, se—seleno, t—threonylcarbamoyl, tm—taurinomethyl. The Venn diagram summarizes data collected from the RNA modification database and contains the 93 post-transcriptional modifications that are found in tRNAs. Some examples mentioned throughout the text are shown with their chemical structure.

Organisms vary in the number of tRNA genes in their genome. For example, the nematode worm *C. elegans*, a commonly used model organism in genetics studies, has 29,647 genes in its nuclear genome, of which 620 code for tRNA.The budding yeast *Saccharomyces cerevisiae* has 275 tRNA genes in its genome.

The human genome has approximately 20,848 protein-coding genes, of which there are 497 nuclear genes encoding cytoplasmic tRNA molecules, and 324 tRNA-derived pseudogenes (tRNA genes thought to be no longer functional). Regions in nuclear chromosomes, very similar in sequence to mitochondrial tRNA genes, have also been identified (tRNA-lookalikes). These tRNA-lookalikes are also considered part of the nuclear mitochondrial DNA (genes transferred from the mitochondria to the nucleus).

As with all eukaryotes, there are 22 mitochondrial tRNA genes in humans. Mutations in some of these genes have been associated with severe diseases like the MELAS syndrome.

Cytoplasmic tRNA genes can be grouped into 49 families according to their anticodon features. These genes are found on all chromosomes, except the 22 and the Y chromosomes. High clustering on 6p is observed (140 tRNA genes), as well as on chromosome 1. Currently, it is unclear why there is so much redundancy within the genome to decode 61 of the 64 possible codons (the other three are stop codons used to terminate translation).

26.1.3: Aminoacyl tRNA Synthetases

Aminoacyl-tRNA synthetases (aaRSs) are universally distributed enzymes that catalyze the esterification of a tRNA to its cognate amino acid (i.e., the amino acid corresponding to the anticodon triplet of the tRNA according to the genetic code). The product of this reaction, an aminoacyl-tRNA (aa-tRNA), is delivered by elongation factors to the ribosome to take part in protein synthesis.

Aminoacyl-tRNA synthetases are named after the aminoacyl-tRNA product generated, as such, methionyl-tRNA synthetase (abbreviated as MetRS) charges tRNAMet with methionine. In eukaryotes, an alternative nomenclature is often employed using the one-letter code of the amino acid (MARS), and a number is added to refer to the cytosolic (MARS1) or the mitochondrial (MARS2) variants. A total of 23 aaRSs have been described so far, one for each of the 20 proteinogenic amino acids (except for lysine, for which there are two) plus pyrrolysyl-tRNA synthetase (PylRS) and phosphoseryl-tRNA synthetase (SepRS), enzymes with a more restricted distribution that are only found in some bacterial and archaeal genomes. It is also worth noting that in eukaryotes the protein synthesis machinery of mitochondria and chloroplasts generally utilize their own, bacterial-like sets of synthetases and tRNAs that are distinct from their cytosolic counterparts.

The aminoacyl-tRNA synthetases catalyze a two-step reaction that leads to the esterification of an amino acid to the 3’ end of a tRNA along with the hydrolysis of one molecule of ATP, yielding aminoacyl-tRNA, AMP, and PPi. In the first step, termed amino acid activation, both the amino acid and ATP bind to the catalytic site of the enzyme, triggering a nucleophilic attack of the α-carboxylate oxygen of the amino acid to the α-phosphate group of the ATP, condensing into aminoacyl-adenylate (aa-AMP), which remains bound to the enzyme, and PPi, which is expelled from the active site, as shown in Figure 26.1.12

Figure 26.1.12: The Aminoacylation Reaction**.** In the first step, the amino acid (blue) is activated with ATP (red) in the synthetase active site (not depicted), forming the aminoacyl-AMP and releasing PPi. The amino acid is transferred to the tRNA (green) and AMP is released (depicted in the image transfer to the 3'-OH characteristic of class I aaRS, while in class II transfer happens at the 2'-OH with a 3'-OH attack in the second step). Gomez, M.A.R., and Ibba M. (2020) RNA, doi: 10.1261/rna.071720.119

Although tRNA is usually not required for this first step, certain synthetases do require the tRNA species for productive amino acid activation. In the second part of the reaction, either the 2′- or 3′-hydroxyl group of the terminal adenine nucleotide attacks the carbonyl carbon of the adenylate, forming aminoacyl-tRNA and AMP (Figure 27.1.12 B). While the two-step aminoacylation reaction is universally conserved, the aaRSs that catalyze it show extensive structural, and in some instances functional, diversity.

The 23 known aaRSs can be divided into two major classes based on the architecture of their active sites (Class I and Class II). In class I synthetases, the catalytic domain bears a dinucleotide or Rossman fold (RF) featuring a five-stranded parallel β-sheet connected by α-helices and is usually located at or near the N-terminus of the protein. This RF contains the highly conserved motifs HIGH and KMSKS, separated by a connecting domain termed connective peptide 1 (CP1), as shown in Figure 26.1.13(panel A). Class II active site architecture is organized as seven-stranded β-sheets flanked by α-helices and features three motifs that show a lesser degree of conservation than those in class I (panel B). Both classes also exhibit pronounced differences in their modes of substrate binding. Class I aaRSs bind the minor groove of the tRNA acceptor stem (with the exceptions of TrpRS and TyrRS) and aminoacylate the 2'-OH of the ribose of the terminal adenosine, while class II approach tRNA from the major groove and transfer amino acid to the 3'-OH (except PheRS). The mode of ATP binding is also different between both classes, being bound in an extended configuration in class I, while class II binds a bent configuration with the γ-phosphate folding back over the adenine ring. The kinetics of the aminoacylation reaction can also be used as a distinctive mechanistic signature, as aminoacyl-tRNA release is the rate-limiting step for class I enzymes (except for IleRS and some GluRS) while for class II it is the amino acid activation rate instead.

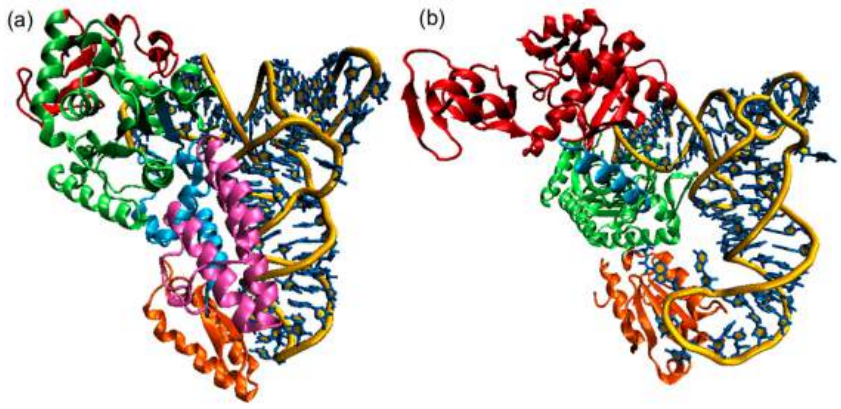
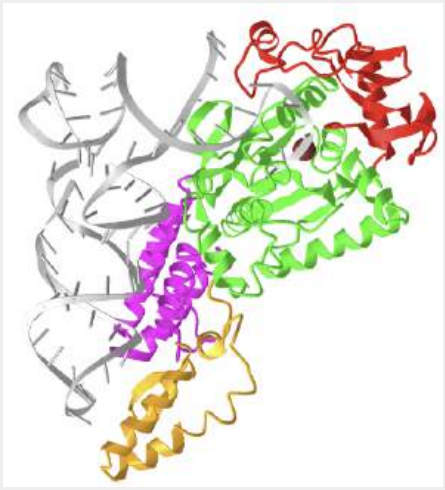
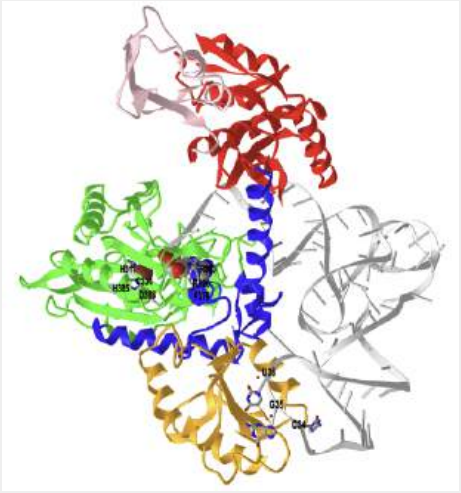


Figure 26.1.13: Structures of Class I and Class II AARSs. Li, R., et. al. (2015) Int J. Mol. Sci. 16(7):15872-15902.

Panel (a) shows the *E. coli*CysRS:tRNA^{Cys} complex. The CP domain (red) and Rossmann fold catalytic domain (green), stem contact fold (cyan), helical bundle domain (magenta), and anticodon binding domain (orange) of CysRS are shown in a ribbon diagram;

Panel (b) A single monomer of the homodimeric *E. coli* ThrRS:tRNA^{Thr} complex. The two N-terminal domains (red), catalytic domain (green), linker (cyan), and anticodon binding domain (orange) of ThrRS are shown in a ribbon diagram. For both structures, the tRNAs are shown in a stick diagram (blue) with a trace of their backbone (yellow).

Figure 26.1.14shows an interactive iCn3D model of the Class I and II aminoacyl-tRNA synthetases

Class I E. Coli Cysteinyl-tRNA synthetase -tRNA(Cys) (1U0B)	class II E. coli threonyl-tRNA synthase - tRNA(Thr) (1QF6)
	
(Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...P4oRWxKvJHFDL6	(Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...6aXwfqXGCpD7c6

The color coding is analogous to Figure 26.1.13with the tRNA shown in gray.

For the Class II E. coli threonyl-tRNA synthase - tRNA(Thr), key conserved amino acids in the active site are shown in CPK-colored sticks and labeled. The anticodon in the tRNA is shown in colored sticks as well. Note the large conformational change in the anticodon loop.

The E. coli threonyl-tRNA synthetase is interesting in that it represses the translation of its own mRNA. A Zn^{2+} ion is involved in binding specificity for the amino acid. It is coordinated by H385, H511, and a water molecule (not shown). R363 interacts with the alpha phosphate, while F379 and R520 align on both sides of the adenine rig. D383 interacts with the amine group of the substrate threonine.

To ensure the faithful translation of the genetic message, synthetases must identify and pair particular tRNAs with their cognate amino acid which relies on the proper recognition of both substrates. This can prove extremely challenging for the synthetases as

not only have they to discriminate the correct tRNA isoacceptor amongst a set of other tRNAs very similar in structure and chemical composition but also be able to select the cognate amino acid amidst an extremely large pool of similar amino acids, both proteinogenic and non-proteinogenic. The evolutionary pressure to maintain fidelity has driven aaRSs to develop an elevated specificity for their substrates, both the tRNA and the amino acid.

In addition, some synthetases have evolved editing activities that specifically target and hydrolyze misactivated amino acids and/or misacylated tRNAs. To date, editing activity has been described in 10 out of the 23 aaRSs. In class I synthetases, this activity is located in the highly conserved CP1 domain, although in some enzymes such as MetRS and LysRS the editing activity resides in the catalytic site. In class II synthetases, however, the editing domains are more idiosyncratic. Editing mechanisms can be divided into two categories, *pre- or post-transfer editing*, with regard to the editing taking place before or after the transfer of the amino acid to the tRNA, as shown in Figure 26.1.15

Pre-transfer editing has been described in both class I and class II aaRSs and takes place after aa-AMP synthesis but before the aminoacyl moiety is transferred to the tRNA. Although the tRNA does not participate in the reaction itself, it has been reported that tRNA binding promotes editing activity in some aaRSs and is a requirement in IleRS and LeuRS. Pre-transfer editing can follow two main pathways. The first one is the selective release of the aa-AMP to the cytosol, where the labile phosphoester bond is spontaneously hydrolyzed. The second route involves the enzymatic breakdown of the product and may happen either in the active site or in an independent editing site.

Post-transfer editing takes place after the transfer of the amino acid to the tRNA and involves the hydrolysis of the ester bond, in a domain separated from the active site. The specific mechanism of editing is idiosyncratic to each synthetase but in general, once formed the aa-tRNA triggers a conformational change, and the 3’ terminus containing the aa is translocated from the active site to the editing site, sometimes traversing distances as large as 40 Å. As the core of the tRNA remains bound to the enzyme, this translocation often involves a rearrangement of the 3’ terminus to relocate to the editing site.

Figure 26.1.15: Aminoacyl tRNA Synthetase Editing Pathways**.** Schematic overview of the editing pathways employed by the synthetases. In the figure above, the events are in italics, while the editing paths are in bold. The pathways are divided between pre-transfer and post-transfer pathways. In the *pre-transfer editing*, the activated non-cognate aminoacyl-adenylate may be released from the enzyme and hydrolyze spontaneously or edited within the active site or a specialized active site. During *post-transfer editing*, the aminoacyl-tRNA can be translocated to the editing site or released and cleared by a dedicated trans-editing factor. When a correct cognate has formed, the resulting aminoacyl-tRNA binds the elongation factors and proceeds to translation in the ribosome. Figure from: Gomez, M.A.R., and Ibba M. (2020) RNA, doi: 10.1261/rna.071720.119

26.1.4: Ribosome Structure

The ribosome is a highly conserved molecular machine. In all organisms, it is composed of two unequal subunits, which consist of a distinct set of ribosomal RNA (rRNA) and ribosomal proteins (RPs) that combine to form a large nucleoprotein complex. The ribosome structures in all living organisms harbor three different tRNA binding sites: The A-site, where decoding occurs and the correct aminoacyl-tRNA (aa-tRNA) is selected based on the mRNA codon displayed, the P-site, which carries the peptidyl-tRNA,

and the E-site, which binds exclusively deacetylated tRNAs that are exiting the ribosome. Thus, during translation the tRNA moves from the A-site through the P- and E-site, where it leaves the ribosome, as shown in Figure 26.1.16

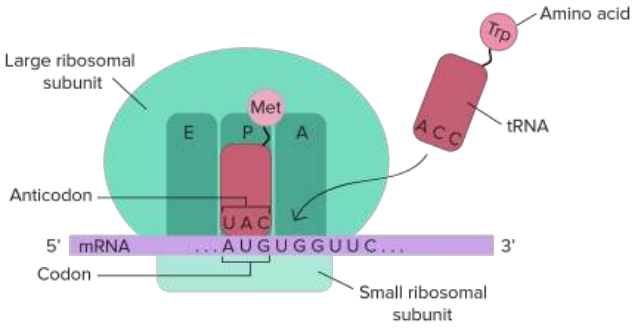


Figure 26.1.16: **Schematic Structure of an Active Ribosome.** The Kahn Academy where it was modified from Openstax College Biology

The mRNA (shown in purple) is assembled between the small subunit and the large subunit of the ribosome (shown in green). tRNA molecules (shown in red) that are loaded with their cognate amino acid (shown in pink) are transitioned through the A-P-E sites of the ribosome during the elongation phase of translation. Movement of the tRNA molecules also shifts the position of the mRNA causing the next three codon bases to line up in the A-site of the ribosome.

The catalytic peptidyl transferase activity occurs when the tRNA molecules are bound in the A- and P-sites, transferring the nascent peptide to the incoming tRNA molecule (Fig. 27.1.15). Ribosomes are ribozymes because the catalytic peptidyl transferase activity that links amino acids together is performed by the rRNA.the complexity of the ribosome structure is reflected in the process of protein synthesis, which can be intersected into three major steps: initiation, elongation, and termination/recycling.

Ribosomes are either free-floating in the cytoplasm or they can be associated with the intracellular membranes that make up the rough endoplasmic reticulum (rER). Proteins translated into the rER will often be transported out of the cell or embedded into the plasma membrane. These processes are illustrated in Figure 26.1.17.

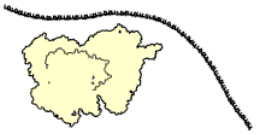


Figure 26.1.17: **A ribosome translating a protein that is secreted into the endoplasmic reticulum.** Figure from: *Bensaccount*

Ribosomes from bacteria, archaea, and eukaryotes in the three-domain system resemble each other to a remarkable degree. They differ in their size, much of the rRNA sequence, and the ratio of protein to RNA. Figure 26.1.18shows the eukaryotic rRNA from the large subunit of the ribosome with highly conserved nucleotide elements (>90% sequence identity) within all of the domains of life, termed universal CNEs or uCNEs indicated. The differences in sequence and structure between eukaryotes and prokaryotes allow some antibiotics to kill bacteria by inhibiting their ribosomes while leaving human ribosomes unaffected.

Figure 26.1.18: **Universally Conserved Nucleotide Elements (uCNEs) in the Large Subunit of the Eukaryotic Ribosome.** rRNA from the yeast large subunit of the ribosome is shown from the crown view (from the subunit interface) with the L1 stalk at the upper left. uCNEs that are ≥90% conserved in sequence in all domains of life are indicated in dark blue. Figure from: Doris, S.M., et. al. (2015) RNA 21:1719-1730

In all species, more than one ribosome may move along a single mRNA chain at one time (as a *polysome*), each "reading" its sequence and producing a corresponding protein molecule. In this way, many proteins can be translated from a single mRNA molecule. Within bacteria, translation is also coupled with transcription, as the two processes are not physically separated from one another. This is illustrated in Figure 26.1.19 In eukaryotic organisms, *polysomes* form during translation. However, transcription and translation are not coupled, as the processes are separated into the nucleus and cytoplasm, respectively.

Figure 26.1.19: Bacterial Transcription and Translation are Coupled. Translation of nascent mRNA in bacteria occurs before the release of the nascent mRNA from the RNA Polymerase and DNA template. Multiple ribosomes assemble on the nascent mRNA and begin the translation of multiple protein products from the newly formed mRNA template. The structure of the ribosomes assembled on the mRNA is referred to as a *polysome*. Figure from: Gonzalezmg1

The mitochondrial ribosomes of eukaryotic cells functionally resemble many features of those in bacteria, reflecting the likely evolutionary origin of mitochondria.

26.1.4.1: Prokaryotic Ribosome Structure

Prokaryotic ribosomes have a mass of about 2500 kDa and a size of 70S (or Svedberg units: A Svedberg unit is a measure of the sedimentation rate in a centrifuge and thus is representative of size). A complete ribosome (70S) can be dissociated into a large subunit (50S) and a small subunit (30S), as shown in Figure 26.1.20 The small subunit is formed by the interactions of 21 different proteins and a 16S RNA molecule, whereas the large subunit contains 34 different proteins and two RNA molecules, a 23S, and a 5S species.

The rate-limiting step in protein synthesis is the formation of the 70S initiation complex which will be discussed in detail in the next section.

Figure 26.1.20: The Small and Large Prokaryotic Ribosomal Subunits*(Left)* Atomic structure of the 30S subunit from *Thermus thermophilus*. Proteins are shown in blue and the single RNA chain is in orange. *(Right)* Atomic structure of the 50S subunit from *Haloarcula marismortui*. Proteins are shown in blue and the two RNA chains are in orange and yellow. The small patch of green in the center of the subunit is the active site where the peptidyl transfer reaction occurs. Figure from: Goodsell, D. (2000) Molecule of the Month

Figure 26.1.21shows an interactive iCn3D model of the E. Coli ribosome (7K00).

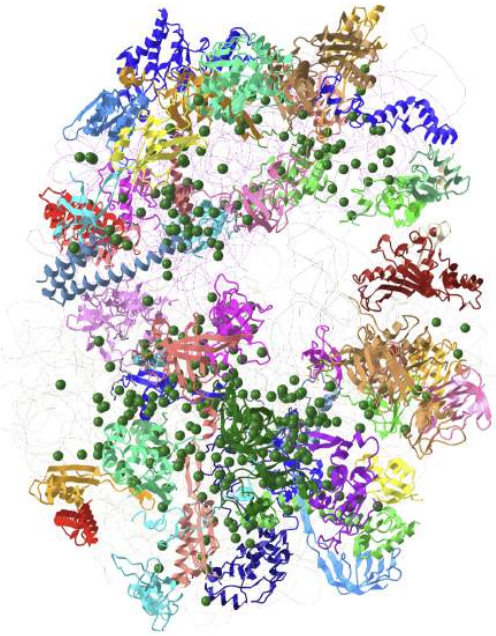


Figure 26.1.21: E. Coli ribosome (7K00). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?51UCkQmuV3BBqeEJ6

The RNA is shown in a faint trace backbone. The proteins are shown as cartoons with different colors. Note the large number of Mg^{2+} ions.

The structure was determined using cryo-EM at high resolution. Some ribosomal proteins have isopeptide bonds (using side chain amine and carboxyl groups) as well as some thioamide backbone replacements for the usual amide links.

26.1.4.2: Eukaryotic Ribosome Structure

Eukaryotic ribosomes are larger than their prokaryotic counterparts at approximately 80S (although there is some modest variation between eukaryotic species). Human cytosolic ribosomes are composed of a large subunit (60S) that contains the 28S, 5.8S, and 5S rRNAs and 47 ribosomal proteins (RPs), and a small subunit (40S) that contains the 18S rRNA and 33 RPs.

The assembly of eukaryotic ribosomal subunits starts in the nucleolus, where RNA polymerase I transcribes the major rRNA precursor (a 45S pre-RNA), from which, after processing and removal of the external and internal transcribed spacers (ETS and ITS), the mature rRNAs are generated, as shown in Figure 26.1.22 The pre-RNA is modified during transcription by small nucleolar ribonucleoproteins (snoRNPs), processed by RNA nucleases, and assembled with numerous RPs. After processing the rRNA precursor, the pre-40S and pre-60S subunits follow separate biogenesis routes. Here we will describe the assembly of the 60S subunit in more detail.

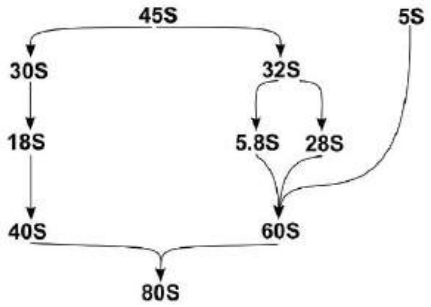


Figure 26.1.22: Schematic Representation of Ribosomal RNA Processing in Eukaryotes. The 45S rRNA transcript is processed by endonucleotide cleavage and splicing to form the 5.8S and 28S components of the large ribosomal subunit and the 18S component of the small ribosomal subunit. The 5S subunit is transcribed independently from the other rRNA components and is incorporated into the large ribosomal subunit. Figure from FeatherPluma

Although the exact assembly of the 60S subunit is not currently known, a model has been postulated that suggests that in the nucleolus, after circularization of rRNA domains, early 60S assembly is carried out sequentially, as shown in Figure 26.1.23 As the transcription of the pre-rRNA proceeds, the rRNA quickly develops a core secondary structure that promotes the interaction of key Assembly Factors (AFs) and RPs before transcriptional termination. Specifically, during this time, the 5.8S portion, ITS2, domains I and II, and partially domain VI are folded in the earliest observed intermediate (state A in Fig. 23). Thus, it appears that the solvent-exposed back side of the large subunit forms like an exoskeleton and construction proceeds by formation of the exit tunnel. This model agrees with a previously suggested model of hierarchical folding based on RP depletion phenotypes. The peptidyl transferase center (PET) is predicted to be one of the later folding steps in the process (state F in Fig. Fig 23) Although the very late-folding peptidyl transferase center is the evolutionarily oldest part of the ribosome, it is likely that folding the solvent side first brings the advantage of providing a stable scaffold for the developing 60S subunit. The folding and assembly of the 40S subunit follow a similar pattern. The two subunits remain unattached until activated in the cytoplasm through the binding of a mRNA transcript with the small subunit. This begins the formation of the initiation complex that will mark the start of the transcriptional process.

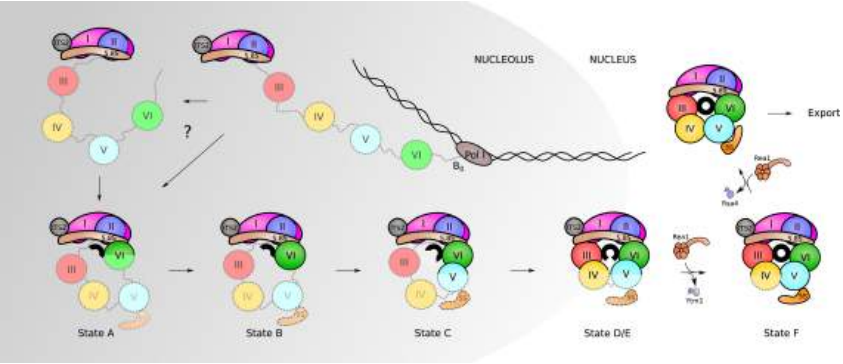


Figure 26.1.23: Assembly Sequence of the Pre-rRNA Domains in the Large Subunit of a Eukaryotic Ribosome. Kater, L., et. al. (2017) Cell 171(7):1599-1610

Assembly of RPs and AFs to the nascent 35S rRNA precursor starts co-transcriptionally. Very early, the pre-rRNA is circularized as domain VI binds to domains I and II and the 5.8S portion of the precursor rRNA. The formation of the PET (displayed here as a black circle) starts with this circularization. Its maturation progresses as rRNA domains fold following this order: VI, V, III, and IV. Full assembly of the PET is only achieved when domain V is completely folded as observed in state F. After that, only a few additional steps need to occur before the particles are exported to the cytoplasm, where they undergo final maturation.

26.1.5: References

This chapter was remixed and adapted from the following resources under creative commons licensing:

1. Wikipedia contributors. (2020, June 27). Wobble base pair. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:47, August 12, 2020, from https://en.Wikipedia.org/w/index.php?title=Wobble_base_pair&oldid=964760055
2. Lorenz, C., Lünse, C., and Mörl, M. (2017) tRNA modifications: Impact on the structure and thermal adaptation. *Biomolecules* 7(2)35. Available at: https://www.mdpi.com/2218-273X/7/2/35/htm
3. Pan, T. (2018) Modifications and functional genomics of human transfer RNA. *Cell Research* 28:395-404. Available at: https://www.nature.com/articles/s41422-018-0013-y#Sec3
4. Bednárová, A., Hanna, M., Durham, I., Van Cleave, T., England, A., Chaudhuri, A., and Krishnan, N. (2017) Lost in translation: Defects in transfer RNA modifications and neurological disorders. *Front. Mol Neurosci.* 10:135. Available at: https://www.researchgate.net/publication/316440980_Lost_in_Translation_Defects_in_Transfer_RNA_Modifications_and_Neurological_Disorders
5. Wikipedia contributors. (2020, May 17). Transfer RNA. In *Wikipedia, The Free Encyclopedia*. Retrieved 23:03, August 15, 2020, from https://en.Wikipedia.org/w/index.php?title=Transfer_RNA&oldid=957227343
6. Gomez, M.A.R., and Ibba M. (2020) Aminoacyl-tRNA Synthetases. *RNA*, doi: 10.1261/rna.071720.119 Available at: https://rnajournal.cshlp.org/content/early/2020/04/17/rna.071720.119.abstract
7. Li, R., Macnamara, L.M., Leuchter, J.D., Alexander, R.W., and Cho, S.S. (2015) MD Simulations of tRNA and Aminoacyl-tRNA Syntetases: Dynamics, Folding, Binding, and Allostery. *Int. J. Mol Sci.* 16(7):15872-15902. Available at: https://www.mdpi.com/1422-0067/16/7/15872

8. Wikipedia contributors. (2020, August 15). Ribosome. In *Wikipedia, The Free Encyclopedia*. Retrieved 05:22, August 16, 2020, from https://en.Wikipedia.org/w/index.php?title=Ribosome&oldid=973144885
9. Kater, L., Thoms, M., Barrio-Garcia, C., Cheng, J., Ismail, S., Ahmed, Y.L., Bange, G., Kressler, D., Berninghausen, O., Sinning, I., Hurt, E., and Beckmann, R. (2017) Visualizing the assembly pathway of nucleolar Pre-60S Ribosomes. Cell 171(7):1599-1610. Available at: https://www.sciencedirect.com/science/article/pii/S0092867417314290
10. Bock, L.V., Kolár, M.H., Grubmüller, H. (2018) Molecular simulations of the ribosome and associated translation factors. *Cur. Op. Struc. Bio.* 49:27-35. Available at: https://www.sciencedirect.com/science/article/pii/S0959440X1730132X
11. Doris, S.M., Smith, D.R., Beamesderfer, J.N., Raphael, B.J., Nathanson, J.A., and Gerbi, S.A. (2015) Universal and domain-specific sequences in 23S-28S ribosomal RNA identified by computational phylogenetics. *RNA* 21:1719-1730. Available at: https://www.researchgate.net/publication/281141702_Universal_and_domain-specific_sequences_in_23S-28S_ribosomal_RNA_identified_by_computational_phylogenetics
12. Aleksashin, M.A., Leppik, M., Hochenberry, A.J., Klepacki, D., Vázquez-Laslop, N., Jewett, M.C., Remme, J., and Mankin A.S. (2019) Assembly and functionality of the ribosome with tethered subunits. *Nature Communications* 10:930. Available at: https://www.nature.com/articles/s41467-019-08892-w#rightslink
13. Wikipedia contributors. (2020, July 3). Formylation. In *Wikipedia, The Free Encyclopedia*. Retrieved 22:03, August 16, 2020, from https://en.Wikipedia.org/w/index.php?title=Formylation&oldid=965827476
14. Gualerzi, C.O., and Pon C.L. (2015) Initiation of mRNA translation in bacteria: structural and dynamic aspects. *Cell Mol Life Sci.* 72:4341-4367. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4611024/
15. Wikipedia contributors. (2020, June 14). Kozak consensus sequence. In *Wikipedia, The Free Encyclopedia*. Retrieved 06:03, August 18, 2020, from https://en.Wikipedia.org/w/index.php?title=Kozak_consensus_sequence&oldid=962444929
16. Knight, J.R.P., Garland, G., Pöyry, T., Mead, E. Vlahov, N., Sfakianos, A., Frosso, S., De-Lima-Hedayioglu, F., Mallucci, G.R., von der Haar, T., Smales, C.M., Sansom, O.J., and Willis, A.E. (2020) Control of translation elongation in health and disease. *Dis. Mod. and Mech.* 13: dmm043208. Available at: https://dmm.biologists.org/content/13/3/dmm043208
17. Adio, S., Sharma, H., Senyushkina, T., Karki, P., Maracci, C., Wohlgemuth, I., Holtkamp, W., Peske, R., and Rodina, M.V. (2018) Dynamics of ribosomes and release factors during translation termination in *E. coli*. *eLife* 7:e34253. Available at: https://elifesciences.org/articles/34252
18. Ge, X., Oliveira, A., Hjort, K., Bergfors, T., Guitiérrez-de-Terán, H., Andersson, D.I., Sanyal, S., and Åqvist, J. (2019) Inhibition of translation termination by small molecules targeting ribosomal release factors. *Scientific Reports* 9: 15424. Available at: https://www.nature.com/articles/s41598-019-51977-1#rightslink
19. Svidritskiy, E., Demo G., Loveland A.B., Xu, C., and Korosteleve, A.A. (2019) Extensive ribosome and RF2 rearrangements during translation termination. *eLife* 8:e46850. Available at: https://elifesciences.org/articles/46850
20. Sauert, M., Temmel, H., and Moll, I. (2015) Heterogeneity of the translational machinery: Variations on a common theme. Biochimie 114:39-47. Available at: https://www.sciencedirect.com/science/article/pii/S0300908414003952

26.2: Protein Synthesis

26.2.1: Prokaryotic Initiation

The small subunit of the ribosome (the 30S) interprets the genetic information by selecting aminoacyl-tRNAs cognate to the mRNA codons in the decoding center. The large subunit (the 50S) carries the catalytic peptidyl transferase center where amino acids are polymerized into a protein. Small and large subunits unite together at the start codon of a gene to form the 70S ribosome and dissociate again at the stop codon upon completing the synthesis of the encoded protein. This process consists of three phases: initiation, elongation, and termination. In this section, we will focus on the initiation of translation.

In bacteria, the initiation phase of protein synthesis involves a limited number of “actors”. Aside from the two ribosomal subunits, key roles are played by the initiator $tRNA^{fmet}$, the translation initiation region (TIR) of the mRNA, and three protein factors – the initiation factors (IFs) IF1, IF2, and IF3 – that ensure speed and accuracy to the overall process. The initiator $tRNA^{fmet}$ contains a methionine residue that has been enzymatically modified to contain an N-terminal formyl group, as shown in Figure 26.2.1. fMet is only used for the initiation of protein synthesis and is thus found only at the N-terminus of the protein. Unmodified methionine is used during the rest translation. Once protein synthesis is completed, the formyl group on methionine may be removed by peptide deformylase, and on occasion, the entire methionine residue can be further removed by the enzyme methionine aminopeptidase.

Figure 26.2.1: Structure of N-formylmethionine.

The TIR sequence within the mRNA contains the start codon and usually an upstream untranslated region that interacts with the small subunit of the ribosome. The bacterial cell produces and expresses a plethora of different mRNAs with different TIR sequences and structures; the efficiency by which these individual transcripts are translated depends not only upon their abundance and stability but also upon the nature of TIR. Thus, unlike the other aforementioned actors that represent constants, the mRNA TIRs represent essentially the only variable in the process of mRNA initiation site selection and can affect translation efficiency.

Although the triplet AUG is by far the most frequent initiation codon found in TIRs, other initiation triplets (i.e., GUG, UUG, AUU, AUC, and AUA) are found in bacteria and the central U is the only universally conserved base of the start codon. Among the aforementioned triplets, those having a 3′-G (i.e., AUG, GUG, and UUG) are recognized equivalently and most efficiently by IF3 during the initiation complex formation.

Another important characteristic of a large number of bacterial mRNA TIRs is the presence of a Shine–Dalgarno (or SD) sequence that is complementary to the 3′ end sequence of 16S rRNA (the anti-SD sequence or aSD). The SD sequence, when present, is usually at an optimal distance of 4–9 nucleotides upstream of the initiation codon, as shown in Figure 26.2.2 While the SD sequence plays an important role in the efficient translation of many mRNA transcripts, it is not essential. Many other mRNA sequences fully lack an SD sequence but are still efficiently transcribed. Thus, the SD sequence is only one example of TIR mechanisms that can play an important role in mRNA binding with the small subunit of the ribosome.

Figure 26.2.2: Shine-Dalgarno Translation Initiation Sequence. Figure modified from *Shakeistone*

Prokaryotic mRNA sequences often share a highly conserved sequence upstream of the start codon known as the Shine-Dalgarno sequence. This consensus sequence is complimentary to the 3′-end of the 16S rRNA sequence in the small subunit of the ribosome.

It is an important feature for the binding and docking of many mRNA molecules with the small ribosomal subunit during transcription initiation.

The three protein initiation factors, IF1, IF2, and IF3, determine the kinetics and fidelity of the overall initiation process. The three IFs are bound, one copy each, to specific sites of the 30S subunit where they assist with the formation of the initiation complex and assembly of the 70S ribosome.

As noted above, the initiator tRNA is first aminoacylated with methionine whose α-NH_2 group is eventually blocked by a specific formyl transferase (TMF) to produce a $tRNA^{fmet}$ molecule. This modification prevents interaction with the elongation factor EF-Tu (which we will see plays an important role in the elongation phase of translation, but not the initiation phase!) Blocking EF-Tu binding ensures instead the recognition and binding of $tRNA^{fmet}$ by initiation factor IF2, effectively docking it with the 30S subunit. Furthermore, $tRNA^{fmet}$ binds with high affinity to the ribosomal P-site, unlike all other aminoacyl-tRNAs that bind to the A-site in a ternary complex with EF-Tu and GTP (details will be presented in the next section). In the P-site, the initiator tRNA must be recognized as correct by the other initiation factors IF3 and IF1.

To form the 30S initiation complex, IF3, and IF2 are the first factors to bind to the 30S subunit forming an unstable 30S-IF3-IF2 complex, as shown in Figure 26.2.3 (panel A). The binding of IF1 causes a conformational change in the 30S subunit stabilizing the complex and allowing the recruitment of the $tRNA^{fmet}$ by IF2. Notably, IF1 binds in the A site of the 30S subunit, where it contacts ribosomal protein S12. Recruitment of the $tRNA^{fmet}$ can also stabilize the mRNA interactions with the 30S subunit through the formation of hydrogen bonds between the codon of the mRNA and the anticodon of the $tRNA^{fmet}$. Note that the binding of mRNA to the 30S subunit is IF-independent and can take place at any time during the 30S assembly process. Two potential routes of mRNA association are shown Figure 26.2.3, panel A, where the mRNA is assembled either prior to or after $tRNA^{fmet}$ recruitment.

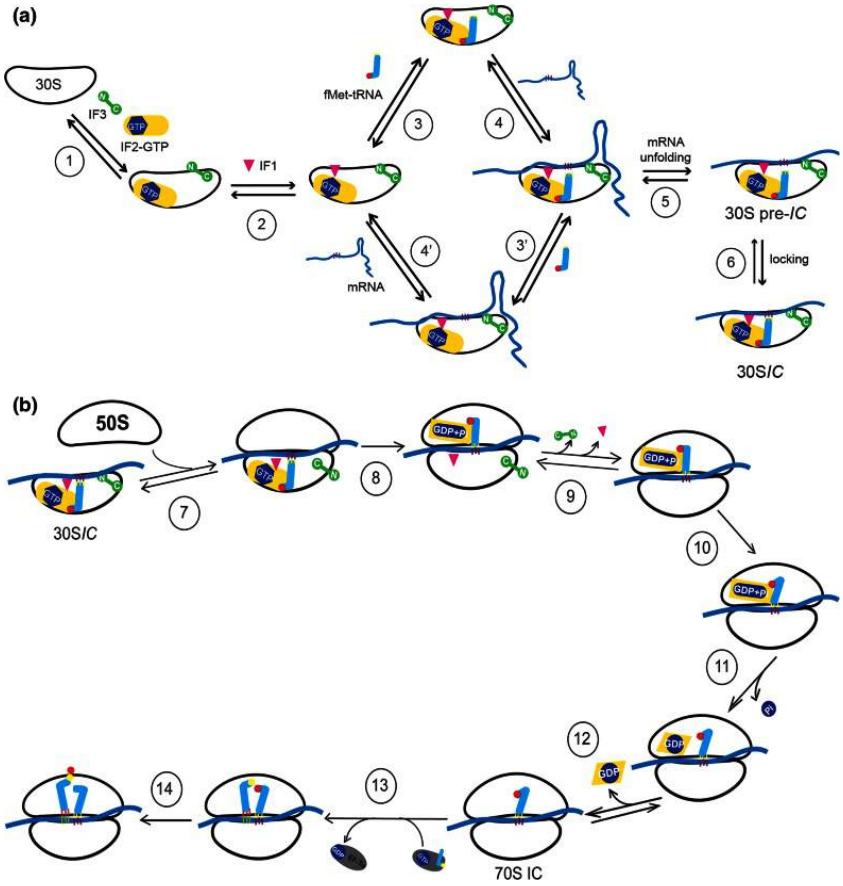


Figure 26.2.3: Initiation of Transcription. Figure from: *Gualerzi, C.O., and Pon, C.L. (2015) Cell Mol Life Sci. 72:4341-4367.*

Step 1: a vacant 30S ribosomal subunit binds IF3 and IF2. *Step 2:* IF1 binds to the 30S subunit in the presence of both IF3 and IF2.

Steps 3 and 3': in the presence of all three factors $tRNA^{fmet}$ is recruited.

Steps 4 and 4':the mRNA is bound with different on and off rates depending on its TIR structure; mRNAs with strong secondary structures are bound more slowly than those having little or no secondary structure.

Step 5: mRNAs containing secondary structures must be unfolded in a process that is facilitated by IF2 bound to GTP and antagonized by IF3.

Step 6: the isomerization of the 30S pre-*IC* allows the P-site codon–anticodon interaction to yield a more stable 30S*IC* from which mRNA and fMet-tRNA are more stably bound.

Step 7: a 30S*IC*, containing IF1, IF2·GTP, IF3 and mRNA whose initiation triplet is P-site decoded by fMet-tRNA, is docked by a 50S subunit.

Step 8: upon contact with the 50S subunit, the GTPase function of IF2 is activated and GTP is rapidly hydrolyzed leaving GDP+Pi bound to IF2.

Step 9: this reversible conformational transition represents the last kinetic checkpoint of translation initiation fidelity by IF3 and IF1, as IF3 and IF1 dissociate from the complex.

Step 10: The first-order isomerization of the IF2-GDP structure causes a shift in the ribosome structure that represents the rate-limiting step in 70S*IC* formation.

Step 11: Pi is dissociated from IF2·GDP.

Step 12: IF2 leaves the ribosome (or moves away from the A-site) clearing the way for EF-Tu binding.

Step 13: the EF-Tu·GTP·aminoacyl-tRNA complex binds to the 70S*IC* and through a number of steps (not represented here) delivers to the ribosomal A-site the aminoacyl-tRNA encoded by the second mRNA codon.

*Step 14:*the $tRNA^{fMet}$ bound in the P-site of the peptidyl transferase center donates its formyl-methionine to the A-site-bound aminoacyl-tRNA to yield the initiation dipeptide fMet-aa. Initiation is then complete and the elongation phase can begin.

Following the recruitment of the mRNA and the $tRNA^{fmet}$ to the 30S initiation complex loaded with the IF2, IF3 and IF1 initiation factors, the 50S subunit is very rapidly docked to yield an initially unstable 70S initiation complex (Fig 27.2.3 b). It should be noted that the IF2 protein is a GTP hydrolase enzyme and, as such, binds with the cofactor GTP prior to the recruitment of the 50S subunit. Contact between the IF2 GTPase activating center with the 50S subunit causes the rapid hydrolysis of GTP to GDP + Pi.

The formation of the 70S complex causes the dissociation of the initiation factors. IF2 is the last factor to be dissociated, leaving the ribosome after having positioned $tRNA^{fMet}$ in the P-site of the 70S initiation complex. It must be placed in the correct orientation to facilitate peptide bond formation. GDP and Pi also dissociate from the complex with the removal of IF2. The elongation factor, EF-G is then free to chaperone the first tRNA into the A-site and the first peptide bond is formed (Step 13 of Fig. 27.2.3 b). This marks the beginning of the elongation phase of protein synthesis.

26.2.2: Eukaryotic Initiation

Eukaryotic translation initiation is more complex than prokaryotic systems and requires the actions of at least 11 eukaryotic initiation factors (eIFs), plus additional auxiliary factors (Table 26.2.1). We will not cover the action of all these eIFs in detail here, but rather focus a few key steps as outlined in Figure 26.2.4

Table 26.2.1: **Comparison of Prokaryotic and Eukaryotic Translation Initiation Factors**

First, the initiator tRNAi is recruited to the small ribosomal subunit (40S) to form a ternary complex with the GTP-bound eukaryotic initiation factor 2 (eIF2). Formation of this 43S pre-initiation complex is strongly enhanced by additional factors, such as eIF3. eIF3 also interacts with the eIF4F complex, which consists of three other initiation factors: eIF4A, eIF4E, and eIF4G. eIF4G is a scaffolding protein that directly associates with both eIF3 and the other two components. eIF4E is the 5′-cap-binding protein. Binding of the mRNA cap by eIF4E is often considered the rate-limiting step of *cap-dependent initiation*, and the concentration of eIF4E is a regulatory nexus of translational control. Certain viruses cleave a portion of eIF4G that binds eIF4E, thus preventing cap-dependent translation to hijack the host machinery in favor of the viral (cap-independent) messages. eIF4A is an ATP-dependent RNA helicase that aids the ribosome by resolving certain secondary structures formed along the mRNA transcript. The poly(A)-binding protein (PABP) also associates with the eIF4F complex via eIF4G, and binds the poly-A tail of most eukaryotic mRNA molecules. This protein has been implicated in playing a role in circularization of the mRNA during translation.The 43S preinitiation complex accompanied by the protein factors moves along the mRNA chain toward its 3′-end, in a process known as ‘scanning’, to reach the start codon (typically AUG). After recognition of the start codon, the large ribosomal subunit (60S) assembles to form the 80S initiation complex, which is ready for elongation. Alternatively, under distinct conditions or on certain transcripts internal initiation can occur in a cap-independent manner at so called *internal ribosome entry sites (IRES).* Eukaroytic translation initiation is shown in Figure 26.2.4

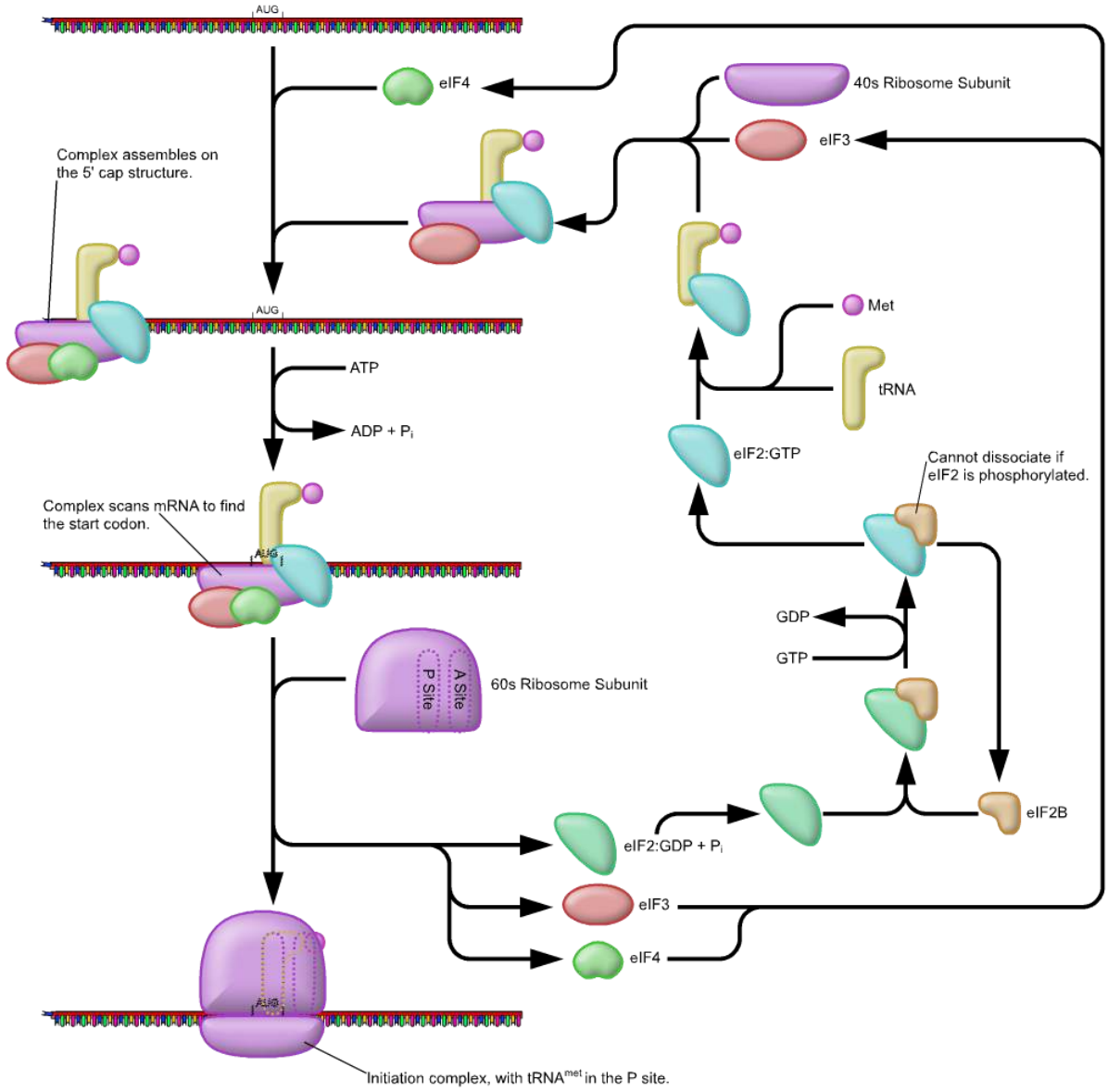


Figure 26.2.4: Eukaryotic Translation Initiation. Figure from: Eukaryotic Translation, Wikiwand

This is a simplified diagram of eukaryotic translation initiation detailing some of the eIFs involved in the process. eIF2 is critical for recruiting the initiation tRNAi to the 40S subunit. eIF3 enhances the activity of eIF2 and also promotes the binding of the 43S pre-initiation complex with the mRNA. eIF3 binds with the mRNA through the interaction of the eIF4 factors and causes the scanning of the pre-initiation complex down the mRNA to locate the start codon (usually AUG). Poly A Binding Proteins (PABPs) bind with the polyA tail sequence of the mRNA and also interact with the eIF4 factors causing the circularization of the mRNA.

As seen in prokaryotic systems with the favored Shine Dalgarno sequence upstream of the start codon within the mRNA sequence, there are also preferred nucleotide sequences within the local vicinity of the start codon in eukaryotic mRNAs, as well. In eukaryotic mRNA, this is known as the **Kozak sequence** (Figure 26.2.5). The sequence was originally defined as 5′-(gcc)gccRccAUGG-3 where:

1. The underlined nucleotides indicate the translation start codon, coding for Methionine.
2. upper-case letters indicate highly conserved bases, *i.e.* the 'AUGG' sequence is constant or rarely, if ever, changes.
3. 'R' indicates that a purine (adenine or guanine) is always observed at this position (with adenine being more frequent according to Kozak rules)
4. a lower-case letter denotes the most common base at a position where the base can nevertheless vary

5. the sequence in parentheses (gcc) is of uncertain significance.

The AUG is the initiation codon encoding a methionine amino acid at the N-terminus of the protein. (Rarely, GUG is used as an initiation codon, but methionine is still the first amino acid as it is the met-tRNA in the initiation complex that binds to the mRNA). Variation within the Kozak sequence alters the “strength” of the translational start site. Kozak sequence strength refers to the favor ability of initiation, affecting how much protein is synthesized from a given mRNA. This is shown in Figure 26.2.5.

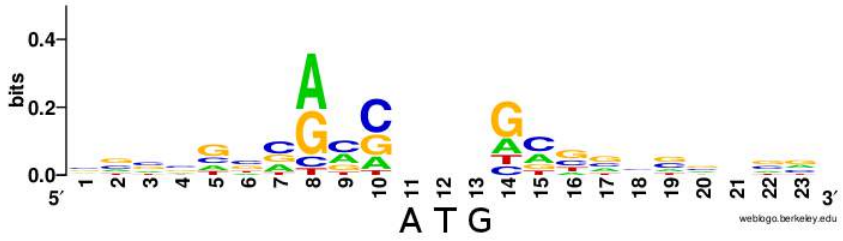


Figure 26.2.5: The Kozak Sequence. The sequence schematic above shows the most conserved bases around the initiation codon from over 10,000 human mRNAs. Larger letters indicate a higher frequency of incorporation. Note the larger size of A and G at the 8 position (−3, Kozak position) and at the G at position 14 which corresponds to (+4) position in the Kozak sequence. Figure from: *Audrey Michel*

26.2.3: The Elongation Phase of Translation

Both prokaryotic and eukaryotic elongation phases of transcription utilize similar elongation factors during the process. Table 26.2.2 provides a summary of their functions.

Table 26.2.2 **Comparison of Prokaryotic and Eukaryotic Translation Elongation Factors**

26.2.4: Prokaryotic Elongation

The prokaryotic elongation phase of transcription requires the activity of three primary elongation factors (EFs), EF-Tu, EF-Ts, and EF-G. During elongation, aminoacyl-tRNAs are delivered to the ribosome in the form of a ternary complex: the tRNA, a translational GTPase (in bacteria: EF-Tu or SelB), and a GTP molecule, as shown in Figure 26.2.6. The tRNA decodes the information on the mRNA by forming hydrogen bonds (H-bonds) between codon and anticodon nucleobases. Remarkably, the free-energy difference between correct (cognate) and incorrect (near-cognate, non-cognate) base pairing alone does not explain the very high fidelity of decoding. Rather, high fidelity is achieved by a two-step decoding process: initial selection leading to GTPase activation and proofreading. In addition to the free-energy difference, kinetic effects contribute to the discrimination. The GTP hydrolysis rate is increased and tRNA rejection rate is decreased by the recognition of the correct codon.

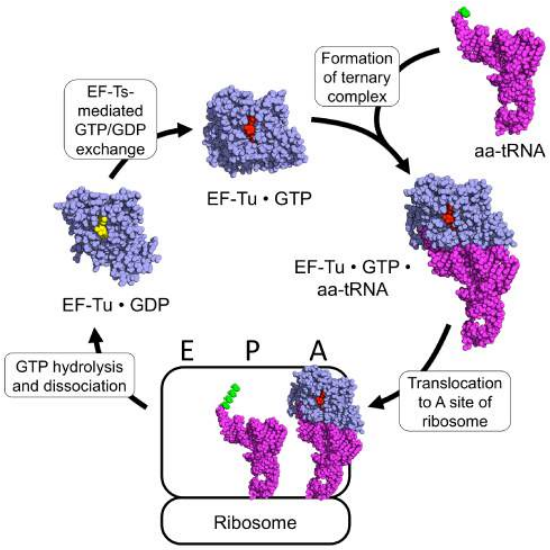


Figure 26.2.6: The EF-Tu tRNA Chaperone Protein. Ef-Tu (shown in blue) is a GTPase enzyme chaperone that when bound to GTP (shown in red), can form a ternary complex with an amino acid-containing tRNA molecule (tRNA shown in purple; amino acid shown in green). The ternary complex translocates to the A-site of the ribosome where the anticodon of the incoming tRNA hybridizes with the codon of the mRNA. An exact match of the codon-anticodon hybridization leads to the hydrolysis of GTP by EF-Tu and its subsequent dissociation from the ribosome (the hydrolysis product GDP is shown in yellow). EF-Tu is unable to release the GDP on its own. It requires an protein-mediated exchange of the GDP for a molecule of GTP using the elongation factor, EF-Ts. Figure from Awchen

Small-subunit nucleotides A1492 and A1493 adopt a flipped-out conformation in the presence of a tRNA and, in this conformation, the tRNA anticodon hydrogen bonds with the codon of the mRNA forming a mini-helix structure as shown in Figure 26.2.7. The flipped out nucleotides A1492, A1493 along with G530 were found to shield the codon–anticodon base pairs from solvent. This shielding of near-cognate base pairs from the solvent is incomplete causing an increase in the free-energy difference between near-cognate and cognate base pairs and more flexibility within the docking region. This reduces the strength of hydrogen bonding between a non-cognate tRNA and causes the inappropriate tRNA to leave the A-site before peptide bond formation can occur. This increases the fidelity and discrimination of tRNA selection, such that only the correct cognate tRNA is incorporated into the A-site.

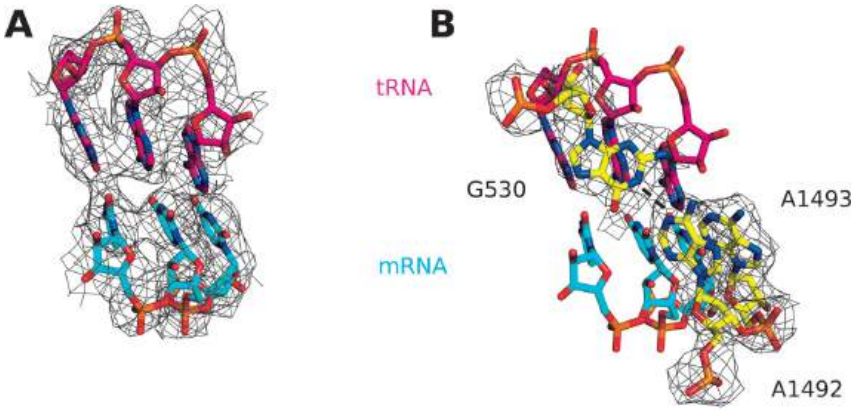


Figure 26.2.7: Cognate Interaction of tRNA and mRNA in *E. coli.* Fislage, M., et. al. (2018) Nuc. Acids Res. 46(11):5861-5874.

Panel (A) shows the interaction of cognate tRNA anticodon (shown in red) with the mRNA codon (shown in blue) when wild type EF-Tu binds and chaperones the tRNA complex to the A-site of the ribosome.Panel (B) is the same figure as shown in (A), with the ribosomal nucleotides A1492 and A1493 depicted in the flipped out conformation.

Interestingly aminoglycosides, a class of antibiotics, bind to the decoding center and lock nucleotides A1492/A1493 in the flipped-out conformation as shown in Figure 26.2.8. In this way aminoglycosides promote the accommodation of near-cognate, thus

wrong, tRNAs into proteins during synthesis causing wide-spread mutagenesis. This is toxic to the bacteria and leads to bacterial cell death.

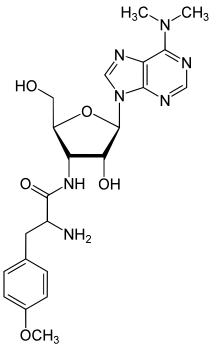


Figure 26.2.8 Structure of the Aminoglycoside Antibiotic, Puromycin. Figure from: *Yikrazuul*

After GTP hydrolysis, the GTPase EF-Tu dissociates from the tRNA. At this point, the EF-Tu is tightly bound with a molecule of GDP and cannot release GDP on its own to be recycled for a second round of tRNA chaperoning. The recharging of EF-Tu is executed by the Elongation Factor Thermo stable (EF-Ts), as shown in Figure 26.2.9. The binding of EF-Ts with EF-Tu-GDP causes a conformational change in EF-Tu that allows the release of GDP. The binding of a new molecule of GTP with the EF-Tu protein causes the dissociation of EF-Ts and fully recharges EF-Tu.

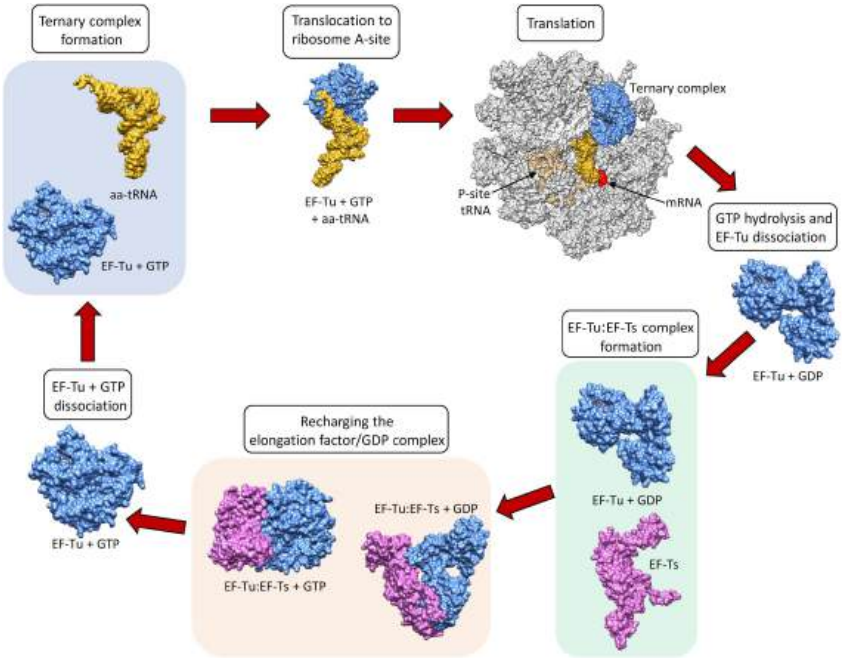


Figure 26.2.9: Recharging of EF-Tu is Mediated by EF-Ts. When EF-Tu is bound with GTP (shown in blue), it can bind an aa-tRNA complex (shown in yellow) causing the translocation of the aa-tRNA to the A-site of the ribosome (shown in grey). The anticodon of the tRNA is positioned in the A-site, such that it can form H-bonds with the codon of the mRNA (shown in red). Binding of the cognate aa-tRNA in the A-site causes GTP hydrolysis by EF-Tu and subsequent dissociation of EF-Tu (GDP) from the ribosome. Following the ribosomal dissociation, EF-Tu(GDP) binds with a molecule of EF-Ts (shown in purple). Binding of EF-Ts causes a conformational change in EF-Tu, such that GDP is released. A molecule of GTP is then incorporated into EF-Tu while it is complexed with EF-Ts. This causes the dissociation of EF-Ts from the complex and the EF-Tu is recharged. *Harvey, K.L, et. al. (2019) Front. Microbiol. 10:2351*

Dissociation of EF-Tu from the ribosome allows the tRNA to move into to the peptidyl transferase center (A-site) on the large subunit. At the core of ribosomal translation is the catalysis of peptide bond formation, as shown in Figure 26.2.10 The current reaction models point to a substrate assisted mechanism. Simulations indicate that the transition state forms due to extensive hydrogen bonding with water molecules and the surrounding rRNA bases and that the C-O bond cleavage takes place after C-N bond formation. Peptide bond formation results in the transition of the amino acid docked on the P-site tRNA to the nascent

growing peptide that is now held on the tRNA in the A-site. Note that this mechanism causes the nascent growing peptide to always grow in the N- to C- direction.

Figure 26.2.10: Proposed Mechanism of Peptide Bond Formation in the Ribosome. *(Left)* tRNA molecules docked in the P- and A-sites are oriented such that the two amino acids move in close proximity. Hydrogen bond formation with the ribosomal rRNA and surrounding water molecules stabilizes the formation of the transition state and enables the amine nitrogen from the A-site amino acid to attack the carbonyl carbon of the P-site amino acid. Formation of the peptide bond occurs prior to the C-O bond breakage forming a transient oxyanion intermediate. *(Right)* The products of the peptidyl transferase reaction include the nascent growing peptide attached to the A-site tRNA and an empty tRNA in the P-site. Figure modified from *Awchen*

Once the peptide bond is formed, the ribosome needs to translocate down the mRNA to make the next mRNA codon available within the A-site. This also requires the shifting of the tRNA molecules, such that the tRNA in the A-site (which is now tethered to the nascent peptide) shifts to the P-site. The P-site tRNA (which is now empty) shifts to the E-site, and if there was an empty tRNA in the E-site, it will shift to exit the ribosome. Shifting the tRNAs and mRNAs within the ribosome core requires the action of the EF-G elongation factor, as shown in Figure 26.2.11

Figure 26.2.11: One Cycle of Elongation. Figure modified from: *Goodsell, D. (2010) Molecule of the Month* (*Left*): During one round of amino acid elongation on a nascent peptide, the EF-Tu protein binds with the cognate aa-tRNA molecule and shuttles it to the A-site of the ribosome. GTP hydrolysis by EF-Tu leads to the hybridization of the anticodon of the tRNA with the codon of the mRNA and causes the dissociation of the EF-Tu (GDP-bound) from the ribosome. *(Center):* Following the dissociation of EF-Tu, the peptide bond is formed leading to the transfer of the nascent peptide from the tRNA in the P-site, to the tRNA in the A-site. (*Right*): Peptide bond formation leads to a conformational change in the ribosome that allows the binding of EF-G (GTP Bound) near the A-site of the ribosome. Rapid hydrolysis of GTP by EF-G causes a large conformational shift in the protein that twists the large subunit of the ribosome and shifts the bound tRNAs from the A- to the P-site; from the P- to the E-site; or from the E-site to exiting the ribosome.

Figure 26.2.12shows an interactive iCn3D model of the eukaryotic 80S ribosome with bound mRNA and tRNAs (6GX3) . (**Very long load time**)

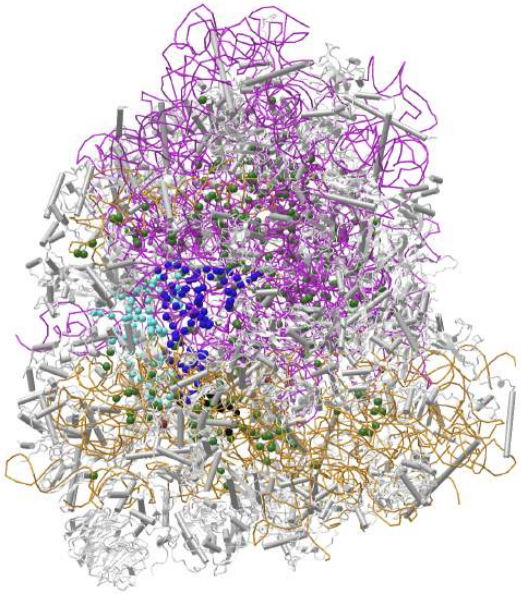


Figure 26.2.12 eukaryotic 80S ribosome with bound mRNA and tRNAs (6GX3) . (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...Ak7ZvtRxaoAQp9 (**Very long load time**)

color coding as follows:

- gray: protein tube
- coiled coils: RNA trace
- black spheres: mRNA
- dark blue spheres: ap/P-site tRNA
- cyan spheres: pe/E-site-tRNA
- green spheres: Mg^{2+}

EF-G is a GTP hydrolase protein that binds to the A-site of the ribosome. The EF-G protein has high flexibility that enables it to act as a hinge. Folding of EF-G is dependent on GTP hydrolysis. Thus, when binding to the ribosome, the fast hydrolysis of GTP acts as a power stroke folding the EF-G protein and causing a conformation shift in the ribosome that enables the translocation of the tRNA residues and the mRNA. Translocation of tRNAs is accompanied by large-scale collective motions of the ribosome: relative rotation of ribosomal subunits and L1-stalk motion, as shown in Figure 26.2.13 The L1 stalk, which is a flexible part of the large subunit, is in contact and moves along with the tRNA from the P to the E site. Once in the EF-G-GDP form, the factor quickly dissociates from the ribosome, opening up the A-site for the recruitment of the next aa-tRNA molecule. The elongation cycle will continue to be repeated until a termination codon is reached.

Figure 26.2.13: Large-scale Motion of the Large Subunit of the Ribosome During Translocation. (a) Pre-translocation structure of the ribosome with tRNAs in A and P sites (green, brown). The L1 stalk of the large subunit is shown in purple. (b) Motions accompanying tRNA translocation. Bock, L.V., Kolár, M.H., Grubmüller, H. (2018) Cur. Op. Struc. Bio. 49:27-35.

26.2.5: Eukaryotic Elongation

The elongation phase in eukaryotic translation is very similar to prokaryotic elongation. Essentially, the mRNA is decoded by the ribosome in a process that requires the selection of each aminoacyl-transfer RNA (aa-tRNA), which is dictated by the mRNA codon in the ribosome acceptor (A) site, peptide bond formation and movement of both tRNAs and the mRNA through the ribosome, as shown in Figure 26.2.14 A new amino acid is incorporated into a nascent peptide at a rate of approximately one every sixth of a second. The first step of this process requires guanosine triphosphate (GTP)-bound eukaryotic elongation factor 1A (eEF1α) to recruit an aa-tRNA to the aminoacyl (A) site, which has an anticodon loop cognate to the codon sequence of the mRNA. The anticodon of this sampling tRNA does not initially base-pair with the A-site codon. Instead, the tRNA dynamically remodels to generate a codon-anticodon helix, which stabilizes the binding of the tRNA-eEF1α-GTP complex to the ribosome A site. This helical structure is energetically favorable for cognate or correct pairing, and so discriminates between the non-cognate or unpaired and single mismatched or near-cognate species. This is important for the accuracy of decoding since it provides a mechanism to reject a non-cognate tRNA that carries an inappropriate amino acid. The pairing of the tRNA and codon induces GTP hydrolysis by eEF1α, which is then evicted from the A site. In parallel with this process, the ribosome undergoes a conformational change that stimulates contact between the 3′ end of the aa-tRNA in the A site and the tRNA carrying the polypeptide chain in the peptidyl (P) site. The shift in position of the two tRNAs [A to the P site and P to the exit (E) site] results in ribosome-catalyzed peptide bond formation and the transfer of the polypeptide to the aa-tRNA, thus extending the polypeptide by one amino acid. The second stage of the elongation cycle requires a GTPase, eukaryotic elongation factor 2 (eEF2), which enters the A-site and, through the hydrolysis of GTP, induces a change in the ribosome conformation. This stimulates ribosome translocation to allow the next aa-tRNA to enter the A-site, thus starting a new cycle of elongation.

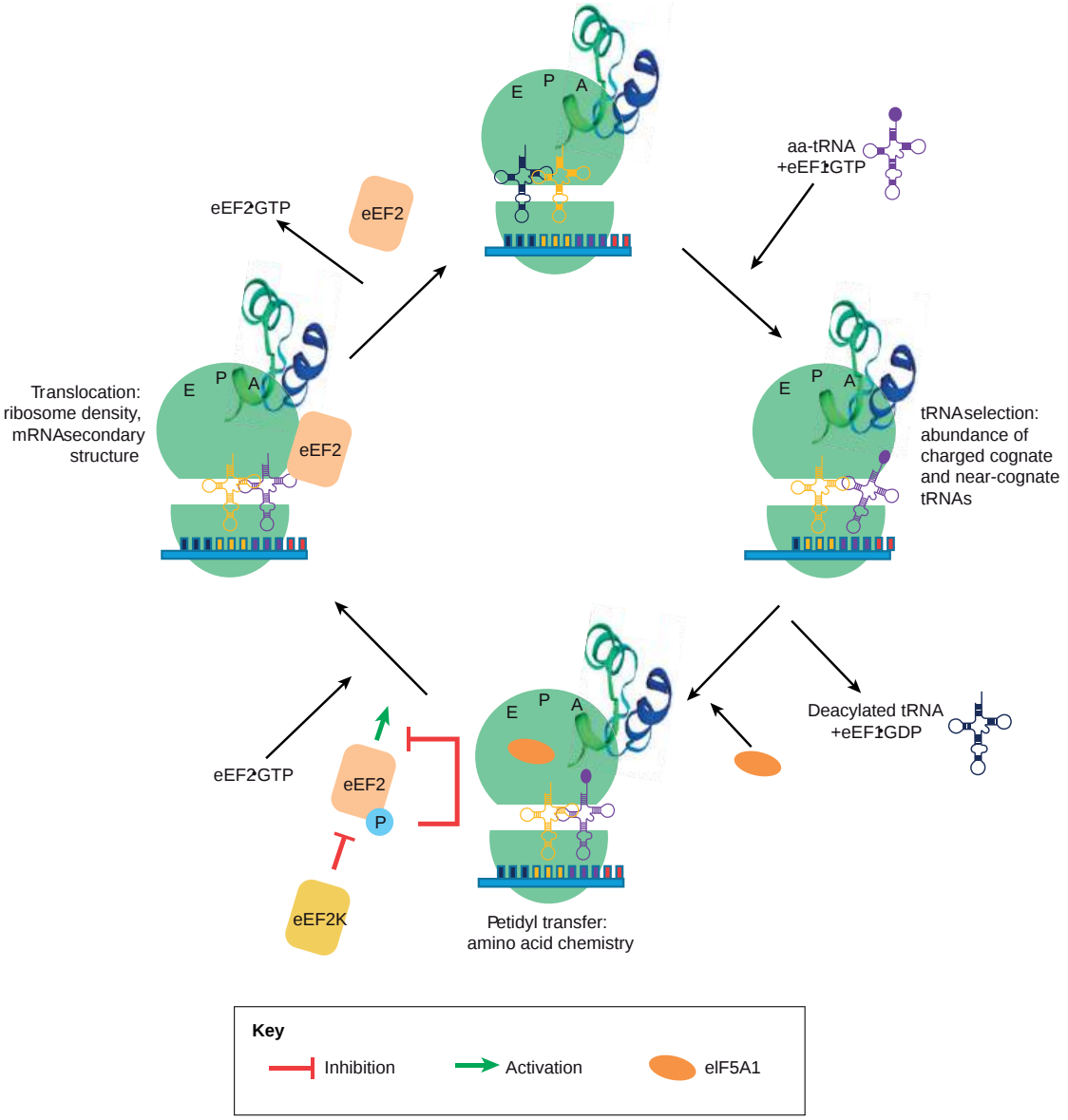


Figure 26.2.14: Eukartyotic Translation Elongation Phase. *Knight, J.R.P., et. al. (2020) Disease Models & Mechanisms 13, dmm043208.*

This schematic represents the four basic steps of eukaryotic translation elongation. The ribosome contains three tRNA-binding sites: the aminoacyl (A), peptidyl (P) and exit (E) sites. In the first step of peptide elongation, the tRNA, which is in a complex with eIF1 and GTP and contains the cognate anticodon to the mRNA coding sequence, enters the A site. Recognition of the tRNA leads to the hydrolysis of GTP and eviction of eEF1 from the A site. In parallel, the deacylated tRNA in the E site is ejected. The A site and the P site tRNAs interact, which allows ribosome-catalyzed peptide bond formation to take place. This involves the transfer of the polypeptide to the aa-tRNA, thus extending the nascent polypeptide by one amino acid. eIF5A allosterically assists in the formation of certain peptide bonds, e.g. proline-proline. eEF2 then enters the A site and, through the hydrolysis of GTP, induces a change in the ribosome conformation and stimulates translocation. The ribosome is then in a correct conformation to accept the next aa-tRNA and commence another cycle of elongation.

26.2.5.1: The Ribosome as a Ribozyme

Protein synthesis from a mRNA template occurs on a ribosome, a nanomachine composed of proteins and ribosomal RNAs (rRNA). Peptide bond formation occurs when another tRNA-amino acid molecule binds to an adjacent codon on mRNA. The tRNA has a cloverleaf tertiary structure with some intrastranded H-bonded secondary structure. The last three nucleotides at the 3' end of the tRNA are CpCpA. The amino acid is esterified to the terminal 3'OH of the terminal A by a protein enzyme, aminoacyl-tRNA synthetase.

Covalent amide bond formation between the second amino acid to the first, forming a dipeptide, occurs at the peptidyl transferase center, located on the larger ribosomal subunit (50S and 60S in bacteria and eukaryotes, respectively). The ribosome ratchets down the mRNA so the dipeptide-tRNA is now at the the P or Peptide site, awaiting a new tRNA-amino acid at the A or Amino site. Figure 26.2.15 below shows a schematic of the ribosome with bound mRNA on the 30S subunit and tRNAs covalently attached to amino acid (or the growing peptide) at the A and P site, respectively.

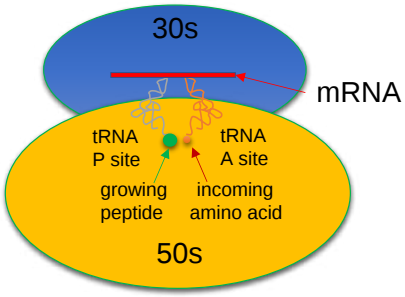


Figure 26.2.15: Ribosome with bound mRNA on the 30S subunit and tRNAs covalently attached to amino acid (or the growing peptide) at the A and P site

A likely mechanism (derived from crystal structures with bound substrates and transition state analogs) for the formation of the amide bond between a growing peptide on the P-site tRNA and the amino acid on the A-site tRNA is shown in Figure 26.2.16 Catalysis does not involve any of the ribosomal proteins (not shown) since none is close enough to the peptidyl transferase center to provide amino acids that could participate in general acid/base catalysis, for example. Hence the rRNA must act as the enzyme (i.e. it is a ribozyme). Initially it was thought that a proximal adenosine with a perturbed pKa could, at physiological pH, be protonated/deprotonated and hence act as a general acid/base in the reaction. However, none was found. The most likely mechanism to stabilize the oxyanion transition state at the electrophilic carbon attack site is precisely located water, which is positioned at the oxyanion hole by H-bonds to uracil 2584 on the rRNA. The cleavage mechanism involves the concerted proton shuffle shown below. In this mechanism, the substrate (Peptide-tRNA) assists its own cleavage in that the 2'OH is in position to initiate the protein shuttle mechanism. (A similar mechanism might occur to facilitate hydrolysis of the fully elongated protein from the P-site tRNA.) Of course all of this requires perfect positioning of the substrates and isn't that what enzymes do best? The main mechanisms for the catalysis of peptide bond formation by the ribosome (as a ribozyme) are intramolecular catalysis and transition state stabilization by the appropriately positioned water molecule.

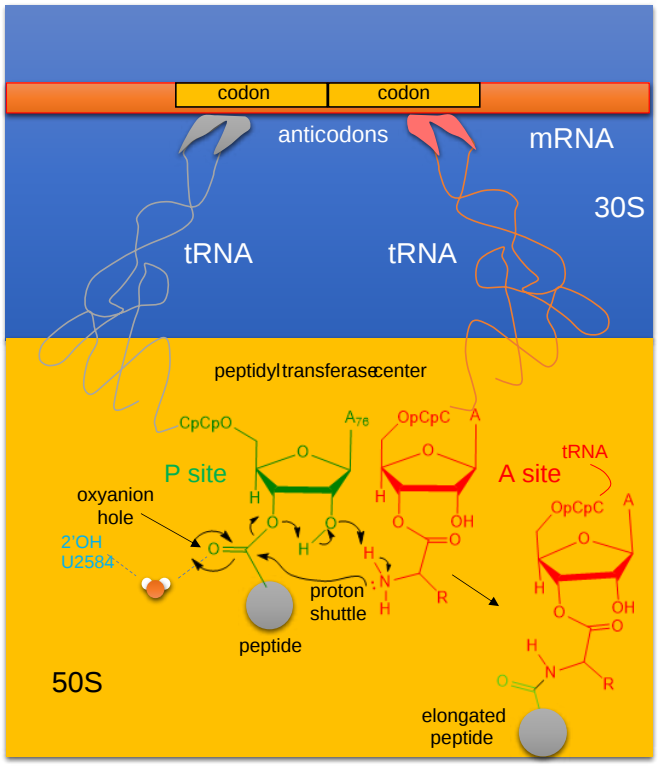


Figure 26.2.16: formation of the amide bond between a growing peptide on the P-site tRNA and the amino acid on the A-site tRNA.

26.2.6: Translation Termination

26.2.6.1: Prokaryotic Termination

Termination of bacterial protein synthesis occurs when a stop codon is presented in the ribosomal A-site and is recognized by a class I release factor, RF1 or RF2. These release factors (RFs) have different but overlapping specificities, where RF1 reads UAA and UAG and RF2 reads UAA and UGA, with strong discrimination against sense codons. The RFs are multi-domain proteins, where binding and stop codon recognition by domain 2 at the decoding site causes the universally conserved GGQ motif of domain 3 to insert into the A-site of the PTC, some 80 Å away from the decoding site. This event triggers hydrolysis of the peptidyl-tRNA bond in the P-site of the PTC, and the nascent peptide chain can then be released via the ribosomal exit tunnel, as shown in Figure 26.2.14 After peptide release, RF1 and RF2 dissociate from the post-termination complex. The dissociation is accelerated by a class II release factor called RF3, which functions as a translational GTPase that binds and hydrolyses GTP in the course of termination.

While RF3 increases the efficiency of peptide hydrolysis, it is not an essential protein for the process. In gene knockout studies, RF3 is dispensable for growth of *Escherichia coli*, and its expression is not conserved in all bacterial lineages. For example, RF3 is not present in the thermophilic model organisms of the *Thermus* and *Thermatoga* genera and in infectious *Chlamydiales* and *Spirochaetae*. This means that both RF1 and RF2 are capable of performing a complete round of termination independently of RF3 or that other GTPases from the elongation or initiation phases of translation can compensate for the action of RF3.

The release factors RF1 and RF2 acquire an open conformation (Figure 26.2.17on the 70S ribosome, which is distinctly different from the closed conformation observed in crystal structures of free RFs. The conformational equilibrium of the free RFs in solution shows that this open conformation is dominating at about 80%.

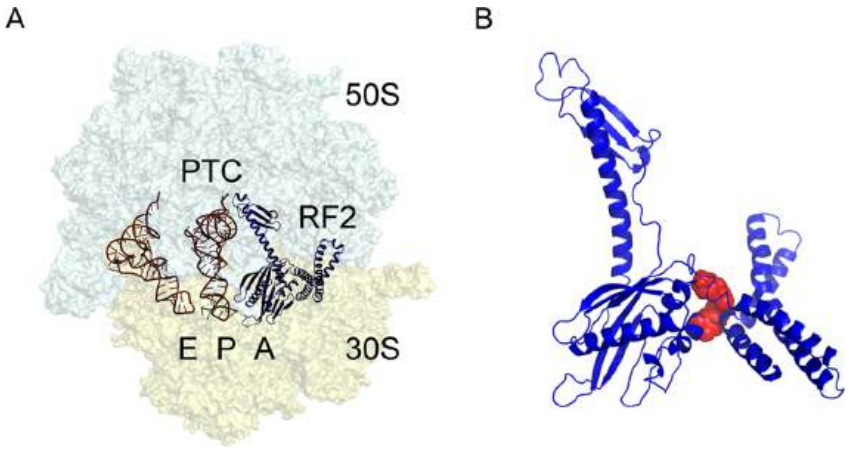


Figure 26.2.17: The bacterial 70S ribosome termination complex with RF2. (A) View of the ribosome termination complex with E- and P-site tRNAs (brown), mRNA (green), and RF2 (dark blue). (B) Close-up view of the hinge region of RF2 between domains 1 and 4 used for virtual screening, where the putative binding region is indicated by a docked ligand (red). *Ge, X., et. al. (2019) Scientific Reports 9:15424.*

During peptide hydrolysis, the RF factors cause rotational and conformational changes within the ribosome that allow the binding of a ribosome recycling factor (RRF) and the EF-G GTPase, which leads to the dissociation of the large subunit from the small subunit and the release of the mRNA, as shown in Figure 26.2.18

Figure 26.2.18: Termination of Translation. quent binding of the ribosome recyclying factor, RRF, and EF-G causes the dissociation of the large and small ribosomal subunits and the release of the mRNA. Figure modified from: *Bock, L.V., Kolár, M.H., Grubmüller, H. (2018) Cur. Op. Struc. Bio. 49:27-35.*

When a stop codon enters the A-site of the ribosome RF1 or RF2 enter the A-site and bind with the mRNA. This leads to the hydrolysis of the protein and release through the exit tunnel. Binding of RF3 and GTP hydrolysis causes the dissociation of the RF factors and conformational change of the ribosome structure. Subse

26.2.7: Eukaryotic Termination

In eukaryotes and archaea, on the other hand, a single omnipotent RF reads all three stop codons. Although the mechanism of translation termination is basically the same, there is neither sequence nor structural homology between the bacterial RFs and the eukaryotic eRF1, apart from the universally conserved GGQ motif which is required for peptide hydrolysis from the tRNA. the eRF3 GTPase coordinates the release of eRF1 following hydrolysis. In Archaea, there is no eRF3 homolog, instead the aEF1A protein mediates this function. The process of eukaryotic ribosomal disassembly and recycling is currently not well understood, but appears to involve an ABC type ATPase called ABCE1. Mitochondria have independent RFs that can recognize standard and non-standard stop codons, and are more homologous with bacterial systems of ribosomal recycling and disassembly.

26.2.8: Summary of Translation

An overall summary of prokaryotic translation is given in Figure 26.2.19

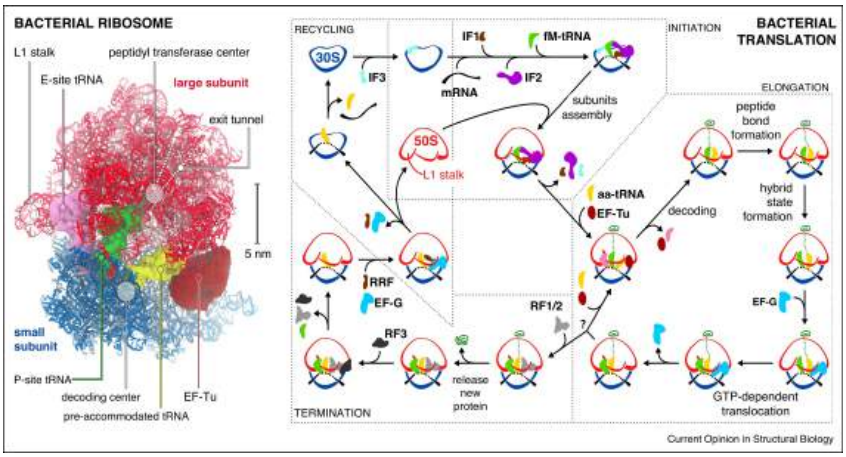


Figure 26.2.19: Summary of Prokaryotic Translation. *Bock, L.V., Kolár, M.H., Grubmüller, H. (2018) Cur. Op. Struc. Bio. 49:27-35.*

Left panel: Structure of the bacterial ribosome in complex with EF-Tu (PDB 5AFI).

Right Panel: Scheme of the bacterial translation cycle. 30S: small subunit; 50S: large subunit; IF1, IF2, IF3: initiation factors; fM-tRNA: N-formylmethionine tRNA; aa-tRNA: aminoacyl tRNA; EF-Tu, EF-G: elongation factors; RF1, RF2, RF3: release factors; RRF: ribosome recycling factor; green trace: nascent protein. The question mark stands for a stop codon recognition.

26.2.9: References

This chapter was remixed and adapted from the following resources under creative commons licensing:

1. Wikipedia contributors. (2020, June 27). Wobble base pair. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:47, August 12, 2020, from https://en.Wikipedia.org/w/index.php?title=Wobble_base_pair&oldid=964760055
2. Lorenz, C., Lünse, C., and Mörl, M. (2017) tRNA modifications: Impact on structure and thermal adaptation. *Biomolecules* 7(2)35. Available at: https://www.mdpi.com/2218-273X/7/2/35/htm
3. Pan, T. (2018) Modifications and functional genomics of human transfer RNA. *Cell Research* 28:395-404. Available at: https://www.nature.com/articles/s41422-018-0013-y#Sec3
4. Bednárová, A., Hanna, M., Durham, I., Van Cleave, T., England, A., Chaudhuri, A., and Krishnan, N. (2017) Lost in translation: Defects in transfer RNA modifications and neurological disorders. *Front. Mol Neurosci.* 10:135. Available at: https://www.researchgate.net/publication/316440980_Lost_in_Translation_Defects_in_Transfer_RNA_Modifications_and_Neurological_Disorders
5. Wikipedia contributors. (2020, May 17). Transfer RNA. In *Wikipedia, The Free Encyclopedia*. Retrieved 23:03, August 15, 2020, from https://en.Wikipedia.org/w/index.php?title=Transfer_RNA&oldid=957227343
6. Gomez, M.A.R., and Ibba M. (2020) Aminoacyl-tRNA Synthetases. *RNA*, doi: 10.1261/rna.071720.119 Available at: https://rnajournal.cshlp.org/content/early/2020/04/17/rna.071720.119.abstract
7. Li, R., Macnamara, L.M., Leuchter, J.D., Alexander, R.W., and Cho, S.S. (2015) MD Simulations of tRNA and Aminoacyl-tRNA Syntetases: Dynamics, Folding, Binding, and Allostery. *Int. J. Mol Sci.* 16(7):15872-15902. Available at: https://www.mdpi.com/1422-0067/16/7/15872
8. Wikipedia contributors. (2020, August 15). Ribosome. In *Wikipedia, The Free Encyclopedia*. Retrieved 05:22, August 16, 2020, from https://en.Wikipedia.org/w/index.php?title=Ribosome&oldid=973144885
9. Kater, L., Thoms, M., Barrio-Garcia, C., Cheng, J., Ismail, S., Ahmed, Y.L., Bange, G., Kressler, D., Berninghausen, O., Sinning, I., Hurt, E., and Beckmann, R. (2017) Visualizing the assembly pathway of nucleolar Pre-60S Ribosomes. Cell 171(7):1599-1610. Available at: https://www.sciencedirect.com/science/article/pii/S0092867417314290
10. Bock, L.V., Kolár, M.H., Grubmüller, H. (2018) Molecular simulations of the ribosome and associated translation factors. *Cur. Op. Struc. Bio.* 49:27-35. Available at: https://www.sciencedirect.com/science/article/pii/S0959440X1730132X
11. Doris, S.M., Smith, D.R., Beamesderfer, J.N., Raphael, B.J., Nathanson, J.A., and Gerbi, S.A. (2015) Universal and domain-specific sequences in 23S-28S ribosomal RNA identified by computational phylogenetics. *RNA* 21:1719-1730. Available at: https://www.researchgate.net/publication/281141702_Universal_and_domain-specific_sequences_in_23S-28S_ribosomal_RNA_identified_by_computational_phylogenetics

12. Aleksashin, M.A., Leppik, M., Hochenberry, A.J., Klepacki, D., Vázquez-Laslop, N., Jewett, M.C., Remme, J., and Mankin A.S. (2019) Assembly and functionality of the ribosome with tethered subunits. *Nature Communications* 10:930. Available at: https://www.nature.com/articles/s41467-019-08892-w#rightslink
13. Wikipedia contributors. (2020, July 3). Formylation. In *Wikipedia, The Free Encyclopedia*. Retrieved 22:03, August 16, 2020, from https://en.Wikipedia.org/w/index.php?title=Formylation&oldid=965827476
14. Gualerzi, C.O., and Pon C.L. (2015) Initiation of mRNA translation in bacteria: structural and dynamic aspects. *Cell Mol Life Sci.* 72:4341-4367. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4611024/
15. Wikipedia contributors. (2020, June 14). Kozak consensus sequence. In *Wikipedia, The Free Encyclopedia*. Retrieved 06:03, August 18, 2020, from https://en.Wikipedia.org/w/index.php?title=Kozak_consensus_sequence&oldid=962444929
16. Knight, J.R.P., Garland, G., Pöyry, T., Mead, E. Vlahov, N., Sfakianos, A., Frosso, S., De-Lima-Hedayioglu, F., Mallucci, G.R., von der Haar, T., Smales, C.M., Sansom, O.J., and Willis, A.E. (2020) Control of translation elongation in health and disease. *Dis. Mod. and Mech.* 13: dmm043208. Available at: https://dmm.biologists.org/content/13/3/dmm043208
17. Adio, S., Sharma, H., Senyushkina, T., Karki, P., Maracci, C., Wohlgemuth, I., Holtkamp, W., Peske, R., and Rodina, M.V. (2018) Dynamics of ribosomes and release factors during translation termination in *E. coli*. *eLife* 7:e34253. Available at: https://elifesciences.org/articles/34252
18. Ge, X., Oliveira, A., Hjort, K., Bergfors, T., Guitiérrez-de-Terán, H., Andersson, D.I., Sanyal, S., and Åqvist, J. (2019) Inhibition of translation termination by small molecules targeting ribosomal release factors. *Scientific Reports* 9: 15424. Available at: https://www.nature.com/articles/s41598-019-51977-1#rightslink
19. Svidritskiy, E., Demo G., Loveland A.B., Xu, C., and Korosteleve, A.A. (2019) Extensive ribosome and RF2 rearrangements during translation termination. *eLife* 8:e46850. Available at: https://elifesciences.org/articles/46850
20. Sauert, M., Temmel, H., and Moll, I. (2015) Heterogeneity of the translational machinery: Variations on a common theme. Biochimie 114:39-47. Available at: https://www.sciencedirect.com/science/article/pii/S0300908414003952

26.3: Translational Regulation and Protein Degradation

Search Fundamentals of Biochemistry

26.3.1: Regulation of Translation

26.3.1.1: Heterogeneity of Ribosome Structure

Over the years, many studies performed in eukaryotes presented evidence that ribosomes can vary in their protein and rRNA complement between different cell types and developmental states. These observations culminated in the postulation of the *'ribosome filter hypothesis'* by Mauro and Edelman in the year 2002. The authors propose that the ribosome composition functions as a translation determination factor. Depending on the RPs and rRNA sequences represented in the respective ribosome, the complex acts like a filter that selects for specific mRNAs and hence modulates translation, as shown in Figure 26.3.1. RP heterogeneity can arise from differential expression of paralogs/homologs of RP proteins within different cell types or occur due to differential post-translational modifications of RPs, such as phosphorylation. The protein-to-rRNA ratio may also slightly vary within ribosomal composition affecting translation efficiency and selectivity.

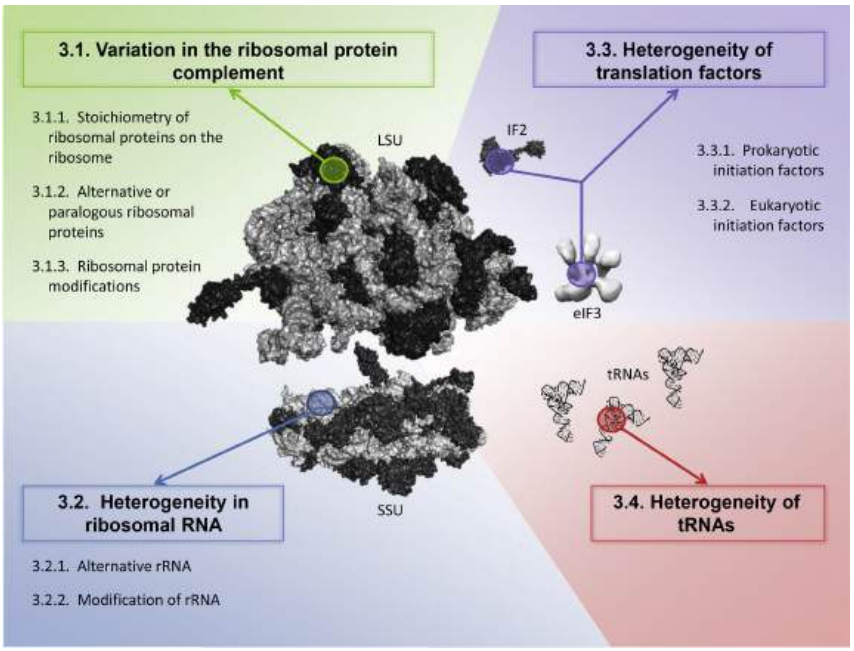


Figure 26.3.1: Components of the translation machinery that have the potential to contribute to functional heterogeneity. Structures were visualized with Polyview 3D software using the following maps: *E. coli* large ribosomal subunit (LSU; pdb accession code: 3d5a, rRNA in light gray, proteins in dark gray); small ribosomal subunit (SSU; pdb accession code: 3d5b); *E. coli* IF2 (pdb accession code: 1g7r); human eIF3 (EMD-2166); yeast $tRNA^{Phe}$ (pdb accession code: 6tna).*Sauert, M., Temmel, H., and Moll, I. (2015) Biochimie 114: 39-47.*

RNA genes are also present in multiple copies throughout the genomes of organisms from all domains of life. For example, the bacteria *Streptomyces coelicolor* harbors six copies of divergent large subunit (LSU) rRNA genes that constitute at least five different LSU rRNA species in a cell. These genes were shown to be differentially transcribed during the morphological development of the organism. Similarly, *B. subtilis* harbors ten *rRNA* operons and their reduction to one copy increased the doubling time as well as the sporulation frequency and the motility of the resulting mutant.

Modification of the rRNA also provides another avenue of ribosomal heterogeneity. Similar to tRNA, rRNA residues can be chemically modified and commonly have 2-OH methylation. The conversion of uridine to pseudouridine is also quite common. In eukaryotes, the modifications are facilitated by snoRNAs and their tissue-specific expression might be a source for ribosome specialization. In light of the increasing evidence, ribosome heterogeneity, though still far from being entirely understood, proves to be an integral mechanism to modulate and fine-tune protein synthesis in response to environmental signals in all organisms.

26.3.1.2: Effects of Sequence and Secondary Structure in mRNA

The amount of protein produced from any given mRNA (i.e., the translational output) is influenced by multiple factors specified by the primary nucleotide sequence. These factors include GC content, codon usage, codon pairs, and secondary structure. For

example, 5'UTR sequences in the mRNA may interact with small miRNAs and lead to RNA interference. miRNA interactions may also target mRNA for degradation(Figure 26.3.2). This process is aided by protein chaperones called argonautes. This antisense-based process involves steps that first process the miRNA so that it can base-pair with a region of its target mRNAs. Once the base pairing occurs, other proteins direct the mRNA to be destroyed by nucleases. Fire and Mello were awarded the 2006 Nobel Prize in Physiology or Medicine for this discovery.

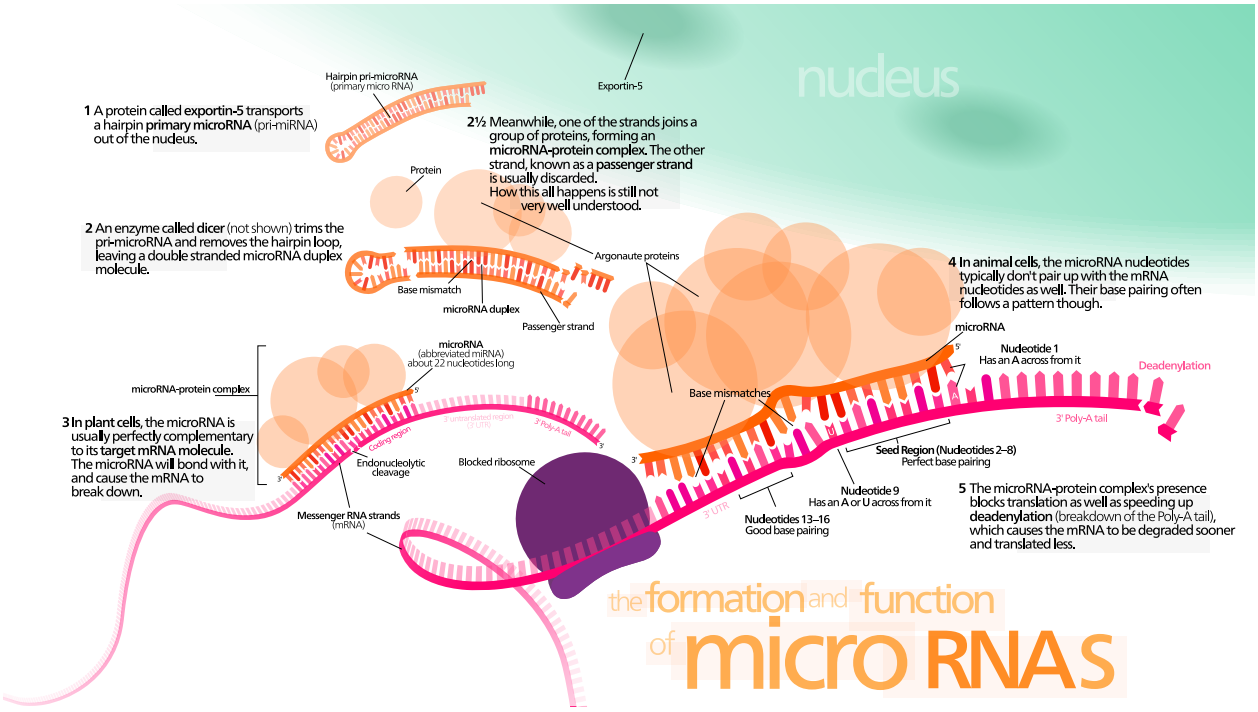


Figure 26.3.2: Role of Micro RNA (miRNA) in the Inhibition of Eukaryotic mRNA Translation. *Wikimedia Commons*

- Step (1) shows how Exportin-5 transports a hairpin primary micro RNA (pri-miRNA) out of the nucleus and into the cytoplasm.
- Step (2) shows how Dicer (not shown) trims the pri-miRNA and removes the hairpin loop. A group of proteins, known as Argonautes, form a miRNA/protein complex.
- Steps (3,4) show how miRNA/protein complex hydrogen bonds with mRNA based on complimentary sequence homology, and blocks translation.
- Step (5) shows the miRNA/protein complex binding speeds up the breakdown of the polyA tail of the mRNA, causing the mRNA to be degraded sooner.

26.3.1.3: Effects of the Nascent Peptide on Ribosome Efficiency

Since the Peptidyl Transferase Center (PTC) is buried within the large subunit, during translation the nascent peptide chain (NC) exits through a 100 Å-long tunnel (Figure 26.3.3). The exit tunnel plays an active role in protein synthesis. Certain peptide sequences specifically interact with tunnel walls and induce ribosome stalling. Furthermore, the exit tunnel is a binding site for a clinically important class of antibiotics known as the macrolides.

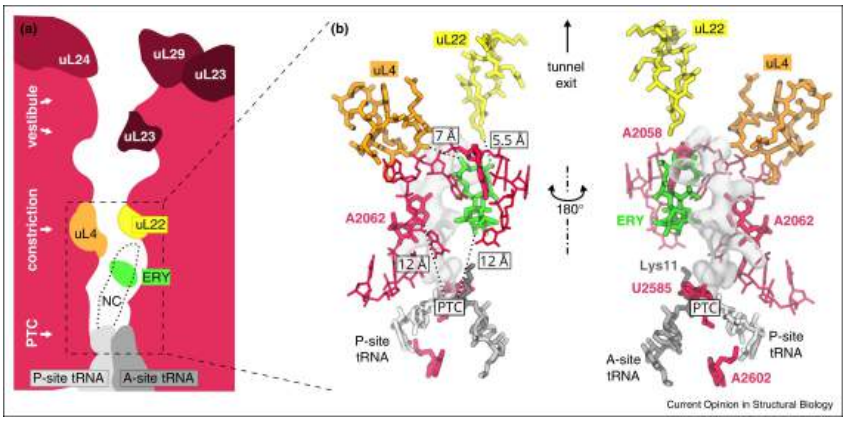


Figure 26.3.3: The Ribosomal Exit Tunnel (a) Scheme of the ribosome exit tunnel with several proteins highlighted. NC, nascent peptide chain; ERY, erythromycin; PTC, peptidyl transferase center. (b) Context of the erythromycin (ERY, in green) binding; figure based on PDB: 5JTE. Several large subunit nucleotides are highlighted in bold red. Two proteins uL4 and uL22 form a constriction site. The nascent peptide is shown as a transparent surface structure. *Bock, L.V., Kolár, M.H., Grubmüller, H. (2018) Cur. Op. Struc. Bio. 49:27-35*

When synthesizing proteins containing proline stretches (i.e. several prolines in a row), ribosomes become stalled. Stalling is alleviated by a specialized elongation factor, EF-P in bacteria. Recently, cryo-EM structures of a ribosome stalled by a proline stretch with and without EF-P were resolved. In simulations of the PTC region, elongation factor P (EF-P) was observed to stabilize the P-site tRNA in a conformation compatible with peptide bond formation, while in the absence of EF-P, the P-site tRNA moved away from the A-site tRNA.

The exit tunnel can accommodate 30–60 AAs, depending on the level of NC compaction. The rate of translation of about 4–22 AA per second in bacteria provides the NC with sufficient time to explore its conformational space and to start folding when still bound to the ribosome-tRNA complex.

26.3.2: Proteasome

The 26S proteasome is the central element of proteostasis regulation in eukaryotic cells. It is required for the degradation of protein factors in multiple cellular pathways and it plays a fundamental role in cell stability. The 26S proteasome has a structural configuration that confines the proteolytic active sites in a location unreachable for native and functional proteins, thus preventing uncontrolled degradation. The proteolytic active sites are found in the interior of a barrel-shaped core particle (CP or 20S). The entrances of the tunnel, placed at the distal ends of the barrel, are commonly occupied by the regulatory particle (RP or 19S), a sophisticated protein assembly that acts as a substrate processing machine. The regulatory particle has the important role of receiving, deubiquitinating, unfolding, and translocating substrates to the CP and it adopts different configurations depending on the activity states they exhibit. This process typically requires ATP hydrolysis. Moreover, conformationally distinct proteasomes may show different subcellular distributions depending on functional requirements in each cell type and environmental situation. Proteasomes are distributed throughout the cell, detected in the cytoplasm and the nucleus, and they can localize to hotspots in distinct intracellular regions or specific sites with high protein metabolism or with specific protein degradation requirements.

The core of the proteasome consists of a symmetrical cylinder-shaped structure composed of four stacked rings, each containing 7 different subunits and is called the 20S proteasome, as shown in Figure 26.3.4 The two outer rings are each composed of seven α-subunits (α1-α7 or PSMA1-7). During proteasome assembly, the α-rings serve as the backbone for the incorporation of β-subunits, followed by the dimerization of two half proteasomes. In mature proteasomes, the α-rings regulate substrate entrance since the α-subunits have hydrophobic loops that close the 20S barrel to prevent the random entry of substrates.

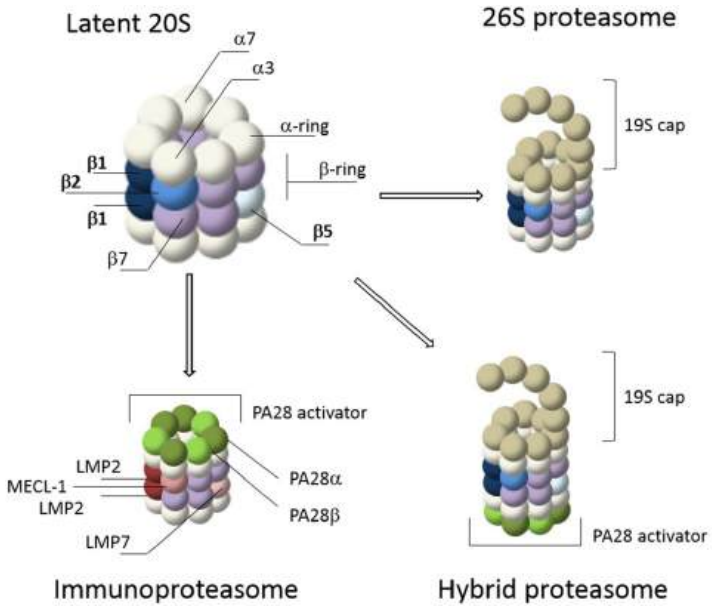


Figure 26.3.4: Sabine et al., Frontiers in Molecular Biosciences, 6 (2019), https://www.frontiersin.org/articles...olb.2019.00056. DOI=10.3389/fmolb.2019.00056. Creative Commons Attribution License (CC BY).

The 20S core of the proteasome consists of 4 stacked rings. The outer rings contain seven α-subunits (white) while the inner rings contain seven β-subunits (purple). The catalytic subunits, β1, β2, and β5, are depicted in shades of blue. Gate opening of the 20S core occurs via capping by proteasome activators such as the 19S cap or PA28. The 19S cap is the most abundant activator and it forms the 26S proteasome together with the 20S core. Different cells have different caps. For example, interferoIFN-γ stimulation induces *de novo* formation of immunoproteasomes, which contains the immune subunits β1i (LMP2), β5i (LMP7), and β2i (MECL-1) (shades of red), as well as proteasome activation by PA28αβ (shades of green). Proteasomes in neural tissue are discussed below.

In general, protein entry can only be established after gate opening by proteasome activators (PAs) such as the 19S cap, after which substrates can enter the interior of the 20S core for degradation. The inner two rings of the 20S barrel consist of the subunits β1-β7 (PSMB1-7). Each β-ring contains 3 catalytic subunits; termed β1, β2, and β5. In mature 20S complexes, the pro-peptides of these catalytic β-subunits are auto-catalytically removed. Upon autocatalytic processing, the N-terminal threonine residues become exposed as the catalytically reactive residues, harboring both the nucleophile (the hydroxyl group) and the catalytic base (the N-terminal amine) involved in peptide bond cleavage. Each catalytic subunit has selectivity toward specific residues. β1 has caspase- or peptidyl-glutamyl peptidase-like activity, preferring cleavage at the C-terminus of acidic residues. β2 has trypsin-like activity and cleaves after basic residues, while β5 has chymotrypsin-like activity and prefers cleavage after hydrophobic residues. Figure 26.3.5 shows an interactive iCn3D model of the Human 20S Proteasome (6RGQ).

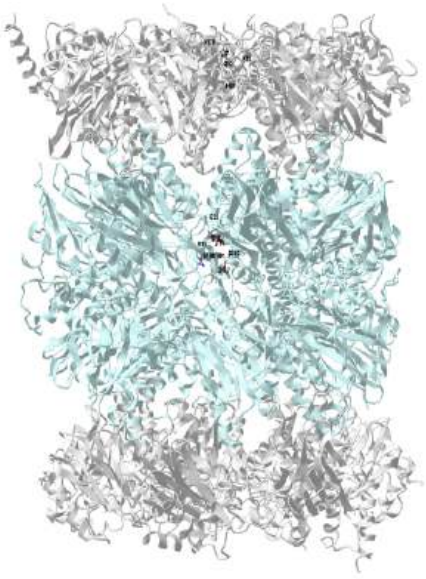


Figure 26.3.5: Human 20S Proteasome (6RGQ). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...9gDp6nj7Co2Ud9

The alpha subunits are shown in light gray and the beta subunits are in light cyan. Active site residue, as defined by databases, are shown for one of the alpha chain (L37, Q53, K55, H68, V170) and one of the beta chains (T1, D17, R19, K33, S130, D167, S170, G171).

A possible generic mechanism for the cleavage of peptide bonds by the beta chain is shown in Figure 26.3.6. The mechanism shows how the newly exposed N-terminal threonine acts as both the nucleophile (the hydroxyl group) and the catalytic base.

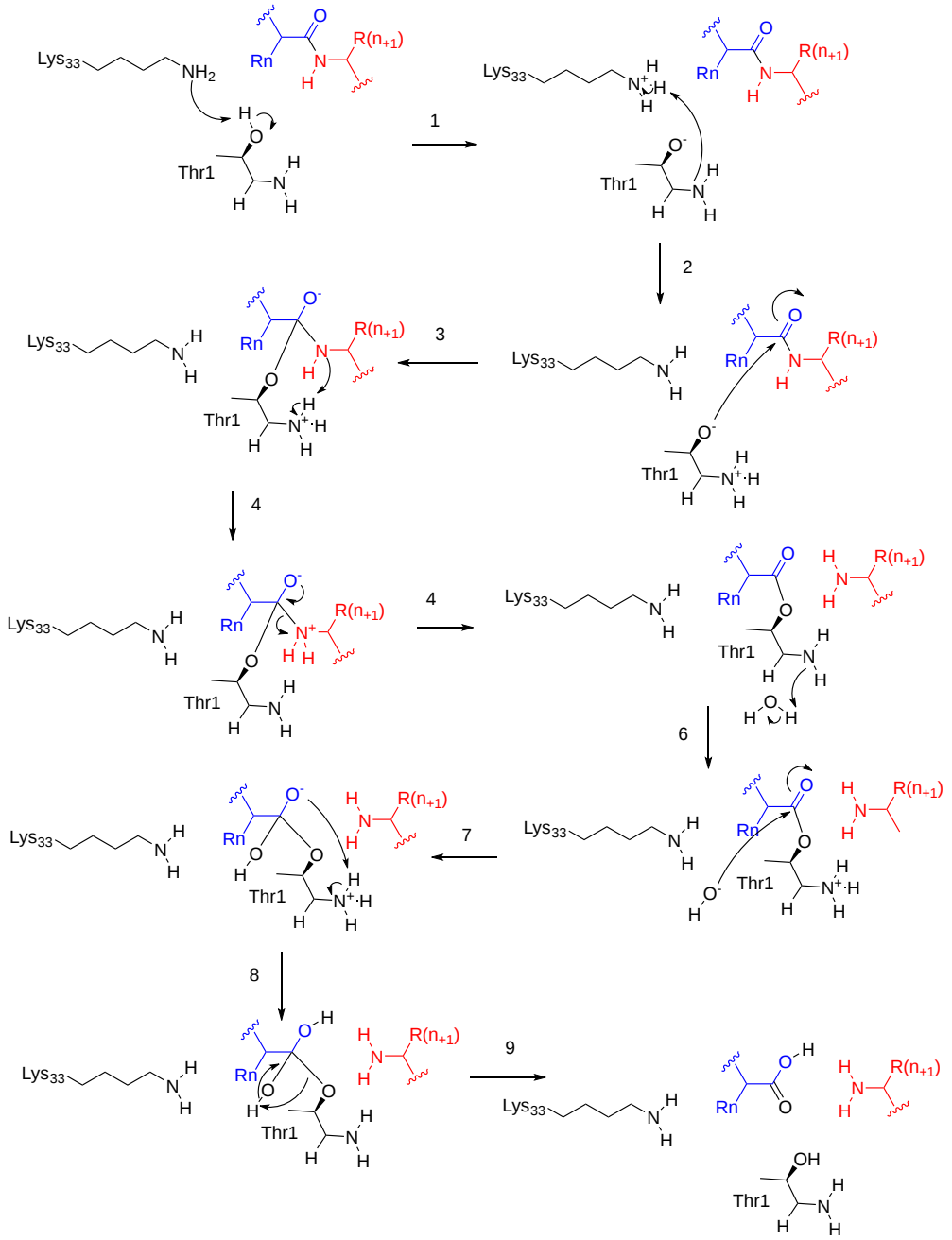


Figure 26.3.6: Possible mechanism for the beta chain proteasome cleavage of proteins. after Arjun Saha, Gabriel Oanca, Dibyendu Mondal, and Arieh Warshel. *The Journal of Physical Chemistry B* **2020***124* (27), 5626-5635. DOI: 10.1021/acs.jpcb.0c04435

Proteins are normally targeted to the proteasome using ubiquitin labels attached covalently to a lysine residue, usually in a chained form that produces a polyubiquitinated protein. The process of protein polyubiquitination is carried out by a highly specialized and diverse enzymatic system, which
includes families of ubiquitin-activating enzymes (E1), ubiquitin-conjugating enzymes (E2), and ubiquitin ligases (E3).

The covalent attachment of ubiquitin to specific target proteins is mainly accomplished by stepwise enzymatic cascade reactions, and ubiquitin is attached to the substrates via the concerted action of ubiquitin-activating enzyme (E1), ubiquitin-conjugating enzyme (E2), and ubiquitin ligase (E3). The attachment of ubiquitin or ubiquitin chains to the substrate is a successive process, as shown in Figure 26.3.7. First, the C-terminal carboxylic acid is activated by adenylation using a molecule of ATP forming an ***adenylate (AMP-) intermediate.*** The ***adenylate*** acts as a good leaving group during the next reaction where an E1-ubiquitin thioester bond is formed between the C-terminal Gly carboxyl group of ubiquitin and the active site Cys of the E1 enzyme. AMP leaves the active site at this point. Note that ATP is used in many reactions to activate carboxylic acid functional groups through the

formation of an adenylate intermediate and that this will be seen as a theme in many different types of reactions throughout this textbook. Once the ubiquitin is docked as a thioester on E1, it can be transferred to a Cys residue of the E2 enzyme to form an E2-ubiquitin thioester-linked intermediate. This enzymatic reaction is known as a ***transesterification***.

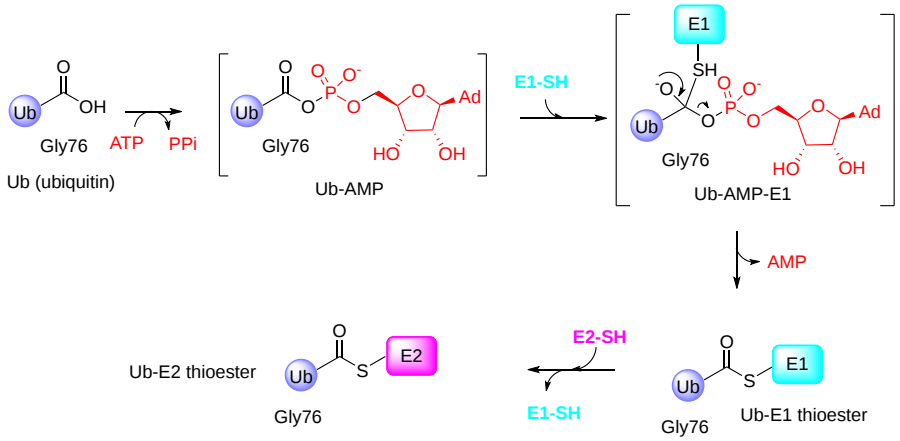


Figure 26.3.7: Enzymatic Activity of the E1 Ubiquitin Activating Enzyme. Conversion of Gly 76 on ubiquitin to E1 and E2 thioesters. After Hann, Z.S., et. al. (2019) *PNAS* 116(31)15475-15484.

Eventually, the E2 transfers the ubiquitin to the substrate protein by E3. Ubiquitin is conjugated to the target protein through an isopeptide bond between its C-terminal glycine (Gly76) and the ε-amino group of a lysine residue. There are three typical ways of linking the ubiquitin with the substrate, as shown in Figure 26.3.8 The first is called mono-ubiquitination, which refers to the modification of one site of a substrate by a single ubiquitin molecule. The second is multi-mono-ubiquitination, which means adding several ubiquitin molecules repetitively to distinct sites (multi-mono-ubiquitination). The third is called polyubiquitination (including linear polyubiquitination and branched polyubiquitination), in which ubiquitin molecules are added to the same site (polyubiquitination, including linear polyubiquitination and branched polyubiquitination) of a substrate. In the second and third ways of linking, the previously attached ubiquitin serves as the “acceptor” of subsequently added ubiquitin. Of course, polyubiquitin chains linked by the same Lys are homogeneous, while those linked at different Lys are heterogeneous or mixed ones.

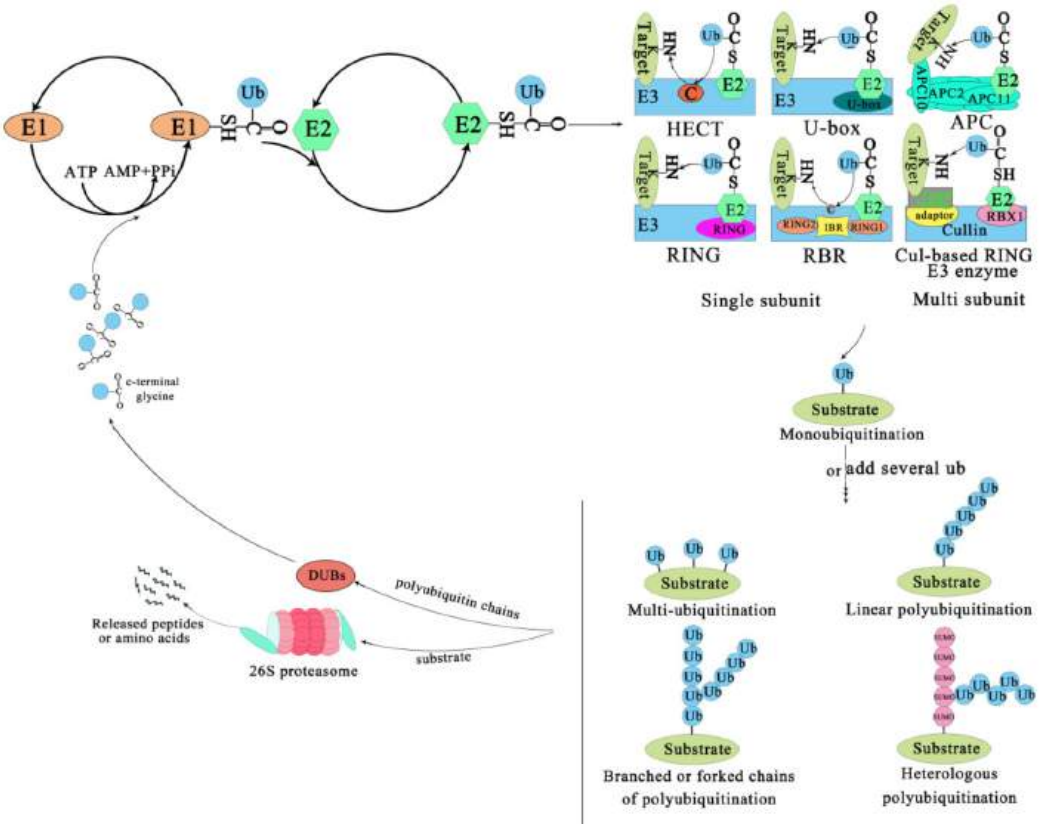


Figure 26.3.8: **The Ubiquitin-Proteasome System.** Liu, W., et.al. (2020) Int J Mol Sci 21(8):2894

The process of ubiquitination from activation to the attachment to the substrate is catalyzed by three major enzymes. The substrates labeled by ubiquitin are degraded by the 26S proteasome or play a non-degradative role in other processes. Abbreviations: APC, Anaphase-promoting complex; DUBs, Deubiquitinating enzymes; E1, Ubiquitin-activating enzyme; E2, Ubiquitin-conjugating

enzyme; E3, Ubiquitin-ligase enzyme; Cul-based, Cullin-RING box1-Ligase; HECT, Homology to E6-AP C Terminus; Ub, Ubiqitin; SUMO, Small ubiquitin-related modifier; RBX1, RING-Box 1; RING, Really interesting new gene; RBR, RING1-IBR(cysteine/histidine-rich region)-RING2.

Subsequently, the substrate complex tagged by the ubiquitin is either degraded by the 26S proteasome or executes nonproteolytic functions, such as the regulation of gene expression, cellular trafficking, or other biological function. In most cases, polyubiquitinated proteins are recognized and degraded by the 26S proteasome, and the ubiquitin or ubiquitin chain is hydrolyzed and freed by deubiquitinating enzymes (DUBs) for reuse in further conjugation cycles after being removed from the substrate protein (Figure 27.3.6).

The family of E3 enzymes is large and diverse. It is estimated that there are 600-700 E3 enzymes in humans, representing approximately 5% of the human genome. Thus, E3 enzymes can be very substrate specific, leading to the specialized degradation of a small subset of proteins within the cell. E2 enzymes are the intersection between E1 and E3 enzymes and help to determine the ubiquitination of specific target proteins by interacting with different types of E3 enzymes.

Figure 26.3.9 shows an interactive iCn3D model of the Yeast proteasome in resting state (C1-a) (6J2X). (long load)

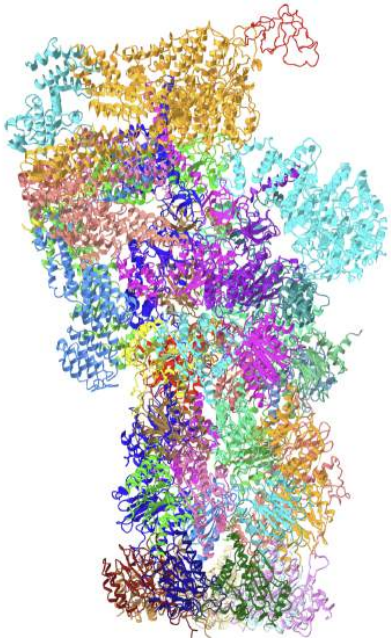


Figure 26.3.9: Yeast proteasome in resting state (C1-a) (6J2X). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?1AmGmkghbZD2zute7

It's too complex to discuss this in much detail. The chains are presented in different colors. You should able to find the alpha and beta chains of the core 20S particle. The regulatory cap proteins comprising the lid are on top of the alpha ring.

26.3.3: References

This chapter was remixed and adapted from the following resources under creative commons licensing:

1. Wikipedia contributors. (2020, June 27). Wobble base pair. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:47, August 12, 2020, from https://en.Wikipedia.org/w/index.php?title=Wobble_base_pair&oldid=964760055
2. Lorenz, C., Lünse, C., and Mörl, M. (2017) tRNA modifications: Impact on structure and thermal adaptation. *Biomolecules* 7(2)35. Available at: https://www.mdpi.com/2218-273X/7/2/35/htm
3. Pan, T. (2018) Modifications and functional genomics of human transfer RNA. *Cell Research* 28:395-404. Available at: https://www.nature.com/articles/s41422-018-0013-y#Sec3
4. Bednárová, A., Hanna, M., Durham, I., Van Cleave, T., England, A., Chaudhuri, A., and Krishnan, N. (2017) Lost in translation: Defects in transfer RNA modifications and neurological disorders. *Front. Mol Neurosci.* 10:135. Available at: https://www.researchgate.net/publication/316440980_Lost_in_Translation_Defects_in_Transfer_RNA_Modifications_and_Neurological_Disorders

5. Wikipedia contributors. (2020, May 17). Transfer RNA. In *Wikipedia, The Free Encyclopedia*. Retrieved 23:03, August 15, 2020, from https://en.Wikipedia.org/w/index.php?title=Transfer_RNA&oldid=957227343
6. Gomez, M.A.R., and Ibba M. (2020) Aminoacyl-tRNA Synthetases. *RNA*, doi: 10.1261/rna.071720.119 Available at: https://rnajournal.cshlp.org/content/early/2020/04/17/rna.071720.119.abstract
7. Li, R., Macnamara, L.M., Leuchter, J.D., Alexander, R.W., and Cho, S.S. (2015) MD Simulations of tRNA and Aminoacyl-tRNA Syntetases: Dynamics, Folding, Binding, and Allostery. *Int. J. Mol Sci.* 16(7):15872-15902. Available at: https://www.mdpi.com/1422-0067/16/7/15872
8. Wikipedia contributors. (2020, August 15). Ribosome. In *Wikipedia, The Free Encyclopedia*. Retrieved 05:22, August 16, 2020, from https://en.Wikipedia.org/w/index.php?title=Ribosome&oldid=973144885
9. Kater, L., Thoms, M., Barrio-Garcia, C., Cheng, J., Ismail, S., Ahmed, Y.L., Bange, G., Kressler, D., Berninghausen, O., Sinning, I., Hurt, E., and Beckmann, R. (2017) Visualizing the assembly pathway of nucleolar Pre-60S Ribosomes. Cell 171(7):1599-1610. Available at: https://www.sciencedirect.com/science/article/pii/S0092867417314290
10. Bock, L.V., Kolár, M.H., Grubmüller, H. (2018) Molecular simulations of the ribosome and associated translation factors. *Cur. Op. Struc. Bio.* 49:27-35. Available at: https://www.sciencedirect.com/science/article/pii/S0959440X1730132X
11. Doris, S.M., Smith, D.R., Beamesderfer, J.N., Raphael, B.J., Nathanson, J.A., and Gerbi, S.A. (2015) Universal and domain-specific sequences in 23S-28S ribosomal RNA identified by computational phylogenetics. *RNA* 21:1719-1730. Available at: https://www.researchgate.net/publication/281141702_Universal_and_domain-specific_sequences_in_23S-28S_ribosomal_RNA_identified_by_computational_phylogenetics
12. Aleksashin, M.A., Leppik, M., Hochenberry, A.J., Klepacki, D., Vázquez-Laslop, N., Jewett, M.C., Remme, J., and Mankin A.S. (2019) Assembly and functionality of the ribosome with tethered subunits. *Nature Communications* 10:930. Available at: https://www.nature.com/articles/s41467-019-08892-w#rightslink
13. Wikipedia contributors. (2020, July 3). Formylation. In *Wikipedia, The Free Encyclopedia*. Retrieved 22:03, August 16, 2020, from https://en.Wikipedia.org/w/index.php?title=Formylation&oldid=965827476
14. Gualerzi, C.O., and Pon C.L. (2015) Initiation of mRNA translation in bacteria: structural and dynamic aspects. *Cell Mol Life Sci.* 72:4341-4367. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4611024/
15. Wikipedia contributors. (2020, June 14). Kozak consensus sequence. In *Wikipedia, The Free Encyclopedia*. Retrieved 06:03, August 18, 2020, from https://en.Wikipedia.org/w/index.php?title=Kozak_consensus_sequence&oldid=962444929
16. Knight, J.R.P., Garland, G., Pöyry, T., Mead, E. Vlahov, N., Sfakianos, A., Frosso, S., De-Lima-Hedayioglu, F., Mallucci, G.R., von der Haar, T., Smales, C.M., Sansom, O.J., and Willis, A.E. (2020) Control of translation elongation in health and disease. *Dis. Mod. and Mech.* 13: dmm043208. Available at: https://dmm.biologists.org/content/13/3/dmm043208
17. Adio, S., Sharma, H., Senyushkina, T., Karki, P., Maracci, C., Wohlgemuth, I., Holtkamp, W., Peske, R., and Rodina, M.V. (2018) Dynamics of ribosomes and release factors during translation termination in *E. coli*. *eLife* 7:e34253. Available at: https://elifesciences.org/articles/34252
18. Ge, X., Oliveira, A., Hjort, K., Bergfors, T., Guitiérrez-de-Terán, H., Andersson, D.I., Sanyal, S., and Åqvist, J. (2019) Inhibition of translation termination by small molecules targeting ribosomal release factors. *Scientific Reports* 9: 15424. Available at: https://www.nature.com/articles/s41598-019-51977-1#rightslink
19. Svidritskiy, E., Demo G., Loveland A.B., Xu, C., and Korosteleve, A.A. (2019) Extensive ribosome and RF2 rearrangements during translation termination. *eLife* 8:e46850. Available at: https://elifesciences.org/articles/46850
20. Sauert, M., Temmel, H., and Moll, I. (2015) Heterogeneity of the translational machinery: Variations on a common theme. Biochimie 114:39-47. Available at: https://www.sciencedirect.com/science/article/pii/S0300908414003952
21. Shen, Q., Wang, G., Li, S., Liu, X., Lu, S., Chen, Z., Song, K., Yan, J., Geng, L, Huang, Z., Huang, W., Chen, G., and Zhang, J. (2016) ASDv3.0: Unraveling allosteric regulation with structural mechanisms and biological networks. Nucleic Acids Research 44(D1):D527-D535. Available at: https://academic.oup.com/nar/article/44/D1/D527/2503129
22. Wikipedia contributors. (2020, January 19). Isozyme. In *Wikipedia, The Free Encyclopedia*. Retrieved 03:42, May 19, 2020, from https://en.Wikipedia.org/w/index.php?title=Isozyme&oldid=936548836
23. Wikipedia contributors. (2020, April 30). COX-2 inhibitor. In *Wikipedia, The Free Encyclopedia*. Retrieved 07:00, May 22, 2020, from en.Wikipedia.org/w/index.php?title=COX-2_inhibitor&oldid=954080651
24. Clarkson, C.W. (2018) Major Side Effects of NSAIDs and COX-2 Selective Inhibitors. *TUSOM Pharmwiki.* Available at: http://tmedweb.tulane.edu/pharmwiki/doku.php/nsaid_side_effects?do=
25. Santos, A.L. and Lindner, A.B. (2017) Protein Porsttranslational Modifications: Roles in Aging and Age-Related Disease. *Oxidative Medicine and Cellular Longevity,* Article ID: 5716409. Available at: https://www.hindawi.com/journals/omcl/2017/5716409/#copyright

26. Wikipedia contributors. (2020, May 12). Protein phosphorylation. In *Wikipedia, The Free Encyclopedia*. Retrieved 17:38, May 25, 2020, from en.Wikipedia.org/w/index.php?title=Protein_phosphorylation&oldid=956228409
27. Szylveszter, K.P., Németh, T., and Mócsai, A. (2019) Tyrosine Kinases in Autoimmune and Inflammatory Skin Diseases. Front. Immunol. 10.3389(2019.01862). Available at: https://www.frontiersin.org/articles/10.3389/fimmu.2019.01862/full
28. Wikipedia contributors. (2020, May 7). Histone. In *Wikipedia, The Free Encyclopedia*. Retrieved 21:12, May 25, 2020, from en.Wikipedia.org/w/index.php?title=Histone&oldid=955458038
29. Wikipedia contributors. (2020, May 13). Mucin. In *Wikipedia, The Free Encyclopedia*. Retrieved 01:10, June 7, 2020, from en.Wikipedia.org/w/index.php?title=Mucin&oldid=956387296
30. Wikipedia contributors. (2020, March 15). Allosteric regulation. In *Wikipedia, The Free Encyclopedia*. Retrieved 04:07, June 7, 2020, from en.Wikipedia.org/w/index.php?title=Allosteric_regulation&oldid=945637073
31. Dixit, Ajay. Dawra, Rajinder K. Dudeja, Vikas. Saluja, Ashok K. (2016). Role of trypsinogen activation in genesis of pancreatitis. *Pancreapedia: Exocrine Pancreas Knowledge Base*, DOI: 10.3998/panc.2016.25
32. Coll-Martinez, B., and Crosas, B. (2019) How the 26S Proteasome Degrades Ubiquitinated Proteins in teh Cell. *Biomolecules* 9(9):395. Retrieved from: https://www.mdpi.com/2218-273X/9/9/395
33. Liu, W., Tang, X., Qi, X., Ghimire, S., Ma, R., Li, S., Zhang, N., and Si H. (2020) The Ubiquitin Conjugating Enzyme: An Important Ubiquitin Transfer Platform in Ubiquitin-Proteosome System. *Int J Mol Sci* 21(8):2894. Retrieved from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7215765/

CHAPTER OVERVIEW

27: Regulation of Gene Expression

27.1: Regulation of Gene Expression in Bacteria
27.2: Regulation of Gene Expression in Eukaryotes

27.1: Regulation of Gene Expression in Bacteria

27.1.1: Introduction

Each nucleated cell in a multicellular organism contains copies of the same DNA. Similarly, all cells in two pure bacterial cultures inoculated from the same starting colony contain the same DNA, except for changes that arise from spontaneous mutations. If each cell in a multicellular organism has the same DNA, then how is it that cells in different parts of the organism's body exhibit different characteristics? Similarly, how is it that the same bacterial cells within two pure cultures exposed to different environmental conditions can exhibit different phenotypes? In both cases, each genetically identical cell does not turn on, or express, the same set of genes. Only a subset of proteins in a cell at a given time is expressed.

Genomic DNA contains both *structural genes*, which encode products that serve as cellular structures or enzymes, and *regulatory genes*, which encode products that regulate gene expression. The expression of a gene is a highly regulated process. Whereas regulating gene expression in multicellular organisms allows for cellular differentiation, in single-celled organisms like prokaryotes, it primarily ensures that a cell's resources are not wasted making proteins that the cell does not need at that time.

Elucidating the mechanisms controlling gene expression is important to the understanding of human health. Malfunctions in this process in humans lead to the development of cancer and other diseases. Understanding the interaction between the gene expression of a pathogen and that of its human host is important for the understanding of a particular infectious disease. Gene regulation involves a complex web of interactions within a given cell among signals from the cell's environment, signaling molecules within the cell, and the cell's DNA. These interactions lead to the expression of some genes and the suppression of others, depending on circumstances.

Prokaryotes and eukaryotes share some similarities in their mechanisms to regulate gene expression; however, gene expression in eukaryotes is more complicated because of the temporal and spatial separation between the processes of transcription and translation. Thus, although most regulation of gene expression occurs through transcriptional control in prokaryotes, regulation of gene expression in eukaryotes occurs at the transcriptional level and post-transcriptionally (after the primary transcript has been made).

In bacteria and archaea, structural proteins with related functions are usually encoded together within the genome in a block called an *operon* and are transcribed together under the control of a single *promoter,* resulting in the formation of a **polycistronic transcript**, as shown in Figure 27.1.1. In this way, regulation of the transcription of all of the structural genes encoding the enzymes that catalyze the many steps in a single biochemical pathway can be controlled simultaneously, because they will either all be needed at the same time, or none will be needed. For example, in *E. coli*, all of the structural genes that encode enzymes needed to use lactose as an energy source are encoded next to each other in the lactose (or *lac*) operon under the control of a single promoter, the *lac* promoter. French scientists François Jacob (1920–2013) and Jacques Monod at the Pasteur Institute were the first to show the organization of bacterial genes into operons, through their studies on the *lac* operon of *E. coli*. For this work, they won the Nobel Prize in Physiology or Medicine in 1965.

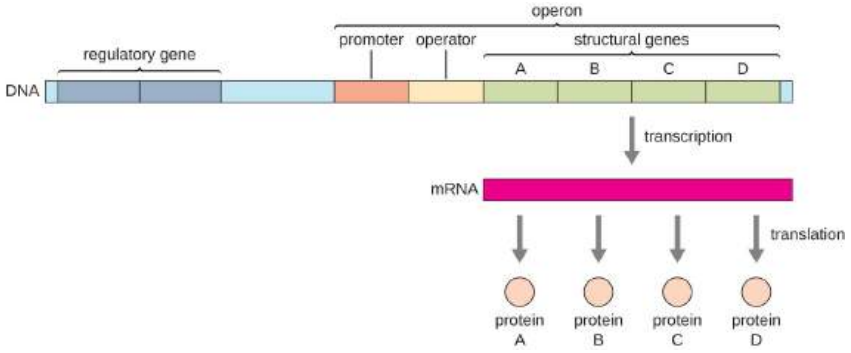


Figure 27.1.1: **Schematic Representation of an Operon.** In prokaryotes, structural genes of related functions are often organized together on the genome and transcribed together under the control of a single promoter. The operon's regulatory region includes both the promoter and the operator. If a repressor binds to the operator, then the structural genes will not be transcribed. Alternatively, activators may bind to the regulatory region, enhancing transcription. *Parker, N., et. al. (2019) Microbiology. Openstax*

Each operon includes DNA sequences that influence its own transcription; these are located in a region called the regulatory region. The regulatory region includes the promoter and the region surrounding the promoter, to which ***transcription factors***, proteins

encoded by regulatory genes, can bind. Transcription factors influence the binding of *RNA polymerase* to the promoter and allow its progression to transcribe structural genes. A ***repressor*** is a transcription factor that suppresses the transcription of a gene in response to an external stimulus by binding to a DNA sequence within the regulatory region called the *operator*, which is located between the RNA polymerase binding site of the promoter and the transcriptional start site of the first structural gene. Repressor binding physically blocks RNA polymerase from transcribing structural genes. Conversely, an *activator* is a transcription factor that increases the transcription of a gene in response to an external stimulus by facilitating RNA polymerase binding to the promoter. An *inducer*, a third type of regulatory molecule, is a small molecule that either activates or represses transcription by interacting with a repressor or an activator.

In prokaryotes, there are examples of operons whose gene products are required rather consistently and whose expression, therefore, is unregulated. Such operons are *constitutively expressed*, meaning they are transcribed and translated continuously to provide the cell with constant intermediate levels of the protein products. Such genes encode enzymes involved in housekeeping functions required for cellular maintenance, including DNA replication, repair, and expression, as well as enzymes involved in core metabolism. In contrast, other prokaryotic operons are expressed only when needed and are regulated by repressors, activators, and inducers.

Prokaryotic operons are commonly controlled by the binding of repressors to operator regions, thereby preventing the transcription of the structural genes. Such operons are classified as either *repressible operons* or inducible operons. Repressible operons, like the tryptophan (*trp*) operon, typically contain genes encoding enzymes required for a biosynthetic pathway. As long as the product of the pathway, like tryptophan, continues to be required by the cell, a repressible operon will continue to be expressed. However, when the product of the biosynthetic pathway begins to accumulate in the cell, removing the need for the cell to continue to make more, the expression of the operon is repressed. Conversely, *inducible operons*, like the *lac* operon of *E. coli*, often contain genes encoding enzymes in a pathway involved in the metabolism of a specific substrate like lactose. These enzymes are only required when that substrate is available, thus expression of the operons is typically induced only in the presence of the substrate.

27.1.2: The trp Operon - A Repressible Operon

E. coli can synthesize tryptophan using enzymes that are encoded by five structural genes located next to each other in the *trp* operon, as shown in Figure 27.1.2 When environmental tryptophan is low, the operon is turned on. This means that transcription is initiated, the genes are expressed, and tryptophan is synthesized. However, if tryptophan is present in the environment, the *trp* operon is turned off. Transcription does not occur and tryptophan is not synthesized.

When tryptophan is not present in the cell, the repressor by itself does not bind to the operator; therefore, the operon is active and tryptophan is synthesized. However, when tryptophan accumulates in the cell, two tryptophan molecules bind to the *trp* repressor molecule, which changes its shape, allowing it to bind to the *trp* operator. This binding of the active form of the *trp* repressor to the operator blocks RNA polymerase from transcribing the structural genes, stopping the expression of the operon. Thus, the actual product of the biosynthetic pathway controlled by the operon regulates the expression of the operon.

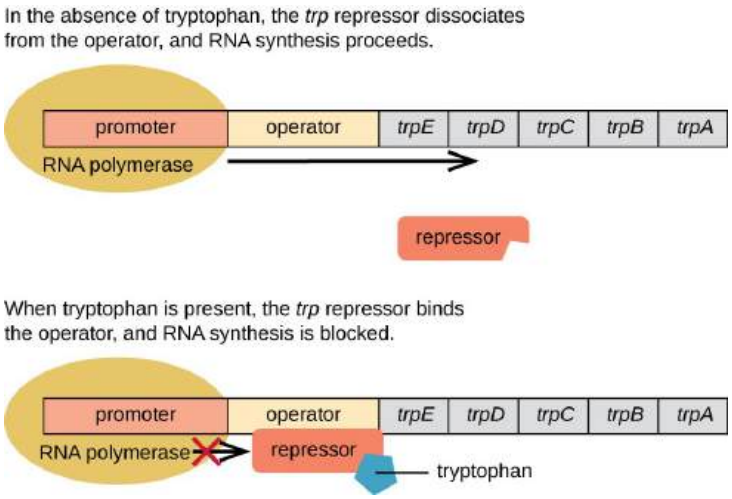


Figure 27.1.2: The Trp Operon. *Parker, N., et. al. (2019) Microbiology. Openstax*

The five structural genes needed to synthesize tryptophan in *E. coli* are located next to each other in the *trp* operon. When tryptophan is absent, the repressor protein does not bind to the operator, and the genes are transcribed. When tryptophan is plentiful, tryptophan binds the repressor protein at the operator sequence. This physically blocks the RNA polymerase from transcribing the tryptophan biosynthesis genes.

Figure 27.1.3 shows an interactive iCn3D model of the E. Coli Trp repressor - operator complex (1TRO).

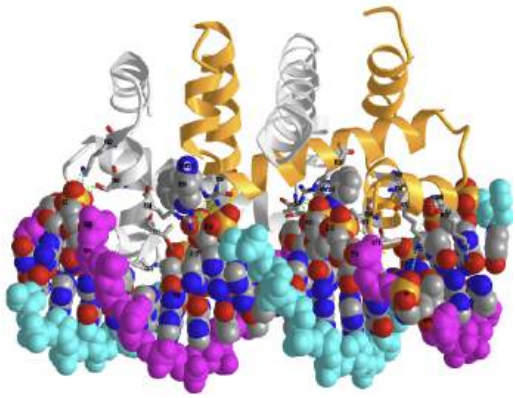


Figure 27.1.3: E. Coli Trp repressor - operator complex (1TRO). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...tjbYwr47gVZjn8

The Trp repressor is shown as a dimer with one subunit gray and the other gold. The backbone of the two DNA strands is shown in spacefill magenta and cyan, except where the bases on the major grove interact with the Trp repressor. The tryptophan in each of the proton monomers is shown in spacefill with CPK colors and labeled. Noncovalent interactions (hydrogen bonds and salt bridges) between the protein and DNA are shown with dotted lines. 6 water-mediated hydrogen bonds to phosphate are not shown. Note that there are few H bonds between the protein and base hydrogen bond donors and acceptors, suggesting that the repressor might bind specifically through geometric interactions with the backbone along with the water-mediated hydrogen bonds.

27.1.3: The Lac Operon: An Inducible Operon

The *lac* operon is an example of an *inducible operon* that is also subject to activation in the absence of glucose. The *lac* operon encodes three structural genes, *lacZ*, *lacY*, and *lacA*, necessary to acquire and process the disaccharide lactose from the environment, as shown in Figure 27.1.4

Figure 27.1.4: Biological Activity of the lac Operon.Figure modified from Esmaeili, A., et. al. (2015) BMC Bioinformatics 16:311

Panel (A) shows a schematic representation of the *lac* operon in *E. coli*. The lac operon has three structural genes, lacZ, lacY, and lacA that encode for β-galactosidase, permease, and galactoside acetyltransferase, respectively. The promoter *(p)* and operator *(o)* sequences that control the expression of the operon are shown. Upstream of the *lac* operon is the *lac* repressor gene, *lacI*, controlled by the *lacI* promoter (*p*).

Panel (B) shows the lac repressor inhibition of the lac operon gene expression in the absence of lactose. The *lac* repressor binds with the operator sequence of the operon and prevents the RNA polymerase enzyme which is bound to the promoter *(p)* from initiating transcription.

Panel (C) shows that in the presence of lactose, some of the lactose is converted into allolactose, which binds and inhibits the activity of the *lac* repressor. The *lac* repressor-allolactose complex cannot bind with the operator region of the operon, freeing the RNA polymerase and causing the initiation of transcription. Expression of the *lac* operon genes enables the breakdown and utilization of lactose as a food source within the organism

The *lacZ* gene encodes the β-galactosidase (β-gal) enzyme responsible for the hydrolysis of lactose into simple sugars glucose and galactose, as shown in Figure 27.1.5. The β-gal enzyme can also mediate the breakdown of the alternate substrate 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (Xgal) (panel B). The breakdown product, 5-bromo-4-chloro-3-hydroxyindole – 1, spontaneously dimerizes to form the intensely blue product, 5,5′-dibromo-4,4′-dichloro-indigo – 2. Thus, Xgal has been a valuable research tool, not only in the study of the enzymatic activity of β-gal but also in the development of the commonly used blue-white DNA cloning system that utilizes the β-gal enzyme as a marker in molecular cloning experiments.

Figure 27.1.5: Reactions Controlled by the Expression of the Lac Operon. Andreas Piehler, Yikrazuul, and NUROtikerr Panel (A) shows the expression of the β-galactosidase enzyme enables the breakdown of lactose into simple sugars, glucose, and galactose for *E. coli* to use as a food resource. Panel (B) shows that the β-galactosidase enzyme also mediates the breakdown of the non-native substrate 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (Xgal). Breakdown product (1) 5-bromo-4-chloro-3-hydroxyindole quickly dimerizes into the intensely blue product (2) 5,5′-dibromo-4,4′-dichloro-indigo making it a useful tool for molecular biology. Panel (C) shows that β-D-1-thiogalactopyranoside (IPTG) can serve as a non-native inducer of the lac operon. It mimics the structure of lactose and binds with the Lac Repressor.

The lac operon contains two more genes, in addition to *lacZ* (Fig. 4). The *lacY* gene encodes a permease that increases the uptake of lactose into the cell and *lacA* encodes a galactoside acetyltransferase (GAT) enzyme. The exact function of GAT during lactose metabolism has not been conclusively elucidated but acetylation is thought to play a role in the transport of the modified sugars.

For the *lac* operon to be expressed, lactose must be present. This makes sense for the cell because it would be energetically wasteful to create the enzymes to process lactose if lactose was not available.

In the absence of lactose, the *lacI* gene is constitutively expressed, expressing the *lac* repressor protein (Fig. 28.2.3 B). The lac repressor binds with an operator region of the *lac* operon and physically prevents RNA polymerase from transcribing the structural genes (Fig. 28.2.3 B). However, when lactose is present, the lactose inside the cell is converted to allolactose. Allolactose serves as an inducer molecule, binding to the repressor and changing its shape so that it is no longer able to bind to the operator DNA (Fig. 28.2.3 C). Removal of the repressor in the presence of lactose allows RNA polymerase to move through the operator region and begin transcription of the *lac* structural genes. In addition to lactose, laboratory experiments have revealed that the non-natural compound Isopropyl β-D-1-thiogalactopyranoside (IPTG) can also bind with the *lac* repressor and cause the expression of *lac* operon (panel C). Similar to Xgal, this compound has also been used as a research tool for molecular cloning.

Figure 27.1.6 shows an interactive iCn3D model of the lactose operon repressor and its complexes with DNA (1LBG).

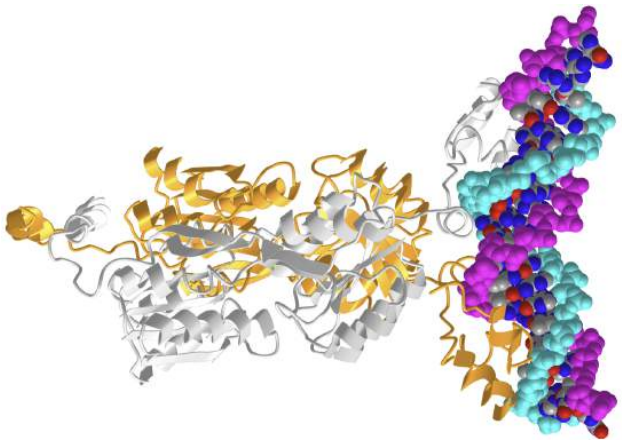


Figure 27.1.6: Lactose operon repressor and its complexes with DNA (1LBG). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...a79MvrAHGpJk99

The resolution of the structure above was insufficient to show the amino acid side chains. Figure 27.1.7 shows an interactive iCn3D model of the NMR solution structure of the dimer of LAC repressor DNA-binding domain complexed to its natural operator O1 (2KEI).

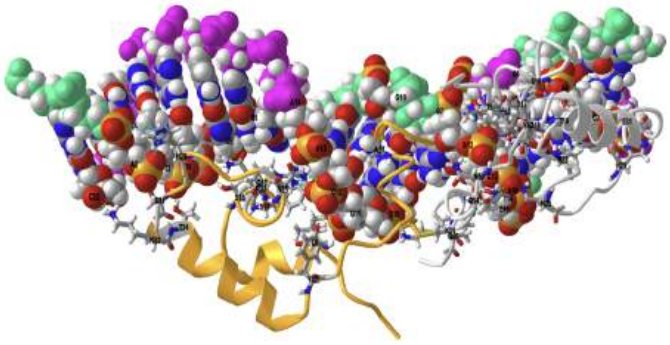


Figure 27.1.7: NMR solution structure of Dimer of LAC repressor DNA-binding domain complexed to its natural operator O1 (2KEI). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...p48VrPFXg6WgD9

Note the presents of white spheres representing hydrogen atoms (these don't appear in crystal structure but do in NMR structures. Color coding is the same as above. Zoom in to see specific interactions between the protein and the exposed DNA base hydrogen bond donors and acceptors. The complex of O1 and O2 shows similar specific and nonspecific contacts, which makes sense given the lambda repressor has similar affinity for those two operator sites. In contrast, one side of the O3 complex shows a loss of protein: DNA interactions, consistent with its lower affinity of its operator O3.

27.1.4: The Lac Operon: Activation by Catabolite Activator Protein

Bacteria typically can use a variety of substrates as carbon sources. However, because glucose is usually preferable to other substrates, bacteria have mechanisms to ensure that alternative substrates are only used when glucose has been depleted. Additionally, bacteria have mechanisms to ensure that the genes encoding enzymes for using alternative substrates are expressed only when the alternative substrate is available. In the 1940s, Jacques Monod was the first to demonstrate a preference for certain substrates over others through his studies of *E. coli*'s growth when cultured in the presence of two different substrates simultaneously. Such studies generated diauxic growth curves, like the one shown in Figure 27.1.8. Although the preferred substrate glucose is used first, *E. coli* grows quickly and the enzymes for lactose metabolism are absent. However, once glucose levels are depleted, growth rates slow, inducing the expression of the enzymes needed for the metabolism of the second substrate, lactose. Notice how the growth rate in lactose is slower, as indicated by the lower steepness of the growth curve.

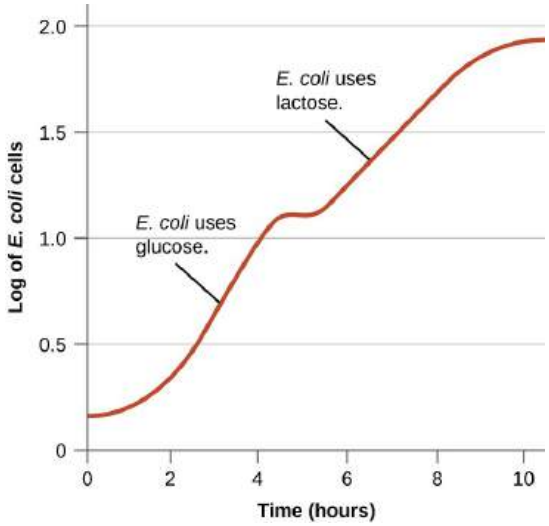


Figure 27.1.8: Utilization of Glucose in *E. Coli.* When grown in the presence of two substrates, *E. coli* uses the preferred substrate (in this case glucose) until it is depleted. Then, enzymes needed for the metabolism of the second substrate are expressed and growth resumes, although at a slower rate. *Parker, N., et. al. (2019) Microbiology. Openstax*

The ability to switch from glucose use to another substrate like lactose is a consequence of the activity of an enzyme called Enzyme IIA (EIIA). When glucose levels drop, cells produce less ATP from catabolism and EIIA becomes phosphorylated. Phosphorylated EIIA activates adenylyl cyclase, an enzyme that converts some of the remaining ATP to cyclic AMP (cAMP), a cyclic derivative of AMP and an important signaling molecule involved in glucose and energy metabolism in *E. coli*, as shown in Figure 27.1.9. As a result, cAMP levels begin to rise in the cell. This is an indicator to the cell, that overall energy levels are low and that ATP is being depleted.

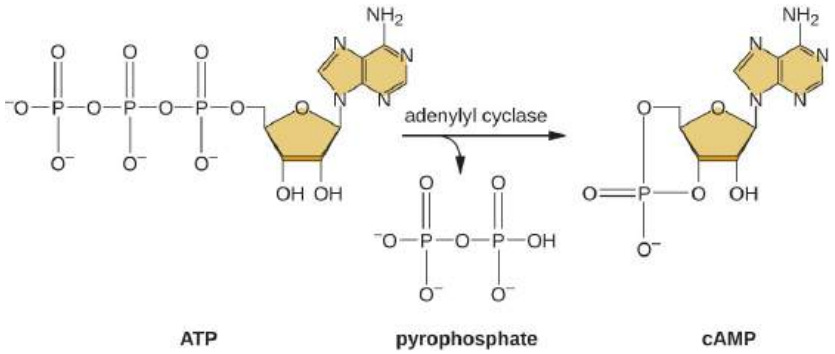


Figure 27.1.9: Conversion of ATP to cAMP. When ATP levels decrease due to the depletion of glucose, some remaining ATP is converted to cAMP by adenylyl cyclase. Thus, increased cAMP levels signal glucose depletion. Figure from: Parker, N., et. al. (2019) Microbiology. Openstax

The *lac* operon also plays a role in this switch from using glucose to using lactose. When glucose is scarce, the accumulating cAMP caused by increased adenylyl cyclase activity binds to catabolite activator protein (CAP), also known as cAMP receptor protein (CRP). The complex binds to the promoter region of the *lac* operon, as shown in Figure 27.1.10. In the regulatory regions of these operons, a CAP binding site is located upstream of the RNA polymerase binding site in the promoter. The binding of the CAP-cAMP complex to this site increases the binding ability of RNA polymerase to the promoter region to initiate the transcription of the structural genes. Thus, in the case of the *lac* operon, for transcription to occur, lactose must be present (removing the lac repressor protein) and glucose levels must be depleted (allowing the binding of an activating protein). When glucose levels are high, there is catabolite repression of operons encoding enzymes for the metabolism of alternative substrates. Because of low cAMP levels under these conditions, there is an insufficient amount of the CAP-cAMP complex to activate the transcription of these operons.

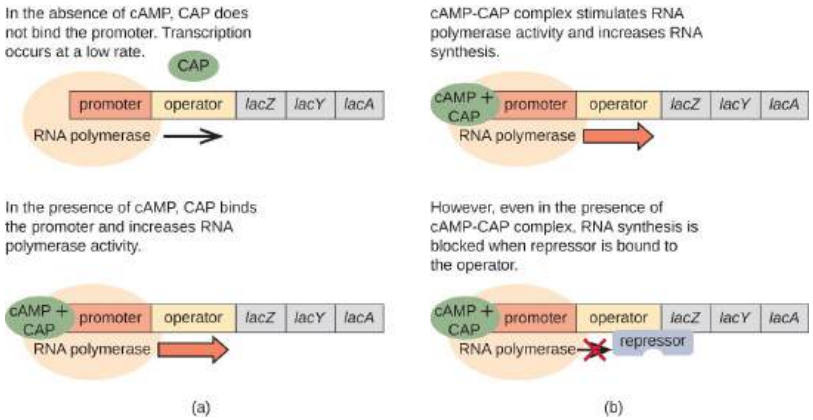


Figure 27.1.10: Effect of CAP on the *Lac* Operon. (a) In the presence of cAMP, CAP binds to the promoters of operons, like the *lac* operon, that encode genes for enzymes for the use of alternate substrates. (b) For the *lac* operon to be expressed, there must be activation by cAMP-CAP as well as the removal of the lac repressor from the operator. Parker, N., et. al. (2019) Microbiology. Openstax

Figure 27.1.11 shows an interactive iCn3D model of the Catabolite activator protein CAP-DNA complex with bound cAMP (2CGP).

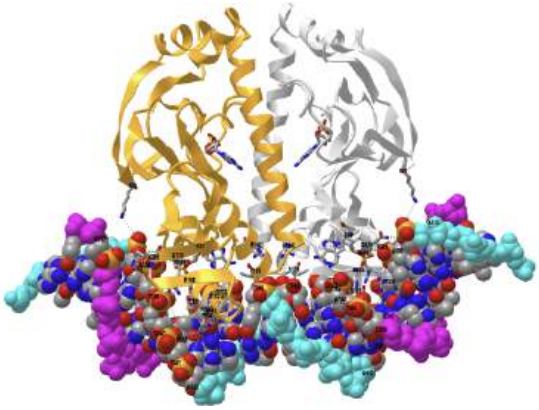


Figure 27.1.11: Catabolite activator protein CAP-DNA complex with bound cAMP (2CGP). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...z6mLhBpPmPMsC6

27.1.5: Global Responses of Prokaryotes

In prokaryotes, several higher levels of gene regulation have the ability to control the transcription of many related operons simultaneously in response to an environmental signal. A group of operons all controlled simultaneously is called a *regulon*.

27.1.5.1: Alarmones

When sensing impending stress, prokaryotes alter the expression of a wide variety of *operons* to respond in coordination. They do this through the production of *alarmones*, which are small intracellular nucleotide derivatives, such as guanosine pentaphosphate (pppGpp), as shown in Figure 27.1.12

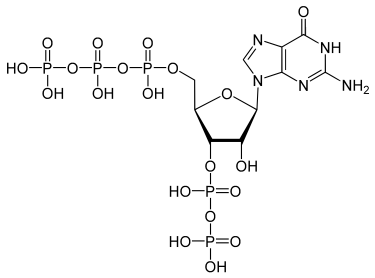


Figure 27.1.12: Structure of Guanosine Pentaphosphate (pppGpp). *Yikrazuul*

Alarmones change which genes are expressed and stimulate the expression of specific stress-response genes. For example, pppGpp signaling is involved in the stringent response in bacteria, causing the inhibition of RNA synthesis when there is a shortage of amino acids present. This causes translation to decrease and the amino acids present are therefore conserved. Furthermore, pppGpp causes the up-regulation of many other genes involved in stress response such as the genes for amino acid uptake (from surrounding media) and biosynthesis.

The use of alarmones to alter gene expression in response to stress appears to be important in pathogenic bacteria, as well. On encountering host defense mechanisms and other harsh conditions during infection, many operons encoding virulence genes are upregulated in response to alarmone signaling. Knowledge of these responses is key to being able to fully understand the infection process of many pathogens and to the development of therapies to counter this process.

27.1.5.2: Quorum Sensing

Quorum sensing (QS) is an intercellular communication mechanism of bacteria used to coordinate the activities of individual cells at the population level in response to surroundings through the production and perception of diffusible signal molecules such as Acyl Homoserine Lactones or small signaling peptides, as shown in Figure 27.1.13 The signal synthase, signal receptor, and signal molecules are three essential elements of the basic QS circuit machinery. Genes encoding signal-generating proteins are also included among the QS target genes. This forms an autoinduction feedback loop to modulate the generation of signal molecules. Several bacterial behaviors including virulence factors expression, secondary metabolites production, biofilm formation, motility, and luminescence are regulated by QS. Through complex regulatory networks, bacteria are capable of expressing corresponding genes according to their population size and of behaving in a coordinated manner.

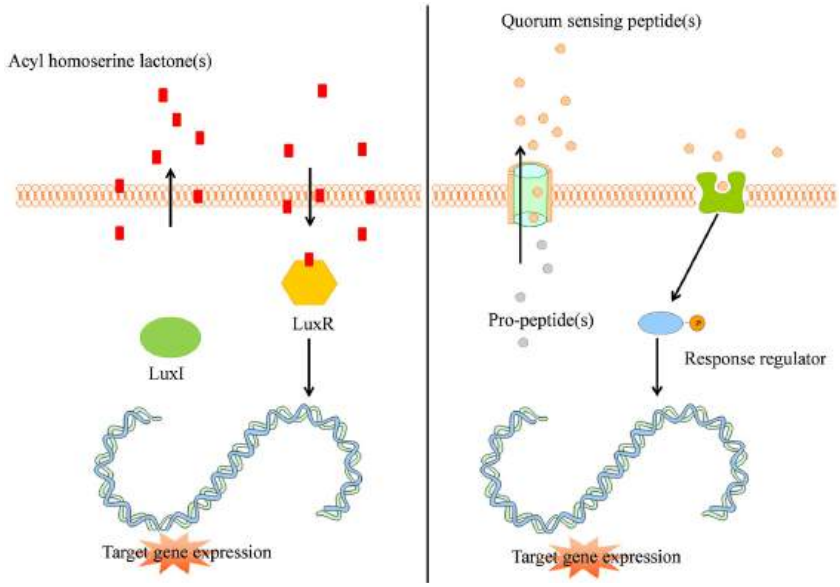


Figure 27.1.13: Examples of Quorum Sensing Pathways. *Verbeke, F., et.al. (2017) Frontiers in Neuroscience 11:183.*

The left panel shows the typical Gram-negative quorum sensing mechanism. Acyl homoserine lactone molecules, synthesized by LuxI, passively pass the bacterial cell membrane and when a sufficient concentration is reached (threshold level) activate the intracellular LuxR which subsequently activates target gene expression in a coordinated way. Note that a single cell is shown for simplicity. However, acyl homoserine lactones will commonly diffuse and target neighboring cells within the colony to mediate a communal or population response within the bacterial colony.

The right panel shows that quorum-sensing peptides are synthesized by the bacterial ribosomes as pro-peptidic proteins and undergo posttranslational modifications during excretion by active transport. The quorum-sensing peptides bind membrane-associated receptors which get autophosphorylated and activate intracellular response regulators via phosphotransfer. These phosphorylated response regulators induce increased target gene expression.

For example, some microbial species, such as *Staphylococcus aureus*, can encase their community within a self-produced matrix of hydrated extracellular polymeric substances that include polysaccharides, proteins, nucleic acids, and lipid molecules. These encasements are known as biofilms. The formation of the *biofilm* on solid surfaces is a step-wise process comprising several stages, as shown in Figure 27.1.14 It starts with the conditioning of the surface through the coating with macromolecules from the aqueous surrounding, which enables the initial reversible adhesion of microorganisms. The next step is the formation of stronger, irreversible attachments to the surface, followed by the proliferation and aggregation of microorganisms into multicellular and multilayered clusters, which actively produce an extracellular matrix. Some cells in the mature biofilms continuously detach and separate from the aggregates, representing a continuous source of planktonic bacteria that can subsequently spread and form new microcolonies.

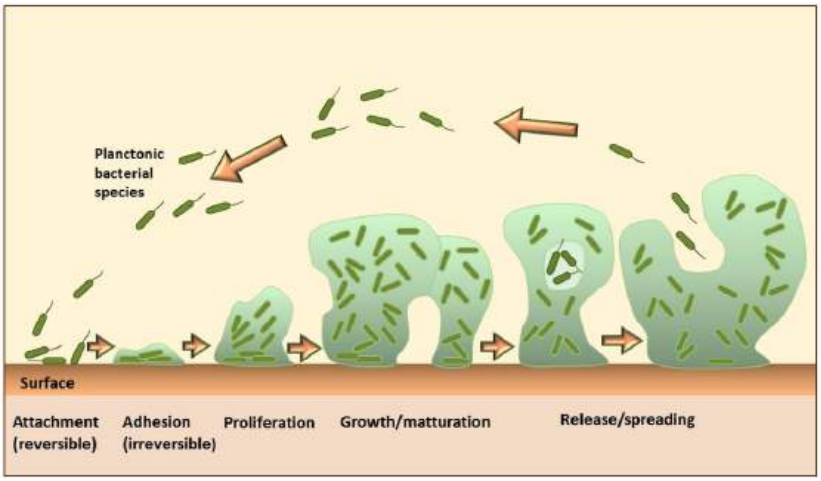


Figure 27.1.14 Schematic drawing of biofilm formation. *Rukavina, Z., and Vanic, Z. (2016) Pharmaceutics 8(2):18.*

Biofilms are a common cause of chronic, nosocomial (originating in a hospital), and medical device-related infections because they can develop either on vital or necrotic tissue as well as on the inert surfaces of different implanted materials. Moreover, biofilms are linked with high-level resistance to antimicrobials, frequent treatment failures, and increased morbidity and mortality. As a consequence, biofilm infections and accompanying diseases have become a major health concern and a serious challenge for both modern medicine and pharmacy. The rough estimation shows that more than 60% of hospital-associated infections are attributable to the biofilms formed on indwelling medical devices, which result in more than one million cases of infected patients annually and more than $1 billion in hospitalization costs per year in the USA.

Biofilm infections share some common characteristics: slow development in one or more hot spots, delayed clinical manifestation, and persistency for months or years, usually with interchanging periods of acute exacerbations and absence of clinical symptoms. Even though they are less aggressive than acute infections, their treatment is challenging to a greater extent. There is upto a 1000-fold decrease in the susceptibility of biofilms to antimicrobial agents and disinfectants as well as resistance to host immune response. Thus, ways to reduce or inhibit biofilm formation are highly sought. The majority of the proposed biofilm-control methods focus on: (i) prevention and minimization of biofilm formation by selection and surface modifications of anti-adhesive materials; (ii) debridement techniques including ultrasound and surgical procedures; (iii) disruption of biofilm QS-signaling system; or (iv) achieving proper drug penetration and delivery to formed biofilms by the use of an electromagnetic field, ultrasound waves, photodynamic activation or specific drug delivery systems.

27.1.5.3: Alternate σ Factors

Since the σ subunit of bacterial RNA polymerase confers specificity as to which promoters should be transcribed, altering the σ factor used is another way for bacteria to quickly and globally change what regulons are transcribed at a given time. The σ factor recognizes sequences within a bacterial promoter, so different σ factors will each recognize slightly different promoter sequences. In this way, when the cell senses specific environmental conditions, it may respond by changing which σ factor it expresses, degrading the old one and producing a new one to transcribe the operons encoding genes whose products will be useful under the new environmental condition. For example, in sporulating bacteria of the genera *Bacillus* and *Clostridium* (which include many pathogens), a group of σ factors controls the expression of the many genes needed for sporulation in response to sporulation-stimulating signals.

27.1.6: Prokaryotic Attenuation and Riboswitches

Although most gene expression is regulated at the level of transcription initiation in prokaryotes, there are also mechanisms to control both the completion of transcription, as well as translation, concurrently. Since their discovery, these mechanisms have been shown to control the completion of transcription and translation of many prokaryotic operons. Because these mechanisms link the regulation of transcription and translation directly, they are specific to prokaryotes, because these processes are physically separated in eukaryotes.

One such regulatory system is attenuation, whereby secondary stem-loop structures formed within the 5’ end of an mRNA being transcribed determine if transcription to complete the synthesis of this mRNA will occur and if this mRNA will be used for translation. Beyond the transcriptional repression mechanism already discussed, attenuation also controls the expression of the *trp*

operon in *E. coli* as shown in Figure 27.1.15 The *trp* operon regulatory region contains a leader sequence called *trpL* between the operator and the first structural gene, which has four stretches of RNA that can base pair with each other in different combinations. When a terminator stem-loop forms, transcription terminates, releasing RNA polymerase from the mRNA. However, when an antiterminator stem-loop forms, this prevents the formation of the terminator stem-loop, so RNA polymerase can transcribe the structural genes.

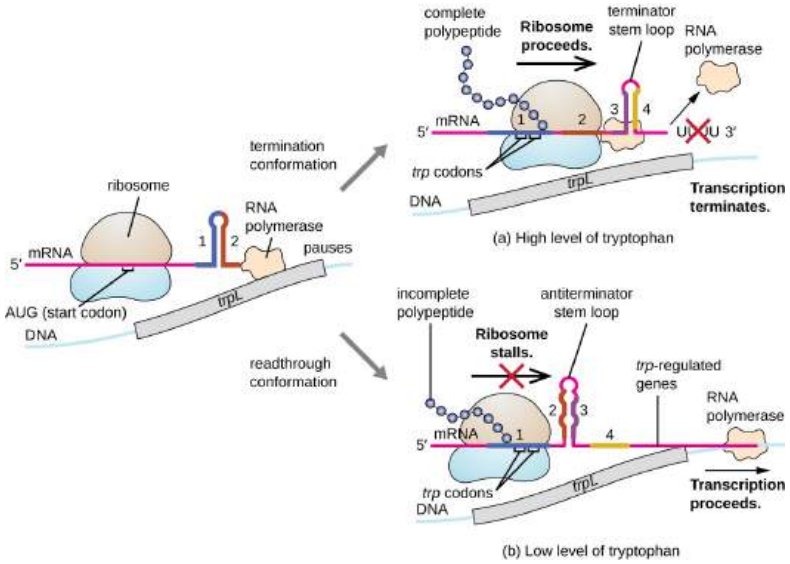


Figure 27.1.15: Attenuation of Transcription and Translation. *Parker, N., et. al. (2019) Microbiology. Openstax*

When tryptophan is plentiful, translation of the short leader peptide encoded by *trpL* proceeds, the terminator loop between regions 3 and 4 forms, and transcription terminates. When tryptophan levels are depleted, translation of the short leader peptide stalls at region 1, allowing regions 2 and 3 to form an antiterminator loop, and RNA polymerase can transcribe the structural genes of the *trp* operon.

A related mechanism of concurrent regulation of transcription and translation in prokaryotes is the use of a riboswitch, a small region of noncoding RNA found within the 5’ end of some prokaryotic mRNA molecules, as shown in Figure 27.1.16 A riboswitch may bind to a small intracellular molecule to stabilize certain secondary structures of the mRNA molecule. The binding of the small molecule determines which stem-loop structure forms, thus influencing the completion of mRNA synthesis and protein synthesis.

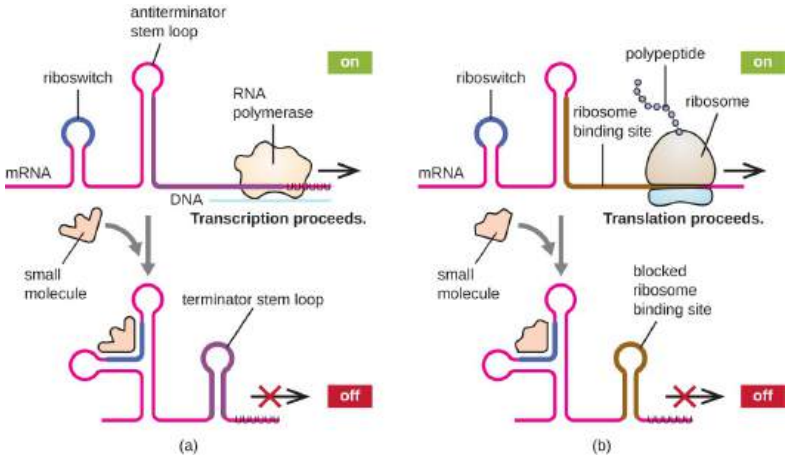
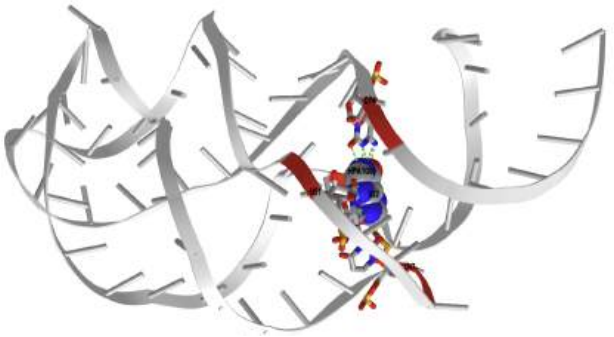


Figure 27.1.16: Riboswitch Form and function. Riboswitches found within prokaryotic mRNA molecules can bind to small intracellular molecules, stabilizing certain RNA structures, and influencing either the completion of the synthesis of the mRNA molecule itself (left) or the protein made using that mRNA (right). *Parker, N., et. al. (2019) Microbiology. Openstax*

Riboswitches found within prokaryotic mRNA molecules can bind to small intracellular molecules, stabilizing certain RNA structures, and influencing either the completion of the synthesis of the mRNA molecule itself (left) or the protein made using that mRNA (right).

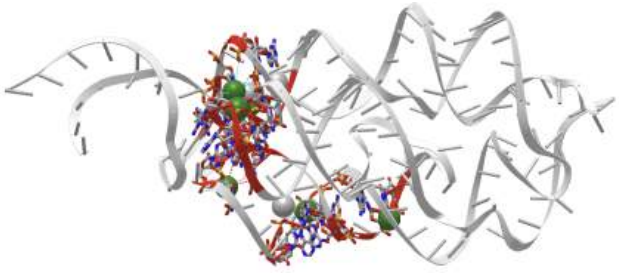
Figure 27.1.17shows interactive iCn3D models of a series of bacterial riboswitches. They are described in the legend below.

Guanine-responsive riboswitch bound to metabolite hypoxanthine (4FE5)



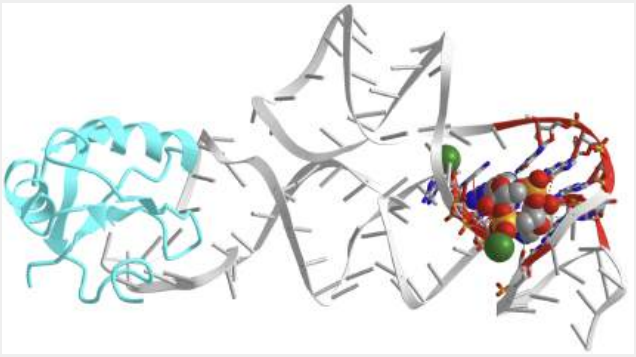
A. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...Nh1Sv7oSSyC6Z6

Divalent cation-sensing regulatory RNA (2QBZ)



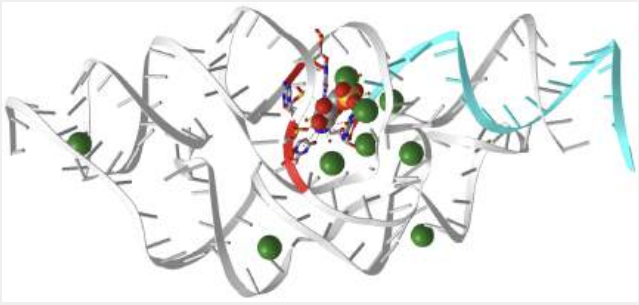
B. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...VbU9Vf6DUA4VE8

Cyclid-di-GMP RNA riboswitch (3IRW)



C. (Copyright; author via source). Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...wVntbdScP3VCF8

GlmS ribozyme bound to glucosamine-6-phosphate (2Z75)



D. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...w9T4K9ffK1NU27

A: Guanine-responsive riboswitch bound to metabolite hypoxanthine (4FE5) - Hypothanine, involved in purine metabolism, is shown bound to RNA representing the 5' untranslated region of the xanthine phosphoribosyltransferase (xbt)/ xanthine-specific purine permease (pbux) genes that lead to transcription termination.

B: The M-box in mycobacterial genes regulating Mg^{2+} transport binds divalent cation. They are transcribed under low Mg^{2+} concentrations. Salt bridges (ion-ion interactions) are shown in cyan and pi-cation interactions in red dotted lines

C: Bis-(3'-5')-cyclic dimeric guanosine monophosphate (c-di-GMP) is a second messenger in bacteria and regulates many cellular processes including the formation of biofilms. The riboswitch shown here is from Vibrio cholerae. The U1 small nuclear ribonucleoprotein A is shown in cyan. Figure 27.1.18 shows a cartoon of the actual riboswitch in the 5' untranslated region of target genes

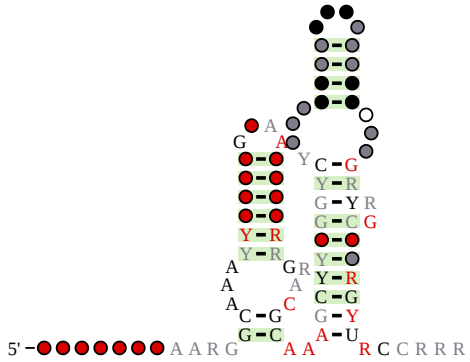


Figure 27.1.18" Consensus secondary structure and sequence conservation of Cyclic di-GMP-I riboswitch

D. This ribozyme is in the 5′ untranslated region of glucosamine-6-phosphate synthase mRNA. The protein enzyme, 2. This protein enzyme catalyzes the conversion of fructose 6-phosphate and glutamine to glucosamine 6-phosphate (GlcN6P) and glutamate. The *glmS* ribozyme in the 5'-untranslated region cleaves itself on binding GlcN6P. This self-cleavage is inhibited by glucose 6-phosphate (Glc6P). Hence high levels of the gene product for the synthase lead to cleavage of its own mRNA. The glmS ribozyme RNA is shown in gray and its substrate RNA in cyan.

27.1.7: References

1. Parker, N., Schneegurt, M., Thi Tu, A-H., Lister, P., Forster, B.M. (2019) Microbiology. *Openstax*. Available at: https://opentextbc.ca/microbiologyopenstax/
2. Chan, K-G., Liu, Y-C., and Chang C-Y. (2015) Inhibiting N-acyl-homoserine lactone synthesis and quenching *Pseudomonas* quinolone quorum sensing to attenuate virulence. *Front. Microbiol*. 6:1173. Available at: https://www.frontiersin.org/articles/10.3389/fmicb.2015.01173/full
3. Rukavina, Z., and Vanic Zeljka. (2016) Current trends in development of liposomes targeting bacterial biofilms. *Pharmaceutics* 8(2):18. Available at: https://www.mdpi.com/1999-4923/8/2/18/htm
4. Wikipedia contributors. (2020, April 18). Guanosine pentaphosphate. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:26, August 23, 2020, from https://en.Wikipedia.org/w/index.php?title=Guanosine_pentaphosphate&oldid=951778776
5. Verbeke, F., De Craemer, S., Debunne, N., Janssens, Y., Wynendaele, E., Van de Wiele, C., and De Spiegeleer. B. (2017) Peptides as quorum sensing molecules: measurement techniques and obtained levels *in vitro* and *in vivo. Frontiers in Neuroscience* 11:183. Available at: https://www.researchgate.net/publication/316055402_Peptides_as_Quorum_Sensing_Molecules_Measurement_Techniques_and_Obtained_Levels_In_vitro_and_In_vivo
6. Yesudhas, D., Batool, M., Anwar, M.A., Panneerselvam, S., and Choi, S. (2017) Proteins recognizing DNA: Structural uniqueness and versatility of DNA-binding domains in Stem Cell Transcription Factors. *Genes* 8(8):192. Available at: https://www.mdpi.com/2073-4425/8/8/192/htm
7. Castellanos, M., Mothi, N., and Muñoz, V. (2020) Eukaryotic transcription factors can track and control their target genes using DNA antennas. *Nature Comm.* 11:540. Available at: https://www.nature.com/articles/s41467-019-14217-8
8. Neideracher, G., Klopf, E., and Schüller, C. (2011) Interplay of dynamic transcription and chromatin remodeling: Lessons from yeast. *Int J Mol Sci* 12(8):4758-4769. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3179130/
9. Kim, S., and Kaang, B-K. (2017) Epigenetic regulation and chromatin remodeling in learning and memory. *Exp. & Mol. Med.* 49:e281. Available at: https://www.nature.com/articles/emm2016140#Fig1
10. Tvardovskly, A., Schwämmle, V., Kempf, S., Rogowska-Wrzesinka, A., and Jensen, O.N. (2016) Accumulation of histone variant H3.3 with age is associated with profound changes in the histone methylation landscape. *Nuc. Acids Res.* 45(16):1093. Available at: https://www.researchgate.net/publication/318987684_Accumulation_of_histone_variant_H33_with_age_is_associated_with_profound_changes_in_the_histone_methylation_landscape
11. Cipolletti, M., Fernandez, V.S., Montalesi, E., Marino, M., Fiochetti, M. (2018) Beyond the antioxidant activity of dietary polyphenols in cancer: The modulation of estrogen receptors (ERs) signaling. *Int J. Mol Sci* 19(9)2624. Available at: https://www.mdpi.com/1422-0067/19/9/2624/htm

12. Griekspoor, A., Zward, W., Neefjes, J., and Michalides, R. (2007) Visualizing the action of steroid hormone receptors in living cells. *Nucl. Recept. Signal*. 5:e003 Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1853070/
13. Mitsis, T., Papargeorgiou, L., Efthimiadou, A., Bacopoulou, F., Vlachakis, D., Chrousos, G.P., Eliopoulos, E. (2020) A comprehensive structural and functional analysis of the ligand binding domain of the nuclear receptor superfamily reveals highly conserved signaling motifs and two distinct canonical forms through evolution. *World Acad Sci J* 1: 264-274, 2019. Available at: https://www.spandidos-publications.com/10.3892/wasj.2020.30
14. Reed, S.M., and Quelle, D.E. (2015) p53 Acetylation: Regulation and consequences. *Cancers* 7(1):30-69. Available at: https://www.mdpi.com/2072-6694/7/1/30/htm.
15. Maclaine, N.J., and Hupp, T.R. (2009) The regulation of p53 by phosphorylation: a model for how distinct signals integrate into the p53 pathway. *Aging* 1(5):490-502. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2806026/
16. Wikipedia contributors. (2020, August 1). Estrogen. In *Wikipedia, The Free Encyclopedia*. Retrieved 01:28, September 6, 2020, from https://en.Wikipedia.org/w/index.php?title=Estrogen&oldid=970560042
17. Kluska, K., Adamczyk, J., and Krezel, A. (2018) Metal binding properties, stability, and reactivity of zinc fingers. *Coord. Chem Rev*. 367:18-64. Available at: https://www.sciencedirect.com/science/article/pii/S0010854517305441
18. Wikipedia contributors. (2020, July 4). Leucine zipper. In *Wikipedia, The Free Encyclopedia*. Retrieved 07:00, September 7, 2020, from https://en.Wikipedia.org/w/index.php?title=Leucine_zipper&oldid=965962667
19. Wikipedia contributors. (2020, April 15). Zinc finger. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:28, September 7, 2020, from https://en.Wikipedia.org/w/index.php?title=Zinc_finger&oldid=951116840
20. Horsthemke, B. (2018) A critical view on transgenrational epigenetic inheritence in humans. *Nat. Comm*. 9:2973. Available at: https://www.nature.com/articles/s41467-018-05445-5#rightslink

---

27.2: Regulation of Gene Expression in Eukaryotes

As seen in Chapter 26, the initiation of transcription requires the assembly of a multitude of transcription factors (TF) localized at the promoter region. Transcription can also utilize far-reaching interactions of enhancers, that bind at a distant DNA site and loop back around to stabilize the RNA polymerase at the promoter. Control of transcriptional initiation is dependent on TF factor activation, TF binding with specific DNA recognition sequences, and chromatin remodeling.

27.2.1: Transcription Factor (TF) Activation

Many TFs are expressed within cells and held in an inactive conformation until the right environmental stimulus is present within the cell. Cellular signaling pathways can cause post-translational protein modifications leading to TF activation or small molecules may physically bind and allosterically modify the protein structure to mediate activation. Here we will use examples from the cell cycle signaling cascade and steroid hormone receptor pathways to highlight some mechanisms of TF activation. A key element to take away from this section is that transcription factor activation is often highly pleiotropic and has many cellular effects. Depending on the cell type and the environmental conditions, different combinations of downstream target genes may be activated or inactivated. Teasing apart these intricacies and the physiological effects that they have within an organism is a major goal of ongoing research.

27.2.1.1: Cell Cycle Regulation by p53

p53 is one of the most studied proteins in science. To date, over 68,000 papers appear in PubMed containing p53 or TP53 in the title and/or abstract. Originally described as an oncogene (since a mutated, functionally altered form of the protein was first characterized), p53 is now recognized as the most frequently inactivated tumor suppressor in human cancers. It is a transcription factor that controls the expression of genes and miRNAs affecting many important cellular processes including proliferation, DNA repair, programmed cell death (apoptosis), autophagy, metabolism, and cell migration, as shown in Figure 27.2.1. Many of those processes are critical to a variety of human pathologies and conditions extending beyond cancer, including ischemia, neurodegenerative diseases, stem cell renewal, aging, and fertility. Notably, p53 also has non-transcriptional functions, ranging from intrinsic nuclease activity to activation of mitochondrial Bak (Bcl-2 homologous antagonist killer) and caspase-independent apoptosis.

As a transcription factor, p53 responds to various genotoxic insults and cellular stresses (e.g., DNA damage or oncogene activation) by inducing or repressing the expression of over a hundred different genes. p53 transcriptional regulation plays a dominant role in causing the arrest of damaged cells, facilitating their repair and survival, or inducing cell death when DNA is damaged irreparably. p53 can also cause cells to become permanently growth arrested, and there is compelling *in vivo* evidence that these “senescent” cells secrete factors that enhance their clearance by the immune system, leading to tumor regression. Through these mechanisms, p53 helps maintain genomic stability within an organism, justifying its long-held nickname “guardian of the genome”. Other p53 gene targets are involved in inhibiting tumor cell angiogenesis, migration, metastasis, and other important processes (such as metabolic reprogramming) that normally promote tumor formation and progression

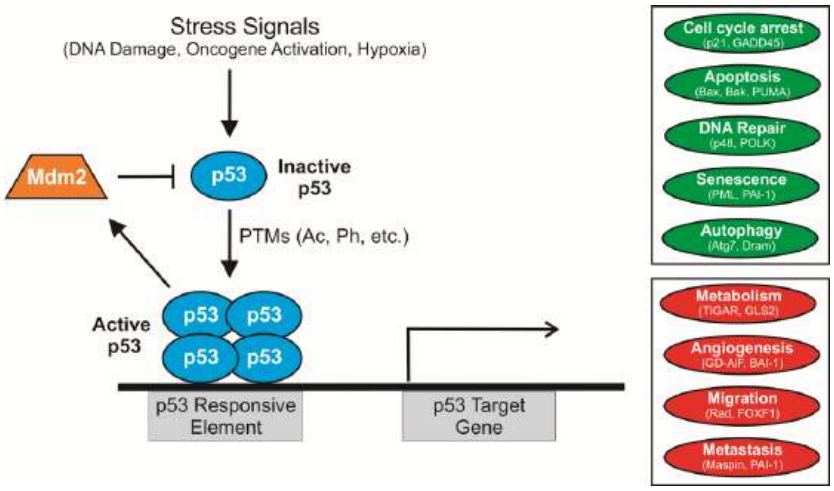


Figure 27.2.1: Cellular stress leads to p53 transcriptional activation of downstream targets. Reed, S.M., and Quelle, D.E. (2015) 7(1):30-69.

Figure 27.2.2 shows an interactive iCn3D model of the human p53 tetramer bound to the natural CDKN1A(p21) p53-response element (3TS8).

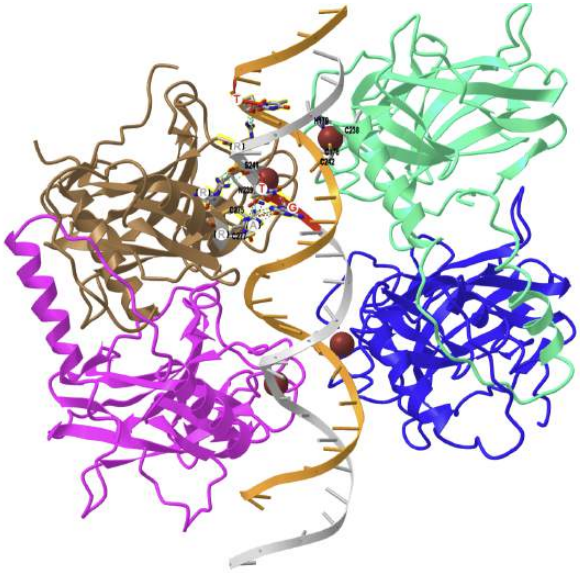


Figure 27.2.2: Human p53 tetramer bound to the natural CDKN1A(p21) p53-response element (3TS8). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...cvfiRBEzfxJuk6

Each monomer of the p53 tetramer is shown in a different color. The noncovalent interactions of the brown monomer with the DNA are shown, with key amino acids and nucleic acid bases shown in CPK-colored sticks. They interact with the Zn^{2+} ions, shown in the light green monomer.

Normally, p53 levels are kept low by its major antagonist, Mdm2, an E3 ubiquitin ligase that is itself a transcriptional target of p53. Stress signals, such as DNA damage, oncogene activation, and hypoxia, promote p53 stability and activity by inducing post-translational modifications (PTMs) and tetramerization of p53. p53 functions as a transcription factor that binds to specific p53 response elements upstream of its target genes. p53 affects many important cellular processes linked to tumor suppression, including the induction (green) of senescence, apoptosis, and DNA repair as well as inhibition (red) of metabolism, angiogenesis, and cell migration. These functions are largely mediated through transcriptional regulation of its targets (examples given).

p53 protein function is regulated post-translationally by coordinated interaction with signaling proteins including protein kinases, acetyltransferases, methyl-transferases, and ubiquitin-like modifying enzymes as shown in Figure 27.2.3 The majority of the sites of covalent modification occur at intrinsically unstructured linear peptide docking motifs that flank the DNA-binding domain of p53 which plays a role in anchoring or in allosterically activating the enzymes that mediate covalent modification of p53. In undamaged cells, the p53 protein has a relatively short half-life and is degraded by a ubiquitin-proteasome-dependent pathway through the action of E3 ubiquitin ligases, such as MDM2 (Fig 28.3.1). Following stress, p53 is phosphorylated at multiple residues, thereby modifying its biochemical functions required for increased activity as a transcription factor. Post-translational modifications help to stabilize the tetramer formation of the protein and enhance the translocation of the protein from the cytoplasm into the nucleus. The tetrameric form of p53 is then functional to bind to DNA in a sequence-specific manner and either activate or repress transcription, depending on the target sequence. Some post-translational modifications, such as acetylation, are DNA-dependent and can play a role in chromatin remodeling and activation of p53 target gene expression.

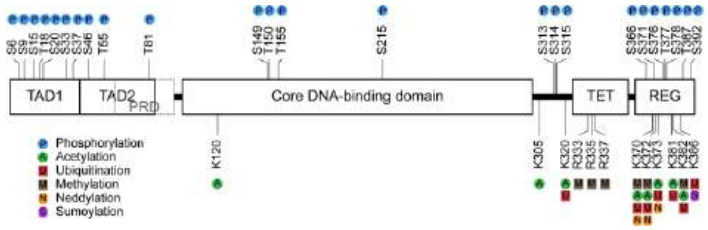


Figure 27.2.3: Sites of Post-Translational Modification on p53. Schematic representation of the 393 amino acid domain structure of human p53 showing the sites of post-translational modification including phosphorylation, acetylation, ubiquitination, methylation, neddylation, and sumoylation. Abbreviations: N-terminal transactivation domain (TAD); proline-rich domain (PRD); tetramerization domain (TET); C-terminal regulatory domain (REG); arginine (R); lysine (K); serine (S); threonine (T). *Maclaine, N.J., and Hupp T.R. (2009) Aging 1(5):490-502*

It should be noted that single-point mutations that modify the ability of the protein to be phosphorylated in one position, typically do not show a decrease in the stabilization or activation of the protein following a damage or stress event. Thus, multiple modifications likely allow for redundancy within this pathway and ensure the activation of the protein following a stress event. Furthermore, the environment within the cell can lead to different p53 phenotypes, such as the activation of growth arrest and DNA repair processes (ie if there is not a lot of damage) or it can lead to the activation of apoptosis or programmed cell death pathways (ie if the damage is too extensive to be repaired).

27.2.1.2: Steroid Hormone Receptors

Steroid hormone receptors (SHRs) belong to the superfamily of *nuclear receptors (NRs)*, which are one of the essential classes of transcriptional factors. NRs play a critical role in all aspects of human development, metabolism. and physiology. Since they generally act as ligand-activated transcription factors, they are an essential component of cell signaling. NRs form an ancient and conserved family that arose early in the metazoan lineage. NR molecular evolution is characterized by major events of gene duplication and gene losses. Phylogenetic analysis revealed a distinct separation of NR ligand binding domains (LBDs) into 4 monophyletic branches, the steroid hormone receptor-like cluster, the thyroid hormone-like receptors cluster, the retinoid X-like and steroidogenic factor-like receptor cluster and the nerve growth factor-like/HNF4 receptor cluster, as shown in Figure 27.2.4

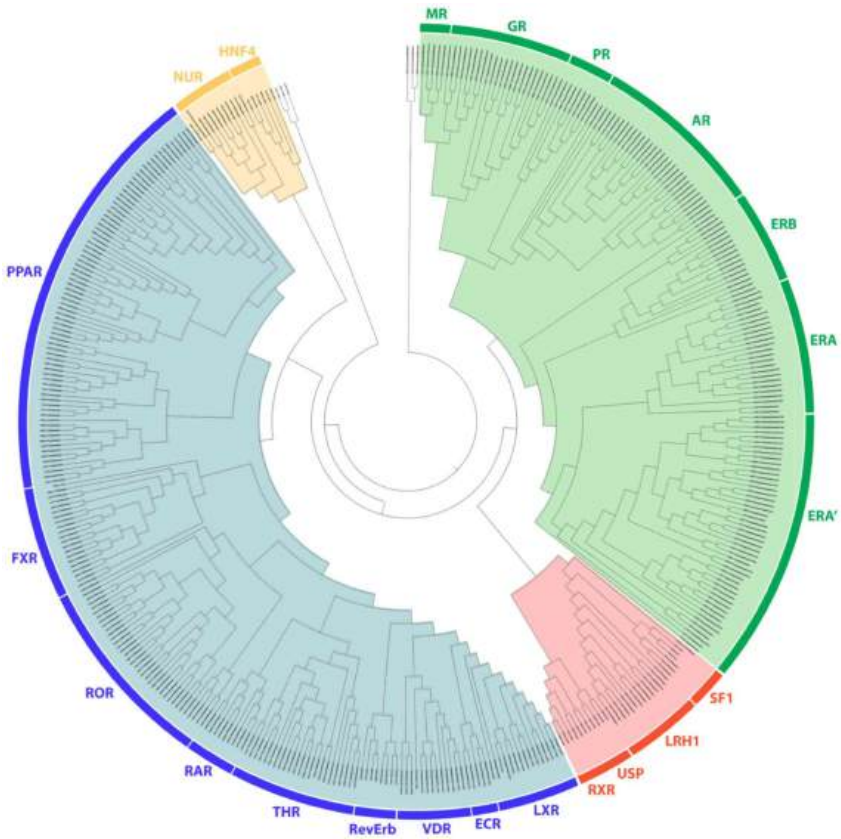


Figure 27.2.4: **Phylogenetic tree of the nuclear receptors' ligand binding domain.** Four distinct monophyletic branches are visible. Those monophyletic branches are divided into subcategories. The phylogenetic trees confidently separate the steroid hormone-like (branch colored green), the retinoid X-like and steroidogenic factor-like receptors cluster (branch colored orange), the thyroid hormone-like receptors cluster (branch colored blue), and the nerve growth factor-like/hepatocyte nuclear factor-4 receptors cluster (branch colored yellow). *Mitsis, T., et. al. (2020) World Acad Sci J 1: 264-274, 2019*

Here we will focus on the Steroid Hormone-Like Receptors branch (SHRs). SHRs plays a key role in many important physiological processes like organ development, metabolite homeostasis, and response to external stimuli. The estrogen receptor comes in two major forms, ERα and ERβ. Other members of this subgroup include the cortisol-binding glucocorticoid receptor (GR), the aldosterone-binding mineralocorticoid receptor (MR), the progesterone receptor (PR), and the dihydrotestosterone (DHT) binding androgen receptor (AR), as shown in Figure \(\PageIndex{5}\ below.

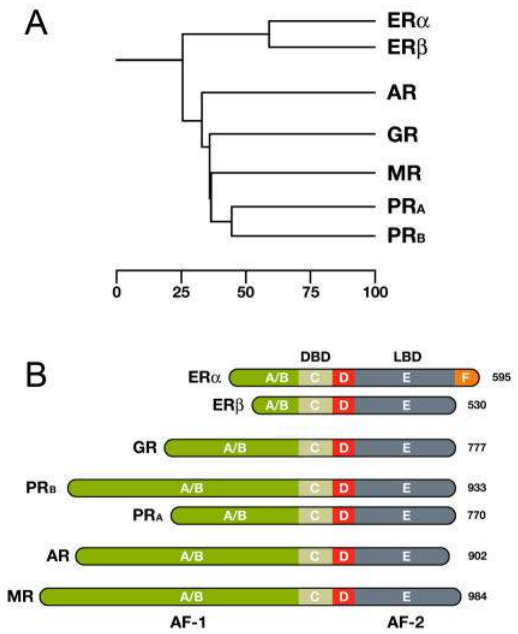


Figure 27.2.5: Overview of Steroid Hormone Receptor Family (SHR). *riekspoor, A., et. al. (2007) Nucl Recept Signal. 5:e003* A. Phylogenetic tree of the Steroid Hormone Receptor (SHR) family showing the evolutionary interrelationships and distance between the various receptors. Based on alignments available at The NucleaRDB [Horn et al., 2001]. B. All steroid receptors are composed of a variable N-terminal domain (A/B) containing the AF-1 transactivation region, a highly conserved DNA Binding Domain (DBD), a flexible hinge region (D), and a C-terminal Ligand Binding Domain (LBD, E) containing the AF-2 transactivation region. The estrogen receptor α is unique in that it contains an additional C-terminal F domain. Numbers represent the length of the receptor in amino acids.

Panel A shows a phylogenetic tree of the Steroid Hormone Receptor (SHR) family showing the evolutionary interrelationships and distance between the various receptors. Based on alignments available at The NucleaRDB [Horn et al., 2001].

Panel B shows that all steroid receptors are composed of a variable N-terminal domain (A/B) containing the AF-1 transactivation region, a highly conserved DNA Binding Domain (DBD), a flexible hinge region (D), and a C-terminal Ligand Binding Domain (LBD, E) containing the AF-2 transactivation region. The estrogen receptor α is unique in that it contains an additional C-terminal F domain. Numbers represent the length of the receptor in amino acids.

The members of the Steroid Hormone Receptor family share a similar, modular architecture, consisting of several independent functional domains (Fig. 5B above). Most conserved is the centrally located DNA binding domain (DBD) containing the characteristic zinc-finger motifs. The DBD is followed by a flexible hinge region and a moderately conserved Ligand Binding Domain (LBD), located at the carboxy-terminal end of the receptor. The estrogen receptor α is unique in that it contains an additional F domain of which the exact function is unclear. The LBD is composed of twelve α-helices (H1-H12) that together fold into a canonical α-helical sandwich. Besides its ligand binding capability, the LBD also plays an important role in nuclear translocation, chaperone binding, receptor dimerization, and coregulator recruitment through its potent ligand-dependent transactivation domain, referred to as AF-2. A second, ligand-independent, transactivation domain is located in the more variable N-terminal part of the receptor, designated as AF-1. To date, no crystal structure of a full-length SHR exists, though structures of the DBD and LBD regions of most SHRs are available. These have helped significantly in understanding the molecular aspects of DNA and ligand binding, but have to some extent also led to biased attention to these parts of the receptor only. For example, many coregulator interaction studies are still performed with the LBD only, while numerous studies have demonstrated that the AF-2 domain often tells only part of the story. With the help of biophysical techniques, however, it is feasible to study the full-length receptor in its native environment.

Most SHRs remain in the cytoplasm of the cell until they are bound with the appropriate steroid as shown in Figure 27.2.6 Steroid binding causes the dimerization of SHRs and localization to the cell nucleus, where the SHRs interact with the DNA at sequence-specific motifs known as Hormone Response Elements (HREs) (Step 5). Many SHRs can also interact with membrane-bound receptors and affect cellular signaling pathways, in addition to the activation of gene expression (step 6).

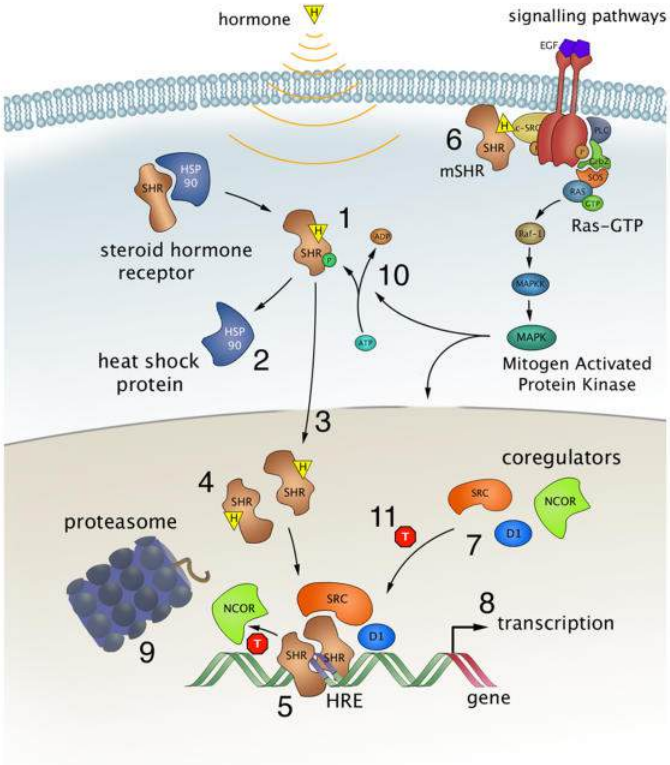


Figure 27.2.6: Steroid Hormone Receptors (SHR) act as hormone-dependent nuclear transcription factors. Upon entering the cell by passive diffusion, the hormone (H) binds the receptor, which is subsequently released from heat shock proteins, and translocates to the nucleus. There, the receptor dimerizes, binds specific sequences in the DNA, called Hormone Responsive Elements or HREs, and recruits a number of coregulators that facilitate gene transcription. *Griekspoor, A., et. al. (2007) Nucl Recept Signal. 5:e003*

Steroid hormones, such as estrogens, reach their target cells via the blood, where they are bound to carrier proteins. Naturally occurring estrogens including estradiol, estrone, estriol, differ primarily in structure on the presence of hydroxyl groups (Fig. 28.3.6). Estradiol is the predominant estrogen during reproductive years both in terms of absolute serum levels as well as in terms of estrogenic activity. During menopause, estrone is the predominant circulating estrogen, and during pregnancy, estriol is the predominant circulating estrogen in terms of serum levels. Another type of estrogen called estetrol (E4) is produced also produced predominantly during pregnancy as shown in Figure 27.2.7. Estrogens function in many physiological processes, including the regulation of the menstrual cycle and reproduction, maintaining bone density, brain function, cholesterol mobilization, and maturation of reproductive organs during development, and they play a role in controlling inflammation.

Figure 27.2.7: Naturally Occurring Estrogens. *Wikipedia (2020) Estrogen.*

Because of their lipophilic nature, it is thought that steroid hormones, such as estrogen, pass the cell membrane by simple diffusion, although some evidence exists that they can also be actively taken up by the endocytosis of carrier protein-bound hormones. For a

long time, it has been assumed that binding of the ligand resulted in a simple on/off switch of the receptor (Fig. 6, step 1). While this is likely the case for typical agonists like estrogen and progesterone, this is not always correct for receptor antagonists, used in drug therapy. These antagonists come in two kinds, so-called partial antagonists (for the estrogen receptors known as SERMs for Selective Estrogen Receptor Modulators) and full antagonists. The partial antagonist can, depending on cell type, act as a SHR agonist or antagonist. In contrast, full antagonists (for ER known as SERDs for Selective Estrogen Receptor Downregulators) always inhibit the receptor, independent of cell type, in part by targeting the receptor for degradation. Binding of either type of antagonist results in major conformational changes within the LBD and in the release from heat shock proteins that thus far had protected the unliganded receptor from unfolding and aggregation (Fig. 6 step 2).

Figure 27.2.8 shows an interactive iCn3D model of the Androgen Receptor DNA-Binding Domain Bound to a Direct Repeat Response Element (1R4I).

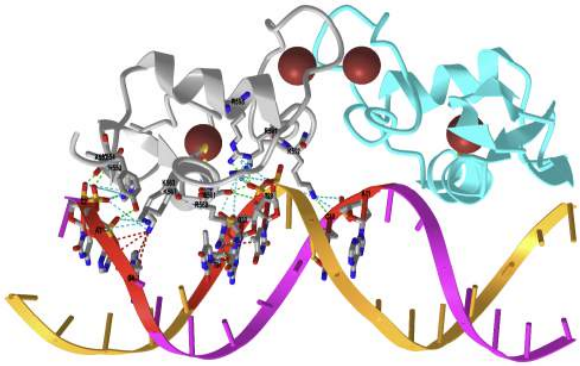


Figure 27.2.8: Androgen Receptor DNA-Binding Domain Bound to a Direct Repeat Response Element (1R4I). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...FKLa7aaZiq6Hz8

27.2.2: Transcription Factor (TF) Recognition and Binding to DNA

TF controls gene expression by binding to their target DNA site to recruit, or block, the transcription machinery onto the promoter region of the gene of interest. Their function relies on the ability to find their target site quickly and selectively. In living cells, TFs are present in nM concentrations and bind the target site with comparable affinity, but they also bind any DNA sequence (nonspecific binding), resulting in millions of low affinity (i.e., $>10^{-6}$ M) competing sites. Nonspecific binding facilitates the search for the target site by three major mechanisms as shown in Figure 27.2.9.

The second scenario is a ‘hopping’ mechanism, in which a TF might hop from one site to another in 3D space by dissociating from its original site and subsequently binding to the new site. This may happen within the same chain and re-association occurs adjacent to the former dissociated site. A third search mechanism is described as ‘intersegmental transfer’. In this scenario, the protein moves between two sites via an intermediate ‘loop’ formed by the DNA and subsequently binds at two different DNA sites. This mechanism applies to TFs with two DNA-binding sites. Proteins with two DNA-binding sites can occasionally bind non-specifically to two locations situated far apart within the DNA strand, that are brought into close contact through the formation of these loops. Such TFs transfer across a point of close contact without dissociating from the DNA.

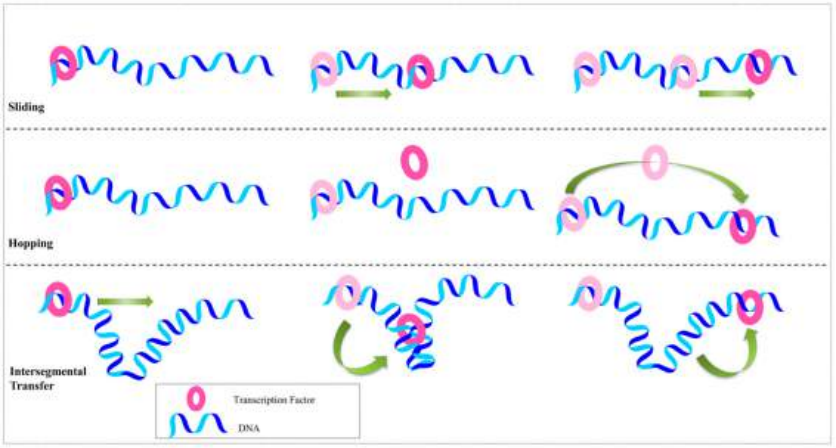


Figure 27.2.9: **Protein-DNA recognition mechanisms.** The main three protein-DNA recognition mechanisms are shown. *Yesudhas, D., et.al. (2017) Genes 8(8):192*

Top: When the transcription factor (pink ring) moves from one site to another by sliding along the DNA and is transferred from one base pair to another without dissociating from the DNA, this mechanism is called sliding.

Center: Hopping occurs when the transcription factor moves on the DNA by dissociating from one site and re-associating with another site.

Bottom: Intersegmental transfer describes the mechanism by which the transcription factor gets transferred through DNA bending or the formation of a DNA loop, resulting in the protein being bound transiently to both sides and subsequently moving from one site to the other.

One of the main scenarios involves a 'sliding' mechanism, in which the protein moves from its initial non-specific site to its actual target site by sliding along the DNA (also known as 1-dimensional (1D) sliding). When the TF starts to move and shift counterions from the phosphate backbone, the same number of counterions binds to the site left free by the protein. The sliding rate is also dependent on the hydrodynamic radius of the protein; the required rotational movement over the DNA backbone is greater for larger proteins, that tend to slide slowly.

Recent Updates: 9/25/23

The sliding model that was proposed by Von Hippel and Berg suggests that DNA-binding protein exists in two interconverting conformations. One is a specific form (O) that can bind to a target DNA sequence, such as an operator in DNA, through specific hydrogen bonds (along with electrostatic interaction) characterized by a low K_D. The other is a nonspecific form (D) that binds mainly with weaker affinity through electrostatic interactions and a high K_D. Nonspecific binding brings the protein to the DNA surface. Dynamic conformational changes from the O to D allow sampling of hydrogen bonds between donor and acceptors in the protein and in the major grove of the protein. The protein can diffuse much more quickly along the DNA to find its target site since the search for the specific target site is now effectively 1D instead of 3D. There is no thermodynamic barrier to sliding since counterions that leave the DNA when bound to the protein rebind behind it as the protein slides. These processes are illustrated in Figure 27.2.10below.

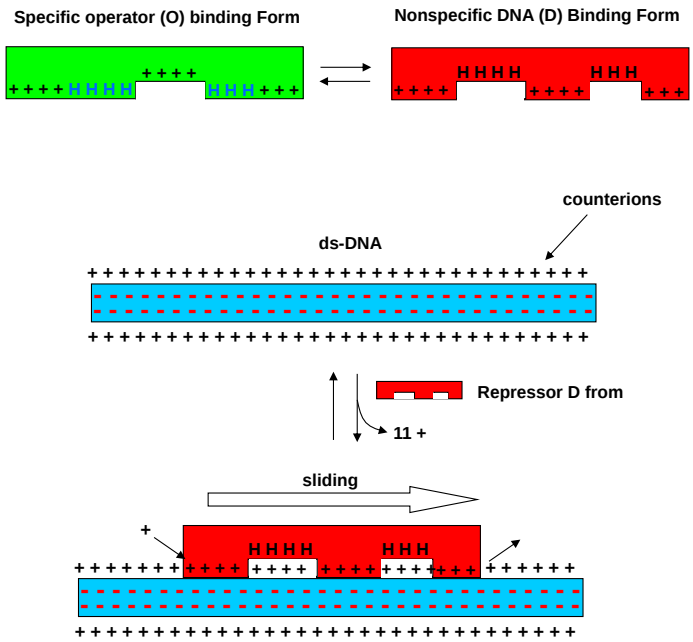


Figure 27.2.10 Two-state model sliding model of DNA protein binding to its specific binding site (operator)

There are a large number of overlapping nonspecific binding sites (let's say each is 6 base pairs in length), which also help drive the nonspecific binding of the protein to the DNA through entropy increases (i.e consider the probability of binding to 1 site on the DNA versus multitudes of overlapping sites). Experiments show that the on-rate for a DNA binding protein for finding its target site increases with increasing length of the DNA molecule the specific site is embedded in. The opposite would be expected given the diffusion rate of large molecules. In fact, the k_{on} was found to be greater than diffusion-controlled limits, which can be explained by the reduced dimensionality of the search for the specific site when the protein is loosely bound through nonspecific electrostatic interactions that enable the 1D search.

Each eukaryotic TF controls tens to hundreds of genes scattered throughout the genome, and expressing each gene needs various TFs simultaneously binding to their sites to form the transcription complex, an extremely rare event in probabilistic terms. As a result, the *in vivo* site occupancy patterns of eukaryotic TFs are more complex than predicted by their *in vitro* site-specific binding profiles and do not strongly correlate with the actual levels of gene expression. An interesting feature highlighted by genome analysis is an accumulation of potential TF binding sites in regions flanking eukaryotic genes. Such clusters of degenerate recognition sites are assumed to be key for transcription control and thus are generally classified as gene regulatory regions (RR). For example, the affinity of the *Drosophila* TF Engrailed to the RRs of its target genes is strongly amplified by long tracts of degenerate consensus repeats that are present in such regions.

Role of Short Tandem Repeats (STRs) in the Genome

As we have mentioned previously, only about 1.5% of the human genome encodes genes for actual proteins. Much of the genome is transcribed at low levels into RNAs, some of which have clearly defined functions (examples include rRNA, mRNAs, and regulatory RNAs). A large part is presumably involved in facilitating the 3D organization of the genome and its dynamic architecture, which determines its replicative and transcriptive access. One poorly understood feature of genomic DNA is **short tandem repeats (STRs)**. These repeats stretch up to 100 nucleotides in length with each repetitive tandem repeat running from just 1 to 6 bases long. They comprise about 6% of the genome (compared to the 1.5% for protein-coding genes), and are found in abundance in chromatin that is transcribed mRNA for proteins. For example, a specific DNA sequence might be GTCACGTGAC while a small STR would be $(CG)_6C(CG)_{11}$

Specifically, STRs surround sites where classic transcription factors (TF) bind. As we described in this and the previous chapter sections, TFs bind through specific DNA binding motifs (like helix-loop-helix or Zn^{2+} fingers) to consensus sequences (such as response elements and enhancers) in the DNA as they function to control transcription. The binding of transcription factors or other proteins to target sequences occurs initially through nonspecific electrostatic interactions, followed by a dimensionally restricted diffusion along the DNA as the protein finds its specific sequence target.

In contrast, the STRs offer little sequence uniqueness for high affinity, low K_D binding sites for specific protein interactions, so the question remains as to how they express function, which their omnipresence suggests they have. Studies by Horton et al (Science 381, 1304 (2023) have shown that classic TFs do indeed bind STRs, albeit at lower affinity (higher K_D) compared to their binding to classic TF-specific sequences. The bind with higher affinity than to nonspecific DNA sequences. Just as nonspecific interactions facilitate TF binding to promoter sites, so do the multiple STRs that straddle the DNA binding element.

If a TF binds its isolated target DNA with a low effective K_D (high affinity) and $-\Delta G^0$ value, a target DNA surrounded by multiple STRs would have an even lower effective K_D (higher affinity) and even more $-\Delta G^0$ value. They do so by increasing the effective on rate (k_{on}) for protein binding. (Remember that $K_D = k_{off}/k_{on}$.) The effective size of the target for the TF becomes greater when it is an “island” in the middle of a STR “sea”. The increased affinity stems in large part from the more favorable entropy of having the TF bind not to just 1 site but effectively to multiple overlapping sites. This also increases the localized concentration of TF near the specific site which drives binding. The k_{off} is not expected to change.

Another effect of the STRs on TF binding is that multiple DNA-binding proteins can bind to the same site through their interactions with STRs at the site, leading to new ways to regulate gene transcription. The group studies just two TFS but sequence analyses suggest that many TF would use a similar mechanism.

27.2.3: Histone Modification and Chromatin Remodeling

Regulation of transcription involves dynamic rearrangements of chromatin structure. Recall that eukaryotic DNA is complexed with histone octamers, which are composed of dimers of the core histones H2A, H2B, H3, and H4. 147 bp of DNA are wrapped 1.65 times around each octamer forming nucleosomes, the basic packaging units of chromatin. Nucleosomes, connected by linker DNA of variable length as “beads on a string”, generate the 11 nm linear structure. The linker histone H1 is positioned at the top of the core histone octamer and enables higher organized compaction of DNA into transcriptionally inactive 30 nm fibers.

To understand the role of chromatin in the regulation of transcription it is important to know where nucleosomes are positioned and how the positioning is achieved. Basically, there are four groups of activities that change chromatin structure during transcription: (1) histone modifications, (2) eviction and repositioning of histones, (3) chromatin remodeling, and (4) histone variant exchange. Histone modifiers introduce post-translational, covalent modifications to histone tails and thereby change the contact between DNA and histones. These modifications govern access to regulatory factors. Histone chaperones aid in the eviction and positioning of histones. A third class of chromatin restructuring factors is ATP-dependent chromatin remodelers. These multi-subunit complexes utilize energy from ATP hydrolysis for various chromatin remodeling activities including nucleosome sliding, nucleosome displacement, and the incorporation and exchange of histone variants.

Post-translational modifications (PTMs) of histone proteins are a primary mechanism that controls chromatin architecture. Over 20 distinct types of histone PTMs have been described, among which the most abundant ones are acetylation and methylation of lysine residues. Histone PTMs can be deposited on and removed from chromatin by different enzymes, known as histone PTM ‘writers’ and ‘erasers’. Histone PTMs exert their regulatory effects via two main mechanisms. First, histone PTMs serve as docking sites for various nuclear proteins—histone PTM ‘readers’—that specifically recognize modified histone residues through their modification-binding domains. Recruitment of these proteins at specific genomic loci promotes key chromatin processes, such as transcriptional regulation and DNA damage repair. Second, some histone PTMs, such as acetylation, directly affect chromatin's higher-order structure and compaction, thereby controlling chromatin accessibility to protein machinery such as those involved in transcription. Chromatin may adopt one of two major states interchangeably. These states are *heterochromatin* and *euchromatin*. *Heterochromatin* is a compact form that is resistant to the binding of various proteins, such as transcriptional machinery. In contrast, *euchromatin* is a relaxed form of chromatin that is open to modifications and transcriptional processes, as shown in Figure 27.2.10 Histone methylation promotes the formation of Heterochromatin whereas, histone acetylation promotes euchromatin.

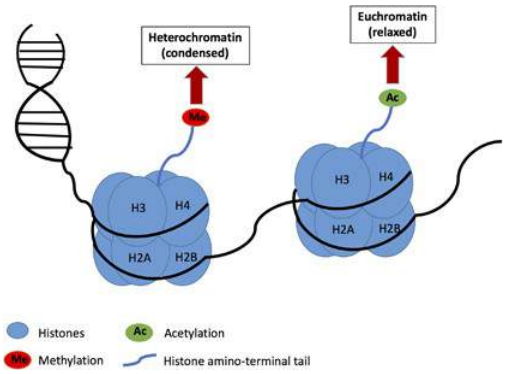


Figure 27.2.11: Schematic drawing of histone methylation and acetylation in relation to chromatin remodeling. *Kim, S., and Kaang, B-K. (2017) Exp & Mol. Med. 49:e281*

The addition of methyl groups to the tails of histone core proteins leads to histone methylation, which in turn leads to the adoption of a condensed state of chromatin called 'heterochromatin.' Heterochromatin blocks transcription machinery from binding to DNA and results in transcriptional repression. The addition of acetyl groups to lysine residues in the N-terminal tails of histones causes histone acetylation, which leads to the adoption of a relaxed state of chromatin called 'euchromatin.' In this state, transcription factors and other proteins can bind to their DNA binding sites and proceed with active transcription.

Chromatin remodeling can also be an ATP-dependent process and involve histone dimer ejection, full nucleosome ejection, nucleosome sliding, and histone variant exchange as shown in Figure 27.2.12 ATP-dependent chromatin remodeling complexes bind to nucleosome cores and the surrounding DNA, and, using energy from ATP hydrolysis, they disrupt the DNA-histone interactions, slide or eject nucleosomes, alter nucleosome structures, and modulate the access of transcription factors to the DNA (Figure 28.3.9). In addition to modulating gene expression, some of the complexes are involved in nucleosome assembly and organization, following transcription at locations in which nucleosomes have been ejected, packing of DNA, following replication, and DNA repair.

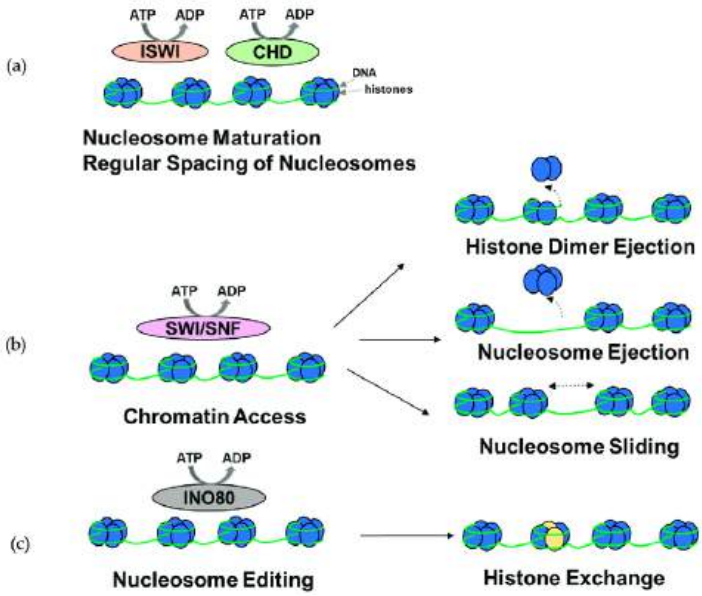


Figure 27.2.12: Overview of the functions of ATP-dependent chromatin remodeling complexes. *Hasan, N., and Ahuja, N. (2019) Cancers 11(12):1859*

Panel (a) shows a subset of ISWI and CHD complexes is involved in nucleosome assembly, maturation, and spacing.

Panel (b) shows SWI/SNF complexes are primarily involved in histone dimer ejection, nucleosome ejection, and nucleosome repositioning through sliding, thus modulating chromatin access.

Panel (c) shows INO80 complexes are involved in histone exchange. It should be noted that the complexes might be involved in other chromatin remodeling functions.

Figure 27.2.13 shows the effects of Histone Variant H3.3 on *C. elegans* Lifespan

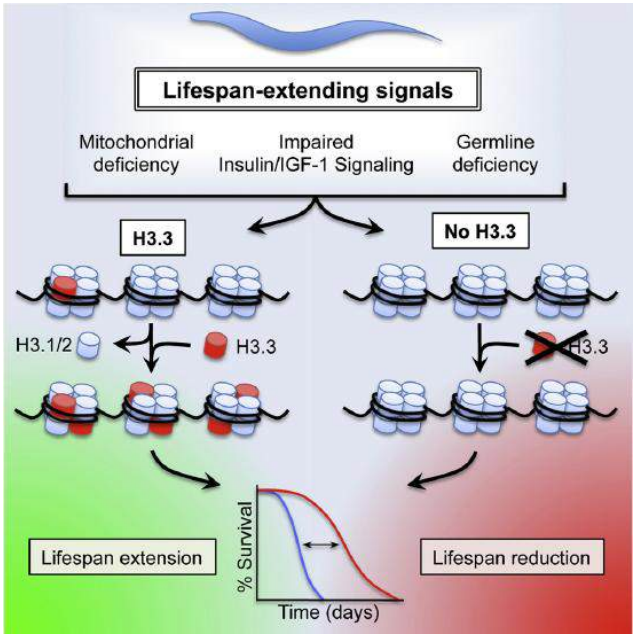


Figure 27.2.13: The Effects of Histone Variant H3.3 on *C. elegans* Lifespan. H3.3 expression increases over time in *C. elegans* during their normal lifespan. In organisms with impaired Inulin/IGF-1 signaling, germline deficiency of H3.3 resulted in significant decreases in lifespan. Piazzesi, A., et. al. (2016) Cell Rep 17(4):987-996.

27.2.4: Protein-DNA Interactions

Proteins use a wide range of DNA-binding structural motifs, such as homeodomain (HD), helix-turn-helix (HTH), and high-mobility group box (HMG) to recognize DNA. HTH is the most common binding motif and can be found in several repressor and activator proteins, as shown in Figure 27.2.14 Despite their structural diversity, these domains participate in a variety of functions that include acting as substrate interaction mediators, enzymes to operate DNA, and transcriptional regulators. Several proteins also contain flexible segments outside the DNA-binding domain to facilitate specific and non-specific interactions. For example, many HD proteins use N-terminal arms and a linker region to interact with DNA. The Encyclopedia of DNA Elements (ENCODE) data suggest that about 99.8% of putative binding motifs of TFs are not bound by their respective TFs in the genome. It is, therefore, clear that the presence of a single binding motif per TF is not adequate for TF binding.

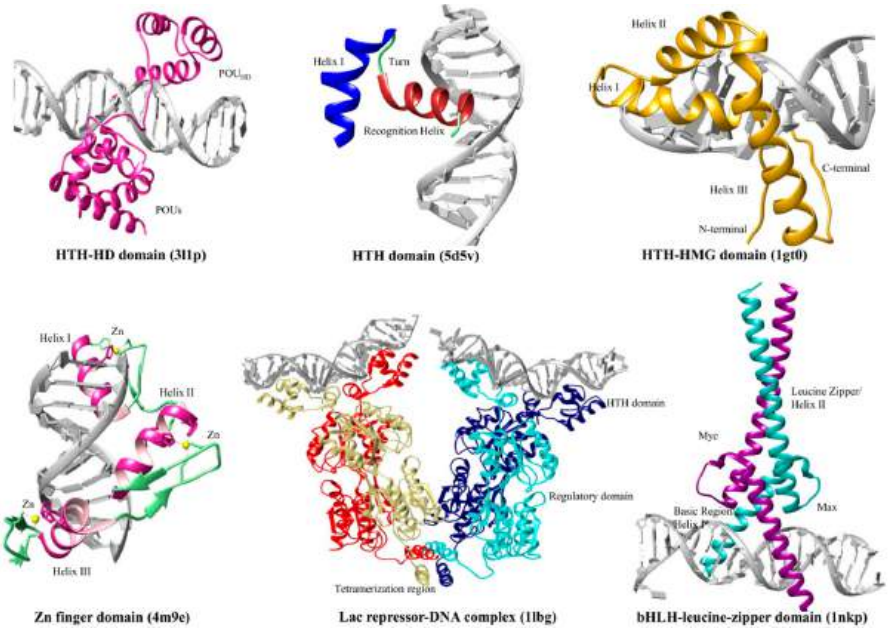
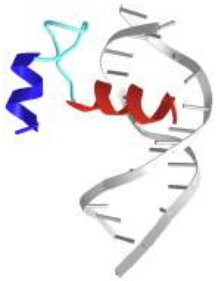


Figure 27.2.14: Representative figures of the transcription factor binding domains. The figure shows the crystal structures of different types of TF domains (3l1p, 4m9e, 5d5v, 1lbg, 1gt0, and 1nkp). The structures were obtained from the Protein Data Bank (PDB) and redrawn using chimera. The respective domains and important regions have been labeled. HTH stands for the helix-turn-helix domain. bHLH stands for the basic helix-loop-helix motif. HD and HMG stand for homeodomain and high-mobility group box domain, respectively. Yesudhas, D., et.al. (2017) Genes 8(8):192

Figure 27.2.14 shows interactive iCn3D models of the transcription factor binding domains as depicted in the figure above. (Copyright; author via source)

Human Hsf1 with Satellite III repeat DNA - HTH Domain (5d5v)



Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...qkHMwpH7L

POU-HMG-DNA ternary complex - HTM-HMG domain (1gt0)

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...b9X7DEGSK

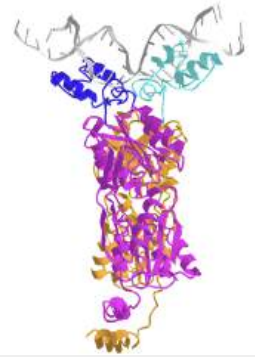
lexHTH-HDdomain(3l1p)

Clickthheimageforapop

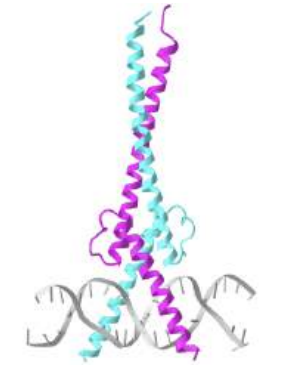
u p o r u s e t h i s e x t e r n a l l i n k : h t t p s : / / s t r u c t u r e .

ncbi.nlm.nih.gov/i...fHmWMvUCiJus77	
Lactose operon repressor and its complexes with DNA and inducer (1lbg) f 4 z i n	Myc-Max and Mad-Max recognizing DNA(1nkp)

cfingerDNA-binding domain in complex with methyl



Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...6v3y24U7N2qYC8



Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...qcTBN4UwxPFH99

ated DNA (4m9e)

Click the image for a pop up pour set

h i s e x t e r n a l l i n k : h t t p s : / / s t r u c t u r e . n c b i . n l m	

.nih.gov/i...T7ZXxMhQ66WNL9	

Most of the search mechanistic studies that try to determine how TFs find their binding sites are limited to naked DNA-protein complexes, which do not reflect the actual crowded environment of a cell. Studies with naked DNA and transcription factors have shown that many DNA-binding proteins travel a long distance by 1D diffusion. However, the search process for eukaryotes must occur in the presence of chromatin, which can hinder protein mobility. In this case, the protein must dissociate from the DNA, enter a 3D mode of diffusion state, and continue the target site searching process. The sliding and intersegmental transfer mechanisms can be explained through the example of the *lac* repressor. The *lac* repressor contains 4 identical monomers (a dimer of dimers) for its DNA binding. The binding sequence of these dimers is symmetric or pseudo-symmetric, and each half is identified by these identical monomers. The HTH domain of the *lac* repressor is the DNA-binding domain that facilitates the interaction with its target site on DNA as shown in Figure 27.2.15

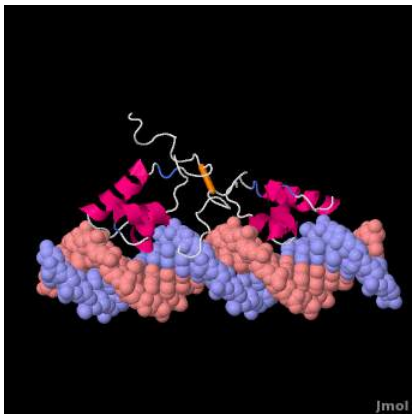


Figure 27.2.15: The Helix-Turn-Helix Motif of the Lac Repressor. Lac repressor binds to DNA non-specifically, enabling it to slide rapidly along the DNA double helix until it encounters the lac operator sequence. The DNA-binding domain employs a helix-turn-helix (HTH) motif (Alpha Helices, Turns). During non-specific binding, the hinge region is disordered. The DNA double helix is depicted as straight in the model when the Lac Repressor binds non-specifically. Upon recognizing the specific operator sequence, the non-specific binding converts to specific binding. During this conversion, the hinge region changes from disordered loops to Alpha Helices, which bind to the minor groove of the DNA. As explained below, this binding stabilizes a kinked ("bent") DNA double helix conformation. *Protopedia – Life in 3D*

As a result of a rapid search (sliding) along the DNA molecule and intersegmental transfer between distant DNA sequences, the lactose repressor finds its target sites faster than the diffusion limit. The section comprised between residues 1–46 of the HTH protein domain, characterized by three α-helices, maintains its secondary structure through specific and non-specific binding. When the repressor binds to a non-specific site, the HTH domain interacts with the DNA backbone and maintains the interaction with its helix region in the major groove juxtaposition. This arrangement facilitates the interaction of the recognition helix with the edges of the DNA bases, enabling the repressor to walk or search for its specific site on the DNA. The C-terminal residues of the DNA-binding domain, residues 47–62, form the hinge region, and are normally disordered during non-specific recognition; however, during specific site recognition, residues 50–58 acquire an α-helix configuration (hinge helix) (Fig. 15 above). The disordered hinge region and the flexibility of the HTH domain allow the protein to move freely along the DNA to search for its target site. In specific binding complexes, the hinge helix of each monomer is located at the symmetrical center of the binding site, thereby causing the hinge helices to interact with each other (intersegmental transfer) to allow better stability. Moreover, DNA bends at the symmetrical center of the specific binding site (37° angle), thereby supporting monomer-monomer interactions.

In addition to the helix-turn-helix structure, the zinc finger motif is also very common, especially in eukaryotic TFs, as shown in Figure 27.2.16 Proteins that contain zinc fingers (zinc finger proteins) are classified into several different structural families. Unlike many other clearly defined supersecondary structures such as Greek keys or β hairpins, there are a number of types of zinc fingers, each with a unique three-dimensional architecture. A particular zinc finger protein's class is determined by this three-dimensional structure, but it can also be recognized based on the primary structure of the protein or the identity of the ligands coordinating the zinc ion. Despite the large variety of these proteins, however, the vast majority typically function as interaction modules that bind DNA, RNA, proteins, or other small, useful molecules. Variations in structures serve primarily to alter the binding specificity of a particular protein. The most common type of zinc finger motif utilizes two Cys and two His residues (CCHH) coordinating the Zn(II) ion to adopt a ββα fold with three hydrophobic residues responsible for the formation of a small hydrophobic core which offers additional stabilization of the zinc finger domain.

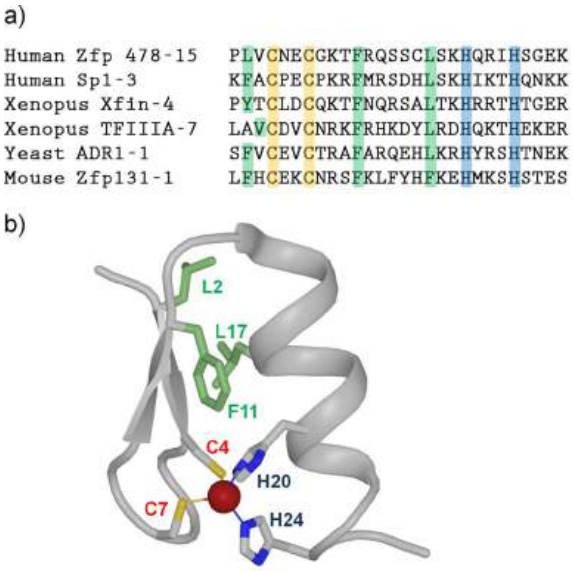


Figure 27.2.16: **Sequence alignments of the CCHH zinc fingers and a representative structure.**Kluska, K., Adamczyk, J., and Krezel, A. (2018) Coord Chem Rev 367:18-64

Panel (a) shows the alignment of the TFIIIA-like zinc finger domains from different organisms. The green color denotes residues that are responsible for the hydrophobic core formation in most CCHH zinc fingers (L17, F11, and L2). Yellow and blue indicate the coordinating Cys and His residues, respectively.

Panel (b) shows the 3D NMR structure of 15-th ZF from zinc finger protein 478 [PDB: 2YRH]

Figure 27.2.17 shows an interactive iCn3D model of C2H2-type zinc finger domain (699-729) from zinc finger protein 473 (2YRH).

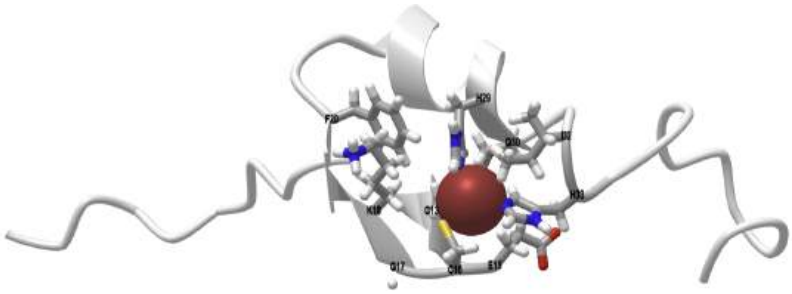


Figure 27.2.17: C2H2-type zinc finger domain (699-729) from zinc finger protein 473 (2YRH). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...M5jGWvYU9NHP78

27.2.4.0.1:

Overall, zinc finger motifs display considerable versatility in binding modes, even between members of the same class (e.g., some bind DNA, others protein), suggesting that they are stable scaffolds that have evolved specialized functions. For example, zinc finger-containing proteins function in gene transcription, translation, mRNA trafficking, cytoskeleton organization, epithelial development, cell adhesion, protein folding, chromatin remodeling, and zinc sensing, to name but a few. Zinc-binding motifs are stable structures, and they rarely undergo conformational changes upon binding their target.

The last binding domain that we will consider in detail here is the helix-loop-helix domain found in Leucine zipper-containing proteins. Specifically, bZIPs (Basic-region leucine zippers) are a class of eukaryotic transcription factors. The bZIP domain is 60 to 80 amino acids in length with a highly conserved DNA binding basic region and a more diversified leucine zipper dimerization region. The two regions form α-helical structures that are connected via a looped region. This forms a core helix-loop-helix (HLH) structure within each monomer of the protein. Two monomers then join through the formation of a leucine zipper junction forming a heterodimeric protein structure. The resulting heterodimer can bind with DNA in a sequence-specific manner through the basic α-helices as shown in Figure 27.2.18

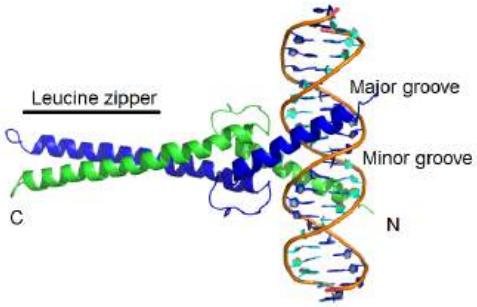


Figure 27.2.18: Leucine Zipper Transcription Factors from the bZIP family. The monomer subunits of a heterodimeric bZIP protein contain a Helix-loop-Helix (HLH) core structure, where one helix forms the leucine zipper with the other monomer, and the basic helices of each monomer interact with the major groove of the target DNA. The helices are held together by a flexible loop region. (One monomer is shown in blue and one monomer is shown in green). Latacca

Specifically, basic residues, such as lysines and arginines, interact in the major groove of the DNA, forming sequence-specific interactions). Most bZIP proteins show a high binding affinity for the ACGT motifs. The bZIP heterodimers exist in a variety of eukaryotes and are more common in organisms with higher evolution complexity.

Figure 27.2.19shows an interactive iCn3D model of the GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted alpha helices (1YSA).

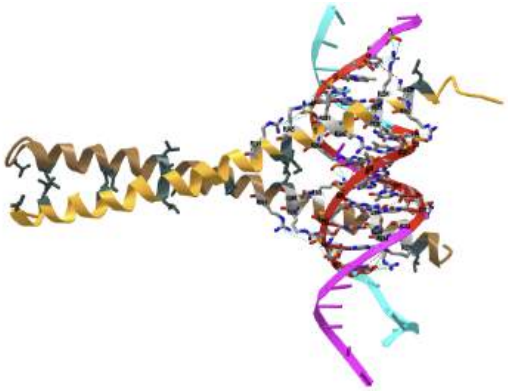


Figure 27.2.19: GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted alpha helices (1YSA). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...hTPDCEQJeLm4z6

27.2.5:

27.2.6: Epigenetics and Transgenerational Inheritance

Even though all somatic cells of a multicellular organism have the same genome, different cell types have different transcriptomes (sets of all expressed RNA molecules), different proteomes (sets of all proteins), and, hence, different functions. Cell differentiation during embryonic development requires the activation and repression of specific sets of genes by the action of cell lineage-defining transcription factors. Within a cell lineage, gene activity states are often maintained over several rounds of cell divisions (a phenomenon called “cellular memory” or “cellular inheritance”). Since the rediscovery of epigenetics some 30 years ago (it was originally proposed by Conrad Hal Waddington in the early 1940s), cellular inheritance has been attributed to gene regulatory feedback loops, chromatin modifications (DNA methylation and histone modifications) as well as long-lived non-coding RNA molecules, which collectively are called the *“epigenome”*. Among the different chromatin modifications, DNA methylation and polycomb-mediated silencing are probably the most stable ones and endow genomes with the ability to impose silencing of transcription of specific sequences even in the presence of all of the factors required for their expression.

27.2.7: Defining Transgenerational Epigenetic Inheritance

The metastability of the epigenome explains why development is both plastic and canalized, as originally proposed by Waddington. Although epigenetics deal only with the cellular inheritance of chromatin and gene expression states, it has been proposed that epigenetic features could also be transmitted through the germline and persist in subsequent generations. The widespread interest in ***"transgenerational epigenetic inheritance"*** is nourished by the hope that epigenetic mechanisms might provide a basis for the inheritance of acquired traits. Yes, Lamarck has never been dead and every so often raises his head, this time with the help of epigenetics.

Although acquired traits concerning body or brain functions can be written down in the epigenome of a cell, they cannot easily be transmitted from one generation to the next. For this to occur, these epigenetic changes would have to manifest in the germ cells as well, which in mammals are separated from somatic cells by the so-called Weismann barrier. Further, the chromatin is extensively reshaped during germ cell differentiation as well as during the development of totipotent cells after fertilization, even though some loci appear to escape epigenetic reprogramming in the germline. Long-lived RNA molecules appear to be less affected by these barriers and therefore more likely to carry epigenetic information across generations, although the mechanisms are largely unsolved.

27.2.8: Evidence for Transgenerational Epigenetic Inheritance

In the past 10 years, numerous reports on transgenerational responses to environmental or metabolic factors in mice and rats have been published. The factors include endocrine disruptors, high-fat diet, obesity, diabetes, undernourishment as well as trauma. These studies investigated DNA methylation, sperm RNA, or both. For example, when male mice are made prediabetic by treatment with streptozotocin it affects the DNA methylation patterns in their resulting sperm, as well as the pancreatic islets of F1 and F2 of the resulting offspring. Furthermore, studies have shown that traumatic stress in early life altered behavioral and metabolic processes in the progeny and that injection of sperm RNAs from traumatized males into fertilized wild-type oocytes reproduced the alterations in the resulting offspring.

In humans, epidemiological studies have linked food supply in the grandparental generation to health outcomes in the grandchildren. An indirect study based on DNA methylation and polymorphism analyses has suggested that sporadic imprinting defects in Prader–Willi syndrome are due to the inheritance of a grandmaternal methylation imprint through the male germline. Because of the uniqueness of these human cohorts, these findings still await independent replication. Most cases of segregation of abnormal DNA methylation patterns in families with rare diseases, however, turned out to be caused by an underlying genetic variant. Thus, studies of this nature must rule out the effects of traditional genetic inheritance as being a factor of the observed phenotypes.

Genetic inheritance alone cannot fully explain why we resemble our parents. In addition to genes, we inherited from our parents the environment and culture, which in parts have been constructed by the previous generations as shown in Figure 27.2.20 A specific form of the environment is our mother's womb, to which we were exposed during the first 9 months of our life. The maternal environment can have long-lasting effects on our health. In the Dutch hunger winter, for example, severe undernourishment affected pregnant women, their unborn offspring, and the offspring's fetal germ cells. The increased incidence of cardiovascular and metabolic disease observed in F1 adults is not due to the transmission of epigenetic information through the maternal germline, but a direct consequence of the exposure in utero, a phenomenon called "fetal programming" or—if fetal germ cells and F2 offspring are affected—"intergenerational inheritance".

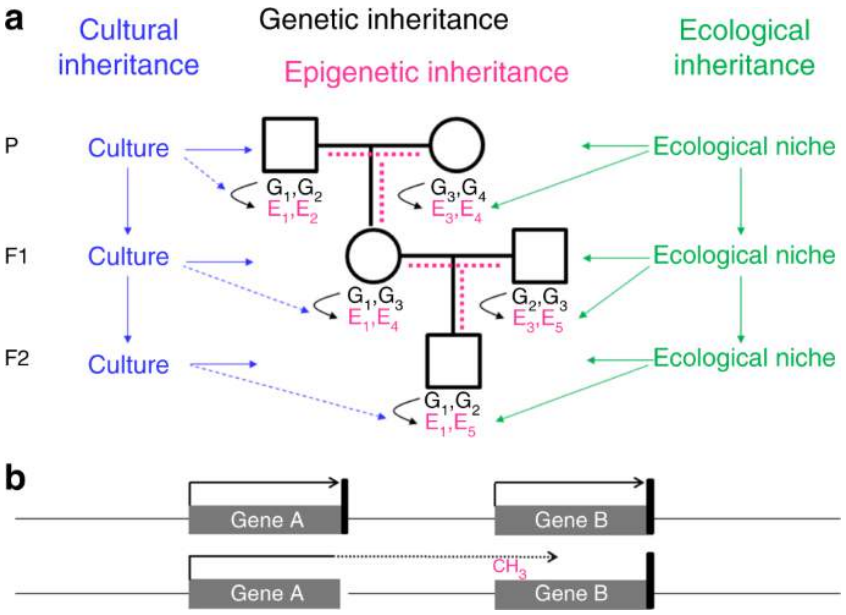


Figure 27.2.20: Transgenerational inheritance systems. *Horsthemke, B. (2018) Nat Comm. 9:2973*

Panel a shows that offspring inherit from their parent's genes (black), the environment (green), and culture (blue). Genes and the environment affect the epigenome (magenta) and the phenotype. Culture also affects the phenotype, but at present, there is no evidence of a direct effect of culture on the epigenome (broken blue lines). It is a matter of debate, how much epigenetic information is inherited through the germline (broken magenta lines). G genetic variant, E epigenetic variant.

Panel b shows that an epimutation (promoter methylation and silencing of gene B in this example) often results from aberrant read-through transcription from a mutant neighboring gene, either in sense orientation as shown here or in antisense orientation. The presence of such a secondary epimutation in several generations of a family mimics transgenerational epigenetic inheritance, although it represents genetic inheritance. Black arrow, transcription; black vertical bar, transcription termination signal; broken arrow, read-through transcription

27.2.9: Roadmap to Proving Transgenerational Epigenetic Inheritance

Here are some steps to show that inheritance is determined by epigenetics and not classical genetics.

- Rule out genetic, ecological, and cultural inheritance. For studies in mice and rats, inbred strains and strictly controlled environments need to be used. When a pregnant female animal is exposed to a specific environmental stimulus, F3 offspring and subsequent generations must be studied to exclude a direct effect of the stimulus on the embryos' somatic cells and germ cells. Even more desirable is the use of *in vitro* fertilization (IVF), embryo transfer, and foster mothers. When a male animal is exposed to an environmental stimulus, F2 offspring must be studied to exclude transient effects on germ cells. To ensure that any phenotype is exclusively transmitted via gametes, IVF must be used, controlling for possible artifacts relating to IVF. In contrast with laboratory animals, it is impossible to rule out ecological and cultural inheritance in humans, but genetic effects should and can be excluded. If an epimutation follows Mendelian inheritance patterns, be cautious: you are more likely looking at a secondary epimutation and genetic inheritance. Study the haplotype background of the epimutation: if in a given family it is always on the same haplotype, you are again most likely dealing with a secondary epimutation. Do whole genome sequencing to search for a genetic variant that might have caused the epimutation and be aware that this variant might be distantly located. Good spots to start looking at are the two neighboring genes, where a mutation might cause transcriptional read-through in sense or antisense orientation into the locus under investigation. Unfortunately, if you don't find anything, you still cannot be 100% sure that a genetic variant does not exist.
- Identify the responsible epigenetic factor in the germ cells. Admittedly, this is easier said than done, especially in female germ cells, which are scarce or unavailable. Be aware that germ cell preparations may be contaminated with somatic cells or somatic DNA. Use swim-up (sperm) or micromanipulation techniques to purify germ cells to the highest purity. Exclude the presence of somatic cells and somatic DNA by molecular testing, for example by methylation analysis of imprinted genes, which are fully methylated or fully unmethylated only in germ cells.

- Demonstrate that the epigenetic factor in the germ cells is responsible for the phenotypic effect in the next generation. If possible, remove the factor from the affected germ cells and demonstrate that the effect is lost. Add the factor to control germ cells and demonstrate that the effect is gained. While RNA molecules can and have been extracted from the sperm of exposed animals and injected into control zygotes, DNA methylation, and histone modifications cannot easily be manipulated (although CRISPR/Cas9-based epigenome editors are being developed and used for this purpose), and all of these experiments can hardly be done in humans. In light of these problems, this might currently be too much to ask for to prove transgenerational epigenetic inheritance in humans, but should, nevertheless, be kept in mind and discussed.

27.2.10: References

1. Parker, N., Schneegurt, M., Thi Tu, A-H., Lister, P., Forster, B.M. (2019) Microbiology. *Openstax*. Available at: https://opentextbc.ca/microbiologyopenstax/
2. Chan, K-G., Liu, Y-C., and Chang C-Y. (2015) Inhibiting N-acyl-homoserine lactone synthesis and quenching *Pseudomonas* quinolone quorum sensing to attenuate virulence. *Front. Microbiol*. 6:1173. Available at: https://www.frontiersin.org/articles/10.3389/fmicb.2015.01173/full
3. Rukavina, Z., and Vanic Zeljka. (2016) Current trends in the development of liposomes targeting bacterial biofilms. *Pharmaceutics* 8(2):18. Available at: https://www.mdpi.com/1999-4923/8/2/18/htm
4. Wikipedia contributors. (2020, April 18). Guanosine pentaphosphate. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:26, August 23, 2020, from https://en.Wikipedia.org/w/index.php?title=Guanosine_pentaphosphate&oldid=951778776
5. Verbeke, F., De Craemer, S., Debunne, N., Janssens, Y., Wynendaele, E., Van de Wiele, C., and De Spiegeleer. B. (2017) Peptides as quorum sensing molecules: measurement techniques and obtained levels *in vitro* and *in vivo. Frontiers in Neuroscience* 11:183. Available at: https://www.researchgate.net/publication/316055402_Peptides_as_Quorum_Sensing_Molecules_Measurement_Techniques_and_Obtained_Levels_In_vitro_and_In_vivo
6. Yesudhas, D., Batool, M., Anwar, M.A., Panneerselvam, S., and Choi, S. (2017) Proteins recognizing DNA: Structural uniqueness and versatility of DNA-binding domains in Stem Cell Transcription Factors. *Genes* 8(8):192. Available at: https://www.mdpi.com/2073-4425/8/8/192/htm
7. Castellanos, M., Mothi, N., and Muñoz, V. (2020) Eukaryotic transcription factors can track and control their target genes using DNA antennas. *Nature Comm.* 11:540. Available at: https://www.nature.com/articles/s41467-019-14217-8
8. Neideracher, G., Klopf, E., and Schüller, C. (2011) Interplay of dynamic transcription and chromatin remodeling: Lessons from yeast. *Int J Mol Sci* 12(8):4758-4769. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3179130/
9. Kim, S., and Kaang, B-K. (2017) Epigenetic regulation and chromatin remodeling in learning and memory. *Exp. & Mol. Med.* 49:e281. Available at: https://www.nature.com/articles/emm2016140#Fig1
10. Tvardovskly, A., Schwämmle, V., Kempf, S., Rogowska-Wrzesinka, A., and Jensen, O.N. (2016) Accumulation of histone variant H3.3 with age is associated with profound changes in the histone methylation landscape. *Nuc. Acids Res.* 45(16):1093. Available at: https://www.researchgate.net/publication/318987684_Accumulation_of_histone_variant_H33_with_age_is_associated_with_profound_changes_in_the_histone_methylation_landscape
11. Cipolletti, M., Fernandez, V.S., Montalesi, E., Marino, M., Fiochetti, M. (2018) Beyond the antioxidant activity of dietary polyphenols in cancer: The modulation of estrogen receptors (ERs) signaling. *Int J. Mol Sci* 19(9)2624. Available at: https://www.mdpi.com/1422-0067/19/9/2624/htm
12. Griekspoor, A., Zward, W., Neefjes, J., and Michalides, R. (2007) Visualizing the action of steroid hormone receptors in living cells. *Nucl. Recept. Signal*. 5:e003 Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1853070/
13. Mitsis, T., Papargeorgiou, L., Efthimiadou, A., Bacopoulou, F., Vlachakis, D., Chrousos, G.P., Eliopoulos, E. (2020) A comprehensive structural and functional analysis of the ligand binding domain of the nuclear receptor superfamily reveals highly conserved signaling motifs and two distinct canonical forms through evolution. *World Acad Sci J* 1: 264-274, 2019. Available at: https://www.spandidos-publications.com/10.3892/wasj.2020.30
14. Reed, S.M., and Quelle, D.E. (2015) p53 Acetylation: Regulation and consequences. *Cancers* 7(1):30-69. Available at: https://www.mdpi.com/2072-6694/7/1/30/htm.
15. Maclaine, N.J., and Hupp, T.R. (2009) The regulation of p53 by phosphorylation: a model for how distinct signals integrate into the p53 pathway. *Aging* 1(5):490-502. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2806026/
16. Wikipedia contributors. (2020, August 1). Estrogen. In *Wikipedia, The Free Encyclopedia*. Retrieved 01:28, September 6, 2020, from https://en.Wikipedia.org/w/index.php?title=Estrogen&oldid=970560042

17. Kluska, K., Adamczyk, J., and Krezel, A. (2018) Metal binding properties, stability, and reactivity of zinc fingers. *Coord. Chem Rev*. 367:18-64. Available at: https://www.sciencedirect.com/science/article/pii/S0010854517305441
18. Wikipedia contributors. (2020, July 4). Leucine zipper. In *Wikipedia, The Free Encyclopedia*. Retrieved 07:00, September 7, 2020, from https://en.Wikipedia.org/w/index.php?title=Leucine_zipper&oldid=965962667
19. Wikipedia contributors. (2020, April 15). Zinc finger. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:28, September 7, 2020, from https://en.Wikipedia.org/w/index.php?title=Zinc_finger&oldid=951116840
20. Horsthemke, B. (2018) A critical view on transgenrational epigenetic inheritence in humans. *Nat. Comm.* 9:2973. Available at: https://www.nature.com/articles/s41467-018-05445-5#rightslink

---

Binding and Noncovalent Interaction Problems

Chapter 1 - Problems

CHAPTER OVERVIEW

Enzyme Kinetics Problems

Topic hierarchy

Carbonic Anydrase Inhibition

WNovak-Beta Jupyter NB - Ex 1: Under Construction

Carbonic Anydrase Inhibition

These questions derive from the Research Literature Module - Carbon Capture Using Carbonic Anhydrase

Question 1

Using the equation below, at what ratio of CO_2/[HCO_3^-] would the rate for the forward reaction (CO_2 sequestration) be cut in half?

$$v_0 = \frac{V_M S}{K_M \left(1 + \frac{I}{Kis}\right) + S} \tag{1}$$

If you need some help, hover over -

Give me a hint!

Answer

$$v_{-I} = \frac{V_M S}{K_M + S} = \frac{(1)(1)}{(1+1)} = 0.5$$

$$v_{+I} = \frac{V_M S}{K_M \left(1 + \frac{[I]}{K_{is}}\right) + S} = \frac{(1)(1)}{1\left(1 + \frac{[I]}{2}\right)} = 0.25 = \frac{1}{1 + \frac{[I]}{2}}$$

$$0.25\left(1 + \frac{[I]}{2}\right) = 1 \tag{2}$$

$$1 + \frac{[I]}{2} = 4$$

$$\frac{[I]}{2} = 3$$

$$[I] = 6$$

Hence it doesn't take much HCO_3^- buildup to inhibit the "capture" of CO_2!

WNovak-Beta Jupyter NB - Ex 1: Under Construction

CHAPTER OVERVIEW

Enzyme Mechanisms Questions

Topic hierarchy

Carbonic Anhydrase - Mechanism

Carbonic Anhydrase - Mechanism

These questions derive from the Research Literature Module - Carbon Capture Using Carbonic Anhydrase

Carbonic Anhydrase - Structure and Mechanism 1

An active site Zn^{2+} appears to bind a water molecule and reduce its pK_a such that the bound form is OH^-. This is illustrated in the left panel of Figure 1 below, which depicts the local environment of the bound Zn^{2+} (coordinated by histidine side chains and an OH^-) in the absence (left) and presence (right) of CO_2. Note that the back histidine is difficult to barely visible (but still evident) in both structures. To assist in viewing the structure, the right panel shows an interactive iCn3D model of Zn- human carbonic anhydrase II at pH 7.8 and 0 atm CO (6LUW).

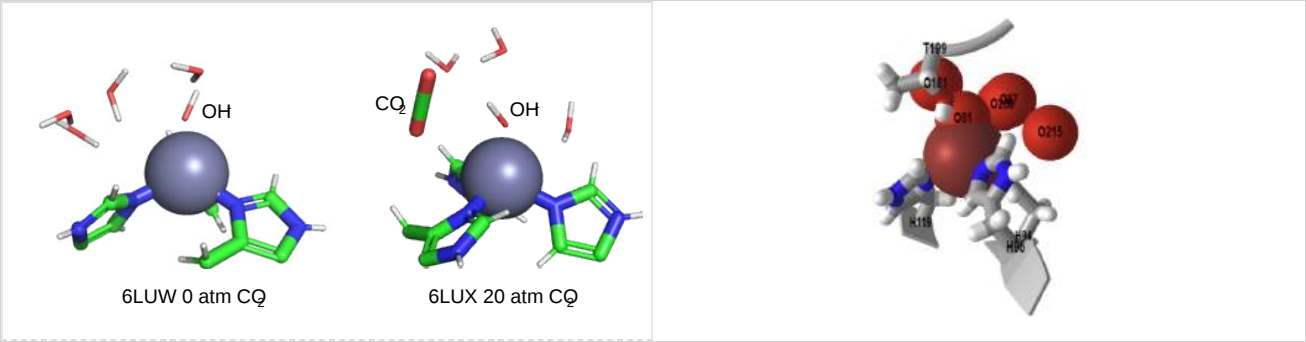


Figure 1: Left panel: Coordination of OH^- to Zn^{2+} in carbonic anhydrase in the absence (left) and presence (right) of substrate CO_2. Right panel: Zn- human carbonic anhydrase II at pH 7.8 and 0 atm CO (6LUW) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...43DYyZFvHJpZn9

? Question \PageIndex1}

What is the coordination geometry of the Zn ion?

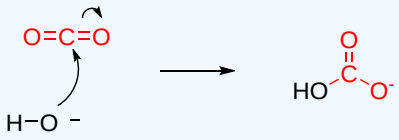
Answer

tetrahedral

? Question 2

Draw a simplified reaction mechanism showing bicarbonate formation from the two reactants, CO_2 and OH^-.

Answer



The enzyme is reversible and in humans is important in CO_2 transport in respiration and maintaining intracellular pH, which is also its key role in most organisms.

CO_2, like the other atmospheric gases O_2 and N_2, are nonpolar and have limited solubility in water. The solubility of these gases in water at 20°C and 1 atm pressure, in g/L and mM, are shown in Table 1 below.

Gas (ref)	solubility (aq) (g/L)	solubility (mM)
CO_2	1.7	38
O_2	0.044	1.3

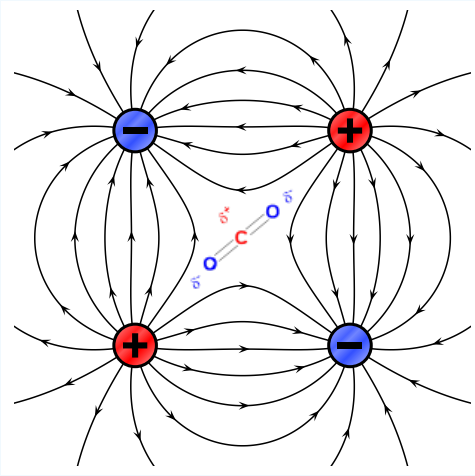
N_2	0.019	0.68

Question 3

Offer reasons that explain the significantly higher (but still low) solubility of CO_2 in water compared to O_2 and N_2.

Answer

CO_2 is considered a nonpolar molecule since it has no net molecular dipole. However, it does have 2 bond dipoles (pointing in opposite directions), so the carbon atom is δ^+ while the Os are δ^-. This probably contributes to its greater solubility than N_2 and O_2 which don't have bond dipoles. CO_2 would not orient itself in a dipole electric field, but it would to some extent in a quadrupole (4 poles) electric field where the positive potentials are oriented north and south and the negative potentials at east and west. CO_2 has a quadrupole moment. In addition, the continued reaction of CO2 and water to form the weak acid carbonic acid would contribute to its higher apparent solubility. It is not clear to the authors if these contributions to solubility are accounted for in the experimental values of solubility.



A quadruple and its associated magnetic field with oriented CO_2. https://commons.wikimedia.org/wiki/F...quadrupole.svg

Question 4

How does the enzyme facilitate the transport of CO_2 in blood? How does it maintain intracellular pH?

Answer

It converts the poorly soluble carbon unit in CO_2 to the strongly soluble bicarbonate anion. HCO_3^-.

A simple explanation for maintaining intracellular pH comes from the chemical equation below.

$CO_2(aq) + H_2O\ (l) \leftrightarrow H_2CO_3\ (aq) + H_2O(l) \leftrightarrow H_3O^+(aq) + HCO_3^-(aq)$.

HCO_3^- is the conjugate base of the weak acid, H_2CO_3 so the system is a classic buffer. For a complete explanation of why the system can act as a buffer at neutral pH even though the pKa of the weak acid is 3.6, see Chapter 2.3 for review.

Let's explore the structure of two different CAs, human carbonic anhydrase II and the carbonic anhydrase from Neisseria gonorrhea.

Human carbonic anhydrase II

The structure of native human carbonic anhydrase II and its catalytic mechanism is shown in Figure 5 below.

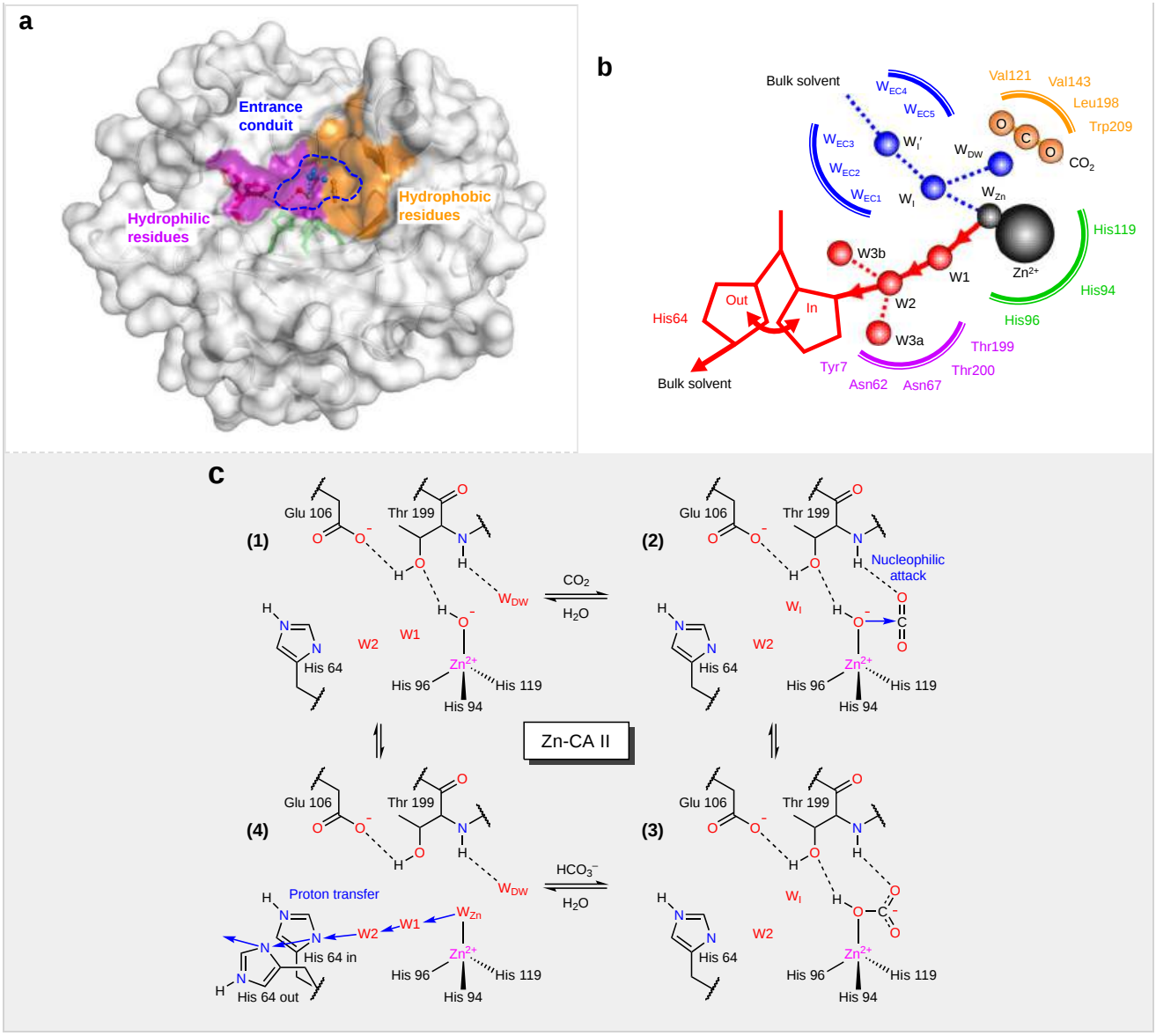


Figure 5: **Structure of native human carbonic anhydrase II (Zn-CA II) and its catalytic mechanism.** Kim, J.K., Lee, C., Lim, S.W. *et al.Nat Commun* **11**, 4557 (2020). https://doi.org/10.1038/s41467-020-18425-5. Creative Commons Attribution 4.0 International License, http://creativecommons.org/licenses/by/4.0/.

Panel a shows the active site consists of the zinc binding site, hydrophobic/hydrophilic regions, and the entrance conduit (EC).

Panel b shows the water networks in the active site that are responsible for the proton transfer (red) and substrate/product/water exchange (blue) during enzyme catalysis.

Panel c shows the forward reaction mechanism of Zn-CA II.

The active site itself lies at the bottom of a deep cavity (15 Å deep) in the protein, which is readily accessible to solvent

An interactive iCn3D model of human carbonic anhydrase II with bound bicarbonate and CO_2 (2VVB) is shown in Figure 6 below.

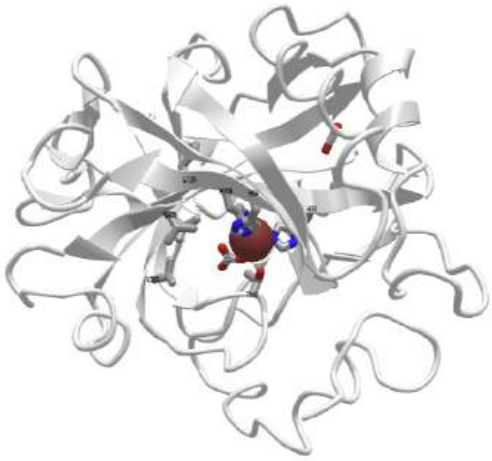


Figure 6: Human carbonic anhydrase II with bound bicarbonate and CO_2 (2VVB) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...ibgmgsm3UWhjX6

The active site residues shown in Figure 6 are labeled and shown as sticks. Bound CO_2 and HCO_3^- are also shown as sticks.

Question 5

This question addresses the Biomolecular Visualization Framework theme(s) Molecular Interactions (MI), Atomic Geometry (AG)

Using iCn3D to show the noncovalent interactions between bicarbonate, Zn, and the protein by using iCn3D. Measure the distance between the hydrogen-bonded atoms in HCO_3^- and Thr 199.

iCn3D instructions

Trackpad and Mouse Controls

•rotate: click and drag (mouse: left click and drag)

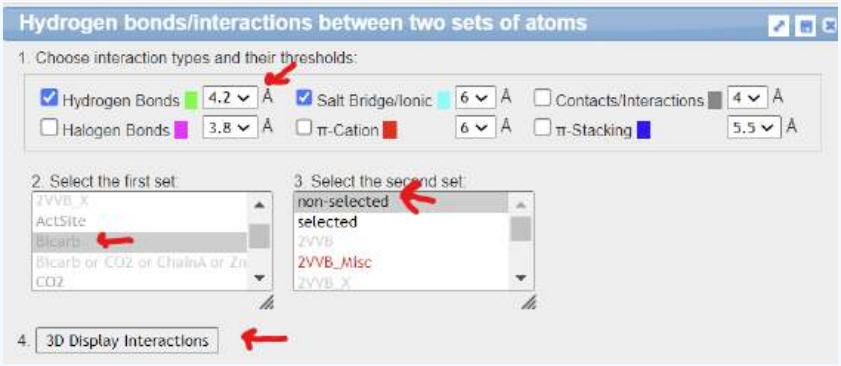
•zoom: pinch and spread (mouse: rotate the scroll wheel)

•translate: two-finger click and drag (mouse: right click and drag)

•Re-center: left click View from the top menu bar, then select “Center Selection”

•Note: ctrl-click on a PC = command-click on Mac; alt-click on PC = option click on Mac

1. Open the external link: https://structure.ncbi.nlm.nih.gov/i...ibgmgsm3UWhjX6
2. From the top menu bar, choose **Analysis**, **Interactions**
3. In the new popup window select the following prompts, then click **4. 3D Display interactions**



4. Close all but the main modeling window.

5. Zoom into the bicarbonate binding site. From the top menu bar, choose **Analysis**, **Distance**, **distance between 2 atoms**, and pick the two atoms (by holding down the Alt key or Option on a Mac) involved in the hydrogen bond between the bicarbonate and the amide

6. Rescale the label size by choosing **Analysis, Label Scale, 0.4**.

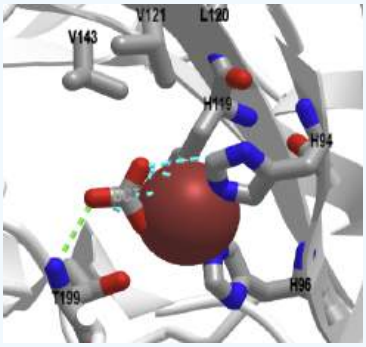
7. From the top menu bar choose **Select, Toggle highlights** to remove yellow and box highligting.

8. Save a PNG file by choosing

9. You can reload the PNG file directly into iCn3D by choosing **File**, **Open File**, i**Cn3D PNG image**

Answer

The blue dotted lines are ion-ion interactions. Note the green hydrogen bond between T199 and the bound bicarbonate. Remember that hydrogen atoms are not shown in PDB files from x-ray structures. The H-bond distance is 3.1 angstroms. For some reason, the H-bond does not show unless the initial constraint distances are moved to 4.2 Angstroms.



? Question 6

What is a likely function of Val, Leu and Trp cluster in CAII (shown in Figures 5 and 6)?

Answer

These side chains are all hydrophobic (as illustrated in gold in Figure 5) and provide a weak binding environment for nonpolar CO_2.

? Question 7

What ligand would likely replace OH^- at low pH values? What would happen to the activity of the enzyme at lower pHs?

Answer

At low pH, ie. at pH values lower than the pKa of the Zn-bound water (which deprotonates to form the OH^- ligand and nucleophile), the ligand and nucleophile would be H_2O. The enzyme would display a lower activity given the weaker nucleophile.

Question 8

From the mechanism shown in Figure 5, does bicarbonate coordinate the Zn^{2+} ion in a monodentate or bidentate manner?

Answer

Monodentate as only 1 coordinate covalent bond forms on an electron pair donation from the bicarbonate to the Zn^{2+}.

Question 9

Write a verbal description of the mechanism of CAII based on Figure 5

Answer

CO_2 binds to the active site through loose association with the cluster of hydrophobic side chains. The Zn^{2+} bound OH^- acting as a nucleophile attacks the central carbon of the CO_2 forming HCO_3^-. The carbonate forms a monodentate interaction with Zn and also a hydrogen bond to Thr 199. The HCO_3^- is then displaced by an incoming water molecule. The other product of the reaction, H^+, moves through a hydrogen bond network of water molecules (W1 and W2) to His 64 and eventually to bulk water. The interactions with substrate and products are weak allowing fast exchange.

Question 10

Why is a proton transfer path needed?

Answer

H^+ is a product of the reaction: CO_2 (g) + H_2O ↔ H_2CO_3 (aq) ↔ HCO_3^- (aq) + H^+ (aq). It must depart to prevent charge build-up, maintain charge balance, and keep the correct electrostatic environment of the active site.

Question 11

Thr 199 plays a key role in the mechanism. State a reason for its importance in the reversible reaction.

Answer

Thr 199 supplies two hydrogen bonds to the bicarbonate. It actually destabilizes bicarbonate bonding with respect to a T199A mutant. In the mutant, carbonate might bind Zn^{2+} in a bidentate fashion, leading to tighter binding and a slower dissociation rate of the product, HCO_3^-.

Question 12

The dissociation constant K_D (or K_{is}) for bicarbonate binding to HCAII is about 77 mM. What kind of inhibitor might it be for the forward reaction?

Answer

Given that it binds in the active site and would prevent binding of the substrate for the forward reaction, CO_2, it is a competitive inhibitor.

Question 13

Using the equation below, at what ratio of $CO_2/[HCO_3^-]$ would the rate for the forward reaction (CO_2 sequestration) be cut in half? Assume the V_M=1, K_M forward reaction is 1, [S] = 1 and K_{is} is 2 (a wide range of values are reported in the Brenda Database.

$$v_0 = \frac{V_M S}{K_M \left(1 + \frac{I}{Kis}\right) + S} \tag{1}$$

If you need some help, hover over -

Give me a hint!

Answer

$$v_{-I} = \frac{V_M S}{K_M + S} = \frac{(1)(1)}{(1+1)} = 0.5$$

$$v_{+I} = \frac{V_M S}{K_M \left(1 + \frac{[I]}{K_{is}}\right) + S} = \frac{(1)(1)}{1\left(1 + \frac{[I]}{2}\right)} = 0.25 = \frac{1}{1 + \frac{[I]}{2}}$$

$$0.25\left(1 + \frac{[I]}{2}\right) = 1 \tag{2}$$

$$1 + \frac{[I]}{2} = 4$$

$$\frac{[I]}{2} = 3$$

$$[I] = 6$$

Hence it doesn't take much HCO_3^- buildup to inhibit the "capture" of CO_2!

Question 14

The rate-limiting step for human CA II is the dissociation of a proton from Zn^{2+}-bound water and not the removal of the resulting proton from the enzyme. What does that imply about the rate of removal of the proton from the enzyme?

Answer

It must be very fast, that is at diffusion-controlled limits through the H-bond channel.

Synthetic mimetics of the active site of CA have been made. These are heteromacrocycles (similar to the heme of hemoglobin) as shown in Figure 9 below.

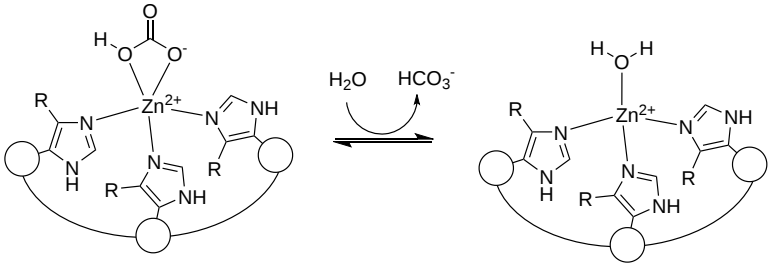


Figure 9:

The macrocycle mimetic has three imidazole groups coordinating zinc.

Question 15

The macrocycle mimetic has three imidazole groups coordinating zinc. Is the bicarbonate coordinated to the Zn^{2+} in a monodentate or bidentate fashion? From the "denticities" of the interactions of bicarbonate and Zn^{2+} for human CAII and the mimetic, which catalyst, CAII or the macrocycle would you expect to have a lower K_D for bicarbonate? How might this affect the rate-limiting step for the mimetic?

Answer

The mimetic is bidentate, is it should bind more tightly to bicarbonate, hindering its dissociation, and hence making "product" inhibition more likely.

Question 16

What is the utility of having both CO_2 and HCO_3^- bind weakly to the enzyme

Answer

Weak binding "permits their rapid exchange. The hydrogen-bonding arrangement in the active site is such that the water or hydroxide ion donates a hydrogen bond to a proximal threonine (Thr199 in hCA II) because the hydroxyl group of this residue is forced to donate its hydrogen in a hydrogen bond to a negatively charged glutamate side chain (Glu106 in hCA II). Site-directed mutations have confirmed this model of the catalytic mechanism. The substrates/products carbon dioxide/bicarbonate are fairly weakly bound against a hydrophobic wall in the active site, which permits their rapid exchange."

CHAPTER OVERVIEW

Global Challenges with Literature-Based Guided Assessments

Topic hierarchy

1. Global Challenges: An Introduction
Carbon Capture Using Carbonic Anhydrase
Trauma and Health

1. Global Challenges: An Introduction

Global Challenges		Topic

Introduction

Instead of presenting a litany of end-of-chapter or end-of-book questions that are not linked in content or concepts, we will present a number of problem-solving assessments linked to research literature that deal with key challenges that face the world today. When the questions are mostly derived from the literature, we'll also call them **Literature-Based Guided Assessments (LGAs)**. Each will focus on a particular biological system (enzyme, pathways, etc) and contain a series of sequential and linked questions on a particular protein, for example, and its function. The problems are summative and hence require an understanding of structure, noncovalent interactions, binding, kinetics, and reaction mechanisms. The modules are guided and have elements of problem-solving and POGIL questions.

Each Module will ultimately focus on the structure and properties of key biomolecules. The modules will:

1. Address ASBMB Core concepts and Learning Objectives (generalized below) through analysis and interpretation of research findings in one or more research publications.

- Energy is required by and transformed in biological systems.
- Macromolecular structure determines function and regulation.
- Information storage and flow are dynamic and interactive.
- Biochemical systems maintain a state of homeostasis, a steady stable state while continually adjusting to conditions, *which requires energy input, organization, and control mechanisms.*
- Evolution plays a pervasive role in shaping the form and function of all biological molecules and organisms.

2. Link to critical problems facing the world (see the Table) which have clearly identified biochemical components. These critical problems include health care disparities, climate change, pandemics, addiction, childhood trauma, food insecurity, biodiversity loss, ecosystem (ocean, soil, forest) health, and misinformation/disinformation. These problems are often linked and not mutually exclusive. Most biochemistry textbooks focus on problems using biomedical examples. Expanding to study key world problems that are not directly biomedical and are underrepresented in textbooks, allows students to apply their acquired knowledge and understanding into different areas.

3. Follow general features found in problem-based learning and in case studies, which provide contextual applications for the detailed learning opportunities found in biochemistry books and courses.

- A broad introduction (text, videos, personal narratives) describing the critical world problem and the relevancy of the selected biochemical system to the problem
- A more detailed description of the selected biochemical system, including links to specific locations in Fundamentals of Biochemistry as well as external resources
- Research literature results (graphs, tables, models, etc), taken from journals that allow derivatives and reuse by appropriate Creative Commons licensing (for example, CC BY 4.0), for interpretation

4. Focus on representative biomacromolecules (protein, nucleic acid, glycan, lipid and combination of them) relevant to the broader problem for which detailed structure/function questions can be explored

5. Explicitly address and link to appropriate BioMolViz framework themes, goals and objectives to the biomolecules key to the LGA.

Relationship of LGAs to BioMolViz and Molecular CaseNet

The completed LGAs will consist of a broad introduction and relevant biochemical research findings woven into a narrative that will include nested questions based on the literature with an ultimate focus on a key biomacromolecule. It will not take the form or

detail of a full case study as found in Molecular CaseNet (headed by Shuchi Dutta and its Steering Committee, which includes Henry Jakubowski, who is also on the Steering Committee of BioMolViz). As the LGAs in Fundamentals of Biochemistry and indeed the whole text, as well as the Molecular CaseNet are free online educational resources (OERs), both communities can freely share resources. Since the LGAs have some attributes of case studies, we hope that contributors to Molecular CaseNet will freely use the LGAs and convert them to more expansive case studies, housed within Molecular CaseNet.

Likewise, the research literature-based questions in the LGAs that focus on biomacromolecule structures will be explicitly linked to the themes, goals and objectives of the BioMolViz literacy framework. However, the specific questions will not be included in the web repository created by BioMolViz. The repository questions have gone through many iterative cycles of construction, revision, external review by expert panels, and validation by actual classroom use. Instead, the questions in Fundamentals of Biochemistry LGAs that target specific biomolecular visualization framework objectives will help to expand knowledge and understanding of biomolecule visual literacy and BioMolViz objectives, which ultimately is the goal of BioMolViz.

A full semester of biochemistry would be necessary to complete a full LGA, as the questions extend from structure, binding, kinetics, mechanism, metabolism, and signal transduction. Yet parts of a complete LGA could be completed after students complete the corresponding chapter in the book. Hence parts of a given LGA will be listed in Volume 5 under the corresponding topic (carbohydrate structure, for example). A link will be provided back to the home LGA from which the questions were derived

Global Challenges as the Bases for Literature-Based Guided Assessments (LGAs)

Here are the world challenges we have selected that we serve as the bases for the LGAs.

World Problems	**Research Literature Modules**
Health Disparities	Type II Diabetes
	Orphan receptors
	Poverty and stress response: Poverty and epigenetics
	PM2.5s
	Pb pollution
	Pollution (air/water)
Climate Change	Thermal tolerance plants
	Carbon Capture
	Photosynthesis, CO2 sequestration
	Heat Stroke
	Biofuels
	Modeling climate change
Pandemics	Vaccine Development
	Ebola
	Malaria
	Emerging Diseases
	Evolution
Addiction	Natural/Synthetic opiates
	Alcohol abuse
Trauma	PTSD
Food Insecurity	photosynthesis
	fertilizers
Loss of Biodiversity	extinction
Ecosystem Health	Soil
	Oceans

World Problems	**Research Literature Modules**
	Forest
Integrity of Information	Western blots and image modifications

Carbon Capture Using Carbonic Anhydrase

Global Challenges	Climate Change Carbon Capture Using Carbonic Anhydrase	Literature-Based Guided Assessment (LGA)
		

Key Words, Concepts: protein structure, structure/function relationships, enzyme kinetics, enzyme mechanisms, reaction mechanisms, Western blot analysis, site-directed mutagenesis, biomolecular visualization, computational modeling, graphic analysis

The Problem

Our climate is changing as planetary temperatures rise from increasing amounts of the greenhouse gas carbon dioxide released into the atmosphere (detailed in Chapter 31) on the burning of fossil fuels. The rate of release is unparalleled in geological history. Present levels (415 ppm) have not been seen for at least 3 million years. Human societies and cultures have had the opportunity to develop in relatively stable climatic conditions. Figure 1 below shows the rise in atmospheric CO_2 over the last 1000 years.

Figure 1: CO_2 levels in the atmosphere over the last 1000 years. Our world in data. https://ourworldindata.org/

The steep rise around 1790 coincides with the start of the Industrial Revolution. The rise in atmospheric CO_2 has led to a corresponding rise in the average global temperatures, as illustrated in Figure 2 below.

Figure 2: Average global temperature changes over the last 1000 years. Our world in data. https://ourworldindata.org/

The per capita emissions of CO_2 across the world derive from the use of coal, oil, and gas, as illustrated in Figure 3 below.

Figure 3: Per capita emission of CO_2 from fossil fuel type

If we want to decrease emissions, we also need to know in which economic sectors fossil fuels are used. The main sources of global energy-related CO_2 emissions by sector are shown in Figure 4:

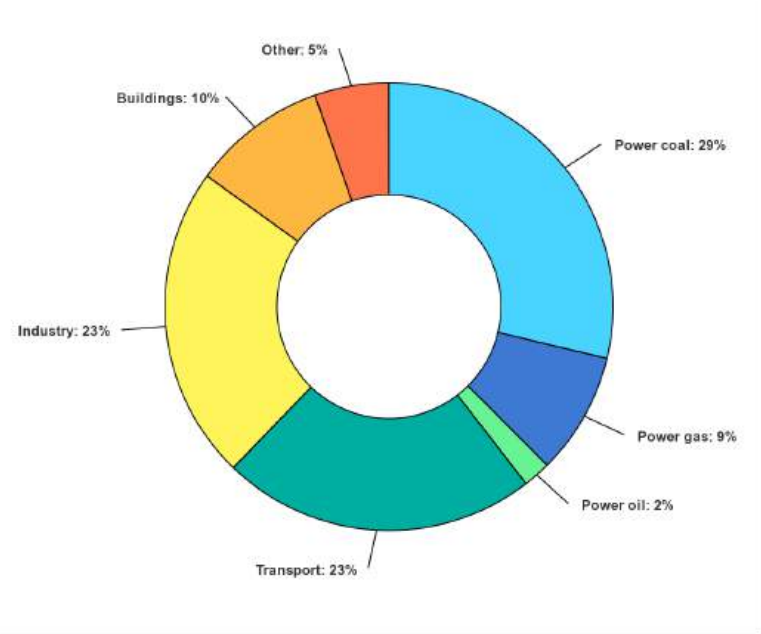


Figure 4: Global energy-related CO_2 emissions by sector. Updated on 10/26/22. https://www.iea.org/data-and-statist...ions-by-sector. IEA. License: CC BY 4.0

Note the use of coal, gas, and oil for energy production (electricity) accounts for 40% of global CO_2 emissions, with transportation (mostly through the use of gasoline and diesel fuel) and industry accounting for about 25% each.

Simple chemistry tells us that there are two ways to decrease the amount of product (in this case CO_2) in a chemical reaction:

- decrease the concentration of reactants (i.e. reduce fossil fuel use)
- remove the product, in this case, CO_2 from the air.

The latter process is called carbon capture or sequestration. It is a daunting process that nature has mastered (through photosynthesis), but it clearly can't keep up with the huge injection of CO_2 in the atmosphere caused by burning fossil fuel.

We simply can't stop using fossil fuels, which would result in huge economic and social unrest. Alternative green fuels (solar, wind, for example) are being rapidly expanded but can't replace fossil fuels for many years. One of the reasons is that fossil fuels are very energy-dense (MJ/kg) compared to other sources of energy. It's also fascinating to look at the energy transitions humans have made over time. Figure 5 below shows the energy transition over a log-time scale (for presentation purposes) as well as the energy densities of individual sources.

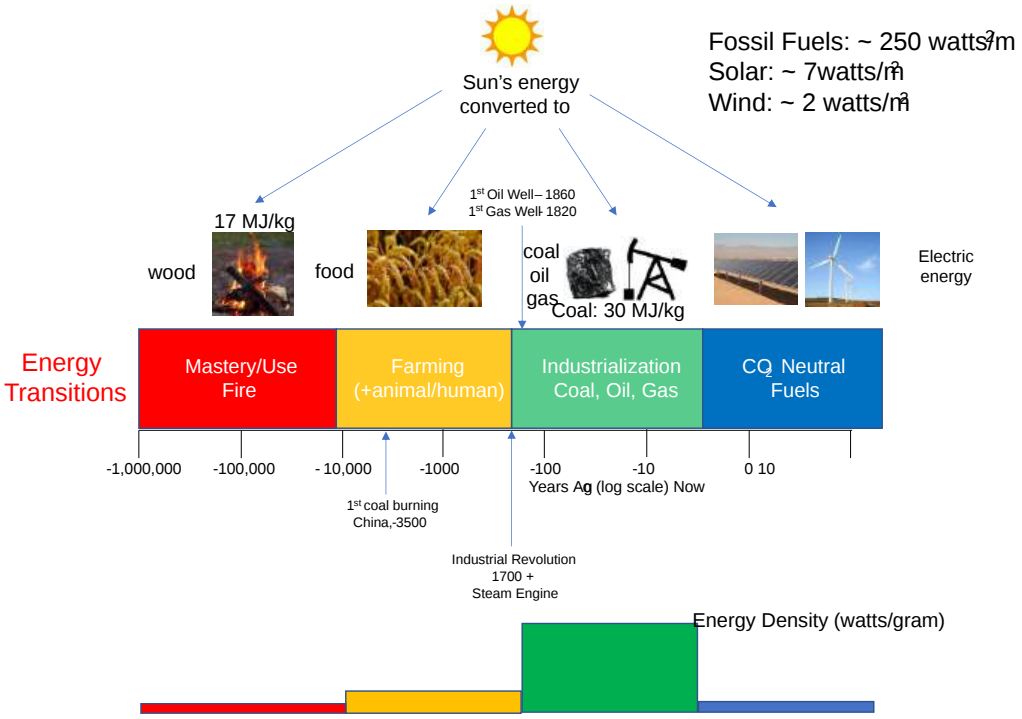


Figure 5: Human-created energy transitions (log time scale) and energy densities of individual sources

New technologies are needed to capture CO_2. We have to move much faster in a new clean energy transition than we have in our entire history. A potentially ideal solution would be to capture CO_2 from power plants before they reach the atmosphere. We will now look at research into an old enzyme, carbonic anhydrase, that is being repurposed for industrial-level carbon capture, carbonic anhydrase.

Carbonic Anhydrase (CA)

We have already encountered this enzyme before (Chapter 6.1). It catalyzes the hydration of CO_2 (g) as shown below.

$$CO_2\,(g) + H_2O \leftrightarrow H_2CO_3\,(aq) \leftrightarrow HCO_3^-\,(aq) + H^+\,(aq)$$

It is among the fastest of all enzymes, with a k_{cat} of 10^6 s^{-1} and a k_{cat}/K_m of 8.3 x 10^7 $M^{-1}s^{-1}$ (reference). It is diffusion controlled in that the rate of diffusion of reactants and products, not the chemical steps, determine the reaction rate. It can convert 10^6 molecules of CO_2(g) to HCO_3^- each second. No wonder scientists and engineers are studying it to capture CO_2. It's a big challenge though to capture CO_2 released on combustion of coal or natural gas in a power plant. Here are two problems that must be overcome:

- The enzyme must be thermostable at elevated temperatures to capture the CO_2 found in high-temperature power plant emissions
- The enzyme is reversible so it will be inhibited by the product HCO_3^-
- The enzyme must be stable to somewhat alkaline conditions (pH of 0.1M $NaHCO_3$ = 8.3)

For carbon capture from fossil fuel emissions, CA is immobilized by surface adsorption, covalent attachment, encapsulation, and entanglement. Immobilized enzymes are typically more thermostable and can be used in flow-through as opposed to solution phase capture. The immobilized enzyme matrix must withstand high temperatures (up to 100°C, and alkaline solvents used to strip the matrix for reuse.

The enzyme is found throughout life and typically has an active site Zn^{2+}. There are 8 families, α, β, γ, δ, ζ η, θ, and ι, with the α family being the most abundant. The α forms are generally active as dimers, but can act as monomers and tetramers.. There are 15 isoforms of the α form in humans and have a prime role in pH regulation. They are found in bacteria, fungi, plants, and algae. β-CAs are found in some types of bacteria, archaea, fungi, some higher plants, and invertebrates. CA in chloroplasts (and mitochondria (algae) are involved in carbon fixation. We will focus our attention on engineering carbonic anhydrase to make them more thermostable, alkali insensitive, and less susceptible to product inhibition by bicarbonate.

Natural enzymes can be isolated and selected for thermal and alkali stability. In addition, new versions selected for these properties can be engineered using directed evolution or site-directed mutagenesis. You wish to increase the thermal stability of a protein using mutagenesis. Essentially you wish to perturb the equilibrium between the folded (native) protein and the unfolded (denatured) protein so as to preferentially stabilized the native state.

? Question 1

Using mutagenesis, what residues might you change in a native protein to make it more stable at higher temperatures?

Answer

? Question 2

What measurements would you make to quantitate the change in thermal stability?

Answer

The actual amino acid composition and more strangely specific dipeptide sequences within a sequence are associated with thermal stability of hyperthermophilic proteins. For example, proteins from two different types of archaea with different optimal growth temperatures show that the one with the higher growth temperature have significantly higher levels of VK, KI, YK, IK, KV, KY, and EV and decreased levels of DA, AD, TD, DD, DT, HD, DH, DR, and DG. Similar experiments have been done in bacterial cells. Using machine learning, the dipeptide sequences K*H, KR, TF, P*M, F**N, V**Y, MW, and WQ were important in themostability where the * denotes a gap in the residues.

Structure and Mechanism

An active site Zn^{2+} appears to bind a water molecule and reduce its pK_a such that the bound form is OH^-. This is illustrated in the left panel of Figure 6 below, which depicts the local environment of the bound Zn^{2+} (coordinated byhistidine side chains and an OH^-) in the absence (left) and presence (right) of CO_2. Note that the back histidine is difficult to barely visible (but still evident) in both structures. To assist in viewing the structure, the right panel shows an interactive iCn3D model of Zn- human carbonic anhydrase II at pH 7.8 and 0 atm CO (6LUW).

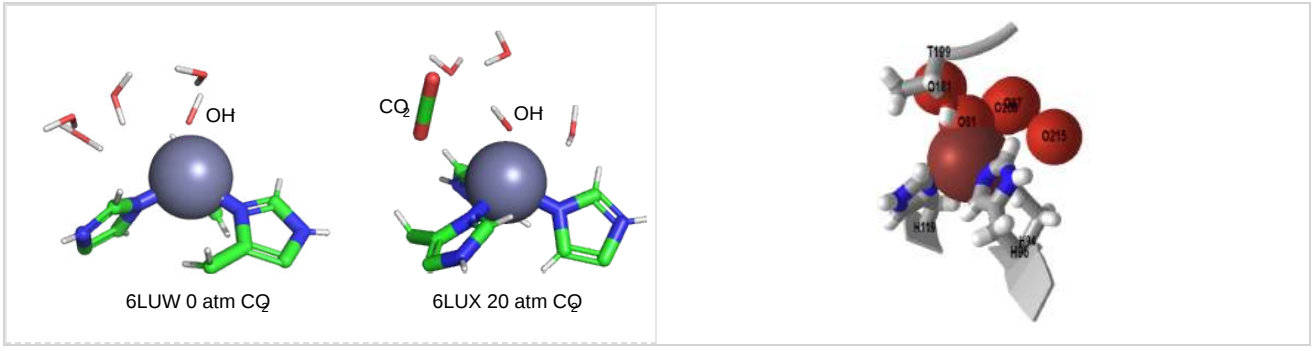


Figure 6: Left panel: Coordination of OH^- to Zn^{2+} in carbonic anhydrase in the absence (left) and presence (right) of substrate CO_2. Right panel: Zn- human carbonic anhydrase II at pH 7.8 and 0 atm CO (6LUW) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...43DYyZFvHJpZn9

? Question 3

What is the coordination geometry of the Zn ion?

Answer

Question 4

Draw a simplified reaction mechanism showing bicarbonate formation from the two reactants, CO_2 and OH^-.

Answer

The enzyme is reversible and in humans is important in CO_2 transport in respiration and maintaining intracellular pH, which is also its key role in most organisms.

CO_2, like the other atmospheric gases O_2 and N_2, are nonpolar and have limited solubility in water. The solubility of these gases in water at 20°C and 1 atm pressure, in g/L and mM, are shown in Table 1 below.

Gas (ref)	solubility (aq) (g/L)	solubility (mM)
CO_2	1.7	38
O_2	0.044	1.3
N_2	0.019	0.68

Question 5

Offer reasons that explain the significantly higher (but still low) solubility of CO_2 in water compared to O_2 and N_2.

Answer

Question 6

How does the enzyme facilitate the transport of CO_2 in blood? How does it maintain intracellular pH?

Answer

We'll explore the structure of two different CAs, human carbonic anhydrase II and the carbonic anhydrase from Neisseria gonorrhea in this guided problem-solving module.

Human carbonic anhydrase II

The structure of native human carbonic anhydrase II and its catalytic mechanism is shown in Figure 7 below.

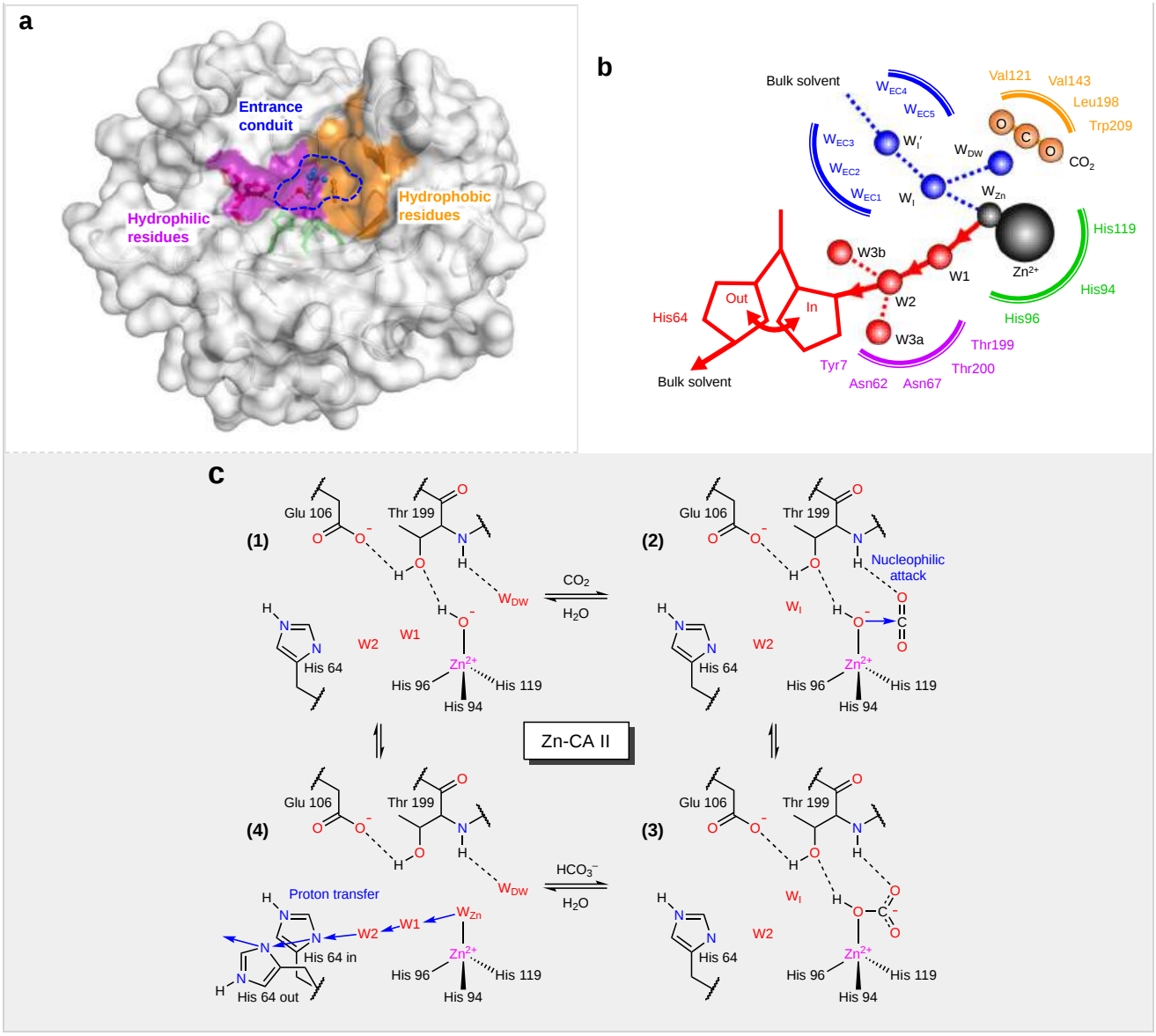


Figure 7: **Structure of native human carbonic anhydrase II (Zn-CA II) and its catalytic mechanism.** Kim, J.K., Lee, C., Lim, S.W. *et al.Nat Commun* **11**, 4557 (2020). https://doi.org/10.1038/s41467-020-18425-5. Creative Commons Attribution 4.0 International License, http://creativecommons.org/licenses/by/4.0/.

Panel a shows the active site consists of the zincbinding site, hydrophobic/hydrophilic regions, and the entrance conduit (EC).

Panel b shows the water networks in the active site that are responsible for the proton transfer (red) and substrate/product/water exchange (blue) during enzyme catalysis.

Panel c shows the forward reaction mechanism of Zn-CA II.

The active site itself lies at the bottom of a deep cavity (15 Å deep) in the protein, which is readily accessible to solvent

An interactive iCn3D model of human carbonic anhydrase II with bound bicarbonate and CO_2 (2VVB) is shown in Figure 8 below.

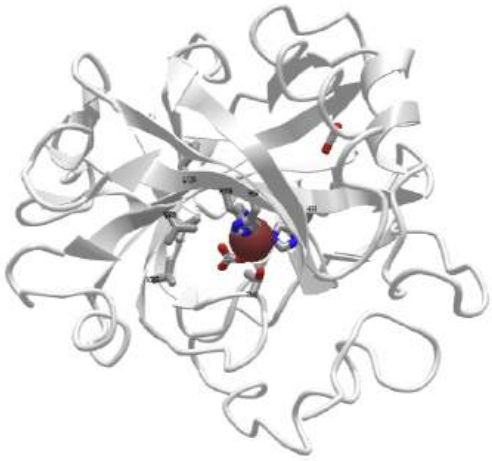


Figure 8: Human carbonic anhydrase II with bound bicarbonate and CO_2 (2VVB) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...ibgmgsm3UWhjX6

The active site residues shown in Figure 8 are labeled and shown as sticks. Bound CO_2 and HCO_3^- are also shown as sticks.

Question 7

This question addresses the Biomolecular Visualization Framework theme(s) Molecular Interactions (MI), Atomic Geometry (AG)

Using iCn3D to show the noncovalent interactions between bicarbonate, Zn, and the protein by using iCn3D. Measure the distance between the hydrogen-bonded atoms in HCO_3^- and Thr 199.

iCn3D instructions

Trackpad and Mouse Controls

•rotate: click and drag (mouse: left click and drag)

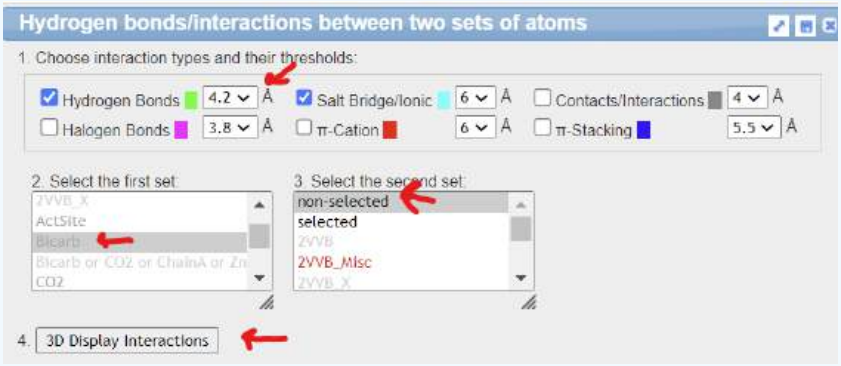
•zoom: pinch and spread (mouse: rotate the scroll wheel)

•translate: two-finger click and drag (mouse: right click and drag)

•Re-center: left click View from the top menu bar, then select “Center Selection”

•Note: ctrl-click on a PC = command-click on Mac; alt-click on PC = option click on Mac

1. Open the external link: https://structure.ncbi.nlm.nih.gov/i...ibgmgsm3UWhjX6
2. From the top menu bar, choose **Analysis**, **Interactions**
3. In the new popup window select the following prompts, then click **4. 3D Display interactions**



4. Close all but the main modeling window.

5. Zoom into the bicarbonate binding site. From the top menu bar, choose **Analysis**, **Distance**, **distance between 2 atoms**, and pick the two atoms (by holding down the Alt key or Option on a Mac) involved in the hydrogen bond between the bicarbonate and the amide

6. Rescale the label size by choosing **Analysis, Label Scale, 0.4**.

7. From the top menu bar choose **Select, Toggle highlights** to remove yellow and box highligting.

8. Save a PNG file by choosing

9. You can reload the PNG file directly into iCn3D by choosing **File**, **Open File**, i**Cn3D PNG image**

Answer

Question 8

What is a likely function of Val, Leu and Trp cluster in CAII (shown in Figures 7 and 8)?

Answer

Question x

What ligand would likely replace OH^- at low pH values? What would happen to the activity of the enzyme at lower pHs?

Answer

Question 9

From the mechanism shown in Figure 7, does bicarbonate coordinate the Zn^{2+} ion in a monodentate or bidentate manner?

Answer

Question 10

Write a verbal description of the mechanism of CAII based on Figure 7

Answer

Question 11

Why is a proton transfer path needed?

Answer

Question 12

Thr 199 plays a key role in the mechanism. State a reason for its importance in the reversible reaction.

Answer

Question 13

The dissociation constant K_D (or K_{is}) for bicarbonate binding to HCAII is about 77 mM. What kind of inhibitor might it be for the forward reaction?

Answer

Question 14

Using the equation below, at what ratio of CO_2/[HCO_3^-] would the rate for the forward reaction (CO_2 sequestration) be cut in half? Assume the V_M=1, K_M forward reaction is 1, [S] = 1 and K_{is} is 2 (a wide range of values are reported in the Brenda Database.

$$v_0 = \frac{V_M S}{K_M \left(1 + \frac{I}{Kis}\right) + S} \quad (1)$$

If you need some help, hover over -

Give me a hint!

Answer

Question 15

The rate-limiting step for human CA II is the dissociation of a proton from Zn^{2+}-bound water and not the removal of the resulting proton from the enzyme. What does that imply about the rate of removal of the proton from the enzyme?

Answer

It must be very fast, that is at diffusion-controlled limits through the H-bond channel.

Synthetic mimetics of the active site of CA have been made. These are heteromacrocycles (similar to the heme of hemoglobin) as shown in Figure 9 below.

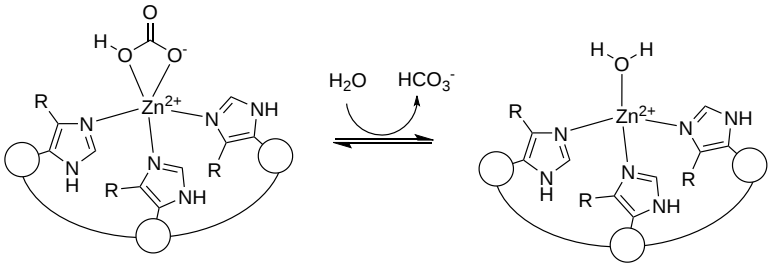


Figure 9:

The macrocycle mimetic has three imidazole groups coordinating zinc.

Question 16

The macrocycle mimetic has three imidazole groups coordinating zinc. Is the bicarbonate coordinated to the Zn^{2+} in a monodentate or bidentate fashion? From the "denticities" of the interactions of bicarbonate and Zn^{2+} for human CAII and the mimetic, which catalyst, CAII or the macrocycle would you expect to have a lower K_D for bicarbonate? How might this affect the rate-limiting step for the mimetic?

Answer

Question 17

What is the utility of having both CO_2 and HCO_3^- bind weakly to the enzyme

Answer

Carbonic anhydrase from Neisseria gonorrhea (*ng*CA)

Data from: Jo, B., Park, T., Park, H. *et al.* Engineering *de novo* disulfide bond in bacterial α-type carbonic anhydrase for thermostable carbon sequestration. *Sci Rep* **6**, 29322 (2016). https://doi.org/10.1038/srep29322. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/

Now that we understand the general chemistry, structure, and reaction mechanism of carbon anhydrase (at least the alpha human CAII form), let's explore efforts to engineer more thermostable variants. One example is the carbonic anhydrase from N. gonorrheas. This CA has been used as a target for mutagenesis to increase thermal stability of the enzyme, through the introduction of new disulfide bonds.

Even though only about 35% of the amino acids are identical, the overall structures are similar. This is illustrated in Figure 10 below.

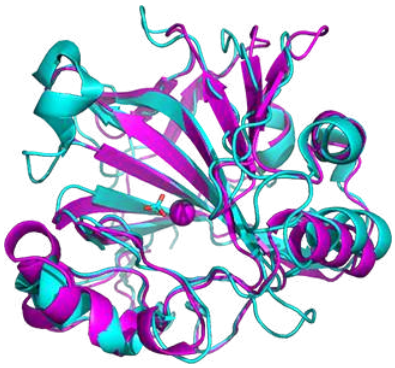


Figure 10: Alignment of the carbonic anhydrase from Neisseria gonorrhea (NG-CA) magenta,1KOQ) and human CA II (cyan, 2VVB)

The active site is mostly conserved compared to human CA II. The Zn^{2+} bound water has a pKa of around 6.5, compared to the value of 7.0 in human CA II. The hydrophobic patch (pocket) is similar, with Phe 93, Leu 153 and Tyr 72 in the NG-CA replacing Phe 95, Phe 176, Phe 70 in human CA II, respectively. The histidine ligands to Zn^{2+} are His92 (94), His94 (96), and His111 (119), where the numbers in parentheses represent Hu CA II. The proton removed from Zn^{2+} bound water is transferred to His 66 (64 in human CA II) and then to His 64.

The single disulfide bond between 181 and C28 is shown in Figure 1 below

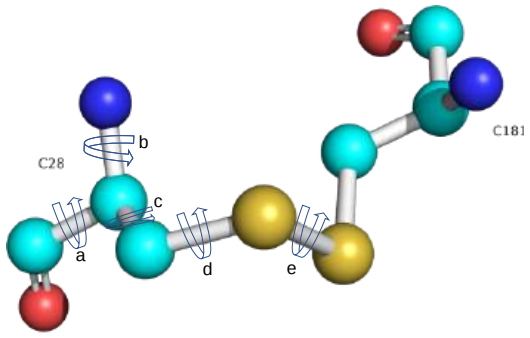


Figure 11: Single disulfide bond between 181 and C28 in wild type Carbonic anhydrase from Neisseria gonorrhea

Question 18

BIOMOLVIZ This question addresses the Biomolecular Visualization Framework theme(s) Atomic Geometry (AG)

Identify the correct torsion angles in Figure *x* above. Verbal definitions of torsional angles in a peptide chain are listed below. The successive atoms after the C_α leading away from the backbone atoms are X_β-X_γ-X_δ-X_ε (in that order). C is the backbone carbonyl C and N is the backbone nitrogen atom.

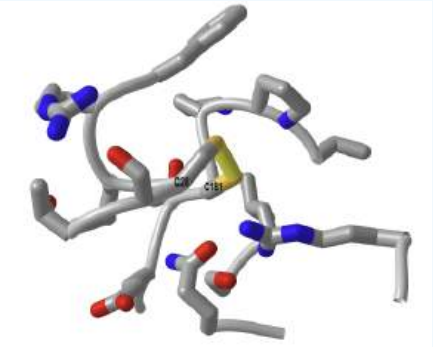
- phi (φ) is the angle of right-handed rotation around N-C_α bond. φ = 0 if the C_α-C bond is cis (eclipsed) to the C-N bond. Values range from -180 to 180 degrees.
- psi (ψ) is the angle of right-handed rotation around C_α -C bond. ψ = 0 if the C-N bond is cis (eclipsed) to the N-C_α bond. Values range from -180 to 180 degrees.
- chi_1 (χ_1) is the rotation around N-C_α-X_β-X_γ
- chi_2 (χ_2) is the rotation around C_α-X_β-X_γ-X_δ
- chi_3 (χ_3) is the rotation around X_β-X_γ-X_δ-X_ε

Answer

Question 19

BIOMOLVIZ This question addresses the Biomolecular Visualization Framework theme(s) Atomic Geometry (AG), Topology and Connectivity (TC)

Figure 12 below shows an interactive iCn3D model of the atoms within 4A of the disulfide bond in Carbonic anhydrase from Neisseria gonorrhea (1KOQ). Rotate the model to determine the approximate chi$_3$ (χ_3) dihedral angle. Hint: site down the S-S bond.



NCBI iCn3D Figure 12: Atoms within 4A of the disulfide bond in Carbonic anhydrase from Neisseria gonorrhea (1KOQ). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...nutjP6EL2PubRA

Answer

We mentioned previously that variants with higher thermal stability are likely to be more rigid and less flexible. Flexibility can be determined through analysis of molecular dynamic simulations and also by examining of the B factor values in PDB file. This number is a measure of the displacement of an atom from a mean The numbers in the last column in the file are called the temperature factors or B-factor. The B-factor describes the mean-square displacement, a measure of the displacement of an atom from an average value. If the atoms are more flexible, the electron density determined in x-ray structures is lower than if the atoms are more fixed, which gives high electron density.

To make stabilizing disulfide bonds, investigators found site chains close enough that when mutated to cysteines could potentially form disulfide bonds. In addition, they search for such residues in surface loops (without alpha and beta structure) which are inherently more flexible. Introducing disulfide bonds into the loop would stabilize it and make it more rigid. Table 2 below shows a description of the double cysteine CA variants in the study.

Variant designation	Position	Wild-type residues	Loop length	Sum of B-factors
T133C/D197C	133, 197	Thr/Asp	63	87.60
P56C/P156C	56, 156	Pro/Pro	99	80.82
N63C/P145C	63, 145	Asn/Pro	81	77.17

Table 2: **Description of the double cysteine CA variants in ngCA**

The locations of the side chair targeted for mutations to cysteine pairs are shown in Figure 13 below.

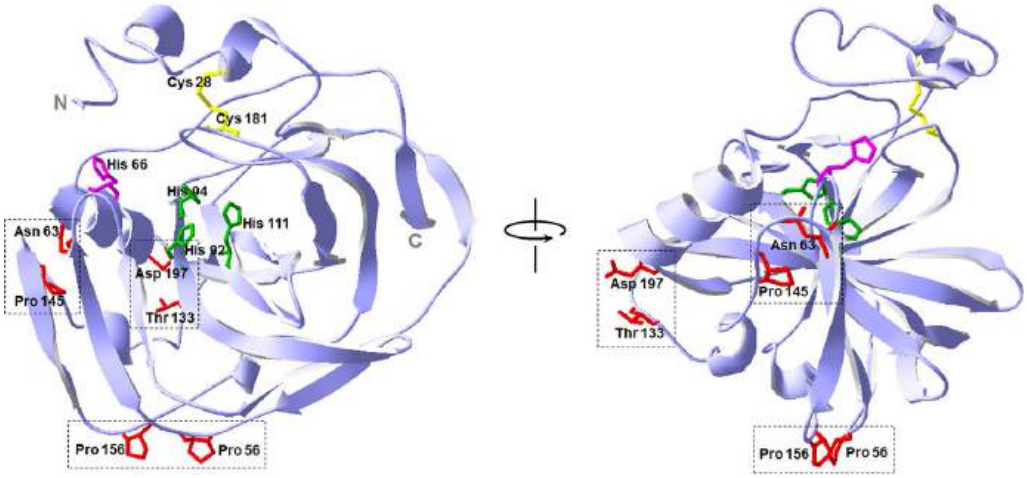
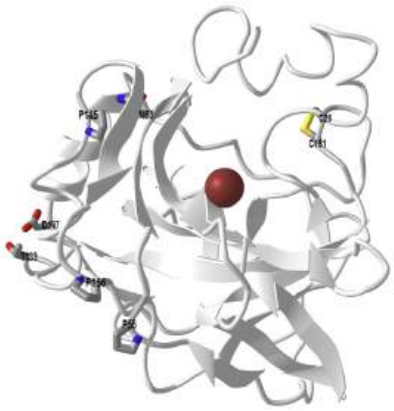


Figure 13**: 3D structure of *ng*CA and location of residue pairs for disulfide engineering**

The zinc (not shown)-coordinating histidine residues in the catalytic active site are shown in green. The proton shuttle histidine residue is shown in magenta. The native disulfide bond is colored yellow.

An interactive iCn3D model of carbonic anhydrase from Neisseria gonorrhea (1KOQ) highlighting the 3 pairs of sidechains for mutations is shown in Figure 14 below.



NCBI iCn3D Figure 14: Carbonic anhydrase (Neisseria gonorrhea) with 3 paired side chains for engineered disulfide (1KOQ) (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...KRP2tFJ1pV1FF6

The mutations were made and the wild-type proteins and three mutants were subjected to SDS-polyacrylamide gel electrophoresis (SDS-PAGE). The stained gels are shown in Figure 15 below.

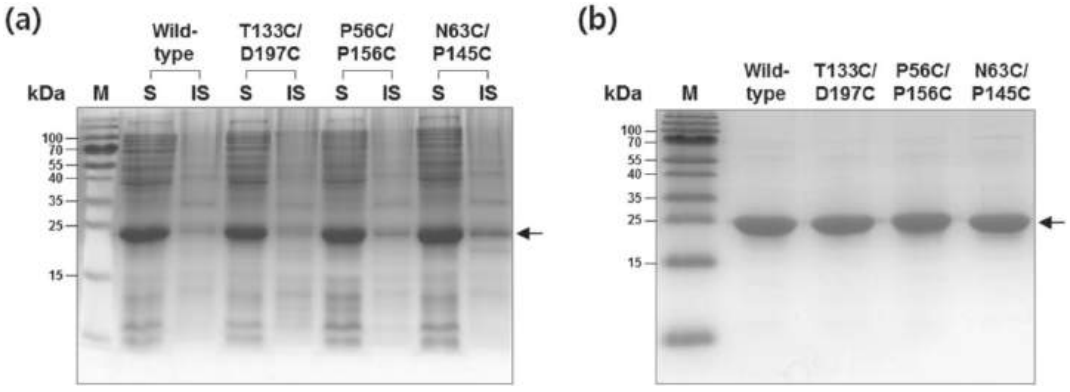


Figure 15: **Expression and purification of disulfide CA variants.**

Panel (**a**) shows the expression of the protein in transformed cells 25 °C after IPTG induction and fractionated into soluble and insoluble fractions. SHuffle strain, an engineered *E. coli* strain that promotes cytoplasmic disulfide bond formation, was used.

Panel (**b**) shows purification results. Each lane was loaded with 4 µg of each purified CA variant. The proteins were visualized with Coomassie blue staining after SDS-PAGE. The arrow indicates the position of the bands corresponding to *ng*CA variants. Lane: M, molecular weight marker; S, soluble fraction; IS, insoluble fraction.

? Question 20

a. Interpret the results of the PAGE gels in Figure 15 above

b. How pure were the proteins based on the PAGE gel result in Panel (b). Can you infer from the gel that the proteins folded correctly?

Answer

The investigators next determined if the expressed and purified wild-type and mutant proteins had the correct number of disulfide bonds. They did this by reacting the proteins in the absence and presence of dithiothreitol with DTNB or 5,5'-dithiobis(2-nitrobenzoic acid), also called Ellman's Reagent. Both structures are shown in Figure 16 below.

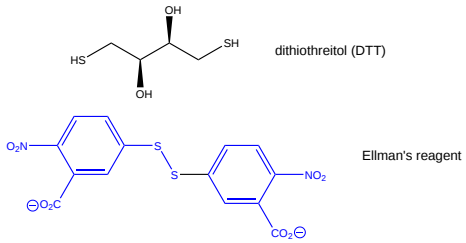


Figure 16: Structures of DTT and DTNB

DTT is a reducing agent that cleaves disulfide.

? Question 21

Draw a mechanism showing the reaction of a disulfide with DTT.

Answer

DTNB reacts with free sulfhydryls to form the 2-nitro-5-thiobenzoic acid anion leaving group that absorbs at 412 nm.

? Question 22

Draw a mechanism showing the reaction of free sulfhydryl (like Cys) with Ellman's reagent

Answer

The results of the reaction of the proteins with Ellmans's agent, in the presence and absence of DTT, are shown in Table 3 below.

CA variant	Free thiol/protein (mol/mol) [a]		Deduced no. S-S bonds
	−DTT	+DTT	
Wild-type	0.06 ± 0.02	1.80 ± 0.14	?
T133C/D197C	0.08 ± 0.02	3.79 ± 0.22	?
P56C/P156C	0.06 ± 0.03	3.89 ± 0.06	?
N63C/P145C	0.08 ± 0.03	3.75 ± 0.05	?

Table 3: Analysis of disulfides in CA using Ellman's reagent. [a]Numbers are represented in mean ± SD.

? Question 23

How many S-S would you deduce from the table are present in the wild-type and mutant enzymes? Did the correct disulfide bonds form? Explain your answers

Answer

Table 4 below shows the catalytic activities of the disulfide CA variants at 25 °C.

CA variant	CO_2 hydration activity			
	Relative esterase activity [a]	$k_{cat} \times 10^{-4}$ (s^{-1})	K_M (mM)	$k_{cat}/K_M \times 10^{-6}$($M^{-1}$ s^{-1})
Wild-type	1.00	1.44	14.2	1.01
T133C/D197C	1.49	1.97	16.7	1.18
P56C/P156C	1.03	1.44	16.9	0.85
N63C/P145C	0.55	0.27	17.3	0.16

Table 4: catalytic activities of the disulfide CA variants at 25 °C [a]The specific activity of the wild-type corresponds to 0.22 U/µmol-enzyme.

? Question 24

Why did the investigators conduct this experiment? Interpret the results

Answer

Now comes the big question: were the investigators able to engineer thermal stability into the carbonic anhydrase? Experimental results to show the thermostability of the disulfide CA mutants are shown in Figure 17 below.

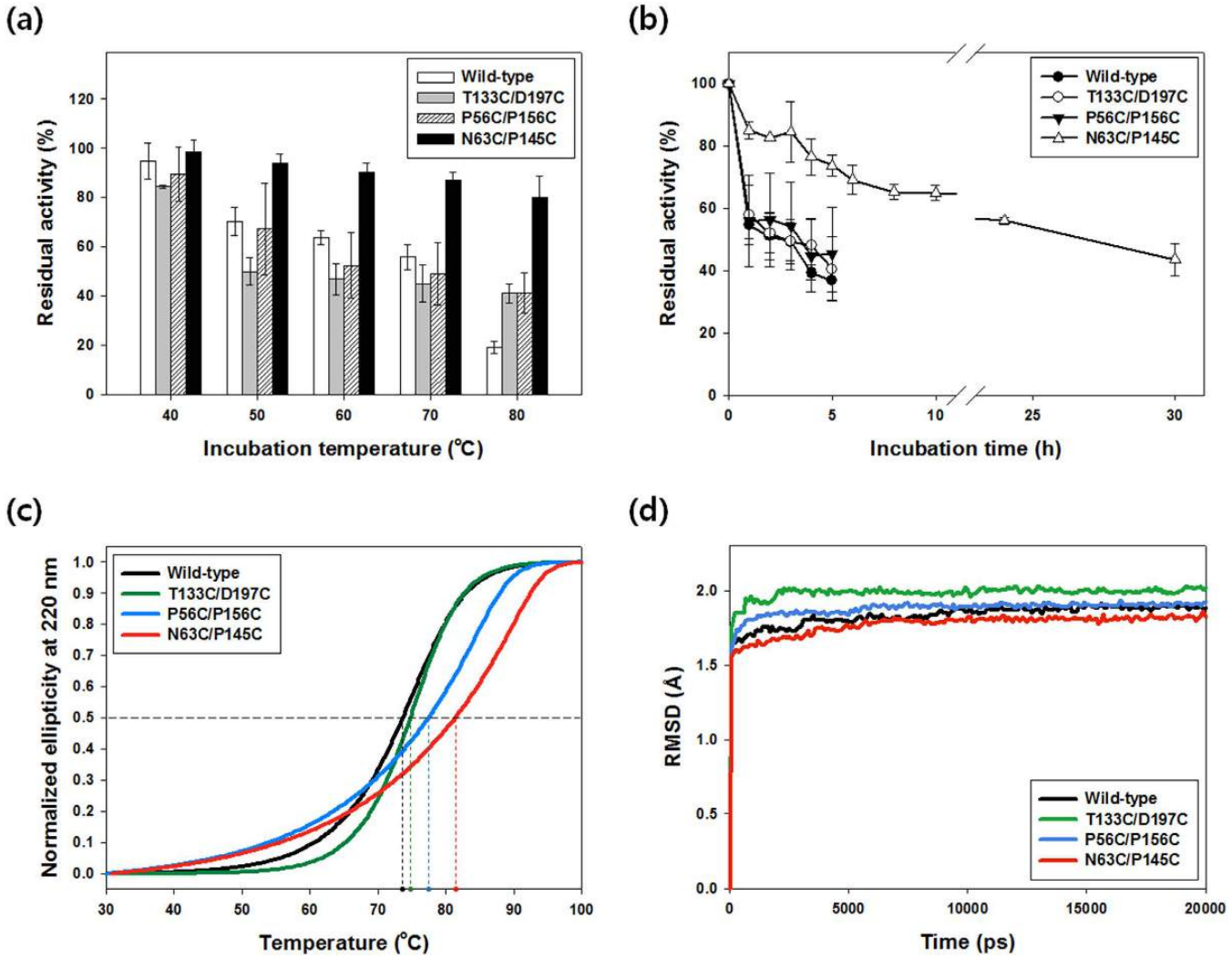


Figure 17: Thermostability of the disulfide CA variants

Panel (**a**) shows short-term kinetic stability. The enzyme solutions (40 µM) were incubated for 30 min at different temperatures, and the residual activities were measured by esterase activity assay. Activities of 100% correspond to untreated samples. Panel (**b**) shows long-term kinetic stability at 70 °C. The half-lives ($t_{1/2}$) of the CA variants were estimated by fitting the experimental data to an exponential decay curve. Each value represents the mean of at least three independent experiments, and the error bars represent the standard deviations.

? Question 25

Analyze the results in Panels (a) an (b). Which protein was most thermostable over the short (30 minute) and long (hours) incubating time at elevated temperatures?

Answer

Panel (c) shows heat-induced denaturation of disulfide CA variants. Temperature-dependent changes of the circular dichroism ellipticity were recorded at 220 nm on CD spectrometer. The denaturation curves were normalized to the fraction of unfolded protein. The horizontal dashed line indicates the point at which the fraction of unfolded protein is 0.5. The vertical dashed lines point to T_M values. Panel (**d**) shows the overall RMSD of disulfide variants from molecular dynamic simulations performed at 400 K for 20 ns.

Question 26

Analyze the results in Panel (c). What do changes in the CD helicity show? Which protein was most thermostable based on T_M values? Is the decrease in enzyme activity in panels (a) and (b) result from the denaturation of the protein?

Answer

Question 27

Analyze the molecular dynamics simulation results in Panel (d)

Answer

You may remember from both introductory chemistry and from Chapter 4.12, that you can calculate the thermodynamic parameters, $\Delta\mathbf{H}^o$ and $\Delta\mathbf{S}^o$ for N ↔D at room temperature from thermal denaturation curves using the van 't Hoff equation.

In this case, K_{eq} values can be calculated from thermal denaturation curves by monitoring change in CD signal at 220 nm, and applying this equation (also from Chapter 3.12).

$$K_{eq} = \frac{[D]_{eq}}{[N]_{eq}} = \frac{f_D}{f_N} = \frac{f_D}{1 - f_D} \tag{2}$$

From this, we can calculate ΔG^0.

$$\Delta G^0 = -R\,T \ln K_{eq} = -R\,T \ln\left[\frac{f_D}{1 - f_D}\right] \tag{3}$$

Knowing K_{eq}, ΔH^0, DS^0 can be calculated as shown below. A semi-log plot of $\ln K_{eq}$ vs 1/T is a straight line with a slope of - $\Delta H^0 R$ and a y-intercept of + $\Delta S^0/R$, where R is the ideal gas constant.

$$\begin{aligned} \Delta G^0 &= \Delta H^0 - T\Delta S^0 = -RT\ln K_{eq} \\ \ln K_{eq} &= -\frac{\Delta H^0 - T\Delta S^0}{RT} \\ \ln K_{eq} &= -\frac{\Delta H^0}{RT} + \frac{\Delta S^0}{R} \end{aligned} \tag{4}$$

The equation below shows that the derivative of equation (8) with respect to 1/T (i.e. the slope of equation 8 plotted as $\ln K_{eq}$ vs 1/T) is indeed $-\Delta H^0/R$. Equation (9) is the van 't Hoff equation, and the calculated value of the enthalpy change is termed the van 't Hoff enthalpy, $\Delta H^0 v_{Hoff}$.

$$\frac{d \ln K_{eq}}{d(1/T)} = -\frac{\Delta H^0}{R} = -\frac{\Delta H^0_{vHoff}}{R} \tag{5}$$

Using this method, the thermodynamic parameters for unfolding of the protein were calculated. The results are shown in Table 4 below.

CA variant	Melting temperature, T_M (°C)	Enthalpy change of unfolding, ΔH (kcal mol^{-1})	Entropy change of unfolding, ΔS (kcal mol^{-1} K^{-1})
Wild-type	73.6	48.8	0.141
T133C/D197C	74.7	52.8	0.153
P56C/P156C	77.4	35.1	0.091
N63C/P145C	81.4	30.0	0.085

Table 4: Thermodynamic parameters for protein unfolding for WT and mutant CAs

? Question 28

Which effects, enthalpy or entropy of unfolding, were associated with the increased thermal stability of the mutants compared to the wild-type protein. Remember were are considering the denaturation reaction, N ↔ D.

Answer

Finally, the enzymatic activity of the wild-type and mutants CAs (using a small ester substrate) were studied as a function of temperature. The relative activity of the wild-type and all 3 disulfide mutants are plotted as a function of temperature in the histogram graphs shown in Figure 18 below.

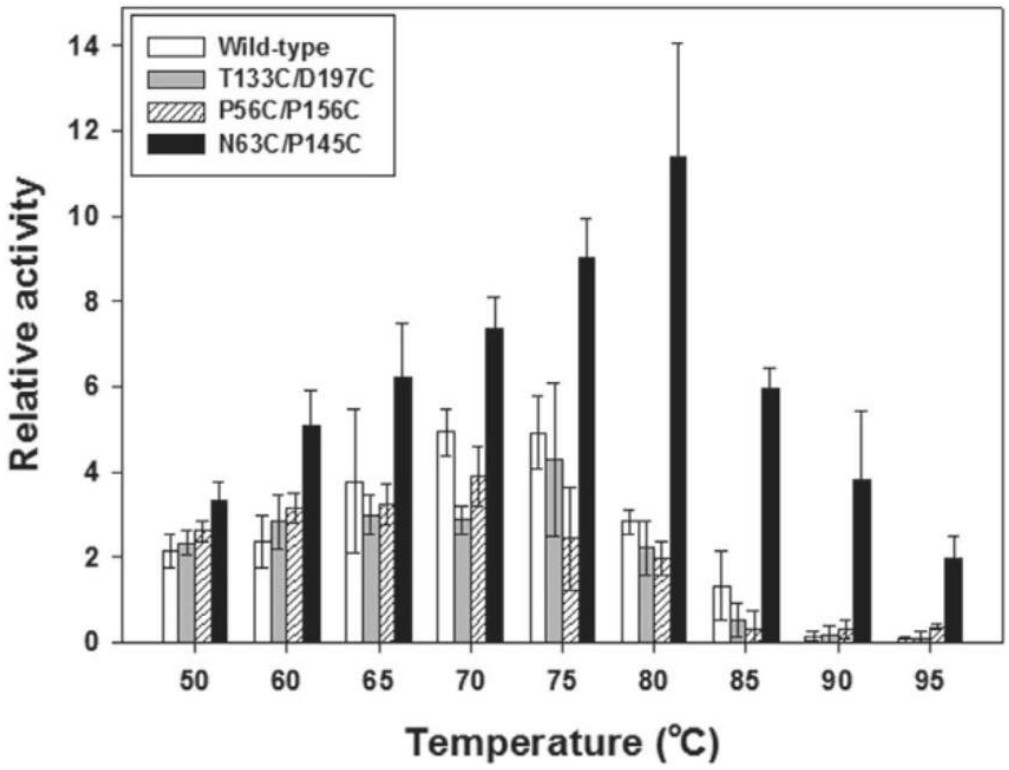


Figure 18: Effect of temperature on the activity of disulfide CA variants.

Esterase activities of disulfide variants were measured at each temperature and normalized to the activity of each enzyme at 25 °C. Each value represents the mean of three independent experiments, and the error bars represent the standard deviations. If you plotted the data as curves, you would get bell-shaped graphs.

? Question 29

Explain why the histogram plots (and line plots if they were drawn) are bell-shaped. Are the results in accordance with the previous results.

Answer

Disulfide engineering

Craig, D.B., Dombkowski, A.A. Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins. *BMC Bioinformatics* **14**, 346 (2013). https://doi.org/10.1186/1471-2105-14-346. Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0),

Several computational programs have been developed to determine amino acid pairs that could be mutated to high-temperature stabilizing disulfide bonds. The stabilizing effects appear largest when the disulfide bond is made within the largest (and most flexible) loops (between 25-75 residues). These loops also had the highest residue B-factors.

In selecting pairs to form engineered disulfide, not only proximity (distance) but also geometry (torsion angles) of the resulting disulfide bond are important. We saw this previously in the analysis of the energy of butane rotamers, as illustrated in Figure 19 below.

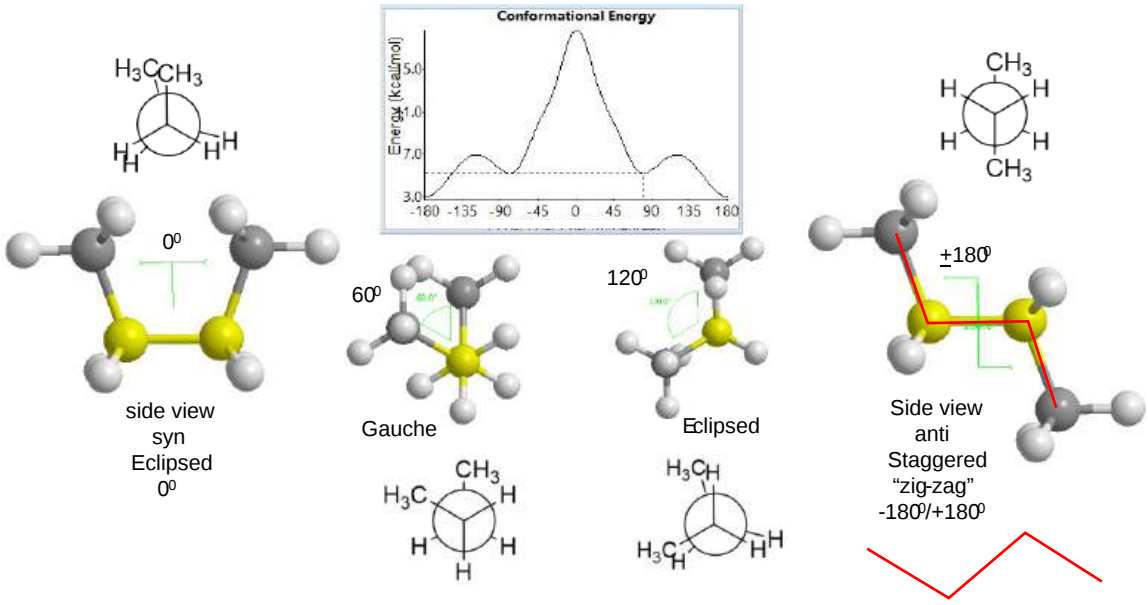


Figure 19: Newman projections for butane

Programs to determine amino acid pairs to mutate for disulfide bond formation test S-S bond torsional stability by determining the torsion angle χ_3 for the S-S bond. Evaluation of a database of many native proteins shows the χ_3 angle are centered in two major peaks at -87 and +97 degrees, as shown in Figure 20 below

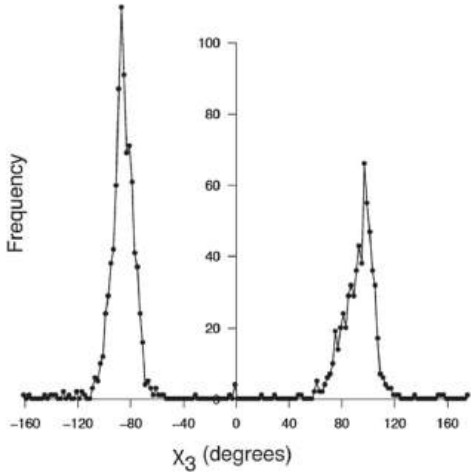


Figure 30: **Distribution of χ_3 torsion angles observed in 1505 native disulfide bonds found in 331 PDB protein structures.** Peaks occur at -87 and +97 degrees. Craig, D.B., Dombkowski, A.A. Disulfide by Design 2.0: a web-based tool for disulfide

? Question 4

Does your calculated value of χ_3 for the native disulfide in carbonic anhydrase from Neisseria gonorrhea follow the observed angles?

Answer

An empirical energy equation was developed to determine the E vs χ_3 dihedral angle for disulfide bonds (Craig, D.B., Dombkowski, A.A., ibid). It is shown below.

$$E(\chi_3) = 4.0\,[1 - \cos(1.957\,[\chi_3 + 87])] \tag{6}$$

A graph of the equation made with Excel is shown in Figure 21 below

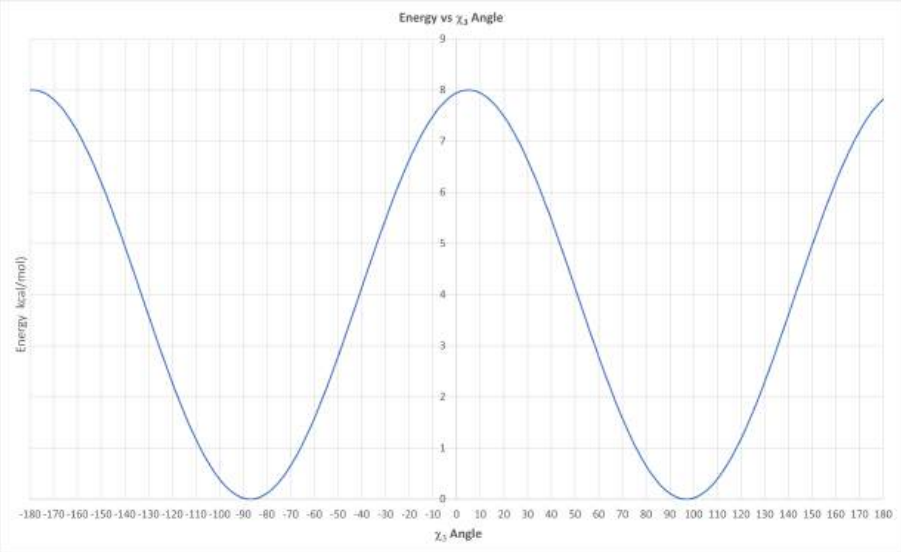


Figure 21: E vs χ_3 dihedral angle for disulfide bonds from empirical equation

Figure 22 shows the distribution of energy values in the 1505 native disulfide bonds using our updated function. The study by Craig and Dombkowski showed that almost all (90%) of disulfides in native proteins in the PDB have an energy < 2.2 kcal/mol, so this metric could be used to determine possible disulfide bond pairs created by mutagenesis.

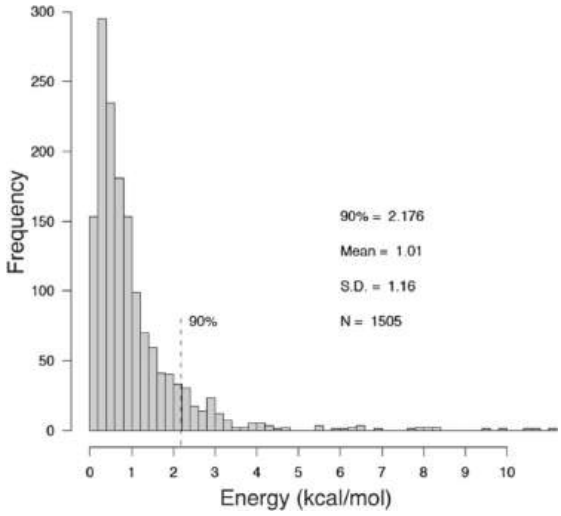


Figure 22: **Distribution of the disulfide bond energy calculated for 1505 native disulfide bonds in our survey set using the DbD2 energy function.** The mean value is 1.0 kcal/mol, and the 90th percentile is 2.2 kcal/mol. (Craig, D.B., Dombkowski, A.A., ibid)

? Question 31

You calculated the approximate value for χ_3 dihedral for the C28-C181 native disulfide bond form in Carbonic anhydrase from Neisseria gonorrhea using iCn3D in Question x above. Determine the approximate energy from the empirical function graph in Figure x above for that χ_3 dihedral. What percent of native disulfide bonds have that particular calculated energy in the 1505 native disulfide surveyed?

Answer

Yet there are other important parameters as well that would affect the energy of the disulfide bond in the protein. Figure 23 below shows a comparison of native residue B-factors in stabilizing and destabilizing engineered disulfide bonds

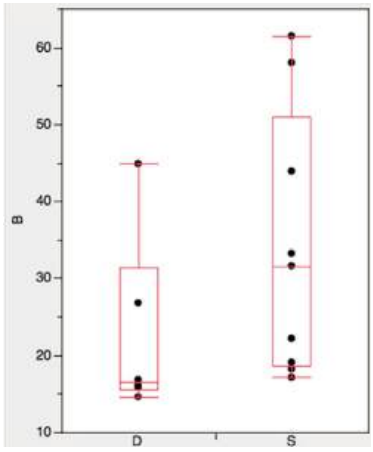


Figure 23: **Comparison of native residue B-factors in stabilizing and destabilizing engineered disulfide bonds.** The native structures associated with engineered disulfides previously reported as stabilizing (S) or destabilizing (D), based on experimental evidence, were analyzed with DbD2. The mean B-factor for residues involved in stabilizing disulfide bonds was 31.6 compared with 16.5 for those involved in destabilizing bonds, $P = 0.066$. (Craig, D.B., Dombkowski, A.A., ibid)

? Question 32

Which proteins were more stabilized by engineered disulfide, those with higher or low B factors.

Answer

Another Paper:

Prediction of disulfide bond engineering sites using a machine learning method

https://www.nature.com/articles/s41598-020-67230-z

Gao, X., Dong, X., Li, X. *et al.* Prediction of disulfide bond engineering sites using a machine learning method. *Sci Rep* **10**, 10330 (2020). https://doi.org/10.1038/s41598-020-67230-z.

The amount of data present in a single PDB file is very large, but it is nothing compared to the collective data in all PDB files. If only we could extract empirical rules from the collective PDB files that govern disulfide bond formation. It turns out we can with machine learning and artificial intelligence that can be used to develop and train predictive algorithms. Machine learning has been used to predict amino acid pairs for cysteine mutations to form engineered disulfide bonds. It recognizes 99% of natural disulfide bonds. residues. It uses these parameters:

- distances between the alpha-carbons and the beta-carbons of the bonded cysteine residues
- three torsion angles around the disulfide bonds (χ^1,χ^{ss},$\chi^{1'}$).

An example of one variable that helps define the stability of disulfide bonds is the distances between the Cα atoms for the disulfide-bonded cysteines, as shown in Figure 24 below.

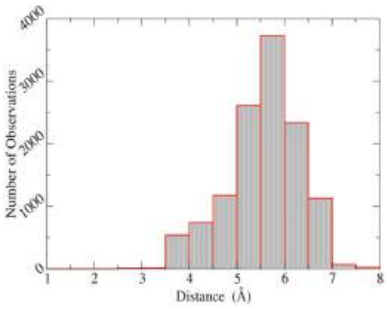


Figure 24: The histogram of distances between Cα atoms of disulfide-bonded cysteines.

The distances range from 3.0 Å and 7.5 Å. Knowing this would constrain the number of choices for paired amino acid side chains for mutagenesis to produce disulfide.

Machine learning can also be used to find other distance constraints to optimize mutagensis experiments. Figure 25 below shows a graph of 10 different distances and their relative importance in determining disulfide bond stability.

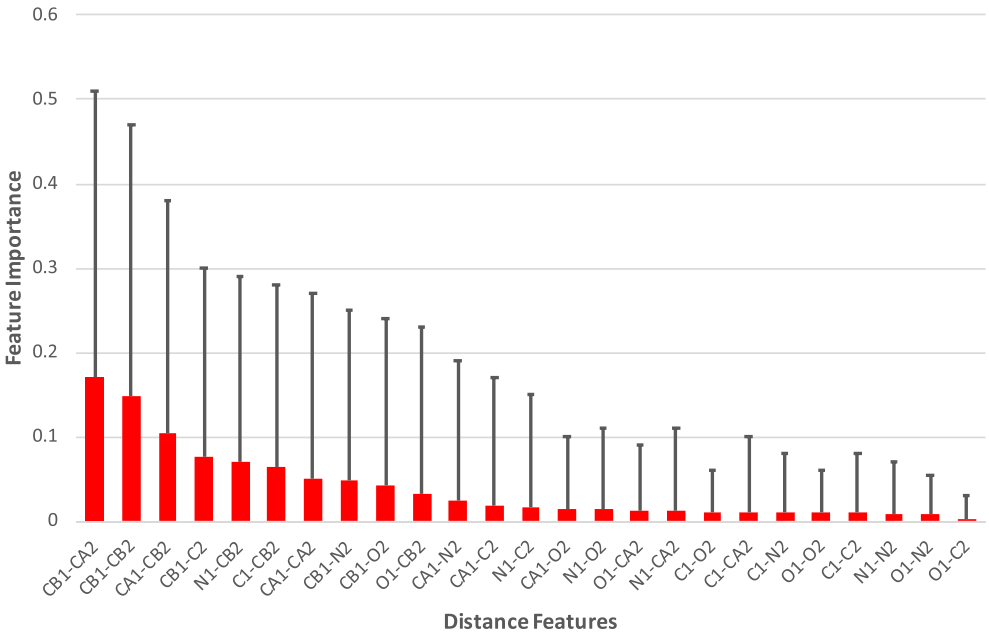


Figure 25: The relevance of the distance features to the classification outcome. Out of the 45 unique distances, 20 distances have negligible influence on the classification performance. The distances between Cβ and main-chain atoms of the pairing residue are important features in disulfide bond classifications.

An interactive iCn3D model of carbonic anhydrase from Neisseria gonorrhea (1KOQ) highlighting two pairs of amino acids identified by machine learning as candidates for mutations to disulfide-bonded cysteines is shown in Figure 26 below.

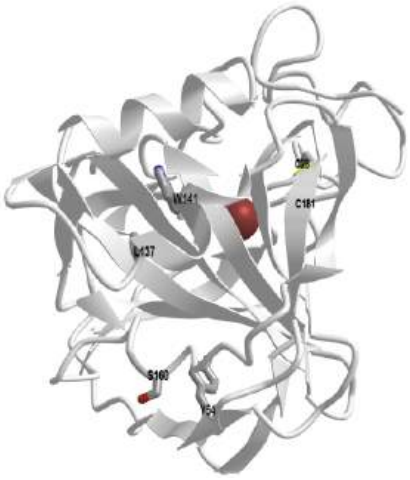


Figure 26: Carbonic anhydrase from Neisseria gonorrhea (1KOQ) highlighting two pairs of amino acids identified by machine learning as candidates for mutations to disulfide-bonded cysteines (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...ZH6Tv1urwkKhz8

? Question 33

This question addresses the Biomolecular Visualization Framework theme(s) Atomic Geometry (AG), Macromolecular Building Blocks (MB)

Using the data in Figure 25 and measurements made using the iCn3D model above (Figure 25) , determine which pair would be most likely engineered into a disulfide bond. Complete the table below with the distances you made using iCn3D.

iCn3D instructions: Open the external link and follow these instructions

Trackpad and Mouse Controls

•rotate: click and drag (mouse: left click and drag)

•zoom: pinch and spread (mouse: rotate the scroll wheel)

•translate: two-finger click and drag (mouse: right click and drag)

•Re-center: left click View from the top menu bar, then select “Center Selection”

•Note: ctrl-click on a PC = command-click on Mac; alt-click on PC = option click on Mac

Instructions

1. Zoom to clearly see the amino acid pair for distance measures
2. From the top menu bar, choose **Analysis**, **Distance**, distance between 2 atoms
3. Pick the appropriate 2 atoms for measure distance by holding down the Alt key and selecting both
4. Record the distances in the table below.

Mutation Pair	CB1-CA2 (Å)	CB1-CB2 (Å)	CA1-CB2 (Å)
1 (L137) - 2 (W141)			
1 (Y54) - 1 (S160)			

Answer

Additional References

1. Himadri Bose and Tulasi Satyanarayana. Microbial Carbonic Anhydrases in Biomimetic Carbon Sequestration for Mitigating Global Warming: Prospects and Perspectives. Front Microbiol. 2017; 8: 1615. doi: 10.3389/fmicb.2017.01615
2. Ayanne de Oliveira Maciel, Paul Christakopoulos, Ulrika Rova, Io Antonopoulou. Carbonic anhydrase to boost CO2 sequestration: Improving carbon capture utilization and storage (CCUS), Chemosphere, 299, 2022, https://doi.org/10.1016/j.chemosphere.2022.134419
3. Craig Bettenhausen. The life-or-death race to improve carbon capture. Craig Bettenhausen. CEN. July 18, 2021
4. Tanvi Sharma, Swati Sharma, Hesam Kamyab, Ashok Kumar Energizing the CO2 utilization by chemo-enzymatic approaches and potentiality of carbonic anhydrases: A review. Journal of Cleaner Production. Volume 247, 20 February 2020, 119138. https://doi.org/10.1016/j.jclepro.2019.119138
5. Mustaffa, N.I.H.; Latif, M.T.; Wurl, O. The Role of Extracellular Carbonic Anhydrase in Biogeochemical Cycling: Recent Advances and Climate Change Responses. *Int. J. Mol. Sci.* **2021**, *22*, 7413. https://doi.org/10.3390/ijms22147413.
6. Jo, B., Park, T., Park, H. et al. Engineering de novo disulfide bond in bacterial α-type carbonic anhydrase for thermostable carbon sequestration. Sci Rep 6, 29322 (2016). https://doi.org/10.1038/srep29322

Global Challenges	Trauma Post Traumatic Stress Disorder and Health	Literature-Based Guided Assessment (LGA)
		

Needed background and learning goals

This LGA is designed for students at the end of a one or two-semester biochemistry class. Any additional background in molecular biology would be helpful. This LGA should help students to:

- describe the complexity and molecular signatures of trauma and PTSD
- analyze transcriptome data and their statistical analyses, a topic not often covered in one-semester biochemistry courses
- create and interpret molecular models to address the structure/function relationships of proteins
- analyze simplified signaling networks and their regulation

This LGA is not only a formative/summative assessment for students, but also is a tutorial on some aspects of signaling and analysis of the transcriptome. Data comes not from just one paper but many.

Introduction

Trauma (and Post Traumatic Stress Disorder (PTSD) arising from it) is a global issue. Many people who have experienced combat, lived in a war zone, witnessed violent events, experienced environmental catastrophes or severe accidents, or who have experienced physical, emotional, or sexual abuse experience some symptoms of PTSD. The young, including small children through young adults (the latter whom society, unfortunately, chooses to send to war) with developing brains are especially susceptible to PTSD. Even one traumatic event may trigger PTSD. From 15-43% of boys and girls have experienced at least one traumatic event. In the U.S., Latinos, African Americans, and Indigenous peoples have higher rates than whites.

Of those boys and girls, 3-15% of girls and 1-6% of boys will likely develop PTSD. Repetitive trauma (such as emotional, sexual, and physical abuse) in children also increases the rates. Girls have a higher rate of PTSD after experiencing trauma. PTSD rates as high as 69-85% have been reported for adults who experienced sexual/physical abuse as children. In those who have experienced war, PTSD has historically been called by different names including soldier's heart, shell shock, and war neurosis. It is not a weakness of character that causes some to get PTSD. Rather it is a a physical, neurological, emotional, and behavioral outcome in some people who have been exposed to events people should never experience.

The Diagnostic and Statistical Manual of Mental Disorders (DSM-5) of the American Psychiatric Association states that PTSD is defined by four symptoms, including re-experiencing (flashbacks), dissociation, avoidance, hyperarousal and hypervigilance, and negative alterations in mood and cognition. Other common features include recurring nightmares and sleep disturbance. Seemingly innocuous clues such as sights and smells can trigger people into flashbacks and other PTSD symptoms. Here are the diagnostic criteria for PTSD from the DSM-5.

Although around 90% of people in the US have had at least one exposure to a traumatizing event, the rate of PTSD in the general population is about 8% but if the trauma results are severe (as in combat veterans and those sexually assaulted), it can rise to 25-35%, or even higher in victims of torture or prolong childhood abuse. Not everyone exposed to such stresses gets diagnosable PTSD. The severity of the trauma, levels, and history of previous stresses, psychological and social support mechanisms, and other traits affect the likelihood of getting PTSD. Genetic factors play a role with 30-40% inheritability reported for women who are

twice as likely to get PTSD and have more severe and persistent symptoms. At a neurobiological level, the mental health effects must involve circuits in the brain that are involved in the learning and extinction of fear. Even mice display sex differences in fear conditioning. Brain areas involved in fear conditioning (prefrontal cortex, amygdala, and the hippocampus) and sex differences in those regions likely play a role in differential susceptibility to PTSD.

As we will see below, PTSD affects all body systems and not just the brain. Abused children typically have poorer general health in later life. In the Adverse Childhood Experiences (ACES) study, Felitti et al. documented that over 50% of 9508 respondents to a question given after a standard health care exam self-reported one episode and 25% more than 1 episode of childhood exposure in these categories: psychological, physical, or sexual abuse; violence against mother; or living with household members who were substance abusers, mentally ill or suicidal, or ever imprisoned. Those who had experienced such episodes in 4 or more categories had 4-12 times the risk for drug abuse, alcoholism, depression, and suicidality. In addition, they smoked at a 2-4 increased rate and had over 50 sexual partners. The study found a strong relationship between childhood exposure to the abuse and dysfunction listed above and the development of adult diseases (heart, lung, and liver diseases as well as cancer) which are among the leading causes of death in adulthood.

Keywords

Here are some references.

- Felitti VJ, Anda RF, Nordenberg D, et al. Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults. The Adverse Childhood Experiences (ACE) Study. *Am J Prev Med*. 1998;14(4):245-258. doi:10.1016/s0749-3797(98)00017-8
- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3181586/).
- https://www.ptsd.va.gov/understand/c...ma%20survivors.
- https://theconversation.com/brain-sc...se-ptsd-115669
- Kritikos, M., Clouston, S.A.P., Huang, C. *et al.* Cortical complexity in World Trade Center responders with chronic posttraumatic stress disorder. *Transl Psychiatry* **11**, 597 (2021). https://doi.org/10.1038/s41398-021-01719-7. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.
- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3181586/
- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5918463/
- https://www.nature.com/articles/s41380-022-01498-7
- ACES: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6220625/

Figure 1 shows brain regions implicated in interpersonal violence (IPV), childhood trauma and sexual assault.

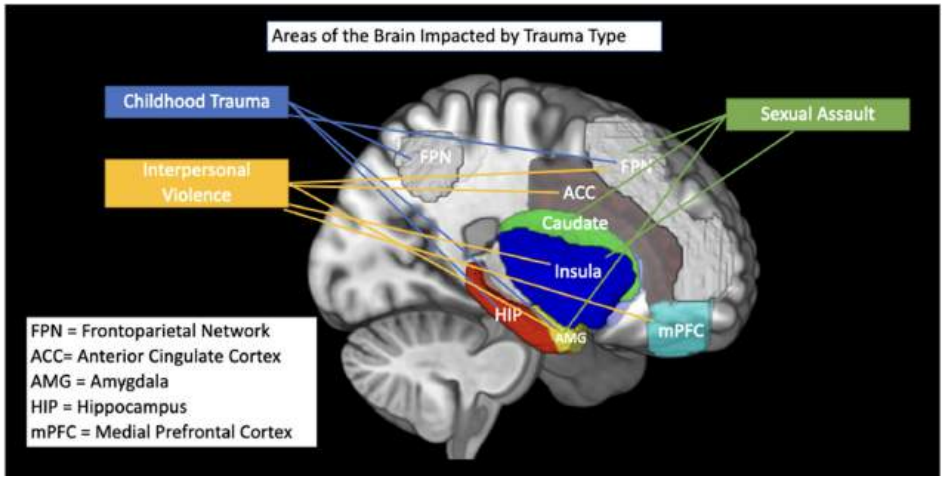


Figure 1: Brain regions implicated in IPV, childhood trauma, and sexual assault. Eder-Moreau Elizabeth et al., Neurobiological Alterations in Females With PTSD: A Systematic Review, Frontiers in Psychiatry, 13, 2022. https://www.frontiersin.org/articles...yt.2022.862476. **Creative Commons Attribution License (CC BY)**.

The general functions of each region are listed below.

- frontoparietal network - the center of cognitive control
- anterior cingulate cortex - cognitive functions as well as emotional expression, control of attention, and regulation of mood
- amygdala - formation of memories (especially emotion-associated ones) and emotion regulation
- hippocampus - involved in the regulation of learning and memory including its consolidation
- medial prefrontal cortex - integrates information from various brain regions and is involved in cognition, emotion, motivation, socialization
- Caudate nucleus - controls movement of the voluntary skeleton and is involved also in learning, memory, emotion and motivation

Brain scans show differences between people with and without a PTSD diagnosis. Changes in the cerebral cortex can be measured using a parameter called fractal dimension (FD) obtained from MRI brain scans. It is determined using multiple attributes that relate to shape complexity. Figure 1 below shows FD values in the brains of PTSD-negative to PTSD-positive first responders from the 9/11 attacks in New York City. PTSD diagnoses were based on interview and their evaluation using the DSM-IV symptom designation (reexperiencing symptoms, avoidance, hyperarousal, and negative life experiences).

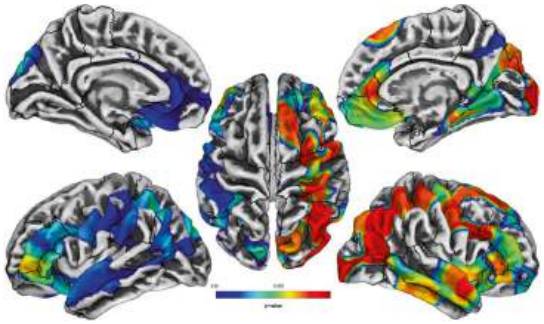


Figure 1: Kritikos, M., Clouston, S.A.P., Huang, C. *et al.* Cortical complexity in World Trade Center responders with chronic posttraumatic stress disorder. *Transl Psychiatry* **11**, 597 (2021). https://doi.org/10.1038/s41398-021-01719-7. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Gray indicates no significant differences with statistically significant differences colored from dark blue (P = 0.05) to green/yellow (P ≈ 0.005) to dark red (P ≈ 0.0005).

A review of T and P values

If you compare two sets of data, and you wish to determine if the two means are equal, you can determine P and T values. For a t-test, you define two hypotheses:

- H_0 or the null hypothesis: the population means are equal ($\mu_1 = \mu_{2)}$ (the two population means are equal)
- H_1 alternative hypothesis: the population means are NOT equal ($\mu_1 \neq \mu_2$)

You then perform a t-test which determines the difference in the means compared to the variation in each sample. Effectively it is the difference expressed in standard error units. The higher the T value, the more likely the population means are different and you can reject the null hypothesis. A T value close to zero supports the null hypothesis that the means are the same. The figure below shows t-values ranging from 3-7.5 showing that the scans from PTSD victims are different than control (combat veterans without PTSD).

You can then calculate the p-value that corresponds to the calculated t-value. We mentioned p-values above. A p-value < 0.05 suggests that the null hypothesis is correct and the same means are the same with a high probability (> 95%). The number <0.05 suggests that the risk of concluding that the means are different is less than 5%. The p values represented in the colored regions in the figure above are all <0.5)

Two Sample t-test Calculator

Another technique, **magnetoencephalography (MEG)** has also been used to compare those with PTSD and controls. A flow of ions across a neural membrane produces a magnetic field, just as does the flow of electrons through a wire. If many neurons

(50,000 - 100,000) neurons are active, a magnetic field can be measured in that region. Changes in the summative magnetic field can be measured at the millisecond time scale. Hence MEG provides direct spatial and temporal measures of brain activity. This compares to functional MRI (fMRI) which is more indirect and involves measuring oxgyen levels from blood flow in brain regions where neurons are active when people think about their trauma or view traumatic images. MEG measurement can be measured when subjects are sleeping as well.

Figure 1 below shows differences in magnetoencephalography (MEG) in the resting state of controls (combat veterans without PTSD) and combat victims with PTSD.

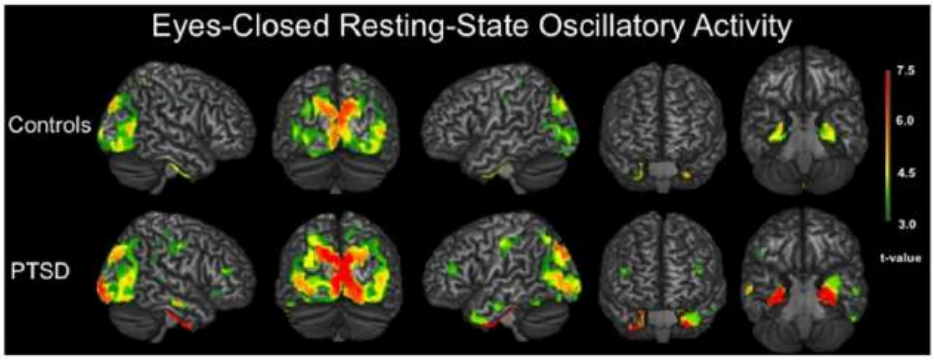


Figure 1: 3D Renditions showing rain regions with significant resting-state neuronal activity in compabte veterans without PTSD (top) and with PTSD (bottom row). Badura-Brack Amy S et al., Resting-State Neurophysiological Abnormalities in Posttraumatic Stress Disorder: A Magnetoencephalography Study. Frontiers in Human Neuroscience. 11, 12-17 (2017). https://www.frontiersin.org/articles...hum.2017.00205. **Creative Commons Attribution License (CC BY)**.

All group-level maps have been thresholded at ($p < 0.001$, corrected); a color scale bar showing respective *t*-values appears on the far right. Brain regions with significant activity in veterans without PTSD included bilateral occipital regions, inferior parietal cortices, postcentral gyri, posterior cingulate, parahippocampal gyri and other medial temporal areas. For veterans with PTSD, bilateral activity was observed in the primary motor and somatosensory cortices (i.e., precentral and postcentral gyri), superior parietal cortices, inferior parietal cortices, occipital regions, dorsolateral prefrontal cortices, inferior temporal sulci, posterior cingulate and bilateral medial temporal structures including parahippocampal gyri, hippocampi and the amygdala.

Figure 1 below shows the results of MEG brain scans documenting brain regions that differ in activity (red stronger in PTSD, blue stronger in control) in combat veterans.

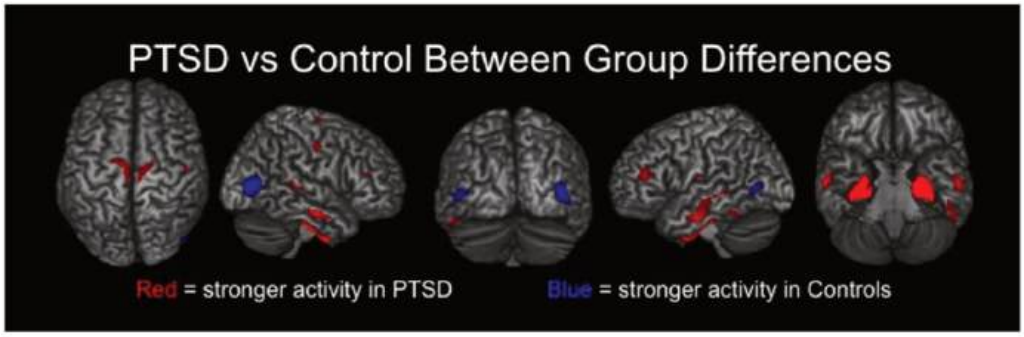


Figure 1: Group differences in resting-state neuronal activity between combat veterans with and without PTSD. Badura-Brack Amy S et al. Ibid.

In the top row, brain areas exhibiting significantly ($p < 0.01$, corrected) stronger activity in veterans with PTSD are shown in red, whereas regions with significantly weaker activity in these patients appear in blue. As shown, veterans with PTSD had stronger resting-state neuronal activity in bilateral medial temporal areas (e.g., amygdalae, parahippocampal), sensorimotor, temporal, and prefrontal regions, along with significantly weaker activity in left and right lateral occipital cortices.

These studies clearly show significant changes in neuronal activity in cohorts with PTSD compared to controls. However, as you will learn below, PTSD leads to widespread systemic changes outside of the nervous system in people with PTSD, with a plethora of biochemical changes and effects. We could focus this section on a specific documented neuronal change to excitatory glutamate

receptors such as the N-methyl-D-aspartate receptor (NMDAR) which is involved in synaptic plasticity critical learning and memory. We'll reserve that for another problem set and instead focus on the systemic changes seen in victims of PTSD to gain a broader understanding of its impact on people.

Effects of trauma on individuals - Soldiers with PTSD

Seid Muhie et al. Signatures of post-traumatic stress disorder in war-zone-exposed veteran and active-duty soldiers. Cell Reports - Medicine. 4, May 16, 2023. DOI:https://doi.org/10.1016/j.xcrm.2023.101045.

Something as complicated as trauma effects on an organism is bound to produce a large number of changes in key metabolites, transcripts, proteins and enzymes. Hence before we focus on one enzyme as an example, let's look at changes in whole systems within traumatized people. Several key studies have focused on active duty and veteran US soldiers with diagnoses of PTSD from their experiences in active combat in Iraq and Afghanistan. There have been large-scale PTSD genome-wide association studies (GWAS). In addition, the study by Muhie et used more defined cohorts and assessment measures:

- The Systems Biology Consortium (SBC) - 340 **veterans** (300 males and 40 females)
- Fort Campbell Cohort (FCC) cohorts - 180 **active-duty** service members (159 males and 21 females); 26 followed 13 month
- All were exposed in PTSD A defined by the military as follows: The person was exposed to: death, threatened death, actual or threatened serious injury, or actual or threatened sexual violence, in the following way(s): Direct exposure. Witnessing the trauma. Learning that the trauma happened to a close relative or close friend.)
- Clinician-Administered PTSD Scale (CAPS)

They analyzed

- whole blood, plasma, serum, buffy-coats (white blood cells)
- 1,305 proteins were assayed in serum samples
- DNA methylation (focusing on *cis*-regulatory sites, probes for DNA within 1,500 bp of the promoter regions)

Changes in gene expression (and other types of -omics data) are often measured using microchips to measure genes, transcripts, and proteins in samples. The processed data is usually color-coded along the visible spectra. This offers a visual impression of whether the expression is increased (red, orange, or pink) or decreased (violet, blue, or green). This visual data can be misleading if changes are expressed on a linear scale. Let's say the expression of a **control (C)** and **sample (S)** are determined and the ratio of S/C is calculated. From a mathematical sense, a linear scale has a potential flaw.

- If the expression of S is doubled with respect to C, then x = S/C =2, a +1 increase in the "signal"
- If the expression of S is halved with respect to C, then x = S/C = 0.5, a -1/2 decrease in the"signal".

We could "normalize" the data and results by using the $\log_{10}(x)$. If expression increased 10 fold, x= S/C=10, and $\log_{10}(10) = +1$. If it decreased 10 fold, then x= S/C=0.1, and $\log_{10}(0.1) = -1$. This gives up and down expression values of equal "visual" weight for the same fold-change in expression.

To better cover changes around the 2-fold increases/decreases (halved) range, $\log_2(x)$ is used instead. This gives a bigger visual change in the calculated expression data, as shown in the table below. A doubling of expression now gives a $\log_2(2) = +1$, the same number as a 10-fold increase gives if expressed as $\log_{10}(10) = +1$

doubling or 2x increase ↑	10-fold increase ↑	halving or 0.5 decrease ↓	10 fold decrease ↓
$\log_2(2) = 1$	$\log_2(10) = 3.3$	$\log_2(0.5) = -1$	$\log_2(0.1) = -3.3$

From the analysis, Muhie et al. found that several major pathways are altered in traumatized soldiers. Figure 1 below shows multi-omics (metabolites, proteins, mRNA, microRNA, and DNA methylation) changes in two major sets of pathways, inflammatory responses/oxidative stress (A, top) and angiogenesis (growth of new blood vessels/epithelial cell dysfunction (B, bottom).

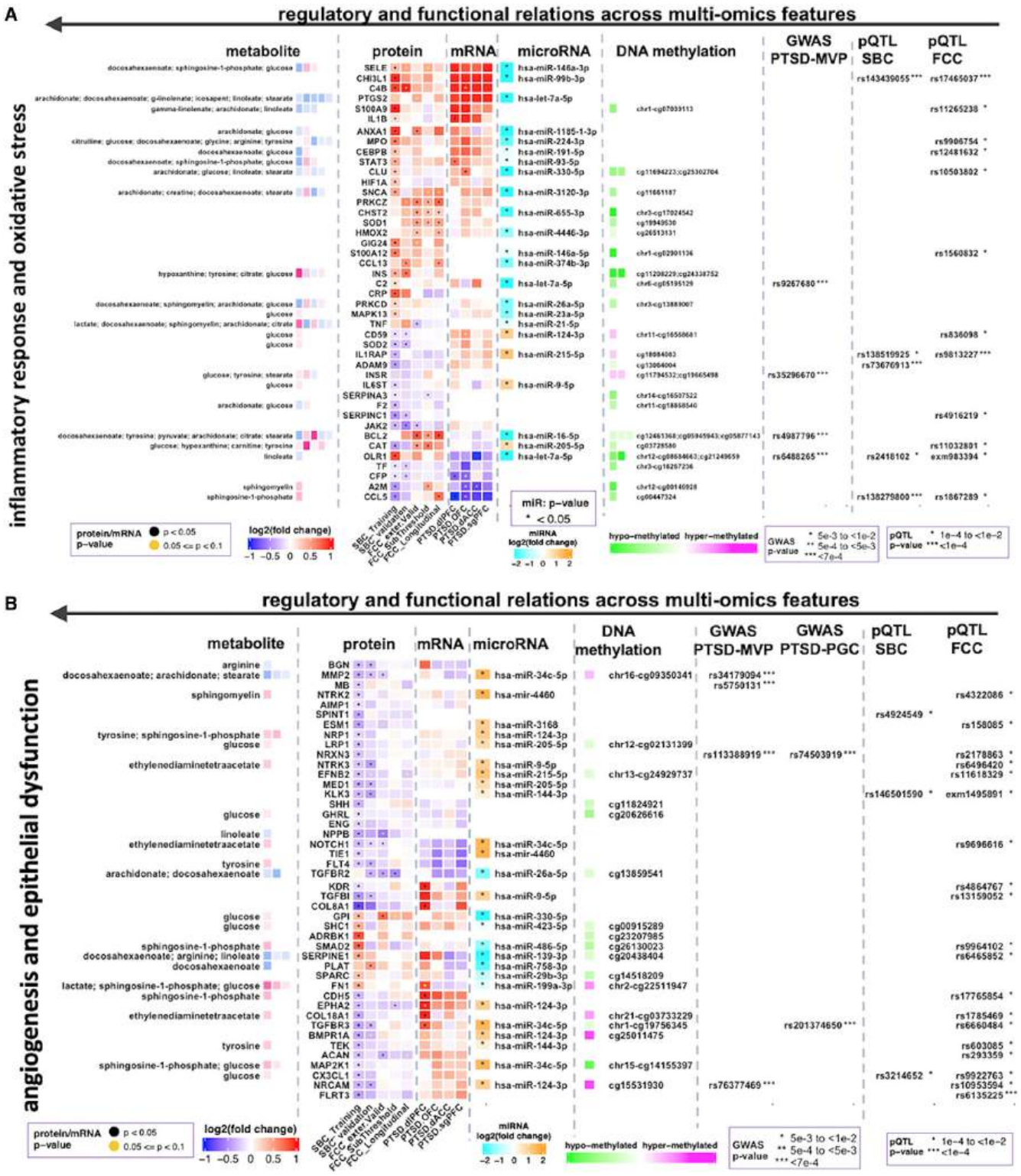


Figure 1: Integrated multi-omics showing regulatory and functional relations (horizontally from right to left) across genetic variants, epigenetic marks, microRNAs, mRNAs proteins, and metabolites. Seid Muhie et al. Signatures of post-traumatic stress disorder in war-zone-exposed veteran and active-duty soldiers. Cell Reports - Medicine. 4, May 16, 2023. DOI:https://doi.org/10.1016/j.xcrm.2023.101045. Creative Commons Attribution (CC BY 4.0).

Panel (A and B) show differentially expressed proteins (DEPs) that were persistent across PTSD cohorts and associated with (A) activated inflammatory response or oxidative stress, (B) impaired angiogenesis, epithelial dysfunction, or cardiovascular

function were integrated with multi-omics datasets and compared across SBC Training (109/109 PTSD+/−), SBC Testing (43/39 PTSD+/−), FCC Validation (47/44 PTSD+/−), FCC Longitudinal (26/26 PTSD+/−), and FCC Subthreshold (68/44 PTSD subclinical/controls) cohorts. Vertical lanes of the protein heatmap correspond to the fold changes of each protein from each group of cohorts (as shown by the labels). SBC: Systems Biology Consortium (veteran cohort), FCC: Fort Campbell (active-duty) Cohort; for brain regions (postmortem mRNA brain data): dlPFC, dorsolateral pre-frontal cortex (PFC); ACC, anterior cingulate cortex (ACC); dACC, dorsal ACC; sgPFC, subgenual PFC; OFC, orbito-frontal cortex

? Exercise 1

Look at the changes in the protein and mRNA for inflammatory response and oxidative stress. Are these expression responses generally increased or decreased?

a. inflammatory response and oxidative stress - proteins

b. inflammatory response and oxidative stress - mRNA

c. angiogenesis and epithelial dysfunction - proteins

d. angiogenesis and epithelial dysfunction - mRNA

Answer

? Exercise 2

What happens to the expression of these genes? What is their function (find it at Uniprot)

a. CDH5

b. SOD2

c. CCL5

Answer

? Exercise 3

Answer these questions on micro-RNA expression.

a. In general compare the expression of mi-RNAs in inflammatory response/oxidative stress and in angiogenesis/epithelial dysfunction

b. Do a web search to find out the role of mi-RNAs 146a (in inflammatory response/oxidative stress).

c. Do a web search to find out the role of mi-RNA 32 (in angiogenesis and epithelial dysfunction).

Answer

The authors state that "Although PTSD has primarily been conceptualized as a brain disease, it is increasingly being recognized as a systemic condition affecting multiple physiological parts and associated with divergent chronic medical conditions. Thus, PTSD

is coming to be seen as a systemic disorder rather than as a purely psychological illness."

Serine/threonine-protein kinase Sgk1 Function

Let's focus on just one enzyme, **serum/glucocorticoid-regulated kinase 1(SGK1)**, also called **serine/threonine-protein kinase**1, that seems to play a significant role in many pathways that are affected by trauma. Let's try to interpret and understand data from several research papers that discuss the role of SGK1 in several signaling pathways. As a Ser/Thr protein kinase, SGK phosphorylates many protein substrates, some of which are protein kinases (to activate/inhibit them), enzymes other than kinases, and transcription factors, as shown in Figure 1 below.

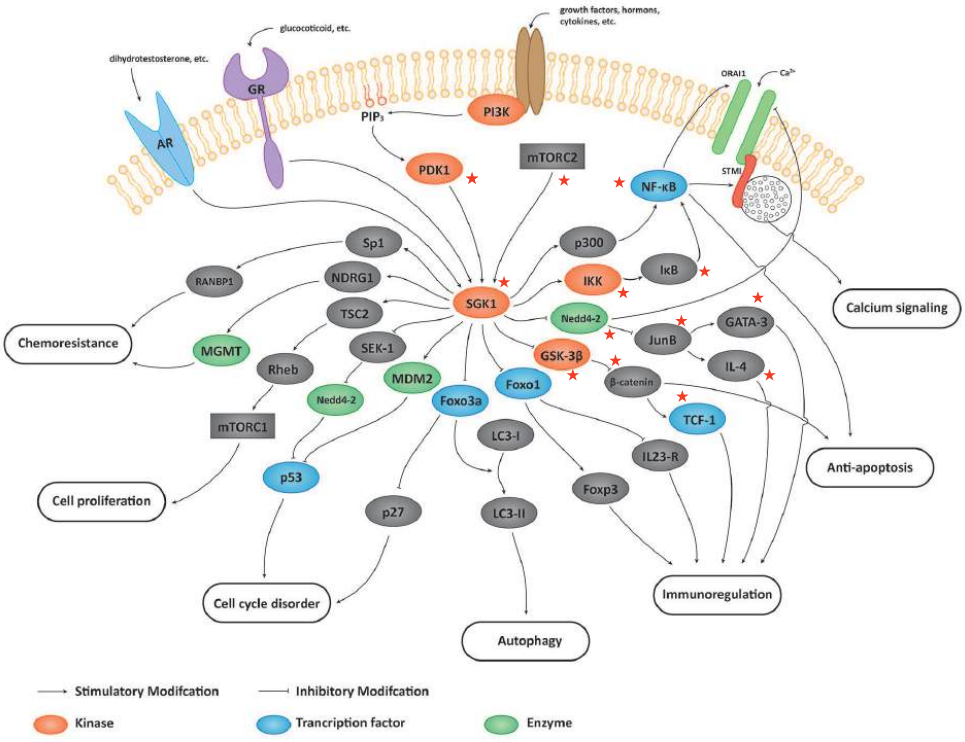


Figure 1: Signaling pathway of SGK1 in oncology. The prospect of serum and glucocorticoid-inducible kinase 1 (SGK1) in cancer therapy: a rising star. Ruizhe Zhu, Gang Yang, Zhe Cao, Kexin Shen, Lianfang Zheng, Jianchun Xiao, Lei You, and Taiping Zhang. Therapeutic Advances in Medical Oncology 2020 10.1177/1758835920940946. Creative Commons Non-Commercial CC BY-NC.

The ones we will discuss below have a red star (★) by them. SGK1 is activated by phosphorylation by two major kinases, PDK1★ and mTORC2★. We will discuss those later when we study the structure/function of the molecule in depth.

1. SGK1 regulates the immune system

The immune system is extremely complicated (for a detailed review see Chapter 5.4). Several major types of immune cells exist:

- Antigen-presenting cells - cells like macrophages, and dendritic and microglial cells in the brain that engulf foreign pathogens or cellular. In addition, virally infected and cancer cells have "tumor" antigens presented on their surfaces for recognition by the immune system.
- B cells - produce antibodies that target foreign extracellular or surface molecules of bacteria (for example).
- T cells - recognize virally-infected or cancer cells by binding to fragments of viral or tumor antigens presented on the cell surface. Two major types are **T Helper** cells that express a protein called CD4. These bind to cells like macrophages that express MHCII protein. Another set of effector T lymphocytes (for example T Cytotoxic, Natural Killer cells) express CD8. These cells express MHCI proteins and kill virus-infected cells and produce antiviral cytokine. Other CD8 cells suppress the immune response.

Signaling among these cells occurs through the release of cytokines from T cells that amplify or dampen immune cell response.

SGK1, downstream of mTORC2 is a regulator of T Helper or T_H cells. There are two major types:

- T_H1: generally produces proinflammatory cytokines (for example interleukin-2 or **IL2**, **interferon-γ**, and tumor necrosis factor β or TNF-β) that are especially involved in immune responses to intracellular pathogens (viruses and bacteria) and in prolonged autoimmune reactions against self (such as in Type I diabetes and MS).
- T_H2: generally produces "antiinflammatory" cytokines such as **IL 4**, **5**, and **13** involved in antibody responses to some extracellular parasites and allergic responses to "persistent" antigens (such as in asthma, dermatitis, etc.) and **IL 10**, which damps down the immune response, in part by turning down expression of T_H1 cytokines.

Each also regulates the other to fine-tune the immune response. They derive from a common precursor T_Hp or T_H0 cell which differentiates into the two"poles" (T cell polarization), T_H1 and T_H2, each with their own phenotype.

To test the role of SGK1 in Tcell activation, the gene was deleted to produce *Sgk1*$^{-/-}$' progeny. Under different "polarizing" sets of conditions, these cells can differentiate from a precursor T_H0 cell into either T_H1 or T_H2 cells. In the absence of polarizing cytokines, T cells from differentiated into TH1 cells.

? Exercise 4

Appropriate factors were added to T_H0 wild-type (WT) or **T-Sgk1**$^{-/-}$ cells to differentiate into the cells shown in the table below. The data is approximate based on graphs in the paper.

Cell Type	**T_H0 WT**	**T_H0 T-Sgk1$^{-/-}$**	**T_H1 WT**	**T_H1 T-Sgk1$^{-/-}$**	**T_H2 WT**	**T_H2 T-Sgk1$^{-/-}$**
IL-4 (ng/mL)	ND	ND	0.20	ND	1.8	0.55

Heikamp, E., Patel, C., Collins, S. *et al.* The AGC kinase SGK1 regulates T_H1 and T_H2 differentiation downstream of the mTORC2 complex. *Nat Immunol* **15**, 457–464 (2014). https://doi.org/10.1038/ni.2867

What conclusions can you draw from these data?

Answer

? Exercise 5

In another set of experiments, the same cell types (stimulated as above) were stained for intracellular **IFN-γ** in CD4+ T cells in the presence of an inhibitor of protein transport across the membrane. The results show that **all the T cells** had substantial IFN-γ expression.

a. Why did they include an inhibitor of protein transport in their experiments.

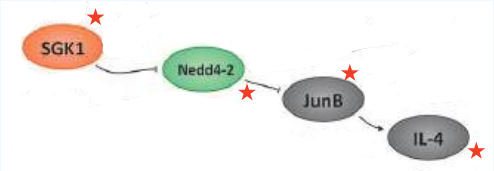
b. How do these results compare to the expression of IL-4 above

c. In a mouse model of asthma, an allergic disease, what might you expect if SGK1 was not present?

Answer

Exercise 6

How does SGK1 affect T_H2 cells and the eventual production of IL-4? It appears to do so through Nedd4 and Jun Bas illustrated in the partial signaling pathway below.



a. Is Nedd4 an enzyme? A kinase? Both?

b. Does SGK1 inhibit or activate Nedd? How specifically?

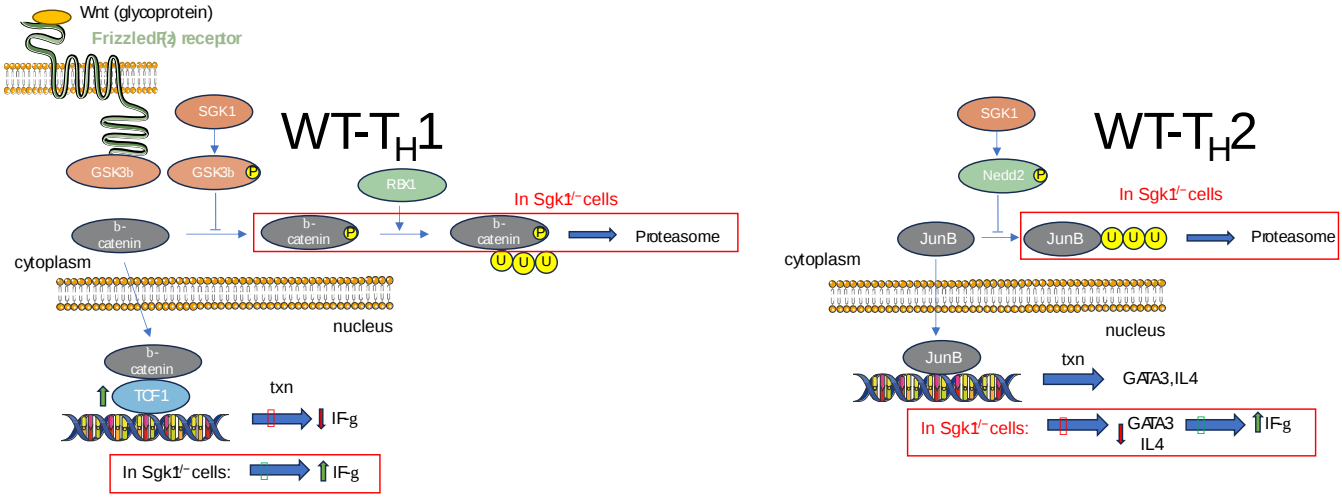
c. Search Uniprot to find out what type of enzyme human Nedd4 is.

d. What is the net effect of SGK1 on Jun B, a transcription factor, and IL-4 concentrations?

e. Jun B is higher in WT T cells cultured to produce T_H2 than in T_H1 cells. Compare Jun B levels in SGK1-deficient T cells under both T_H1- and T_H2-polarizing conditions.

Answer

Here is an expanded pathway showing signaling effects in wildtype and *Sgk1*$^{-/-}$ cells.



Exercise 7

Summarize how the lack of SGK1 causes β-catenin and JunB degradation.

Answer

From these brief questions, it should be clear that the effects of changing the expression of just one protein (SGK1 in this case) are really complicated. SGK1 is involved in:

- the control of protein synthesis and proliferation in endothelial cells (and hence the vasculature)
- major depressive disorders
- development and progression of tumors
- promotion of energy production in detached cells
- skeletal muscle homeostasis
- stress-induced cognitive impairment and PTSD
- symptoms in Parkinson's disease animal models

So let's look at one case of SGK1's involvement in brain function.

---

2. SGK1 affects brain glial cells in neurodegenerative diseases

The protein is widely expressed in the brain with increasing levels for a variety of brain diseases. Its name derives from the fact that its expression increased in rat tumor cells with the addition of serum and glucocorticoids. Parkinson's disease (PD) is associated with the death of dopaminergic neurons (DA). Damaged neurons, extracellular protein aggregates, and microbial agents are removed by microglial cells, which are a specialized type of macrophage in the central nervous system. For instance, they can initiate cell death in damaged cells.

SGK1 is elevated in the midbrain and increased expression leads to death of DA neurons. The increased expression appears to occur in the microglial cell. Hence the inhibition of microglial cell SGK1 would be a useful treatment in PD and other brain neurodegenerative diseases.

The inflammatory activities of microglial arise from the activation of a key transcription factor, NFκB, whose activation leads to the expression of inflammatory cytokines and the inflammasome. NFκB is perhaps the key regulator of the inflammatory response, which as shown above is elevated in PTSD. It is a dimer of two proteins P50/P65 (also called NFkB1/RelA), which binds to promoter regions for inflammatory genes in the nucleus. Cytoplasmic NFkB1 is inactive when bound to an inhibitor, IKB, which keeps it from translocating to the nucleus. Inflammatory signals lead to the phosphorylation of IKB by IkB kinase (or IKK, a dimer of IKKα and IKKβ), leading to the dissociation of IKB and the translocation of the NFκB into the nucleus and subsequent expression of inflammatory cytokines.

SGK1 increases NFκB signaling and the deleterious inflammatory response in glial cells as illustrated in Figure 1 below. For simplicity, both NFκB and IKKβ are shown as monomers

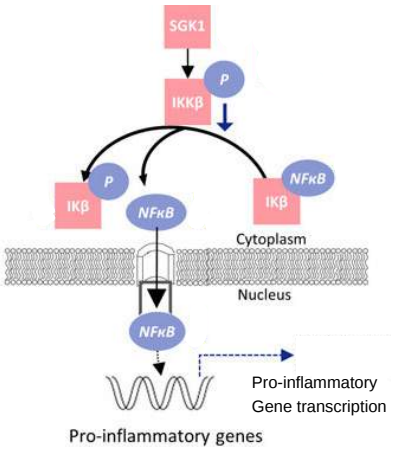


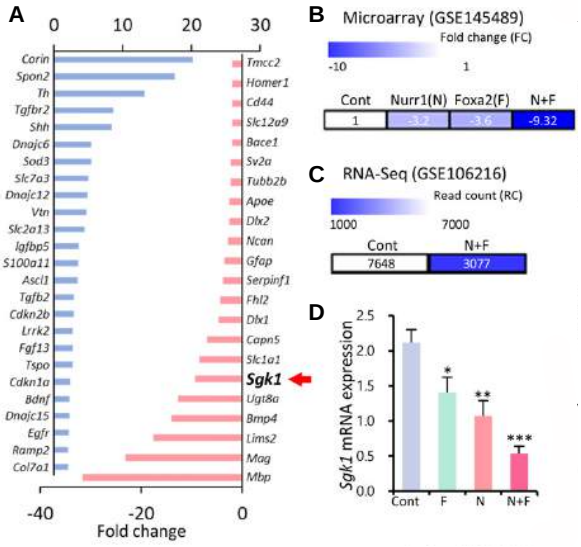
Figure 1: Modified from Kwon et al. **SGK1 inhibition in glia ameliorates pathologies and symptoms in Parkinson disease animal models.** EMBO Molecular Medicine, 2021. https://doi.org/10.15252/emmm.202013076. **Published under the terms of the CC BY 4.0 license**

? Exercise 8

From the simplified pathway diagram above, how does SGK1 promote inflammation?

Answer

Hence inflammation can be decreased if SGK1 can be negatively regulated at the transcription or post-transcriptional levels. In glial cells, two other proteins, Nurr1 (N) and Foxa2 (F), are involved in the regulation of SGK1 transcription. To test their effects, glial cells were transduced (genes added) with the N and F genes, and in a mock control without them. The results are shown in the figure below.



Panel A shows that Sgk1 is significantly downregulated (drop in expression, red arrow, red rectangle) in the glial cells transduced with Nurr1 (N) and Foxa2 (F).

Panel B and Panel show fold change (FC) or actual level changes (red count).

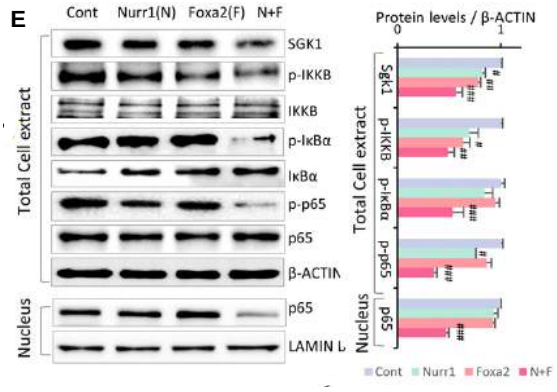
Panel D shows RT-PCR mRNA expression levels.

? Exercise 9

What are the effects of the addition of the gene for N, F, and N+F compared to the sham control?

Answer

Panel E below shows a Western blot of PAGE gels showing protein levels and graphs of the Western blot data showing quantitative values. P65 is one of the subunits of NFKB. β-actin gene is a "housing-keeping" gene whose levels are not expected to change.



Exercise 10

What happens to the concentration of these proteins on the treatment of cells with N+F? The small red letter p indicates that a phospho group is attached as a posttranslational modification. (p65 is the name of that protein and the p indicates no particular phosphorylation state.) Interpret the results.

SGK1	p-IKKB	p-IKBα	p65 (NFKB) total cell	p65 (NFKB) nucleus	β-actin
?	?	?	?	?	?

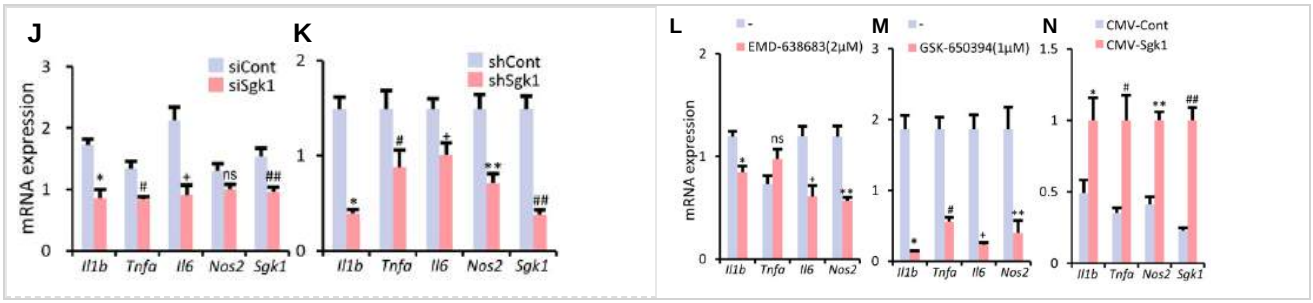
Answer

See the table below.

SGK1	p-IKKB	p-IKBα	p65 (NFKB) total cell	p65 (NFKB) nucleus	β-actin
decreases	decreases	decreased	constant	decreased	constant

The N + F-mediated downregulation of NFκB signaling is attained by inhibiting IκB phosphorylation (E) and thus blocking the release of NFκB from the NFκB-IκB inhibitory complex

What happens to the actual expression of inflammatory cytokines if inhibitory RNA (that block SGK1 translation by binding to SGK1 mRNA) are added to cells or if an "extra" gene is added to overexpress SGK1? The results for inhibition are shown in Panels J-M and for overexpression in N below.



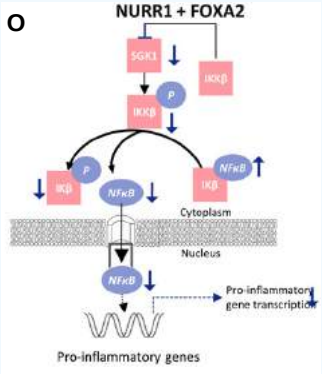
Exercise 11

Interpret the results

Answer

Exercise 12

The figure below summarizes the effect of NURR1, FOXA2 on SGK1-mediated expression of cytokines in glial cells. Summarize these effects in words.



Answer

Structure/Function of SGK1

Structure information is derived from:

Zhao B, Lehr R, Smallwood AM, Ho TF, Maley K, Randall T, Head MS, Koretke KK, Schnackenberg CG. Crystal structure of the kinase domain of serum and glucocorticoid-regulated kinase 1 in complex with AMP PNP. Protein Sci. 2007 Dec;16(12):2761-9. doi: 10.1110/ps.073161707. Epub 2007 Oct 26. PMID: 17965184; PMCID: PMC2222817. NO CREATIVE COMMONS PERMISSIONS

Serum and glucocorticoid-regulated kinase 1, SGK 1, a member of the **AGC** kinase family which includes protein kinases A (PKA), G (PKG), and C (PKC), is a serine/threonine kinase. It has 3 domains, an N-terminal PK-like domain, a kinase domain, and a C-terminal hydrophobic domain. Its catalytic domain is around 50% homologous to other members.

The protein is regulated at the transcriptional and also at the post-transcription level, the latter by phosphorylation. Two are required. One occurs at Ser 422 in the C-terminal hydrophobic domain by mTORC2. This allows its recruitment to membrane phosphatidylinositol 3, 4, 5-trisphosphate-dependent kinase (PDK1), which phosphorylates SGK 1 in a second location, Thr 256, located in a conformationally flexible activation loop in the catalytic domain, leading to the formation of a competent active site in the enzyme.

All AGC kinases bind ATP and phosphorylate Ser/Thr side chains on target proteins. Hence the conformations of the catalytic domain of the active protein kinases are similar. Their inactive structures diverge more, which allows them to bind to specific target domains or proteins in a process that provides specificity for the phosphorylation. Such is the case with SGK 1.

Figure 6s shows an interactive iCn3D model of the catalytic domain of inactive serum and glucocorticoid-regulated kinase 1 (SGK 1) in complex with AMP-PNP (2R5T).

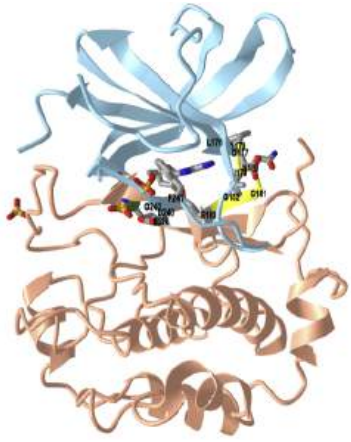


Figure 4: Catalytic Domain of Inactive Serum and Glucocorticoid- Regulated Kinase 1 in Complex with AMP-PNP (2R5T).. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...695vFs5jzQjH67

As with other AGC kinases, the catalytic domain of SGK1 has an N-terminal (light blue, mostly antiparallel beta strands) lobe and C-terminal (salmon, mostly loops and alpha helices) lobe, with ATP and its substrate analog AMP-PNP binding in the groove between them. Key amino acids involved in catalysis are shown. The ATP binding site consists of residues 104-136 with key catalytic residues at 218-230.

The first phosphorylation occurs at Ser 422 found in the hydrophobic/aromatic motif FLGFSYA. This sequence is conserved in AGC kinases, which have the general consensus sequence FXXFS/TF/Y. This entire motif is not seen in the structure above. The second phosphorylation occurs at Thr 256 (CPK-colored sticks, labeled).

Figure 6 shows another interactive iCn3D model of the inactive SGK 1 and AMP-PNP highlight activation and catalytic features (2R5T).

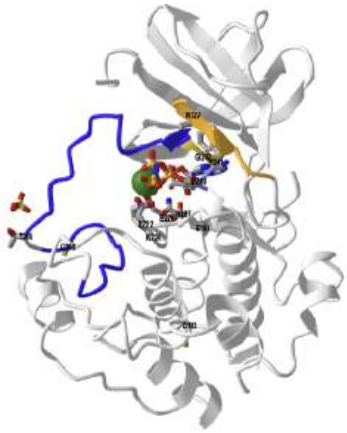


Figure 4: Inactive SGK 1 and AMP-PNP highlighting activation and catalytic features (2R5T). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?MioLpk9FqXoVEbBp8

The activation loop is shown in blue. Activation of the kinase requires phosphorylation of the activation loop Thr 256 (shown in sticks and labeled) by PDK1. (Note that it is next to a SO_4^{2-}, part of ammonium sulfate used in the crystallization process.) Key residues involved in the binding of AMP-PNP in the active site are shown as CPK-colored sticks and labeled.

Catalysis requires binding of substrate, in this case ATP (or in the model above AMP-PNP) and any activating conformational changes that position catalytic residues around the substrate:

- **Glu 146 and Lys 127**: E146 (E91 in PKA), is found in the N-terminal lobe αC-Helix of most AGC kinases. It interacts with a key catalytic residue K127 (K72 in PKA) in the catalytic loop. However, the equivalent αC-Helix is **not** found in the inactive SGK1. Rather it forms a beta-strand. This is perhaps the main feature differentiating inactive SGK1 (along with the activation loop) from other AGC kinases. In AGC kinases, the activation loop is connected to the N-lobe through the αC helix.
- **Lys 72** which stabilizes the β phosphate of ATP.
- **Lys127, Glu183, Asp222, Lys224, Glu226, and Asn227**: These mostly charged side chains interact with the phosphates in the substrate (analog)
- **Asp 240, Phe 241, and Gly 242 (DFG motif)**: This motif is in the activation loop and helps in catalysis by positioning a molecule of ATP bound to magnesium or manganese for phosphoryl transfer which helps the SGK1 to switch between DFG-out (inactive) and DFG-in
- **Cys193** and **Cys258**: These might also affect the activity.

? Exercise 13

Zoom into the iCn3D model and localize E146 and K127. What did you find? Offer an explanation. With what does K127 interact?

Answer

? Exercise 14

Zoom into the iCn3D model and answer the following questions.

a. What is the role of K127 in binding and in catalysis?

b. What is the role of K224?

c. E226:

d. N227

e. E183

f. D222

Answer

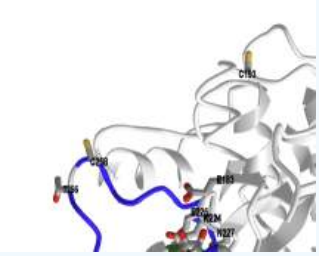
Cys 258 near T256 and it forms a S-S with S193 to form a covalent homodimer :

Exercise 15

Locate Cys 258 and C193 in the iCn3D above. Propose a mechanism in which they might alter activity.

Answer

A disulfide bond between these two Cys on different monomers results in a covalent homodimer. One of the Cys is near the Thr 256 which is phosphorylated by PDK1 to activate SGK 1. This might not occur in the covalent dimer. Also, oxidative modification of Cys 258 might affect the binding site on SGK1 for PDK1 altering its post-translation phosphorylation and the activity of the enzyme.



Exercise 16

One clue to conformational flexibility is to color code the model based on the B-factor.

Use this link and do the following in iCn3D:

- Choose **Analysis, Seq. and Annotations**, and then the Details tab.
- Under Proteins, click 2R5T_A to highlight the protein
- Choose Color, B-Factor, Original

What do the B-factors tell you about the conformational flexibility of the inactive protein?

Answer

Exercise 17

Make an iCn3D model of 2R5T.

- Select these amino acids by choosing **Analysis, Seq and Annotations**, and then the details tab: Ile104, Val112, Ala125, Val160, Leu176, Asp177, Tyr178, Ile179, Leu229, Thr239 (Click/sweep each amino acids to highlight them)
- **Select, save selection**, and give it a name
- **Style, Sidechains, Stick**
- **Color, Atom**
- **Analysis, Label, Per residue and number**
- **Analysis, lab scale, 3.0**
- **Style, background, transparen**t
- **Select, Clear selection** (to remove yellow highlights)

- **Files, Share Link, Lifetime short link,** then copy it and save it in a file documenting your work

What role do these amino acids play in the SGK 1 catalytic domain?

Answer

Investigators studied the phosphorylation and the activity of the protein *in vitro*. The structure reported above and used in the phosphorylation studies contained planned mutations.

? Exercise 18

Investigators make these mutations in the protein. Explain why they did.

a. S74A, S78A, S397A, S401A.

b. S422D mutation.

c. R192A

Answer

The inactive SGK1 structure, with its unique conformation in the αC and activation loop regions, may indicate that αC and the activation loop are important for substrate recognition and binding and thus are two key regulatory elements within the kinase domain. Upon activation, SGK1 will adopt a conformation similar to other active forms of AGC family kinases.

The amino acid sequence of SGK1 provides a plausible explanation for the disordered αC helix. Six successive large hydrophilic residues (KKKEEK, residues 136–141) are located at the beginning of the segment corresponding to αC, while there are only three hydrophilic residues in this region in AKT2 and four in PKA. These large hydrophilic residues are generally quite solvent-exposed and may make the peptide chain more flexible.

? Exercise 19

What happens on the phosphorylation of Thr 256 in the activation loop of SGK1? A comparison with analogous residue, Thr 197) in Protein Kinase A (PKA) would be helpful to know.

- In PKA, Thr 197 becomes phosphorylated and interacts with Arg165 and Lys189 in active PKA
- In SGK 1, Thr 256 becomes phosphorylated and likely interacts with Arg 221 and Lys 245 (the equivalent of Arg 165 and Lys 189 in PKA) in active SGK1.

a. What type of noncovalent interactions occur between pThr197 and Arg 165/Lys 189?

b. Measure the distance between Thr 256 and Arg 221/Lys 245 in this rendered structure of SGK 1 and describe what must happened to these residues in the likely active state of SGK 1.

In the iCn3D model,

- Choose **Analysis, Seq and Annotation**, the Detail tab, and then select Arg 221/Lys 245 with your mouse.
- **Select, Save Selection**, and name it R221K245.
- **Style, Side Chains, Sticks**
- **Color, Atom**

- **Analysis, Label, Per Residue and Number**
- **Analysis, Distance, Between 2 sets**
- Pick the predefined T256 as 1 set, and R221K245 as the second
- Choose Display to see the distance between them in Angstroms.

Answer

Bashir A. Akhoon, Neha S. Gandhi, Rakesh Pandey, Computational insights into the active structure of SGK1 and its implication for ligand design, Biochimie,
Volume 165, 2019, Pages 57-66, ISSN 0300-9084, https://doi.org/10.1016/j.biochi.2019.07.007. NO CREATIVE COMMONS PERMISSIONS

Additional Results

1. Transcriptomic studies of four prefrontal cortex subregions from postmortem tissue of people with PTSD show that PTSD has different molecular traits than major depressive disorders. 32 unique genes were identified. A few keys are listed below.

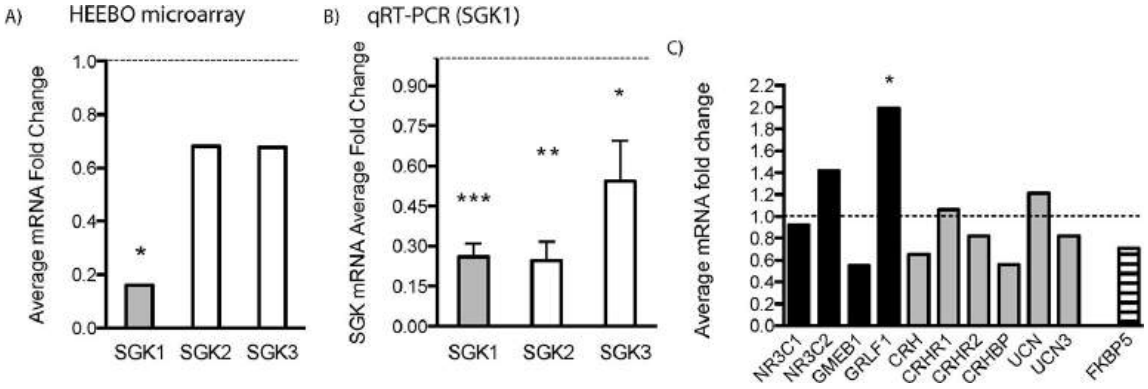
- ELFN1 (downregulated): Synaptic adhesion, allosteric regulator of metabotropic glutamate (excitatory) receptors. It's required for formation of synapses.
- GABA (downregulated): GAD2 encoding glutamic acid decarboxylase-2 required for GABA synthesis, and SLC32A1 which is a vesicular transporter for GABA and glycine. These suggest a downregulation of GABA (an inhibitory neurotransmitter) signaling
- UBA7 (downregulated): ubiquitin-like activating enzyme involved in the expression of inflammatory genes.
- FKBP5 (upregulated): a glucocorticoid receptor chaperone. Key driver on females

---

Licznerski P, Duric V, Banasr M, Alavian KN, Ota KT, Kang HJ, et al. (2015) Decreased SGK1 Expression and Function Contributes to Behavioral Deficits Induced by Traumatic Stress. PLoS Biol 13(10): e1002282. https://doi.org/10.1371/journal.pbio.1002282. Creative Commons CC0 public domain

This group looked at SGK1 expression in the postmortem prefrontal cortex of PTSD subjects (who died with PTSD) and in rat models of PTSD when rats were subjected to stimuli that resulted in symptoms of helplessness (exposure to acute foot shock stress without the possibility of escape) and anhedonia (lose of ability to experience pleasure as measured by decreased preference for a sweetened sucrose solution).

Results of the studies of mRNA levels for SGK1 and two isozymes, SGK 2 and 3 from postmortem samples are shown below.



Panel (A) shows microarray gene expression analysis of postmortem dorsolateral PFC (DLPFC) samples collected from patients with PTSD. The * indicates significant results. The dashed line indicates healthy controls.

Panel (B) shows real-time qPCR was conducted to verify the microarray findings. Data are expressed as a mean fold change ± standard error of the mean (SEM) switch student's *t* test values of $*p < 0.05$, $**p < 0.01$, $***p < 0.001$].

Panel (C) shows expression levels of stress- and glucocorticoid-regulated genes were also examined. Changes are shown as a mean fold change. Asterisk indicates a significant *p*-value ($*p < 0.05$, FDR adjusted). NR3C1, glucocorticoid receptor; NR3C2, mineralocorticoid receptor; GMEB1, glucocorticoid modulatory element binding protein 1; GRLF1, glucocorticoid receptor DNA binding factor 1; CRH, corticotropin-releasing hormone; CRHR1, corticotropin-releasing hormone receptor 1; CRHR2, corticotropin-releasing hormone receptor 2; CRHBP, corticotropin-releasing hormone binding protein; UCN, urocortin; UCN3, stresscopin (urocortin 3); FKBP5, FK506 binding protein 5.

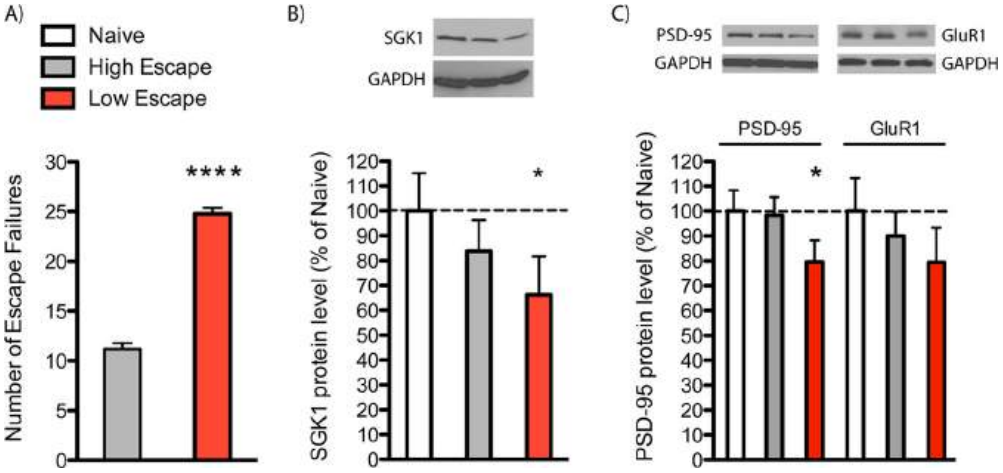
? Exercise 20

Interpret the results.

Answer

Rat brain SGK1 from either low or high escapers from shock (a model for learned helplessness) was run on PAGE gels. The results are shown below.

dfdfd



Panel (A): Rats were exposed to inescapable shock (day 1), tested in active avoidance (AA) (day 4), and then sacrificed (day 8) as indicated. Animals were separated into low- and high-escape groups according to their performance in AA test. Data are the mean ± SEM [t(12) = 14.39, Student's *t* test, ****$p < 0.0001$].

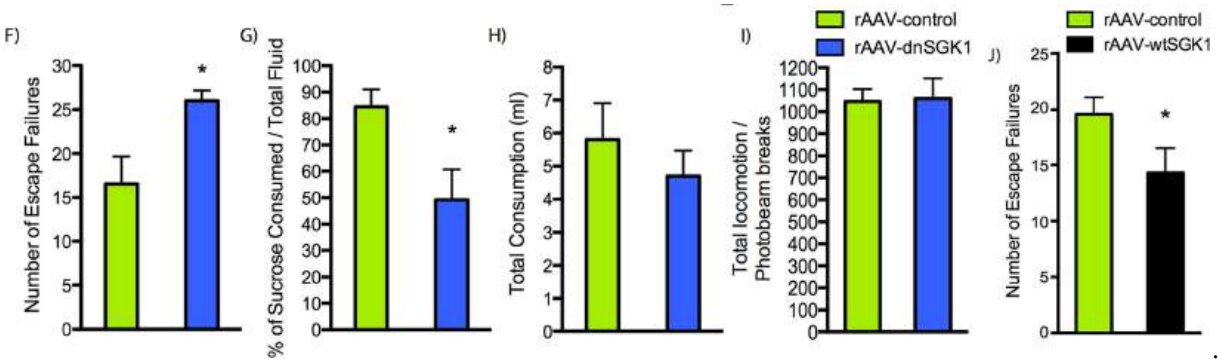
Panel (B): SGK1 and (C) PSD-95 protein levels in dissected PFC are decreased in the low-escape group. Data are mean ± SEM percent change over control group (naive, $n = 5$; high escape, $n = 5$; low escape, $n = 9$).

? Exercise 21

Interpret the results. Why did they run GAPDH (glyceraldehyde-3-phosphate dehydrogenase) on the gels? PSD-95 is the postsynaptic density protein-95 and GluR1 is the glutamate receptor gene product.

Answer

To determine the role of SGK1 in the behavior of live animals, investigators made a dominant negative mutated form of SGK1 (S422A) which was referred to as dnSGK1. The gene was added to an adeno-associated virus (AAV) to express the protein in live animals and then observed their behavior. Some results are shown below.



Effects of rAAV-dnSGK1 infusion into medial PFC on behavior were tested in (F) AA (escape failures), (G) sucrose preference, (H) total fluid consumption, and (I) locomotor activity. Data are shown as mean ± SEM (controls $n = 11$; dnSGK1 $n = 9$). (t(18) = 2.61 for AA and t(18) = 2.795 for SPT, Student's *t* test, *$p < 0.05$]. (J) AA (escape failures) for rAAV-wtSGK1 injected rats (controls $n = 10$; wtSGK1 $n = 10$). (t(18) = 1.933, one-tailed Student's *t* test, *$p < 0.05$.)

? Exercise 22

Answer these questions.

a. What structure/activity changes in SGK1 would you expect in the mutant?

b. Interpret the results

Answer

Now let's return to whole genome/transcriptome studies. To gain new insight, it would be useful to analyze a more specific set of cells to get clues about the effects of PTSD on their biochemistry. The first study mentioned above looked at "omics" analyses of whole blood, plasma, serum, and buffy-coats. A newer approach is to study specific cell types from people with and without PTSD. We've clearly showed that PTSD is a systems disease, but at its heart are neurological effects, so it makes sense to study neurons (not just post-mortem brain tissue or do whole brain scans). Instead of isolating neurons from people, investigators have made glutamineric neurons in culture from stem cells isolated from people so the effects of PTSD can be studied in "petri dishes".

How can you mimic PTSD conditions in culture? One way is to add glucocorticoids (like the drug dexamethasone or hydrocortisone) to the cells from donors with PTSD and as controls, those who don't. We've seen repeatedly in the above examples the effects of changes in SGK1, which as its name implies, is affected by glucocorticoids. Transcriptomic studies to determine changes in gene expression have been done to compare the following sets of cells and PTSD mimics.

1. **Peripheral blood mononuclear cells (PBMCs) from combat veterans with PTSD and controls without PTSD**

This study used **dexamethasone (DEX)** as the cell "stressor". To interpret the graphs you need to know some definitions which are likely unfamiliar to biochemistry students who haven't yet studied large scale omic analyses.

- **FDR** (**False Discovery Rate).** This is the expected # false predictions/ # total predictions, so it's a statistical measure of the likeliness that the results are **not** what is hypothesized (or the null hypothesis). It's like a p-value except for multiple tests. It is the rate at which features that are determined to be significant are in fact **not**. For example, an FDR of 0.05 means that 5% of all features determined to be significant are NOT (so 95 out of 100 are). Hence the lower the FDR, the high likely that the features are significant (in this case that they are associated with PTSD). The table below shows some examples of different FDR values. Often they are expressed as -log FDR(fx) so the larger this value, the more likely an outcome is significant.

FDR (fx)	FDR (% determined significant that aren't)	-log FDR (fx)
0.01	1%	2
0.05	5%	1.3
0.1	10%	1
0.25	25%	0.60

- **DEGS: differentially expressed genes**
- LogFC: differences observed between vehicle and hydrocortisone exposure (for example)

Results from the PBMC studies +/- DEX are shown in the figure below.

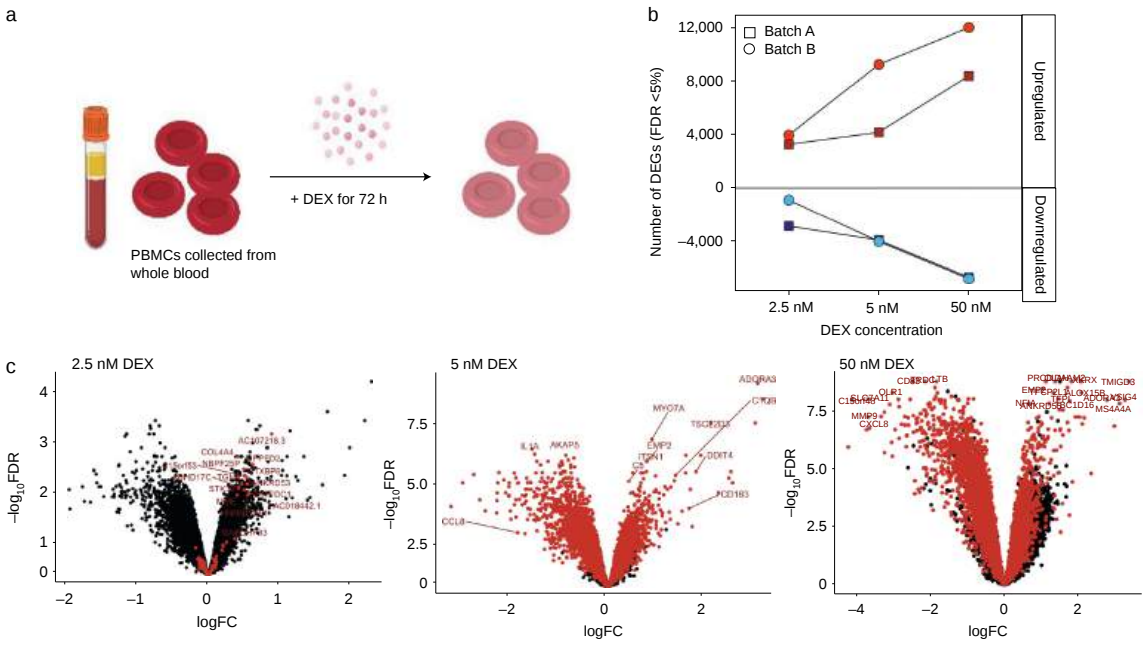


Fig. 1: Transcriptional response to DEX in PBMCs.

Panel a shows how PBMCs from 20 PTSD cases and 20 combat-exposed controls were treated with DEX for 72 h and RNA-seq was performed.

Pane b shows the number of **differentially expressed genes (DEGs)** observed in batch A (this study) and batch B (another published study) are upregulated and downregulated (*y* axis) across three different concentrations of DEX conditioning (*x* axis).

Panel c shows Meta-analysis of expression logFC (differences observed between vehicle and DEX exposure, x axis) was plotted against $-\log_{10}$(FDR) for each gene. Red points indicate significant DEGs in the meta-analysis. (Since the graph in the paper shows logFC and not $\log_{10}$FC, we will assume, as in our previous discussions, that logFC implies $\log_2$FC, which is customary.) Panel c is called a volcano (scatter) plot that shows the statistical significance ($-\log_{10}$FDR or otherwise P value) vs a measure of the magnitude of the change (here logFC)

? Exercise 23

Which part of Figure C, 50 nM DEX shows the genes are most statistically downregulated? Upregulated? Interpret the overall result

Answer

Changes were high in genes for immune signaling (this make sense since most nucleated cells in the blood are immune cells) and also in glucocorticoid sensitivity.

2. Neurons from combat veterans with PTSD and controls that without PTSD

Next investigators did the same experiments using glutaminergic neurons derived from pluripotent stem cells induced with neurogenin 2 (NGN2) to produce NGN2-neurons with glutamate receptors, glutamate transporters, and the neurotransmitter glutamate.

Results from the NGN2 neutrons +/- hydrocortisone (a glucocorticoid) are shown in the figure below.

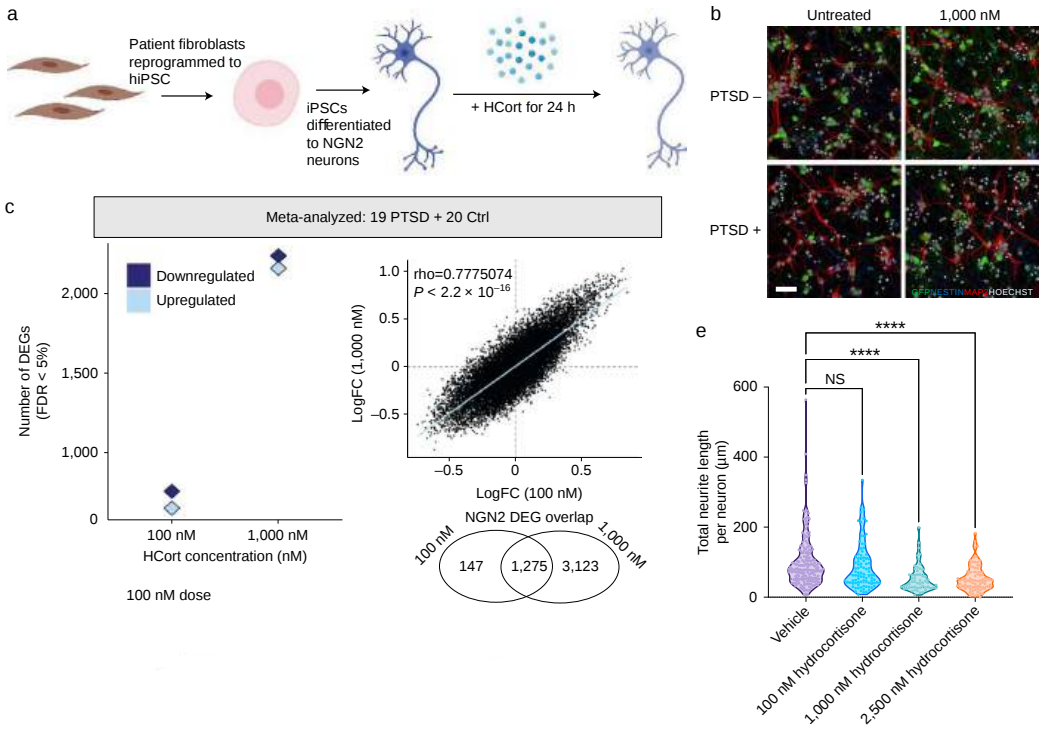


Fig. 2: Gene expression changes to HCort in hiPSC-derived neurons.

Panel a, hiPSC-derived *NGN2*-neurons were treated with HCort for 24 h and RNA-seq performed.

Panel b, *NGN2*-neurons stained for neuronal markers NESTIN and MAP2, nucleic marker HOECHST and green fluorescent protein to confirm neuronal identity and morphology across all conditions.

Panel c, Meta-analyzed DEGs in response to increasing concentrations of HCort shows robust changes in *NGN2*-neurons. A comparative analysis of transcriptome-wide $\log_2$FC in response to different concentrations of HCort in *NGN2*-neurons shows similar responses, indicating a conserved response across all donors to HCort in *NGN2*-neurons.

Panel e, Morphological analysis of neurite outgrowth in day 7 *NGN2*-neurons showing a dose-dependent decrease in neurite outgrowth with HCort exposure.

? Exercise 24

What conclusions can be derived from

- Panel B
- Panel C
- Panel E

Answer

Some genes/pathways that were underexpressed include acetylcholine signaling, ubiquitin ligation, specific cell type development and differentiation. Some that were overexpressed include protein methylation, immune cell function, histone modification, and regulation of gene expression.

3. Direct comparison of gene expression in neurons from active duty/veterans with +/- PTSD

Results from these studies are shown below.

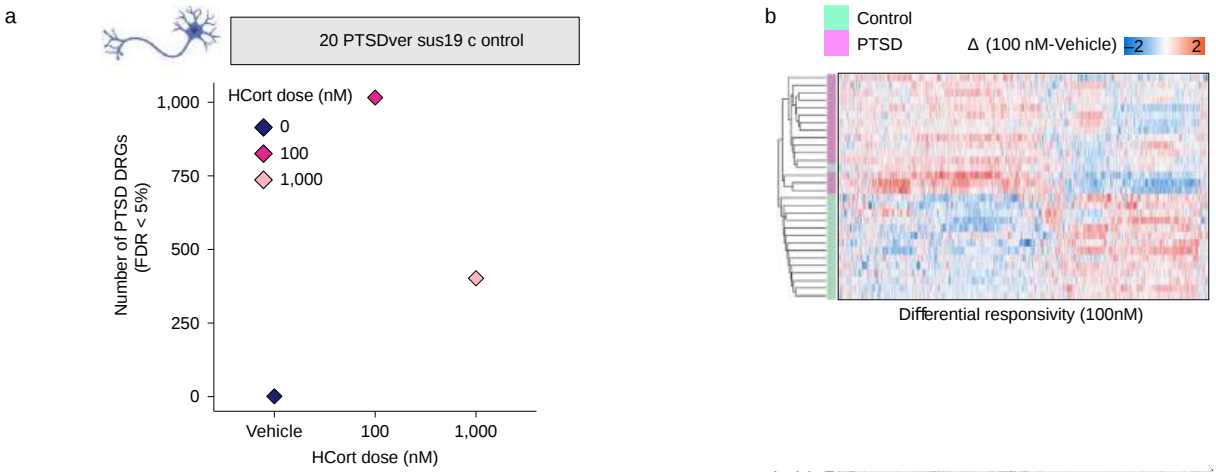


Fig. 4: PTSD(+) specific responses to HCort in *NGN2*-neurons.

Panel a shows genes that differ in their response to HCort in PTSD(+) donors compared with PTSD(–) donors, here termed DRGs, were detected in both the 100 nM and 1,000 nM dose, indicating PTSD diagnosis-specific responses to HCort.

Panel b, Significant *NGN2*-DRGs correctly classify PTSD(+) from PTSD(–) participants using an "unsupervised (machine learning) approach" (does not require selection of input parameters a prior. Just raw data is used).

? Exercise 25

Interpret Panels a and b.

Answer

Here is a summary from the paper:

"Both blood and neuronal glucocorticoid responses were significantly enriched for immune response genes; neuronal glucocorticoid responses were also associated with brain development and neurodevelopmental disorder genes. Although a PTSD diagnosis-specific signature was not detectable at baseline in either cell type, glucocorticoid hypersensitivity in PTSD was observed. These findings are consistent with the glucocorticoid hypersensitivity hypothesis; for example, patients with PTSD have altered blood sensitivity to glucocorticoids[6] and perturbations in glucocorticoid receptor signaling have been shown for PTSD in PM brain tissue[10].

Stress impacts the risk to psychiatric disorders throughout the lifespan—across prenatal development[62,63], childhood[52] and adulthood[52,64,65]. Moreover, the significant and positive relationship between our observed associations in *NGN2*-neurons and previously demonstrated transcriptional PTSD signatures in PM brains suggests that some impacts of glucocorticoid exposure may persist through to adulthood. Notably, glucocorticoid stimulation is not a specific model for PTSD; stress response is comorbid in many psychiatric disorders. HCort-responsive genes therefore likely represent aspects of stress response shared across PTSD and other neuropsychiatric disorders, such as the shared impact on social cognition between PTSD and ASD[66,67]. Combined analysis of context-dependent hiPSC models with cross-lifespan datasets, such as PM brains, may uncover long-term glucocorticoid-dependent PTSD signatures with which to refine hallmarks of PTSD susceptibility following combat exposure."

Other related studies:

- SGK1 knockdown in the medial prefrontal cortex reduces resistance to stress-induced memory impairment

- Glucocorticoids Can Induce PTSD-Like Memory Impairments in Mice

---

https://www.nature.com/articles/s41380-022-01498-7

Lee, B., Pothula, S., Wu, M. *et al*. Positive modulation of N-methyl-D-aspartate receptors in the mPFC reduces the spontaneous recovery of fear. *Mol Psychiatry* **27**, 2580–2589 (2022). https://doi.org/10.1038/s41380-022-01498-7. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Epigenetic effects of trauma

The studies - a note of caution

Epigenetic effects of trauma would have the potential to be passed on to offspring and could lead to multigenerational trauma effects. It is very difficult to conduct quality studies to show effects in subsequent generations. Subjects have to be identified and consent to participation in the study. Few studies have been done on the biochemical effects of childhood trauma on victims and intergenerational effects. Perhaps the best studies derive from war victims (soldiers, civilian populations, mostly from WWII) and their descendants. It is also difficult to disentangle possible epigenetic effects from environmental causes such as those that might arise from children brought up by a parent who has or has had PTSD. Studies have also been done after events such as 9/11 and even the Covid pandemic. Animal trauma models have also been explored and even those have ethical concerns.

Epigenetics - short review

Trauma might also be encoded by a myriad of epigenetic changes that may or may not be passed onto offspring. Figure 1 below shows different types of epigenetic changes in trauma victims.

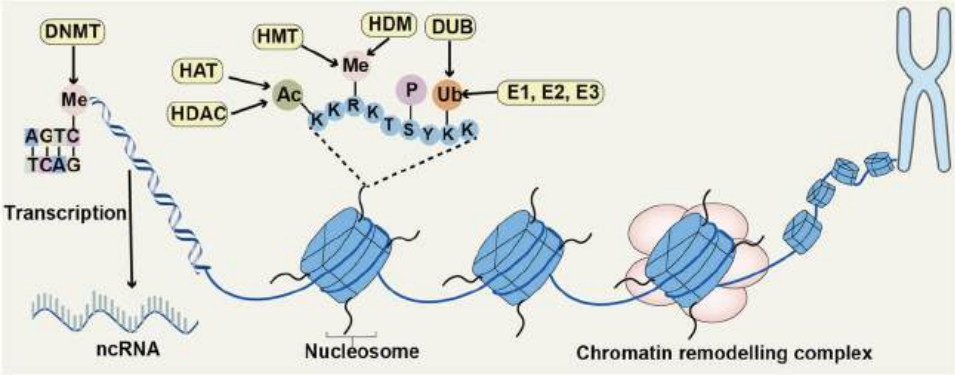


Figure 1: Four different epigenetic regulatory mechanisms. Wu, YL., Lin, ZJ., Li, CC. *et al*. Epigenetic regulation in metabolic diseases: mechanisms and advances in clinical study. *Sig Transduct Target Ther* **8**, 98 (2023). https://doi.org/10.1038/s41392-023-01333-7. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

The figure shows 4 kinds of epigenetic changes.

DNA methylation

This occurs on cytosine bases in CpG motifs (island). Methylation is abundant in telomers, centromers, repetitive sequences, and on the inactivated X chromosomes in females. It can affect genome stability and silencing (of expression). 5meC is the most common but other modifications (hydroyxmethy, formyl, and carboxyl) also occur at the 5' position of cytosine bases. The

methylation is catalyzed by several different DNA methyltransferases (DNMTs) using *S*-adenosylmethionine (SAM) as the methyl donor. DNMT1 maintains methylation states while DNMT3A and DNMT3B create new meCs.

Histone modifications

Two copies each of the positively charged histones H2A, H2B, H3, and H4 comprise the nucleosome core around which negatively charged dsDNA wraps. The histones can be post-translationally modified by acetylation, methylation, lactylation, phosphorylation, dopaminylation, and ubiquitination. Acetylation removes the + charge on histones and decreases DNA binding affinity. These modifications also add binding sites for proteins which affect the overall packing of chromatin. Overall, these modifications can change the accessibility of sites on chromatin required for gene transcription and hence they can regulate gene transcription.

Acetylations are catalyzed by histone acetyltransferases (HATs) and reversed by histone deacetylases (HDACs). Most modifications occur on the N-terminal sections of H3 and H4. Proteins with bromodomains act as histone acetylation "readers" and are involved in transcriptional regulation and chromatin remodeling. The N-termini of H3 and H4 are also the site of methylation of lysine or arginine residues, which can produce mono-, di-, or trimethylated lysine residues. Some methylations activate transcription, others inhibit it. The enzymes involved are lysine methyl transferase (KMTs) and lysine demethylases (KDMs), while protein arginine methyltransferases (PRMTs) modify arginines.

Chromatin remodeling

DNA wound around a nucleosome is not accessible to RNA polymerase (about the size of the nucleosome core) for transcription. Even worse for accessibility is condensed chromation. Remodeling of the chromatin, which requires ATP hydrolysis, can open it for transcription by increasing accessibility. Several families of remodelers (SWI/SNF, ISWI, chromodomain helicase DNA binding or CHD and inositol requiring 80 (INO80) can remodel chromatin.

Noncoding RNAs (ncRNAs)

Two types are described below

- endogenous miRNA of about 23 NT length can inhibit gene expression by inhibiting mRNA translation. RNA polymerase II is used in their synthesis and the inhibitor act to inhibit translation through the RNA-induced silencing complex (RISC) and binding to the 3′-untranslated regions (3′-UTRs) of specific mRNAs. The synthesis and hence expression of specific miRNA can hence be regulated by the epigenetic mechanism described above. More directly, but they can also regulate the key enzymes involved in epigenetic modifications including HDACs and DNMTs. diabetes.
- lncRNA are longer than 200 nucleotides and include long intergenic non-coding RNAs (lincRNAs), antisense RNA, and others. They can affect chromatin structure and even enzyme activity through specific binding interactions.

Here are some references that describe the epigenetic effects of trauma.

- A review of epigenetic contributions to post-traumatic stress disorder
- Epigenetic Modifications in Stress Response Genes Associated With Childhood Trauma
- Methyl-CpG binding protein 2 expression is associated with symptom severity in patients with PTSD in a sex-dependent manner
- Epigenetics of childhood trauma: Long-term sequelae and potential for treatment
- GABAergic mechanisms regulated by miR-33 encode state-dependent fear
- Epigenetic mechanisms in fear conditioning: Implications for treating post-traumatic stress disorder
- Intergenerational transmission of trauma effects: putative role of epigenetic mechanisms
- The public reception of putative epigenetic mechanisms in the transgenerational effects of trauma

CHAPTER OVERVIEW

iCn3D for Biomolecular Visualization Learning Themes and Goals

Topic hierarchy

Biomolviz- Constructing iCn3D Models to Target Biomolecular Visualization Skills
BioMolViz Framework
BioMolViz Theme: Alternate Renderings (AR)
BioMolViz Theme: Atomic Geometry (AG)
BioMolViz Theme: Construction and Annotation (CA)
BioMolViz Theme: Ligands and Modifications (LM)
BioMolViz Theme: Macromolecular Assemblies (MA)
BioMolViz Theme: Macromolecular Building Blocks (MB)
BioMolViz Theme: Molecular Dynamics (MD)
BioMolViz Theme: Molecular Interactions (MI)
BioMolViz Theme: Structural Model Skepticism (SK)
BioMolViz Theme: Structure-Function Relationship (SF)
BioMolViz Theme: Symmetry/ Asymmetry Recognition (SA)
BioMolViz Theme: Topology and Connectivity (TC)

Biomolviz- Constructing iCn3D Models to Target Biomolecular Visualization Skills

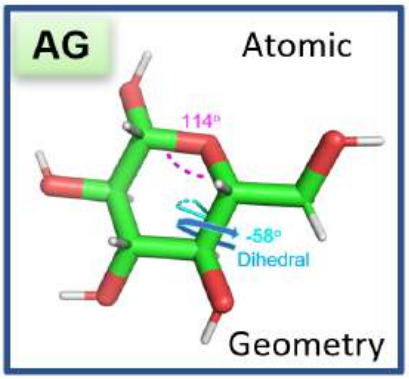
It is not only important to visualize pre-rendered models of biomolecules, but it is also important to be able to create them to address key aspects of structure and function. These efforts should be guided by a clear set of learning goals and objectives that target student understanding of structure and function. So it is fortunate that clear learning themes, goals and objectives are articulated in a Biomolecular Visualization Framework created by BioMolViz,

BIOMOLVIZ

Promoting Molecular Visualization Literacy

The BMV framework is used with permission from BioMolViz.Org

Copy the appropriate row when assigning a theme, goal, and objective to a designated iCn3D or other biomolecular visualization assessment

Atomic Geometry (AG)		Three-atom and four-atom (dihedral) angles, metal size and metal-ligand geometries, steric clashes

AG1. Students can describe the ideal geometry for a given atom within a molecule and deviations from the ideal geometry due to neighboring interactions.

AG1.01 Students can identify atomic geometry/hybridization for a given atom. (Novice)

AG1.02 Students can measure bond angles for a given atom. (Novice)

AG1.03 Students can identify deviations from the ideal bond angles. (Amateur)

AG1.04 Students can explain deviations from the ideal bond angles due to local effects. (Amateur, Expert)

AG1.05 Students can predict the effect of deviations from ideal bond angles on the structure and function of a macromolecule. (Expert)

AG1.06 Students can identify the geometric features of bonds (e.g., peptide bond, glycosidic, phosphoester).

AG2. Students can compare and contrast different structural conformations with regard to energy, the addition of substituents, and the impact on the structure/function of a macromolecule.

AG2.01 Students can describe different conformations that a structure can adopt using visualization tools. (Amateur)

AG2.02 Students can describe different conformations of atoms about a bond using visualization tools. (Novice)

AG2.03 Students can distinguish energetically favorable and unfavorable conformations that a structure can adopt. (Amateur)

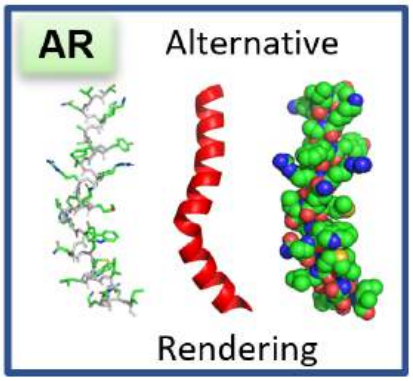
AG2.04 Students can predict the effect of a given substituent on the structure and function of a macromolecule (e.g., substituent on a carbohydrate/ligand, R groups/rotamers, phosphorylation, methylation of nucleic acids, post-translational modifications). (Expert)

AG3. Students can describe dihedral/torsion angles in biomolecules.

AG3.01 Students can identify a dihedral/torsion angle in a three-dimensional representation of a molecule. (Novice)

AG3.02 Students can identify the planes between which a dihedral/torsion angle exists within a three-dimensional representation of a macromolecule. (Novice)

AG3.03 Students can identify phi, psi, and omega torsion/dihedral angles in a three-dimensional representation of a protein. (Amateur)

Alternate Renderings (AR)		Rendering of a macromolecular structure such as a protein or nucleic acid structure in various ways from the simplest possible way (connections between alpha carbons) to illustration of secondary structure (ribbons) to surface rendering and space filling.

AR1. Students can interpret or create molecular images that convey features such as secondary structure, CPK coloring, and active sites.

AR1.01 Students can manipulate rendered structures to illustrate molecular properties. (Novice)

AR1.02 REMOVED (integrated with SF2.02)

AR1.03 Students can describe or label structural differences among multiple structures. (Amateur, Expert)

AR1.04 Students can infer information from rendering a structure in different ways. (Novice, Amateur, Expert)

AR1.05 Students can create renderings that distinguish secondary structural features. (Novice)

AR1.06 Students can create an information rich rendering of a structure that depicts structural features found in the literature. (Amateur)

AR1.07 Students can create an information rich rendering of a structure containing ligands, covalent modifications, and noncanonical amino acids or nucleotides. (Amateur, Expert)

AR1.08 Students can use molecular visualization to tell a story about a macromolecular structure. (Expert)

AR1.09 REMOVED (integrated with MI1.02)

AR1.10 Students can convert textbook images of small molecules into 3D representations in a molecular visualization program. (Amateur)

AR2. Students can choose the best rendering of a macromolecule to use in a given situation.

AR2.01 Students can recognize that a cartoon rendering is a summary of the detail in a line rendering. (Novice, Amateur)

AR2.02 Students can describe the atoms and their representations in different renderings (e.g., coloring, showing hydrogens/double bonds). (Novice)

AR2.03 Students can identify or create a suitable rendering, or combination of renderings, for a specific purpose (e.g., a surface rendering overlaid with a cartoon to highlight the van der Waals surface alongside secondary structure, or active site sticks shown over a cartoon). (Novice, Amateur)

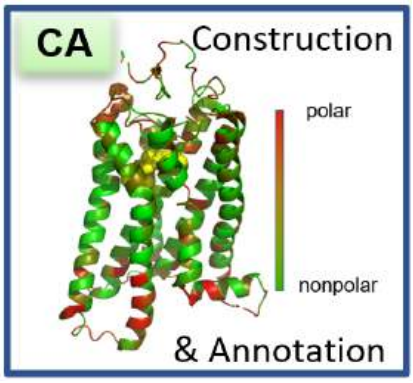
AR2.04 Students can identify the limitations in various renderings of molecular structures. (Amateur)

AR2.05 Students can understand the level of detail of different molecular representations. (Novice, Amateur, Expert)

AR2.06 Students can transition comfortably between equivalent 2D and 3D renderings of biomolecules. (Novice, Amateur, Expert)

AR2.07 Students can use and interpret color in the context of macromolecules to clarify and/or highlight features (e.g., coloring amino acids differently by property, different molecules uniquely in a complex, protein chains, secondary structure).

(Novice)

Construction and Annotation (CA)		Ability to build macromolecular models, either physical or computerized, and, where possible, add commentary, either written or verbal, to tell a molecular story.

CA1. Students can compose information-rich renderings of macromolecule-ligand interactions.

CA1.01 Students can construct and annotate a model of a macromolecule bound to a ligand. (Amateur)

CA1.02 Students can construct a model of a macromolecule bound to a ligand and identify the types of molecular interactions. (Amateur)

CA1.03 Students can construct a model of a macromolecule bound to a ligand and assess the importance of molecular interactions. (Expert)

CA1.04 Students can produce a model of a macromolecule based on a known structure of a related macromolecule. (Amateur, Expert)

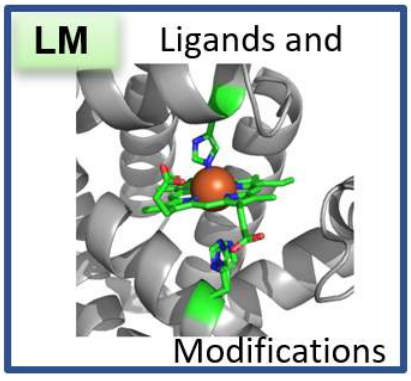
CA2. Students can compose a rendering to predict the cellular location of a protein (e.g., extracellular, membrane associated, or cytoplasmic) based on the properties and orientations of functional groups.

CA2.01 Students can design a rendering that conveys properties such as polarity, charge, secondary structure, etc. to suggest the cellular location of a macromolecule. (Amateur)

CA2.02 Students can create protein images with colored polar/nonpolar residues to determine whether they fold with a hydrophobic core. (Amateur)

CA2.03 Students can create images to display polar/nonpolar residues and propose a role for the protein and/or how it interacts with its environment - and that the predictions would be plausible based on the protein. (Amateur)

CA2.04 Students can make accurate predictions of the location/function of the protein that incorporates additional protein features, such as transmembrane helices, apparent docking surfaces, etc. (Expert)

Ligands and Modifications (LM)		Metals and metal clusters, additions such as glycosylation, phosphorylation, lipid attachment, methylation etc.

LM1. Students can identify ligands and modified building blocks (e.g., hydroxyproline, aminosaccharides, modified nucleobase) within a rendered structure.

LM1.01 Students can use the annotation associated with a pdb file to identify and locate ligands and modified building blocks in a given biomolecule. (Amateur)

LM1.02 Students can visually identify non-protein chemical components in a given rendered structure. (Amateur)

LM1.03 Students can distinguish between nucleic acid and ligands (e.g., metal ions) in a given nucleic acid superstructure. (Amateur)

LM1.04 Students can explain how a ligand in a given rendered structure associates with the biomolecule (e.g., covalent interaction with residue X). (Amateur)

LM1.05 Students can locate/identify ligands and modified building blocks in unannotated structures and describe their role. (Expert)

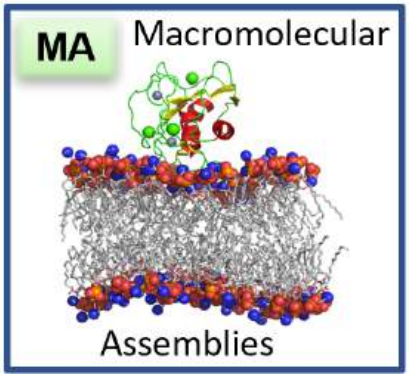
LM2. Students can describe the impact of a ligand or modified building block on the structure/function of a macromolecule.

LM2.01 Students can look at a given rendered structure and describe how the presence of a specific ligand or modified building block alters the structure of that biomolecule. (Amateur)

LM2.02 Students can explain how the removal of a particular ligand or modified building block would alter the structure of a given biomolecule. (Expert)

LM2.03 Students can use molecular visualization tools to predict how a specified ligand or modified building block contributes to the function of a given protein. (Amateur, Expert)

LM2.04 Students can predict how a ligand or modified building block contributes to the function of a protein for which the structure has been newly solved. (Expert)

Macromolecular Assemblies (MA)		Polypeptides, oligosaccharides, and nucleic acid and lipid superstructures (e.g. protein–nucleic acid complexes, lipid membrane-associated proteins)

MA1. Students can describe various macromolecular assemblies.

MA1.01 Students can identify individual biomolecules in a macromolecular assembly. (Novice, Amateur, Expert)

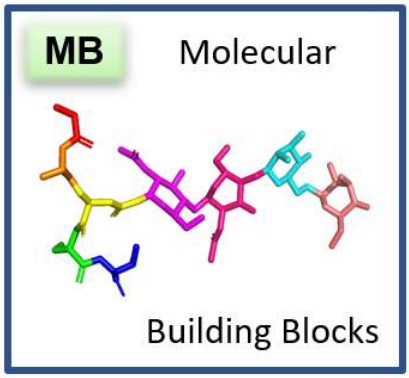
MA1.02 Students can describe functions of individual biomolecules within a macromolecular assembly. (Novice, Amateur, Expert)

MA1.03 Students recognize the various lipid ultrastructures (e.g., micelles, bicelles, vesicles, and lipid bilayers) in a 3D structure. (Novice)

MA2. Students can compose information-rich renderings of macromolecular assemblies.

MA2.01 Students can render a macromolecular assembly to highlight individual structures. (Amateur)

MA2.02 Students can render a macromolecular assembly to illustrate structural features (e.g., binding interfaces, symmetry, tertiary structure, etc.). (Novice, Amateur, Expert)

Macromolecular Building Blocks (MB)		Recognition of native amino acids, nucleotides, sugars, and other biomonomer units/building blocks. Understanding of their physical and chemical properties, particularly regarding functional groups.

MB1. Students can identify individual building blocks of biological polymers.

MB1.01 Given a rendered structure of a biological polymer, students can identify the ends of a biological polymer. (Novice, Amateur, Expert)

MB1.02 Given a rendered structure, students can divide the polymer into its individual building blocks. (Novice)

MB1.03 Given a rendered structure, students can identify the individual building blocks. (Novice)

MB2. Students can describe the contributions different individual building blocks make in determining the 3-D shape of the polymer.

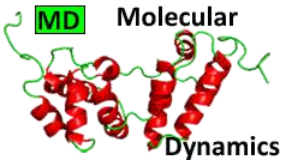
MB2.01 Students can describe the physical/chemical properties of an individual building block/functional group in a rendered structure of a polymer. (Amateur)

MB2.02 Students can describe the significance of the location of individual building blocks within the 3D structure of a polymer (protein, carbohydrate, or nucleic acid). (Novice, Amateur, Expert)

MB2.03 Students can identify physical/chemical properties of individual building blocks/functional groups in different local environments. (Amateur)

MB2.04 Using a visualized structure, students can identify stereochemistry (e.g., in carbohydrate, lipid, and protein structures). (Amateur)

MB2.05 Students can modify/mutate a building block to change the 3D structure of a polymer (protein, carbohydrate, or nucleic acid). (Amateur, Expert)

Molecular Dynamics (MD)		Animated motion simulating conformational changes involved in ligand binding or catalysis, or other molecular motion/dynamics.

MD1. Students can describe the impact of the dynamic motion of a biomolecule on its function.

MD1.01 Students can recognize that biological molecules have different conformations. (Novice, Amateur)

MD1.02 Students can correlate molecular movement with function. (Novice, Amateur, Expert)

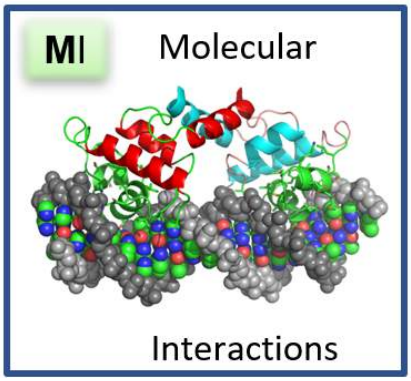
MD2. Students can predict limits to macromolecular movement.

MD2.01 Students can locate potential regions of flexibility and inflexibility in the structure of a biomolecule. (Novice, Amateur)

MD2.02 Students can recognize acceptable/unacceptable movement within a macromolecule by determining whether the movement is within allowable bond angles. (Expert)

MD2.03 Students can recognize acceptable/unacceptable movement within a macromolecule by determining whether the movement results in steric hindrance. (Amateur)

MD2.04 Students can recognize acceptable/unacceptable movement within a macromolecule by considering the atomic packing constraints. (Expert)

Molecular Interactions (MI)		Covalent and noncovalent bonding governing ligand binding and subunit-subunit interactions.

MI1. Students can predict the existence of an interaction using structural and environmental information (e.g. bond lengths, charges, pH, dielectric constant).

MI1.01 Students can distinguish between covalent and noncovalent interactions. (Novice)

MI1.02 Students can identify different noncovalent interactions (e.g., hydrogen bonds, ionic interactions, van der Waals contacts, induced dipole) given a 3D structure. (Amateur)

MI1.03 Students can predict whether a functional group (region) would be a hydrogen bond donor or acceptor. (Amateur)

MI1.04 Students can render the 3D structure of a biomolecule so as to demonstrate the ionic interactions and/or charge distribution of the different non-covalent interactions. (Amateur)

MI1.05 As it relates to a particular rendered structure, students can rank the relative strengths of covalent and noncovalent interactions. (Amateur)

MI2. Students can evaluate the effect of the local environment on various molecular interactions.

MI2.01 Students can identify regions of a biomolecule that are exposed to or shielded from solvent. (Novice)

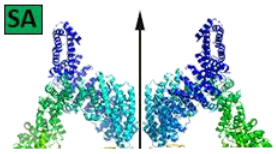
MI2.02 Students can identify other molecules in the local environment (e.g., solvent, salt ions, metals, detergents, other small molecules) that impact a molecular interaction of interest. (Novice)

MI2.03 Students can predict the impact of other molecules in the local environment (e.g., solvent, salt ions, metals, detergents, other small molecules) on a molecular interaction of interest. (Amateur)

MI2.04 Students can predict the pKa of an ionizable group based on the influence of its local three-dimensional environment. (Amateur)

MI2.05 Students can propose a change to the local environment that would yield a desired change in a molecular interaction. (Expert)

MI2.06 Using molecular visualization tools, students can determine which intermolecular force is most critical to stabilizing a given interaction. (Expert)

Symmetry/ Asymmetry Recognition (SA)		Recognition of symmetry elements within both single chain and multi-chain macromolecules.

SA1. Students can identify symmetric or asymmetric features in rendered molecules.

SA1.01 Students can identify symmetric features in a rendered molecule (shown in fixed orientation). (Novice)

SA1.02 Students can rotate a single macromolecule, multi-chain macromolecules (e.g., homo- or heteromers), complexes of macromolecules, and supramolecular assemblies to identify axes of symmetry. (Amateur)

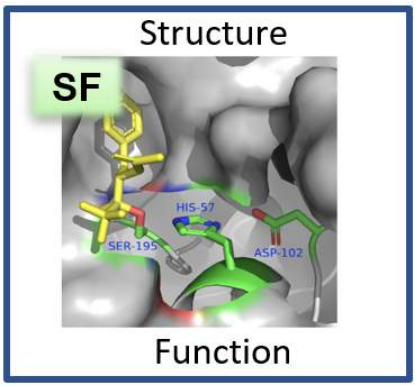
SA1.03 Students can identify symmetric and asymmetric features in rendered molecules after coloring a given rendered molecule to reveal structural features (charge, hydrophobicity, etc.). (Amateur)

SA2. Students can hypothesize the functional significance of symmetry or asymmetry in rendered molecules.

SA2.01 Students can explain the functional significance of rotational axes of symmetry (or asymmetry) in a given rendered molecule. (Novice, Amateur, Expert)

SA2.02 Students can predict functional significance of symmetry (or asymmetry) in a given rendered molecule. (Amateur, Expert)

Structure-Function Relationship (SF)



Active/binding sites, microenvironments, nucleophiles, redox centers, etc. (please also see LM2.03)

SF1. Students can evaluate biomolecular interaction sites using molecular visualization tools.

SF1.01 Students can identify functionally relevant cofactors, ligands or substrates associated with a macromolecule and describe their role (e.g., an active site magnesium ion). (Amateur, Expert)

SF1.02 Students recognize that the size and shape of the ligand must match the size and shape of the binding site. (Novice, Amateur)

SF1.03 Students recognize that the polarity or electrostatic potential of a surface complements that of the ligand or substrate. (Novice, Amateur)

SF1.04 Students recognize that the hydrophobicity of a surface complements that of the ligand or substrate. (Novice, Amateur)

SF1.05 REMOVED (integrated with SF1.03)

SF1.06 Students can use docking software to predict how the surface properties of a macromolecule guide and allow the binding of a ligand or substrate. (Amateur)

SF2. Using molecular visualization, students can predict the function of biomolecules.

SF2.01 Students can recognize structurally related molecules. (Novice)

SF2.02 Students can superimpose structurally related molecules. (Novice, Amateur)

SF2.03 Students can identify functionally relevant features of a macromolecule (e.g., an active site cysteine, a functional loop). (Amateur)

SF2.04 Students can predict molecular function given a binding site. (Amateur, Expert)

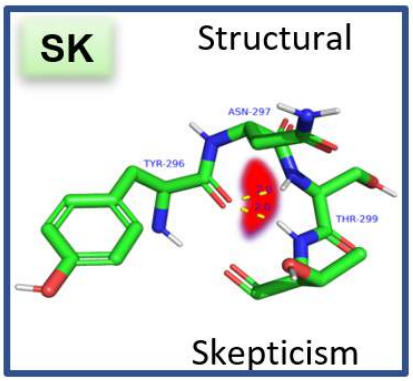
SF3. Using molecular visualization, students can predict the function of an altered macromolecule.

SF3.01 Students can structurally alter a macromolecule. (Novice)

SF3.02 Students can propose structural alterations to test interactions in a macromolecule. (Amateur)

SF3.03 Students can predict the impact of a structural alteration on the function of a macromolecule. (Amateur, Expert)

Structural Model Skepticism (SK)



Recognition of the limitations of models to describe the structure of macromolecules.

SK1. Students can critique the limitations of a structural model of a macromolecule.

SK1.01 Students can explain that the pdb file is a model based on data and that, as a model, it has limitations. (Novice, Amateur)

SK1.02 Students associate resolution with reliability of atom positions. (Amateur)

SK1.03 Students can identify building blocks (for example, amino acid side chains) whose orientation in a biopolymer is uncertain. (Expert)

SK1.04 Students can evaluate the flexibility/disorder of various regions of a macromolecular structure. (Novice, Amateur, Expert)

SK1.05 Students can reconcile inconsistent numbering of individual building blocks among species and structure files. (Novice)

SK1.06 Students can utilize a Ramachandran plot/steric clashes to interpret the validity of a structure. (Amateur, Expert)

SK1.07 Students can describe the limitations of a macromolecule-ligand docking simulation. (Amateur, Expert)

SK2. Students can evaluate the quality of 3D models including features that are open to alternate interpretations based on molecular visualization and PDB flat files.

SK2.01 Students can evaluate a crystal structure for crystal packing effects. (Novice, Amateur, Expert)

SK2.02 Students can resolve differences between the asymmetric unit and the functional biological assembly. (Expert)

SK2.03 Students can differentiate functional ligands (with biological/biochemistry role) from nonfunctional ligands (most solvents, salts, ions, and crystallization agents). (Novice, Amateur, Expert)

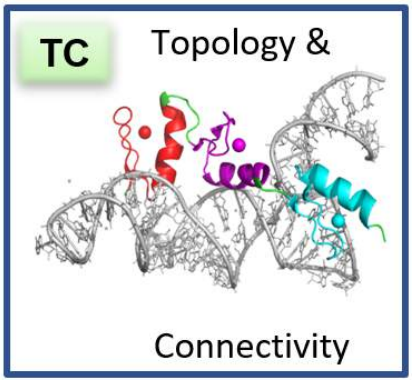
SK3. Students can discuss the value of experimentally altering a biomolecule to facilitate structure determination.

SK3.01 Students can identify non-native structural features. (Amateur)

SK3.02 Students can propose molecular modifications to facilitate structure determination. (Amateur, Expert)

SK3.03 Students can propose a purpose for the introduction of non-native structural features to facilitate structure determination. (Amateur, Expert)

Topology and Connectivity (TC)



Following the chain direction through the molecule, translating between 2D topology mapping and 3D rendering.

TC1. Students can describe or illustrate the linkages between building blocks within a macromolecule.

TC1.01 Students can trace the backbone of a macromolecule in three dimensions. (Novice, Amateur)

TC1.02 Students can use appropriate terms to describe the linkages/bonds/interactions that join individual building blocks together in a macromolecule or macromolecular assembly. (Novice, Amateur)

TC1.03 Given a virtual model of individual building blocks, students can predict the types of linkages/bonds/interactions that are possible or favorable. (Amateur)

TC1.04 Given individual building blocks, students can appropriately connect them to create a biological polymer (e.g., drawing carbohydrate linkages, a small peptide). (Amateur, Expert)

TC2. Students can describe the overall shape and common motifs within a 3D macromolecular structure.

TC2.01 Using molecular visualization software, students can describe the three-dimensional structure of a macromolecule, including overall shape and common structural motifs. (Novice, Amateur, Expert)

TC2.02 Students can identify common domains/motifs within a macromolecule. (Amateur, Expert)

TC2.03 Students can identify connectivity features between domains or subunits in a macromolecular structure. (Amateur)

TC2.04 Students can identify interactions between domains or subunits in a macromolecular structure. (Amateur, Expert)

TC2.05 Students can describe how domains/motifs in a macromolecule work together to achieve a concerted function in the cell. (Amateur, Expert)

TC2.06 Students can identify the levels of protein structure (e.g., parse a tertiary/quaternary structure into a series of secondary structures/motifs) and the ways in which they are connected from a three-dimensional structure. (Novice, Amateur, Expert)

TC3. Students can explain how any given biomolecular interaction site can be made by a variety of topologies.

TC3.01 Students can recognize that the groups that comprise a functional site only require proper arrangement in three-dimensional space rather than a particular order or position in the linear sequence. (Amateur)

TC3.02 Students can recognize similarities and differences in two similar - but not identical - three dimensional structures. (Amateur)

TC3.03 Students can describe dissimilar portions of homologous proteins as arising from genetic insertions/deletions/rearrangements. (Amateur)

BioMolViz Theme: Alternate Renderings (AR)

It is not only important to visualize pre-rendered models of biomolecules, but it is also important to be able to create them to address key aspects of structure and function. These efforts should be guided by a clear set of learning goals and objectives that target student understanding of structure and function. So it is fortunate that clear learning themes, goals and objectives are articulated in a Biomolecular Visualization Framework created by BioMolViz,

BioMolViz Theme: Atomic Geometry (AG)

It is not only important to visualize pre-rendered models of biomolecules, but it is also important to be able to create them to address key aspects of structure and function. These efforts should be guided by a clear set of learning goals and objectives that target student understanding of structure and function. So it is fortunate that clear learning themes, goals and objectives are articulated in a Biomolecular Visualization Framework created by BioMolViz,

BioMolViz Theme: Construction and Annotation (CA)

It is not only important to visualize pre-rendered models of biomolecules, but it is also important to be able to create them to address key aspects of structure and function. These efforts should be guided by a clear set of learning goals and objectives that target student understanding of structure and function. So it is fortunate that clear learning themes, goals and objectives are articulated in a Biomolecular Visualization Framework created by BioMolViz,

BioMolViz Theme: Ligands and Modifications (LM)

It is not only important to visualize pre-rendered models of biomolecules, but it is also important to be able to create them to address key aspects of structure and function. These efforts should be guided by a clear set of learning goals and objectives that target student understanding of structure and function. So it is fortunate that clear learning themes, goals and objectives are articulated in a Biomolecular Visualization Framework created by BioMolViz,

BioMolViz Theme: Macromolecular Assemblies (MA)

It is not only important to visualize pre-rendered models of biomolecules, but it is also important to be able to create them to address key aspects of structure and function. These efforts should be guided by a clear set of learning goals and objectives that target student understanding of structure and function. So it is fortunate that clear learning themes, goals and objectives are articulated in a Biomolecular Visualization Framework created by BioMolViz,

BioMolViz Theme: Macromolecular Building Blocks (MB)

It is not only important to visualize pre-rendered models of biomolecules, but it is also important to be able to create them to address key aspects of structure and function. These efforts should be guided by a clear set of learning goals and objectives that target student understanding of structure and function. So it is fortunate that clear learning themes, goals and objectives are articulated in a Biomolecular Visualization Framework created by BioMolViz,

BioMolViz Theme: Molecular Dynamics (MD)

It is not only important to visualize pre-rendered models of biomolecules, but it is also important to be able to create them to address key aspects of structure and function. These efforts should be guided by a clear set of learning goals and objectives that target student understanding of structure and function. So it is fortunate that clear learning themes, goals and objectives are articulated in a Biomolecular Visualization Framework created by BioMolViz,

BioMolViz Theme: Molecular Interactions (MI)

It is not only important to visualize pre-rendered models of biomolecules, but it is also important to be able to create them to address key aspects of structure and function. These efforts should be guided by a clear set of learning goals and objectives that target student understanding of structure and function. So it is fortunate that clear learning themes, goals and objectives are articulated in a Biomolecular Visualization Framework created by BioMolViz,

BioMolViz Theme: Structural Model Skepticism (SK)

It is not only important to visualize pre-rendered models of biomolecules, but it is also important to be able to create them to address key aspects of structure and function. These efforts should be guided by a clear set of learning goals and objectives that target student understanding of structure and function. So it is fortunate that clear learning themes, goals and objectives are articulated in a Biomolecular Visualization Framework created by BioMolViz,

BioMolViz Theme: Structure-Function Relationship (SF)

It is not only important to visualize pre-rendered models of biomolecules, but it is also important to be able to create them to address key aspects of structure and function. These efforts should be guided by a clear set of learning goals and objectives that target student understanding of structure and function. So it is fortunate that clear learning themes, goals and objectives are articulated in a Biomolecular Visualization Framework created by BioMolViz,

BioMolViz Theme: Symmetry/ Asymmetry Recognition (SA)

It is not only important to visualize pre-rendered models of biomolecules, but it is also important to be able to create them to address key aspects of structure and function. These efforts should be guided by a clear set of learning goals and objectives that target student understanding of structure and function. So it is fortunate that clear learning themes, goals and objectives are articulated in a Biomolecular Visualization Framework created by BioMolViz,

BioMolViz Theme: Topology and Connectivity (TC)

It is not only important to visualize pre-rendered models of biomolecules, but it is also important to be able to create them to address key aspects of structure and function. These efforts should be guided by a clear set of learning goals and objectives that target student understanding of structure and function. So it is fortunate that clear learning themes, goals and objectives are articulated in a Biomolecular Visualization Framework created by BioMolViz,

CHAPTER OVERVIEW

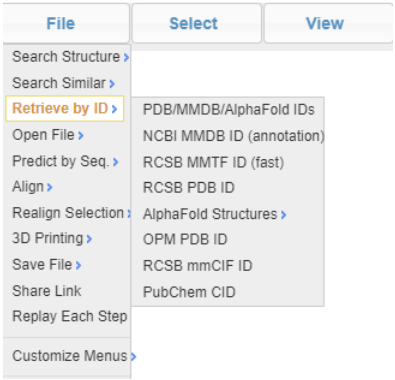
iCn3D Molecular Modeling Tutorials

Topic hierarchy

iCn3D Basics: File Types
iCn3D Basics: Saving Files
iCn3D Basics: Analysis Menu
iCn3D Basics: Mouse Commands
iCn3D Basics: Selecting and Viewing with a mouse
iCn3D Basics: Style and Color
iCn3D Intro Tutorial A: Modeling a Short Peptide in a Protein
iCn3D Intro Tutorial B: Rendering a Protein
iCn3D Intro Tutorial C: Finding Pockets in Proteins
iCn3D Intro Tutorial D: Modeling Psychoactive Drugs in Target Proteins
iCn3D Skill: Alternative Rendering and Saving Files
iCn3D Skill: Displays surface of a protein - Superoxide Dismutase
iCn3D Skill: Aligning two structures
iCn3D Skill: Analysis of Noncovalent Interactions
iCn3D Skill: Creating and Saving Selections
iCn3D Skill: Mutations
iCn3D Skill: Selection through Sequence and Annotations
iCn3D Skill: Showing a Protein-Protein Interface
iCn3D Tutorial: Binding interactions of SARS-Cov-2 Spike receptor domain with ACE2
iCn3D Tutorial: Overlay many ACE2 receptor binding domain analogous to SARS-Co2 RBD using Blast
iCn3D Tutorial: Protein Kinase B (AKT)
iCn3D Tutorial Question: Intermediate Problem - Cyclooxygenase II

iCn3D Basics: File Types

File, Retrieve by ID



MMDB

- MMDB (Molecular Modeling Database) files from the NCBI
- derived from PDB atomic coordinates but with …
- Database information (quaternary struct, molecular interactions, SNPs, conserved domains, clinical variants – i.e related structure info, not just xyz coord

PDB

- xyz coordinates

RCSB MMTF ID (fast)

- Great for very big structures that otherwise too slow in loading; Few modeling options.

AlphaFold Structure

- Computationally determined structures
- Uses Uniprot or RCSB ID.

OPM PDB ID

- Get structures on membrane proteins

Other Membrane Protein Links

•Eukaryotic peripheral membrane proteins (PerMemDB)

•Orientation of proteins in membranes (OPM)

•PDBTM: Protein Data Bank of Transmembrane Proteins

•mpstruct: Membrane Proteins of Known Structure

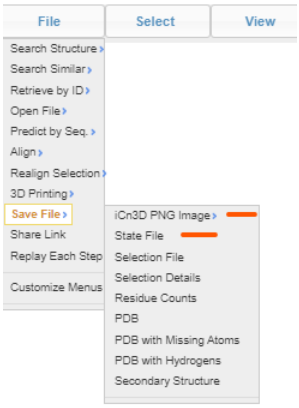
•Membranome: Membrane proteome of single helix membrane proteins

•MemProtMD: A database of membrane proteins embedded in lipid bilayers with lipids obtained in Molecular Dynamics simulation

iCn3D Basics: Saving Files

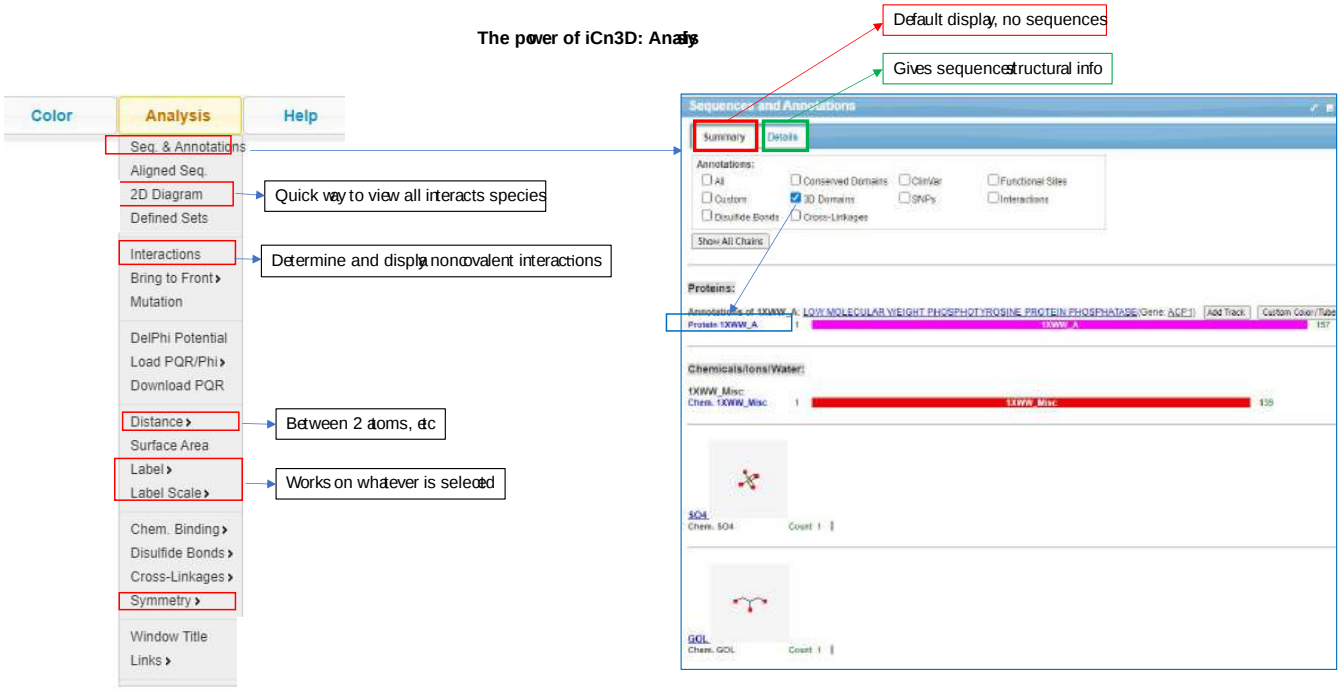
File, Save File

- In a new iCn3D window choose **Open File, iCn3D PNG image** and see the same file you started with.
- Likewise, in a new iCn3D window choose **Open File, State/Script File** and see the same file you started with.
- They can be sent to others to open as well



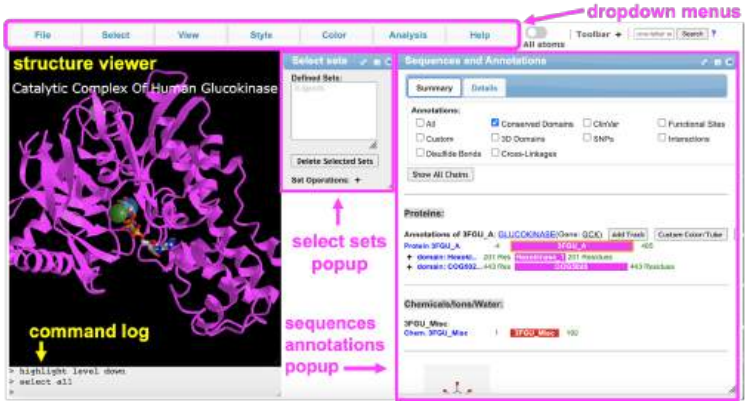
iCn3D Basics: Analysis Menu

Analysis Menu



iCn3D Basics: Mouse Commands

iCn3D modeling screen



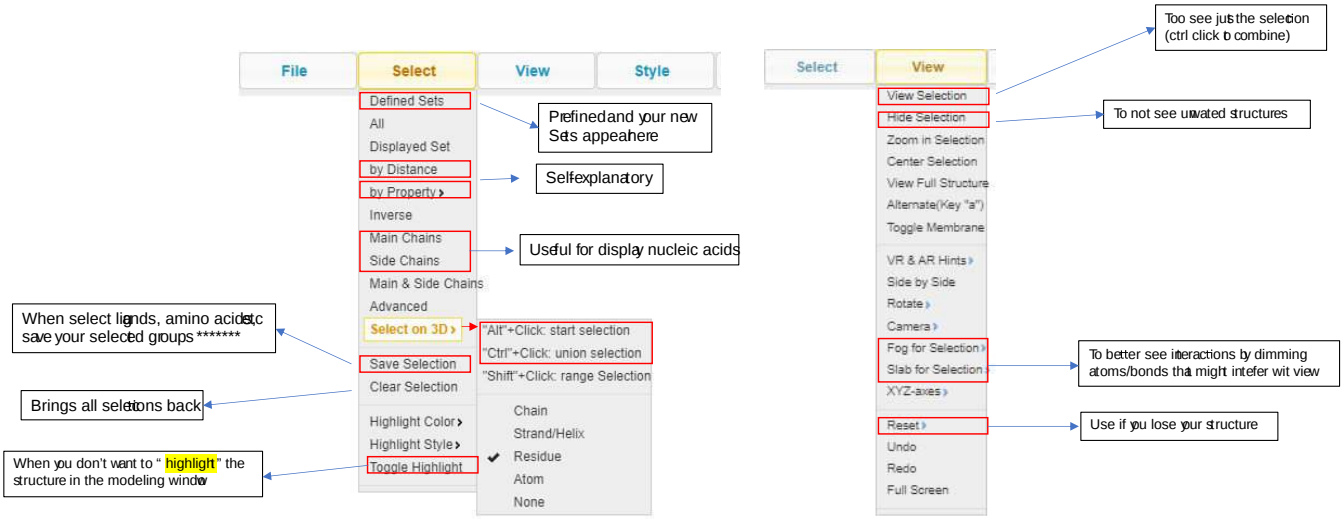
Mouse commands

rotate: click and drag (*mouse*: left click and drag; *keyboard*: j, i, l, and m keys)
zoom: pinch and spread (*mouse*: rotate the scroll wheel; *keyboard*: x and z keys)
translate: two finger click and drag (*mouse*: right click and drag)
Re-center: left click View from the top menu bar, then select "Center Selection"

Note: ctrl click on a PC = command click on Mac
alt click on PC = option click on Mac

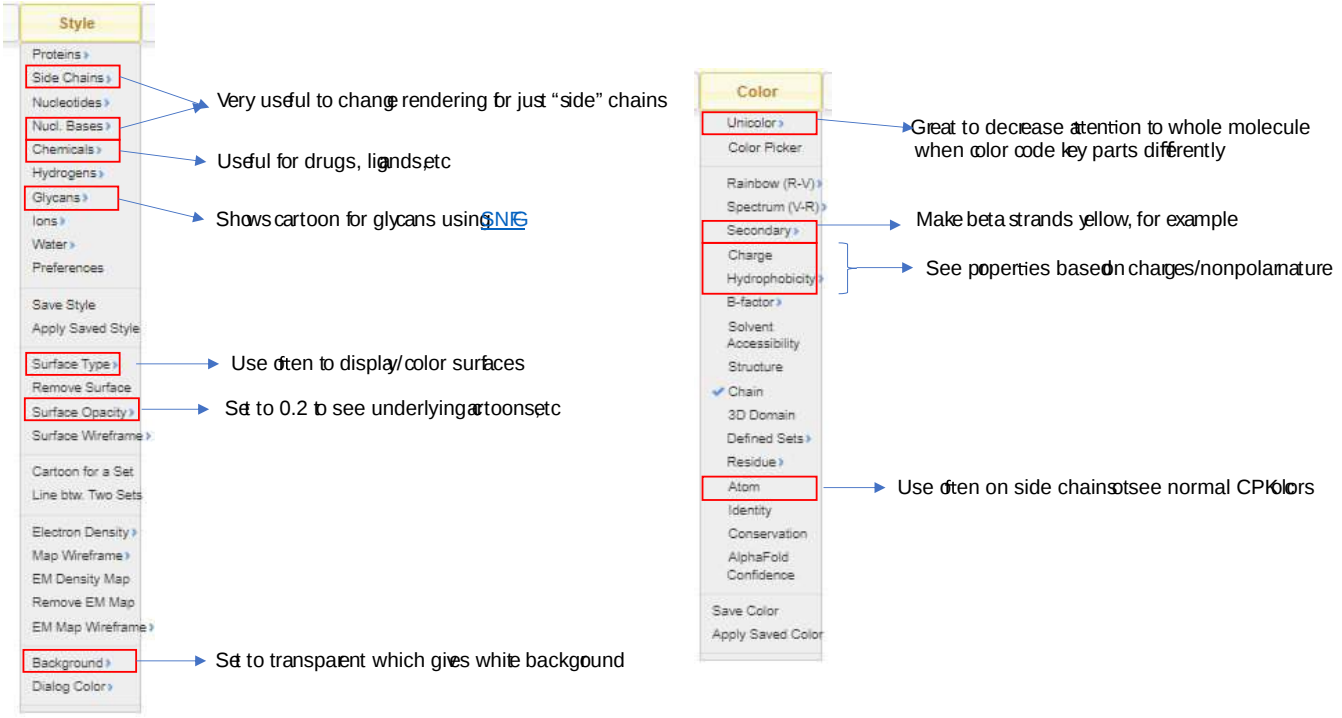
iCn3D Basics: Selecting and Viewing with a mouse

Selecting with the mouse (left) and viewing selection (right)



iCn3D Basics: Style and Color

Style and Color



iCn3D Intro Tutorial A: Modeling a Short Peptide in a Protein

A. Modeling short sections of a protein chain

Pick one of the small protein fragments below for modeling using iCn3D

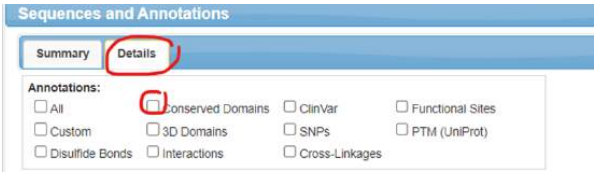
PDB	Description of protein (all small fragments)
2YW8	Crystal structure of human RUN and FYVE domain-containing protein
6EEY	human Scribble PDZ4 R1110G Mutant
2PA1	PDZ domain of human PDLIM2 bound to a C-terminal extension from human beta-tropomyosin
3A03	Hox11L1 homeodomain
3IWL	cisplatin bound to a human copper chaperone (monomer)
5Z2S	DUX4-HD2 domain
6L1C	PHF20L1 Tudor1 Y24L mutant
3D2N	MBNL1 tandem zinc finger 1 and 2 domain
3RD2	NIP45 SUMO-like Domain 2
7NZC	SH3 domain of POSH (Plenty of SH3 Domains protein)
1I2T	HUMAN HYPERPLASTIC DISCS PROTEIN: AN ORTHOLOG OF THE C-TERMINAL DOMAIN OF POLY(A)-BINDING PROTEIN
1NTE	CRYSTAL STRUCTURE ANALYSIS OF THE SECOND PDZ DOMAIN OF SYNTENI
2Y9U	Structural basis of p63a SAM domain mutants involved in AEC syndrome
2FMA	Alzheimer's Amyloid Precursor Protein (APP) Copper Binding Domain in 'small unit cell' form, atomic resolution
4OU0	Crystal Structure of RPA32C
1ZT3	C-terminal domain of Insulin-like Growth Factor Binding Protein-1 isolated from human amniotic fluid
2E3H	Crystal structure of the CLIP-170 CAP-Gly domain 2
1L9L	GRANULYSIN FROM HUMAN CYTOLYTIC T LYMPHOCYTES
5EFM	Beclin 1 Flexible-helical Domian (FHD) (141-171)

2BZX	Atomic model of CrkL-SH3C monomer
1NHL	SNAP-23N Structure
7UW7	Crystal structure of the Human TRIP12 WWE domain (isoform 2) in complex with ADP
4N7F	3rd WW domain of human Nedd4-1
2Q9V	C890S mutant of the 4th PDZ domain of human membrane-associated guanylate kinase
6T9Q	second, C-terminal repeat of the DNA-binding domain of human TImeless
1WVN	domain 3 of human alpha polyC binding protein
3I8Z	human chromobox homolog 4 (CBX4)
2F60	Dihydrolipoamide Dehydrogenase (E3)-Binding Domain of Human E3-Binding Protein
7FGN	FAF1 UBL1
5UM3	V122L mutant of human UBR-box domain from UBR2
2FJZ	Alzheimer's Amyloid Precursor Protein (APP) copper-binding domain (residues 133 to 189) in 'small unit cell' form, metal-free

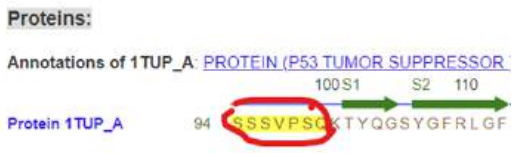
1. Open https://www.ncbi.nlm.nih.gov/Structure/icn3d/full.html
2. Input in your assigned pdb code and select Load Biological Unit

3. Choose **Analysis, Seq and Annotation**

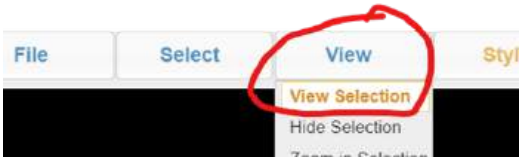
4. Choose Details tab and uncheck conserved domains



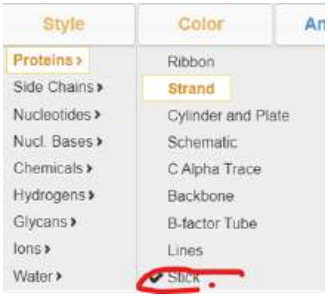
5. With your mouse, select, hold, and sweep between the first 5-10 amino acids (given in single letter code) as illustrated below. When you select them, they will turn yellow.



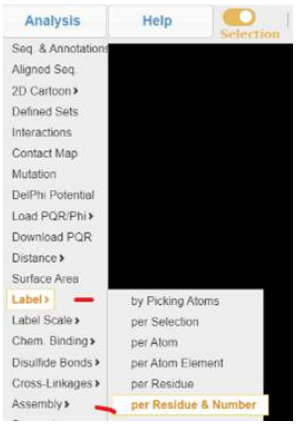
6. Choose **View, View Selection** (to limit view to what you want



7. Choose **Style, Proteins, Sticks** to see all the bonds

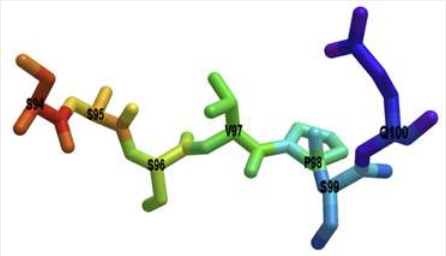
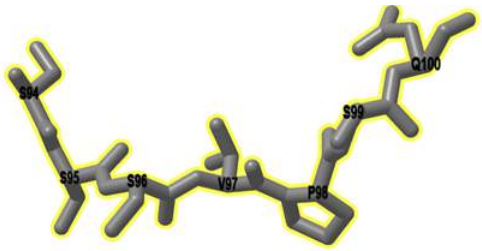
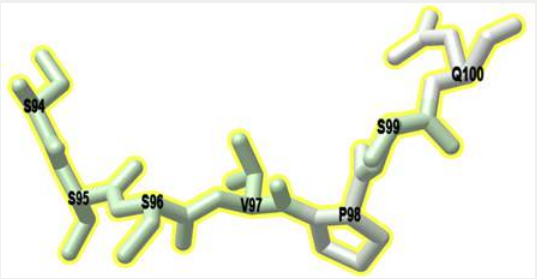
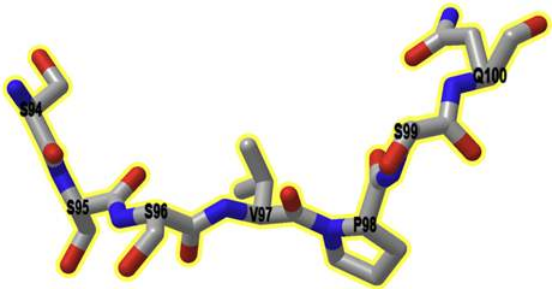


8. Change the background from black by choosing from top menu bar **Style, Background, Transparent**



9. Choose, **Analysis, Label, Per residue/#;** then **Analysis, Label Scale, 2**

10. Next, color your model as shown below in different ways as described in the table below. Then take a screen capture of the selection and replace the image in the table cell with your own

Color	Paste snip of renderings as shown below.
Spectrum, Selection to better see each amino acids in selection	
Charge Gray if no charges (ignore yellow highlight)	
Hydrophobicity (if nonpolar like oil)	
Atom (red oxygen, blue nitrogen, yellow Sulfur	

iCn3D Intro Tutorial B: Rendering a Protein

B. Rendering Full Proteins

Pick one of the proteins below that has both alpha helices and beta sheets. You will then change the protein style (rendering) to see the same protein in different ways to illustrate different properties of the proteins.

Monomeric proteins with alpha (helices) and beta (sheets)

1HDO	Human biliverdin IX beta reductase: NADP complex
2HC2	Engineered protein tyrosine phosphatase beta catalytic domain
5ZUN	Crystal structure of human monoacylglycerol lipase in complex with compound 3l
1X3S	Crystal structure of human Rab18 in complex with Gppnhp
4IN0	Crystal Structure of human splicing factor dim2/TXNL4B
1KGD	Crystal Structure of the Guanylate Kinase-like Domain of Human CASK
5KQL	Co-crystal structure of LMW-PTP in complex with 2-oxo-1-phenyl-2-(phenylamino)ethanesulfonic acid
1QGV	HUMAN SPLICEOSOMAL PROTEIN U5-15KD
1MF7	INTEGRIN ALPHA M I DOMAIN
4RQR	Crystal Structure of Human Glutaredoxin with MESNA
4JKA	Open and closed forms of R1865A human PRP8 RNase H-like domain with bound Co ion
5C4M	RhoA GDP with novel switch II conformation
6P0J	Crystal structure of GDP-bound human RalA
4MMM	Human Pdrx5 complex with a ligand BP7
4M6IJ	Crystal structure of human dihydrofolate reductase (DHFR) bound to NADPH
3M9J	Crystal structure of human thioredoxin C69/73S double mutant, reduced form

1. Load ID **4LPK**, Crystal Structure of K-Ras protein with a small molecule, GDP, bound
2. Choose **Color, Secondary, Sheet in Yellow**
3. **Style, Background, White**

4. Choose **Style, Proteins** and display as ribbon, cylinder and plate, C alpha trace, backbone, lines and sphere
5. Paste your results in the table below.

6.

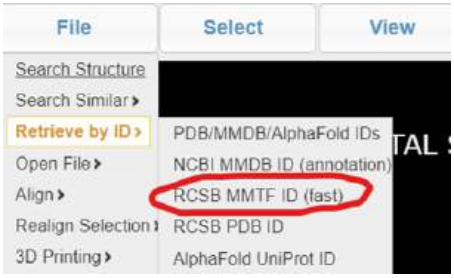
Results (replace image with yours)

PDB ID, description:

ribbon	
Cylinder and plate	
C alpha trace	
backbone	
line	
sphere	

7. Model a protein dimer that has two subunits. Use the pdb code for 1LFD (CRYSTAL STRUCTURE OF THE ACTIVE RAS PROTEIN COMPLEXED WITH THE RAS-INTERACTING DOMAIN OF RALGDS). Paste your favorite image below.

8. Model a huge structure, the human rhinovirus 14 (causes colds, PDB: 4RHV). It so big you have to load it in a different way, as shown below. Paste your favorite image below.



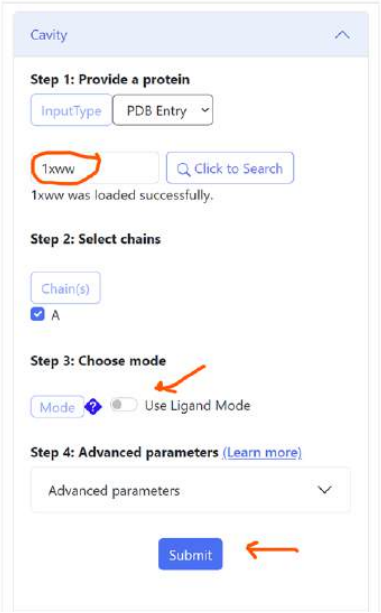
iCn3D Intro Tutorial C: Finding Pockets in Proteins

Small molecules that bind to larger proteins must have shape AND charge complementarity with the binding pocket in the protein. You can put a small molecule into an appropriate-sized pocket in a protein. You can't put a positively charged small molecule into a pocket lined with a positive charge. Let's find the pockets in a small protein, LMWPTP, a phosphatase that cleaves a negatively charged phosphate group (PO_3^{-2}) proteins (pdb 1xww). It also binds the small sulfate ion (SO_4^{-2}) in the same pocket.

Finding Pockets

Let's find the pocket where the ligand could bind using a free program called CavityPlus. 2022

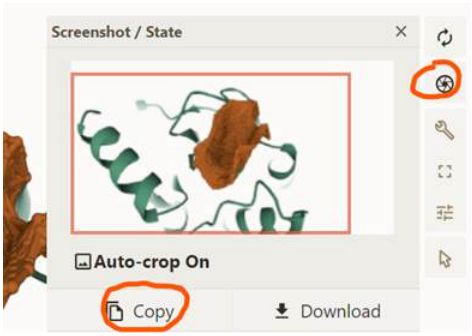
1. Load http://www.pkumdl.cn:8000/cavityplus/computation.php#/ and select **Start Computing**
2. Input 1xww, then select **Click to Search**. Wait until the structure loads to continue.
3. Then simply choose **Submit**. (Make sure that Use Ligand Mode is **not** selected)



bnbnbn

4. After the run, you will see a new window open on the left-hand-side with the protein and the top #1 Cavity highlighted. Site 1 is the presumptive location for the binding of SO_4^{2-}. Use your mouse to rotate the protein to better see the cavity. To see a list of the amino acids lining the binding pocket surface, and the surface area and volume of the cavity, select **…** under **More**. They will appear in the Residue row.

5. Copy and Paste into the table below the list of amino acids comprising the pocket into the table below. Then take a screen snip as shown in the image to the right (Note: if you can unzip the downloaded file, you could select the Download Results link and use other programs to view the results).



Amino acids in pocket	Image snip

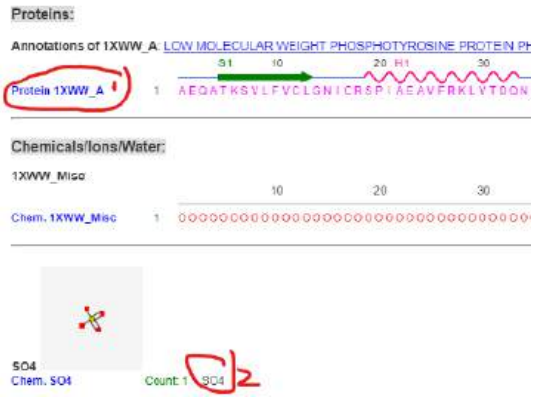
Viewing Small Molecule in a Binding Pocket

Now let's model the phosphate (PO_3^{-2})/sulfate (SO_4^{-2}) binding site in the phosphatase using iCn3D.

1. load 1XWW in iCn3D

2. Render the proteins as follows:

- **Analyses, Sequences and annotation**, Details Tab, uncheck conserved domain
- Click 1XWW_A 1st, then in the top menu bar choose **Select, Save Selection** and name it phosphatase



- **Color, Charge**
- **Style, Surface Opacity, Fast Transparency, 3**
- **Style, surface type, molecular Surface**

3. Next render the SO4 as follows:

- Choose SO4 (2) with the mouse, then **Select, Save Selection** and name it SO4
- **Style, Chemicals, Sphere**

4. Snip and paste an image of the protein with the surface display and the bound SO_4^{-2} in spheres.

5. Optional: To see actual interactions between SO_4^{-2} and the protein

- **Style, Remove surface**
- **Analysis, Interactions**,
- In popup window, choose for 1st set – sulfate; choose for the 2nd set – phosphatase; click 3D Display Interactions; Snip table with types/colors on interactions
- **View Selection; Style, Sidechains, Stick; Color, by Atom**,
- **Analysis, Label, Per Residue and number; Analysis, Label Scale, 2**;
- **Style, background, white**
- Snip an image of the interaction legend and modeled interactions, and paste below.

iCn3D Intro Tutorial D: Modeling Psychoactive Drugs in Target Proteins

Use iCn3D to model the binding of one of the psychoactive/analgesic drug to their receptor (either the 5HT or cannabinoid receptor). Paste the final model in the space shown.

Serotonin (5-hydroxytryptamine) receptors

5-HTreceptors are indirectly involved in the mechanism of action of antidepressant drugs. Most antidepressant drugs like Prozac increase the concentration of 5-HT in the extracellular brain synapse by inhibiting its reuptake into neurons. Prozac doesn't bind to the 5HT receptor but to a membrane protein that removes 5HT from synaptic region.

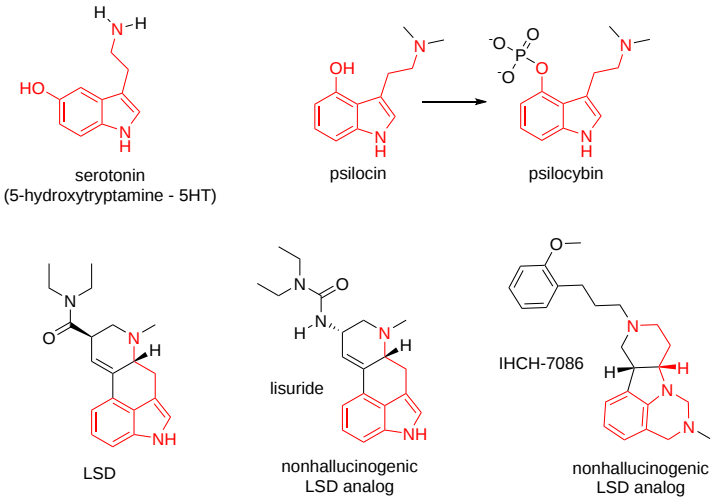
There are 7 types of 5HT receptors and each has different biochemical effects

Different drugs target different 5HT receptors.

Let's focus drugs that target 5-HT_{2A} act as antidepressants, but also lead to hallucination.

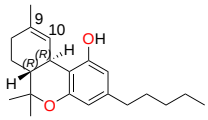
Serotonin (5HT) 2A receptor: These molecules bind to it.

- serotonin (5HT), the physiological agonist, 7WC4 (does not cause hallucinations)
- psilocin, 7WC5 , hallucinogenic, a metabolite of psilocibin)
- LSD, 7WC6, hallucinogenic
- lisuride, 7WC7, non-hallucinogenic
- IHCH-7086, nonhallucinogenic

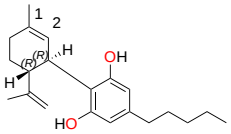


THC and CBD receptors

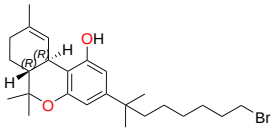
These molecules bind to them.



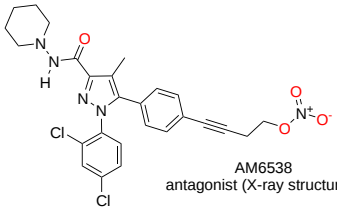
Δ9-tetrahydrocannabinol
THC



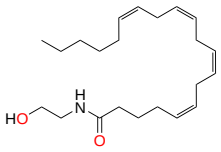
cannabidiol
CBD (AM11542)



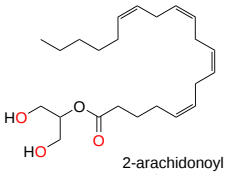
AM11542
agonist



AM6538
antagonist (X-ray structures)



ananamide
endogenous agonist



2-arachidonoyl glycerol
endogenous agonist

- Human CB1 in complex with agonist AM11542 (5XRA)
- class A GPCR Cannabinoid Receptor-Gi Complex Structure with bound agonist (6KPF) - The agonist is AM12033, which is similar to AM11542
- CBD-bound full-length rat TRPV2 in nanodiscs (6U88)

Your model:

PDB ID:

"DRUG"

RECEPTOR:

Rendered Image

---

iCn3D Skill: Alternative Rendering and Saving Files

Structure

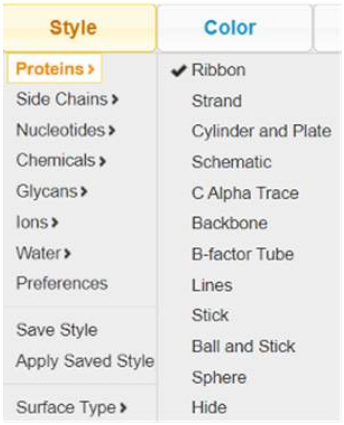
- PDB ID: 1xww
- Protein: Low molecular weight protein tyrosine phosphatase
- Activity: hydrolyzes Tyr-OPO_3^{2-} phosphoester bond
- Description: single chain, bound SO_4^{2-} (competitive inhibitor), bound glycerol (nonspecific stabilizer)

Alternative Rendering

Load Structure and Mouse/Trackpad Controls

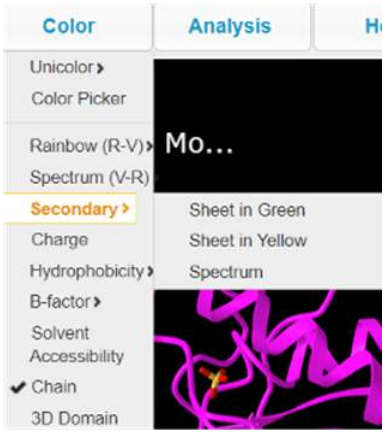
- Open iCn3D - https://www.ncbi.nlm.nih.gov/Structure/icn3d/full.html
- For a simple menu, use the dropdown: File > Customize Menus > Simple Menus.
- In the *Please input MMDB or PDB*, enter 1xww. Press enter or click load biological unit.
- Default render is ribbon (cartoon) with black background and small molecules shown as sticks. Hover over objects with the mouse to reveal their identity.

1. From the top menu bar, choose **Style**, **Protein**, then try some of the available choices:



2. For your favorite protein styles, select **Color** by left-clicking on the top menu, then pick available choices. Try:

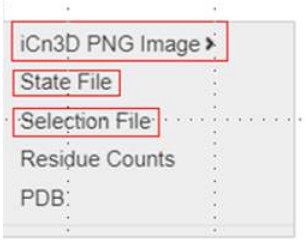
- **Secondary**, **Sheets in Yellow**
- **Charge**
- **Hydrophobicity**



3. Under **Style**, choose **Protein → Ribbon**. Under **Color**, choose **Secondary, Sheets in Yellow** before the next step.
4. To view sidechains, **Style, Side Chains, Stick** (they will remain the same color as secondary structure for now)
5. Color for SO_4^{2-} and bound glycerol will default to CPK coloring (key below)
6. Convert back to cartoon (**Style, Side Chains, Hide**)

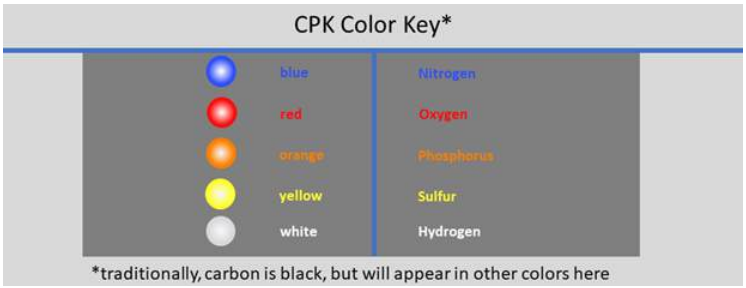
Saving Files

7. Style, Background, Transparent
8. Saving Files: There are several ways to save your work. The first option below saves a PNG image, the second creates a share link
   a. **File, Save Files, iCn3D PNG image, original size**; Give it a name. Can be reloaded in iCn3D with File, Load, iCn3D PNG IMAGE
   b. **File, Share Link, Save Lifelong Short URL**. Copy and paste this link to share your work.



Pre-Rendered Model Link

To check your work (or if you got stuck during any of the steps above) catch up using this link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?CgfEnF27TN7aYQpr6



iCn3D Skill: Displays surface of a protein - Superoxide Dismutase

A. The molecular surface of superoxide dismutase

PDB ID: **2sod**

- Description
- superoxide dismutase
- dimer with 2 Cu^{2+} ions bound in each subunit
- catalyses the reaction of $O_2^- + O_2^- + 2H^+ \rightarrow H_2O_2 + O_2$

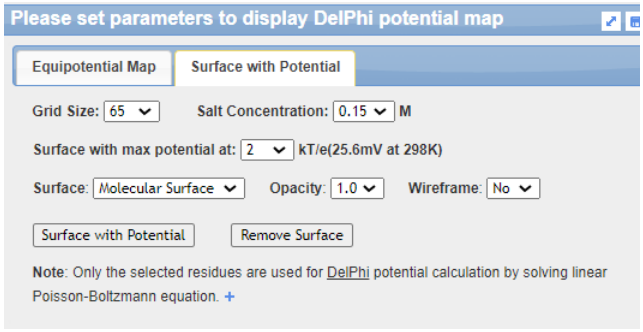
Instructions

1. **File, Retrieve by ID,** MMDB
2. **Style, Surface Opacity, Fast Transparency, 0.2**
3. **Select, Select on 3D, Chain.** Alt Click on one chain
4. **Color, Secondary, Sheets in Yellow**
5. **Style, Surface Type, Molecular** (takes a while to render)
6. Now Alt Click on the other chain
7. **Color, Wimley White Hydrophobicity*** (takes a while to render)
8. **Style, Surface Type, Molecular** (takes a while to render)
9. **Style, Background, White**
10. **File, Share Link, Save short url**

B. The electrostatic surface potential of superoxide dismutase

Background: Superoxide (O_2^-) is a toxic free radical and hence dangerous. To capture it as effectively as possible, the surface distribution of charged side changes is such that a positve electrostatic potential surrounds each active site.

1. **File, Retrieve by ID,** MMDB
2. **Select, Select on 3D, Chain.** Alt Click on one chain, Control-click on the second. Alternatively, choose A**nalysis, Sequence and Annotations, Details Tabs (uncheck conserved domains)**, and click one chain (blue font) and then control-click the second so both are highlighted.
3. In the Details tab of **Analysis, Sequence and Annotations,** sweep both Cu ions (with mouse) to select them, the **Color, Yellow**
4. **Analysis, DelPhi Potential, DelPhi Potential**
5. Choose the **Surface with Potential Map** tab. Chose the options below



6. **Select Surface with Potential.** Look a the blue (positive potential) surround the yellow Zn ions in both active sites.
7. **Style, Surface Type, Molecular** (takes a while to render)
8. **Style, Background, White**
9. **File, Share Link, Save short url**

iCn3D Skill: Aligning two structures

Superimpose and annotate active (phosphorylated) form of cyclin-dependent kinase 2 (1JST) and inactive (1FIN)

Description

1JST (active)

- PHOSPHORYLATED CYCLIN-DEPENDENT KINASE-2 BOUND TO CYCLIN A. Has bound Mn2+ and ATPγS.
- pT160 is on the regulatory T-loop of CDK2. It is mostly buried and neutralized by 3 Arg side chains. Compared to unphosphorylated CDK2–CyclinA, the T-loop moves by as much as 7 Å,
- T loop about 147-163

1FIN (inactive):

- CYCLIN A-CYCLIN-DEPENDENT KINASE 2 COMPLEX. Has a bound ATP as well as Cyclin A (B and D chains)

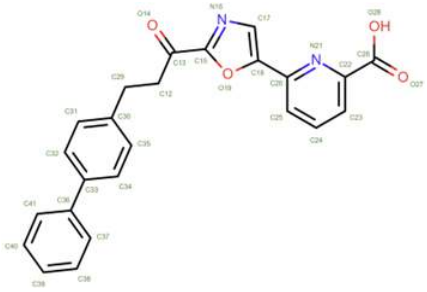
Instructions

1. **File, Align, Structure to Structure**
2. Input **1JST** and 1FIN
3. Choose All Matching Molecules Superimposed
4. In the Select residues in aligned alignment window, click the blue font 1JST-C to select the entire chain. **Color, Green, Dark Green** (green for go = active)
5. Likewise, select 1FIN-A. **Color, Yellow, Yellow** (so we can color the T loop red)
6. Alt Click in the modeling window MN and ctrl Click ATP. **Select, Save Selection, MN_ATP**
7. **Style, Chemicals, Sphere**
8. In long sequences, capture with mouse approx T loop 147-163. **Select, Save Selection**, T loops; Color red
9. In the long sequences, hover over and then click to select in 1JST_C x160 (this is phospho-Thr160) and then in 1FIN_A T160 (i.e. Thr 160). **Select, Save Selection,** Thr160s
10. **Style, Sidechains, Sticks**
11. **Color, Atom**
12. In the Defined Sets, click then ctrl click all of the selections you made: MnATP, Thr160s, align_1FIN_A, and align_1JST_C. Then save as **000All**. Using 000 at the start of the name put the combined selection at the top of the Defined Sets window. This will be important later.
13. Make sure 000All is selected, then **View, View Selection**
14. **Style, Background, White**
15. **File, Share Link, Lifelong Short Link,**
16. Paste the link into a new window, much as a student would. **Analysis, Defined Set**, then click 000All at the top to highlight the superimposed chains. Once all is highlighted, toggle between the two states by clicking the Alternative (Key “a”) menu button directly under the File button.
17. Now try View, Side by Side to see both (**don’t** toggle this form)

iCn3D Skill: Analysis of Noncovalent Interactions

Structure

- PDB ID: 3K83
- Protein: Inhibitor (below, abbreviated F278458) bound to a humanized variant of Fatty Acid Amide Hydrolase (FAH)

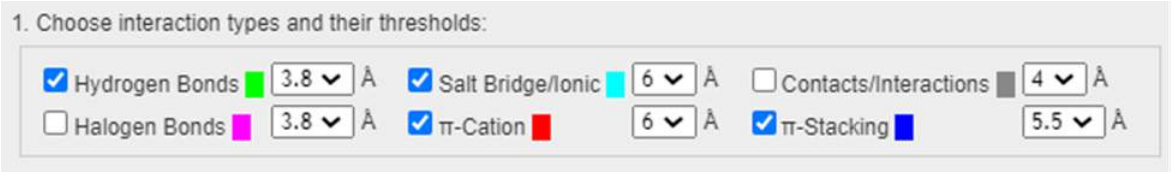


- Activity: FAH Catalyzes the hydrolysis of endogenous amidated lipids like the sleep-inducing lipid oleamide, the endocannabinoid anandamide, and other fatty amides, regulating the signaling functions of these molecules
- Description: The functional unit is a dimer, each with an active site. A chloride binds between the two subunits

In this activity you will see key noncovalent interactions between the inhibitor and (FAH). You will pick one inhibitor and view its noncovalent interactions with one subunit

Modeling Instructions

1. Open a new iCn3D window, 3K83 at: https://www.ncbi.nlm.nih.gov/Structure/icn3d/full.html
2. **Select, Select on 3D** to ensure default is **Residue**; With mouse, zoom and center to Alt-Click the inhibitor (F278458) in the magenta subunit (option-click on a Mac)**.**
3. **Select, Save Selection,** name it **Drug**
4. **Analysis, Interactions.**
5. Under 1. Choose interaction types and their thresholds, Check only the noncovalent interactions shown below. Make sure the Contacts/Interactions is unchecked as it shows all interactions between contacting Van der Waals surfaces including many hydrophobic interactions, so it is quite cluttered.



6. Select the Drug as the first set (under 2. Select the first set); choose 3K83 (under 3. Select the second set)
5. Click on the box: 4. 3D Display Interactions. This will display interacting side chains. Drag the HBonds/Interactions window to the bottom right away from the molecular display. Note the coloring of the dotted lines:

Green - hydrogen bonds

Red - π-cation

Blue - π stacking

6. **Select, Save Selection,** name it **Interactions**
7. **Style, Sidechains, Stick** (so the next step will color side chains)
8. **Color, atom (*Note*:** this drug is covalently bound to Serine241)
9. In defined sets Ctrl click Drug and Interaction to highlight all
10. **View, View Selection** (to only see Drug and Interactions)
11. **Select, Toggle Highlight** (to remove highlights if necessary)

Note: Once you run interactions, iCn3D adds many new additions to the Defined sets window. Explore these to learn different ways to examine the interactions.

Optional (if you'd like to add labels, recolor the background, and obtain the share link as in the previous activity)

12. **Analysis, Label, Per Residue & No**
13. **Analysis, Label Scale,** pick a number that works for you
14. **Style, Background, Transparent**
15. **File, Share Link,** copy short link

Pre-Rendered Model Link

To check your work (or if you got stuck during any of the steps above) view the model using this link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?pDr5EBZmo3TyAbTP6

Note: You can alter interactions on the 3D model, or examine 2D displays of them. Try:

- Click **5. Reset** at the bottom of the HBonds/Interactions window. Add Contacts/ Interactions (mostly nonpolar) by clicking on the box. Choose Drug as the first set. Choose 3K83 as the second set. Then click **4. 3D Display Interactions** to show.
- Next, try **4. 2D Interaction Network**. In the popup that opens, click a colored line in the 2D window to highlight a specific interaction.

If using the Pre-Rendered model: From the top menu, select: **Analysis → Interactions**. In the popup choose **5. Reset**, select 3K83 (full protein) as set 1 and drug as set 2, then **4. 2D Interaction Network**. Click colored lines in the 2D window.

---

iCn3D Skill: Creating and Saving Selections

Structure

- PDB ID: 3UBB
- Protein: Rhomboid intramembrane serine protease GlpG (3UBB) with phosphonofluoridate inhibitor
- Activity: Integral membrane serine protease
- Description: Single chain transmembrane protease from *E. coli* bound to a phosphonofluoridate inhibitor, which is covalently bonded to the catalytic serine. Red and blue dots ("dummy" atoms) indicate extracellular and intracellular membranes, respectively. Uses a catalytic **dyad** composed of serine (S201) and histidine (H254).

Load Structure

1. Open iCn3D - https://www.ncbi.nlm.nih.gov/Structure/icn3d/full.html
2. In the *Please input MMDB or PDB,* enter 3UBB. Press enter or click load biological unit.

Selecting using: 1. The structure viewer window with the Mouse and 2. The sequences and annotations menu (Key Learning Objectives)

This section will show how you can select residues, chains, etc in the Modeling window.

3. **Select, Select on 3D** to ensure default is **Residue**; With cursor, hover over the inhibitor (name 3UB) and Alt Click (option click on Mac) it. A yellow halo will appear around it.
4. **Select, Save Selection**, name it Inhibitor.
5. Now, use the top menu to open the sequences and annotations tab: **Analysis, Seq. and Annotations**
6. In the sequences and annotations window, uncheck "Conserved Domains," and then click the **Details tab**. Click individually on the one letter code for S201 and H254. On selection, they will turn yellow.

(Note: To see the scroll bar in sequences/annotations in a Mac, choose Systems preferences in your operating system [not the iCn3D settings], General, show scroll bars and check always; see iCn3D About notes).

7. **Select, Save Selection,** name CatDyad

Rendering (Optional, can use the highlighted link below instead)

8. **Select**, **defined sets** (Note that "defined sets" brings up a complete list of objects, many selections are pre-built into iCn3D to get you started with a model.)
9. In Selected Sets, click 3UBB_A.
10. **Color, Unicolor, Gray, Light Gray**
11. In Selected Sets, click CatDyad
12. We want to show the side chain of the catalytic dyad, so use the top menu: **Style, SideChains, Sticks**
13. Recolor to CPK coloring: **Color, Atom**
14. **Analysis, Label, Per Residue and Number**
15. **Analysis, Label Scale,** pick number that works for you
16. **Style, Background, Transparent**
17. Remove any active selections (yellow glow) by Select, Toggle highlight

Pre-Rendered Model Link

https://structure.ncbi.nlm.nih.gov/icn3d/share.html?fhT3dwckYg8i5XJj8

The short URL above may be used to catch up for the next section of the tutorial

Selecting sphere within around 5Å of the inhibitor (Key Learning Objective)

Our goal is to find all atoms with 5Å from the inhibitor, this time without showing interactions. In the next step, we will designate 2 sets of objects. Set 1 will be the inhibitor. Set 2 will be nearby residues/bound molecules. In our case, Set 2 will be defined as the protein.

18. **Select, by Distance**
19. For the first set, select inhibitor; for Set 2 click 3UBB
20. For **Set 2. Sphere with a radius to 5 Å**
21. **Click Display; Close the Select by distance window**
22. Now save the highlighted groups 5Å from the inhibitor through **Select, Save Selection,** Name 5AfromInhib
23. **Style, Side chains, Stick**
24. **Color, Atom**
25. Select both the inhibitor and the surrounding residues: In **Defined Sets,** Click Inhibitor, and Ctrl+Click 5AfromInhib (Command+Click on a Mac)
26. Display only the active site for clarity: **View, View Selection**
27. **Analysis, Label, Per Residue & Number**
28. To see water, **Style, Water, Sphere**
29. **File, Share Link,** copy short link

Pre-Rendered Model Link

To check your work (or if you got stuck during any of the steps above) view the model using this link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?sqZf4Zvqn5zWK21d9 This link shows the membrane. Choose **View, Toggle Membrane** to hide it.

---

iCn3D Skill: Mutations

Analysis: Mutation

A simple way to change 1 amino acid to another

PDB ID: **1xww**

Description: human B-form low molecular weight protein tyrosine phosphatase (has single domain) with bound sulfate (SO4) and glycerol (GOL)

Instructions

1. **File, Retrieve by ID,** MMDB
2. **Analysis, Mutation**
3. Input the desired mutation C12S this way (1XWW_A_12_S) where A is the A chain, 12 is the aa # and S is the new mutated amino acid residue Ser. Note that C12 is the wild-type active site nucleophile in phosphatases and changing it to Ser will effectively abolish phosphatase activity. To view you could use 3 different methods. We'll use one.
4. 3D with scap: toggle back and forth between structures using the key “a”
5. **Interaction**: shows the mutation in 3D and change in interactions
6. PDB: shows structure and exports a PDB file within 10 A of the mutation
7. Choose Interactions and say wow!
8. **Color, Atom**
9. **Style, Background, White**
10. Toggle back and forth between the structures with the letter “a”
11. **File, Share Link, Lifelong Short Link**

iCn3D Skill: Selection through Sequence and Annotations

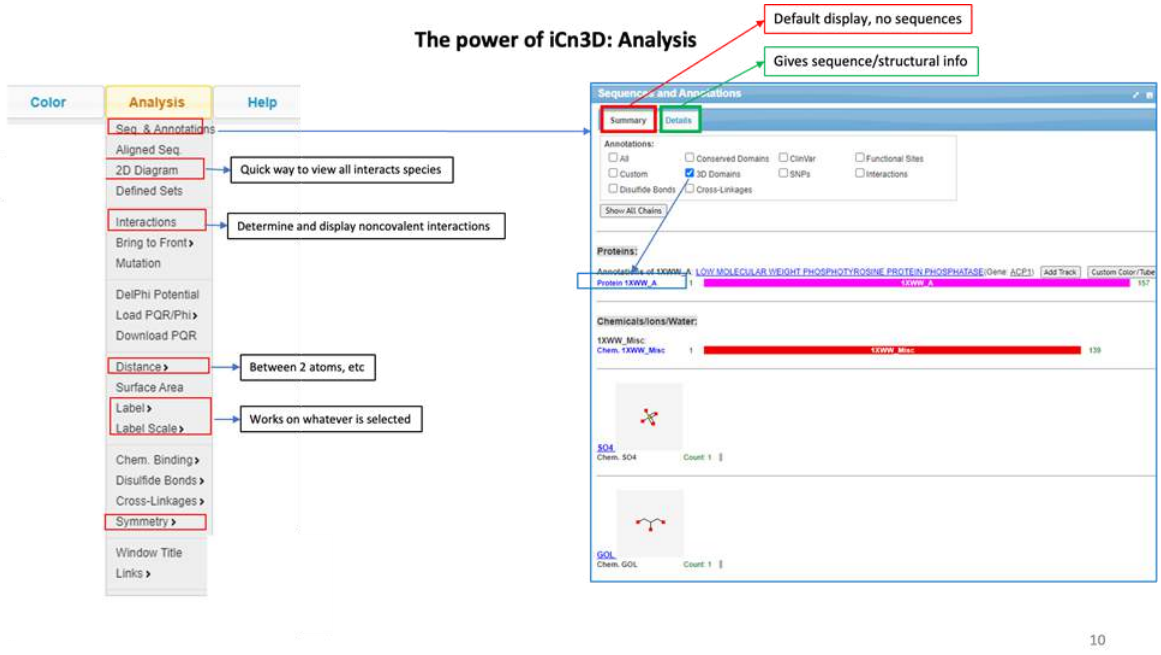
Structure

- MMDB ID: 1xww
- Protein: Low molecular weight protein tyrosine phosphatase
- Activity: hydrolyzes Tyr-OPO_3^{2-} phosphoester bond
- Description: single chain, bound SO_4^{2-} (competitive inhibitor), bound glycerol (nonspecific stabilizer)

Load Structure and Mouse/Trackpad Controls

- Open iCn3D - https://www.ncbi.nlm.nih.gov/Structure/icn3d/full.html
- For a simple menu, use the dropdown: File > Customize Menus > Simple Menus.
- In the *Please input MMDB or PDB*, enter 1xww. Press enter or click load biological unit.
- Default render is ribbon (cartoon) with black background and small molecules shown as sticks. Hover over objects with the mouse to reveal their identity.

Figure: *The Sequences and Annotations Menu*



For this part, we will be using the model from Activity 1. To load the premade model from Part 1, use this link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?CgfEnF27TN7aYQpr6

From the literature, it is known that the active site is a nucleophilic cysteine (C12). It is part of the phosphate-binding loop (P-loop, AA 12-18: sequence CLGNICR). Let's find, select, and render these amino acids.

Modeling Instructions

1. Under **Analysis** (top menu bar), choose **Sequence and Annotations**
2. Choose **Details tab, uncheck Conserved Domains**

Before we continue, look at the built-in choices you have for selection:

Annotations:

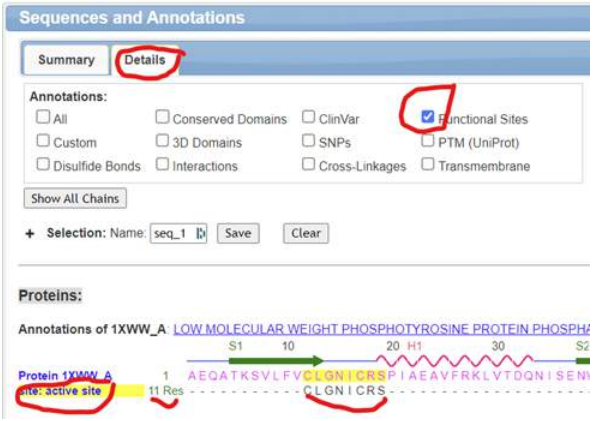
☐ All	☐ Conserved Domains	☐ ClinVar	☐ Functional Sites
☐ Custom	☐ 3D Domains	☐ SNPs	☐ Interactions
☐ Disulfide Bonds	☐ Cross-Linkages		

3. In the Sequences and Annotation window, click Protein 1XWW_A
4. Under **Select** (top menu bar), choose **Toggle Highlights**
5. Hover over C12 in the sequence (in Seq and Annot window), click and hold down the mouse key, and sweep over C12-C18 to select the P loop
6. **Select, Save Selection**, name it: Ploop
7. Within this highlighted selection**, Style, Side Chains, Sticks**
8. **Color, Atom**
9. **Analysis, Label, Per Residue & Number**
10. **Analysis, Label Scale,** pick number that works for you
11. **Analysis, Label, Change Label Color** (globally). Click in the text box and a Color box will pop up, choose from a palette, then **Display**. (Alternatively, pick a hex code**).**
12. In the Sequences and Annotation window, click SO4
13. **Style, Chemicals, Sphere** to change the sulfate to a space filling rendering
14. Files, Share Link, Copy Short URL

Pre-Rendered Model Link

To check your work (or if you got stuck during any of the steps above) catch up using this link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?QhtGuE8pkaJpGs1X9

Note: For some enzymes, iCn3D can automatically display key active site and binding residues. These can be seen as shown by selecting the items indicated in the left figure.



iCn3D Skill: Showing a Protein-Protein Interface

Protein:Protein Interface

PDB ID: 3S9D

- binary complex between interferon alpha 2 (IFNa2) and its recetor IFNAR2
- Cyan: IFN, blue IFN receptor

Instructions

1. **Retrieve by ID,** MMDB
2. **Analysis, Seq. & Annotationsm, Details Tab.**
3. **Select magenta IFN alpha chain:** 3s9d_A
4. **Select, Save Selections** INF
5. Repeat with blue 3s9d-b. Save Select INFR
6. **Select, INF; Select by Distance, within 5A.** Set 1: selected; Set 2: nonselected; **Display**
7. **Select, Save Selection:** 5Ang from IFN
8. **Select 5A from INF: Style, side chains stick, color atom**
9. repeat for INFR
10. **Select Select by Distance, within 5A.** Set 1: IFNR; Set 2: nonselected; Display
11. **Select, Save Selection**: 5Ang from IFNR
12. **Select, Side Chains, sticks; Colore atom**
13. **Select 5A from IFNR: Style, side chain sticks. Color, charge, Style, surface type, vanderWaals**
14. **Analysis, Interactions, Set 1: 5A from INF; Set 2: 5A from INFR**
15. **Select, Save, Selections, Inteactions**
16. **Style, Background, White**
17. **File, Share Link, Lifelong**

iCn3D Tutorial: Binding interactions of SARS-Cov-2 Spike receptor domain with ACE2

A. 1st Way

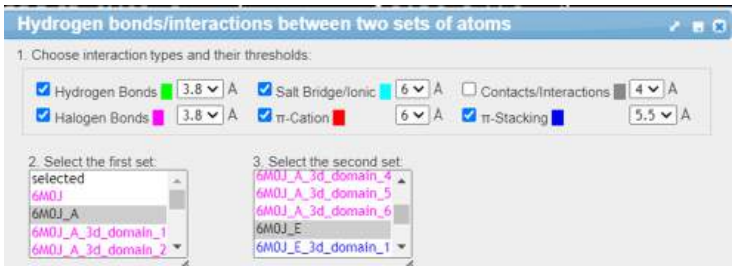
PDB ID: **6M0J**

Description: SARS-CoV-2 spike receptor-binding domain bound with ACE2

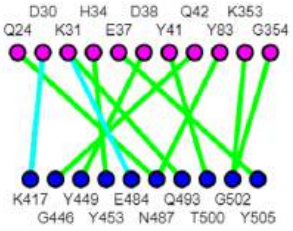
- Angiotensin-converting enzyme 2, chain A, denoted 6M0J_A, pink
- Spike Protein S1, chain B, denoted 6M0J_E, blue

Instructions

1. **File, Retrieve by ID,** MMDB
2. **Seq. & Annotations, Uncheck Conserved Domains**
3. **Analysis, Interactions**, then select the choices in the image below. (Leave out Contact/Interactions to decrease clutter)



4. Click 3D Display Interactions
5. Without closing Interactions window, choose Select, Save Selection, name it BindingInterface
6. Click **2D interaction network** to get the image below



7. In the Selected Sets window, choose BindingInterface.
8. **View, View Selection**
9. **Select, Select Side Chain**
10. **Color, Atom**
11. In the Selected Sets window, choose BindingInterface.
12. **Analysis, Label, Per Residue & Number**
13. **Style, Background, White**
14. **File, Share Link, Lifelong Short Link**

B. 2nd Way

Instructions

1. **File, Retrieve by ID,** MMDB
2. **Seq. & Annotations, Uncheck Conserved Domains, Details Tab**
3. Check **Interactions**

4. From right hand side window sequences: Under 6M0J_A, select the blue label Interact .E, Save as A_with_E
5. Under 6M0J_E, select the blue label Interact .A, Save as E_with_A
6. **Analysis, Interactions.**
7. Choose A_with_E and E_with_A
8. **3D Display Interactions**
9. **Select, Sidechains**
10. **Color, Atom**
11. In H Bonds/ Interactions window, choose Highlight Interactions in Table (can toggle on/off individual interactions),
12. Try the 2D Interaction Map and the Buried Surface Area
13. **Style, Background, White**
14. **File, Share Link, Lifelong Short Link**

---

iCn3D Tutorial: Overlay many ACE2 receptor binding domain analogous to SARS-Co2 RBD using Blast

Use BlastP to align receptor-binding domains

PDB ID: **6M0J**

- SARS-CoV-2 spike receptor-binding domain bound with ACE2
- Angiotensin-converting enzyme 2, chain A, denoted 6M0J_A, pink
- Spike Protein S1, chain B, denoted 6M0J_E, blue

Instructions

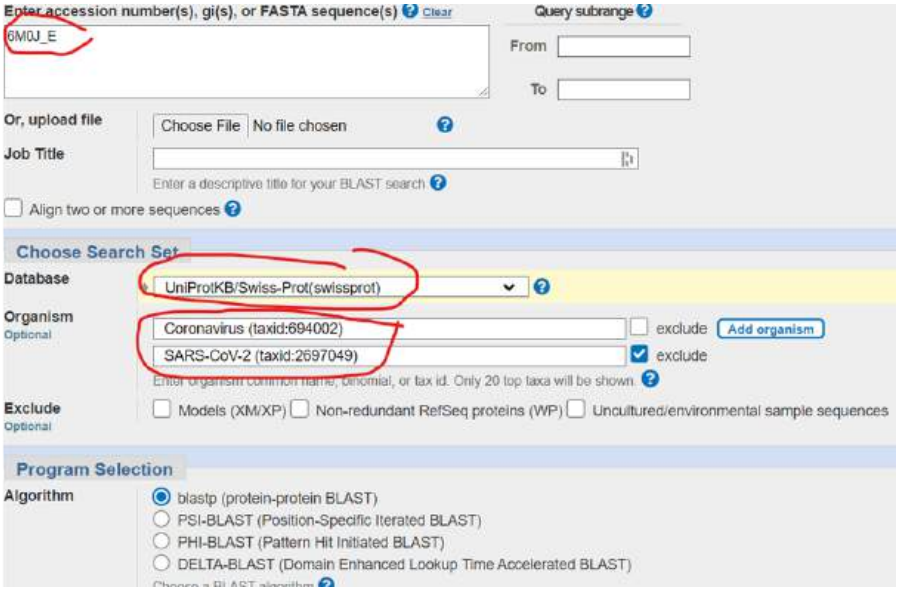
1. **File, Retrieve by ID,** MMDB
2. **Seq. & Annotations.** In Summary tab check Custom and Conserved Domains
3. In Summary tab, click on 6M0J_E
4. **View, View Selection**
5. In Details tab, Add track.
6. Choose the FASTA Alignment Tab in the new window
7. Cut and paste the aligned FASTA sequences file you made using the instructions on the next page or to save time just copy/paste the same FASTA Alignment file below into the iCn3D Fasta window.
8. Enter 319 for the Position of the first residue in Sequences & Annotations window (# is right there); **Check Color sequence by identity**:
9. Choose Add Tracks. The RBD appears color coded with the darkest rd 100% identity and the darkest blue 0% (See F377, K387)
10. **Style, Background, White**
11. **File, Share Link, Lifelong Short Link (this might not work if you choose too many proteins)**

Instructions to use Blast

Use of BlastP for COV-2 ACE2 Receptor Binding Domain (RBD)

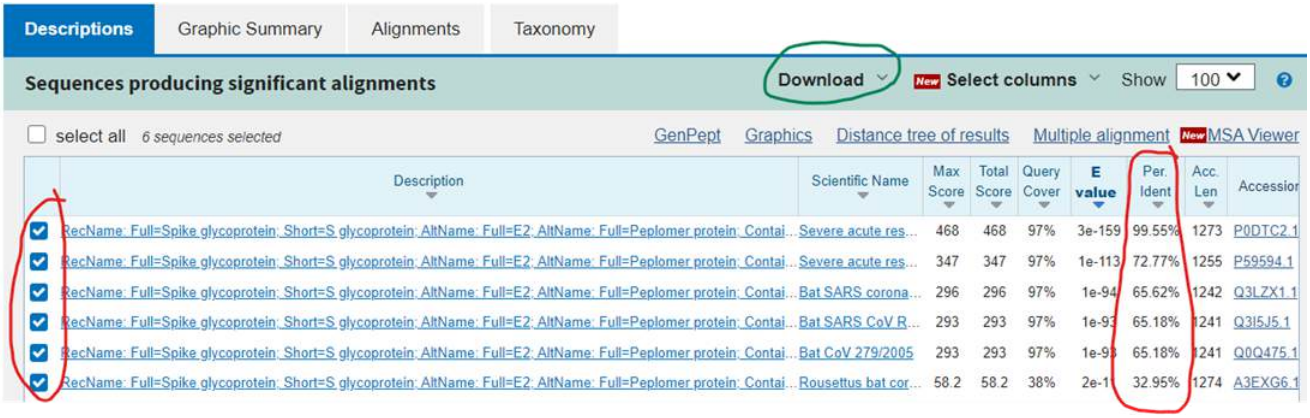
Find sequences homologous to the structure of the RBSthrough BlastP:

Instructions - use this image for help



1. Go to BlastP
2. Enter this Accession number (in this case PDB ID with chain: **6M0J_E** for the RBD
3. Under Choose Search Set

4. **Database:** Choose UniProt
5. **Organism:** input Coronavirus and click it when it appears in the search
6. Add Organism and then input SARS-CoV-2 and click it when it appears in the search; Click Exclude (since we already know the structure of the RBD of CoV-2)
7. Choose Blastp and then BLAST
8. After a few minutes you will get the general output shown below .



9. Check just the ones with percent identity >30% and choose Download, **Fasta Aligned Sequences**
10. **Copy the resulting text file and use in iCn3D** as described in the previous page**.**

---

iCn3D Tutorial: Protein Kinase B (AKT)

Protein Kinase B (AKT)

A more complicated protein with multiple domains and at least two major conformations with different PDB structures: Protein Kinase B (aka AKT)

PDB files:

- 3CQW (active)
- 3O96 (inactive)
- 1UNQ PH domain

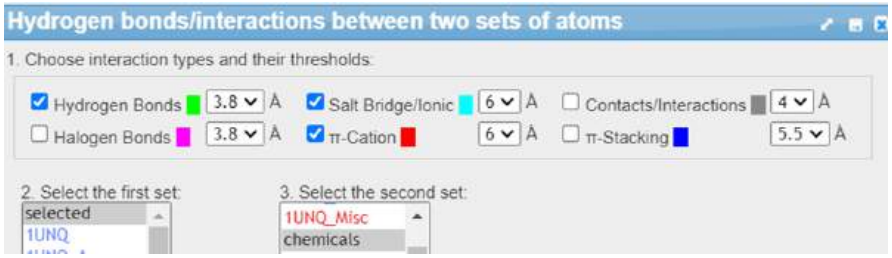
A. Exploration of AKT – Domain Structure

- PDB ID: **1UNQ**
- Description: Pleckstrin Homology Domain Of Protein Kinase B/Akt Bound To Ins(1,3,4,5)-Tetrakisphophate

1. Open new iCn3D
2. **File, Retrieve by ID,** MMDB
3. Alt Click the ligand**. Style, Chemicals, Sphere**
4. **Select, Select on 3D, check Chain,** alt click the protein
5. **Color, Secondary, Sheets in Yellow**

With the protein still selected:

1. **Analysis, Interactions**
2. For the second set choose chemicals
3. Check just the noncovalent interactions shown below



4. Choose 3D Display Interactions. This will display interacting side chains
5. **Selections, Save Selections,** name it Interactions
6. **Style, Sidechains, Stick** (so next step will just color side chains)
7. **Color, atom**
8. In defined sets choose Interactions. **Analysis, Label, Per Residue & No**.
9. **Analysis, Label Scale, 3**
10. **Style, Background, White**
11. **File, Share Link,** copy short link

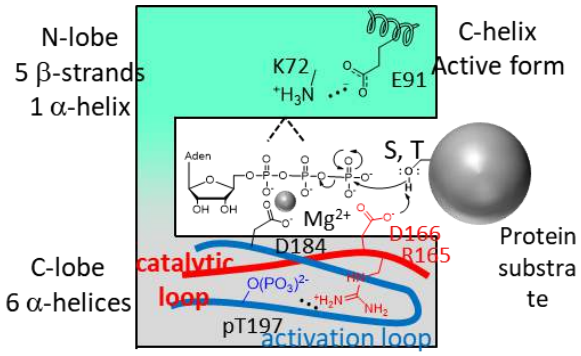
B. Exploration of AKT – Self-defined aa, sequences, Render PH-domain, N-lobe, C-Lobe, Catalytic Loop, Activation Loop

PDB ID: **3O96 (inactive)**

Description:

- Crystal Structure of Human AKT1 with an Allosteric Inhibitor
- RAC-alpha serine/threonine-protein kinase; Chain A
- Allosteric inhibitor IQO

Cartoon structure



Instructions

1. **File, Retrieve by ID,** MMDB
2. **Analysis, Seq. & Annotations, Uncheck Conserved Domains**
3. **Analysis, Defined Sets**. Click 1st domain (PH = top 3O96_A_3d_domain)
4. **Style, Protein, C-Alpha Trace.**
5. **Color , Unicolor, Yellow**
6. **Defined Sets**. Choose 2st domain (N-Lobe). **Color, unicolor, Cyan**
7. Defined Sets. Choose 3st domain (C-Lobe). **Color, unicolor, Grey, Light Gray**
8. **Seq. & Annotations,** Details Tab**,** Uncheck Conserved Domains.
9. Select by hand the catalytic loop (table below) by using the horizontal scroll to find and select 271V-287H. Hover over each amino acids and click to select it or use your mouse to sweep out a rectangle over all the range of amino acids.
10. **Select, Save Selection, CatLoop**
11. **Color, Unicolor, Red**
12. Repeat for **Activation Loop** and save the selection. **Color Unicolor, Blue**
13. **Style, Background, White**
14. **File, Share Link, Lifelong Short Link**

Numbering or generic kinase and AKT

SITE	**Generic** Kinase	**Akt (+108)**
N lobe	K72	K180
N lobe	E91	E199
C lobe, cat loop	R165	R273
C lobe, cat loop	D166	D274
C lobe, act loop	D184	D292
C lobe, act loop	T197	T305 (NO SEQ)
Approx Cat Loop	163-179	271V—287H

Approx Act loop Start DFG (292-294) to APE	184-200	292-319 308Tmiss toAPE end 319

C. Compare 3CQW (active) and 3O96 (inactive) – Superimpose two AKT structures

PDB ID: **3CQW** and **3O96**

Description

3O96 (inactive): Crystal Structure of Human AKT1 with an Allosteric Inhibitor

- RAC-alpha serine/threonine-protein kinase; Chain A
- Allosteric inhibitor IQO

3CQW (active):

- Crystal Structure of Akt-1 complexed with substrate peptide (hence active form) and inhibitor
- Peptide: Glycogen synthase kinase-3 beta peptide

Instructions

1. **File, Align, Structure to Structure**
2. Input **3CQW** and **3O96**
3. Choose All Matching Molecules Superimposed
4. In the top of the right hand alignment window, save the aligned state as presented (al_seq1). Just the aligned sequences will be highlighted.
5. View, View Selection
6. **Style, Background, White**
7. **File, Share Link, Lifelong Short Link**
8. Toggle between the two states by clicking the toggle menu button just underneath the File menu button

iCn3D Tutorial Question: Intermediate Problem - Cyclooxygenase II

Structure

- PDB ID: 4PH9
- Protein: Ibuprofen (IBP) bound to cyclooxygenase-2
- Activity: COX-2 catalyzes the conversion of arachidonic acid(AA) to prostaglandin G2 (PGG2), and is a target of non-steroidal anti-inflammatory drugs (NSAIDs) and COX-2 selective inhibitors (coxibs). Arachidonic acid (AA), not shown in this structure, binds in a "L" shape

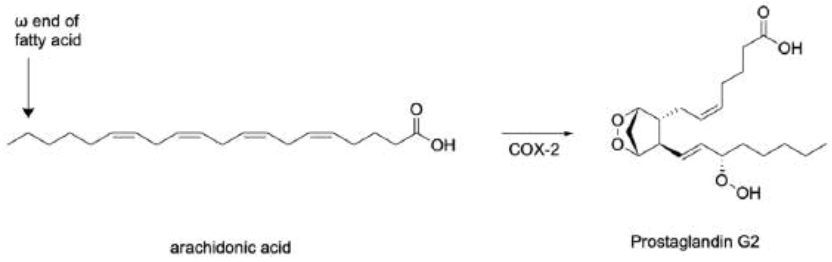
Key amino acids

- Arg-121 and Tyr-356 are close to the carboxylate of AA
- Phe205, Phe209, Val228, Val344, Phe381, and Leu534 form a hydrophobic groove for the ω-end of AA.
- Ser 530, which is above this, gets acetylated by aspirin
- Tyr 385, near C13 in AA, forms a free radical which removes a single electron from C13

For more information on the mechanism of COX-2, scroll down to the end of the chapter section in Fundamentals of Biochemistry.

Description: Biological dimer with heme, ibuprofen (IBP), and many other ligands bound. The initial display is messy! Your task is to clearly render a specific structural feature.

Assessment: The enzyme cyclooxygenase-2 (COX-2) produces arachidonic acid from prostaglandin G2, as shown in the reaction below. Prostaglandin G2 is an important metabolite in inflammation, so inhibition of (COX-2) reduces inflammation.



Potential visualization activities for students:

A. Identify the noncovalent interactions of COX-2 with a heme in 1 subunit
B. Model and describe the noncovalent interactions of COX-2 with ibuprofen
C. Model the interactions at the dimer interface of the protein; identify two amino acids on different subunits that are participating in a [hydrogen bond / ionic interaction]
D. Show the key active site residues of the enzyme

Steps for Discussion and/or Modeling

STEP 1: As a table, or small group at one table, choose one from A-D above.

STEP 2: Load the model 4PH9 at https://www.ncbi.nlm.nih.gov/Structure/icn3d/full.html

With your group, broadly discuss the steps students would need to perform to accomplish the activity you chose in STEP 1. Consider:

- for your course, would it make sense for them to start with the model as it is loaded?
- If you were to pre-render the model and provide them a shared link, which steps would you perform ahead of time? Which steps of modeling are important to their learning

STEP 3: Open the sequences and annotations tab:

- Analysis → Sequences & Annotations
- Uncheck "conserved domains," and then check "functional sites"

STEP 4: Discuss how the built in iCn3D features may help you create a model quickly.

Pre-Rendered Model Links

1. Identify the noncovalent interactions of COX-2 with a heme in 1 subunit

- https://structure.ncbi.nlm.nih.gov/icn3d/share.html?ffnfemgp1ZJZxa4u9 w/o preset
- https://structure.ncbi.nlm.nih.gov/icn3d/share.html?8PaBfJKGZoKR5y5g8 w/preset

2. Model and describe the noncovalent interactions of COX-2 with ibuprofen

- https://structure.ncbi.nlm.nih.gov/icn3d/share.html?XHxMrn48zrVbKKCy6 w/o presents
- https://structure.ncbi.nlm.nih.gov/icn3d/share.html?sJHhAqhPDgT57kTV8 w/preset

3. Model the interactions at the dimer interface of the protein; identify two amino acids on different subunits that are participating in a [hydrogen bond / ionic interaction]

- https://structure.ncbi.nlm.nih.gov/icn3d/share.html?V3Zs4zCAoTM2Mk8A8 w/preset
- https://structure.ncbi.nlm.nih.gov/icn3d/share.html?K27jkv5n25rhXm9B8 surface

4. Show the key active site residues of the enzyme

- https://structure.ncbi.nlm.nih.gov/icn3d/share.html?A3mmUMd7KznoymrN6 (w/o preset)
- https://structure.ncbi.nlm.nih.gov/icn3d/share.html?fznyAtJnTkTmuNreA (3KRK with bound arachadonic acid)

CHAPTER OVERVIEW

Literature Based Guided Assessments (LGAs)

Topic hierarchy

Voltage-Gated Sodium Channel: LGA - Students 082423

Voltage-Gated Sodium Channel: LGA - Students 082423

Written by Henry Jakubowski, Emily Schmitt Lavin, Arthur Sikora, and Subhasish Chatterjee

Literature-Based Guided Assessment (LGA)		Voltage-Gated Sodium Channel

Introduction

Eukaryotic voltage-gated sodium (Na_V) channels generate and sustain action potentials in nerve and muscle cells by moving Na^+ ions from the outside to the inside of the cell. This increases and makes positive the transmembrane potential of the cell, which at rest is approximately -70 mV (more negative inside). Once activated, the channel undergoes a fast inactivation (1-2 ms), without which the firing of nerves and muscles becomes dysregulated, a potentially lethal effect. Please view the information in Chapter 11.3 on the voltage-gated sodium channel before you do these guided assessment activities.

The questions below are derived from a paper from Jiang et al. on the structure and properties of the **α-scorpion toxin LqhIII** (MW 7,000) bound to rat cardiac sodium channel **Na_V1.5** (MW 227,000) by. (Jiang, D., Tonggu, L., Gamal El-Din, T.M. *et al.* Structural basis for voltage-sensor trapping of the cardiac sodium channel by a deathstalker scorpion toxin. *Nat Commun* **12**, 128 (2021). https://doi.org/10.1038/s41467-020-20078-3. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/)

The **Lqh** toxin is made by **L**eiurus **q**uinquestriatus ***h****ebraeus*. It is found in North Africa, the Middle East, and Western India and is shown below.

htt.://en.wikipedia.org/wiki/Deathstalker.

The study shows how the deathstalker scorpion (LqhIII) toxin inhibits the fast inactivation of cardiac sodium channels (Nav1.5). In other words, you will see how the toxin keeps the channel open longer than it would be open in its absence.

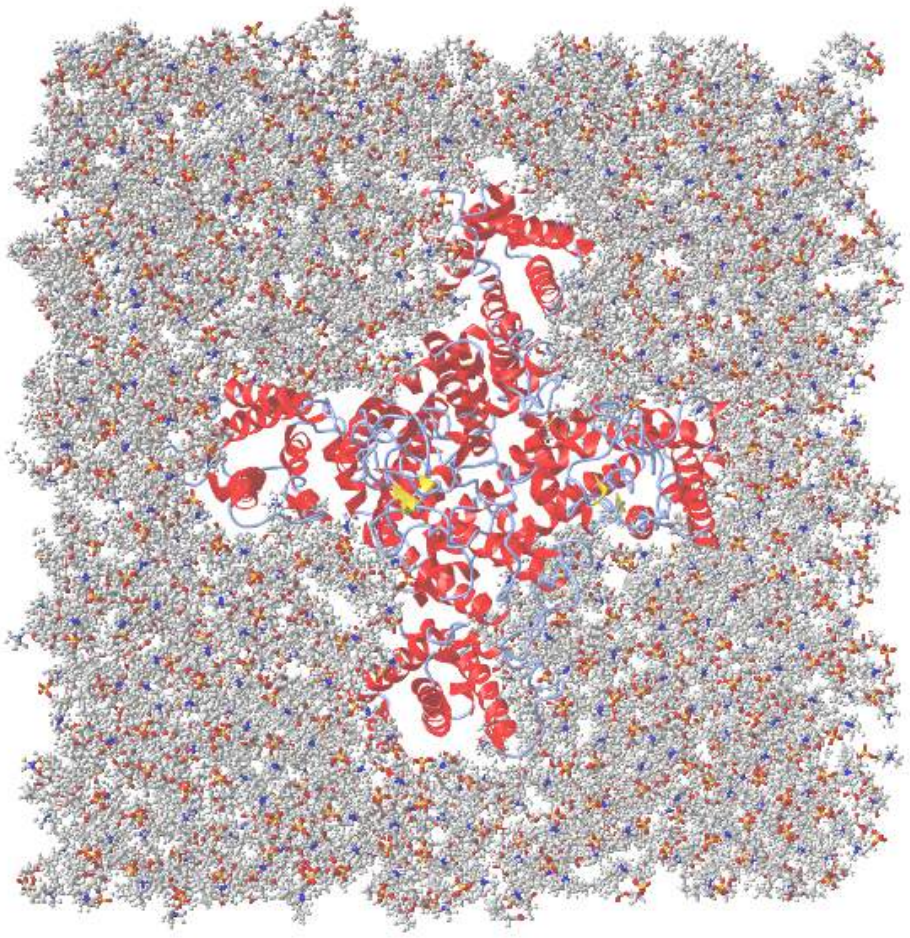
- In the absence of toxin, the sodium channel **Na_V1.5** returns in 1-2 ms to an inactive state when an 'inactivation gate" moves to occlude the open pore.
- The α-scorpion toxin LqhIII inhibits the return of the channel to the inactive state. Since the toxin inhibits the channel's fast inactivation of Na^+ ion flow into the cell, the channel stays open longer.

The toxin leads to the inhibition of the normal fast inhibition (inactivation) of the channel. Hence the channel stays open (activated) longer. This is analogous to the statement that the enemy of my enemy is my friend!

Voltage-Gated Sodium Channel.1: Techniques to study the Na Channel

For the experiments described in the paper, the rat sodium cardiac channel $Na_V1.5$ was purified, its structure determined, and its functional properties (the regulated movement of Na^+ to the inside of the cell - electrophysiology) in human epithelial cells measured. The protein is expressed in rat cardiac cells and is found in the cell membrane of the cell. The iCn3D image below shows the alpha chain of rat cardiac $Na_V1.5$ (6UZ3) embedded in a simple bilayer (DMPC) to model how it might appear in a cardiac cell membrane bilayer.

Data from
https://api.libretexts.org/endpoint/bounce/https://bio.libretexts.org/@api/deki/files/75773/6



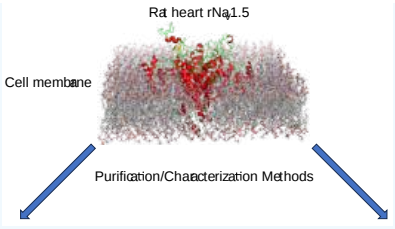
```
> load url
https://api.libretexts.org/endpoint/bounce/https://bio.libretexts.org/@api/deki/files/75778/6UZ3_
NaV1pt5inDMPCCharmm.pdb | type pdb
>
```

PDB coordinates based on S. Jo, T. Kim, V.G. Iyer, and W. Im (2008). CHARMM-GUI: A Web-based Graphical User Interface for CHARMM. *J. Comput. Chem. 29:1859-1865*. S. Jo, T. Kim, and W. Im (2007) Automated Builder and Database of Protein/Membrane Complexes for Molecular Dynamics Simulations. *PLoS ONE 2(9):e880*

? Exercise *[Math Processing Error]*

Complete the flow chart below to show two different approaches that could be used to purify the protein and prepare it for structural and functional studies. On the left show steps you could use to directly purify the protein from rat hearts. On the right use the rat gene (Scn5a) for the channel as the starting point for purification. Here are some links for review if needed.

- Chapter 3.4: Protein Purification
- Overview of Protein Expression
- Recombinant protein expression



Answer

From heart tissue:

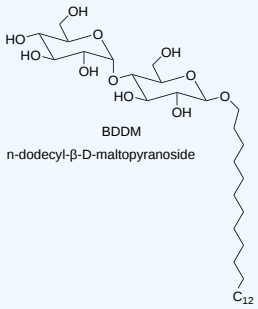
From heart tissue DNA

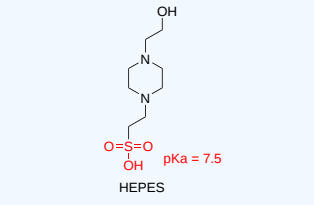
Here are some review links:

- Overview of Protein Expression
- Recombinant protein expression

? Exercise *[Math Processing Error]*

Key components of the buffer solution used to purify the channel include HEPES and 1% (w/v) n-dodecyl-β-D-maltopyranoside. Their structures are shown below. Describe the role of each.





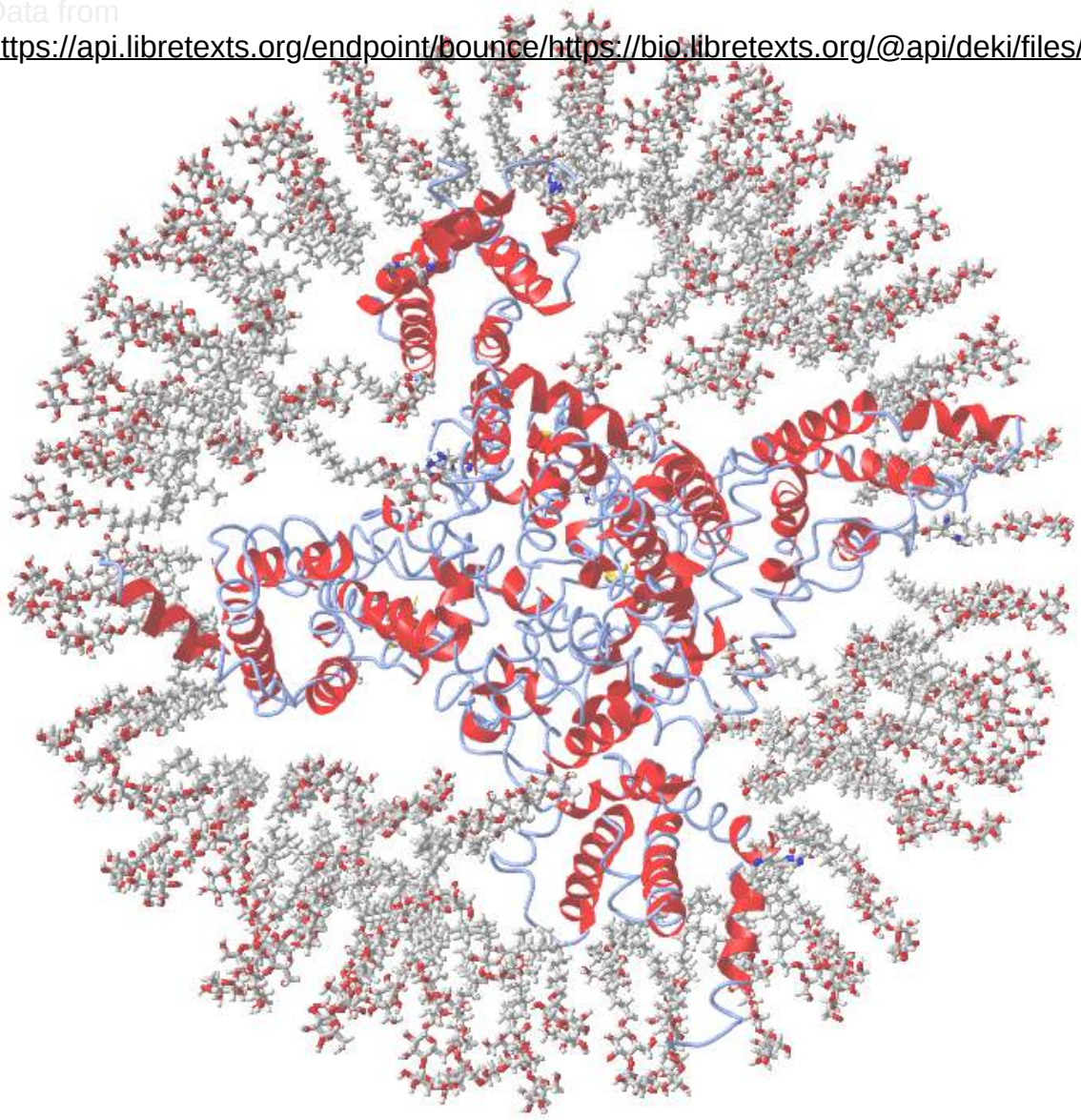
Answer

.

? Exercise *[Math Processing Error]*

The following iCn3D shows the interaction of the purified sodium channel with n-**d**o**d**ecyl-**β**-D-**m**alto**p**yranoside (BDDM). Explain the differences between the iCn3D models showing the protein in a bilayer and interacting with the BDDM.

Data from
https://api.libretexts.org/endpoint/bounce/https://bio.libretexts.org/@api/deki/files/75768/6



Feedback

```
> load url
https://api.libretexts.org/endpoint/bounce/https://bio.libretexts.org/@api/deki/files/75768/6uz3_
ratcardNaVinMicellePymolCIFDiscPDB.pdb | type pdb
>
```

Answer

.

? Exercise *[Math Processing Error]*

The elution of the complex on a size exclusion column (Panel A) and the analyses for the eluted fractions by SDS-PAGE (Panel B) are shown in the figure below. The α-scorpion toxin LqhIII: rat cardiac sodium channel $Na_V1.5$ complex elutes in the area of the first peak shown in blue.

a. How do molecules separate on size exclusion chromatography?

b. Compare the molecular weights of the first peak to the second complex peak in Panel A.

c. Which band(s) in Panel B likely represent the rat cardiac sodium channel $Na_V1.5$ based on the intensity of the stained band? The toxin LqhIII is the lowest band. (bands at 17.5K and 12.5 are FGF12b and calmodulin, respectively, which were added to stabilize the channel**)**

d. Why do the channel and toxin elute together in the size exclusion column shown in Panel A, but are separate bands in PAGE gel in Panel B?

e. How could you get the complex to separate as two peaks, the free $Na_V1.5$ channel, and the free α-scorpion toxin LqhIII?

f. To get information on the receptor, go to Uniprot and paste in rat cardiac sodium channel $Na_V1.5$ into the search box. Go to Sequence and Isoform in the left panel and find the actual MW. Knowing this, what are the major bands at about 170K and 60K?

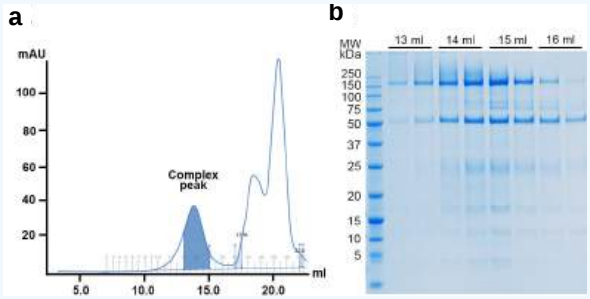


Figure: Purification of the **r**ecombinant NaV1.5C/LqhIII complex. a. Representative size-exclusion chromatography profile of purified **r**Nav1.5C/LqhIII. Peak fractions collected for cryo-EM grid preparation are shown in blue. b. SDS-PAGE of the size exclusion peak fractions stained by Coomassie blue.

Answer

.

CryoEM was used to determine the structure of the sodium channel:toxin complex. Here is a short YouTube video that describes the technique. Also review the appropriate part of Chapter 3.3: Analyses and structural predictions of protein structure.

? Exercise *[Math Processing Error]*

Describe the temperature conditions for protein samples in cryoEM. What is the reported resolution of cryo EM structure? How does this compare to X-ray structures? What are some advantages of using cryoEM over X-ray crystallography and NMR to determine the structure of proteins?

Answer

.

Molecular dynamics was also used to probe the conformational changes in the structure of the complex on the picosecond (10^{-12} s) to nanosecond (10^{-9} s) time scale. For a review of molecular dynamics, see Chapter 3.3: Analyses and structural predictions of protein structure. It can be used to probe dynamic changes in protein structure which cryoEM can't.

? Exercise *[Math Processing Error]*

Answers these multiple choice questions (created by AIPDF through ChatGPT4 -paid version using this prompt: Write 5 question for a biochemistry major about the use of molecular dynamics and the finding in the paper)

1. What was one of the primary uses of molecular dynamics in this research?
   - A) Predicting the behavior of NaV channels without toxins.
   - B) Analyzing hydration and Na^+ permeation through the rNaV1.5C/LqhIII complex.
   - C) Studying the interaction between different toxins.
   - D) Predicting the behavior of potassium channels.

2. In the molecular dynamics simulation analysis, what was aligned to the initial position for each snapshot?
   - A) The α-toxin LqhIII.
   - B) The voltage-sensing domain IV.
   - C) The Cα atoms from pore transmembrane helices.

- D) The fast inactivation gate.

3. Approximately how long were the unrestrained "production" simulations generated?

- A) 10.35 ns.
- B) 5000 steps.
- C) 300 ns.
- D) 2 fs.

4. Based on the molecular dynamics analyses, what was observed about the activation gate structure of the rNaV1.5C/LqhIII complex?

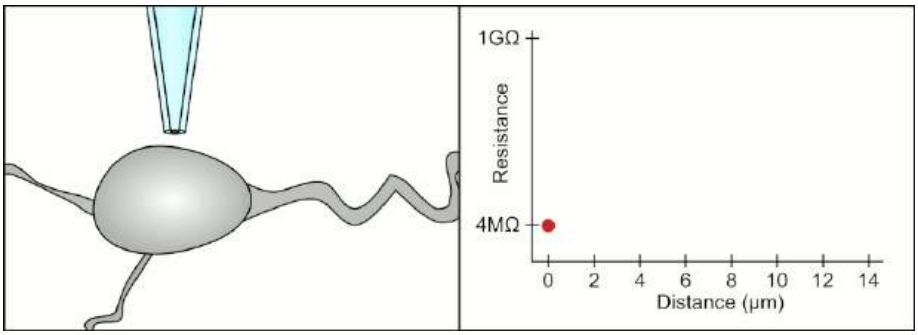
- A) It was fully open for Na^+ conductance.
- B) It was functionally closed for Na^+ conductance.
- C) It was in a metastable state.
- D) It showed no significant change from the rNaV1.5C structure.

Answer

.

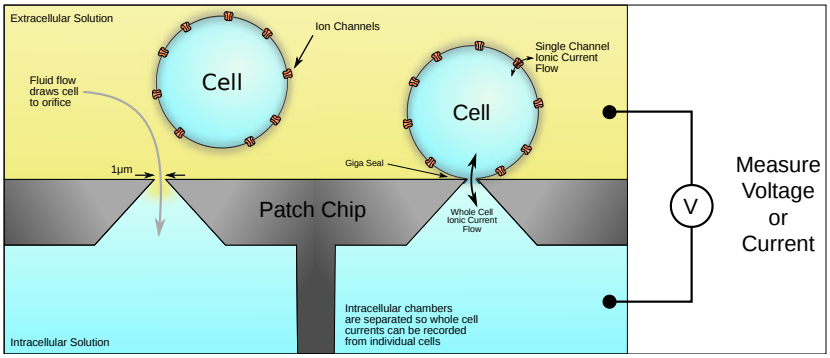
The authors used two types of electrophysiological techniques, patch clamp, and voltage clamp. Here is some brief background.

In a whole-cell patch clamp experiment, a pipet is placed on a cell, and suction is applied until a tight seal, indicated by a sharp rise in electrical resistance (gigaohm level) is made. This is illustrated in the figure below.



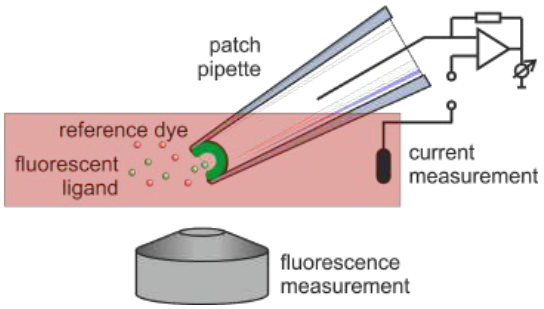
Patch Clamp Resistance. Formation of gigaseal. Holst. https://en.wikipedia.org/wiki/Automa..._Animation.gif. CC BY-SA 3.0

The cell can then be connected to a patch clamp chip in such a way that transmembrane potential or current can be measured on single-channel ion flow. This is illustrated in the figure below.



Holst. Patch Clamp Chip. Batch clamp chip showing a gigaseal, whole-cell recording configuration, and the ion channel and whole cell current. https://en.wikipedia.org/wiki/Automated_patch_clamp#/media/File:Patch_Clamp_Chip.svg. CC BY-SA 3.0

In patch-clamp fluorometry, part of the cell membrane is sucked into the tip with the seal intact. Fluorescent ligands can be applied to one side of the membrane that contains an ion channel and current measurements were made as illustrated in the figure below. Alternatively, as in this paper, side chains in the S4 voltage sensor were labeled with a fluorophore, and changes in fluorescence were observed with changes in membrane potential.



Patch-Clamp Fluorometry. https://www.uniklinikum-jena.de/phys...n/Methods.html

? Exercise *[Math Processing Error]*

Answers these general multiple-choice questions (created by AIPDF through ChatGPT4 -paid version using this prompt: Write five multiple-choice questions about the use of patch clamp techniques to measure sodium currents in cells)

1. What is the primary purpose of the patch-clamp technique in cellular electrophysiology?
   - A) To visualize cell structures.
   - B) To measure the concentration of sodium ions inside cells.
   - C) To record ion currents across cell membranes.
   - D) To stimulate cellular growth.

3. In a typical neuron at resting potential (-70 mV) and in this study (epithelial cells transformed with the rat channel, what is the direction of the sodium current when sodium channels open?
   - A) Inward, into the cell.
   - B) Outward, out of the cell.
   - C) There is no movement of sodium.
   - D) Both inward and outward simultaneously.

4. Which of the following factors can influence the magnitude and direction of sodium currents measured using patch-clamp techniques?
   - A) The concentration of potassium ions outside the cell.
   - B) The voltage across the cell membrane.
   - C) The pH of the cell cytoplasm.
   - D) The size of the cell.

5. Why might a researcher use drugs or toxins during a patch-clamp experiment measuring sodium currents?
   - A) To increase the size of the cell.
   - B) To modulate or block sodium channels and observe the effects.
   - C) To change the color of the cell.
   - D) To stimulate cell division.

Answer

.

? Exercise *[Math Processing Error]*

Answer these general multiple-choice questions about patch-clamp fluorometry. (created by AIPDF through ChatGPT4 -paid version using this prompt: write 5 multiple choice questions of patch clamp fluorometry in which key amino acids in a membrane protein are labeled with a fluorophore)

1. What is the primary advantage of combining patch-clamp with fluorometry in studying membrane proteins?
   - A) It allows simultaneous measurement of electrical activity and conformational changes.
   - B) It increases the fluorescence of all amino acids.
   - C) It enhances the electrical activity of the protein.
   - D) It allows visualization of the entire cell in detail.

2. Why are specific amino acids in a membrane protein labeled with a fluorophore in patch-clamp fluorometry?
   - A) To increase the size of the protein.
   - B) To change the electrical properties of the protein.
   - C) To detect specific conformational changes in the protein during activity.
   - D) To make the protein more soluble in water.

3. Which property of the fluorophore is crucial for patch-clamp fluorometry?
   - A) Its electrical charge.
   - B) Its sensitivity to changes in the local environment or protein conformation.
   - C) Its ability to increase protein activity.
   - D) Its color in visible light.

4. In which scenario would patch-clamp fluorometry be especially useful?
   - A) When studying the overall shape of a cell.
   - B) When investigating the relationship between ion channel gating and conformational changes.
   - C) When trying to increase the fluorescence of a solution.
   - D) When observing the movement of proteins inside the cell.

5. What is a critical consideration when choosing a fluorophore for labeling amino acids in patch-clamp fluorometry?
   - A) The taste of the fluorophore.
   - B) The electrical conductivity of the fluorophore.
   - C) The photostability and brightness of the fluorophore.
   - D) The size of the fluorophore molecule.

Answer

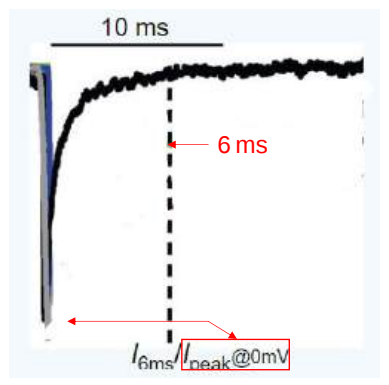
.

Voltage-Gated Sodium Channel.2: Nonstructural Lab Studies of LqhIII Toxin Effects on Rat Sodium Channel $Na_V1.5$ ($rNa_V1.5_C$)

HEK293S GnTI$^-$ (epithelial-like) cells were transformed with the rat cardiac sodium channel $Na_V1.5$ ($rNa_V1.5_C$). The cells were then studied in the absence and presence of the toxin at varying times after toxin addition and at various concentrations of the toxin. The opening and closing of the channel were determined by measuring changes in the Na^+ currents into the cell on channel opening.

? Exercise *[Math Processing Error]*

The resting potential of a cell is around -70 mV (more negative inside). When the transmembrane potential is depolarized by raising the transmembrane potential to around -55 mV or even more positive, the Na^+ channels are activated, and an inward Na^+ current (black line in a modified form of Figure 1a from the paper below) which goes **downward** by convention) through the channel occurs. This is followed by a quick inactivation of the channel and the return to the baseline flow of ions. In the experiment below, the potential was raised from -100 mV (channel closed) to 0 mV (channel open). What is happening to the $Na_V1.5$ Na^+ channel during these 10 ms? What is special about the current at 6 ms (indicated by the dashed vertical line)

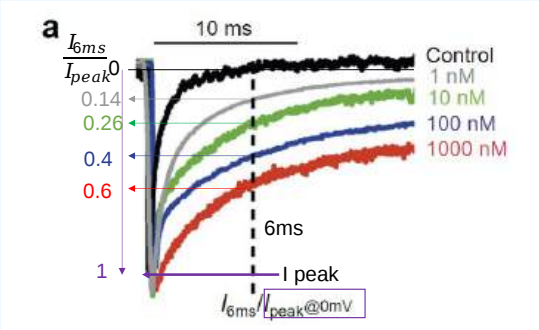


Answer

Initially, the change in voltage opens the $rNa_V1.5_C$ and allows an inward flow of Na^+ ions as evidenced by the vertical drop. As described in the introduction, conformational changes in the $Na_V1.5_C$ (closing of the inactivation loop), follow which closes the channel and stops the current, so the current returns to baseline within about 6 ms. The channel is in the inactive state by around 6 ms.

? Exercise *[Math Processing Error]*

Figure 1a from the paper (modified) below shows a series of lines of different colored (black to red) representing Na^+ currents obtained at 0 (black line) and increasing concentrations (gray through red) of the LqhIII scorpion toxin. Let's assume that the downward I_{peak} =1. The values of I_{6ms}/I_{peak}, calculated from the approximate values shown on the graph, are also shown on the vertical axis Does the toxin alter the immediate response of the cells after the channel was activated? What effect does increasing [toxin] have on the response of the cell? Offer a structural explanation of how the toxin affects the cell by suggesting changes in the toxin-bound structure.



Answer

The toxin at any dose does not substantially affect the initial opening of the channel since the size of the current does not change at first. Then the toxin inhibits the rapid inactivation of the channel within the first 6 ms, leading to a prolonged inward Na^+ current. The inhibition of the inactivation is dose-dependent on the concentration of the toxin. The toxin does not completely block the quick inactivation of the channel as the lines don't return to the black baseline in the time interval measured. This suggests that the toxin binds to the channel and either partially occludes it or prevents the normal conformational change in the $Na_V1.5_C$ that causes a quick deactivation of the channel.

Since Na^+ ions still flow through the channel but at a lower rate than the open state, the toxin-bound channel likely represents a 4th, or partially-opened state of the channel.

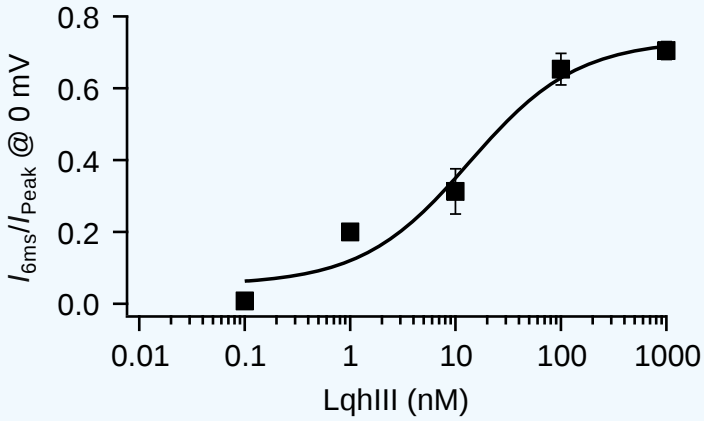
? Exercise *[Math Processing Error]*

Figure 1a (left) from the paper below shows the dependency of the inhibition of the quick inactivation of the channel on the log of the LqhIII concentration. When the transmembrane potential is set to 0 mV (as in this experiment), the channel should open and the current would be maximal. The data points in the graph are close to the ones estimated in the graph from the previous question.

a. In the absence of the toxin, what should the current I be at 6 ms compared to the maximal Na^+ current? That is, what would be the value of I_{6ms}/I_{peak}?

b. Is the channel completely inactivated in the presence of the toxin?

c. At 6 ms, what concentration of toxin (nM) causes 50% inhibition of the maximum effect of the inhibitor on the normal rapid inactivation of the channel?

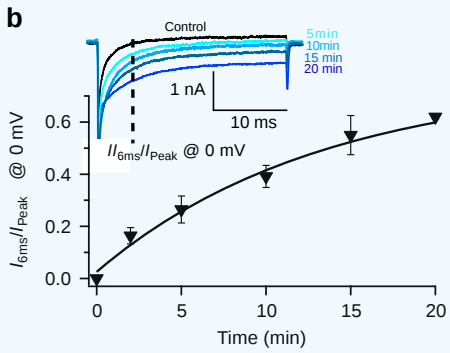


Answer

.

? Exercise *[Math Processing Error]*

In the next experiment, cells were kept at -120 mV at one fixed concentration (100 nM) of toxin. The toxin was left to incubate with the cells for various times up to 20 min. After the incubation time, the transmembrane potential was changed to **0 mV** to activate the channel, and inward Na^+ currents were measured. The results are shown in the top inset graph in Figure 1b from the paper. (Note: It is unclear from the paper if the control was determined at 0 min with 100 nM toxin or no toxin.)



a. Did the length of time cells were pre-incubated with the 100 nM toxin affect Na^+ currents after depolarization of the cells? How did the effects on the cells depend on the preincubation time?

b. Describe and explain these results

Answer

.

Here is another interesting feature of toxin binding. The toxin binds to a site on the resting state of the $Na_V1.5_C$ with high affinity. When the cell becomes depolarized (made more + inside the cell), the affinity for the toxin decreases so it starts to dissociate. The affinity of the toxin for $Na_V1.5_C$ decreases with increasing + transmembrane potential. At very high positive potentials (+100 mV) it appears not to bind.

? Exercise *[Math Processing Error]*

What might account for the decreasing affinity of the bound toxin for the $Na_V1.5_C$ with an increasing transmembrane potential?

Answer

.

? Exercise *[Math Processing Error]*

Time course experiments were conducted on the complex at 100 nM of LqhIII scorpion toxin. A three-pulse protocol can be applied to alter membrane potentials:

- 1^{st}: a pulse from −120 mV to +100 mV for the indicated times then
- 2^{nd}: a 50-ms hyperpolarizing pulse (make membrane potential very negative, perhaps around -100 mV)
- 3^{rd}: a pulse of 50 ms to 0 mV

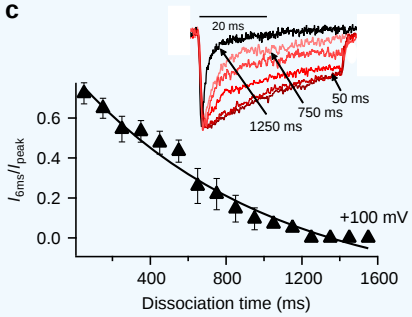
Note that steps 2 and 3 both occur with 0.1 s. What is the purpose of each pulse?

Answer

.

? Exercise *[Math Processing Error]*

Figure 1c from the paper below shows results for a set of 3-pulse designed to allow recovery of the fast inactivation which was blocked by the previous toxin binding. Note that the transmembrane potential for the first pulse was +100 mV



a. Describe what happens to the channel and complex.

b. Explain the results

c. Summarize thermodynamic and kinetic features that make the toxin so effective.

Answer

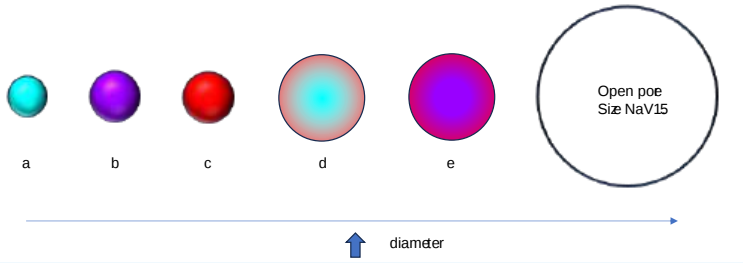
.

Voltage-Gated Sodium Channel.3: Structural Studies of LqhIII Toxin Effects on Rat Sodium Channel $Na_V1.5$ ($rNa_V1.5_C$) - CryoEM

Before we discuss in detail the structure of the sodium channel and its complex with the toxin, let's look at an important attribute of molecules that helps determine their function, the actual size of the species involved.

? Exercise *[Math Processing Error]*

The Na^+ and the K^+ voltage-gated ion channels must have an open pore when the channel protein is active (Na^+ ions move across the channel). Extracellular Na^+ and the K^+ ions don't exist as "naked" ions but they are hydrated by water in an aqueous extracellular environment. The figure below shows the relative sizes of these Group I cations and their hydrated forms, in comparison to the diameter of the open NaV1.5 pore in the channel. Answer the following questions. The red sphere (c) represents the calculated value of the diameter of water assuming its volume when it is bound to a protein is 25 $Å^3$.



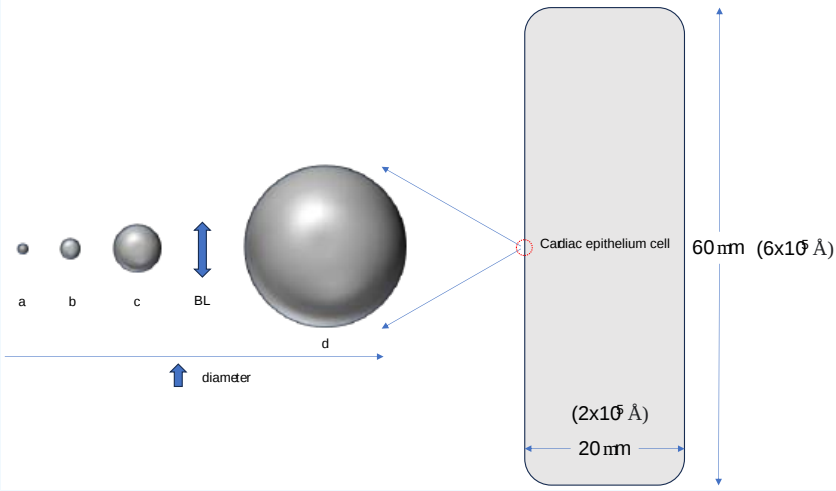
1. Which represents the "naked" (nonhydrated) size of the K^+ ion?
2. Which represents the hydrated Na^+ ion?
3. Based on the pore size alone, which of these species could diffuse through the pore?

Answer

.

Exercise *[Math Processing Error]*

The approximate relative sizes of the hydrated Na^+ ion, pore opening, toxin, and the NaV1.5 protein are shown in the figure below along with the relative width of the bilayer (BL). A cardiac epithelium cell is shown as a rectangle to the right. A red dot  (not visible in the large figure) in the membrane surrounded by the red-dotted circle represents a single NaV1.5 channel.



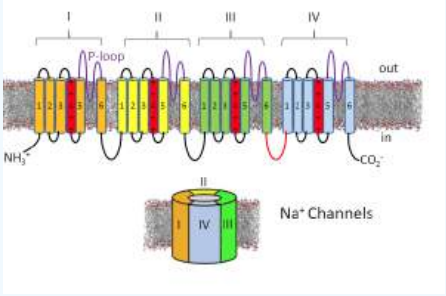
1. Which likely represents the pore?
2. Which represents NaV1.5 channel protein?
3. Which represents the toxin?

Answer

.

Exercise *[Math Processing Error]*

If you hadn't read the paper, where would you consider the most likely location for a toxin to bind to affect the function of Nav1.5? Circle the most likely toxin binding site in the schematic below. Based on the paper, where did it bind? Redraw the toxin in the correct location based on the paper.



Answer

.

? Exercise *[Math Processing Error]*

How is this toxin (LqIII) different in terms of its binding location from other toxins that interfere with the functioning of Nav1.5? What is the effect of the toxin on the channel and the symptoms of this venom?

Answer

.

? Exercise *[Math Processing Error]*

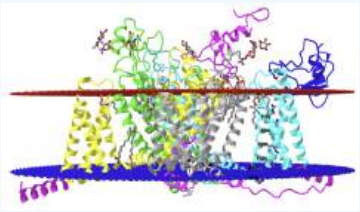
What amino acids should be present in the S4 segment of Nav1.4 and why?

Answer

.

? Exercise *[Math Processing Error]*

The figure below shows an interactive iCn3D model of the rat sodium channel $Na_V1.5$ bound to the LqhIII toxin (7k18).



Rat sodium channel $Na_V1.5$ bound to the LqhIII toxin (7k18). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...jpz8CFQajMKJe7

Domains I-IV are shown in gray, yellow, green, and cyan, respectively. The magenta parts are just extensions/connectors of these domains. The LqhIII toxin is shown in dark blue. The Ile-Phe-Met (IFM) sequence of the inactivation gate that is involved in closing the pore in fast inactivation is shown as black spheres and labeled.

It is important to know that some stretches of amino acids are missing from the structure. These include AAs 1-120 (N-terminus), 298-301, 425-501, 602-608, and 1781-1838.

a. To which domain does the LqhIII toxin bind?

b. Is the binding site close to the pore-forming segments of the domain or the voltage-sensitive segments?

c. Does the layer of red spheres represent the outer (extracellular) or inner (intracellular) leaflet of the membrane?

d. Offer reasons that parts of the protein are missing from the structure.

Answer

.

? Exercise *[Math Processing Error]*

The IFM motif has been shown to be conserved across all voltage-gated sodium channels.

a. What role does it play in these channels and in Nav1.5 and why?

b. The cause of Paroxysomal Extreme Pain Disorder (PEPD), an extremely rare disease with only 15 known affected families, appears to be mutations in the IFM motif which leads to increased sensations of pain. What is a likely effect of the mutations in the IFM motif?

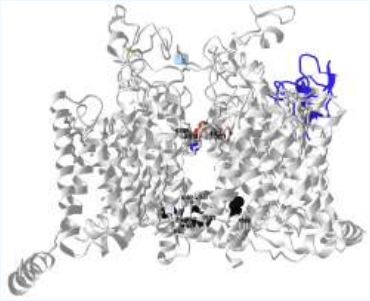
https://en.wikipedia.org/wiki/Paroxy..._pain_disorder

Answer

.

? Exercise *[Math Processing Error]*

The figure below shows a different interactive iCn3D model of the rat sodium channel $Na_V1.5$ bound to the LqhIII toxin (7k18) without a membrane representation for clarity. It shows the selectivity filter DEKA (spacefill, CPK colors), the inactivation gate IFM and the IFM "internal receptor" F1651, L1660, and N1662 (spacefill, CPK colors), and a ring of hydrophobic residues V413, L941, I1471, and I1773 (spacefill, black) that in the closed state completely seal off the cytoplasmic opening in the pore.



Rat sodium channel $Na_V1.5$ bound to the LqhIII toxin without a membrane representation (7k18). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...pPzakHu6Ew6WH9

After viewing the structure from all angles, do you think that the toxin:$Na_V1.5$ complex looks closed, open, or inactivated form? Explain

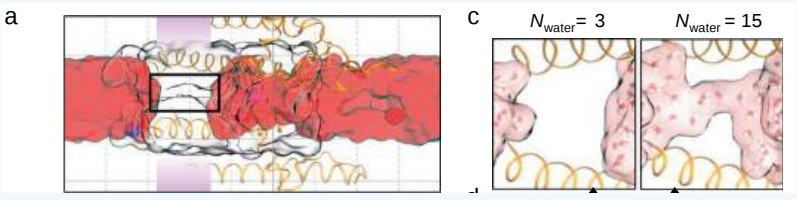
Answer

.

Voltage-Gated Sodium Channel.4: Structural Studies of LqhIII Toxin Effects on Rat Sodium Channel $Na_V1.5$ ($rNa_V1.5_C$) - Molecular Dynamic Simulations

? Exercise *[Math Processing Error]*

Now let's look at some data to see if the pore is really open, partially open, or closed in the toxin:channel complex. One clue is if the structure shows water in the pore as the Na^+ ions must be hydrated to pass through the pore (the opposite case is seen with K^+ channels when K^+ pass through stripped of water). **Molecular dynamic (MD)** simulations were done on the $Na_V1.5$ with and without the toxin to simulate the environment in the channel opening. The results of the MD simulations are shown below in Figures 6 a and c from the paper.



Molecular dynamics analysis of hydration and Na^+ permeation through the $rNa_V1.5_C$/LqhIII complex.

Panel a shows a side view of $rNa_V1.5_C$ (orange ribbons; domains II and IV) from MD simulations highlighting Na^+ ions (blue spheres), the water-occupied volume within a cylinder of radius 8.5 Å (red surface), and the protein-occupied volume within a cylinder of radius 12 Å (colorless surface). The cavity within the pore is outlined with a black rectangle. The region of the intracellular activation gate is shown as a purple band.

Panel c shows molecular representations of the gate containing N_{water} = 3 (left) or 15 (right) water molecules

Based on these studies, do you believe the pore is closed, open, or partially open?

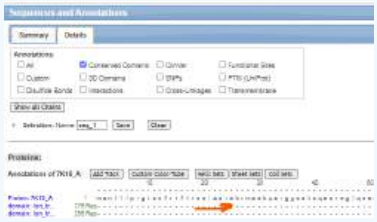
Answer

.

Voltage-Gated Sodium Channel.5: Detailed Structural Analyses of LqhIII Toxin Effects on Rat Sodium Channel $Na_V1.5$ ($rNa_V1.5_C$)

? Exercise *[Math Processing Error]*

From the iCn3D model, write the sequence of the S4 segment that contains the Arg side chains and describe the properties of the amino acids in the sequence. Do this by scrolling along the sequence window in iCn3D (shown below) until you find the labeled Arg shown in the model.



Answer

.

? Exercise *[Math Processing Error]*

Is the helix amphiphilic? That is, are the Arg side chains all on one face of the helix and the nonpolar on the other? To find out, copy and paste the sequence of S4 (above) in this helical wheel predictor and run the program.

Answer

.

✓ Exercise *[Math Processing Error]*

Make a simplified view of the iCn3D model by hiding Domains 1-III to more readily see the contributions of S5 and S6 of Domain IV to the pore.

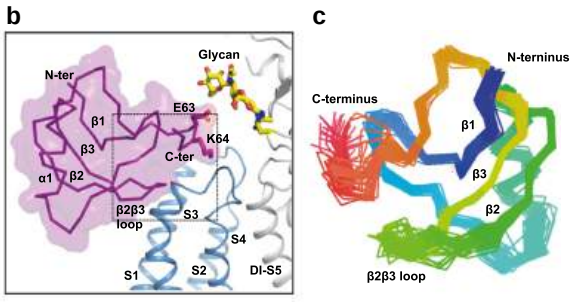
- open iCn3D and load 7K18.
- With your mouse or trackpad, choose **Sequence and Annotation** in the top menu bar
- Choose the Details tab
- Ctrl-Click the two sequences highlighted in yellow below for Domain IV and the toxin.
- Choose **View** from the top menu bar and then **View Selection**
- Choose **Style, Background, Transparent**

Solution

.

Voltage-Gated Sodium Channel.6: **Noncovalent Interactions of LqhIII and Domain IV/VS**

Now let's look at the actual interaction of the toxin with the Domain IV/VS of the channel. A closeup showing the interaction site is shown in Figures 3 b and c from the paper below. Panel C next to it shows the NMR-solution structures of the toxin in the absence of the channel. Each structure determined is represented by a single color line color codes red at the C-terminus to blue at the N-terminus.



Panel b: CryoEM structure of the $rNa_V1.5_C$ Domain IV/VS and LqhIII complex; Panel c: NMR structure of free LqhIII

? Exercise *[Math Processing Error]*

Using this iCn3D model, describe the secondary structure of the bound toxin. How many pairs of cysteine residues are in the LqhIII toxin. Identify which cysteines are involved in the disulfide bonds. What effect do the disulfide bonds have on the beta sheet structure?

Answer

.

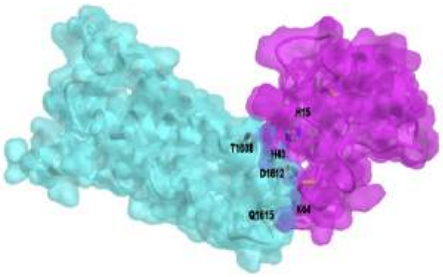
? Exercise *[Math Processing Error]*

What sections of the toxin in panel B make the closest interactions with the Domain IV/VS of the channel? Describe their conformation flexibility in the free toxin.

Answer

.

The figure below shows an interactive iCn3D model of a surface rending of Domain IV of the rat sodium channel $Na_V1.5$ bound to the LqhIII toxin (7k18).



Surface rending of Domain IV of the rat sodium channel $Na_V1.5$ bound to the LqhIII toxin (7k18). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...x5QRPEctT5zsW7.

The molecular surface and underlying secondary structure of the LqhIII toxin are shown in magenta, with key residues H15, H43, and K64 shown as CPK-colored sticks and labeled. Domain IV/VS of the channel is shown in cyan, with key amino acid side chains T1608, D1612, and Q1615 shown as colored and labeled sticks.

? Exercise *[Math Processing Error]*

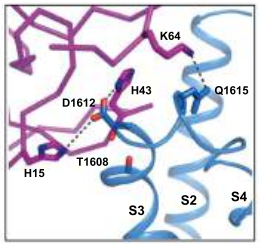
Comment on the shape and possible side chain interactions that contribute to high-affinity binding of the inhibitor to Domain IV/VS.

Answer

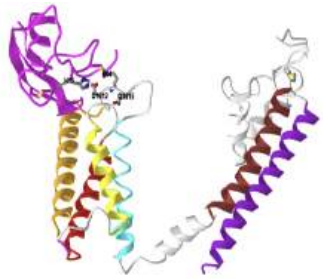
.

Figure 3d from the paper below shows the detailed interactions between LqhIII and *D*IV-VS. Key residues shown in sticks were labeled. Interaction surfaces of the *D*IV-VS (blue) and the LqhIII (purple). Key residues for the interaction are shown in yellow shading and embedded stick

d



The figure below shows an interactive iCn3D model of the rat sodium channel $Na_V1.5$ Domain IV bound to the LqhIII toxin (7k18).



Rat cardiac sodium channel NaV1.5C Domain IV/LqhIII toxin complex (7k18). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?JAjUQAZDZoQ5wiVX9

The toxin is shown in magenta. The segments are colored as follows: S1 is red, S2 is orange, S3 is yellow, S4 is cyan, S5 is brown and S6 is . Key amino acid pairs involved in the binding of the toxin to Domain IV are shown in sticks and labeled.

? Exercise *[Math Processing Error]*

In summary, name and locate the amino acid residues that serve the following roles in the LhqIII toxin: DIV-VS interactions.

1. Which amino acids in the toxin interact with D1612 (the paper describes the interaction as pincers surrounding D1612).
2. The conserved negatively charged residue in the Nav1.5 channel
3. What position is Thr in and what is thought to be its role in the mechanism?

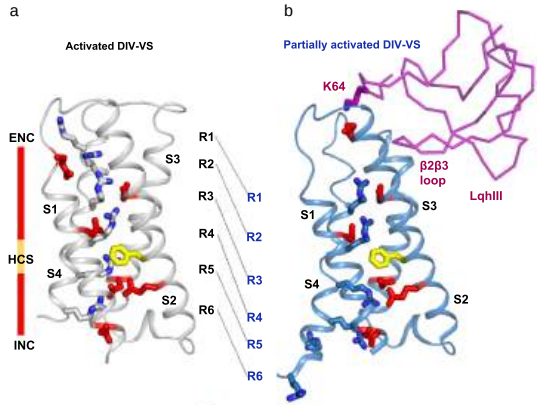
Answer

.

Voltage-Gated Sodium Channel.6.1: Comparison of Activated DomIV-Voltage sensor (VS) with Toxin-bound Partially activated DomIV-VS

Now we'll try to understand Figure 4, Conformational Change of DIV-VS, from the paper and pay special attention to the section of the text, "An intermediate-activated state of DIV-VS trapped by LqhIII" - Let's dissect Figure 4 A and B.

Figure 4a/4b from the paper below shows the c**onformational change of *Domain* IV-Voltage sensor (DIV-VS)** comparing the activated and partially activated state with the bound toxin.



Panel a shows the activated domain IV with key Arg side chains in S4. Panel B shows the partially activated domain IV with the same key Arg side chains in S4. The bound LqhIII is shown as a purple chain. In panels a and b, the:

1. activated Nav1.5 DIV-VS (the voltage sensing domain) is in grey (fig 4a)
2. the intermediate-activated Nav1.5DIV-VS is in blue (Fig 4b)
3. side chains of gating charges of Arg are shown in grey and blue sticks in 4a and in shades of blue sticks in 4b. Side chains in the ENC are shown in red, in the HCS in yellow, and in the INC in red;
4. the shift of each gating charge was indicated by black dashed lines between the structures in panels a and b.

The shift from the cytoplasmic to extracellular parts of the channel is shown in the region between the two panels. The black Rs in the activated DIV-VS(panel A) are further up in the diagram (towards the extracellular region) and further down in the partially activated DIV-VS bound to the toxin.

? Exercise *[Math Processing Error]*

Locate the 6 arginines? (R1-R6) with the blue indicating the N atoms in the positively charged Arg side chain of S4 in Domain IV in one of the iCn3D models above. What do the following abbreviations mean? ENC, HCS, and INC.

Answer

.

? Exercise *[Math Processing Error]*

From Figure 4a to 4b, explain from an electrostatic viewpoint how the movement of the Args towards the extracellular region would promote the movement of Na^+ ions inward. Explain how the movement of Na^+ ions would be diminished in the presence of the toxin.

Answer

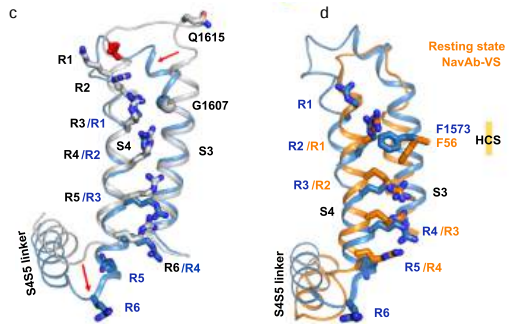
.

Voltage-Gated Sodium Channel.6.2: Comparison of Active and Intermediate-Activated, and Intermediate/Resting state

Now consider Figures 4 C and D from the paper below:

4c: Superposition of NaV1.5 *D*IV-VS between the fully activated state and toxin-bound intermediate-activated state. Red arrows indicate the conformational changes.

4d: Superposition of the intermediate-activated NaV1.5 *D*IV-VS and resting-state NaVAb-VS



? Exercise *[Math Processing Error]*

a. Locate the region in Figure c above that shifts the most from the fully activated to intermediate-activated state of the DIV-VS.

b. Describe the difference shown in Figure d between the intermediate-activated *D*IV-VS structure (blue) upon the resting state Na_VAb-VS structure (orange)

c. What do these differences imply about the conformational states of the apo and toxin-bound channel?

Answer

.

Voltage-Gated Sodium Channel.7: Summary

? Exercise *[Math Processing Error]*

Why might the mode of action be specific for cardiac muscle cells as compared to other toxins that act on sodium channels in skeletal and nerve cells?

Answer

.

After this guided research literature module, you can hopefully better understand the findings in the paper which are summarized in this abstract:

"Voltage-gated sodium (Na_V) channels initiate action potentials in excitable cells, and their function is altered by potent gating-modifier toxins. The α-toxin LqhIII from the deathstalker scorpion inhibits fast inactivation of cardiac $Na_V1.5$ channels with IC_{50} = 11.4 nM. Here we reveal the structure of LqhIII bound to $Na_V1.5$ at 3.3 Å resolution by cryo-EM. LqhIII anchors on top of voltage-sensing domain IV, wedged between the S1-S2 and S3-S4 linkers, which traps the gating charges of the S4 segment in a unique intermediate-activated state stabilized by four ion-pairs. This conformational change is propagated inward to weaken binding of the fast inactivation gate and favor opening the activation gate. However, these changes do not permit Na^+ permeation, revealing why LqhIII slows inactivation of Na_V channels but does not open them. Our results provide important insights into the structural basis for gating-modifier toxin binding, voltage-sensor trapping, and fast inactivation of Na_V channels."

Voltage-Gated Sodium Channel.8: Extensions

1. Interesting sidebar: https://www.sciencedirect.com/science/article/pii/S0021925819308300?via%3Dihub

Chlorotoxin (Cltx) is a 36-amino acid peptide that was originally isolated from *Leiurus quinquestriatus* venom (14) and has been shown to inhibit small conductance Cl− channels in colonic epithelial cells (14, 15). Cltx also inhibits Cl− fluxes across glioma membranes (13, 16). Immunohistochemical studies show that Cltx specifically and selectively binds to glioma cells (17) and radiolabeled Cltx targets tumor cells in mice bearing xenografted glioma tumors. Glioma cell migration and invasion into fetal brain aggregates is significantly reduced by Cltx (13). A recent survey of over 200 tissue biopsies from patients with various malignancies suggests that Cltx binds to the surface of gliomas and other embryologically related tumors of neuroectodermal origin (18) but not to normal brain.

2. Deathstalker scorpion venom also contains chlorotoxin - This is a very interesting story.

https://www.acs.org/molecule-of-the-week/archive/c/chlorotoxin.html#:~:text=Strichartz%20at%20Harvard%20Medical%20School,diagnosing%20and%20treating%20some%20cancers.

Note and remember LqhIII is an alpha toxin

3. https://www.sciencedaily.com/releases/2010/08/100811125947.htm

4. How much is deathstalker venom worth?

- 94$ for 5 ug agitoxin; https://www.scbt.com/p/agitoxin-78207-24-6
- https://news.stanford.edu/2019/06/10/healing-compounds-scorpion-venom/ : “For the past 45 years, Possani has focused on identifying compounds with pharmacological potential in scorpion venom. His group has previously uncovered potent antibiotics, insecticides and anti-malarial agents hidden in the arachnid’s poison.”
- https://www.timesnownews.com/viral/reason-why-scorpion-venom-is-one-of-the-most-expensive-liquids-in-the-world-will-surprise-you-article-99409440 : $39 million dollars a gallon
- The venom of the deathstalker scorpion, one of the most dangerous scorpions on the planet, costs $39 million dollars a gallon, making it the most expensive liquid on Earth.Apr 11, 2023

5. Looks like a great review article: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7277529/

6. Scorpion Venom: Detriments and Benefits

- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7277529/pdf/biomedicines-08-00118.pdf
- https://hal.science/hal-03616273/document
- https://www.rcsb.org/structure/1bmr
- https://www.researchgate.net/publication/308013755_Scorpion-Toxins_Lqh_III

7. Lookfor possible therapeutic potential here https://www.venomdoc.com/

8. Very cool venom graphic - https://www.venomdoc.com/new-page-2

9. The Toxicogenomic Multiverse: Convergent Recruitment of Proteins Into Animal Venoms: https://static1.squarespace.com/static/55a239e2e4b0b3a7ae106f25/t/59814ceae6f2e10bc7ada5f1/1501646072821/2009_Fry_Toxicogenomic_multiverse.pdf

10. The deathstalker scorpion venom alone has been found to have several different kinds of toxins including chlorotoxin (inhibit chloride channels), charybdotoxin (inhibit potassium channels), and agitoxins (affect sodium channels).. https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/leiurus-quinquestriatus. Chlorotoxin was found to selectively bind to glioma cells and serve as a marker for glioblastoma. https://www.acs.org/molecule-of-the-week/archive/c/chlorotoxin.html#:~:text=Strichartz%20at%20Harvard%20Medical%20School,diagnosing%20and%20treating%20some%20cancers. This feature of the scorpion venom was developed by J.M. Olson at Fred Hutchinson Cancer Center (Seattle) as a product called Tumor Paint https://www.fredhutch.org/en/news/center-news/2014/09/tumor-paint-US-trial.html

Membrane Structure, Diffusion and Transport Problems

Metabolism - Carbohydrate Problems

Metabolism - Lipids: Problems

test pGW

Molecular Modeling Questions

CHAPTER OVERVIEW

Signal Transduction Problems

NMDA Receptor - Under Construction
Ras Questions

NMDA Receptor - Under Construction

Under Construction

Glutamatergic transmission has been implicated in the pathophysiology of PTSD, particularly in the effects of N-methyl-D-aspartate receptor (NMDAR) signaling on the synaptic plasticity underlying learning and memory [13]. NMDARs comprise two GluN1 subunits and two GluN2 (A-D) or GluN3 (A, B) subunits. In adult forebrain regions, GluN2A and GluN2B are the main subunits forming receptor complexes with GluN1 at excitatory synapses. GluN2B-containing NMDARs play a preferential role in inducing synaptic plasticity, which is critical for the extinction of fear memories [14, 15]. Systemic injection of GluN2B-specific NMDAR antagonists ((RS)-3-(2-carboxypiperazin-4-yl)-propyl-1-phosphonic acid, ifenprodil) can impair the retention of fear extinction learning. GluN2B-containing NMDARs in both the amygdala and medial prefrontal cortex (mPFC) are also involved in reducing fear during extinction, whereas GluN2A-containing NMDARs play a greater role in the initial formation and/or stabilization of learned fear [15]. Rodent studies demonstrate that GluN2B subunit-containing NMDARs play pivotal roles in fear extinction learning.

---

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4263351/"tetrameric complexes mainly composed of NMDA receptor GluNR1 and GluNR2 subunits with the NR2 subunits modifying the activity of the receptor. "

N-methyl-d-aspartate (NMDA) receptors are Hebbian-like coincidence detectors, requiring binding of glycine and glutamate in combination with the relief of voltage-dependent magnesium block to open an ion conductive pore across the membrane bilayer.

NMDA receptors are Hebbian-like coincidence detectors, requiring the binding of glycine and glutamate to GluN1 and GluN2 subunits[9], respectively, combined with membrane depolarization to relieve magnesium block [10, 11]. Activation of the receptor opens a cation-selective, calcium permeable channel, thus causing further depolarization of the cell membrane and influx of calcium [12]. NMDA receptors are obligatory heterotetrameric assemblies[13, 14], typically composed of two glycine-binding GluN1 subunits and two glutamate-binding GluN2A-D subunits, with the GluN1/GluN2A/GluN2B complex the predominate receptor at hippocampal synapses[15]. Glycine-and d-serine-binding GluN3 subunits are additional subunits, expressed throughout the nervous system but with roles less well defined in comparison to the GluN1/GluN2 assemblies. A hallmark of NMDA receptors, by contrast with AMPA and kainate receptors, is a wide spectrum of allosteric modulation, from nanomolar concentrations of zinc, to the small molecule ifenprodil, polyamines and protons[16] and to voltage-dependent ion channel block by MK-801, ketamine and memantine[17].

The coordinates and structure factors for the structure have been deposited in the Protein Data Bank under accession code 4TLL and 4TLM for Structure 1 and Structure 2, respectively.

These receptors have two GluN1 subunits and two GluN2 (A-D) or GluN3 (A, B) subunits. In the forebrain, GluN2A and B form complex with GluN1 at synapses, with the B subunit playing a role in synaptic plasticity. Synaptic plasticity is necessary to remove "hard-wired" fear circuits. A goal of PTSD therapies is the **extinction** of the previously acquired feared memories through learning. Learning, and more specifically extinction, requires synaptic plasticity.

If one goal of PTSD treatment is the extinction of fear memories, then drugs that target GluN2B are potentially useful. Antagonist (such asd ifenprodil) of GluN2B seem to decrease the ability to extinguish fear retention memories. In other words, the learning and synaptic plasticity need to attenuate the fear memories are inhibited by the antagonist. GluN2B in the NMDAR receptors in the amygdala and medial prefrontal cortex appear to be involved in reducing feat during extinction. IN contact, GluN2A seem to be an important role in forming and stabilizing the learned fear response.

"Glutamatergic transmission has been implicated in the pathophysiology of PTSD, particularly in the effects of N-methyl-D-aspartate receptor (NMDAR) signaling on the synaptic plasticity underlying learning and memory [13]. NMDARs comprise two GluN1 subunits and two GluN2 (A-D) or GluN3 (A, B) subunits. In adult forebrain regions, GluN2A and GluN2B are the main subunits forming receptor complexes with GluN1 at excitatory synapses. GluN2B-containing NMDARs play a preferential role in inducing synaptic plasticity, which is critical for the extinction of fear memories [14, 15]. Systemic injection of GluN2B-specific NMDAR antagonists ((RS)-3-(2-carboxypiperazin-4-yl)-propyl-1-phosphonic acid, ifenprodil) can impair the retention of fear extinction learning. GluN2B-containing NMDARs in both the amygdala and medial prefrontal cortex (mPFC) are also involved in

reducing fear during extinction, whereas GluN2A-containing NMDARs play a greater role in the initial formation and/or stabilization of learned fear [15]. Rodent studies demonstrate that GluN2B subunit-containing NMDARs play pivotal roles in fear extinction learning."

Describe how the NMDA receptor functions, and how it implements the Hebbian model of learning at the synaptic level.

http://charlesfrye.github.io/Foundat...roscience//29/

https://www.pnas.org/doi/full/10.1073/pnas.95.12.7145

https://pubmed.ncbi.nlm.nih.gov/15888440/

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4263351/

CC *** https://www.nature.com/articles/s41380-022-01498-7

Ras Questions

Introduction Questions

1. Where is Ras located in the cell? What causes Ras to be localized here?
2. How does Ras get activated?
3. What are the three dimensional differences between GDP and GTP bound Ras. Why is the GTP bound form “active?”
4. What does Ras bound to GTP bind to? What are the net effects of this in terms of signal transduction?
5. What reaction does Ras catalyze as an enzyme?
6. What type of reaction is a GAP catalyzing? GEF?
    a. Are they opposite reactions? What is different (enzymatic control)
7. GAP is an acronym for GTPase activating protein. What GTPase is being activated?
8. Would RAS “turn off” without a GAP present? Why is this critical?
9. What molecule (GTP or GDP) would be bound to RAS if there were no GAPs or GEFs present? Why?
10. Why does a GAP increase the enzymatic activity RAS? What does a GAP provide that aids in the chemistry of the reaction?

Phos

https://www.ncbi.nlm.nih.gov/pmc/art...%20Ras%20cycle.

Paper about targeting phosphorylation and cancer therapy: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8642438/

1. Ras can be phosphorylated at Tyr 32 and Tyr64 by Src kinase, which leads to inhibition of binding to Raf and increased GTP hydrolysis. Provide a rationale for why this post-translational modification plays a role in the overall function of Ras. (Not sure what to ask here- there seems to be a number of kinases involved that do different things as explained in the paper above)
2. There are mutations in Ras, including Gly 12, Gly13, and Gln 61 which impair GTPase activity and GAP-mediated GTP hydrolysis. Predict what changes this would cause in the cell.
3. Think about designing a small molecule drug that affects the Ras signaling pathway and treats cancer. What proteins could you target? For each protein target, should the drug increase or decrease the activity of its target? Explain your answers.

Fig 1

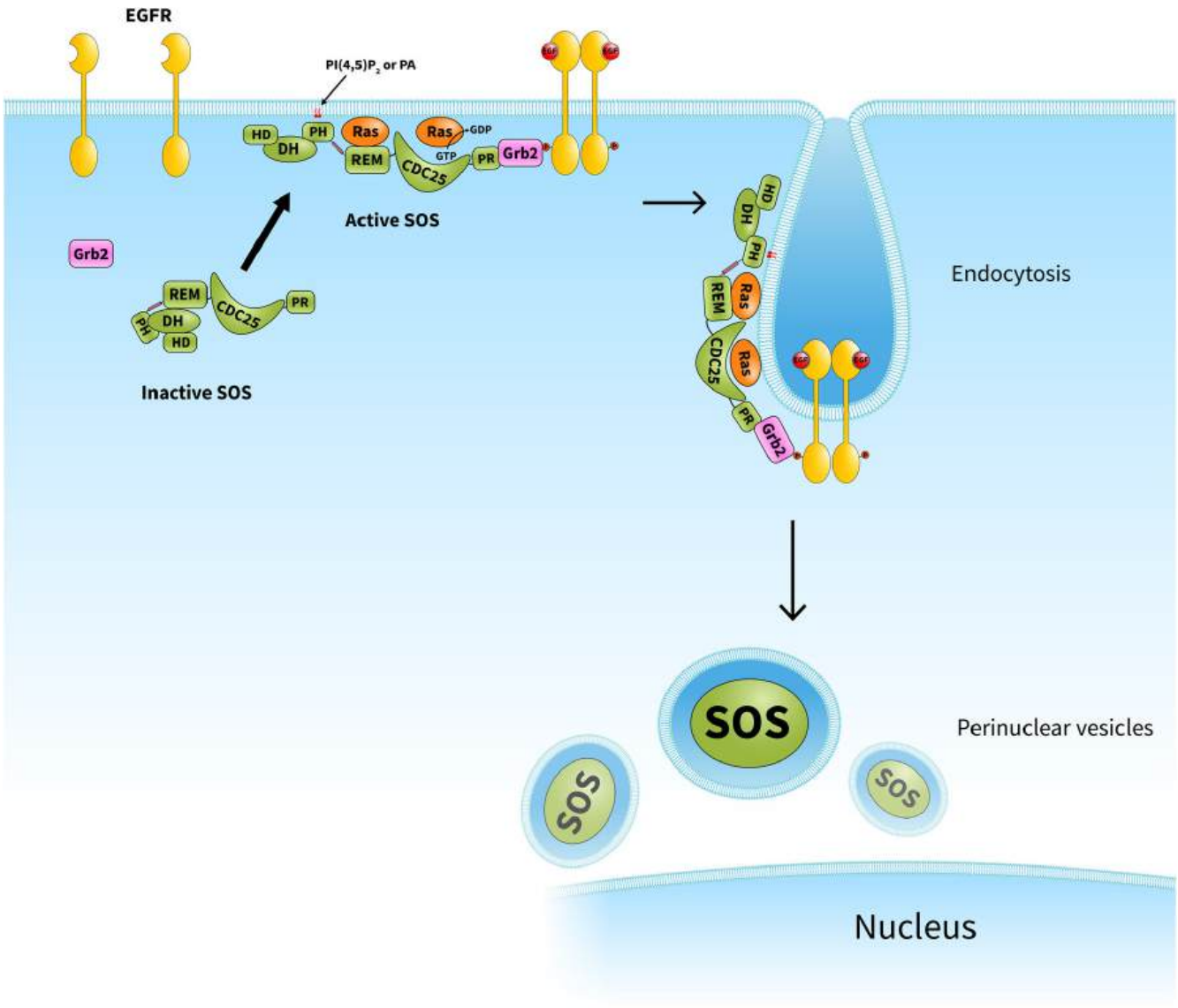
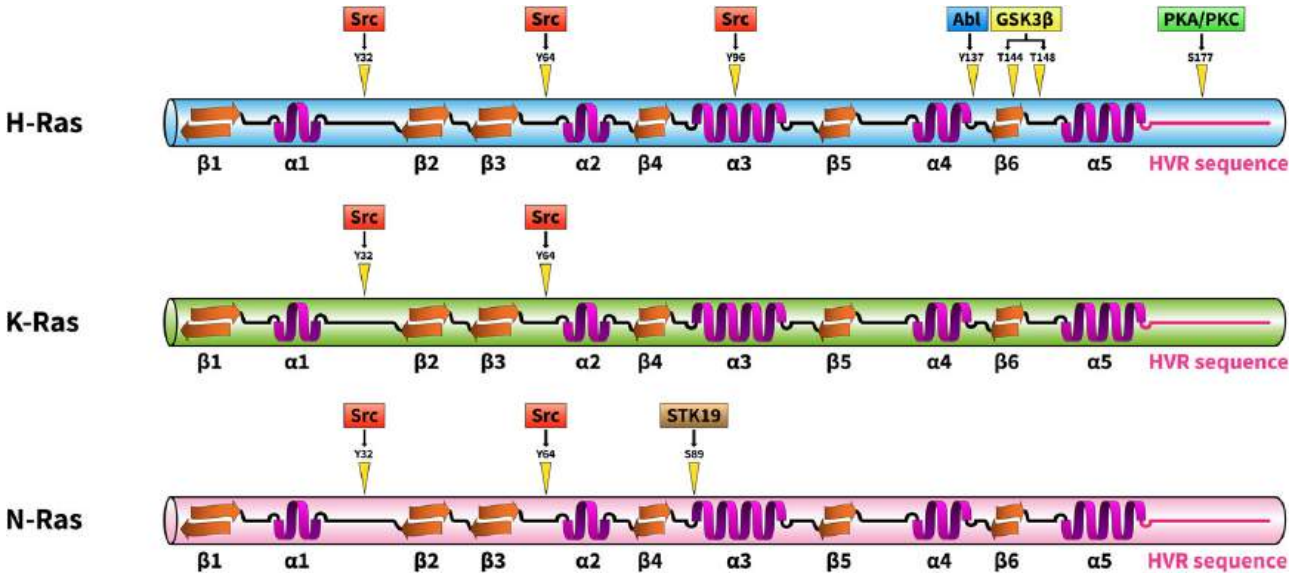
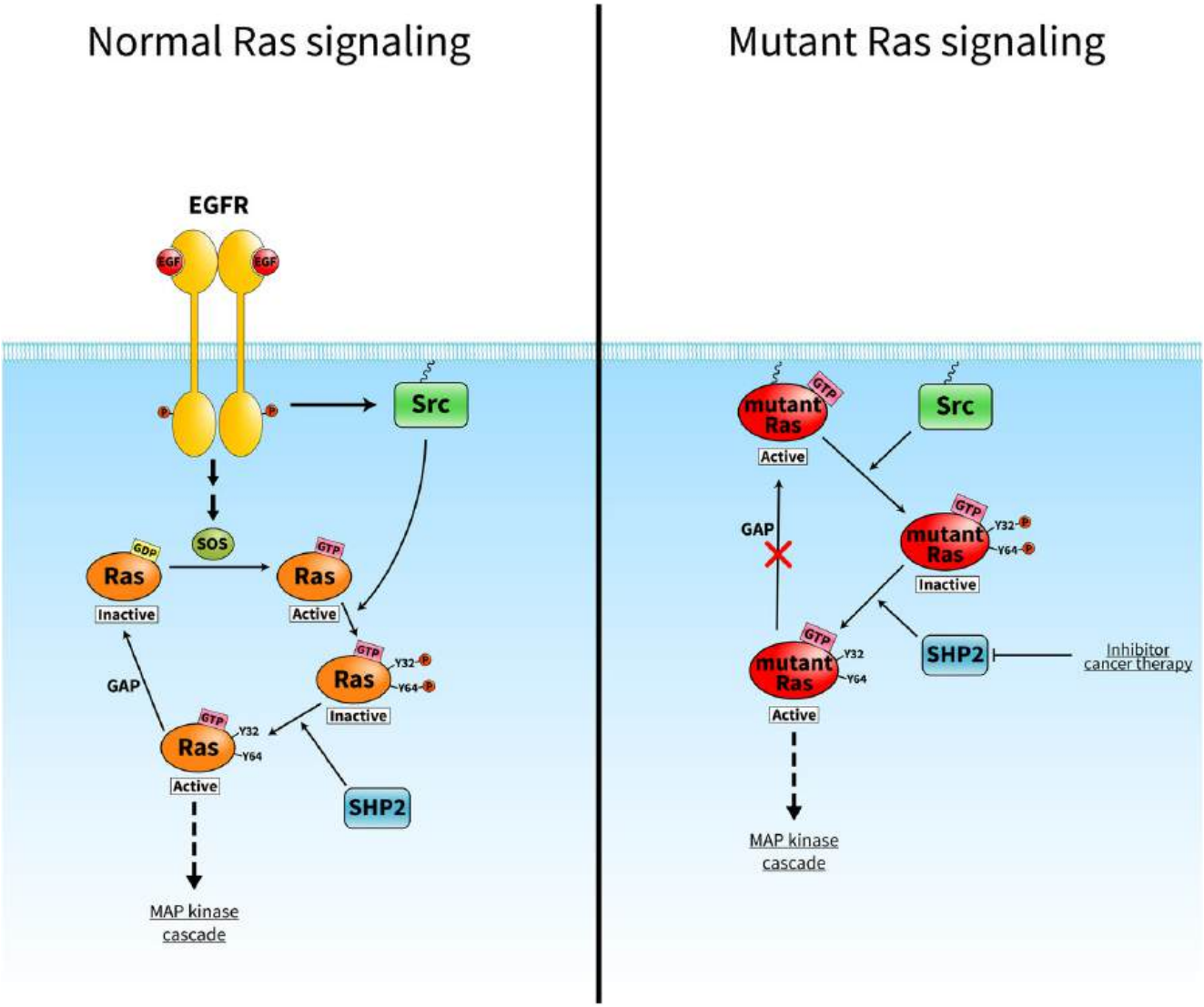

EGFR
PI(4,5)P2 or PA
HD
PH
DH
Ras
REM
Ras
GDP
GTP
CDC25
PR
Grb2
Active SOS
Grb2
REM
CDC25
PR
PH
DH
HD
Inactive SOS
Endocytosis
Perinuclear vesicles
SOS
SOS
SOS
Nucleus

Fig 2


Src
Y32
Src
Y64
Src
Y96
Abl
Y137
GSK3β
T144
T148
PKA/PKC
S177
H-Ras
β1
α1
β2
β3
α2
β4
α3
β5
α4
β6
α5
HVR sequence
Src
Y32
Src
Y64
K-Ras
β1
α1
β2
β3
α2
β4
α3
β5
α4
β6
α5
HVR sequence
Src
Y32
Src
Y64
STK19
S89
N-Ras
β1
α1
β2
β3
α2
β4
α3
β5
α4
β6
α5
HVR sequence

Fiog 3



Structure/Function - Carbohydrate Problems

Structure/Function - Lipid Problems

Structure/Function - Nucleic Acid Problems

CHAPTER OVERVIEW

Structure/Function - Protein Problems

Topic hierarchy

Disulfide Bonds
LGA: Voltage-Gated Sodium Channel - Students 082423
LGA: Protein Stability - Carbonic Anhydase

Disulfide Bonds

Disulfide engineering

The first set of questions below are based on this reference as noted: Craig, D.B., Dombkowski, A.A. Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins. *BMC Bioinformatics* **14**, 346 (2013). https://doi.org/10.1186/1471-2105-14-346. Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0)

Several computational programs have been developed to determine amino acid pairs that could be mutated to high-temperature stabilizing disulfide bonds. The stabilizing effects appear largest when the disulfide bond is made within the largest (and most flexible) loops (between 25-75 residues). These loops also had the highest residue B-factors.

In selecting pairs to form engineered disulfide, not only proximity (distance) but also geometry (torsion angles) of the resulting disulfide bond are important. We saw this previously in the energy analysis of butane rotamers, as illustrated in Figure 1 below.

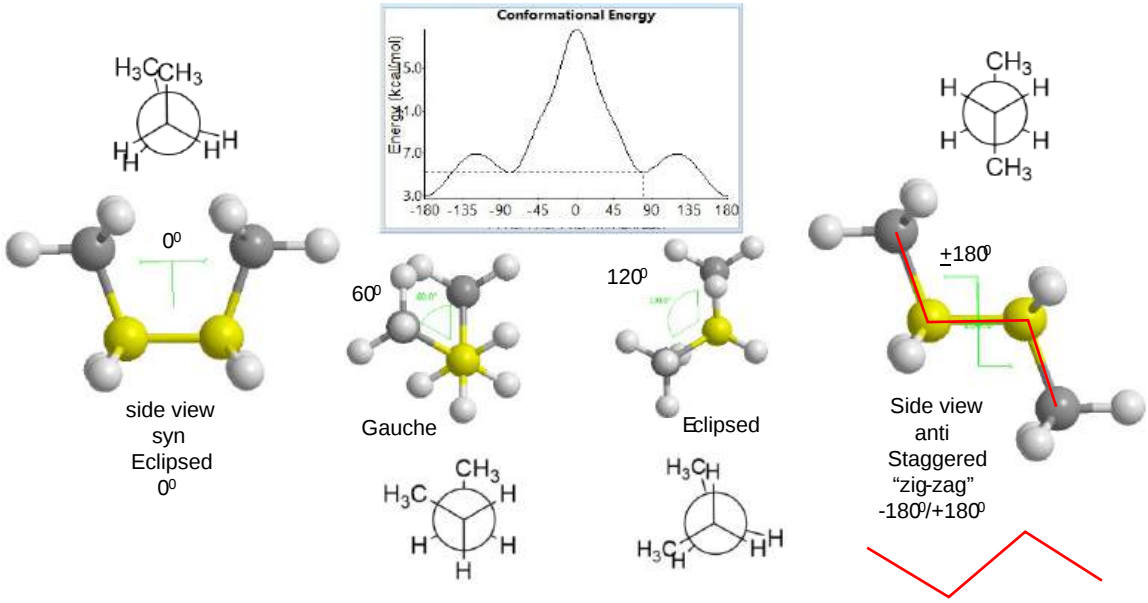


Figure 1: Newman projections for butane

Programs to determine amino acid pairs to mutate for disulfide bond formation test S-S bond torsional stability by determining the torsion angle χ_3 for the S-S bond. Evaluation of a database of many native proteins shows the χ_3 angle are centered in two major peaks at -87 and +97 degrees, as shown in Figure 2 below

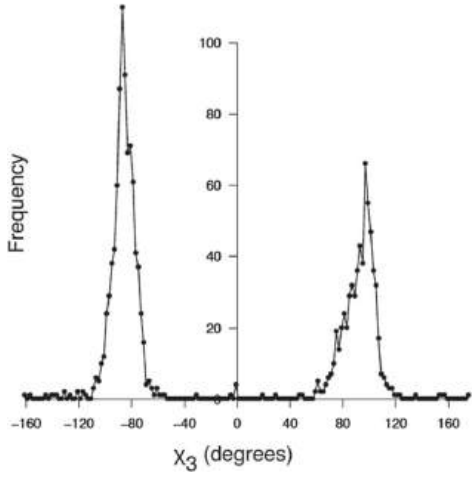


Figure 2: **Distribution of χ_3 torsion angles observed in 1505 native disulfide bonds found in 331 PDB protein structures.** Peaks occur at -87 and +97 degrees. Craig, D.B., Dombkowski, A.A. Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins. *BMC Bioinformatics* **14**, 346 (2013). https://doi.org/10.1186/1471-2105-14-346. Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0),

? Question 1

Does your calculated value of χ_3 for the native disulfide in carbonic anhydrase from Neisseria gonorrhea follow the observed angles?

Answer

At about 90^o so yes it does.

An empirical energy equation was developed to determine the E vs χ_3 dihedral angle for disulfide bonds. The equation is shown below (Craig, D.B., Dombkowski, A.A. ibid.).

$$E(\chi_3) = 4.0\,[1 - \cos(1.957\,[\chi_3 + 87])] \tag{1}$$

A graph of the equation made with Excel is shown in Figure 3 below

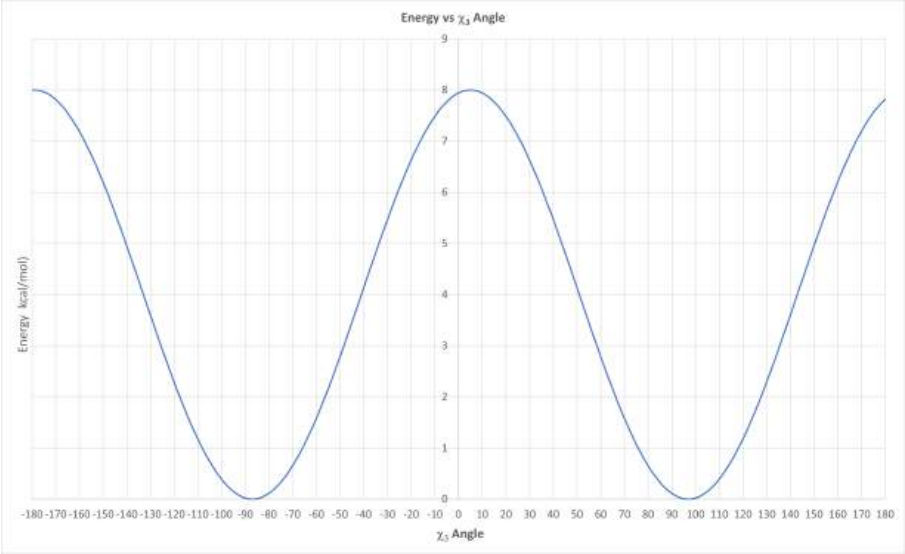


Figure 3: E vs χ_3 dihedral angle for disulfide bonds from empirical equation

Figure 4 below shows the distribution of energy values in the 1505 native disulfide bonds using our updated function. The study by Craig and Dombkowski showed that almost all (90%) of disulfides in native proteins in the PDB have an energy < 2.2 kcal/mol, so this metric could be used to determine possible disulfide bond pairs created by mutagenesis.

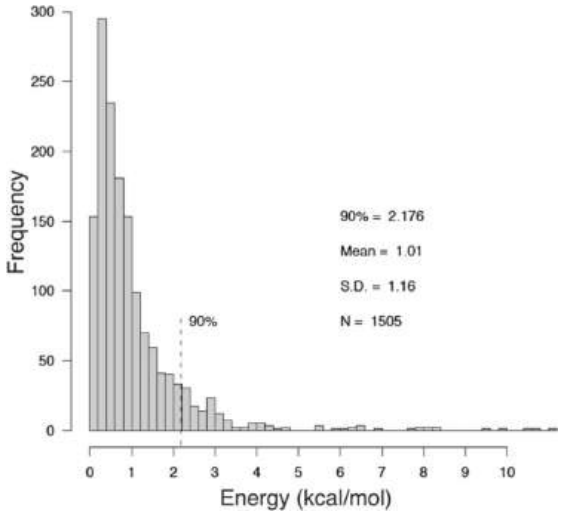


Figure 4: **Distribution of the disulfide bond energy calculated for 1505 native disulfide bonds.** The mean value is 1.0 kcal/mol, and the 90th percentile is 2.2 kcal/mol. Craig, D.B., Dombkowski, A.A. ibid.

? Question 2

You calculated the approximate value for χ_3 dihedral for the C28-C181 native disulfide bond form in Carbonic anhydrase from Neisseria gonorrhea using iCn3D in Question x above. Determine the approximate energy from the empirical function graph in Figure x above for that χ_3 dihedral. What percent of native disulfide bonds have that particular calculated energy in the 1505 native disulfide surveyed?

Answer

The estimated χ_3 dihedral was 90^o. The energy for that χ_3 angle would be approximately $\leq$0.1kcal/mol, which reflects the energies of about 150/1505 or 10% of the disulfides in the database. It is hence in the most stable range of disulfides, based only on the χ_3 dihedral angle.

Yet there are other important parameters as well that would affect the energy of the disulfide bond in the protein. Figure 5 below shows a comparison of native residue B-factors in stabilizing and destabilizing engineered disulfide bonds

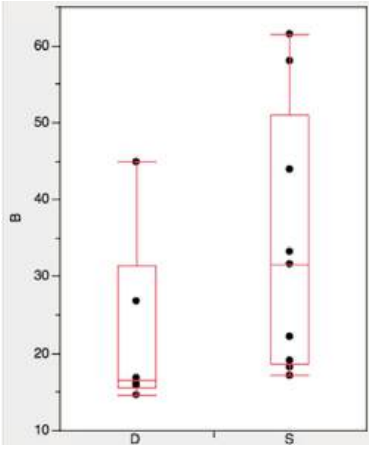


Figure 5: **Comparison of native residue B-factors in stabilizing and destabilizing engineered disulfide bonds.** Analysis of previously reported proteins with stabilizing (S) or destabilizing (D) engineered disulfide bonds. The mean B-factor for residues in proteins with stabilizing disulfide bonds was 31.6 compared with 16.5 for those involved in destabilizing bonds. Craig, D.B., Dombkowski, A.A. ibid.

? Question 3

Which proteins were more stabilized by engineered disulfide, those with higher or low B factors.

Answer

Proteins with engineered disulfides that increase stability have higher B-factors. This makes sense in that proteins with higher B-factors and hence mobility would be predicted to alter conformation and potentially denature more readily.

Prediction of disulfide bond engineering sites using a machine learning method

This set of questions based on this reference: Gao, X., Dong, X., Li, X. *et al.* Prediction of disulfide bond engineering sites using a machine learning method. *Sci Rep* **10**, 10330 (2020). https://doi.org/10.1038/s41598-020-67230-z. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/

The amount of data present in a single PDB file is very large, but it is nothing compared to the collective data in all PDB files. If only we could extract empirical rules from the collective PDB files that govern disulfide bond formation. It turns out we can with machine learning and artificial intelligence that can be used to develop and train predictive algorithms. Machine learning has been used to predict amino acid pairs for cysteine mutations to form engineered disulfide bonds. It recognizes 99% of natural disulfide bonds. residues. It uses these parameters:

- distances between the alpha-carbons and the beta-carbons of the bonded cysteine residues
- three torsion angles around the disulfide bonds (χ^1,χ^{ss},$\chi^{1'}$).

An example of one variable that helps define the stability of disulfide bonds is the distances between the Cα atoms for the disulfide-bonded cysteines, as shown in Figure 6 below.

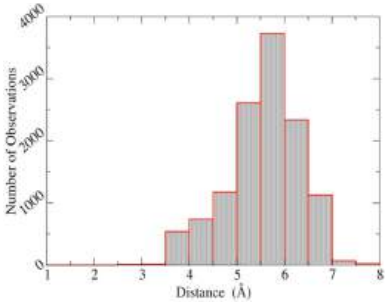
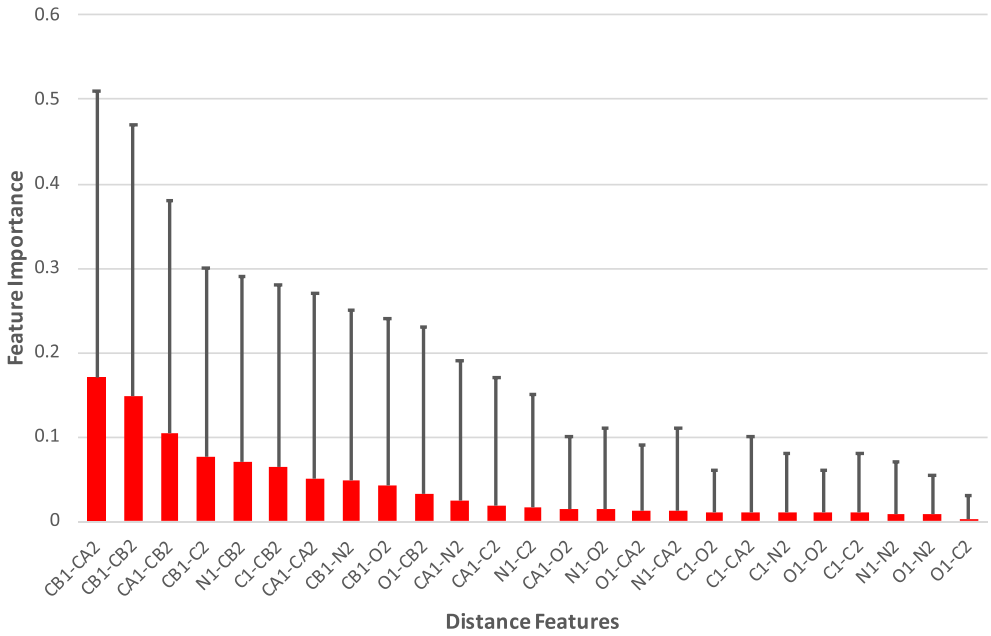


Figure 6: The histogram of distances between Cα atoms of disulfide-bonded cysteines. Gao, X., Dong, X., Li, X. *et al.* Prediction of disulfide bond engineering sites using a machine learning method. *Sci Rep* **10**, 10330 (2020). https://doi.org/10.1038/s41598-020-67230-z. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/

The distances range from 3.0 Å and 7.5 Å. Knowing this would constrain the number of choices for paired amino acid side chains for mutagenesis to produce disulfide.

Machine learning can also be used to find other distance constraints to optimize mutagensis experiments. Figure 7 below shows a graph of 10 different distances and their relative importance in determining disulfide bond stability.



An interactive iCn3D model of carbonic anhydrase from Neisseria gonorrhea (1KOQ) highlighting two pairs of amino acids identified by machine learning as candidates for mutations to disulfide-bonded cysteines is shown in Figure 8 below.

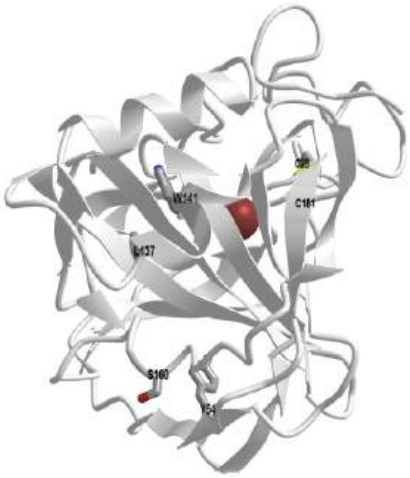


Figure 8: Carbonic anhydrase from Neisseria gonorrhea (1KOQ) highlighting two pairs of amino acids identified by machine learning as candidates for mutations to disulfide-bonded cysteines (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...ZH6Tv1urwkKhz8

Question 4

This question addresses the Biomolecular Visualization Framework theme(s) Atomic Geometry (AG), Macromolecular Building Blocks (MB)

Using the data in Figure 25 and measurements made using the iCn3D model above (Figure 25) , determine which pair would be most likely engineered into a disulfide bond. Complete the table below with the distances you made using iCn3D.

iCn3D instructions: Open the external link and follow these instructions

Trackpad and Mouse Controls

•rotate: click and drag (mouse: left click and drag)

•zoom: pinch and spread (mouse: rotate the scroll wheel)

•translate: two-finger click and drag (mouse: right click and drag)

•Re-center: left click View from the top menu bar, then select “Center Selection”

•Note: ctrl-click on a PC = command-click on Mac; alt-click on PC = option click on Mac

Instructions

1. Zoom to clearly see the amino acid pair for distance measures
2. From the top menu bar, choose **Analysis**, **Distance**, distance between 2 atoms
3. Pick the appropriate 2 atoms for measure distance by holding down the Alt key and selecting both
4. Record the distances in the table below.

Mutation Pair	CB1-CA2 (Å)	CB1-CB2 (Å)	CA1-CB2 (Å)
1 (L137) - 2 (W141)			
1 (Y54) - 1 (S160)			

Answer

Mutation Pair	CB1-CA2 (Å)	CB1-CB2 (Å)	CA1-CB2 (Å)
1 (L137) - 2 (W141)	6	5.2	5.3
1 (Y54) - 1 (S160)	4.1	4.1	5.0

LGA: Voltage-Gated Sodium Channel - Students 082423

Written by Henry Jakubowski, Emily Schmitt Lavin, Arthur Sikora, and Subhasish Chatterjee

LGA.1: Introduction

Eukaryotic voltage-gated sodium (Na_V) channels generate and sustain action potentials in nerve and muscle cells by moving Na^+ ions from the outside to the inside of the cell. This increases and makes positive the transmembrane potential of the cell, which at rest is approximately -70 mV (more negative inside). Once activated, the channel undergoes a fast inactivation (1-2 ms), without which the firing of nerves and muscles becomes dysregulated, a potentially lethal effect. Please view the information in Chapter 11.3 on the voltage-gated sodium channel before you do these guided assessment activities.

The questions below are derived from a paper from Jiang et al. on the structure and properties of the **α-scorpion toxin LqhIII** (MW 7,000) bound to rat cardiac sodium channel **Na_V1.5** (MW 227,000) by. (Jiang, D., Tonggu, L., Gamal El-Din, T.M. *et al*. Structural basis for voltage-sensor trapping of the cardiac sodium channel by a deathstalker scorpion toxin. *Nat Commun* **12**, 128 (2021). https://doi.org/10.1038/s41467-020-20078-3. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/)

The **Lqh** toxin is made by **L**eiurus **q**uinquestriatus ***h****ebraeus*. It is found in North Africa, the Middle East, and Western India and is shown below.

htt.://en.wikipedia.org/wiki/Deathstalker.

The study shows how the deathstalker scorpion (LqhIII) toxin inhibits the fast inactivation of cardiac sodium channels (Nav1.5). In other words, you will see how the toxin keeps the channel open longer than it would be open in its absence.

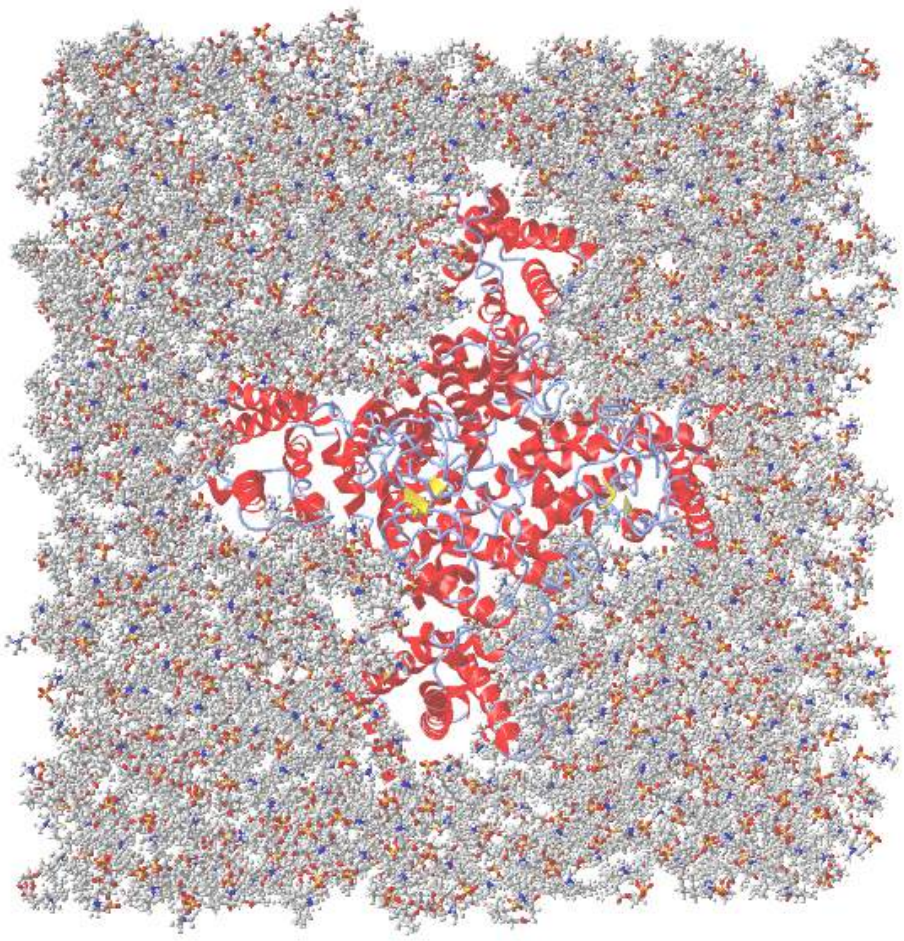
- In the absence of toxin, the sodium channel **Na_V1.5** returns in 1-2 ms to an inactive state when an 'inactivation gate" moves to occlude the open pore.
- The α-scorpion toxin LqhIII inhibits the return of the channel to the inactive state. Since the toxin inhibits the channel's fast inactivation of Na^+ ion flow into the cell, the channel stays open longer.

The toxin leads to the inhibition of the normal fast inhibition (inactivation) of the channel. Hence the channel stays open (activated) longer. This is analogous to the statement that the enemy of my enemy is my friend!

LGA.2: Techniques to study the Na Channel

For the experiments described in the paper, the rat sodium cardiac channel $Na_V1.5$ was purified, its structure determined, and its functional properties (the regulated movement of Na^+ to the inside of the cell - electrophysiology) in human epithelial cells measured. The protein is expressed in rat cardiac cells and is found in the cell membrane of the cell. The iCn3D image below shows the alpha chain of rat cardiac $Na_V1.5$ (6UZ3) embedded in a simple bilayer (DMPC) to model how it might appear in a cardiac cell membrane bilayer.

Data from
https://api.libretexts.org/endpoint/bounce/https://bio.libretexts.org/@api/deki/files/75773/6



Feedback

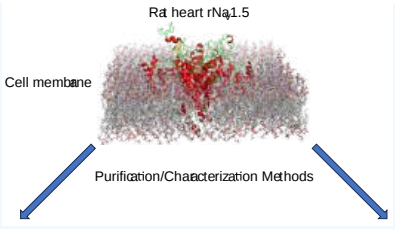
```
> load url
https://api.libretexts.org/endpoint/bounce/https://bio.libretexts.org/@api/deki/files/75778/6UZ3_
NaV1pt5inDMPCCharmm.pdb | type pdb
>
```

PDB coordinates based on S. Jo, T. Kim, V.G. Iyer, and W. Im (2008). CHARMM-GUI: A Web-based Graphical User Interface for CHARMM. *J. Comput. Chem. 29:1859-1865*. S. Jo, T. Kim, and W. Im (2007) Automated Builder and Database of Protein/Membrane Complexes for Molecular Dynamics Simulations. *PLoS ONE 2(9):e880*

? Exercise *LGA.1*

Complete the flow chart below to show two different approaches that could be used to purify the protein and prepare it for structural and functional studies. On the left show steps you could use to directly purify the protein from rat hearts. On the right use the rat gene (Scn5a) for the channel as the starting point for purification. Here are some links for review if needed.

- Chapter 3.4: Protein Purification
- Overview of Protein Expression
- Recombinant protein expression



Answer

From heart tissue:

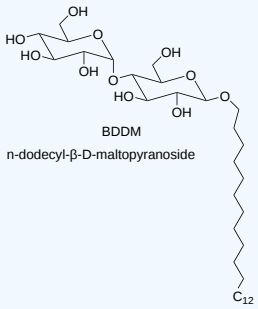
From heart tissue DNA

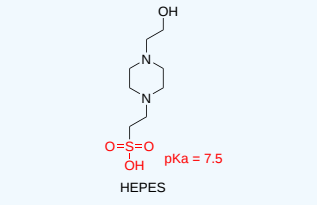
Here are some review links:

- Overview of Protein Expression
- Recombinant protein expression

Exercise $LGA.2$

Key components of the buffer solution used to purify the channel include HEPES and 1% (w/v) n-dodecyl-β-D-maltopyranoside. Their structures are shown below. Describe the role of each.





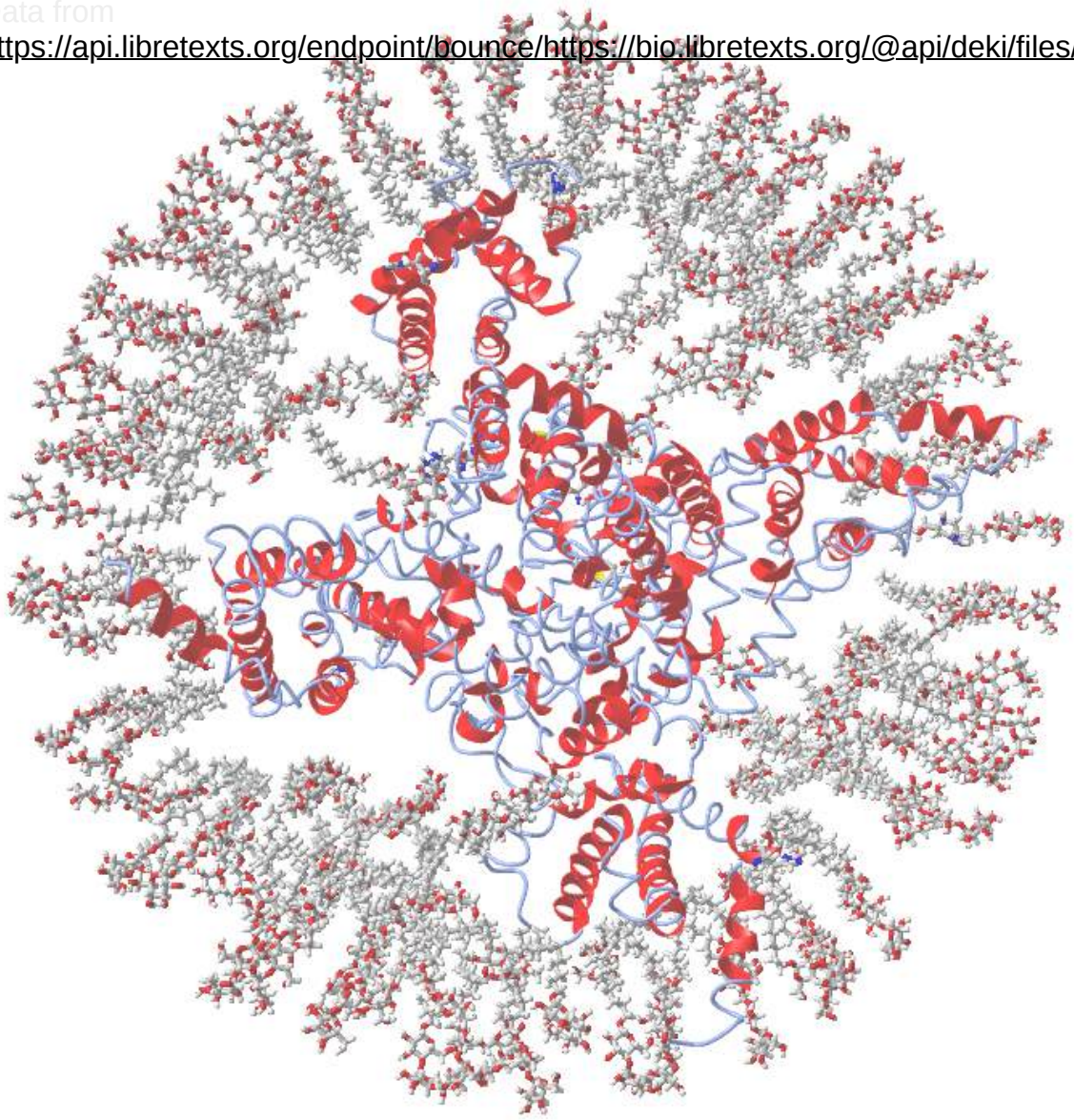
Answer

.

Exercise $LGA.3$

The following iCn3D shows the interaction of the purified sodium channel with n-**d**o**d**ecyl-**β**-D-**m**alto**p**yranoside (BDDM). Explain the differences between the iCn3D models showing the protein in a bilayer and interacting with the BDDM.

Data from
https://api.libretexts.org/endpoint/bounce/https://bio.libretexts.org/@api/deki/files/75768/6



```
> load url
https://api.libretexts.org/endpoint/bounce/https://bio.libretexts.org/@api/deki/files/75768/6uz3_
ratcardNaVinMicellePymolCIFDiscPDB.pdb | type pdb
>
```

Answer

.

? Exercise $LGA.4$

The elution of the complex on a size exclusion column (Panel A) and the analyses for the eluted fractions by SDS-PAGE (Panel B) are shown in the figure below. The α-scorpion toxin LqhIII: rat cardiac sodium channel $Na_V1.5$ complex elutes in the area of the first peak shown in blue.

a. How do molecules separate on size exclusion chromatography?

b. Compare the molecular weights of the first peak to the second complex peak in Panel A.

c. Which band(s) in Panel B likely represent the rat cardiac sodium channel $Na_V1.5$ based on the intensity of the stained band? The toxin LqhIII is the lowest band. (bands at 17.5K and 12.5 are FGF12b and calmodulin, respectively, which were added to stabilize the channel**)**

d. Why do the channel and toxin elute together in the size exclusion column shown in Panel A, but are separate bands in PAGE gel in Panel B?

e. How could you get the complex to separate as two peaks, the free $Na_V1.5$ channel, and the free α-scorpion toxin LqhIII?

f. To get information on the receptor, go to Uniprot and paste in rat cardiac sodium channel $Na_V1.5$ into the search box. Go to Sequence and Isoform in the left panel and find the actual MW. Knowing this, what are the major bands at about 170K and 60K?

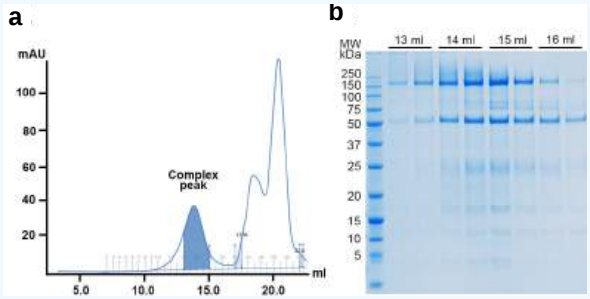


Figure: Purification of the **r**ecombinant NaV1.5C/LqhIII complex. a. Representative size-exclusion chromatography profile of purified **r**Nav1.5C/LqhIII. Peak fractions collected for cryo-EM grid preparation are shown in blue. b. SDS-PAGE of the size exclusion peak fractions stained by Coomassie blue.

Answer

.

CryoEM was used to determine the structure of the sodium channel:toxin complex. Here is a short YouTube video that describes the technique. Also review the appropriate part of Chapter 3.3: Analyses and structural predictions of protein structure.

? Exercise $LGA.5$

Describe the temperature conditions for protein samples in cryoEM. What is the reported resolution of cryo EM structure? How does this compare to X-ray structures? What are some advantages of using cryoEM over X-ray crystallography and NMR to determine the structure of proteins?

Answer

.

Molecular dynamics was also used to probe the conformational changes in the structure of the complex on the picosecond (10^{-12} s) to nanosecond (10^{-9} s) time scale. For a review of molecular dynamics, see Chapter 3.3: Analyses and structural predictions of protein structure. It can be used to probe dynamic changes in protein structure which cryoEM can't.

? Exercise $LGA.6$

Answers these multiple choice questions (created by AIPDF through ChatGPT4 -paid version using this prompt: Write 5 question for a biochemistry major about the use of molecular dynamics and the finding in the paper)

1. What was one of the primary uses of molecular dynamics in this research?
   - A) Predicting the behavior of NaV channels without toxins.
   - B) Analyzing hydration and Na^+ permeation through the rNaV1.5C/LqhIII complex.
   - C) Studying the interaction between different toxins.
   - D) Predicting the behavior of potassium channels.

2. In the molecular dynamics simulation analysis, what was aligned to the initial position for each snapshot?
   - A) The α-toxin LqhIII.
   - B) The voltage-sensing domain IV.
   - C) The Cα atoms from pore transmembrane helices.

- D) The fast inactivation gate.

3. Approximately how long were the unrestrained "production" simulations generated?
   - A) 10.35 ns.
   - B) 5000 steps.
   - C) 300 ns.
   - D) 2 fs.

4. Based on the molecular dynamics analyses, what was observed about the activation gate structure of the rNaV1.5C/LqhIII complex?
   - A) It was fully open for Na^+ conductance.
   - B) It was functionally closed for Na^+ conductance.
   - C) It was in a metastable state.
   - D) It showed no significant change from the rNaV1.5C structure.

Answer

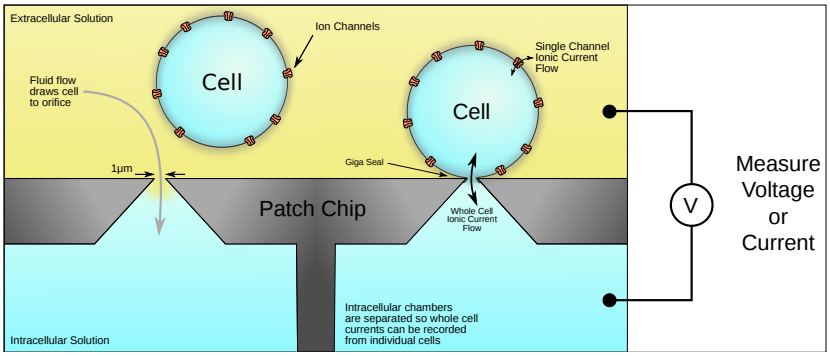
.

The authors used two types of electrophysiological techniques, patch clamp, and voltage clamp. Here is some brief background.

In a whole-cell patch clamp experiment, a pipet is placed on a cell, and suction is applied until a tight seal, indicated by a sharp rise in electrical resistance (gigaohm level) is made. This is illustrated in the figure below.

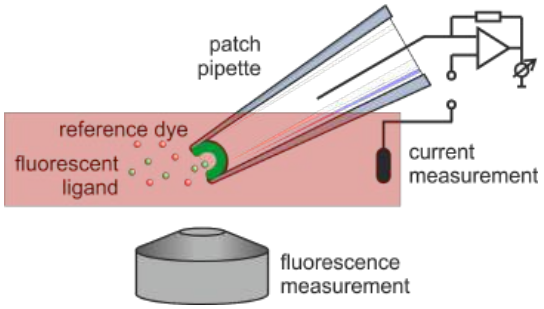
Patch Clamp Resistance. Formation of gigaseal. Holst. https://en.wikipedia.org/wiki/Automa..._Animation.gif. CC BY-SA 3.0

The cell can then be connected to a patch clamp chip in such a way that transmembrane potential or current can be measured on single-channel ion flow. This is illustrated in the figure below.



Holst. Patch Clamp Chip. Batch clamp chip showing a gigaseal, whole-cell recording configuration, and the ion channel and whole cell current. https://en.wikipedia.org/wiki/Automated_patch_clamp#/media/File:Patch_Clamp_Chip.svg. CC BY-SA 3.0

In patch-clamp fluorometry, part of the cell membrane is sucked into the tip with the seal intact. Fluorescent ligands can be applied to one side of the membrane that contains an ion channel and current measurements were made as illustrated in the figure below. Alternatively, as in this paper, side chains in the S4 voltage sensor were labeled with a fluorophore, and changes in fluorescence were observed with changes in membrane potential.



Exercise $LGA.7$

Answers these general multiple-choice questions (created by AIPDF through ChatGPT4 -paid version using this prompt: Write five multiple-choice questions about the use of patch clamp techniques to measure sodium currents in cells)

1. What is the primary purpose of the patch-clamp technique in cellular electrophysiology?
   - A) To visualize cell structures.
   - B) To measure the concentration of sodium ions inside cells.
   - C) To record ion currents across cell membranes.
   - D) To stimulate cellular growth.

3. In a typical neuron at resting potential (-70 mV) and in this study (epithelial cells transformed with the rat channel, what is the direction of the sodium current when sodium channels open?
   - A) Inward, into the cell.
   - B) Outward, out of the cell.
   - C) There is no movement of sodium.
   - D) Both inward and outward simultaneously.

4. Which of the following factors can influence the magnitude and direction of sodium currents measured using patch-clamp techniques?
   - A) The concentration of potassium ions outside the cell.
   - B) The voltage across the cell membrane.
   - C) The pH of the cell cytoplasm.
   - D) The size of the cell.

5. Why might a researcher use drugs or toxins during a patch-clamp experiment measuring sodium currents?
   - A) To increase the size of the cell.
   - B) To modulate or block sodium channels and observe the effects.
   - C) To change the color of the cell.
   - D) To stimulate cell division.

Answer

.

Exercise $LGA.8$

Answer these general multiple-choice questions about patch-clamp fluorometry. (created by AIPDF through ChatGPT4 -paid version using this prompt: write 5 multiple choice questions of patch clamp fluorometry in which key amino acids in a membrane protein are labeled with a fluorophore)

1. What is the primary advantage of combining patch-clamp with fluorometry in studying membrane proteins?
   - A) It allows simultaneous measurement of electrical activity and conformational changes.
   - B) It increases the fluorescence of all amino acids.
   - C) It enhances the electrical activity of the protein.
   - D) It allows visualization of the entire cell in detail.

2. Why are specific amino acids in a membrane protein labeled with a fluorophore in patch-clamp fluorometry?
   - A) To increase the size of the protein.
   - B) To change the electrical properties of the protein.
   - C) To detect specific conformational changes in the protein during activity.
   - D) To make the protein more soluble in water.

3. Which property of the fluorophore is crucial for patch-clamp fluorometry?
   - A) Its electrical charge.
   - B) Its sensitivity to changes in the local environment or protein conformation.
   - C) Its ability to increase protein activity.
   - D) Its color in visible light.

4. In which scenario would patch-clamp fluorometry be especially useful?
   - A) When studying the overall shape of a cell.
   - B) When investigating the relationship between ion channel gating and conformational changes.
   - C) When trying to increase the fluorescence of a solution.
   - D) When observing the movement of proteins inside the cell.

5. What is a critical consideration when choosing a fluorophore for labeling amino acids in patch-clamp fluorometry?
   - A) The taste of the fluorophore.
   - B) The electrical conductivity of the fluorophore.
   - C) The photostability and brightness of the fluorophore.
   - D) The size of the fluorophore molecule.

Answer

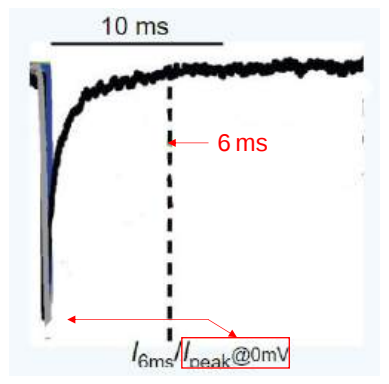
.

LGA.3: Nonstructural Lab Studies of LqhIII Toxin Effects on Rat Sodium Channel $Na_V1.5$ ($rNa_V1.5_C$)

HEK293S GnTI$^-$ (epithelial-like) cells were transformed with the rat cardiac sodium channel $Na_V1.5$ ($rNa_V1.5_C$). The cells were then studied in the absence and presence of the toxin at varying times after toxin addition and at various concentrations of the toxin. The opening and closing of the channel were determined by measuring changes in the Na^+ currents into the cell on channel opening.

? Exercise $LGA.9$

The resting potential of a cell is around -70 mV (more negative inside). When the transmembrane potential is depolarized by raising the transmembrane potential to around -55 mV or even more positive, the Na^+ channels are activated, and an inward Na^+ current (black line in a modified form of Figure 1a from the paper below) which goes **downward** by convention) through the channel occurs. This is followed by a quick inactivation of the channel and the return to the baseline flow of ions. In the experiment below, the potential was raised from -100 mV (channel closed) to 0 mV (channel open). What is happening to the $Na_V1.5$ Na^+ channel during these 10 ms? What is special about the current at 6 ms (indicated by the dashed vertical line)

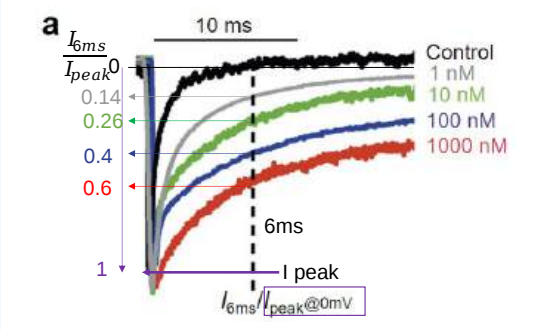


Answer

Initially, the change in voltage opens the $rNa_V1.5_C$ and allows an inward flow of Na^+ ions as evidenced by the vertical drop. As described in the introduction, conformational changes in the $Na_V1.5_C$ (closing of the inactivation loop), follow which closes the channel and stops the current, so the current returns to baseline within about 6 ms. The channel is in the inactive state by around 6 ms.

? Exercise $LGA.10$

Figure 1a from the paper (modified) below shows a series of lines of different colored (black to red) representing Na^+ currents obtained at 0 (black line) and increasing concentrations (gray through red) of the LqhIII scorpion toxin. Let's assume that the downward I_{peak} =1. The values of I_{6ms}/I_{peak}, calculated from the approximate values shown on the graph, are also shown on the vertical axis Does the toxin alter the immediate response of the cells after the channel was activated? What effect does increasing [toxin] have on the response of the cell? Offer a structural explanation of how the toxin affects the cell by suggesting changes in the toxin-bound structure.



Answer

The toxin at any dose does not substantially affect the initial opening of the channel since the size of the current does not change at first. Then the toxin inhibits the rapid inactivation of the channel within the first 6 ms, leading to a prolonged inward Na^+ current. The inhibition of the inactivation is dose-dependent on the concentration of the toxin. The toxin does not completely block the quick inactivation of the channel as the lines don't return to the black baseline in the time interval measured. This suggests that the toxin binds to the channel and either partially occludes it or prevents the normal conformational change in the $Na_V1.5_C$ that causes a quick deactivation of the channel.

Since Na^+ ions still flow through the channel but at a lower rate than the open state, the toxin-bound channel likely represents a 4th, or partially-opened state of the channel.

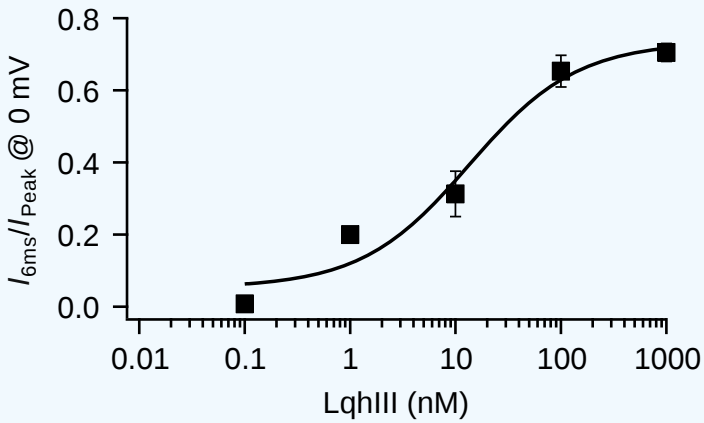
? Exercise $LGA.11$

Figure 1a (left) from the paper below shows the dependency of the inhibition of the quick inactivation of the channel on the log of the LqhIII concentration. When the transmembrane potential is set to 0 mV (as in this experiment), the channel should open and the current would be maximal. The data points in the graph are close to the ones estimated in the graph from the previous question.

a. In the absence of the toxin, what should the current I be at 6 ms compared to the maximal Na^+ current? That is, what would be the value of I_{6ms}/I_{peak}?

b. Is the channel completely inactivated in the presence of the toxin?

c. At 6 ms, what concentration of toxin (nM) causes 50% inhibition of the maximum effect of the inhibitor on the normal rapid inactivation of the channel?

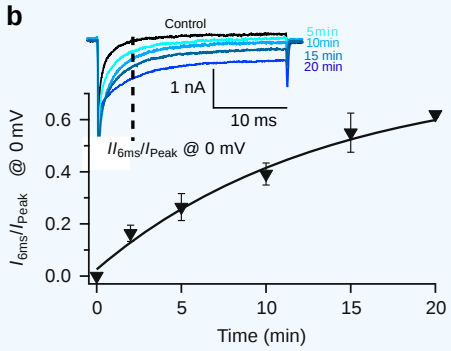


Answer

.

Exercise $LGA.12$

In the next experiment, cells were kept at -120 mV at one fixed concentration (100 nM) of toxin. The toxin was left to incubate with the cells for various times up to 20 min. After the incubation time, the transmembrane potential was changed to **0 mV** to activate the channel, and inward Na^+ currents were measured. The results are shown in the top inset graph in Figure 1b from the paper. (Note: It is unclear from the paper if the control was determined at 0 min with 100 nM toxin or no toxin.)



a. Did the length of time cells were pre-incubated with the 100 nM toxin affect Na^+ currents after depolarization of the cells? How did the effects on the cells depend on the preincubation time?

b. Describe and explain these results

Answer

.

Here is another interesting feature of toxin binding. The toxin binds to a site on the resting state of the $Na_V1.5_C$ with high affinity. When the cell becomes depolarized (made more + inside the cell), the affinity for the toxin decreases so it starts to dissociate. The affinity of the toxin for $Na_V1.5_C$ decreases with increasing + transmembrane potential. At very high positive potentials (+100 mV) it appears not to bind.

Exercise $LGA.13$

What might account for the decreasing affinity of the bound toxin for the $Na_V1.5_C$ with an increasing transmembrane potential?

Answer

.

Exercise $LGA.14$

Time course experiments were conducted on the complex at 100 nM of LqhIII scorpion toxin. A three-pulse protocol can be applied to alter membrane potentials:

- 1^{st}: a pulse from −120 mV to +100 mV for the indicated times then
- 2^{nd}: a 50-ms hyperpolarizing pulse (make membrane potential very negative, perhaps around -100 mV)
- 3^{rd}: a pulse of 50 ms to 0 mV

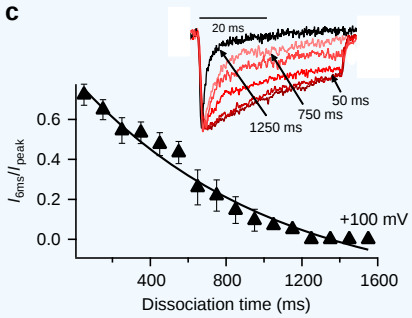
Note that steps 2 and 3 both occur with 0.1 s. What is the purpose of each pulse?

Answer

.

Exercise $LGA.15$

Figure 1c from the paper below shows results for a set of 3-pulse designed to allow recovery of the fast inactivation which was blocked by the previous toxin binding. Note that the transmembrane potential for the first pulse was +100 mV



a. Describe what happens to the channel and complex.

b. Explain the results

c. Summarize thermodynamic and kinetic features that make the toxin so effective.

Answer

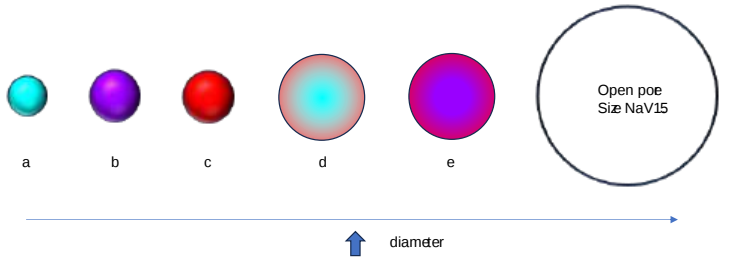
.

LGA.4: Structural Studies of LqhIII Toxin Effects on Rat Sodium Channel $Na_V1.5$ ($rNa_V1.5_C$) - CryoEM

Before we discuss in detail the structure of the sodium channel and its complex with the toxin, let's look at an important attribute of molecules that helps determine their function, the actual size of the species involved.

Exercise $LGA.16$

The Na^+ and the K^+ voltage-gated ion channels must have an open pore when the channel protein is active (Na^+ ions move across the channel). Extracellular Na^+ and the K^+ ions don't exist as "naked" ions but they are hydrated by water in an aqueous extracellular environment. The figure below shows the relative sizes of these Group I cations and their hydrated forms, in comparison to the diameter of the open NaV1.5 pore in the channel. Answer the following questions. The red sphere (c) represents the calculated value of the diameter of water assuming its volume when it is bound to a protein is 25 $Å^3$.



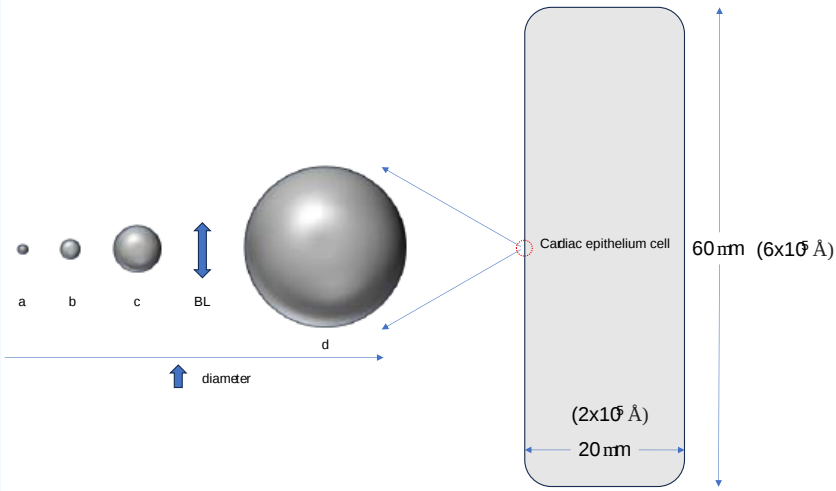
1. Which represents the "naked" (nonhydrated) size of the K^+ ion?
2. Which represents the hydrated Na^+ ion?
3. Based on the pore size alone, which of these species could diffuse through the pore?

Answer

.

Exercise $LGA.17$

The approximate relative sizes of the hydrated Na^+ ion, pore opening, toxin, and the NaV1.5 protein are shown in the figure below along with the relative width of the bilayer (BL). A cardiac epithelium cell is shown as a rectangle to the right. A red dot  (not visible in the large figure) in the membrane surrounded by the red-dotted circle represents a single NaV1.5 channel.



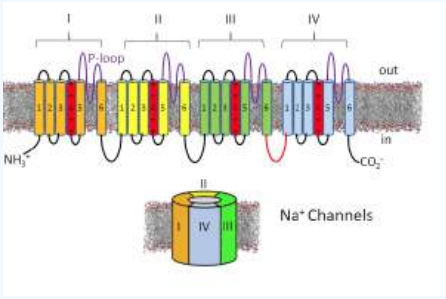
1. Which likely represents the pore?
2. Which represents NaV1.5 channel protein?
3. Which represents the toxin?

Answer

.

Exercise $LGA.18$

If you hadn't read the paper, where would you consider the most likely location for a toxin to bind to affect the function of Nav1.5? Circle the most likely toxin binding site in the schematic below. Based on the paper, where did it bind? Redraw the toxin in the correct location based on the paper.



Answer

.

Exercise $LGA.19$

How is this toxin (LqIII) different in terms of its binding location from other toxins that interfere with the functioning of Nav1.5? What is the effect of the toxin on the channel and the symptoms of this venom?

Answer

.

Exercise $LGA.20$

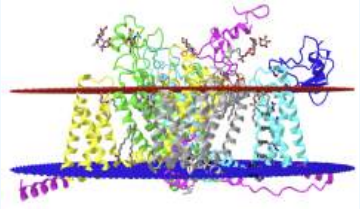
What amino acids should be present in the S4 segment of Nav1.4 and why?

Answer

.

Exercise $LGA.21$

The figure below shows an interactive iCn3D model of the rat sodium channel $Na_V1.5$ bound to the LqhIII toxin (7k18).



Rat sodium channel $Na_V1.5$ bound to the LqhIII toxin (7k18). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...jpz8CFQajMKJe7

Domains I-IV are shown in gray, yellow, green, and cyan, respectively. The magenta parts are just extensions/connectors of these domains. The LqhIII toxin is shown in dark blue. The Ile-Phe-Met (IFM) sequence of the inactivation gate that is involved in closing the pore in fast inactivation is shown as black spheres and labeled.

It is important to know that some stretches of amino acids are missing from the structure. These include AAs 1-120 (N-terminus), 298-301, 425-501, 602-608, and 1781-1838.

a. To which domain does the LqhIII toxin bind?

b. Is the binding site close to the pore-forming segments of the domain or the voltage-sensitive segments?

c. Does the layer of red spheres represent the outer (extracellular) or inner (intracellular) leaflet of the membrane?

d. Offer reasons that parts of the protein are missing from the structure.

Answer

.

Exercise $LGA.22$

The IFM motif has been shown to be conserved across all voltage-gated sodium channels.

a. What role does it play in these channels and in Nav1.5 and why?

b. The cause of Paroxysomal Extreme Pain Disorder (PEPD), an extremely rare disease with only 15 known affected families, appears to be mutations in the IFM motif which leads to increased sensations of pain. What is a likely effect of the mutations in the IFM motif?

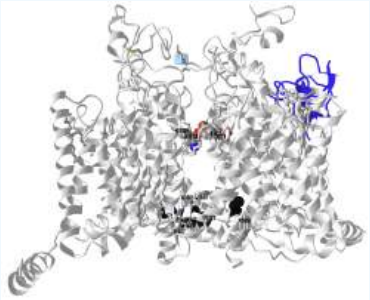
https://en.wikipedia.org/wiki/Paroxy..._pain_disorder

Answer

.

Exercise $LGA.23$

The figure below shows a different interactive iCn3D model of the rat sodium channel $Na_V1.5$ bound to the LqhIII toxin (7k18) without a membrane representation for clarity. It shows the selectivity filter DEKA (spacefill, CPK colors), the inactivation gate IFM and the IFM "internal receptor" F1651, L1660, and N1662 (spacefill, CPK colors), and a ring of hydrophobic residues V413, L941, I1471, and I1773 (spacefill, black) that in the closed state completely seal off the cytoplasmic opening in the pore.



Rat sodium channel $Na_V1.5$ bound to the LqhIII toxin without a membrane representation (7k18). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...pPzakHu6Ew6WH9

After viewing the structure from all angles, do you think that the toxin:$Na_V1.5$ complex looks closed, open, or inactivated form? Explain

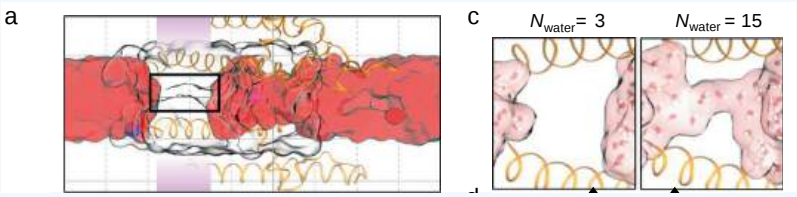
Answer

.

LGA.5: Structural Studies of LqhIII Toxin Effects on Rat Sodium Channel $Na_V1.5$ ($rNa_V1.5_C$) - Molecular Dynamic Simulations

Exercise $LGA.24$

Now let's look at some data to see if the pore is really open, partially open, or closed in the toxin:channel complex. One clue is if the structure shows water in the pore as the Na^+ ions must be hydrated to pass through the pore (the opposite case is seen with K^+ channels when K^+ pass through stripped of water). **Molecular dynamic (MD)** simulations were done on the $Na_V1.5$ with and without the toxin to simulate the environment in the channel opening. The results of the MD simulations are shown below in Figures 6 a and c from the paper.



Molecular dynamics analysis of hydration and Na^+ permeation through the $rNa_V1.5_C$/LqhIII complex.

Panel a shows a side view of $rNa_V1.5_C$ (orange ribbons; domains II and IV) from MD simulations highlighting Na^+ ions (blue spheres), the water-occupied volume within a cylinder of radius 8.5 Å (red surface), and the protein-occupied volume within a cylinder of radius 12 Å (colorless surface). The cavity within the pore is outlined with a black rectangle. The region of the intracellular activation gate is shown as a purple band.

Panel c shows molecular representations of the gate containing N_{water} = 3 (left) or 15 (right) water molecules

Based on these studies, do you believe the pore is closed, open, or partially open?

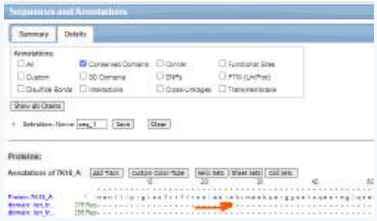
Answer

.

LGA.6: Detailed Structural Analyses of LqhIII Toxin Effects on Rat Sodium Channel $Na_V1.5$ ($rNa_V1.5_C$)

Exercise $LGA.25$

From the iCn3D model, write the sequence of the S4 segment that contains the Arg side chains and describe the properties of the amino acids in the sequence. Do this by scrolling along the sequence window in iCn3D (shown below) until you find the labeled Arg shown in the model.



Answer

.

Exercise $LGA.26$

Is the helix amphiphilic? That is, are the Arg side chains all on one face of the helix and the nonpolar on the other? To find out, copy and paste the sequence of S4 (above) in this helical wheel predictor and run the program.

Answer

.

Exercise $LGA.27$

Make a simplified view of the iCn3D model by hiding Domains 1-III to more readily see the contributions of S5 and S6 of Domain IV to the pore.

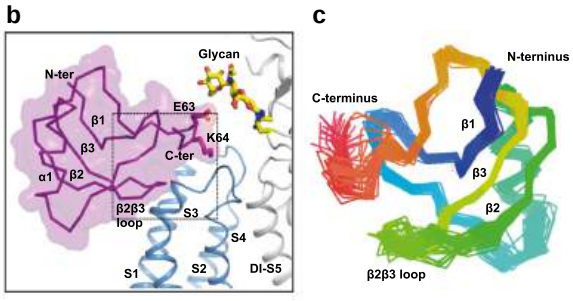
- open iCn3D and load 7K18.
- With your mouse or trackpad, choose **Sequence and Annotation** in the top menu bar
- Choose the Details tab
- Ctrl-Click the two sequences highlighted in yellow below for Domain IV and the toxin.
- Choose **View** from the top menu bar and then **View Selection**
- Choose **Style, Background, Transparent**

Solution

.

LGA.7: Noncovalent Interactions of LqhIII and Domain IV/VS

Now let's look at the actual interaction of the toxin with the Domain IV/VS of the channel. A closeup showing the interaction site is shown in Figures 3 b and c from the paper below. Panel C next to it shows the NMR-solution structures of the toxin in the absence of the channel. Each structure determined is represented by a single color line color codes red at the C-terminus to blue at the N-terminus.



Panel b: CryoEM structure of the rNa$_V$1.5$_C$ Domain IV/VS and LqhIII complex; Panel c: NMR structure of free LqhIII

? Exercise *LGA.28*

Using this iCn3D model, describe the secondary structure of the bound toxin. How many pairs of cysteine residues are in the LqhIII toxin. Identify which cysteines are involved in the disulfide bonds. What effect do the disulfide bonds have on the beta sheet structure?

Answer

.

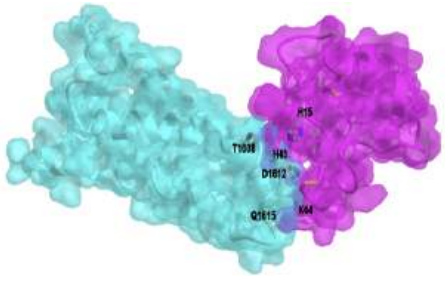
? Exercise *LGA.1*

What sections of the toxin in panel B make the closest interactions with the Domain IV/VS of the channel? Describe their conformation flexibility in the free toxin.

Answer

.

The figure below shows an interactive iCn3D model of a surface rending of Domain IV of the rat sodium channel Na$_V$1.5 bound to the LqhIII toxin (7k18).



Surface rending of Domain IV of the rat sodium channel $Na_V1.5$ bound to the LqhIII toxin (7k18). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...x5QRPEctT5zsW7.

The molecular surface and underlying secondary structure of the LqhIII toxin are shown in magenta, with key residues H15, H43, and K64 shown as CPK-colored sticks and labeled. Domain IV/VS of the channel is shown in cyan, with key amino acid side chains T1608, D1612, and Q1615 shown as colored and labeled sticks.

? Exercise *LGA*.29

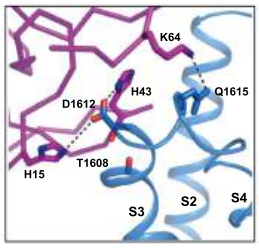
Comment on the shape and possible side chain interactions that contribute to high-affinity binding of the inhibitor to Domain IV/VS.

Answer

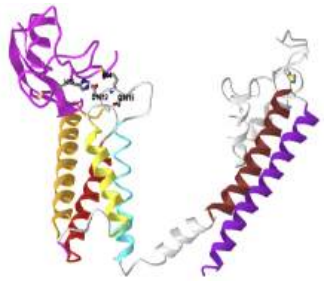
.

Figure 3d from the paper below shows the detailed interactions between LqhIII and *D*IV-VS. Key residues shown in sticks were labeled. Interaction surfaces of the *D*IV-VS (blue) and the LqhIII (purple). Key residues for the interaction are shown in yellow shading and embedded stick

d



The figure below shows an interactive iCn3D model of the rat sodium channel $Na_V1.5$ Domain IV bound to the LqhIII toxin (7k18).



Rat cardiac sodium channel NaV1.5C Domain IV/LqhIII toxin complex (7k18). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?JAjUQAZDZoQ5wiVX9

The toxin is shown in magenta. The segments are colored as follows: S1 is red, S2 is orange, S3 is yellow, S4 is cyan, S5 is brown and S6 is . Key amino acid pairs involved in the binding of the toxin to Domain IV are shown in sticks and labeled.

? Exercise *LGA*.30

In summary, name and locate the amino acid residues that serve the following roles in the LhqIII toxin: DIV-VS interactions.

1. Which amino acids in the toxin interact with D1612 (the paper describes the interaction as pincers surrounding D1612).
2. The conserved negatively charged residue in the Nav1.5 channel
3. What position is Thr in and what is thought to be its role in the mechanism?

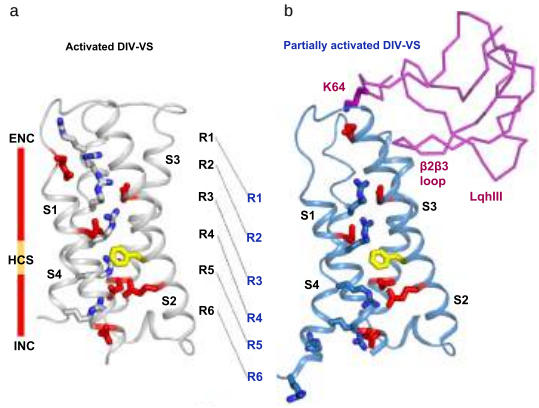
Answer

.

LGA.7.1: Comparison of Activated DomIV-Voltage sensor (VS) with Toxin-bound Partially activated DomIV-VS

Now we'll try to understand Figure 4, Conformational Change of DIV-VS, from the paper and pay special attention to the section of the text, "An intermediate-activated state of DIV-VS trapped by LqhIII" - Let's dissect Figure 4 A and B.

Figure 4a/4b from the paper below shows the c**onformational change of *Domain* IV-Voltage sensor (DIV-VS)** comparing the activated and partially activated state with the bound toxin.



Panel a shows the activated domain IV with key Arg side chains in S4. Panel B shows the partially activated domain IV with the same key Arg side chains in S4. The bound LqhIII is shown as a purple chain. In panels a and b, the:

1. activated Nav1.5 DIV-VS (the voltage sensing domain) is in grey (fig 4a)
2. the intermediate-activated Nav1.5DIV-VS is in blue (Fig 4b)
3. side chains of gating charges of Arg are shown in grey and blue sticks in 4a and in shades of blue sticks in 4b. Side chains in the ENC are shown in red, in the HCS in yellow, and in the INC in red;
4. the shift of each gating charge was indicated by black dashed lines between the structures in panels a and b.

The shift from the cytoplasmic to extracellular parts of the channel is shown in the region between the two panels. The black Rs in the activated DIV-VS(panel A) are further up in the diagram (towards the extracellular region) and further down in the partially activated DIV-VS bound to the toxin.

? Exercise *LGA.31*

Locate the 6 arginines? (R1-R6) with the blue indicating the N atoms in the positively charged Arg side chain of S4 in Domain IV in one of the iCn3D models above. What do the following abbreviations mean? ENC, HCS, and INC.

Answer

.

? Exercise *LGA.29*

From Figure 4a to 4b, explain from an electrostatic viewpoint how the movement of the Args towards the extracellular region would promote the movement of Na^+ ions inward. Explain how the movement of Na^+ ions would be diminished in the presence of the toxin.

Answer

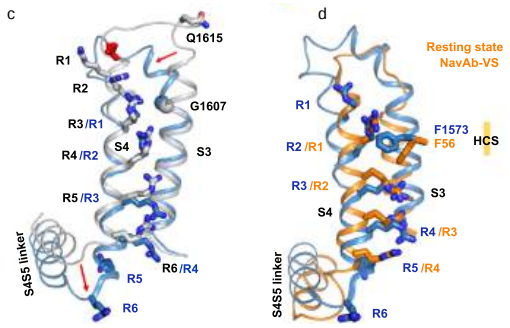
.

LGA.7.2: Comparison of Active and Intermediate-Activated, and Intermediate/Resting state

Now consider Figures 4 C and D from the paper below:

4c: Superposition of NaV1.5 *D*IV-VS between the fully activated state and toxin-bound intermediate-activated state. Red arrows indicate the conformational changes.

4d: Superposition of the intermediate-activated NaV1.5 *D*IV-VS and resting-state NaVAb-VS



? Exercise $LGA.32$

a. Locate the region in Figure c above that shifts the most from the fully activated to intermediate-activated state of the DIV-VS.

b. Describe the difference shown in Figure d between the intermediate-activated *D*IV-VS structure (blue) upon the resting state Na_VAb-VS structure (orange)

c. What do these differences imply about the conformational states of the apo and toxin-bound channel?

Answer

.

LGA.8: Summary

? Exercise $LGA.33$

Why might the mode of action be specific for cardiac muscle cells as compared to other toxins that act on sodium channels in skeletal and nerve cells?

Answer

.

After this guided research literature module, you can hopefully better understand the findings in the paper which are summarized in this abstract:

"Voltage-gated sodium (Na_V) channels initiate action potentials in excitable cells, and their function is altered by potent gating-modifier toxins. The α-toxin LqhIII from the deathstalker scorpion inhibits fast inactivation of cardiac $Na_V1.5$ channels with IC_{50} = 11.4 nM. Here we reveal the structure of LqhIII bound to $Na_V1.5$ at 3.3 Å resolution by cryo-EM. LqhIII anchors on top of voltage-sensing domain IV, wedged between the S1-S2 and S3-S4 linkers, which traps the gating charges of the S4 segment in a unique intermediate-activated state stabilized by four ion-pairs. This conformational change is propagated inward to weaken binding of the fast inactivation gate and favor opening the activation gate. However, these changes do not permit Na^+ permeation, revealing why LqhIII slows inactivation of Na_V channels but does not open them. Our results provide important insights into the structural basis for gating-modifier toxin binding, voltage-sensor trapping, and fast inactivation of Na_V channels."

LGA.9: Extensions

1. Interesting sidebar: https://www.sciencedirect.com/science/article/pii/S0021925819308300?via%3Dihub

Chlorotoxin (Cltx) is a 36-amino acid peptide that was originally isolated from *Leiurus quinquestriatus* venom (14) and has been shown to inhibit small conductance Cl− channels in colonic epithelial cells (14, 15). Cltx also inhibits Cl− fluxes across glioma membranes (13, 16). Immunohistochemical studies show that Cltx specifically and selectively binds to glioma cells (17) and radiolabeled Cltx targets tumor cells in mice bearing xenografted glioma tumors. Glioma cell migration and invasion into fetal brain aggregates is significantly reduced by Cltx (13). A recent survey of over 200 tissue biopsies from patients with various malignancies suggests that Cltx binds to the surface of gliomas and other embryologically related tumors of neuroectodermal origin (18) but not to normal brain.

2. Deathstalker scorpion venom also contains chlorotoxin - This is a very interesting story.

https://www.acs.org/molecule-of-the-week/archive/c/chlorotoxin.html#:~:text=Strichartz%20at%20Harvard%20Medical%20School,diagnosing%20and%20treating%20some%20cancers.

Note and remember LqhIII is an alpha toxin

3. https://www.sciencedaily.com/releases/2010/08/100811125947.htm
4. How much is deathstalker venom worth?
    - 94$ for 5 ug agitoxin; https://www.scbt.com/p/agitoxin-78207-24-6
    - https://news.stanford.edu/2019/06/10/healing-compounds-scorpion-venom/ : “For the past 45 years, Possani has focused on identifying compounds with pharmacological potential in scorpion venom. His group has previously uncovered potent antibiotics, insecticides and anti-malarial agents hidden in the arachnid’s poison.”
    - https://www.timesnownews.com/viral/reason-why-scorpion-venom-is-one-of-the-most-expensive-liquids-in-the-world-will-surprise-you-article-99409440 : $39 million dollars a gallon
    - The venom of the deathstalker scorpion, one of the most dangerous scorpions on the planet, costs $39 million dollars a gallon, making it the most expensive liquid on Earth.Apr 11, 2023
5. Looks like a great review article: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7277529/
6. Scorpion Venom: Detriments and Benefits

- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7277529/pdf/biomedicines-08-00118.pdf
- https://hal.science/hal-03616273/document
- https://www.rcsb.org/structure/1bmr
- https://www.researchgate.net/publication/308013755_Scorpion-Toxins_Lqh_III

7. Lookfor possible therapeutic potential here https://www.venomdoc.com/
8. Very cool venom graphic - https://www.venomdoc.com/new-page-2
9. The Toxicogenomic Multiverse: Convergent Recruitment of Proteins Into Animal Venoms: https://static1.squarespace.com/static/55a239e2e4b0b3a7ae106f25/t/59814ceae6f2e10bc7ada5f1/1501646072821/2009_Fry_Toxicogenomic_multiverse.pdf
10. The deathstalker scorpion venom alone has been found to have several different kinds of toxins including chlorotoxin (inhibit chloride channels), charybdotoxin (inhibit potassium channels), and agitoxins (affect sodium channels).. https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/leiurus-quinquestriatus. Chlorotoxin was found to selectively bind to glioma cells and serve as a marker for glioblastoma. https://www.acs.org/molecule-of-the-week/archive/c/chlorotoxin.html#:~:text=Strichartz%20at%20Harvard%20Medical%20School,diagnosing%20and%20treating%20some%20cancers. This feature of the scorpion venom was developed by J.M. Olson at Fred Hutchinson Cancer Center (Seattle) as a product called Tumor Paint https://www.fredhutch.org/en/news/center-news/2014/09/tumor-paint-US-trial.html

---

LGA: Protein Stability - Carbonic Anhydrase

Carbonic Anhydrase

Engineered Stability

We have already encountered this enzyme before (Chapter 6.1). It catalyzes the hydration of CO_2 (g) as shown below.

$$CO_2\ (g) + H_2O \leftrightarrow H_2CO_3\ (aq) \leftrightarrow HCO_3^-\ (aq) + H^+\ (aq)$$

It is among the fastest of all enzymes, with a k_{cat} of 10^6 s^{-1} and a k_{cat}/K_m of 8.3 x 10^7 $M^{-1}s^{-1}$ (reference). It is diffusion controlled in that the rate of diffusion of reactants and products, not the chemical steps, determine the reaction rate. It can convert 10^6 molecules of CO_2(g) to HCO_3^- each second. No wonder scientists and engineers are studying it to capture CO_2. It's a big challenge though to capture CO_2 released on combustion of coal or natural gas in a power plant. Here are two problems that must be overcome:

- The enzyme must be thermostable at elevated temperatures to capture the CO_2 found in high-temperature power plant emissions
- The enzyme is reversible so it will be inhibited by the product HCO_3^-
- The enzyme must be stable to somewhat alkaline conditions (pH of 0.1M $NaHCO_3$ = 8.3)

For carbon capture from fossil fuel emissions, CA is immobilized by surface adsorption, covalent attachment, encapsulation, and entanglement. Immobilized enzymes are typically more thermostable and can be used in flow-through as opposed to solution phase capture. The immobilized enzyme matrix must withstand high temperatures (up to 100°C, and alkaline solvents used to strip the matrix for reuse.

The enzyme is found throughout life and typically has an active site Zn^{2+}. There are 8 families, α, β, γ, δ, ζ η, θ, and ι, with the α family being the most abundant. The α forms are generally active as dimers, but can act as monomers and tetramers.. There are 15 isoforms of the α form in humans and have a prime role in pH regulation. They are found in bacteria, fungi, plants, and algae. β-CAs are found in some types of bacteria, archaea, fungi, some higher plants, and invertebrates. CA in chloroplasts (and mitochondria (algae) are involved in carbon fixation. We will focus our attention on engineering carbonic anhydrase to make them more thermostable, alkali insensitive, and less susceptible to product inhibition by bicarbonate.

Natural enzymes can be isolated and selected for thermal and alkali stability. In addition, new versions selected for these properties can be engineered using directed evolution or site-directed mutagenesis. You wish to increase the thermal stability of a protein using mutagenesis. Essentially you wish to perturb the equilibrium between the folded (native) protein and the unfolded (denatured) protein so as to preferentially stabilized the native state.

? Question *LGA.1*

Using mutagenesis, what residues might you change in a native protein to make it more stable at higher temperatures?

Answer

A characteristic of the native state of the protein is its conformational stability compared to the conformational flexibility of the many possible denatured states. In addition, the protein must undergo conformational changes as it unfolds. Hence anything that restricts conformational flexibility might preferentially stabilize the native state. These would include changing single or pairs of side chains to allow the formation of more salt bridges and intrachain disulfide bonds, as well as hydrogen bonds. Loops with greater flexibility, as determined by B-factors in the crystal structure files, or by molecular dynamic simulations, could be changed to contain a disulfide, which would clearly stabilize a flexible loop.

? Question *LGA.2*

What measurements would you make to quantitate the change in thermal stability?

Answer

Measures a signal that changes with increasing temperature. The signal can be enzyme activity, or more easily a spectroscopic signal such as absorbance at 280 nm or fluorescence as a function of temperature. Alternatively, the stability at room temperature could be measured using urea as a perturbant. These are discussed in Chapter 4.12.

The actual amino acid composition and more strangely specific dipeptide sequences within a sequence are associated with thermal stability of hyperthermophilic proteins. For example, proteins from two different types of archaea with different optimal growth temperatures show that the one with the higher growth temperature have significantly higher levels of VK, KI, YK, IK, KV, KY, and EV and decreased levels of DA, AD, TD, DD, DT, HD, DH, DR, and DG. Similar experiments have been done in bacterial cells. Using machine learning, the dipeptide sequences K*H, KR, TF, P*M, F**N, V**Y, MW, and WQ were important in themostability where the * denotes a gap in the residues.

Carbonic anhydrase from Neisseria gonorrhea (*ng*CA)

Data from: Jo, B., Park, T., Park, H. *et al.* Engineering *de novo* disulfide bond in bacterial α-type carbonic anhydrase for thermostable carbon sequestration. *Sci Rep* **6**, 29322 (2016). https://doi.org/10.1038/srep29322. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/

Now that we understand the general chemistry, structure, and reaction mechanism of carbon anhydrase (at least the alpha human CAII form), let's explore efforts to engineer more thermostable variants. One example is the carbonic anhydrase from N. gonorrheas. This CA has been used as a target for mutagenesis to increase thermal stability of the enzyme, through the introduction of new disulfide bonds.

Even though only about 35% of the amino acids are identical, the overall structures are similar. This is illustrated in Figure LGA. 10 below.

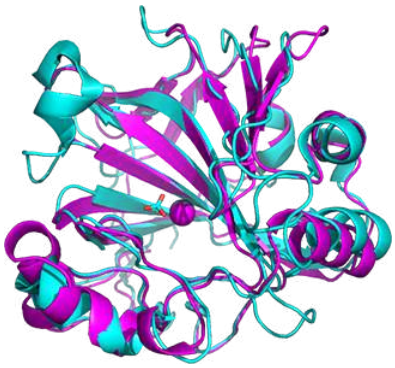


Figure LGA. 10: Alignment of the carbonic anhydrase from Neisseria gonorrhea (NG-CA) magenta,1KOQ) and human CA II (cyan, 2VVB)

The active site is mostly conserved compared to human CA II. The Zn^{2+} bound water has a pKa of around 6.5, compared to the value of 7.0 in human CA II. The hydrophobic patch (pocket) is similar, with Phe 93, Leu 153 and Tyr 72 in the NG-CA replacing Phe 95, Phe 176, Phe 70 in human CA II, respectively. The histidine ligands to Zn^{2+} are His92 (94), His94 (96), and His111 (119), where the numbers in parentheses represent Hu CA II. The proton removed from Zn^{2+} bound water is transferred to His 66 (64 in human CA II) and then to His 64.

The single disulfide bond between 181 and C28 is shown in Figure LGA. 1 below

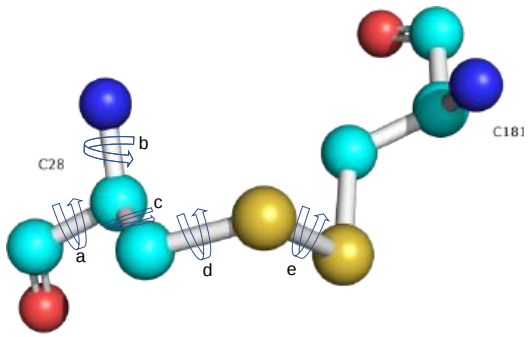


Figure LGA. 11: Single disulfide bond between 181 and C28 in wild type Carbonic anhydrase from Neisseria gonorrhea

Question $LGA.18$

This question addresses the Biomolecular Visualization Framework theme(s) Atomic Geometry (AG)

Identify the correct torsion angles in Figure $LGA.x$ above. Verbal definitions of torsional angles in a peptide chain are listed below. The successive atoms after the C_α leading away from the backbone atoms are X_β-X_γ-X_δ-X_ε (in that order). C is the backbone carbonyl C and N is the backbone nitrogen atom.

- phi (φ) is the angle of right-handed rotation around N-C_α bond. φ = 0 if the C_α-C bond is cis (eclipsed) to the C-N bond. Values range from -180 to 180 degrees.
- psi (ψ) is the angle of right-handed rotation around C_α -C bond. ψ = 0 if the C-N bond is cis (eclipsed) to the N-C_α bond. Values range from -180 to 180 degrees.
- chi_1 (χ_1) is the rotation around N-C_α-X_β-X_γ
- chi_2 (χ_2) is the rotation around C_α-X_β-X_γ-X_δ
- chi_3 (χ_3) is the rotation around X_β-X_γ-X_δ-X_ε

Answer

phi (Φ) = b, psi (Ψ) = a, chi_1 (χ_1) = c, chi_2 (χ_2) = d, and chi_3 (χ_3) = e

Question $LGA.19$

This question addresses the Biomolecular Visualization Framework theme(s) Atomic Geometry (AG), Topology and Connectivity (TC)

Figure $LGA.12$ below shows an interactive iCn3D model of the atoms within 4A of the disulfide bond in Carbonic anhydrase from Neisseria gonorrhea (1KOQ). Rotate the model to determine the approximate chi_3 (χ_3) dihedral angle. Hint: site down the S-S bond.

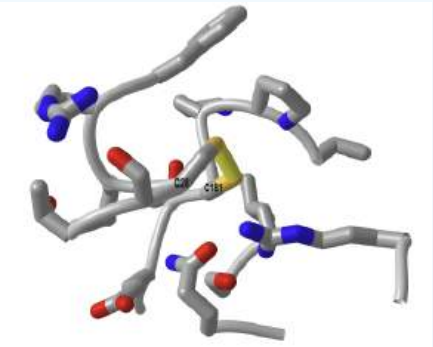
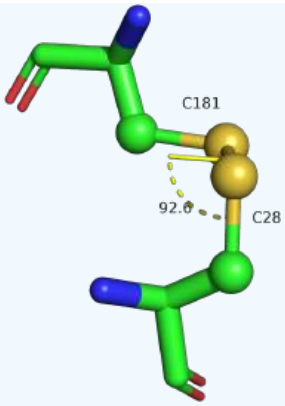


Figure $LGA.12$: Atoms within 4A of the disulfide bond in Carbonic anhydrase from Neisseria gonorrhea (1KOQ). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...nutjP6EL2PubRA

Answer

The visually estimated chi_3 (χ_3) angle for rotation around the S-S bond is $90^{o.}$ Here is the actual angle (image made with Pymol)



We mentioned previously that variants with higher thermal stability are likely to be more rigid and less flexible. Flexibility can be determined through analysis of molecular dynamic simulations and also by examining of the B factor values in PDB file. This number is a measure of the displacement of an atom from a mean The numbers in the last column in the file are called the temperature factors or B-factor. The B-factor describes the mean-square displacement, a measure of the displacement of an atom from an average value. If the atoms are more flexible, the electron density determined in x-ray structures is lower than if the atoms are more fixed, which gives high electron density.

To make stabilizing disulfide bonds, investigators found site chains close enough that when mutated to cysteines could potentially form disulfide bonds. In addition, they search for such residues in surface loops (without alpha and beta structure) which are inherently more flexible. Introducing disulfide bonds into the loop would stabilize it and make it more rigid. Table $LGA.2$ below shows a description of the double cysteine CA variants in the study.

Variant designation	Position	Wild-type residues	Loop length	Sum of B-factors
T133C/D197C	133, 197	Thr/Asp	63	87.60
P56C/P156C	56, 156	Pro/Pro	99	80.82
N63C/P145C	63, 145	Asn/Pro	81	77.17

Table $LGA.2$: **Description of the double cysteine CA variants in ngCA**

The locations of the side chair targeted for mutations to cysteine pairs are shown in Figure $LGA.13$ below.

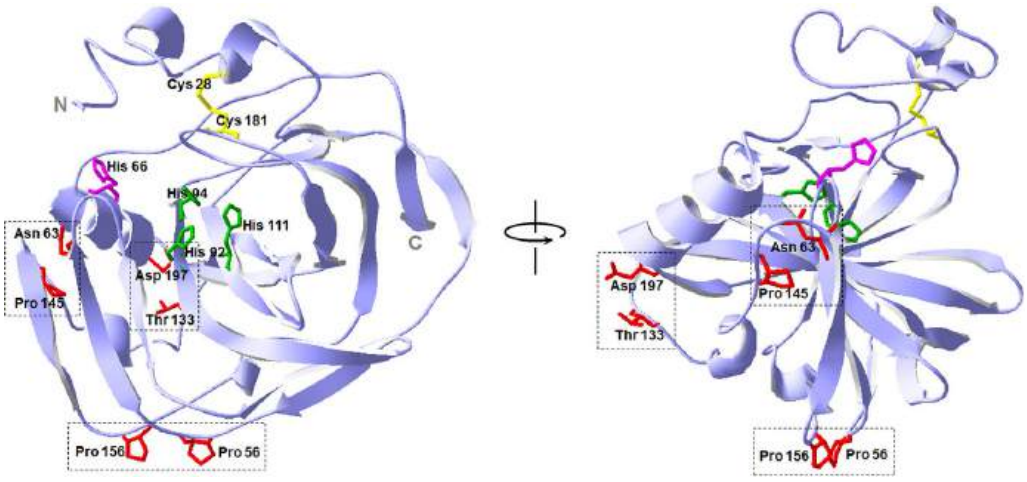


Figure $LGA.13$**: 3D structure of *ng*CA and location of residue pairs for disulfide engineering**

The zinc (not shown)-coordinating histidine residues in the catalytic active site are shown in green. The proton shuttle histidine residue is shown in magenta. The native disulfide bond is colored yellow.

An interactive iCn3D model of carbonic anhydrase from Neisseria gonorrhea (1KOQ) highlighting the 3 pairs of sidechains for mutations is shown in Figure $LGA.14$ below.

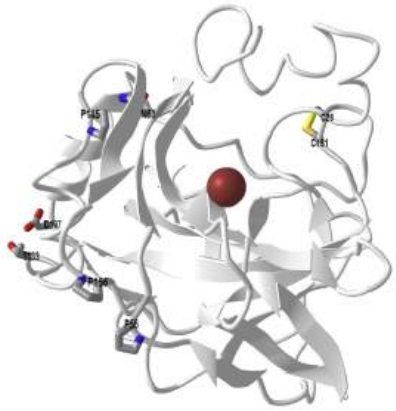


Figure $LGA.14$: Carbonic anhydrase (Neisseria gonorrhea) with 3 paired side chains for engineered disulfide (1KOQ) (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...KRP2tFJ1pV1FF6

The mutations were made and the wild-type proteins and three mutants were subjected to SDS-polyacrylamide gel electrophoresis (SDS-PAGE). The stained gels are shown in Figure $LGA.15$ below.

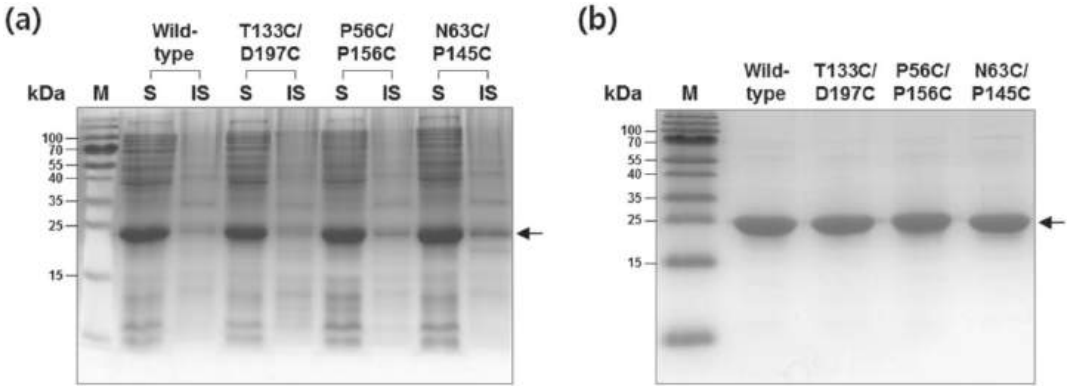


Figure $LGA.15$: **Expression and purification of disulfide CA variants.**

Panel (**a**) shows the expression of the protein in transformed cells 25 °C after IPTG induction and fractionated into soluble and insoluble fractions. SHuffle strain, an engineered *E. coli* strain that promotes cytoplasmic disulfide bond formation, was used.

Panel (**b**) shows purification results. Each lane was loaded with 4 µg of each purified CA variant. The proteins were visualized with Coomassie blue staining after SDS-PAGE. The arrow indicates the position of the bands corresponding to *ng*CA variants. Lane: M, molecular weight marker; S, soluble fraction; IS, insoluble fraction.

? Question $LGA.20$

a. Interpret the results of the PAGE gels in Figure $LGA.15$ above

b. How pure were the proteins based on the PAGE gel result in Panel (b). Can you infer from the gel that the proteins folded correctly?

Answer

a. It appears to show that a small fraction of the wild type and each mutant, especially the N63C/P145C pair, was found in the insoluble fraction. This might result from improper folding of the proteins in E. Coli, leading to hydrophobic side chain surface exposure and aggregation into insoluble "inclusion bodies". This appears to be just a minor issue.

b. All the proteins appear very pure with some very small levels of contamination in the N63C/P145C. The PAGE results show the protein all have the same molecular weight but whether they folded to a native state with activity can not be determined. Nor can it be determined if the disulfide pairs in the mutants were made are if they were, were correctly paired.

The investigators next determined if the expressed and purified wild-type and mutant proteins had the correct number of disulfide bonds. They did this by reacting the proteins in the absence and presence of dithiothreitol with DTNB or 5,5'-dithiobis(2-nitrobenzoic acid), also called Ellman's Reagent. Both structures are shown in Figure $LGA.16$ below.

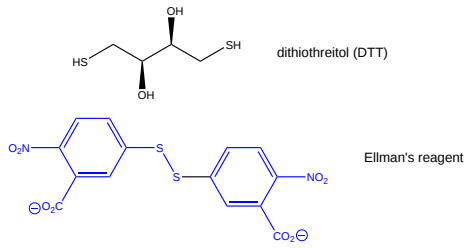


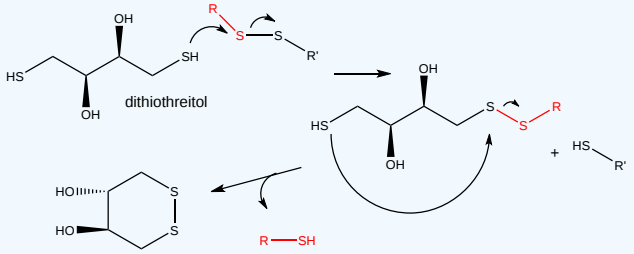
Figure $LGA.16$: Structures of DTT and DTNB

DTT is a reducing agent that cleaves disulfide.

Question $LGA.21$

Draw a mechanism showing the reaction of a disulfide with DTT.

Answer



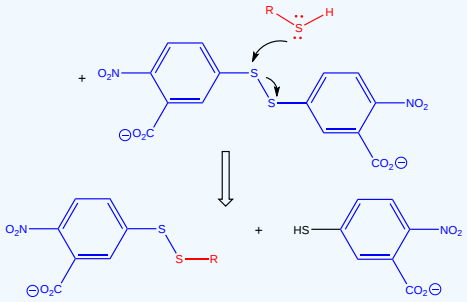
Only surface and not buried free cysteines will be labeled unless the protein is unfolded to expose all the cysteines.

DTNB reacts with free sulfhydryls to form the 2-nitro-5-thiobenzoic acid anion leaving group that absorbs at 412 nm.

Question $LGA.22$

Draw a mechanism showing the reaction of free sulfhydryl (like Cys) with Ellman's reagent

Answer



Only surface and not buried free cysteines will be labeled unless the protein is unfolded to expose all the cysteines.

The results of the reaction of the proteins with Ellmans's agent, in the presence and absence of DTT, are shown in Table $LGA.3$ below.

CA variant	Free thiol/protein (mol/mol) [a]		Deduced no. S-S bonds
	−DTT	**+DTT**	
Wild-type	0.06 ± 0.02	1.80 ± 0.14	?
T133C/D197C	0.08 ± 0.02	3.79 ± 0.22	?
P56C/P156C	0.06 ± 0.03	3.89 ± 0.06	?
N63C/P145C	0.08 ± 0.03	3.75 ± 0.05	?

Table $LGA.3$: Analysis of disulfides in CA using Ellman's reagent. [a]Numbers are represented in mean ± SD.

? Question $LGA.23$

How many S-S would you deduce from the table are present in the wild-type and mutant enzymes? Did the correct disulfide bonds form? Explain your answers

Answer

DTT reduces the disulfide in protein. For each disulfide, two free Cys side chains are made. The molar ratio of CysSH/CA for the wild-type is 1.8 in the presence of DTT. The value is very close to the expected value of 2. For the mutants, the ratio is about 3.8 in the presence of DTT, suggesting 4 free Cys consistent with 2 disulfide bonds.

Table $LGA.4$ below shows the catalytic activities of the disulfide CA variants at 25 °C.

CA variant	CO_2 hydration activity			
	Relative esterase activity [a]	$k_{cat} \times 10^{-4}$ (s^{-1})	K_M (mM)	$k_{cat}/K_M \times 10^{-6}$($M^{-1}$ s^{-1})
Wild-type	1.00	1.44	14.2	1.01
T133C/D197C	1.49	1.97	16.7	1.18
P56C/P156C	1.03	1.44	16.9	0.85
N63C/P145C	0.55	0.27	17.3	0.16

Table $LGA.4$: catalytic activities of the disulfide CA variants at 25 °C [a]The specific activity of the wild-type corresponds to 0.22 U/µmol-enzyme.

? Question $LGA.24$

Why did the investigators conduct this experiment? Interpret the results

Answer

All of the previous results suggest that the mutant proteins were made and had the correct number of double bonds, but the experiments could not tell if the bond pairs were correct. For example, did the T133C/D197C contain a native (C28-C181) and mutant (C133-C197) bond and not another combination? Activity is an excellent predictor of structure. All but one mutant retained nominal activity, as evidenced by a comparison of the rat constants. The N63C/P145C had a 6x lower k_{cat}, but even then it is close to diffusion-controlled.

Now comes the big question: were the investigators able to engineer thermal stability into the carbonic anhydrase? Experimental results to show the thermostability of the disulfide CA mutants are shown in Figure $LGA.17$ below.

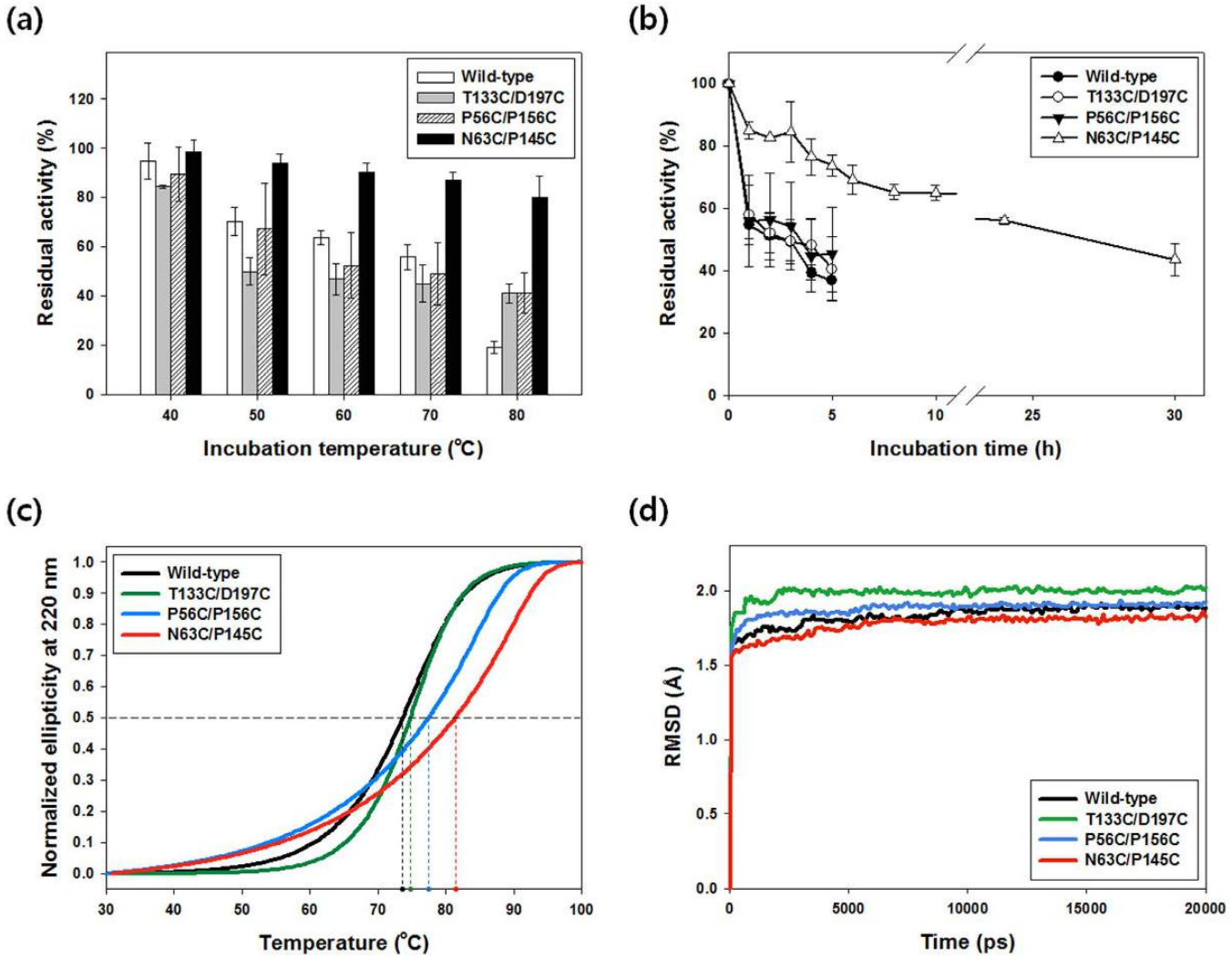


Figure $LGA.17$: Thermostability of the disulfide CA variants

Panel (**a**) shows short-term kinetic stability. The enzyme solutions (40 µM) were incubated for 30 min at different temperatures, and the residual activities were measured by esterase activity assay. Activities of 100% correspond to untreated samples. Panel (**b**) shows long-term kinetic stability at 70 °C. The half-lives ($t_{1/2}$) of the CA variants were estimated by fitting the experimental data to an exponential decay curve. Each value represents the mean of at least three independent experiments, and the error bars represent the standard deviations.

? Question $LGA.25$

Analyze the results in Panels (a) an (b). Which protein was most thermostable over the short (30 minute) and long (hours) incubating time at elevated temperatures?

Answer

At 80 °C, all the mutants showed increases thermostability to short term (30 minute) heating, but one, N63C/P145C was exceptionally thermostable. Longer time courses for heating at 70°C showed that the N63C/P145C was again far more stable over time. Its $t_{1/2}$ was 31.4 h, compared to the values between 4-6 h for the others.

Panel (c) shows heat-induced denaturation of disulfide CA variants. Temperature-dependent changes of the circular dichroism ellipticity were recorded at 220 nm on CD spectrometer. The denaturation curves were normalized to the fraction of unfolded protein. The horizontal dashed line indicates the point at which the fraction of unfolded protein is 0.5. The vertical dashed lines

point to T_M values. Panel (**d**) shows the overall RMSD of disulfide variants from molecular dynamic simulations performed at 400 K for 20 ns.

Question $LGA.26$

Analyze the results in Panel (c). What do changes in the CD helicity show? Which protein was most thermostable based on T_M values? Is the decrease in enzyme activity in panels (a) and (b) result from the denaturation of the protein?

Answer

CD measurements can give a measure of the retention of secondary structure (alpha helices and beta structure) on denaturation. The CD spectrum for different secondary structures is shown below (From Chapter 3.5).

The curves were normalized to fit on a 0-1 scale on the y axis, which then gives a measure of percent denaturation. The temperature half-way up is the T_M, or "melting temperature, at which an equilibrium mixture would contain half native and half denatured protein (true for a small protein with no intermediates). The T_M values were for the wild-type, T133C/D197C, P56C/P156C, and N63C/P145C mutants 73.6 °C 74.7 °C, 77.4 °C, and 81.4 °C. These parallel the $t_{1/2}$ values for enzyme activities, and support the idea that denaturation led to inactivation of the enzyme.

Question $LGA.27$

Analyze the molecular dynamics simulation results in Panel (d)

Answer

The molecular dynamic simulations for all the proteins soon reach equilibrium values as indicated in the plateaus of average room mean square deviation of the protein backbone. The overall molecular root-mean-square deviation (RMSD) of N63C/P145C was the lowest, indicating that it was most rigid. This is in accord with the idea that increased flexibility destabilizes a protein and engineering a disulfide into makes it more rigid and hence more stable to temperature increases. The results are in accordance with the other experiments that show the N63C/P145C was the most thermostable.

" In addition, T133C/D197C showed the highest values in both the overall and the residual RMSD (Fig 3D). This may explain and correlate with the increased activity of T133C/D197C (Ta) and the increased ΔS of unfolding"

You may remember from both introductory chemistry and from Chapter 4.12, that you can calculate the thermodynamic parameters, $\Delta \mathbf{H}^o$ and $\Delta \mathbf{S}^o$ for N ↔D at room temperature from thermal denaturation curves using the van 't Hoff equation.

In this case, K_{eq} values can be calculated from thermal denaturation curves by monitoring change in CD signal at 220 nm, and applying this equation (also from Chapter 3.12).

$$K_{eq} = \frac{[D]_{eq}}{[N]_{eq}} = \frac{f_D}{f_N} = \frac{f_D}{1 - f_D} \tag{LGA.1}$$

From this, we can calculate ΔG^0.

$$\Delta G^0 = -R\,T \ln K_{eq} = -R\,T \ln\left[\frac{f_D}{1-f_D}\right] \tag{LGA.2}$$

Knowing K_{eq}, ΔH^0, DS^0 can be calculated as shown below. A semi-log plot of $\ln K_{eq}$ vs 1/T is a straight line with a slope of - $\Delta H^0 R$ and a y-intercept of + $\Delta S^0/R$, where R is the ideal gas constant.

$$\begin{aligned} \Delta G^0 = \Delta H^0 - T\Delta S^0 = -RT\ln K_{eq} \\ \ln K_{eq} = -\frac{\Delta H^0 - T\Delta S^0}{RT} \\ \ln K_{eq} = -\frac{\Delta H^0}{RT} + \frac{\Delta S^0}{R} \end{aligned} \tag{LGA.3}$$

The equation below shows that the derivative of equation (8) with respect to 1/T (i.e. the slope of equation 8 plotted as $\ln K_{eq}$ vs 1/T) is indeed $-\Delta H^0/R$. Equation (9) is the van 't Hoff equation, and the calculated value of the enthalpy change is termed the van 't Hoff enthalpy, $\Delta H^0 v_{Hoff}$.

$$\frac{d\ln K_{eq}}{d(1/T)} = -\frac{\Delta H^0}{R} = -\frac{\Delta H^0_{vHoff}}{R} \tag{LGA.4}$$

Using this method, the thermodynamic parameters for unfolding of the protein were calculated. The results are shown in Table $LGA.4$ below.

CA variant	Melting temperature, T_M (°C)	Enthalpy change of unfolding, ΔH (kcal mol^{-1})	Entropy change of unfolding, ΔS (kcal mol^{-1} K^{-1})
Wild-type	73.6	48.8	0.141
T133C/D197C	74.7	52.8	0.153
P56C/P156C	77.4	35.1	0.091
N63C/P145C	81.4	30.0	0.085

Table $LGA.4$: Thermodynamic parameters for protein unfolding for WT and mutant CAs

? Question $LGA.28$

Which effects, enthalpy or entropy of unfolding, were associated with the increased thermal stability of the mutants compared to the wild-type protein. Remember were are considering the denaturation reaction, N ↔ D.

Answer

For the reaction N ↔ D, the ΔH^0 values were all positive, indicating the enthalpy changes favored the native state, not the denatured state.

In contrast, the other two mutants were enthalpically destabilized compared to the wild-type as their ΔH^0 were less positive so compared to the wild-type. The prime stabilizer of the native state was the lower entropy (hence a less negative and favored $-T\Delta S^0$ for the denaturation reaction. This makes sense in these mutants are more rigid and would experience less loss of "conformational entropy).

P56C/P156C and N63C/P145C exhibited lower ΔH (destabilizing) and ΔS (stabilizing), showing that the decreased entropic change of unfolding (i.e., the loss of conformational entropy of the unfolded state) by the disulfide bridge was the primary factor for the thermostabilization. These results are not surprising because design strategies aiming 'entropic stabilization' such as disulfide engineering do not always result in engineered proteins ideally with lower ΔS and unchanged ΔH.

These results are in accord with the observation that N63C/P145C was the most thermostable variant and that T133C/D197C showed the highest values in both the overall and the residual RMSD (Fig. 3d). This may explain and correlate with the increased activity of T133C/D197C and the increased ΔS of unfolding.

Finally, the enzymatic activity of the wild-type and mutants CAs (using a small ester substrate) were studied as a function of temperature. The relative activity of the wild-type and all 3 disulfide mutants are plotted as a function of temperature in the histogram graphs shown in Figure $LGA.18$ below.

Figure $LGA.18$: Effect of temperature on the activity of disulfide CA variants.

Esterase activities of disulfide variants were measured at each temperature and normalized to the activity of each enzyme at 25 °C. Each value represents the mean of three independent experiments, and the error bars represent the standard deviations. If you plotted the data as curves, you would get bell-shaped graphs.

? Question $LGA.29$

Explain why the histogram plots (and line plots if they were drawn) are bell-shaped. Are the results in accordance with the previous results.

Answer

Yes. Most chemical reaction show an increase in rate with increasing temperatures until competing reactions take precedence. For an enzyme-catalyzed reaction, that competing reaction is denaturation, which decreases the rate.

Yes the graphs are in accord with the previous results. The N63C/P145C certainly stands out as the best mutant. The authors write that "considering the shifted optimal temperature and the thermoactivation as well as the enhanced thermostability, the disulfide engineered α-type CA with Cys^{63}-Cys^{145} can be a promising biocatalyst for efficient CO_2 sequestration performed under high temperature conditions."

D

CHAPTER OVERVIEW

28: Biosignaling - Capstone Volume I

28.1: General Features of Signal Transduction
28.2: At the cell membrane- receptors and receptor enzymes
28.3: The Next step - The Kinome and Activation of Kinases at the Cell Membrane
28.4: The next step - Downsteam intracellular signaling
28.5: Small G proteins, GAPs and GEFs
28.6: Phosphatases
28.7: Calcium Signaling
28.8: Receptor Guanylyl Cyclases, cGMP, and Protein Kinase G
28.9: Gated Ion Channels - Neural Signaling
28.10: Integrins- Bidirectional Cell Adhesion Receptors
28.11: Signaling by Steroid Hormones
28.12: mTOR and Nutrient Signaling
28.13: Regulation of the Cell Cycle by Protein Kinases
28.14: Programmed Cell Death
28.15: Signaling in Microorganisms
28.16: Signaling in Plants
28.17: Signal Transduction - Vision and Olfaction
28.18: Signal Transduction - Taste (Gustation)
28.19: Signal Transduction - Temperature
28.20: Signal Transduction - Pressure

28.1: General Features of Signal Transduction

28.1.1: Introduction to Cell Signaling

Cell signaling is at the heart of biology. A cell must know how to respond to chemical signals in its environment. These signals control every aspect of cell life and interactions. A cell must sense when to grow, divide and die. It must sense the presence of foreign and toxic molecules. It must defend itself. The membrane represents the divide between the outside and inside world. Signals must cross that divide and this most often happens without the signaling molecule entering the cell. Just the signal itself is transduced across the membrane. And it doesn't stop there. Internal signaling, also across internal membranes of organelles in eukaryotes, propagates spatially and temporally across the cytoplasm in all cells (prokaryotes, Archaea, and eukaryotes). It's impossible to describe the myriad of processes that occur in cell signaling in a single chapter, but we will do our best to present the common features used across cells.

In addition, signals in complex organisms must be integrated within tissues and organs, between organs (brain and liver for example), and within the entire organisms. Mathematical modeling is critical in understanding the complexity of these interconnected interactions. Consider the flee, fight, or freeze responses. What if you were walking down the street and suddenly saw a tiger approaching you? Most would flee. In that process, the neural, muscular, and metabolic systems must be integrated. The resting state is followed by the activation of the flight/flee response, the maintenance of the fleeing state, and then the return to the resting state. An alternative is the freeze response, which in some situations would be adaptive.

Perhaps the most complex signaling occurs in the great communication networks in the neural and immune systems. Think of it. You can experience a traumatic or emotionally-charged event once and remember it forever. Ordinary events leave less permanent traces. The biochemical changes accompanying long and short-term memory are fascinating

Let's consider general principles for signal transduction across membranes of any cell that must respond to its environment. Typically the agent that signals a cell to respond is a molecular signal. Signaling can also be mediated by pressure (for touch and hearing) or light for vision. The chemical signal binds either to a cell surface receptor or to a cytoplasmic receptor if the signaling agent is hydrophobic and can transit the membrane bilayer.

28.1.2: The logic of signaling

To a first approximation, you could consider a cell as a black box, as shown in Figure 28.1.1, with input (an external molecule for example) and resulting output signals within a cell, such as activation of gene transcription, trafficking of molecules within the cell, chemical reactions, or even export of molecules from the cell. There could be one or more signals, and one or more outputs. The inputs might arrive gradually and reach a threshold before triggering an output, or the input might be abrupt leading to an immediate output. This Figure 28.1.1:

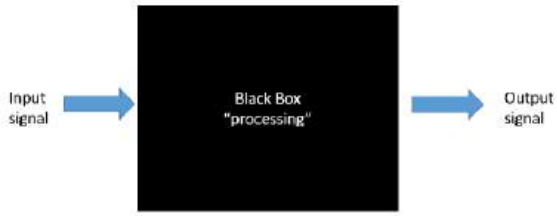
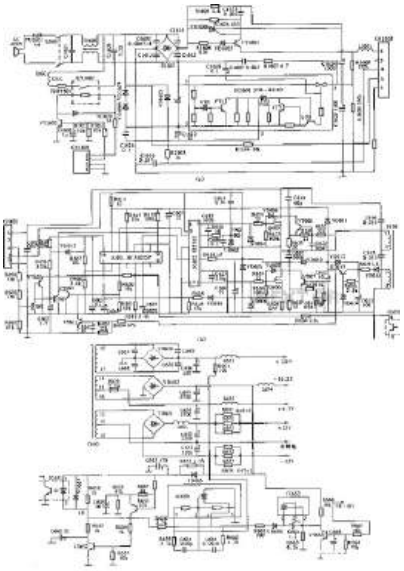
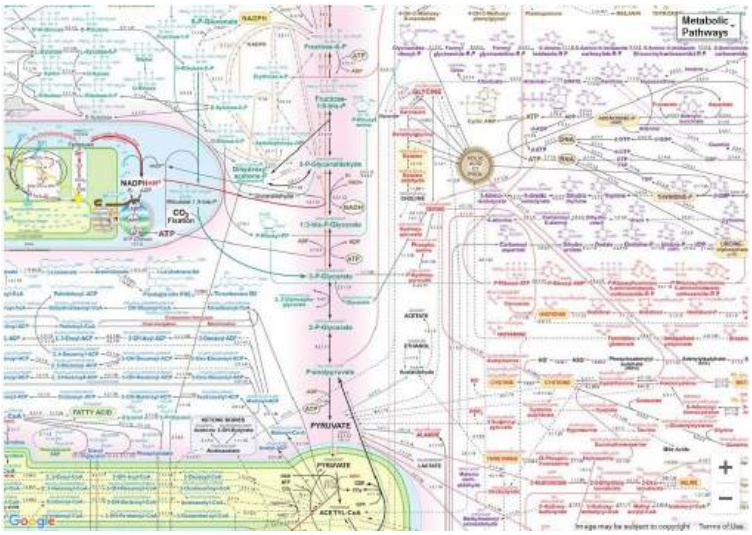


Figure 28.1.1: A cell as a black box processing unit.

Note the use of terminology taken from the world of electronics. Figure 28.1.2 shows the similarity in the complexity of an electronic wiring diagram (left) to the interconnected metabolic pathways of a cell (right).



Electronic wiring diagram



Metabolic pathway map

Figure 28.1.2: Comparison of an electronic wiring diagram and a metabolic pathway map. https://www.sigmaaldrich.com/US/en/t...c-pathways-map

The analogy is much greater than you might think. Cell signaling researchers have adopted the language and symbols of electronics as they consider how two inputs could lead to different output signals. Figure 28.1.3 shows a truth table with different Boolean operator logic gates and symbols used by electrical engineers that also apply to signals to cells and their outputs. Two input signals, A and B, arrive at different logic gates named AND, NOR, OR, XOR AND, and NAND. 0 indicates no signal (either input or output) while 1 indicates a signal (either input or output).

INPUT AND OUTPUT		INPUT NOR OUTPUT		INPUT OR OUTPUT		INPUT XOR OUTPUT		INPUT NAND OUTPUT	
IN	OUT	IN	OUT	IN	OUT	IN	OUT	IN	OUT
A B		A B		A B		A B		A B	
0 0	0	0 0	1	0 0	0	0 0	0	0 0	0
1 0	0	1 0	0	1 0	1	1 0	1	1 0	1
0 1	0	0 1	0	0 1	1	0 1	1	0 1	1
1 1	1	1 1	0	1 1	1	1 1	0	1 1	0

Figure 28.1.3: Boolean operator logic gates

Each of these logic gates leads to different outputs:

- **AND** gates require both A and B input signals for an output;
- **OR** gates require either A or B or both for an output signal;
- **NOR** gates require neither A or B for an output signal;
- **NAND** gates normally have an output signal unless both inputs A and B are present;
- **XOR** (Exclusionary OR) gates give a signal if the two inputs A and B differ.

An example of an analogous biological AND gate is N-WASP protein [a homolog to the Wiskott-Aldrich syndrome protein (WASP)]. This protein regulates actin polymerization and binds multiple signals. Stimulation with two proteins, Cdc42 and phosphatidylinositol (4,5)-bisphosphate (PIP2), activates polymerization.

Researchers are designing protein logic gates by creating a series of heterodimeric proteins. Consider the following heterodimers: A:A', B:B', and C:C', where the second member in each pair is different from the first and bound reversibly through noncovalent interactions (indicated by the colon :) Other versions could exist, such as A:C'. An AND gate for the formation of an A:C' dimer would be made by making the covalent, single molecules A'-B and B'-C. The A:C' dimer could form only in the presence of both A'-B and B'-C. This is illustrated in Figure 28.1.4

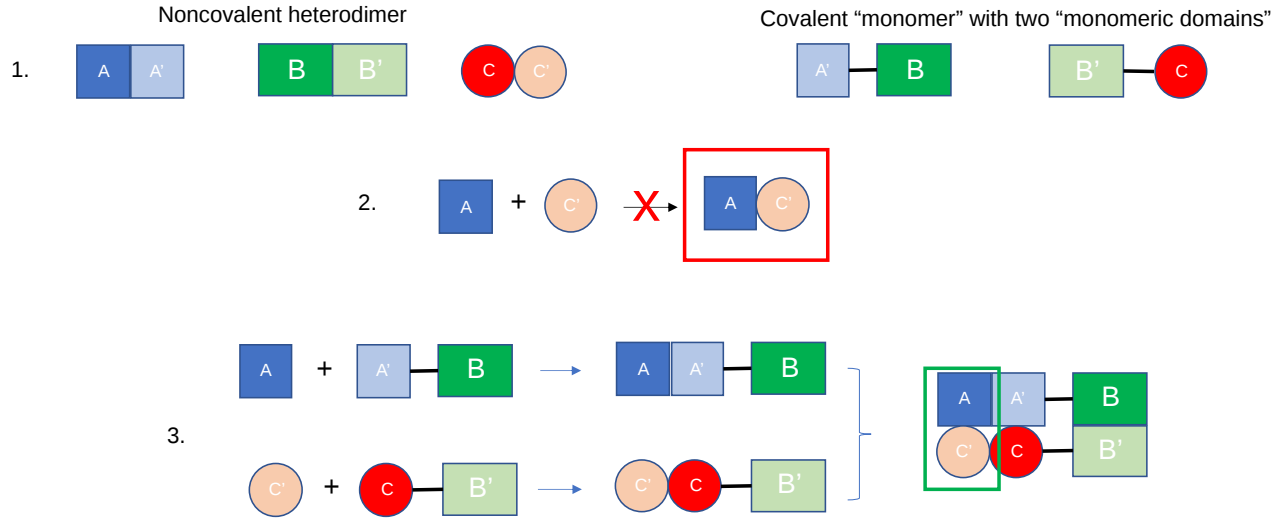


Figure 28.1.4: Designing heterodimeric protein logic gates

Signals and their responses must occur at the right time under the right conditions and for the right duration. Under opposing sets of conditions (for example well feed and starving), opposing signaling pathways, mediated by different signaling molecules, must be mutually integrated and regulated so one is turned on and one shut off. Methods must be in place to terminate signal effects.

Given the complexity of signaling pathways, it's hard to know how to present the material in a single chapter. Binding initiates and mediates almost all biological events. That binding must also be specific to avoid off-target effects. If you were to predict what biomolecule could confer specificity in the binding of a signal and allow changes from the unbound to bound state, you would certainly pick proteins. Protein:ligand binding is key to understanding biosignaling. Other types of biomolecules such as lipids and nucleic acid are involved and are of clear importance, but they are typically involved downstream from the initial protein:signal binding event. Hence we will focus mostly on signaling proteins and conformational/activity changes that occur on signal binding.

To simplify protein involvement in signaling, we can assume signaling proteins have an active and inactive state. The states are interconvertible (reversible). Let's assume the signaling protein is an enzyme. The activity of the enzyme depends on many factors as described in Figure (\PageIndex{5}\) below. The green color represents the active enzyme, while red indicated inactive.

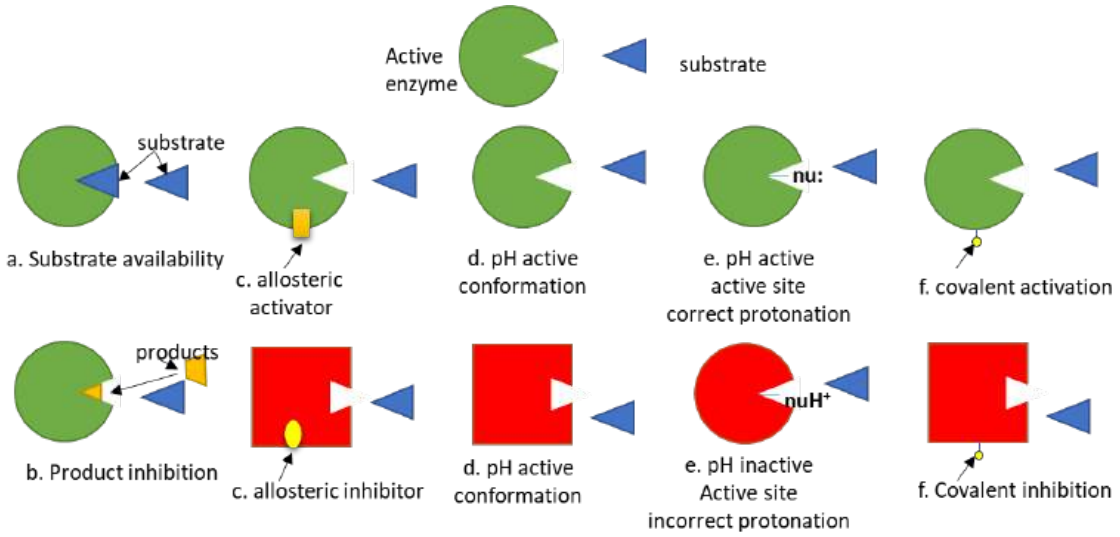


Figure 28.1.5: Ways to change the activity of an enzyme

Of course, proteins that are not enzymes (for example transcription factors) can be regulated similarly.

In addition, the amount (and localization as well) of a signaling protein can regulate the activity of the protein, as illustrated in Figure (\PageIndex{6}\) below.

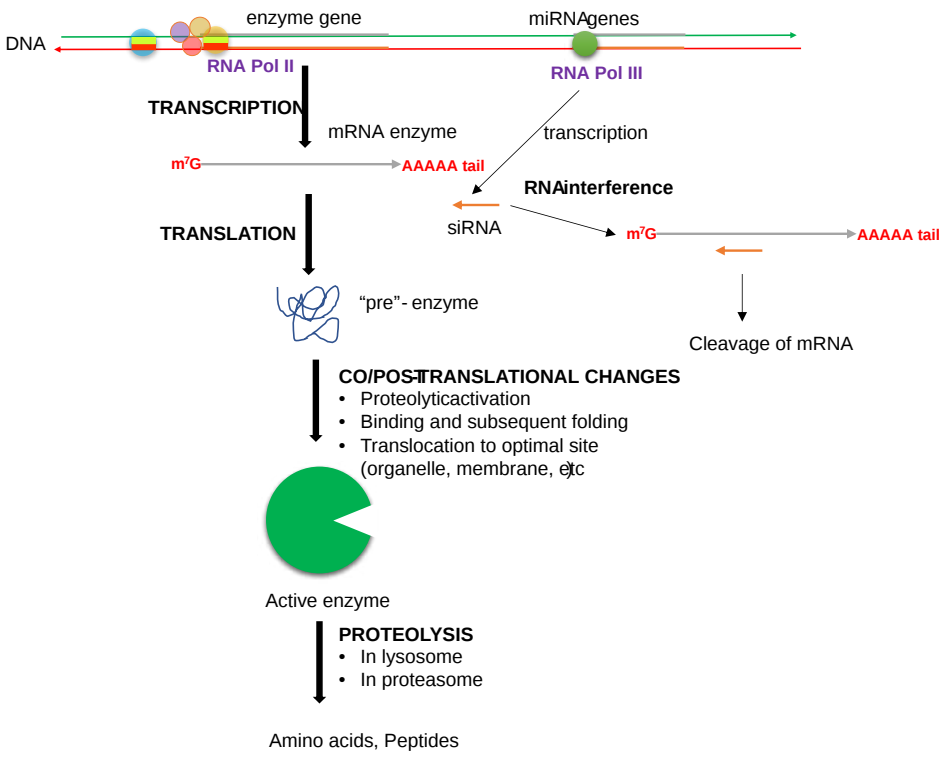


Figure 28.1.6: Ways to change the amount of an enzyme

Each signaling protein can be considered to be a node in a larger pathway consisting of interconnected nodes. This chapter on cell signaling hence can be viewed as a capstone to Unit 1, which explores the structure and function of biomolecules and as a prelude to the study of whole metabolic pathways.

28.1.3: Second Messengers and Signal Translocation

The chemical species that trigger signaling typically binds to a target protein transmembrane receptor on the surface of a cell but does not itself enter the cell. Just the signal enters the cell. If you were to hypothesize how that might happen, you would predict two mechanisms:

- an integral transmembrane receptor undergoes a conformational change on binding that propagates to its cytoplasmic domain, which can then interact either with other cytoplasmic proteins or cytoplasmic domains of other membrane proteins, transmitting the symbol into the cytoplasm;
- the bound conformation of the receptor has enzyme activity and can catalyze chemical reactions on the luminal side of the membrane. This could include the chemical modification of proteins or the synthesis of new small molecules from cell metabolites. These new small molecules are called **second messengers.** They could be formed in the membrane bilayer or the cytoplasm.

There are several diverse types of second messengers. Two common ones are cyclic derivatives of small nucleotides, including cyclic AMP (cAMP) derived from ATP and cyclic GMP (cGMP) derived from GTP. Ca^{2+} ions are usually found in low concentrations in the cytoplasm as they are pumped into internal organelles such as the endoplasmic reticulum and mitochondrion. Signaling processes can release Ca^{2+} ions in waves within the cell. Membrane lipids are also processed to form second messengers. Membrane phospholipids can be cleaved by cell signaling-activated lipases to form free arachidonic acid, sphingosine, diacylglycerol, and inositol-trisphosphate, which can act as second messengers. Redox signaling in the cell can also occur through hydroperoxides acting as second messengers.

What if the response of the cell requires gene transcription? Somehow the signal has to translocate from the cell membrane through the cytoplasm through the nuclear membrane into the nucleus. Hence a series of translocations of multiple downstream signaling events must occur.

We have already seen how newly synthesized proteins have signal sequences that target them to specific cellular locations like the cell membrane, mitochondria, nucleus (nuclear localization sequences - NLS and the small GTP binding protein RAN), or for export. Proteins involved in signaling also can move throughout the cell as part of the signaling process. Likewise, we have seen how cytoplasmic proteins can be targeted to membranes by attachment of fatty acids or isoprenoids.

28.1.4: Post-translational modification of signaling proteins

Nature has chosen the post-translational modifications (PTM) of proteins as a ubiquitous way to alter the signaling states of proteins. As we have seen previously, PTMs can alter protein conformation. They can also present new binding interfaces that allow interaction with other signaling proteins. The main (but not the only) PTM used for signaling is the reversible phosphorylation of the OH-containing amino acid side chains (Tyr, Ser, and Thr) as well as histidine (mostly in prokaryotes), so we will focus on them. Enzymes that catalyse the phosphorylation of proteins are called **protein kinases.** Reversibility is important since if phosphorylation of a target protein is associated with a specific signaling change (either activation or inhibition), then dephosphorylation can easily reverse the signaling event. Enzymes that dephosphorylate phosphoproteins are called **protein phosphatases**. Figure 28.1.7shows the generic reaction of protein kinases and phosphatases.

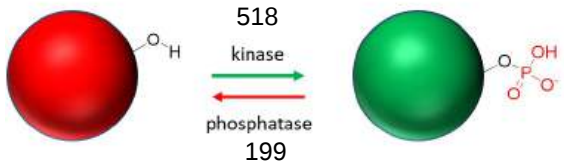


Figure 28.1.7: Kinases and phosphatases

There appear to be 518 protein kinases and 199 protein phosphatases encoded in the human genome. Why so many? If just one protein kinase existed, one mutation in it would be disastrous. In addition, the large number of kinases and phosphatase allows for great control in the specificity of these enzymes for their target protein substrates.

Kinases

Kinases are a class of enzymes that use ATP to phosphorylate molecules within the cell.

The names given to **kinases** show the substrate which is phosphorylated by the enzyme. For example:

- hexokinase - an enzyme that uses ATP to phosphorylate hexoses.
- protein kinase - enzymes that use ATP to phosphorylate proteins within the cell. (Note: Hexokinase is a protein, but is not a protein kinase).
- phosphorylase kinase: an enzyme that uses ATP to phosphorylate the protein phosphorylase within the cell

If a protein is phosphorylated by a kinase, the phosphate group must eventually be removed by a **phosphatase** through hydrolysis. If it wasn't, the phosphorylated protein would be in a constant state of either being activated or inhibited. Kinases and phosphatases regulate all aspects of cellular function. About 1-2% of the entire genome encodes kinases and phosphatases.

Kinases can be classified in many ways. One is substrate specificity: Eukaryotes have different kinases that phosphorylate serine/threonine or tyrosine side chains. Prokaryotes also have His and Asp kinases, but these are unrelated structurally to the eukaryotic kinases. There are 11 structurally different families of eukaryotic kinases, which all fold to a similar active site with an activation loop and catalytic loop between which substrates (ATP and the OH-containing side chain) bind. Simple, single-cell eukaryotic cells (like yeast) have predominantly cytoplasmic Ser/Thr kinases, while more complex eukaryotic cells (like human cells) have many Tyr kinases. These include the membrane-receptor Tyr kinases and the cytoplasmic Src kinases.

Manning et al. have analyzed the entire human genome (DNA and transcripts) and have identified 518 different protein kinases, which cluster into 7 main families as shown in Table 28.1.1: below. Family membership was determined by sequence comparisons of catalytic domains. The entire repertoire of kinases in the genome is called the **kinome**. Alterations in 218 of these appear to be associated with human diseases.

Name	Description
AGC	Contain PKA, PKG, and PKC families
CAMK	Ca2+/CAM-dependent PK
CKI	Casein kinase 1
CMGC	Contain CDK, MAPK,GSK3, CLK families
STE	homologs of yeast sterile 7, 11, 20 kinases; MAP Kinase
PTK	Protein tyrosine kinase
PTKL	Protein tyrosine kinase-like
RGC	Receptor guanylate kinase

Table 28.1.1: The human kinome

Phosphatases

There are three main families of phosphatases, the phospho-Tyr phosphatases (PTP), the phospho-Ser/Thr phosphatases, and those that cleave both. Of all phosphorylation sites, most (86%) are on Ser, 12% involve Thr, and about 2% on Tyr. They can also be categorized by their molecular sizes, inhibitors, divalent cation requirements, etc. In contrast to kinases which differ in the structure of their catalytic domains, many phosphatases (PPs below) gain specificity by binding protein cofactors which facilitate translocation and binding to specific phosphoproteins. The active phosphatase hence often consists of a complex of the phosphatase catalytic subunit and a regulatory subunit. Regulatory subunits for Tyr phosphatases may contain a SH2 domain allowing binding of the binary complex to autophosphorylated membrane receptor Tyr kinases.

Given this background, we can now start to explore signal transduction in a methodical way. There are two major ways to organize our discussions:

- start from the binding of the molecular signal at the cell membrane and trace the signaling events inward into the cell, potentially all the way to the nucleus and gene expression
- describe recurring motifs found in most pathways.

We will use a combination of both, but it makes sense to start with the signaling proteins at the cell membrane. Next, we will focus on second messengers. We'll follow that with detailed explorations of kinases and phosphatases and specific signaling pathways.

28.2: At the cell membrane- receptors and receptor enzymes

28.2.1: Introduction

We are going to start a more detailed description of cell signaling where it begins, at the cell membrane, and move inward, toward intracellular organs, where we will end at the nucleus and changes in gene expression mediated by the signal. Of course, this end is somewhat arbitrary as the signal could propagate from the nucleus back to the membrane where newly synthesized membrane protein might be inserted or even exported, as in the case of secreted antibodies. It is difficult enough to keep track of all the molecular players, let alone add onto that their initial location in the cell and their final location if they translocate. Signaling can also be more daunting for those with a more chemistry focus, who find the details of cellular structure and trafficking a bit daunting. Figure 28.2.1 shows a truncated view of the cell membrane, membrane proteins, and some of the organelles that we will visit throughout this chapter. To stay localized we will repeat the figure or variants of it several times in this section.

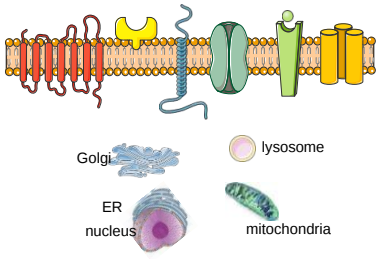


Figure 28.2.1: Location of signaling events. Reference: Nucleus - https://commons.wikimedia.org/wiki/File:0318_Nucleus.jpg; Mitochondria: https://upload.wikimedia.org/Wikipedia/commons/f/f2/Blausen_0644_Mitochondria.png; Golgi: https://smart.servier.com/smart_image/golgi-apparatus-3/. Creative Commons Attribution 3.0 Unported License; Proteins, lysosome: https://smart.servier.com/. Creative Commons Attribution 3.0 Unported LicenseNucleus - https://commons.wikimedia.org/wiki/File:0318_Nucleus.jpg; Mitochondria: https://upload.wikimedia.org/Wikipedia/commons/f/f2/Blausen_0644_Mitochondria.png; Golgi: https://smart.servier.com/smart_image/golgi-apparatus-3/. Creative Commons Attribution 3.0 Unported License; Proteins, lysosome: https://smart.servier.com/. Creative Commons Attribution 3.0 Unported License

28.2.2: Receptors at Cell Membrane

Let's start at the location where signaling almost invariable begins, in the cell membrane, where signals (hormones, neurotransmitters, nutrients, other cells) bind to cell surface transmembrane proteins shown in the red box in Figure 28.2.2

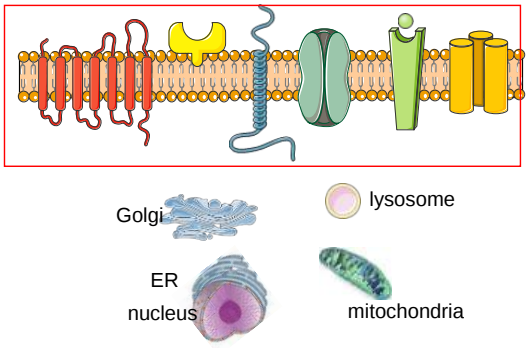


Figure 28.2.2: Signaling events at the cell membrane

We have already discussed integral and peripheral membrane proteins in Chapter 11. When occupied by a ligand signal, the signaling event mediated by a transmembrane receptor might be a change in the membrane protein conformation, which propagates to its intracellular domain. Alternatively, the receptor might change its conformation and become a ligand-gated kinase (or possibly a phosphatase) or a ligand-gated channel or pore. We will discuss signal-gated ion channels in Chapter 28.9 - Neural Signaling.

Ultimately, intracellular enzymes are activated in cells in response to the external molecular signal. This provides amplification of the initial signal since a single activated enzyme undergoes multiple rounds of catalysis before it would become inactivated. If the multiple products (a second messenger or phosphorylated proteins for example) that are formed activate multiple addition enzymes, the signal is further amplified.

28.2.3: Receptors with no kinase or transport activity - G Protein-Coupled Receptors (GPCRs)

One major type of signaling receptor is the **G-protein coupled receptor (GPCR)**. Over **800** GPCRs are encoded in the human genome which represents almost 4-5% of the number of protein-coding genes. The proteins below to five major families: rhodopsin, secretin, glutamate, adhesion, and frizzled/taste2. They don't express enzymatic activity but on binding, they can activate enzymes inside the cell by interacting through their cytoplasmic domains with **G proteins** (GTP binding proteins) in the cytoplasm. GPCRs have been called serpentine receptors as their single amino acid chain proteins have 7 transmembrane-spanning α- helices. All of the GPCRs have similar yet slightly different structures, allowing them to interact with specific ligands. Many of the GPCRs bind to unknown ligands and hence are called **orphan receptors**. We will explore a few GPCRs in more detail below.

28.2.3.1: GPCRs that modulate the membrane enzyme adenylyl cyclase:

β-adrenergic receptor

The β-adrenergic receptor is a prototypical GPCR. Found in muscle, liver, and fat cells, it binds epinephrine and adrenaline, which leads to energy mobilization and muscle activation (i.e. flight or fight response). The mechanism of activation of a GPCR is illustrated using the beta-adrenergic receptor as an example. The unoccupied adrenergic receptor is associated with a **heterotrimeric G protein**, which contains an α, β, and γ subunits. GDP is usually bound to the α subunit. Figure 28.2.3 shows a cartoon of the GPCR interacting with the heterotrimeric G protein.

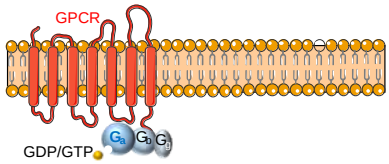


Figure 28.2.3: GPCR-Heterotrimeric complex in the bilayer

When the hormone is bound to the receptor, a conformation change is propagated to the cytoplasmic domain of the GPCR altering its interaction with $G_{\alpha\beta\gamma}$. This causes a conformational change in the G_α subunit, which leads to an exchange of bound GDP with GTP in the G_α subunit, promoting the dissociation of the G_α-GTP from $G_{\beta\gamma}$. G_α-GTP then binds to and activates an adjacent membrane enzyme called **adenylate cyclase**, which produces a second messenger by converting ATP to **cyclic AMP (cAMP)**. Figure 28.2.4 shows steps in the generic GPCR activation cycle.

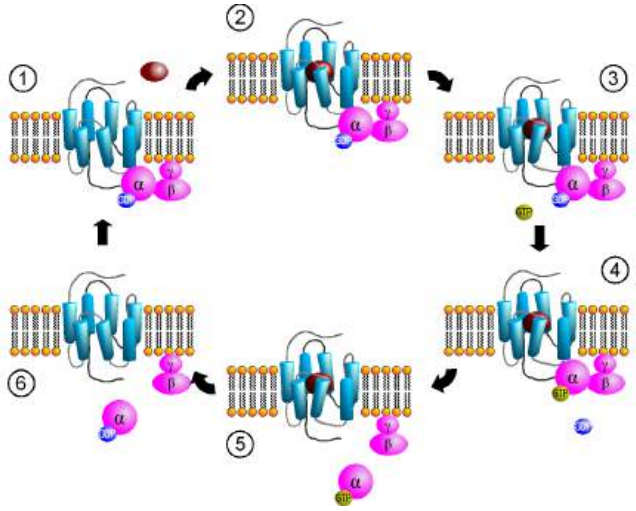


Figure 28.2.4: Activation cycle of G-proteins by G-protein-coupled receptors. https://commons.wikimedia.org/wiki/F...PCR-Zyklus.png Creative Commons Attribution-Share Alike 3.0 Unported

The primary message binds to the GPCR (1 leading to state 2. The cytoplasmic domain conformational changes allow GTP exchange in the G_α subunit as state 3 goes to 4. In 5 the G_α-GTP complex dissociates from $G_{\beta\gamma}$, which remains in the membrane. The G_α-GTP subunit is held and localized to the membrane (not evident in the above figure) through a lipid anchor attached through a post-translational modification.

Remember, the GPCR has no ligand-gated enzymatic activity. Yet it indirectly leads to the activation of a membrane enzyme, adenylate cyclase, when the dissociated G_α-GTP binds to the cyclase. As long as GTP remains bound to the G_α subunit, it will continue to modulate the activity of adenylate cyclase. A built-in regulatory mechanism does exist in the protein since the G_α subunit has **GTPase** activity. The GTP will eventually hydrolyze, and the GDP-G_α subunit will lose affinity for its bound partner (adenylate cyclase), and return to the heterotrimeric G protein associated with the unbound receptor. GPCRs bind the signaling ligand (primary message) in a binding cavity localized at the extracellular face and between four of the transmembrane helices.

The activity and structure of GPCRs have been studied using natural ligands (hormones and neurotransmitters), as well as agonists, partial agonists, inverse agonists, and antagonists. As discussed previously, agonists bind to the natural ligand binding site and elicit the same or a partial response (partial agonist). Inverse agonists bind and lower the response of a constitutively active receptor, and antagonists bind and prevent the normal response of an agonist. About 35% of pharmaceutical drugs target GPCRs, but target only about 15% of the ~800 human GPCRs. The orphan GPCRs are increasingly targets for drug development. Most hormones and neurotransmitters work through GPCRs. In addition, our primary senses of vision, smell (olfaction), and taste (gustation) work through GPCRs.

Figure 28.2.5 shows the structure of the beta 2-andrenergic:Gs complex with bound agonist. No membrane is shown.

Figure 28.2.5: Structure of the beta-andrenergic:Gs complex with bound agonist.

The GPCR is shown in cyan. The 7 transmembrane helices should be obvious. The ligand is bound between them. The G_α subunit is shown in magenta, the G_β in dark blue, and the G_γ in orange/brown. The gray subunit is a Camelid antibody VHH fragment (a single-chain nanoantibody used to stabilize the conformation for crystallization.)

The biggest conformational change that occurs when the GPCR binds an agonist is an outward movement (14 Å) at the intracellular domain of transmembrane segment 6 (TM6) and an extension of the TM5 helix. This leads to a movement of the G_α's alpha-helical domain, enabling GTP exchange for GTP. Of course, multiple reactions determine the fraction of the G_α in the active GTP-bound state. These would include the relative cellular concentrations of free cytoplasmic GDP and GTP, their K_D values for the G_α, the rate constant for the hydrolysis of bound GTP, and the rate constants for the GDP ↔ GTP exchange.

Figure 28.2.6 shows an interactive iCn3D model of the **Crystal structure of the beta2 adrenergic receptor-Gs protein complex 3SN6**

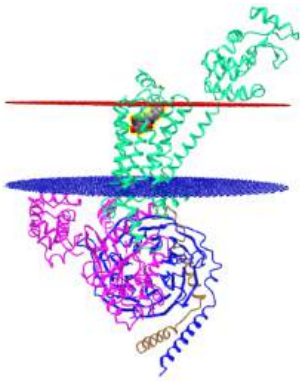


Figure 28.2.6: **β-2 adrenergic receptor-Gs protein complex 3SN6** (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...8R4zKZCfKZy248

The GPCR is shown in green, the G_α in magenta, Gβ in blue, and the Gγ in brown. An agonist (spacefill) is shown in CPK colors near the outer leaflet.

Some bacterial toxins work by inactivating the GTPase activity of the G_α subunit, keeping it in the "stuck" position. For example, cholera toxin, an enzyme released by Vibrio cholera, catalyzes the ADP ribosylation of an Arg in the G_α subunit by transferring everything but the nicotinamide from NAD^+ to the Arg residue.

Since the G_α subunit stimulated the activity of adenylyl cyclases, it is often named a stimulatory G_α protein or $G_{s\alpha}$. Figure 28.2.7 shows an interactive iCn3D model of the adenylyl cyclase activator $G_{s\alpha}$ with GTP-γ-S (1azt)

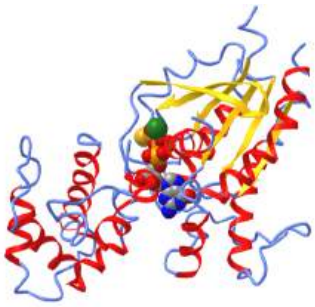


Figure 28.2.7: adenylyl cyclase activator $G_{s\alpha}$ with GTP-γ-S (1azt) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...T7MBqYnaxAH3SA

Now we can explore how the occupied GPCR, which again has no enzymatic activity, activates the enzyme adenylyl cyclase. Figure 28.2.8 shows a cartoon of the $G_{s\alpha}$ subunit bound to adenylyl cyclase as the $G_{\beta\gamma}$ heterodimer remains associated with the membrane

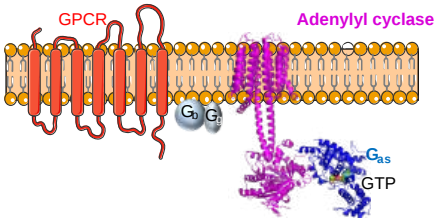


Figure 28.2.8:

Adenylate cyclase converts ATP into the second messenger **cyclic AMP** (cAMP) as shown in Figure 28.2.9. The figure also shows how cAMP is broken down into AMP by the enzyme **cAMP-specific 3',5'-cyclic phosphodiesterase (PDE).** The latter step is necessary to control the lifetime of the second messenger cAMP.

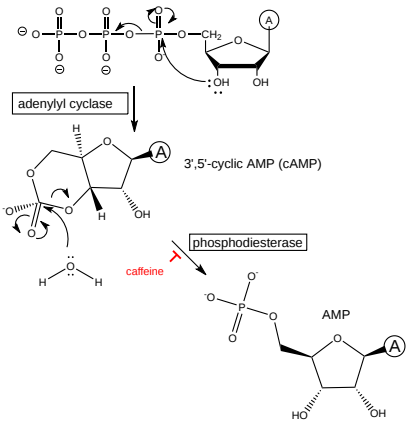


Figure 28.2.9: Formation of cAMP by adenylyl cyclase and its degradation by phosphodiesterase

Why cAMP?

You might ask why cAMP and not just AMP is nature's choice for GPCR signaling. AMP is a very important metabolic species. High concentrations of it signal an energy-depleted state. AMP is used in another signaling process to mobilize a response to adjust the energy state of a cell using a protein called AMP-Protein Kinase, which we will describe later. Many enzymes are also allosterically regulated by AMP.

Figure 28.2.10shows an interactive iCn3D model of the membrane adenylyl cyclase bound to an activated stimulatory $G_{s\alpha}$ protein 6R3Q determined by cryo-EM.

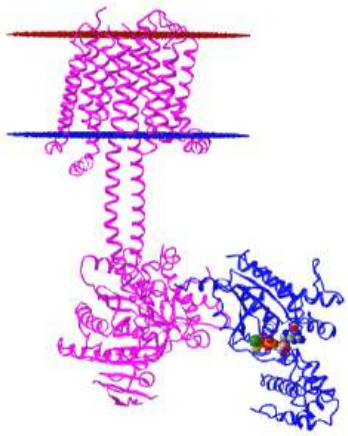


Figure 28.2.10: membrane adenylyl cyclase bound to an activated stimulatory G protein 6R3Q (Copyright; author via source). Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...foXY38deJSMDC6

The carboxyl-terminal cytoplasmic domain has the catalytic and allosteric sites.

Cannabinoid Receptors

In contrast to the beta-adrenergic receptor which mediates activation of adenylate cyclase through $G_{s\alpha}$, some G_{α} subunits inhibit adenylate cyclase when bound. These G_{α} subunits are called $G_{i/o\alpha}$ in contrast to the stimulatory subunits, $G_{s\alpha}$. Also, G_{α} subunits interact with many proteins other than adenylate cyclase. Examples include cannabinoid receptors.

Cannabinoid receptors are named after the exogenous and psychoactive drug Δ^9-tetrahydrocannabinol (**THC**) that binds to the receptor. THC is the major phytocannabinoid (from plants) found in the Cannabis sativa plant and the marijuana-derived from it. The other main cannabinoid in the plant is cannabidiol (**CBD).** Phytocannabinoids bind to two types of human cannabinoid (CB) receptors, CB1 and CB2. They have 44% amino acid and 68% homologies in the entire protein and the transmembrane domain, respectively.

The phytocannabinoids exert their effect through binding to human CB1 and CB2, whose endogenous ligands are two fatty acids derivatives called anandamide (**AEA**) and 2-arachidonoyl glycerol (**2-AG**). The structure of THC, CBD, and the two major

endogenous ligands are shown in Figure 28.2.11

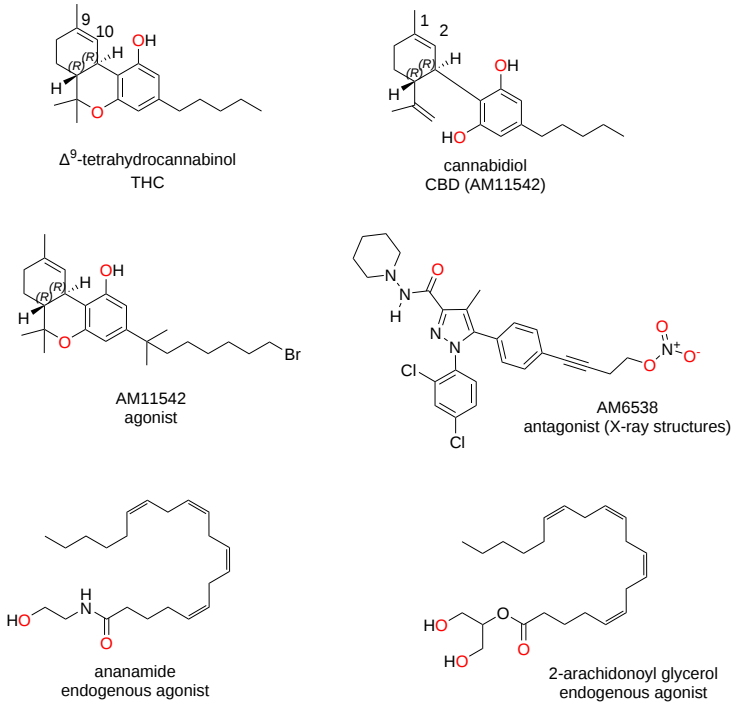


Figure 28.2.11: Structure of agonist and antagonist for cannabinoid receptors

The cannabinoids from Cannabis sativa have a monoterpene isoprenyl group (C10) and a pentyl side chain (C5). The ligands are largely hydrophobic and probably access their binding site in the receptor mainly by lateral movement in the membrane. The receptors differ most in the N-terminal extracellular loop which is also involved in ligand binding.

Figure 28.2.12 shows an interactive iCn3D model of the class A GPCR Cannabinoid Receptor-Gi Complex Structure with bound agonist (6KPF)

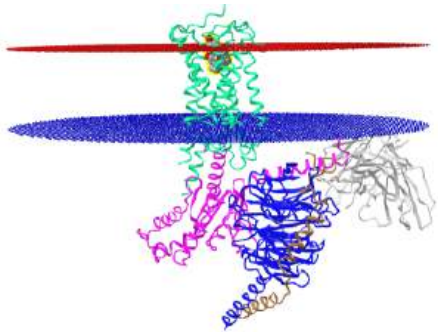


Figure 28.2.12: Class A GPCR Cannabinoid Receptor-GiComplex Structures (6KPF) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...B16mYTKPMhXmg8

The gray protein is a nanobody used to stabilize the protein during crystallization. The agonist is AM12033, which is similar to AM11542 in the figure above, but with a -C≡N terminus instead of a bromide.

THC and CBD

Cannabis sativa contains the psychoactive drug, Δ^9-tetrahydrocannabinol (**THC**), which is a partial agonist for CB1 (binds with reported K_i values of 10 or 53 nm)and CB2 (K_i = 40 nm). Its psychoactive effects on mental activity as well as pain and appetite are well known. In contrast, cannabidiol (**CBD)** is the main, non-psychoactive cannabinoid. It has a much lower affinity for the recombinant CB1 (K_i = 1.5 **μM**) and CB2 (K_i = 3.7 **μM**). It appears to be a partial antagonist for CB1 and a

weak inverse agonist for CB2. It has also been shown that CBD is a negative allosteric modulator of the agonistic effects of THC and 2AG. The actual psychotropic effects of combining THC and CBD are complicated and not understood well.

Figure 28.2.13shows an interactive iCn3D model of human CB1 in complex with agonist AM11542 (5XRA)

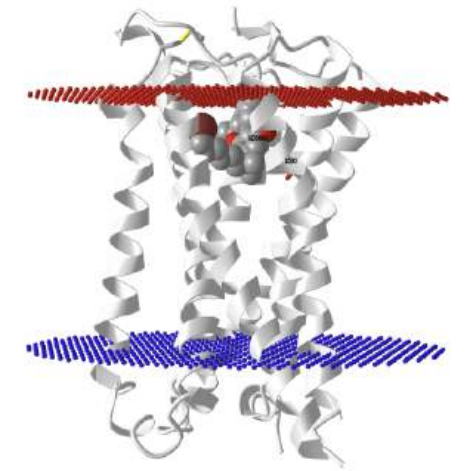


Figure 28.2.13: Human CB1 in complex with agonist AM11542 (5XRA). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...wDv9fXNPaMhwm9

The complex shows a significant conformational change in the structure compared to the receptor bound to an antagonist. This includes a large reduction (50%) in the volume of the binding pocket and an increase in the surface area of the receptor that would bind to G proteins.

How much of a receptor is bound with a cannabinoid depends on the concentration of the cannabinoid ligand and the K_i for the drug. The amount of THC and CBD depends on the genetics of the plant, which has been engineered to greatly decrease THC production (in the hemp plant used for nonpharmacological commercial properties) or increase either THC or CBD production at the expense of the other.

The synthesis of THC and CBD proceeds through a common precursor, CBGA (cannabigerol acid). Two key flavoproteins, Δ^9-tetrahydrocannabinolic acid synthase (THCAS) and cannabidiolic acid synthase (CBDAS) convert this common precursor CBGA into two new precursors, Δ^9-THCA and CBDA, respectively. This final synthetic step involves an oxidative cyclization reaction using O_2 and produces H_2O_2. Spontaneous, non-catalyzed decarboxylation and rearrangements of Δ^9-THCA and CBDA lead to the final products, THC and CBD. This last process occurs on exposure to heat, which occurs in smoking and baking, and at a slower rate during storage.

Commercially used preparations of THC and CBD for medicinal purposes vary widely in concentrations. THC concentration ranges for pain management (<5-10%) are much lower than those for psychotropic effects (<15%), with values of 21% or often found in "recreational" cannabis. High-potency THC strains can contain up to 25-30% THC by dry weight. For strands modified for CBD production, the maximal amount is about 25%. Even though CBD appears to be a partial antagonist for CB1, it appears that the ratio of THC:CBD is important in modulating the "high" or intoxicating state of THC. A ratio of THC:CBD of just over 1:1 leads to synergism or enhancement of the acute effects of THC whereas ratios of THC:CBD of 1:2 to 1:6 seem to have the least intoxicating effects. However, CBD decreases psychotic symptoms of THC and also decreases memory changes associated with THC. CBD also is an allosteric modulator of the μ-opioid receptor.

At present, there are no structures of CBD bound to its cannabinoid receptors. Figure 28.2.14 shows an interactive iCn3D model of the CBD-bound full-length rat transient receptor potential vanilloid 2 (TRPV2) in nanodiscs (6U88)

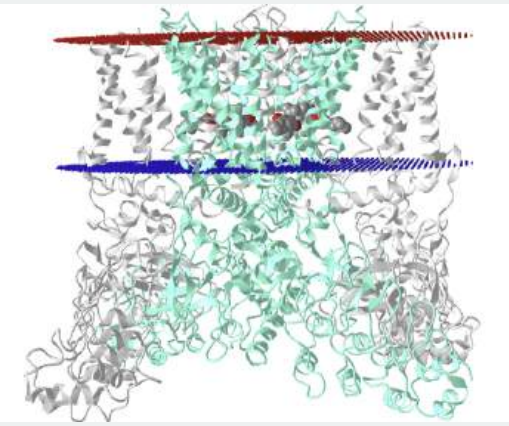


Figure 28.2.14: CBD-bound full-length rat TRPV2 in nanodiscs (6U88)**.** (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...BJcXBriq49LtQA

This receptor is a calcium-permeable non-selective cation channel that is activated at high temperatures (50^0C) but not vanilloids like capsaicin. Hence it acts as a "noxious high temperature" receptor.

Now let's make it even more complicated. There are more than 20 different G_α-like proteins known in 4 major families.

- G_s and G_i regulate adenylyl cyclase
- G_q activates phospholipase Cβ (described below). There are 4 members given these strange names: G_q, G_{11}, G_{14}, and $G_{15/16}$
- $G_{12/13}$ activate small GTPase protein (described in Chapter 28.5)

The G_α protein involved in light sensation is named transducin, while those involved in odorant detection and taste are called $G_{olfactory}$ and $G_{gustatory}$, respectively.

As we add more variants of each signaling component, the origin of the complexity of signaling systems becomes evident. For example, the neurotransmitter serotonin binds to its receptor, a GPCR, and instead of gating the protein open to ion flow (as with other ligand-gated ion channels in the activation of neurons as we will see in Chapter 28.9), it interacts with two different alpha subunits, G_s, which leads to activation of adenylyl cyclase, and G_{12}, which interacts with other small GTP binding proteins called GEFs (we will also see these later).

28.2.3.2: GPCRs modulate the activity of the membrane enzyme phospholipase C.

These receptors use the same mechanism for activation of the membrane enzyme adenylyl cyclase. When the primary signal is bound to the GPCR, a conformation change is propagated to the cytoplasmic domain of the GPCR, altering its interaction with $G_{\alpha\beta\gamma}$ in which the alpha subunit is a member of the $G_{\alpha(q)}$ family. This causes a conformational change in the G_α subunit which leads to an exchange of bound GDP with GTP in the G_α subunit, promoting the dissociation of the G_α-GTP. G_α-GTP then binds to and activates an adjacent membrane enzyme, **phospholipase C (PLC)**, which cleaves membrane phospholipids to produce two, second messengers, **diacylglycerol** and **inositol 1,4,5-trisphosphate (IP_3)**. Their structures are shown in Figure 28.2.15

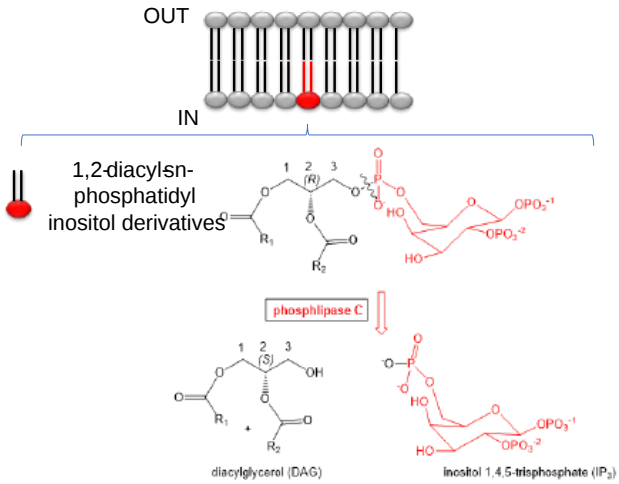


Figure 28.2.15:Activation of phospholipase C

There appear to be 13 kinds of mammalian phospholipase Cs divided into six isotypes (β, γ, δ, ε, ζ, η). Phospholipase C is also named 1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase.

Figure 28.2.16shows an interactive iCn3D model of the $G_{\alpha(q)}$-phospholipase C-β 3 structure (4GNK). The magenta structure is the $G_{\alpha(q)}$ protein with bound GDP in spacefill. The cyan structure is the **Pleckstrin Homology (PH)** domain of the protein. This domain targets proteins to inositol phospholipids in the membrane but appears not to have this function in this protein.

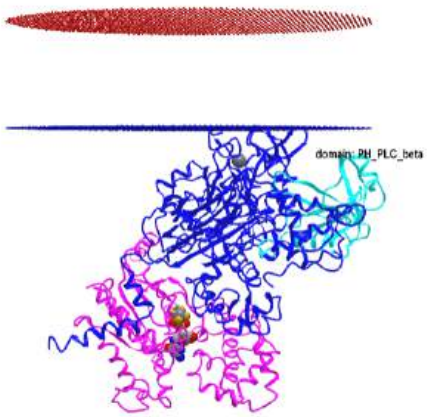


Figure 28.2.16: $G_{\alpha(q)}$-phospholipase C-β-3 structure (4GNK) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?5kiejFu8KgzXsiP9A

Note that in contrast to adenylyl cyclase, PLC is a peripheral, not an integral membrane protein. It is found in the cytoplasm as well as associated with the inner leaflet of the cell membrane where its main activities, regulating and cleaving PIP_2, occur. PLC localizes to lipid rafts enriched in PIP_2. Figure 28.2.17shows the domain structure of phospholipase C β3

Figure 28.2.17 Domain structure of phospholipase C β3

The N-terminal PH_14 represents the **Pleckstrin Homology (PH)** domain, which is among the top 15 of all domains in the human genome. All PLCs except PLCζ have this domain. Note that in another example of complexity, PLC-β binding to the inner leaflet does **not** require PIP2. The iCn3D model above indeed shows binding to the membrane seems to depend on adjacent structures and not the PH domain (cyan) specifically.

Tabel 28.2.1below characteristics of common signals that signal through GPCRs.

signal	vasopressin	epinephrine	light	odorant	odorant	sweet tastant
receptor	VR	β-adrenergic	rhodopsin	odorant receptor 1	odorant receptor 2	sweet receptor

Ga-like subunit	Gi	Gs	transducin	Golfactory	Golfactory	Ggustatory
coupled enzyme	adenylate cyclase	adenylate cyclase	phosphodiesterase	phospholipase C	adenylate cyclase	adenylate cyclase
2nd messenger	decrease cAMP	increase cAMP	decrease cGMP	increase IP3	increase cAMP	increase cAMP
protein affected	decrease PrK-A	increase PrK-A	dec. Ca, Na perm.	inc. Ca perm	inc.Ca, Na perm	dec. K perm

Tabel 28.2.1: Characteristics of common signals that work through GPCRs.

Changeux and Edelstein reviewed the MHC model 40 years after its conception and support its application to signal transduction processes. They include in signaling molecules not only hemoglobin, but regulatory enzymes (aspartate transcarbamylase, phosphofructokinase, LDH, glycogen phosphorylase), membrane receptors (acetylcholine receptor, rhodopsin), and nuclear receptors (lac repressor, steroid hormone receptors). In all these signaling proteins, residue distant from the "active" site participates in binding to allosteric ligands. Often the allosteric site is on a separate domain that can be cleaved from the protein and still maintain allosteric ligand binding properties. The proteins also consist of multiple subunits easily related by distinct symmetry axes. Allosteric ligands often bind in cavities in subunit interfaces along symmetry axes. In general, crystal structure analyses show that low-affinity T and high-affinity R forms of the signaling proteins exist but are accompanied by minor tertiary structure changes in individual subunits (i.e. perfect symmetry in all subunits is not preserved on the binding of allosteric ligand). For neurotransmitter membrane receptors, these two states can be correlated with an open and closed state (for ion flux), and open conformations of these proteins can often be found in mutant forms. However, for many ligand-gated ion channels and G-protein coupled receptors (serpentine), kinetic analyses show more complicated forms than can be represented by a simple two-state (R and T) model. High-resolution microscopy shows evidence for nonsymmetrical quaternary structural changes. These changes can be observed in the absence of ligand, which gives support to the MWC concept that allosteric ligands select certain conformational states, leading to equilibrium shifts in the unliganded receptor to the more high-affinity state. More refined methods of structural analysis will presumably show more evidence of subtle tertiary changes in the proteins that are preludes to quaternary structural changes. Yet the simplicity of the MWC model for explaining many features of signaling proteins remains.

28.2.4: Receptors with signal-gated kinase activity - **Receptor Tyrosine Kinases (RTK)**

Why bother binding a primary message to a GPCR and going through multiple steps before the activation of a membrane enzyme like adenylate cyclase? Wouldn't it be easier and more efficient to have the membrane receptor a ligand-activated enzyme? Such is the case with special membrane receptors called **Receptor Tyrosine Kinases (RTKs)**. There are about 90 tyrosine kinases in the human genome of which 58 are RTKs. Figure 28.2.18shows the family domain structure of the RTKs.

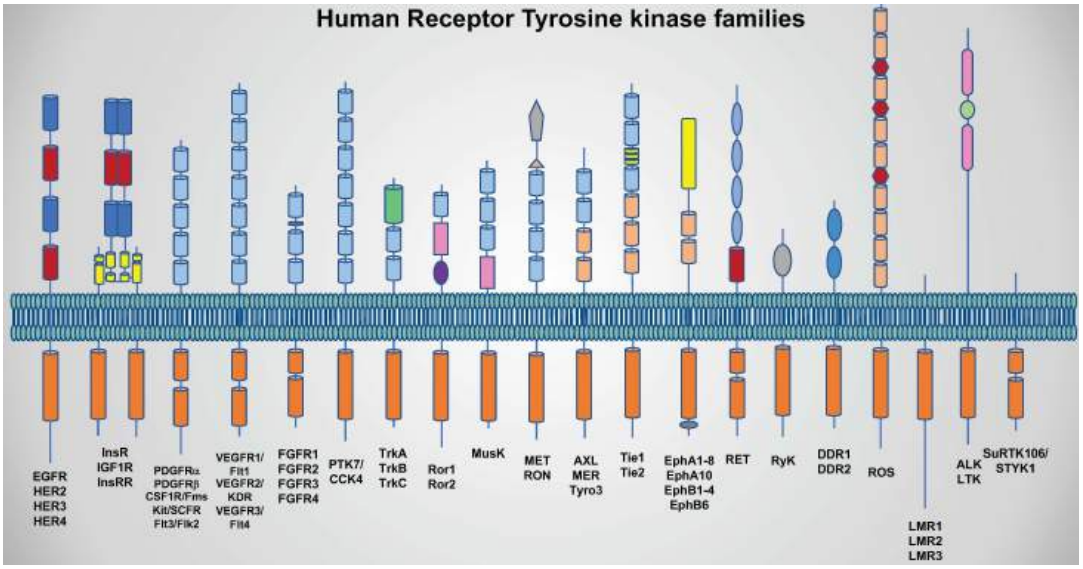


Figure 28.2.18: Family domain structure of the RTKs. Yamaoka, T.; Kusumoto, S.; Ando, K.; Ohba, M.; Ohmori, T. Receptor Tyrosine Kinase-Targeted Cancer Therapy. *Int. J. Mol. Sci.* **2018**, *19*, 3491. https://doi.org/10.3390/ijms19113491. **Creative Commons Attribution License**

Note that the insulin receptor (InsR) is a dimer of two monomeric insulin receptor chains. Figure 28.2.19s shows in more detail the domain structure of the epidermal growth factor receptor (EGFR).

Figure 28.2.19: Domain structure of the EGFR domain.

The domain (red/brown) immediately after the blue domain is the transmembrane domain. Furin is a cellular endoprotease. The green represents L domains which comprise the ligand binding site. Each L domain consists of a single-stranded right-hand beta-helix.

Here is the cascade of events for signaling through EGFR: The transmembrane has ligand-dependent tyrosine kinase activity. Binding of the hormone EGF causes receptor dimerization bringing the intracellular kinase domains (yellow PK_Try_Ser_Thr) together activating them. When active, they can phosphorylate each other (autophosphorylation) or other proteins. When the receptor is autophosphorylated, other proteins can bind to the cytoplasmic domain of the receptor Tyr kinase where they are phosphorylated. The target substrates phosphorylated by the receptor Tyr kinases are proteins with a common 100 amino acid domain called **SH** for **src homology**, based on structural homology to another cytoplasmic protein, **Src**. Src is an intracellular Tyr kinase activated when it binds through 2 SH domains to an autophosphorylated receptor Tyr kinase. Specifically, the SH2 domain has been shown to bind tyrosine-phosphorylated peptides. These domains target proteins to the autophosphorylated receptor Tyr kinase. Many proteins involved in signal transduction have SH2 domains. Some of these proteins also have catalytic domains with kinase activity. Others have phosphatase, transcription factor. or scaffolding domains.

Figure 28.2.20 shows the hormone-depended dimerization of RTKs, their autophosphorylation, and the recruitment of proteins with SH2 domains.

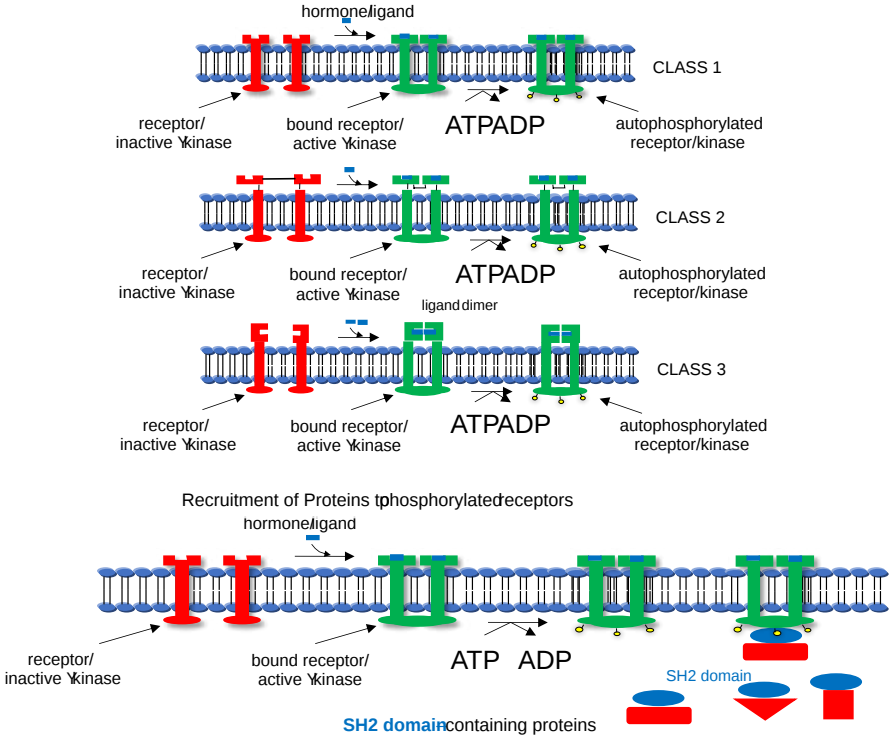


Figure 28.2.20: Activation of receptor tyrosine kinases (RTKs)

It is easier to envision how a GPCR is activated by binding its target hormone than for RTKs. GPCRs are single-chain proteins that pass through the membrane using 7-transmembrane helices. RTKs have a single transmembrane helix. The dimeric form of the RTK has some additional flexibility in the short region between the extracellular and transmembrane domains, allowing for the conformational changes necessary for the activation of the intracellular kinase domains.

The crystal structure of the full EGFR is not known given the difficulties in crystallizing membrane proteins that span the membrane with a single alpha helix. However, separate structures of the dimeric extracellular domain and the intracellular kinase domains are known.

Figure 28.2.21 shows an interactive iCn3D model of the dimeric extracellular and transmembrane domains of the epidermal growth factor receptor (3NJP).

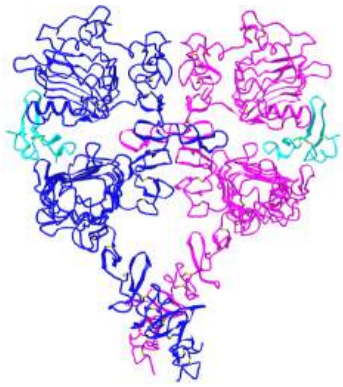


Figure 28.2.21: Dimeric extracellular and transmembrane domains of the epidermal growth factor receptor (3NJP) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...iw1V4T8rsfxDZ7

The two EGFR are shown in blue and magenta. The two bound EGFs are shown in cyan.

Figure 28.2.22shows an interactive iCn3D model of a dimer of the intracellular dimeric EGFR kinase domains in complex with an ATP analog-peptide conjugate (2GS6)

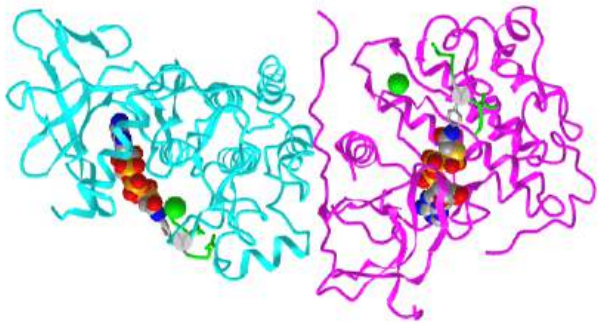


Figure 28.2.22: Dimeric intracellular EGFR kinase domains in complex with an ATP analog-peptide conjugate (2GS6) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...iw1V4T8rsfxDZ7

The two EGFR kinase domains are shown in cyan and magenta. The ATP analogs in each domain (spacefill) are thiophosphoric acid O-((adenosyl-phospho)phospho_-S-acetamidyldiester. The peptide substrates (green stick) are 13mers with a tyrosine (sticks, labeled Y, minus the OHs) connected to the ATP analog.

We've introduced the essential cell membrane players in signal transduction.

- GPCRs which are not enzymes but which activate the bound heterotrimer G prortein G_α subunit, which then can activate or inhibit the integral membrane protein adenylate cylase or activate the membrane-bound enyzme protein kinase C. These enzymes produce second messengers cAMP (adenylate cyclase) and diacylglycerol (DAG)and IP_3 (phospholipase C)
- Receptor tyrosine kinases, ligand-activated receptor kinases, which can,on ligand-induced dimerization, autophosphorylate themselves or other target protein in the cell.

In the next chapter section we will explore the next downstream effects in signaling, mediated by the second messengers cAMP, DAG and IP_3 and the substrates phosphorylated by the ligand-active receptor tyrosine kinases.

28.3: The Next step - The Kinome and Activation of Kinases at the Cell Membrane

In the last section, we studied how the cell surface GPCRs indirectly activate membrane-bound enzymes adenylyl cyclase and phospholipase C (the GPCRs) and how RTKs become active enzymes (kinases) themselves on ligand-induced dimerization. We will now explore what happens next in signaling by examining the products of these activated enzymes and how they continue the signaling process. The products of adenylyl cyclase (cAMP) and phospholipase C (diacylglycerol - DAG - and IP_3) are second messengers. These activate key protein kinases in the cell. The products of the activated RTKs are phosphoproteins (including itself) that are phosphorylated on tyrosines by the RTK. Target proteins bind to the autophosphorylated RTKs through SH2 domains on the target protein. First, we will explore protein kinases and their mechanism in general.

28.3.1: The Kinome

There are 518 different protein kinases (about 1.7% of the human genome) and as a group, they are the key players in most eukaryotic signaling pathways. Collectively they are called the **kinome**. They help regulate every aspect of cell life including metabolism, cell growth, differentiation, and division, as well as programmed cell death. In humans, most (388) are Ser/Thr kinases. 90 are Tyr kinases and 40 are classified as atypical. In this section, we will focus on the **AGC** kinases (60 in total) which include Protein Kinase **A** (PK**A**), Protein Kinase **C** (PK**C**), and also Protein Kinase B (PKB, better known as **AKT**. We'll wait for another section to explore Protein Kinase **G** (PK**G**) which is activated by **cGMP**, not AMP. We'll also study the receptor Tyr kinases (RTKs)

Luckily kinases can be grouped into families based on structure and mechanism. The ones directly activated by the second messengers cAMP (Protein Kinase A) and DAG (Protein Kinase C) are members of one family of kinases the AGC Protein Kinase family. The clustered families of protein kinases are shown in Figure 28.3.1.

28.3.2:

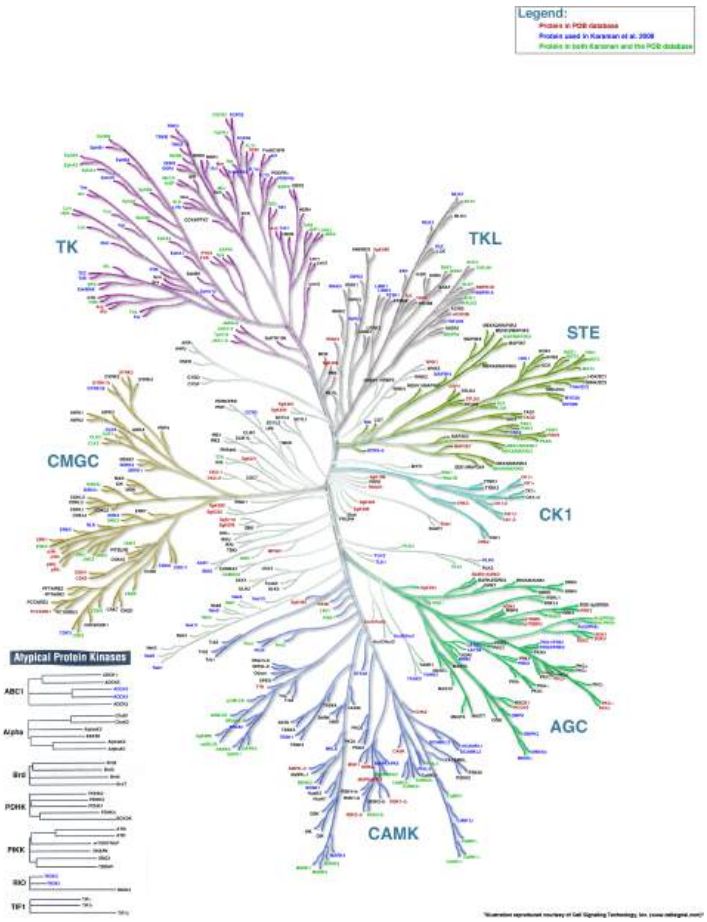


Figure 28.3.1: The family of protein kinases. https://peerj.com/articles/126/. Illustration reproduced courtesy of Cell Signaling Technology, Inc. (www.cellsignal.com)

Luckily, most of the kinases, which catalyze the phosphorylation of the OH-containing side chains (Tyr, Ser, and Thr) of proteins have similar catalytic sites. Given that phosphorylation of proteins often acts as an on/off switch for protein activity and function, it makes great sense that the catalytic site of protein kinases is not available for substrate binding and/or catalysis in the off state. Protein kinases have **catalytic** and **activation loops** and conformation changes on activation of protein kinases usually involve movement of the activation loop and rearrangement of the catalytic residues. These conformational changes are often initiated by phosphorylation by another upstream active protein kinase of key groups in the activation loop! Figure 28.3.2 shows another representation of the different families of protein kinase and their common structural features.

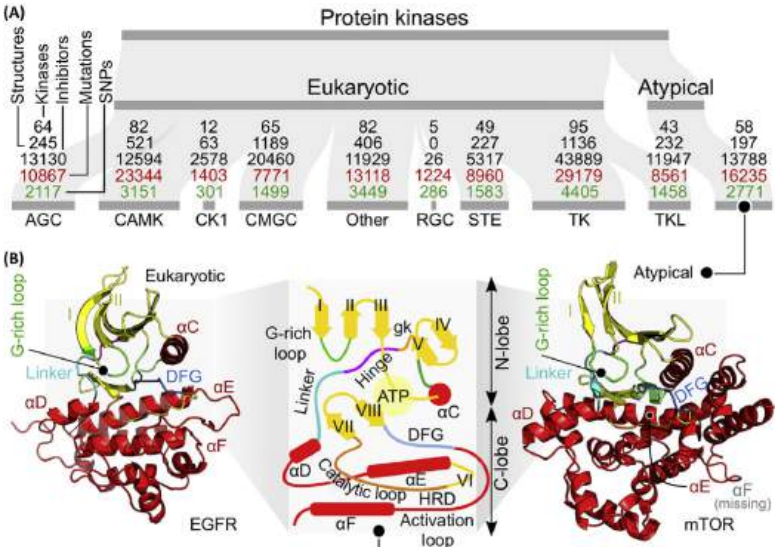


Figure 28.3.2: Family organization and common structural features of protein kinases. Trends in Pharmacological Sciences, November 2019, Vol. 40, No. 11 https://doi.org/10.1016/j.tips.2019.09.002"https://doi.org/10.1016/j.tips.2019.09.002. Creative Commons Attribution (CC BY 4.0)

Pane (A) shows the nine eukaryotic and one atypical protein kinase groups with their numbers of kinases, structures, and inhibitors (shown in black) as well as mutations (in red) and SNPs (in green). Here, the atypical protein kinases are defined as all non-eukaryotic protein kinases including also the nonprotein kinase-like kinases. The middle of Panel (B) shows the activation and catalytic loops and other common features found in the active sites of protein kinases. The left and right parts of Panel (B) show two different kinases. The protein kinase domain of epidermal growth factor receptor (EGFR) is shown on the left and the Ser/Thr protein kinase domain of **mTOR** (**m**ammalian or mechanistic **T**arget **O**f **R**apamycin), a protein we will explore in a separate section, is shown on the right.

28.3.3: AGC Kinases

These **serine/threonine kinases** are activated by the second messengers c**A**MP (**A** Kinase), c**G**MP (**G** kinase, which we will see later), and by diacylglycerol (DAG) formed by activation of phospholipase **C**. Figure 28.3.3 shows the generic structure of active kinase domains (including tyrosine kinases)

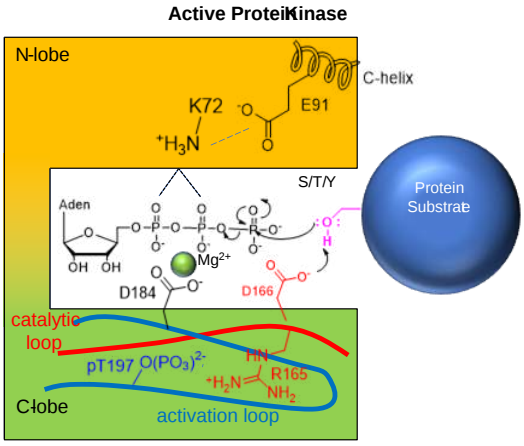


Figure 28.3.3: Generic structure of the active form of ACG kinases (Adapted from the superlative text, Cell Signaling by Lim, Mayer, and Pawson)

A generic kinase has an N-terminal and C-terminal lobe between which the substrates ATP and the target protein containing the key serine, threonine, or tyrosine side chain bind. In the active conformation, Glu 91 (E91) in the αC-helix in the N-lobe forms a salt-bridge (ion-ion interaction) with Lys72 to position it so it can stabilize the binding of the ATP through ion-dipole and H-bond interactions. In the active state, Thr 197 (T197) in the activation loop has been phosphorylated (represented by pT197) by an upstream kinase allowing a less flexible conformation of the loop. This facilitates optimal repositioning of active site side chains in the catalytic loop allowing substrate access and catalysis. Asp 166 (D166) in the catalytic loop acts as a general base in abstracting a proton from the target Ser, Thr or Tyr. Arginine 165 (R165) in the catalytic loop forms a salt bridge to phosphotyrosine 197 (pT197), stabilizing the catalysis-competent conformation of the activation loop. Aspartic acid 184 (D184) plays a key role in stabilizing the Mg^{2+} that itself stabilized the negative charges on the phosphates of ATP and in the developing transition state.

Most kinases are also classified as RD kinases since they have a key arginine (R)- aspartic acid (D) sequence in the catalytic loop. The normal protein substrate for an AGC kinase would be a protein with a Ser or Thr projecting into the active site for phosphorylation by bound ATP. The normal sequence for phosphorylation in protein targets of PKA is Arg-Arg-X-Ser/Thr-X (or more fully as XR(R/K)X(S/T)B where B is a hydrophobic amino acid). Again we will focus on three AGC kinases here, PKA, PKC, and AKT (PKB), and defer our discussion of PKG to another section.

28.3.4: cAMP-dependent Protein Kinase A - PKA

Before we study PKA, the formation of the second messenger cAMP is presented in review in Figure 28.3.4

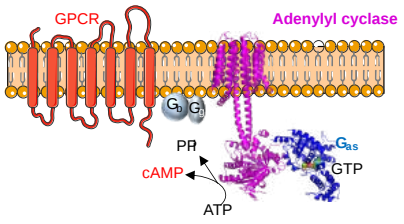


Figure 28.3.4: Synthesis of cAMP on activation of GPCRs and adenylyl cyclase

PKA contains a kinase domain with short N- and C-terminal extensions, making it a prototypical and simple model for study. In the absence of the second messenger cAMP, this prototypic AGC kinase exists as a holo-heterotetramer (or dimer of a heterodimer). It contains two catalytic kinase subunits (C) and two regulatory subunits (R). cAMP produced upon GPCRs activation of adenylyl cyclase binds to the regulatory subunits, causing them to dissociate from the heterotetramer, freeing the catalytic subunits for activity. This is illustrated in Figure 28.3.5

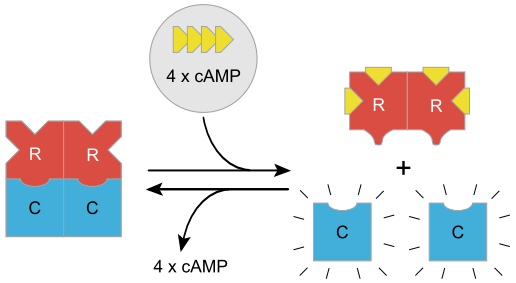


Figure 28.3.5: Activation of Protein kinase A by the second messenger cAMP

The cAMP serves as an allosteric activator of protein kinase A.

Figure 28.3.6 shows the actual structures of the protein kinase A **RIIb tetrameric holoenzyme (3TNP)**

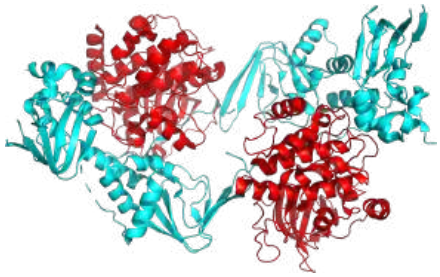


Figure 28.3.6: **Protein Kinase A RIIb tetrameric holoenzyme (3TNP)**

The inhibited catalytic subunits are shown in red and the two regulatory subunits are shown in cyan.

Figure 28.3.7 shows only one catalytic subunit of protein kinase A as it goes from the more closed inactive holoenzyme (tetramer, 3tnp) containing the regulatory subunits (not shown), to the more open apo-form (monomer, 1J3HA), without the regulatory subunit. The conformational change opens the free catalytic subunit to the substrate (protein and ATP) binding.

Figure 28.3.7: Conformational change on the conversion of the catalytic subunit of protein kinase A from the more closed inactive holoenzyme (tetramer, 3tnp) containing the regulatory subunits to the more open apo-form (monomer, 1J3HA) without the regulatory subunit

Side chains required for MgATP binding and phosphoryl transfer are pre-formed in the apo form but some changes still occur on the binding of substrate.

Figure 28.3.8 shows an interactive iCn3D model of the mouse catalytic subunit of cAMP-dependent protein kinase complexed with MnATP and a peptide inhibitor (1ATP)

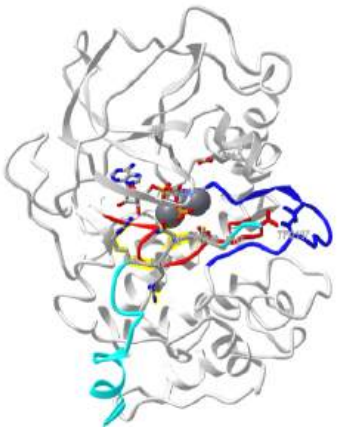


Figure 28.3.8: Catalytic subunit of cAMP-dependent protein kinase complexed with MnATP and a peptide inhibitor (1ATP) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...B8Xyks9kMdbZN9

Zoom into the model to see each of the labeled residues. All of the features shown in Figure 28.3.3 are displayed in this model. The activation loop is shown in dark blue and the catalytic loop is shown in red. The peptide inhibitor is shown in cyan. The normal

target motif for phosphorylation by Protein Kinase A (Arg-Arg-X-Ser/Thr-X) has been replaced in this model by Arg-Arg-X-Ala-X. Why does this make the peptide an inhibitor instead of a substrate?

Many primary signals activate adenylate cyclase through GPCR and produce cAMP as a second messenger. These include corticotrophin, dopamine, epinephrine (β-adrenergic), follicle-stimulating hormone, glucagon, many odorants, prostaglandins E1and E2, and many tastants. All of these would activate protein kinase A. Some enzymes regulated by cAMP-dependent phosphorylation by PKA are shown in Table 28.3.1 below

Enzyme	Pathway
Glycogen Synthase	Glycogen synthesis
Phosphorylase Kinase	Glycogen breakdown
Pyruvate Kinase	Glycolysis
Pyruvate Dehydrogenase	Pyruvate to acetyl-CoA
Hormone-sensitive Lipase	Triacylglycerol breakdown
Tyrosine Hydroxylase	Synthesis of DOPA, dopamine, norepinephrine
Histone H1	Nucleosome formation with DNA
Histone H2B	Nucleosome formation with DNA
Protein phosphatase 1 Inhibitor 1	Regulation of protein dephosphorylation
CREB	cAMP regulation of gene expression
PKA consensus sequence	XR(R/K)X(S/T)B (B = hydrophobic amino acid)

Table 28.3.1: Proteins phosphorylated by Protein Kinase A.

An example of how epinephrine (a flight/fight hormone) can lead to the breakdown of glycogen (your main carbohydrate reserves in muscle and liver) is shown in Figure 28.3.9

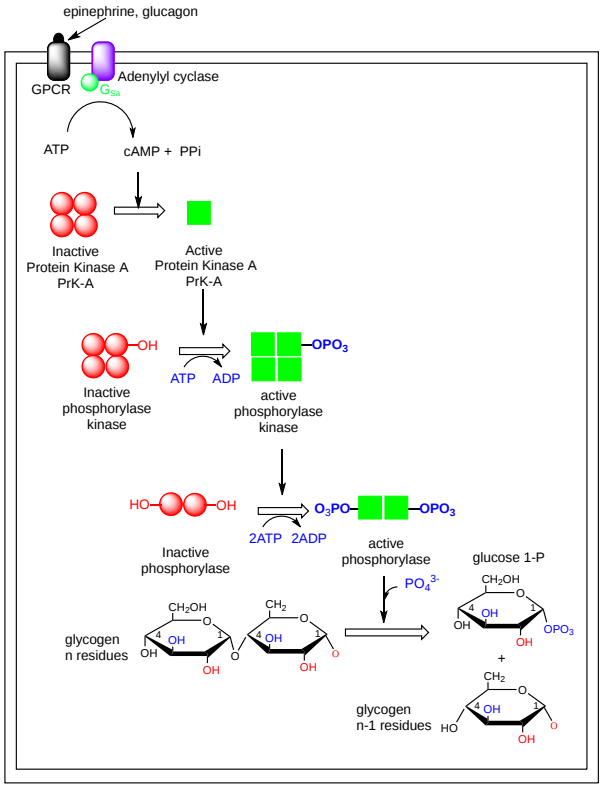


Figure 28.3.9: Cascade activation of glycogen phosphorylase through GPCR and cAMP signaling.

A cascade of events starts with the binding of the hormone to its receptor, followed by the activation of adenylate cyclase and the formation of the second messenger cAMP. This activates PKA, which phosphorylates and activates the enzyme **phosphorylase kinase**. Following the naming convention we discussed earlier, **phosphorylase kinase** is a protein kinase whose target enzyme is another enzyme called glycogen phosphorylase, which is NOT a kinase. When active, glycogen phosphorylase uses inorganic phosphate (P_i) as a nucleophile in a phosphorolysis reaction to cleave glucose from the end of glycogen polymers forming glucose-1-phosphate. This is the first step in the mobilization of glycogen as an energy reserve. No wonder it is so tightly regulated by this complicated signaling pathway.

28.3.5: The second messengers DAG and IP_3 and their activation of Protein Kinase C

The activation of protein kinase C (PKC, a Ser/Thr kinase member of the ACG protein kinase family) is very similar to that of protein kinase A. To start, an extracellular signal molecule binds to a GPCR receptor (again with no intrinsic enzyme activity), causing a conformational change in the receptor that propagates through the membrane to its intracellular domain. That then activates the exchange of GTP for GDP in the alpha subunit of the bound heterotrimeric G protein, which contains the special $G_{\alpha q}$ subunit. The $G_{\alpha q}$ subunit dissociates and binds to the membrane protein enzyme **phospholipase C** (not adenylyl cyclase). Once activated, it cleaves the phospho-head group from the membrane phosphatidyl inositol - 4,5-bisphosphate (PIP2) into two, second messengers - diacylglycerol and inositol trisphosphate (IP_3). These products are shown in Figure 28.3.10

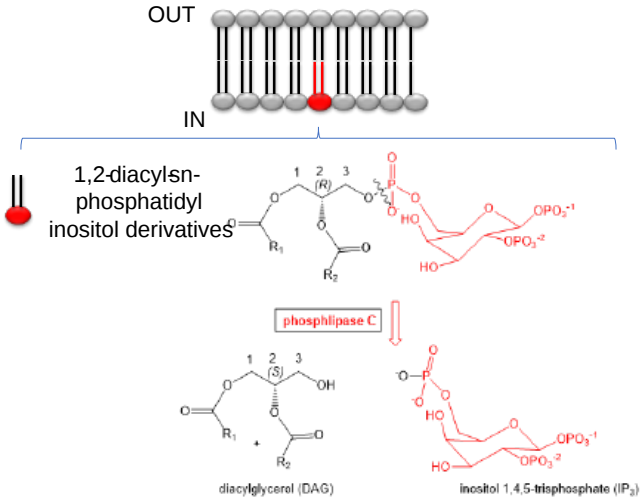


Figure 28.3.10: Cleavage of phosphatidyl inositol - 4,5-bisphosphate (PIP2) into two, second messengers - diacylglycerol and inositol trisphosphate (IP_3)

The formation of the second messengers IP_3 and DAG is presented n Figure 28.3.11 Note that phospholipase C is a peripheral membrane protein bound to the inner leaflet of the cell membrane.

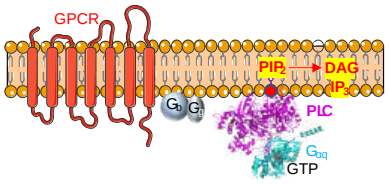


Figure 28.3.11: Synthesis of DAG and IP_3 from PIP_2 by phospholipase C

Diacylglycerol binds to and activates protein kinase C (PKC). The IP_3 binds to ligand-gated receptor/Ca^{2+} channels on internal membranes, leading to an influx of calcium ions into the cytoplasm. The released calcium ions also activate PKC. These steps are illustrated in Figure 28.3.12 in more detail below. The activation of the peripheral membrane phospholipase C (PLC) is very similar to that of the integral membrane protein adenylyl cyclase, both of which produce second messengers.

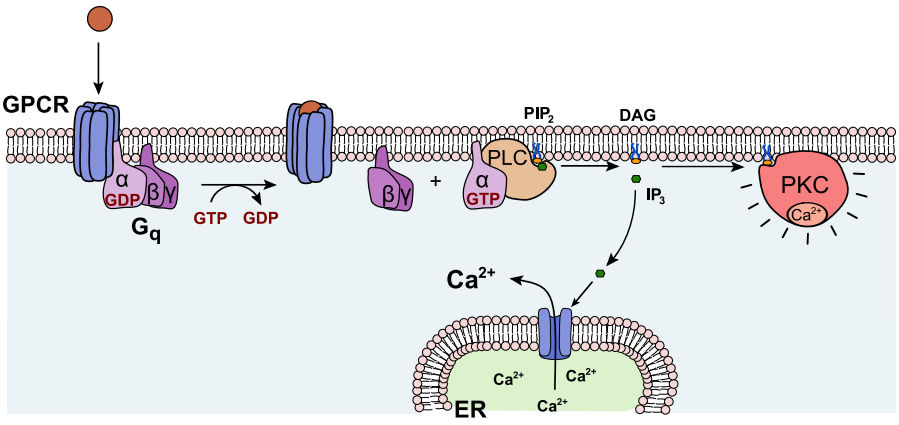


Figure 28.3.12 GPCR activation of phospholipase C, generation of second messengers IP_3, DAG, and Ca^{2+} ions, and downstream activation of Protein Kinase C

You might also expect the regulation of the activation of protein kinase C (PKC) activity to be similar to the regulation of the activation of protein kinase A (PKA). It is, but with a major difference. In contrast to the structure of PKA, which cycles between an inactive R_2C_2 (R is the regulatory and C the catalytic subunit) and an active separated form, PKC is a single chain that has a regulatory, domain, and catalytic domain as shown in Figure 28.3.13 Variants of PKC are also shown.

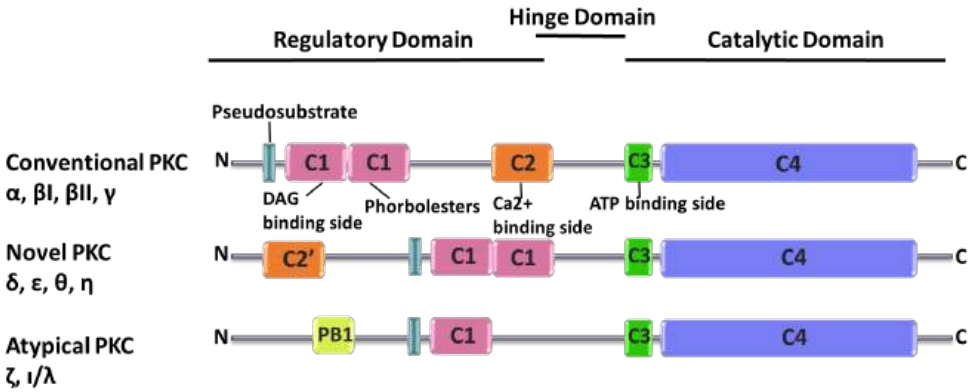


Figure 28.3.13: Domain structure of conventional, novel, and atypical protein kinase Cs. *Kaleli et al. Cells* **2020**, *9*(3), 553; https://doi.org/10.3390/cells9030553https://www.mdpi.com/651098Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

Protein kinase **C** was named "**C**" because of its activation by **C**a^{2+} ions, but you could also remember it because it requires the upstream activation of phospholipase **C**.

There are more than 500 PKCs divided into 15 subgroups and their downstream functions lead to gene expression. They all have 4 common domains, C1-C4. C1 and C2 are the functional regulatory domain and C3 and C4 are in the overall catalytic domain. They have the following activities.

- C1 binds diacylglycerol (DAG) and phorbol esters (commonly known activators or PKC)
- C2 binds Ca^{2+}, another second messenger, which activates the protein; Novel PKCs use DAG and phosphatidyl ethanolamine but not Ca^{2+} to activate the protein, while atypical ones use only DAG.
- C3 binds ATP
- C4 binds target proteins for phosphorylation.

The structure of the phorbol 12-myristate 13-acetate, which activates PKC, is shown in Figure 28.3.14

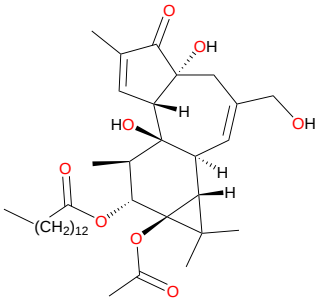


Figure 28.3.14: Phorbol 12-myristate 13-acetate

In a novel way to keep the protein inactive, it contains a small peptide sequence (a pseudosubstrate) that mimics the amino acids around the target phosphorylation site (with the target Ser/Thr). This binds in the active site in the inactive form of PKC like a competitive inhibitor and prevents PKC activity.

PKCα is found in the cytoplasm, cell membrane, nucleus, and mitochondria. Its activation by DAG suggests it becomes localized to membrane surfaces. The protein is conformationally very flexible (note the hinge domain in Figure 28.3.13) so it has been difficult to get detailed structures.

Figure 28.3.15 shows an interactive iCn3D model of the PKC (alpha)-C2 domain complexed with Ca^{2+} and PtdIns(4,5)P2 (IP3) (PDBID 3GPE). The hydrogen bonds and ion-ion interaction between the C2 domain and IP3 are detailed and labeled.

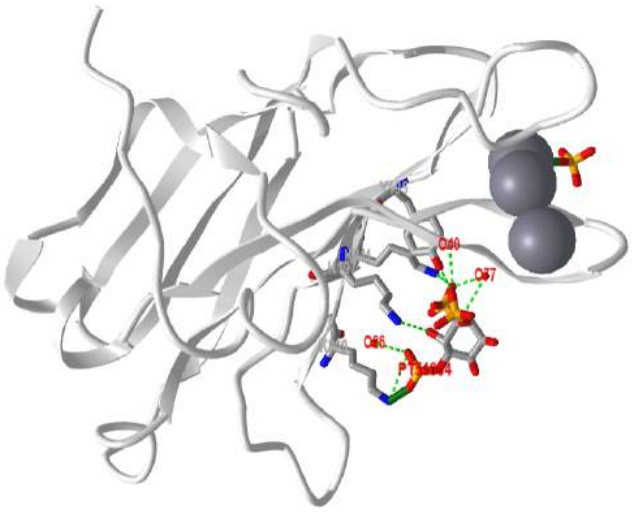


Figure 28.3.15: PKC (alpha)-C2 domain complexed with Ca^{2+} and PtdIns(4,5)P2 (IP3) (3GPE) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...gX24dYC1NhSXi7

Protein kinase C is targeted to the membrane (as shown in Figure 28.3.12), but it lacks the Pleckstrin Homology (PH) which most phospholipase Cs use to interact with PIP_2 in the membrane.

Before activation of PKC by DAG and Ca^{2+}, it must be phosphorylated sequentially. Before these phosphorylations, PKC loosely binds to the membrane with the activation loop open to phosphorylation by kinases. These are the subsequent steps:

1. phosphorylation of Thr500 (PKC βII)) in the activation loop of PKC by an upstream kinase PDK1 (a kinase which also phosphorylates other AGC kinases such as AKT discussed below). This kinase can bind to the exposed activation loop
2. autophosphorylation at the C-terminus of PKC, which leads to conformational positioning of side chains needed for catalysis and substrate binding, and access to the substrate binding site

PKC engages in a very dynamic cycle. It starts as the inactive cytoplasmic form that is autoinhibited by its pseudosubstrate sequence. It then moves to the membrane where the autoinhibition is relieved. All of this requires flexibility which makes it difficult to determine its structure. Figure 28.3.16 shows how the inactive cytoplasmic form of PKC becomes activated at the cell membrane.

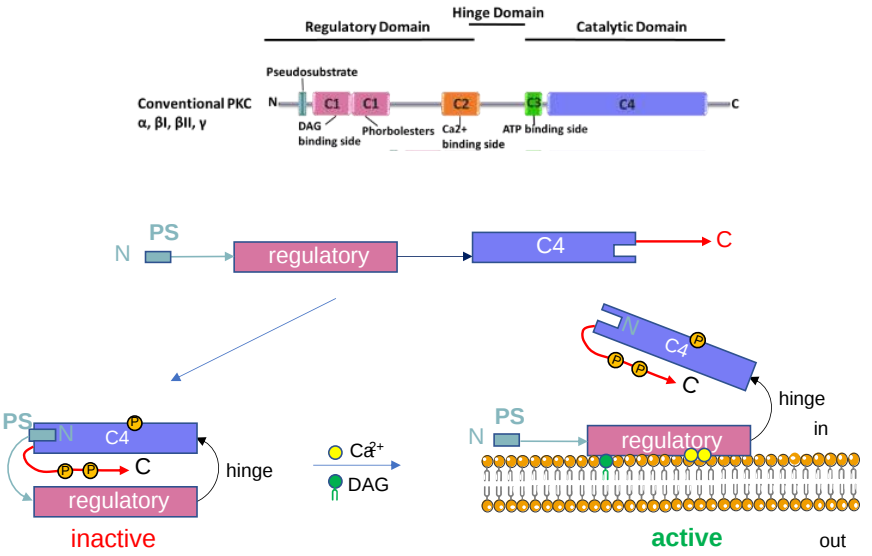


Figure 28.3.16: Conversion of the cytoplasmic inactive PKC to its membrane-bound active form (after Tatyana I. Igumenova. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4979571/)

The regulatory domain, which contains the C1 (DAG binding) and C2 (Ca^{2+}) domains binds to the membrane freeing and activating the kinase domain on the release of the bound internal pseudosubstrate.

Figure 28.3.17 the optimal consensus sequence flanking both sides of the phosphorylation site in target proteins (based on model peptides) and the sequence of the internal pseudosubstrate motif for several PKCs.

PKC isozymes	Position												
	−7	−6	−5	−4	−3	−2	−1	0	1	2	3	4	5
PKCα													
optimal.	R	R	R	R	R	K	G	**S**	F	R	R	K	A
pseudo.	N	R	F	A	R	K	G	**A**	L	R	Q	K	N
PKCβI	F	K	L	K	R	K	G	**S**	F	K	K	F	A
	V	R	F	A	R	K	G	**A**	L	R	Q	K	N
PKCβII	Y	K	L	K	R	K	G	**S**	F	K	K	K	A
	V	R	F	A	R	K	G	**A**	L	R	Q	K	N

Figure 28.3.17: Optimal consensus substrate and actual internal pseudosubstrate sequences for PKC. Nishikawa et al. JBC. 272, 952–960, 1997. https://creativecommons.org/licenses/by/4.0/

Boxed amino acids show structural similarity between the target and internal pseudosubstrate. Position 0 indicates the serine that is phosphorylated in the target. Note that it is replaced with alanine in the pseudosubstrate. Also, note the abundance of positively charged side chains.

Figure 28.3.16 also shows that phosphorylation of key side chains facilitates the activation of the enzyme. The upstream kinase **3-phosphoinositide-dependent protein kinase 1 (PDPK1 or PDK1**) phosphorylates the key Ser/Thr in the activation loop as we discussed above. PDK1 is a "master" Ser/Thr kinase which phosphorylates and activates many proteins, including PKA, PKC, and protein kinase B (which we will explore below).

Figure 28.3.18 shows an interactive iCn3D model of the Protein Kinase C beta II (3PFQ)

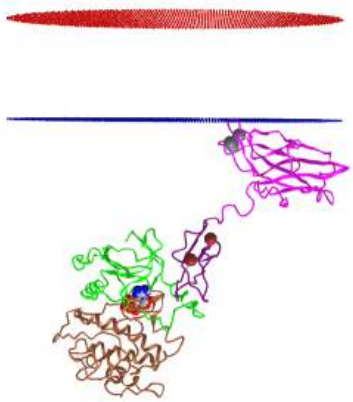


Figure 28.3.18: Protein Kinase C beta II (3PFQ) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...ym5qJhtdWsdN96

The C2 domain (magenta)has two bound Ca^{2+} ions (gray spheres) and interacts with the bottom leaflet of the cell membrane. The C1 domain (purple) has two bound Zn^{2+} ions. The N lobe of the kinase domain is shown in green and the C lobe is shown in brown. A nonhydrolyzable ATP analog, AMPPNP (ANP), is shown in spacefill between the N and C lobes. There is an additional NFD helix preceding the C-terminal tail which can adopt two positions, one which confers low activity and the other high PKC activity. The Phe629 in this region is out of the active site in the low activity form and in it and interacting with the adenine of ATP in the high activity form.

Some signals that activate phospholipase C and make IP3 and diacylglycerol include acetylcholine (a different class than the type located at the neuromuscular junction that we discussed in the last chapter section), angiotensin II, glutamate, histamine, oxytocin, platelet-derived growth factor, vasopressin, gonadotropin-releasing hormone, and thyrotropin-releasing hormone.

In the inactive form of PKC, the arginine-rich basic autoinhibitory pseudosubstrate interacts with acidic side chains in the substrate binding site.An acidic patch in the substrate-binding site (Figure 6.2). When PKC is activated by phosphorylation of the regulatory domain, it phosphorylates arginine-rich sites in protein substrates. This also releases pseudosubstrates from some target inactive protein kinases, which allows them to become active kinases in turn. PKC physiological substrates include receptors, cytoskeleton proteins, protein kinases, proteases, and nuclear proteins

The Ca^{2+} ions also act as second messengers. The calcium ions bind to the **calcium-modulatory** protein, **calmodulin**, which binds to and activates the **calmodulin-dependent kinase (CAM-PK)**, which we will discuss later. Some kinases regulated by calcium and calmodulin include myosin light chain kinase, PI-3 kinase, and CAM-dependent kinases. Ca/CAM also regulates other proteins which include: adenylate cyclase (brain), Ca-dependent Na channel, cAMP phosphodiesterase, calcineurin (phosphoprotein phosphatase 2B), cAMP gated olfactory channels, NO synthase, and plasma membrane Ca/ATPase.

28.3.6: Protein Phosphorylation by activated Receptor Tyrosine Kinases (RTKs)

Figure 28.3.19s shows the dimeric structure of RTKs driven by extracellular signal binding.

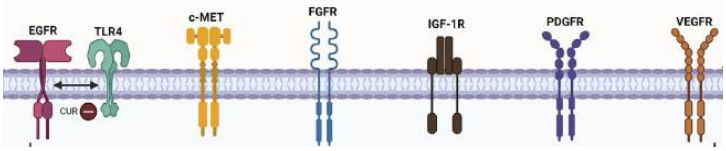


Figure 28.3.19: Dimeric structure of selected RTKs driven by extracellular signal binding. Sareshma Sudhesh Dev et al. Front. Pharmacol., 15 November 2021 | https://doi.org/10.3389/fphar.2021.772510. Creative Commons Attribution License (CC BY).

Table 28.3.2below shows the classification of RTKs into classes and families.

TABLE 1 | Classification of RTKs according to family.

Class	Family	Members
I	EGFR	EGFR, ERBB2, ERBB3, ERBB4
II	Insulin R	INSR IGFR
III	PDGFR	PDGFRα, PDGFRβ, M-CSFR, KIT, FLT3L
IV	VEGFR	VEGFR1, VEGFR2, VEGFR3
V	FGFR	FGFR1, FGFR2, FGFR3, FGFR4
VI	CCK	CCK4
VII	NGFR	TRKA, TRKB, TRKC
VIII	HGFR	MET, RON
IX	EPHR	EPHA1–6, EPHB1–6
X	AXL	AXL, MER, TYRO3
XI	TIE	TIE, TEK
XII	RYK	RYK
XIII	DDR	DDR1, DDR2
XIV	RET	RET
XV	ROS	ROS
XVI	LTK	LTK, ALK
XVII	ROR	ROR1, ROR2
XVIII	MUSK	MUSK
XIX	LMR	AATYK1, AATYK2, AATYK3
XX	Undetermined	RTK106

Table 28.3.2EGFR: epidermal growth factor receptor; InsR: insulin receptor; PDGFR: platelet-derived growth factor receptor; VEGFR: vascular endothelial growth factor receptor; FGFR:fibroblast growth factor receptor; CCK: colon carcinoma kinase; NGFR, nerve growth factor receptor; HGFR: hepatocyte growth factor receptor; EphR: ephrin receptor; Axl: from the Greek word anex-elekto, or uncontrolled, a Tyro3 protein tyrosine kinase; TIE:tyrosine kinase receptor in endothelial cells; RYK: receptor related to tyrosine kinases; DDR: discoidin domain receptor; Ret: rearranged during transfection; ROS: RPTK, expressed in some epithelial cell types; LTK: leukocyte tyrosine kinase; ROR: receptor orphan; MuSK: muscle-specific kinase; LMR: Lemur. Sareshma Sudhesh Dev et al. Front. Pharmacol., 15 November 2021 | https://doi.org/10.3389/fphar.2021.772510. Creative Commons Attribution License (CC BY).

We introduced the activation of RTKs in the previous section and indicated that ligand-induced dimerization led to their activation. Let's expand on that now. As in the case of PKC, the intracellular kinase domains of RTKs are inhibited by specific structures in their chains (cis-autoinhibited). These include the activation loop but in addition C-terminal sequences and the sequence region linking the C-terminal domain with the transmembrane domain. This is called the juxtamembrane region. All of these must be phosphorylated for the activation of kinase activity. The autoinhibition is released on ligand binding and dimerization. The kinase domains can also be allosterically activated by the kinase domain of the dimer. Each domain phosphorylates the cytoplasmic domain of the other, so it's called trans-phosphorylation (i.e. a kinase domain on one monomer does not autophosphorylate itself, which would be called cis-phosphorylation). The now active kinase domains recruit target proteins containing SH2 domains (which bind p-Tyr peptides/proteins) in a fashion that propagates signaling. These include proteins involved in other signaling pathways including MAPK and phosphoinositide-3-kinase (PI3K)/Akt pathways which we will discuss later. A particular phospholipase Cγ (PLC-γ) is also activated by an RTK.

Cancers can arise if the RTK signaling pathways, which control cell growth and division, are overactive. Figure 28.3.20shows four different mechanisms for the expression and/or activation of RTKs that could lead to cancer.

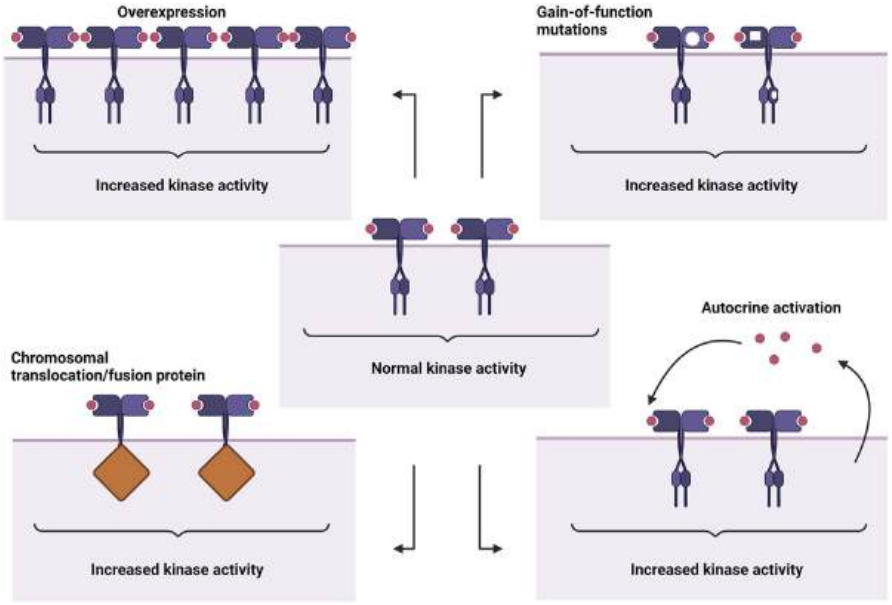


Figure 28.3.20: Possible mechanisms for expression and/or activation of RTKs in cancer cells. Dev et al. Front. Pharmacol., 15 November 2021 | https://doi.org/10.3389/fphar.2021.772510. Creative Commons Attribution License (CC BY).

We will focus on the epidermal growth factor receptor (EGFR) for most of the remaining discussion. One of the most interesting questions is how ligand binding in the extracellular domain of the biotopic integral membrane proteins leads to an intracellular signal in the cytoplasmic domain. It's difficult to imagine such an activation propagating through a single transmembrane helix. We'll discuss how ligand-promoted dimerization of the receptor appears to be the main mechanism of activation of RTKs.

Figure 28.3.21shows the domain structure of three different RTKs, including the insulin receptor (IR) family. The insulin receptor is also synthesized as a single chain but undergoes proteolysis and interchain disulfide bond formation to give the "dimeric" structure shown below.

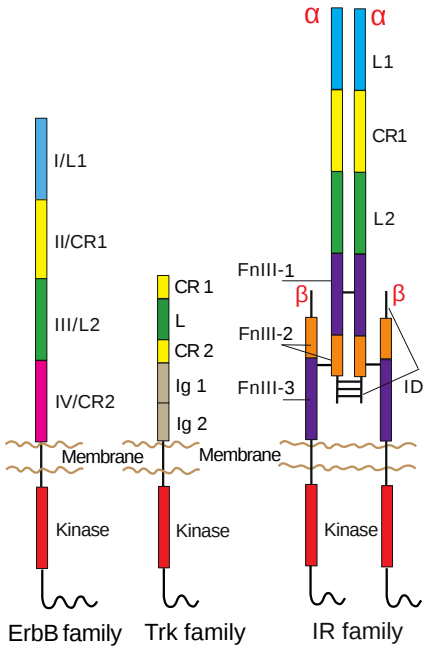


Figure 28.3.21: L,leucine-rich; CR, cysteine-rich; Ig, immunoglobulin-like; FnIII, fibronectin type III; ID, insert domain. The L1, CR1, L2, and CR2 domains of the ErbB family are alternatively termed Domains I–IV. Ichiro N. MaruyamaCells. Cells. 2014 Jun; 3(2): 304-330. doi: 10.3390/cells3020304. (http://creativecommons.org/licenses/by/3.0/). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4092861/

The EGFR is a member of the **ErbB** family which consists of 4 members: EGFR (also called Erb1 or HER1), ErbB2 (HER2), ErbB3 (HER3), and ErbB4 (HER4). The name ErbB arises from the avian **ER**ythro**B**lastosis oncogene **B**). They are also called HER after the **H**uman **E**pidermal growth factor **R**eceptor). The HER2 gene is often dysregulated in breast cancer. Members of the ErbB family have unique numbers and positions of tyrosine in the C-terminal kinase domains. EGFR has 20, of which 12 are phosphorylated. The EGFR is also a bit unique in that it has only one tyrosine in the activation loop that is phosphorylated but the tyrosine itself is not required for kinase activity. Although we suggested earlier that RTKs are activated on dimerization, studies show that RTKs However, an increasing number of studies demonstrate that RTKs exist as inactive dimers in the absence of the ligand.

As shown in Figure 28.3.21, ErbB receptors have 4 extracellular domains, a transmembrane domain, the juxtamembrane region (about 40 residues), the cytoplasmic kinase domain and a C-terminal extension that gets autophosphorylated and which binds downstream target protein through their SH2 domains. Extracellular domains I (L1) and III/L2 have β-helix solenoid secondary motifs that bind the ligand. Domains II/CR1 and IV/CR2 are cysteine-rich disulfide bonds. Some fraction (>80% through cross-linking studies) of the ErbB receptors exist as dimers at the cell surface in the absence of ligand, a finding in contrast to the older view that ligand binding is required for dimerization

The structure of the extracellular domains of the free ErbB and ligand-bound EGFR receptors show conformational changes that are required for dimer formation. In the absence of the ligand, a "tether" arm, denoted by an open triangle in domain IV in Figure 28.3.22 is close to a buried "dimerization" arm (asterisk *) in domain II of the extracellular regions of EGFR, ErbB3, and ErbB4, effectively inhibiting dimerization. When the ligand is bound, domains I and III interact, freeing the dimerization domains to interact (** in EGFR). These features are illustrated in Figure 28.3.22

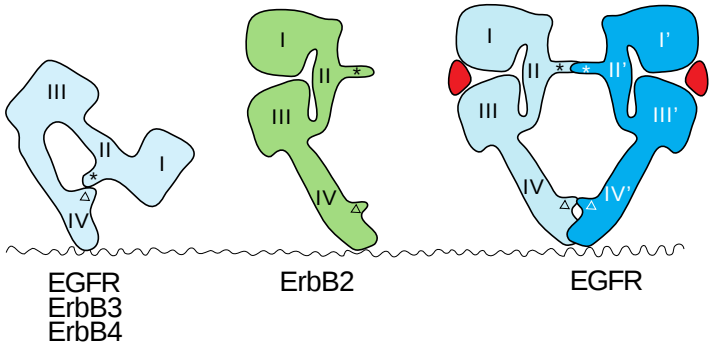


Figure 28.3.12 Schematic representations of the structures of the extracellular regions of the ErbB family. EGFR, ErbB3, and ErbB4 adopt the tethered conformation in the absence of ligand, while ErbB2 adopts an extended, or untethered, conformation that resembles the ligand-activated, dimerization-competent EGFR protomer in the ligand-bound form of the EGFR dimer, shown at the right. The 'dimerization arm' and 'tethering arm' are shown by an asterisk * and an open triangle, respectively. Ligands are shown in red. Domains I–IV correspond to the domains shown in Figure 28.3.21 Not drawn to scale. Ichiro N. MaruyamaCells. Cells. 2014 Jun; 3(2): 304–330. doi: 10.3390/cells3020304' (http://creativecommons.org/licenses/by/3.0/). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4092861/

In addition, the tether arms on both monomers now interact in the dimer. Note that the ligand binding region is not involved in dimer formation and is not in the dimer interface, as it is with other RTKs where the ligand is involved in direct contact in the dimer interface. Mutations in the II/IV domains that inhibit their interactions do not lead to receptor activation so ligand binding is still required. ErbB2 exists in an extended conformation (no ligand is known for it) so it is free to interact with another ErbB chain in a heterodimer.

Structural studies now suggest that EGFR kinase dimer has a symmetrical inactive conformation in which the activation loop is packed and occluding the active site. In addition, it has an asymmetrical active one in which the activation loop is unpacked and the active site is open. How is this transmitted across the subunits? It appears that the C-lobe of the "activator/donor" kinase interacts with the N-lobe of the adjacent "receiver/acceptor" kinase which activates it through a conformational change. When the ligand binds, the inactive dimer dissociates, and the asymmetric active dimer results.

Two models have been proposed for ligand-gated activation of the Erb dimers: the dimerization and rotation/twist models, as shown in Figure 28.3.23

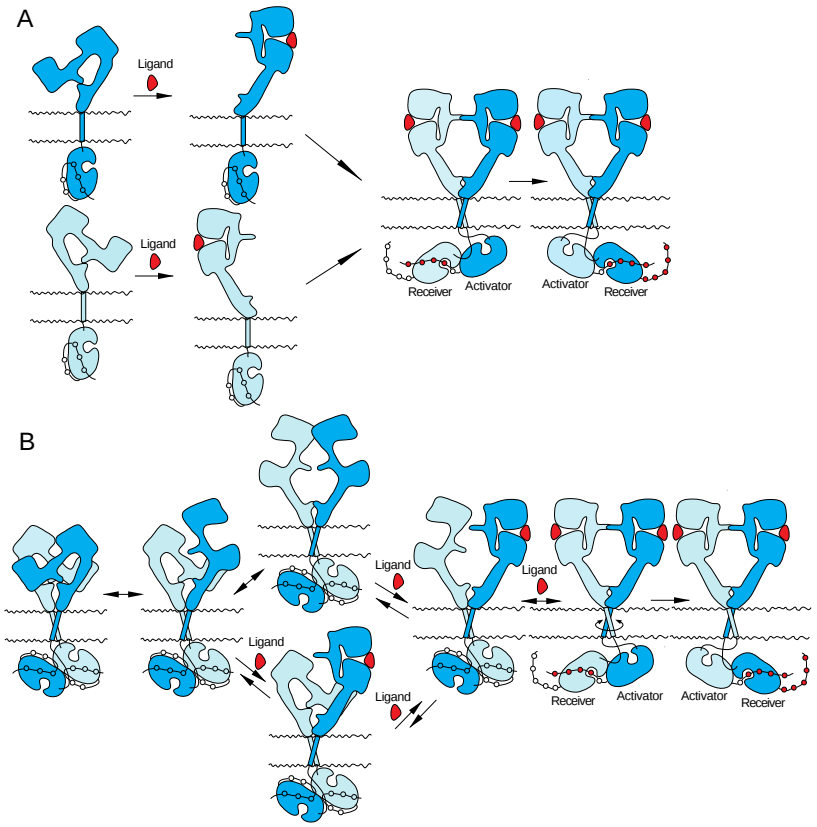


Figure 28.3.23: Models for ligand-induced activation of the ErbB family. (**A**) 'Dimerization' model. (**B**) 'Rotation/twist' model. Not drawn to scale. Ichiro N. MaruyamaCells. Cells. 2014 Jun; 3(2): 304–330. doi: 10.3390/cells3020304' (http://creativecommons.org/licenses/by/3.0/). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4092861/

- A. Dimerization model: Ligand binding to the I and III extracellular domains of a monomer lead to dimerization by causing the release of the tether arm between I and III, allowing the dimerization arms to become free and interact with each other. This causes the kinases domains to adopt the active state
- B. Rotation/Twist model: The receptor is already a dimer in the unliganded state with the extracellular region untethered and the intracellular kinase domains in an inactive symmetric state. Ligand binding causes the two dimerization arms to extend, causing a twist in the transmembrane segment. This causes the kinase domains to adopt the active asymmetric state with the activator kinase domain of one monomer activating the receiver kinase domain of the second monomer in the dimer, with each forming an extended conformation.

Figure 28.3.24 shows an interactive iCn3D model of the intracellular dimeric EGFR kinase domains in complex with an ATP analog-peptide conjugate (2GS6)

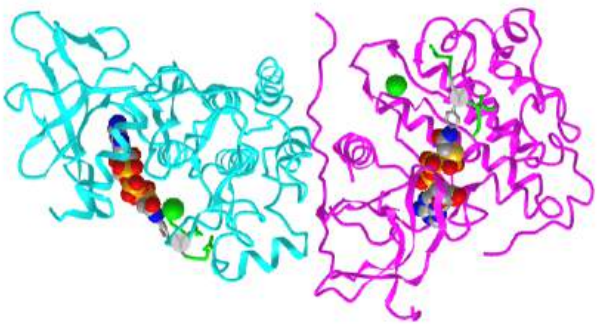


Figure 28.3.24: Dimeric intracellular EGFR kinase domains in complex with an ATP analog-peptide conjugate (2GS6) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...iw1V4T8rsfxDZ7

The two EGFR kinase domains are shown in cyan and magenta. The ATP analogs in each domain (spacefill) are thiophosphoric acid O-((adenosyl-phospho)phospho_-S-acetamidyldiester. The peptide substrates (green stick) are 13-mers with a tyrosine (sticks,

labeled Y, minus the OHs) connected to the ATP analog.

Figure 28.3.25 shows an interactive iCn3D model of a single EGFR kinase domain showing N and C terminal lobes in complex with an ATP analog-peptide conjugate (2GS6)

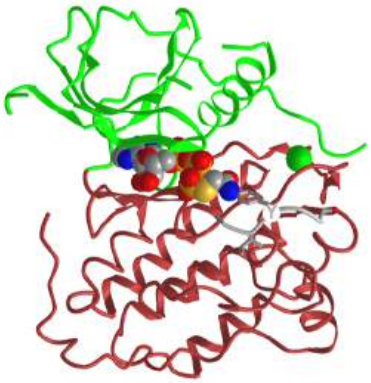


Figure 28.3.25: A single EGFR kinase domain showing N and C terminal lobes in complex with an ATP analog-peptide conjugate (2GS6) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...VLu1jPJYJNoih9

The ATP part of the ATP-peptide conjugate is sandwiched between N-lobe (green) and the C-lobe (brown) just as in PKA. The ATP in the peptide conjugate is shown in spacefill while the peptide is shown in gray. The tyrosine (Y) linked to the peptide is labeled.

Figure 28.3.26 shows an interactive iCn3D model of showing the similarity in structure between the PKA kinase domain (1J3H A chain) and the EGFR kinase domain (2GS6)

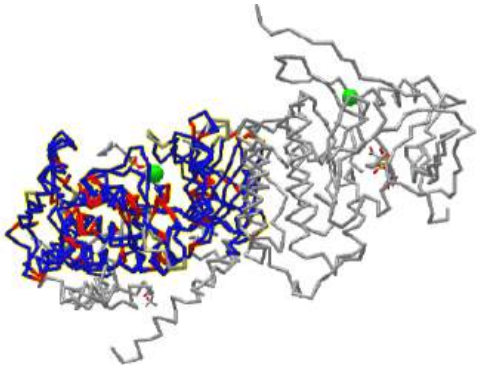


Figure 28.3.26: Alignment of the PKA kinase domain (1J3H A chain) and the EGFR kinase domain (2GS6) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?mVfpwk3Y7EDcM5mq9

The gray structure by itself is the second kinase domain of the EGFR dimer. The superimposed chains are on the other side. Red indicated identical residue and blue nonconserved in the structural alignment. Zoom in on the aligned sequences in blue and red to show how similar the kinase domains are.

In the next chapter section, we will explore the next downstream effects in signaling, mediated by the second messengers cAMP, DAG, and IP_3 and the substrates phosphorylated by the ligand-active receptor tyrosine kinases.

EGFR and HER2 in breast cancer

The HER2 receptor can form homo or heterodimers with single chain ErbB1-ErbB4 (same as HER1-HER4). HER receptors exist as both monomers and dimers, either homo- or heterodimers. Rapid dimerization of HER1, HER3, or HER4 occurs if they form heterodimers with HER2. In addition, any ErbB dimer with HER2 leads to strong intracellular signaling compared to other HER heterodimers. Ligand binding to HERI, HER3, or HER4 induces rapid receptor dimerization, with a marked preference for HER2 as a dimer partner. Since noncancer cells have very little HER2, correspondingly few heterodimers of HER2 around found. If HER2 is overexpressed as in HER2+ breast cancer cells, more heterodimers are found, and anomalously-high levels of signaling occur. This resulted in a poorer prognosis for HER2+ breast cancers in the past.

A revolution in breast cancer therapy has changed that situation. Humanized antibodies (human antibodies made in mice cells) to HER2, called trastuzumab (Herceptin) are now used in therapy. Since the antibodies are derived from human genes, they are not targeted as foreign by the immune system. In early-stage HER2+ breast cancers, the antibody trastuzumab (Herceptin) is administered intravenously periodically for one year. The antibody selectively binds to HER2 on breast cancer cells. Once bound, the Fc portion of the bound antibody signals the immune system, leading to the recruitment of immune cells and modulators to the tumor cell, leading to its destruction.

If the cancers are in a later stage or if tumor cells are found in lymph nodes, a variant of the anti-HER2 antibody can be given in which a chemotherapeutic drug is covalently attached to the antibody. The structure of the drug Ado-trastuzumab emtansine (T-DM1), trade name Kaycyla, is shown in Figure 28.3.27.

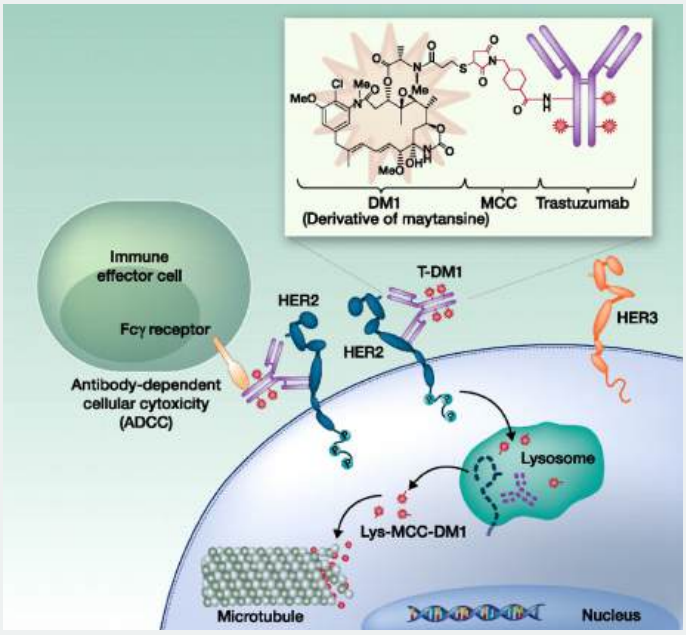


Figure 28.3.27: Ado-trastuzumab emtansine (T-DM1) and its interaction with HER2 positive turmor cells. Sadeghi et al. Pharmacogenomics and Personalized Medicine. 2014 Oct 15;7:329-38. doi: 10.2147/PGPM.S47524. Creative Commons Attribution – Non Commercial (unported, v3.0) License. http://creativecommons.org/licenses/by-nc/3.0/.

When the antibody-drug complex binds to the receptor on cancer cells, it is taken up into the cell. The antibody is degraded and the toxic drug is released into the cell sparing all other types of cells (only ones with the receptor on it are targeted). The released drug binds to microtubules comprising part of the internal cytoskeleton of the cell and prevents changes necessary for cell division. HER2 has a very low level of expression in noncancer cells so side effects can occur. However, these are much less server than traditional chemotherapy, which targets all dividing cells. The antibody acts as a homing device bringing the attached chemotherapy predominantly to tumor cells. It can be likened to a smart bomb or a cruise missile guided to just one target.

A better image of DMI and the linker connecting to the antibody is shown in Figure 28.3.28 The toxic chemotherapy drug is chemically linked to the antibody through a non-reducible thioether linker, N-succinimidyl-4-(N-maleimidomethyl) cyclohexane-1-carboxylate (SMCC).

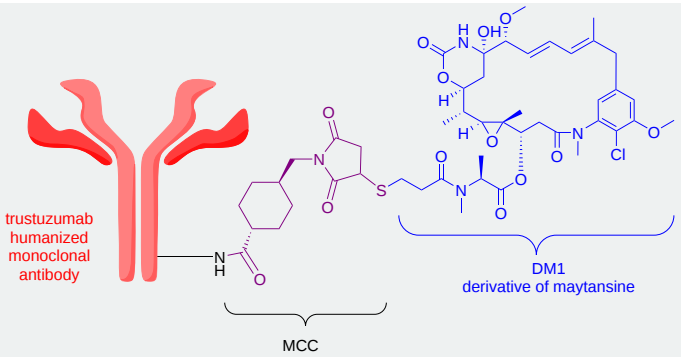


Figure 28.3.28: Ado-trastuzumab emtansine (T-DM1) conjugate used in HER2+ cancers

Figure 28.3.29shows a mechanism for the actual cross-linking reaction. The reaction of the maytansine derivative through its free thiol and the human antibody (trastuzumab) through a free amine should be readily understandable base on the reactions present in Chapter 2.

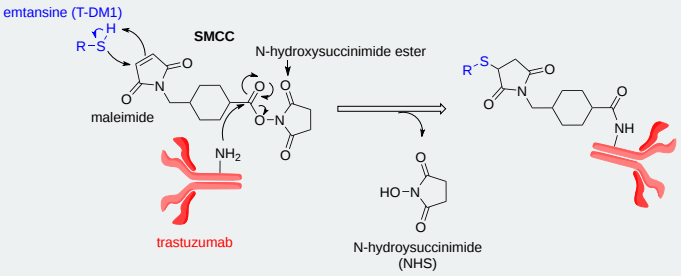


Figure 28.3.29: Cross linking reactions to connect emtansine and trastuzumab

The final linker after the departure of the N-hydroxysuccinimide is designated MCC.

The C-MET receptor, illustrated as an example in Figure 19, is an RTK involved in cell proliferation and survival, and as such, alterations in its expression can lead to cancer. Its mature form results from selective proteolysis by furin. Its physiological ligand is hepatocyte growth factor (HGF), a multisubunit protein in its mature form. It also binds a naturally occurring smaller splicing variant of HGF called NK1. The structure of C-MET bound to both HGF and NK1 has been solved. Binding of either lead to dimerization of the C-MET receptor, as illustrated in the cartoon representation of Figure 28.3.30

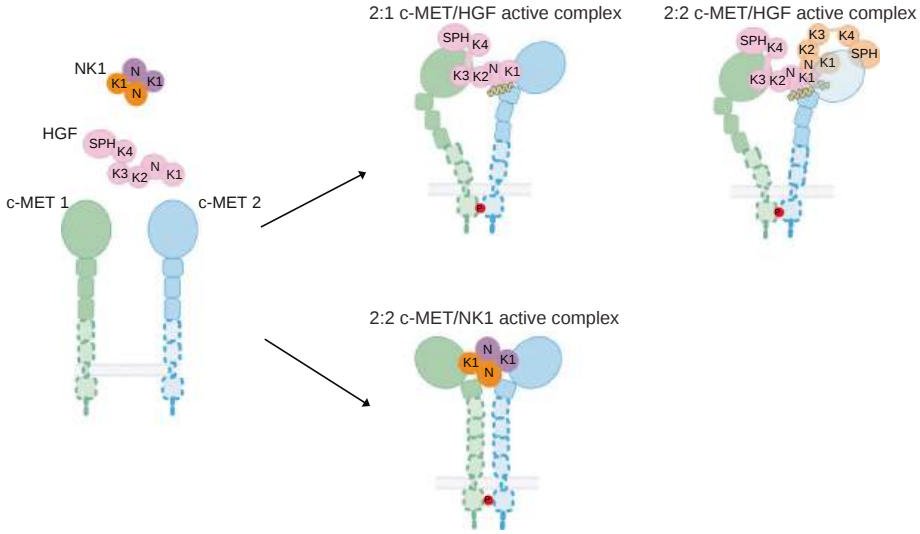


Figure 28.3.30: Cartoon representation of a working model for HGF and NK1-induced c-MET activation. HGF induces an asymmetric c-MET signaling complex, whereas NK1 induces a symmetric c-MET signaling complex. The bindings of a second HGF and heparin strengthen the 2:1 c-MET/HGF active complex. Uchikawa et al. Nature Communications (2021) 12:4074 | https://doi.org/10.1038/s41467-021-24367-3 . Creative Commons Attribution 4.0 International License, http://creativecommons.org/ licenses/by/4.0/.

The glycosaminoglycan heparin enhances c-MET activation by HGF. This is illustrated in Figure 30 above. Heparin binds between the N domain of HGF I and the IPT1 domain of c-MET II, facilitating the interactions between the two domains. In addition, a long enough heparin chain could bridge both HGF I and II, further strengthening the full complex. Structural representations of the c-MET:HGF asymmetric dimer are shown in cartoon form in Figure 28.3.31

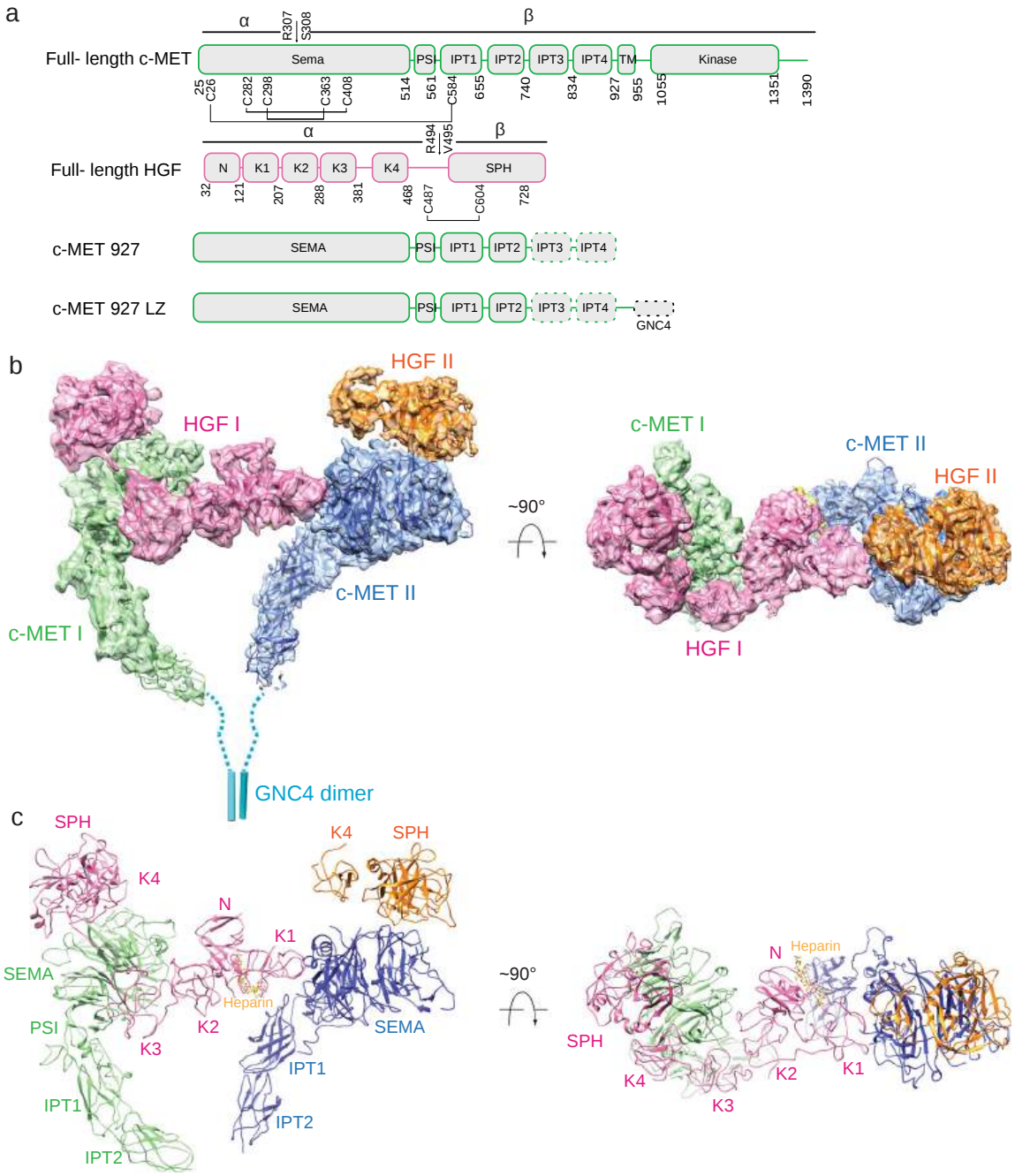


Figure 28.3.31: Overall structure of the c-MET/HGF holo-complex.

Panel (a) shows the domain structures of human c-MET and HGF. The proteolytic cleavage site of c-MET is located between Arg307 and Ser308. The proteolytic cleavage site of HGF is located between Arg494 and Val495. C-MET927-LZ and full-length HGF were used for structural determination in this study. The dash boxes indicate the domains that were unsolved in cryo-EM maps. Panel (b) shows the 3D reconstruction of the 2:2 c-MET/HGF holo-complex and the corresponding ribbon representation of this complex fitted into the cryo-EM map at 4.8 Å resolution, shown in two orthogonal views. Panel (c) shows the ribbon representation of the c-MET/HGF holo-complex shown in two orthogonal views

Figure 28.3.32 Shows the structure of the c-MET/NK1 symmetrical dimer.

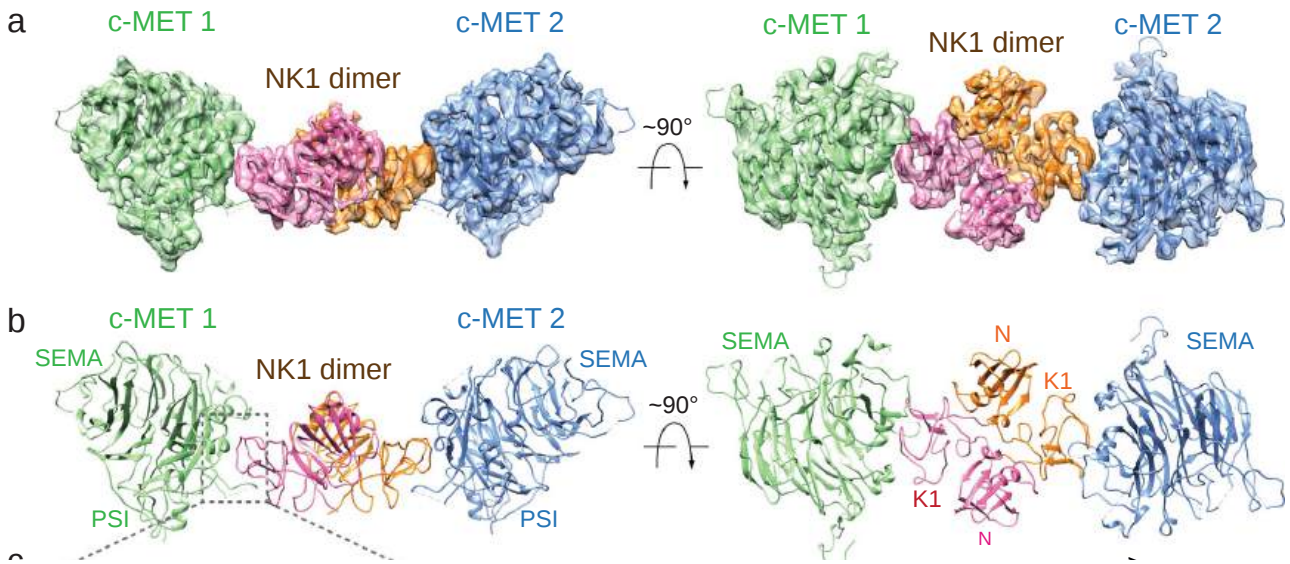


Figure 28.3.32: Overall structure of the c-MET/NK1 complex. a 3D reconstruction of the 2:2 c-MET/NK1 complex, and the corresponding ribbon representation of this complex fitted into the cryo-EM map at 5 Å resolution, shown in two orthogonal views. b The ribbon representation of the 2:2 c-MET/NK1 complex is shown in two orthogonal views.

Nuclear RTKs

What makes signal transduction so complicated yet interesting is the unexpected. It turns out that some RTKs (EGFR, VEGFR, FGFR, IR, and NGFR have been found in the nucleus. It's experimentally easy to localize proteins in cells using immunofluorescence microscopy. It's hard to determine their functions. ErbB-2 is one that is also found in the nucleus. Kinase inhibitors blocked the expression of ErbB-2 in the nucleus suggesting that its kinase activity is required for it to translocate to the nucleus. The structure of the membrane forms of ErbB-2, EGFR, and ErbB-3 appear to be the same as the structure of the nuclear forms.

The carboxy-terminal ends of EGFR and ErbB-4 can activate gene transcription as measured by the expression of luciferase (a fluorescent protein) reporter proteins. Hence they appear to act as transcription factors that bind DNA to promote gene expression. For example, nuclear EGFR increases the expression of cyclin D1 which drives progression through the cell cycle. ErbB-2 appears to activate transcription from the promoter of the gene for cyclooxygenase 2 (COX-2). The protein COX-2 leads to the synthesis of inflammatory prostaglandins. It also increases blood vessel growth and is dysregulated in tumors.

Since proteomic analysis of these growth factor receptors shows no DNA binding motifs or domains, they must promote gene transcription through binding to other protein transcription factors in the nucleus.

28.3.7: Back to AGC Kinase - AKT (Protein Kinase B)

We've just explored the:

- activation of Protein Kinase A (an AGC Kinase) through GPCR signaling, activation of the integral membrane protein adenylyl cyclase, production of the second messenger cAMP, which binds to inactive PKA and leads to its activation
- activation of Protein Kinase C, (an AGC Kinase) again through GPCR signaling, which leads to activation of the peripheral membrane protein phospholipase C, production of the second messengers DAG and IP3 from PIP_2, and activation of PKC at the membrane.
- activation of receptor tyrosine kinases (RTKs) leading to autophosphorylation of the cytoplasmic kinase domain, followed by recruitment of downstream signaling proteins through binding the phosphorylated RTKs through the downstream protein's SH2 domain

Now let's explore another AGC kinase called **AKT** or **protein kinase B (PKB)** that links signaling through RTKs to phosphoinositol-related signaling in a fashion similar to the link between phospholipase C and Protein Kinase C activities. In the process, we will introduce in this section our first nonprotein kinase involved in signaling. It's a **lipid** kinase called **phosphoinositide 3-kinase (P13K)** and it's very important.

Since Protein Kinase B is usually referred to as AKT, we will stick with that abbreviation. As with other AGC kinases, AKT is a Ser/Thr protein kinase. There are three variants, AKT1, AKT2, and AKT3. These are involved in metabolism, growth, and proliferation so they are key players in signaling. It is a key player in the uptake of glucose into cells as it regulates the translocation of the glucose transporter SLC2A4/GLUT4 to the cell surface in response to insulin.

You would expect aberrant expression of these would lead to cancer. The abbreviation AKT appears to derive from "a serine/threonine protein kinase encoded by the oncogene in the transforming retrovirus isolated from the thymoma cell line AKT-8, which is derived from the Stock A Strain k AKR".

As with Protein Kinase C as phospholipase C, AKT is recruited to the inner leaflet of membranes. Recruitment is mediated by its binding through its pleckstrin homology (PH) domain to phosphatidylinositol (3,4,5)-trisphosphate, a modified form of PIP_2, abbreviated either as PtdIns(3,4,5)P_3 or more simply as **PIP_3**. Note that PIP_2 has 3 phosphate groups while PIP_3 has 4.) The mechanism of membrane recruitment is similar to that of phospholipase C, which also binds membrane PIP_2 through its pleckstrin homology (PH). (This is in contrast to PKC which is recruited through its C1 and C2 domains.)

PIP_3 is generated in the membrane from PIP_2 by the enzyme **phosphoinositide 3-kinase (P13K)**, a **lipid** kinase, whose own activation occurs through stimulation of insulin and growth factors receptor tyrosine kinases (RTKs). Class 1 PI3K has a regulatory/adapter subunit (p85) and a 110 kDa catalytic subunit (p110). The regulatory subunit has SH2 domains which recruit it to autophosphorylated RTKs.

Figure 28.3.33 shows how membrane PIP_2 can be converted to the second messengers DAG and IP_3 by phospholipase C, or to PIP_3 by the enzyme phosphoinositide 3-kinase (P13K), which is a **lipid** kinase.

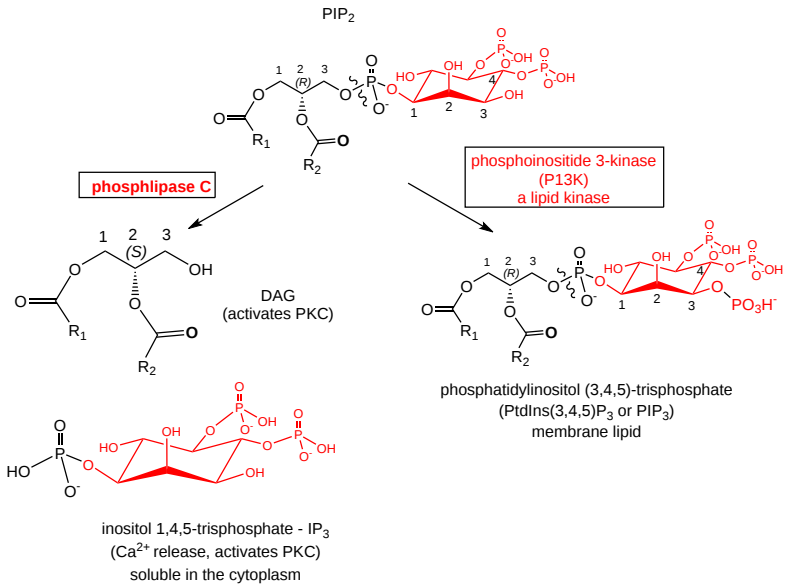


Figure 28.3.33: Conversion of membrane PIP_2 to the second messengers DAG and IP_3 by phospholipase C, or to PIP_3 by phosphoinositide 3-kinase (P13K)

The binding of AKT (PKB) to inner leaflet PIP_3 through its PH domain causes a conformational change that activates AKT for phosphorylation by **phosphoinositide-dependent kinase 1 (PDK1)** a membrane protein kinase. Once activated. AKT dissociates from the membrane and acts enzymatically in the cytosol and nucleus. The overall activation of AKT is shown in Figure 28.3.34

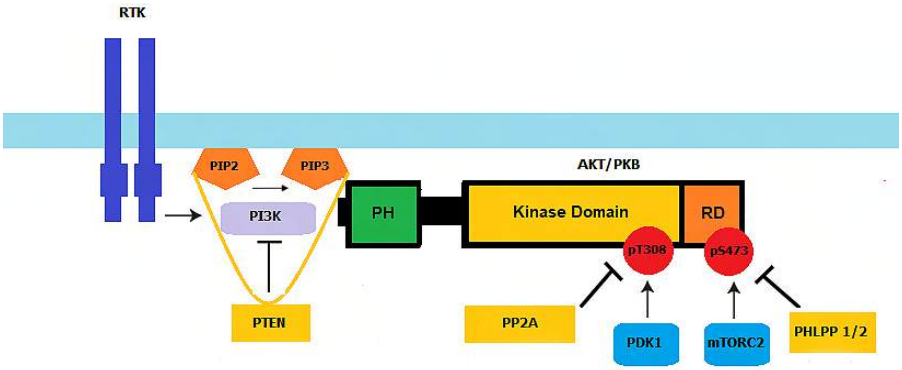


Figure 28.3.34: Activation of AKT through RTK signaling and PI3K activation https://commons.wikimedia.org/wiki/F...activation.png

.

The blunt arrows in the figure above show inhibition by the enzymes indicated. These (PTEN, PP2A, and PHLPP 1/2) are phosphatases.

- PTEN is lipid phosphatase (phosphatidylinositol 3,4,5-trisphosphate 3-phosphatase), which removes a phosphate from the lipid PIP_3.
- PP2A and PHLPP 1/2 are protein phosphatase that removes phosphates added by the kinase PDK1 in the kinase domain and mTORC2 in the C-terminal domain.

Figure 28.3.35shows an interactive iCn3D model of AKT bound to a novel allosteric inhibitor is shown below (3o96).

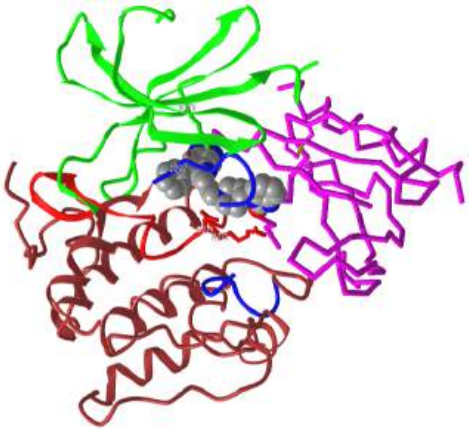


Figure 28.3.35: AKT bound to a novel allosteric inhibitor is shown below (3o96) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...wTWYdpU1Y5JnP9

The PH domain (at N-terminus, bind to PIP3) is shown in magenta, the N-lobe in green, and the C-lobe in brown. The activation loop is in dark blue and the catalytic loop is in red. The blue activation loop is shown in two parts as the interior part, which contains T305 (equivalent to T197 in AGC kinases without a PH domain) is too disordered to resolve. Table 28.3.1 shows the numbering of key amino acids and features of generic AGC kinase and the corresponding numbers in AKT. They are different since AKT has an N-terminal PH domain.

Table 28.3.1

SITE	Generic	Akt (+108)
N lobe	K72	K180
N lobe	E91	E199
C lobe, cat loop	R165	R273
C lobe, cat loop	D166	D274
C lobe, act loop	D184	D292

SITE	Generic	Akt (+108)
C lobe, act loop	T197	T308* missing in this structure)
Approx Cat Loop	163-179	271V—287H
Approx Act loop Start DFG (292-294) to APE	184-200	292-308 308Tmiss toAPE end 319

The allosteric inhibitor shown in the structure above is especially interesting in that it requires both the PH domain as well as the kinase domains for its effect.

Figure 28.3.36 shows an interactive iCn3D model of the structural overlap between the inactive form (shown above, 3o96 containing a bound allosteric inhibitor) with an active form of AKT(3cqw), which has a bound substrate (from glycogen synthase kinase-3 beta, yellow spacefill).

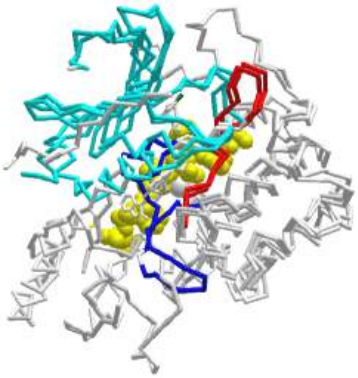


Figure 28.3.36: Alignment of inactive AKT (3o96 with bound allosteric inhibitor) and active AKT(3cqw with bound substrate) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...V2LbSdXLdTA9NA

The bound decapeptide substrate (GRPRTT**S**FAE) in the active form becomes phosphorylated on the **Ser** chain by active AKT. The N-lobe is shown in cyan, while the catalytic loop is in red and the activation loop is in blue. By pressing the "a" key you can toggle between the inactive 3o96 form and the active 3cqw forms. Note the large change in the blue activation loop.

Figure 28.3.37 below shows a series of coupled equilibria reactions that regulate the activity of AKT1.

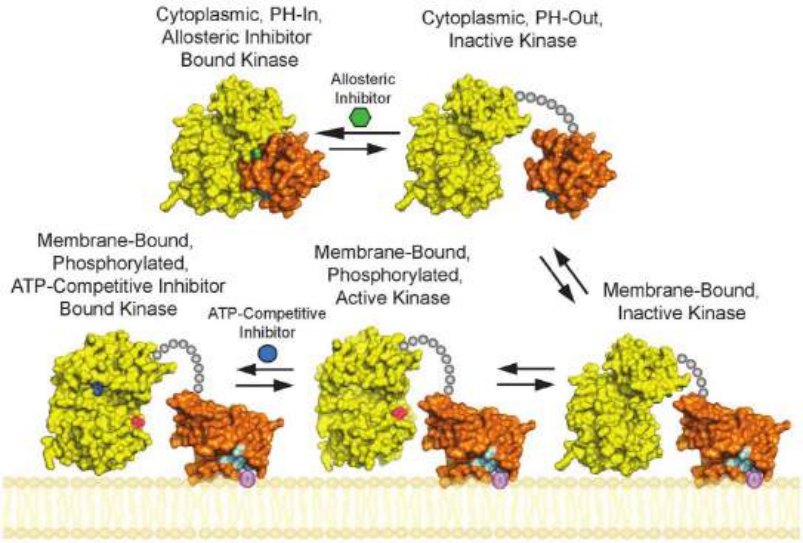


Figure 28.3.37: **Citation:** Wu W-I, Voegtli WC, Sturgis HL, Dizon FP, Vigers GPA, Brandhuber BJ (2010) Crystal Structure of Human AKT1 with an Allosteric Inhibitor Reveals a New Mode of Kinase Inhibition. PLoS ONE 5(9): e12913. https://doi.org/10.1371/journal.pone.0012913. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

The top part of the figure shows the cytoplasmic, nonmembrane-bound form of the enzyme. The PH domain is shown in orange. The top-right figure shows the inactive N- and C-lobes of the kinase loosely interacting with an "out" (or away) conformation of the PH domain with respect to the kinase domains. In the presence of the allosteric inhibitor (green hexagon, green bound ligand), the kinase domains tightly interact with the PH domain in the "in" conformation.

The bottom three structures show AKT bound to the membrane through the interaction of the PH domain with PIP3 (purple). The bottom right kinase domains are identical in representation to the two in the top part of the figure, showing that they are inactive. Only when bound to the membrane is AKT phosphorylated on Thr 308 (red in the bottom middle figure), which activates the enzyme. The bottom left and middle structures show the yellow kinase domains in a different conformation (3cqw), both of which are phosphorylated (red). The active form can bind ATP and protein substrates for phosphorylation. It can also bind ATP analogs which would competitively inhibit the active form of the enzyme by occupying the ATP binding site.

The activation loop in the inhibited form is missing part of its sequence which reflects its disorder. In this state, the loops partially occludes substrate interactions. On phosphorylation of Ser 308 in the activation loop, the loop adopts a different conformation which allows less restricted access to the active. The loop itself on phosphorylation has more local ordering as it shifts away from the active site as seen in the iCn3D model above. Another change decreases inhibitory noncovalent interactions of activation loop amino acids with catalytic residues, which increases catalytic efficiency. In summary, these two types of changes in the activation loop lead to more access by substrates and enhanced catalysis of bound substrates.

Additional regulation of the kinase occurs through the PH domain which adds additional conditions on Akt conformational changes and subsequent activity. The PH domain "appears to lock the kinase in an inactive conformation and the kinase domain disrupts the phospholipid binding site of the PH domain".

Figure 28.3.38 shows an interactive iCn3D model of the separate AKT Pleckstrin Homology (PH) domain bound to the inner member through just the head group of PIP3 (inositol (1,3,4,5)-tetrakisphosphate) (1unq). Waters H bonded to the ligands is not shown.

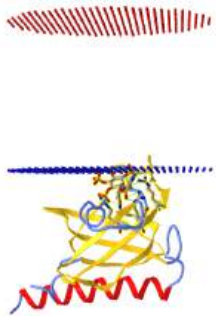


Figure 28.3.38 AKT Pleckstrin Homology (PH) domain bound to inositol (1,3,4,5)-tetrakisphosphate (1unq) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...h9LwZyV9WrJsN6

28.4: The next step - Downsteam intracellular signaling

28.4.1: Intracellular signaling from activated PKA and PKC

We discussed the basics of the activation of the protein kinase A holoenzyme (R_4C_4) by cAMP binding to the regulatory subunit, which frees the catalytic subunit C for activity. Likewise, we discussed the activation of PKC at the cell membrane by DAG, Ca^{2+} ions, and phosphorylation of key Ser/Thr in the protein. Where in the cell are the downstream protein targets of activated PKA and PKC? This is a much simpler question for RTKs since downstream signaling proteins come to them through the interaction of their SH2 domains with the autophosphorylated RTKs. For activated PKA and PKC, it turns out that their location is controlled by **scaffolding proteins**, which bind them either before their activation or after.

Let's discuss a particularly important scaffolding protein, the **A-kinase-anchoring protein (AKAP)**. There are 13 classes of these containing 50 different members. These proteins bind PKA through its regulatory subunit, where cAMP can mediate the activation of the holoenzyme (R_4C_4). In addition, AKAPs can bind other signaling proteins including PKC and phosphatases, the latter of which in turn counter-regulate signaling by phosphoproteins. For example, the bound phosphatases can dephosphorylate PKC to deactivate it as well as other downstream phosphoproteins. AKAPs can also bind phosphodiesterase, the enzyme that converts cAMP to AMP, returning signaling to baseline levels. AKAPs localize key signaling enzymes to sites where biologically appropriate protein targets are localized. In addition, they decrease indiscriminate phosphorylation of other off-target proteins elsewhere in the cell. They may also allosterically regulate the activity of bound signaling proteins.

There are at least 50 **A-kinase anchoring proteins** or **A-kinase anchor proteins** (**AKAPs**) that bind the regulatory subunit of protein kinase A (PKA) and localize PKA to specific sites in the cell. By binding multiple signaling enzymes at specific sites, they integrate signaling pathways mediated by cAMP (for example) with others mediated by PKC (again for example).

Here are some examples of AKAPs in humans (from UniProt). Note that one (12) also binds PKC

- **1, mitochondrial:** Binds to type I and II regulatory subunits of protein kinase A and anchors them to the cytoplasmic face of the mitochondrial outer membrane;
- **6:** Binds to type II regulatory subunits of protein kinase A and anchors/targets them to the nuclear membrane or sarcoplasmic reticulum;
- **7 isoforms alpha and beta:** Targets the cAMP-dependent protein kinase (PKA) to the plasma membrane, and permits functional coupling to the L-type calcium channel;
- **7 isoform gamma:** targets cAMP-dependent protein kinase (PKA) to the cellular membrane or cytoskeletal structures;
- **8:** Acts as an anchor for a PKA-signaling complex onto mitotic chromosomes, which is required for the maintenance of chromosomes in a condensed form throughout mitosis;
- **8-like:** Required for cell cycle G2/M transition and histone deacetylation during mitosis and recruitment of signaling enzymes into the nucleus;
- **9:** assembles several protein kinases and phosphatases on the centrosome and Golgi apparatus;
- **12:** Anchoring protein that mediates the subcellular compartmentation of protein kinase A (PKA) and protein kinase C (PKC)
- **17A:** Splice factor regulating alternative splice site selection for certain mRNA precursors. Mediates the regulation of pre-mRNA splicing in a PKA-dependent manner.

Figure 28.4.1 illustrates the localization/binding of signaling enzyme (PKA, PKA substrates, PDE, other kinases) to AKAPs.

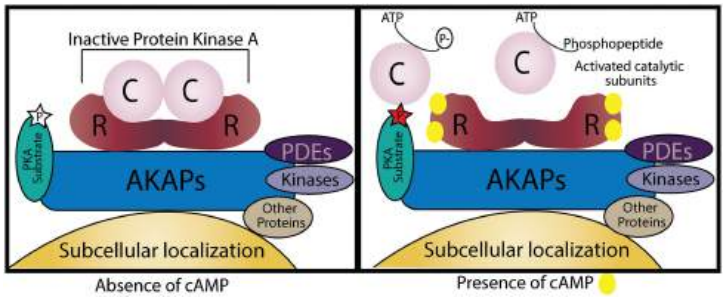


Figure 28.4.1: Localization/binding of signaling enzyme (PKA, PKA substrates, PDE, other kinases). Calejo et al. Frontiers in Pharmacology 6:192 (2015) DOI: 10.3389/fphar.2015.00192. License CC BY 4.0

Note that some AKAPs can also bind PKA substrates, facilitating their phosphorylation and minimizing the phosphorylation of the wrong targets.

AKAPs use an amphiphilic helix to interact with the R_2 regulatory dimer of the PKA. Some AKAPs bind to just one of the regulatory subunits. Note that some AKAPs can also bind PKA substrates, facilitating their phosphorylation and minimizing the phosphorylation of the wrong targets.

Figure 28.4.2 shows specific AKAP complexes in the heart that could be targeted for drug therapies.

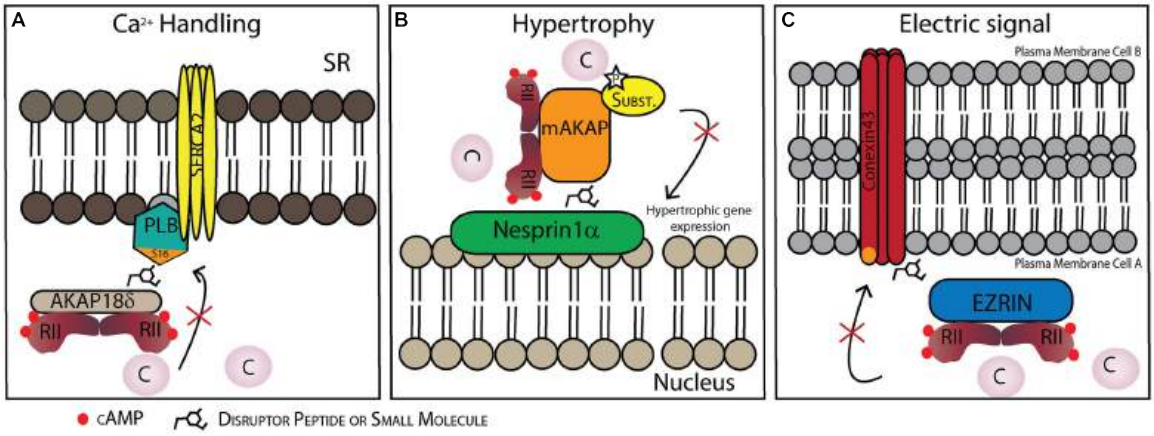


Figure 28.4.2: AKAP complexes in the heart.

Panel (A): Disruption of the AKAP18γ/δ-PLB (another phospholipase involved in signaling) interaction prevents PLB phosphorylation on Ser16 and dislocation from SERCA2 **(Sarcoplasmic/endoplasmic reticulum calcium ATPase 2)**. This inhibits SERCA2 activation and consequently Ca^{2+} uptake into the sarcoplasmic reticulum

Panel (B): Disruption of the nesprin-1α /mAKAP interaction promotes AKAP/PKA complex dissociation from the perinuclear membrane and might be a strategy to reduce hypertrophy. Nesprin 1 is a protein that forms a linking network between organelles and the actin cytoskeleton to maintain the subcellular spatial organization.

Panel (C): Disruption of the connexin 43-ezrin interaction could prevent PKA-mediated phosphorylation increasing inter-cardiomyocyte conductivity which could be cardioprotective following myocardial infarction damage. Connexin is a gap junction protein. Ezrin is involved in the connections of major cytoskeletal structures to the plasma membrane.

To add to the complexity of PKA activation and signaling, there are different forms of the regulatory subunits of the holoenzyme PKA. These include RIalpha (RIA), RIbeta (RIB), RIIalpha (RIIA), and RIIbeta (RIIB). They have different affinities for cAMP, the catalytic subunits Cs, and different AKAPs.

Figure 28.4.3 shows an interactive iCn3D model of the amphiphilic anchoring peptide AKAP-IS for AKAP binding to the docking and dimerization (D/D) domain of the RIIalpha regulatory subunit of PKA (2IZX)

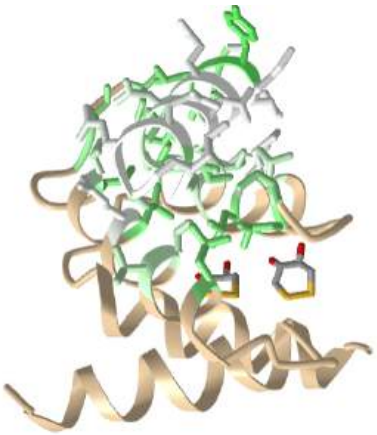


Figure 28.4.3: Amphiphilic anchoring peptide AKAP-IS for AKAP binding to the docking and dimerization (D/D) domain of the RIIalpha regulatory subunit of PKA (2IZX) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...pWKiYgQeGA32e6

The brown represents the RII dimer D/D domains of the regulatory subunit. The anchoring peptide AKAP-IS is shown in gray. In both, the side chains involved in binding of the peptide to the regulatory subunit domains are shown as sticks and colored coded based on the hydrophobicity of the side chains. Green indicates the most hydrophobic. Rotate the model carefully to differentiate the side chains and note that the hydrophobic face of the peptide is interacting with hydrophobic side chains projecting into a groove made by the two RII dimer D/D domains. Polar side chains in AKAP help target the correct isoform of the R subunit.

In addition to binding to some AKAP scaffolds, PKC also binds to **Receptors for Activated C-Kinases (RACKs)**, heat shock proteins (HSP), importing, and annexins (AnxA1, A2, A5, and A6). The interactions of activated PKC with RACK1 and downstream events are shown in Figure 28.4.4

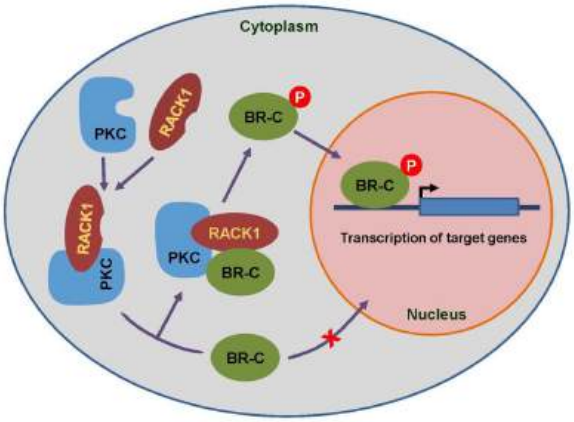


Figure 28.4.4: Interactions of PKC with RACK1 and downstream signaling. Meng et al. PLoS ONE · (2014) DOI: 10.1371/journal.pone.0109111. Creative Commons Attribution License - https://creativecommons.org/licenses/by/4.0/

The insect protein BR-C (Broad Complex) has a DNA binding domain (two zinc fingers) domain for the activation of gene transcription and a BTB) domain that allows binding to RACK1. On binding the PKC:RACK1 complex, BR-C is phosphorylated at Ser373 and Thr406, after which it is translocated into the nucleus where it binds DNA and activates gene transcription

The binding of PKC to RACK1 stabilizes PKC for the phosphorylation of targets. PKC binds to RACK through its C2 regulatory domains. Binding may be to specific forms of PKC including unphosphorylated, inactive, and activated phosphorylated forms, as well as to specific isozymes of PKC. RACK1 may also recruit PKC to the ribosome and it inhibits the activity of SRC kinases which we will discuss later. PKC activity occurs in many cellular locations, including the cell membrane, nucleus, Golgi apparatus, mitochondria, and cytosol. RACKS also bind and recruit other signaling proteins including PLCγ, Src, and integrins. In addition to interactions of PKC with RACK through the C2 domain, PKC can localize through its C1 domain.

Structure of RACK1

RACK1 (317 amino acids) has a very interesting structure. It is a member of a family of over 100 proteins that have tryptophan-aspartate (WD) repeats that are 44-60 amino acids and ending in WD or a variant. It is homologous to the beta subunit of the

heterodimeric Gαβγ signaling protein. RACK 1 interacts with proteins through a 7-bladed propeller structure that allow the binding of proteins with SH2 domains, plextrin homology (PH) domains, and C2 domains (PKCs). Figure 28.4.5 shows the WD repeats in human RACK1. Note that the N-terminal end of the WD repeat often is glycine-histidine (GH).

Human RACK1

```
wd 1   MTEQMTLRGTLKGHNGWVTQIATTPQFPDMILSASRDKTIIMWKLTRD
wd 2   ETNYGIPQRALRGHSHFVSDVVISSDGQFALSGSWDGTLRLWDLTTG
wd 3   TTTRRFVGHTKDVLSVAFSSDNRQIVSGSRDKTIKLWNTLGVC
wd 4   KYTVQDESHSEWVSCVRFSPNSSNPIIVSCGWDKLVKVWNLANCKLK
wd 5   TNHIGHTGYLNTVTVSPDGSLCASGGKDGQAMLWDLNEG
wd 6   KHLYTLDGGDIINALCFSPNRYWLCAATGPSIKIWDLEGKIIVDEL
wd 7   KQEVISTSSKAEPPQCTSLAWSADGQTLFAGYTDNLVRVWQVTIGTR
```

Figure 28.4.5: Sequence of RACK1 depicting WD-repeats. R. Adams, D.R., Ron, D. & Kiely, P.A. RACK1, A multifaceted scaffolding protein: Structure and function. *Cell Commun Signal* **9,** 22 (2011). https://doi.org/10.1186/1478-811X-9-22; https://creativecommons.org/licenses/by/2.0)

Figure 28.4.6 a model for the interaction of RACK1 and PKC-βII.

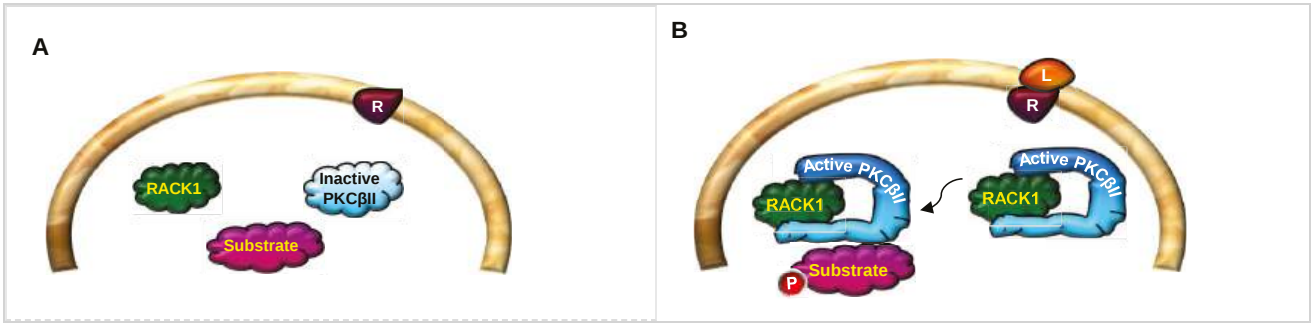


Figure 28.4.6: Model for PKC-βII and RACK1 interaction. Adams et al. Cell Communication and Signaling 2011, 9:22 http://www.biosignaling.com/content/9/1/22. Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0)

Panel (A) shows the resting state with no interaction between RACK1 and inactive PKCbII. Panel (B) shows how activation of PKC-βII leads to its interaction with RACK1. Substrate binding and phosphorylation follow. R is a receptor and L is its ligand.

Figure 28.4.7 shows how RACK1 can translate into the nucleus after ligand (L) gated activation of GPCRs (R) through adenylyl cyclase production of cAM and activation of Protein Kinase A.

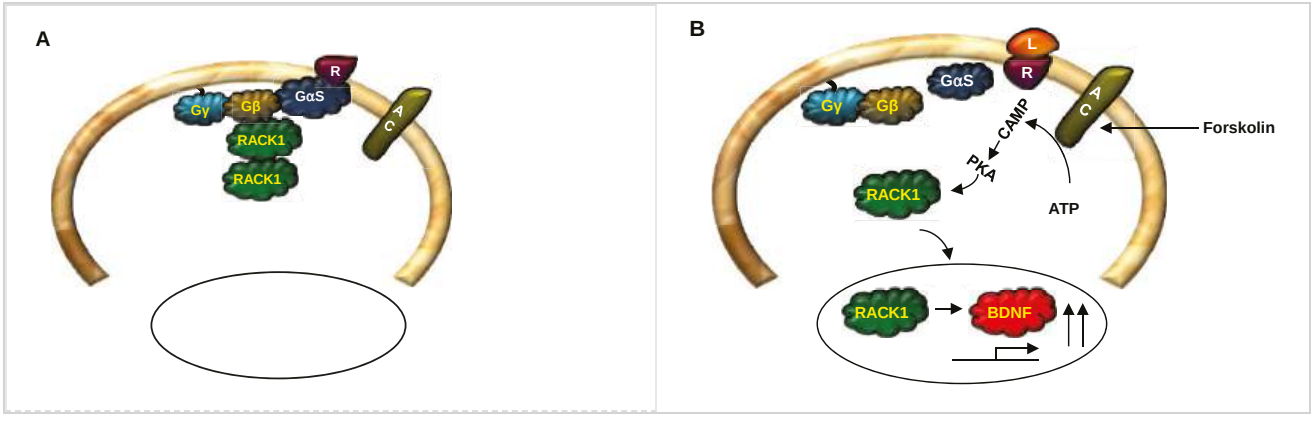


Figure 28.4.7 Model for cAMP/PKA-mediated nuclear translocation of RACK1. Adams et al. Cell Communication and Signaling 2011, 9:22 http://www.biosignaling.com/content/9/1/22. Creative Commons Attribution License. (http://creativecommons.org/licenses/by/2.0)

Panel (A) shows the resting state of RACK, which forms homodimers and heterodimers with the homologous Gβ subunit of the Gαβγ complex. Panel (B) shows how activation of PKA leads to dissociation of RACK which then can translate into the nucleus, where it leads to increased transcription of brain-derived neurotrophic factor (BDNF).

Figure 28.4.8 shows an interactive iCn3D model of a the human Rack1 (4AOW) color coded as in Figure 28.4.5

Figure 28.4.7: human Rack1 (4AOW) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...7WDaKw2dxGK87A

28.4.2: Downstream signaling from activated receptor tyrosine kinases

To review once again, when receptor tyrosine kinases get activated by binding a primary messenger such as a growth factor, the receptors dimerize, activating their cytoplasmic kinase domains. The activated RTK then autophosphorylates itself. The phosphorylated intracellular domain provides a docking site for other cellular "adaptor" proteins that contain an SH2 domain. The bound adaptor protein binds other signaling molecules within the cell leading to downstream propagation of the signal. Figure 28.4.8 shows some RTKs and downstream signaling events. We have already discussed the activation of the lipid kinase phosphoinositide 3-kinase (P13K) which leads to the activation of Akt (PKB). In the rest of this section, we will focus on the next step after the activation of RTKs. We focus on the epidermal growth factor receptor (EGFR, ErbB1) again.

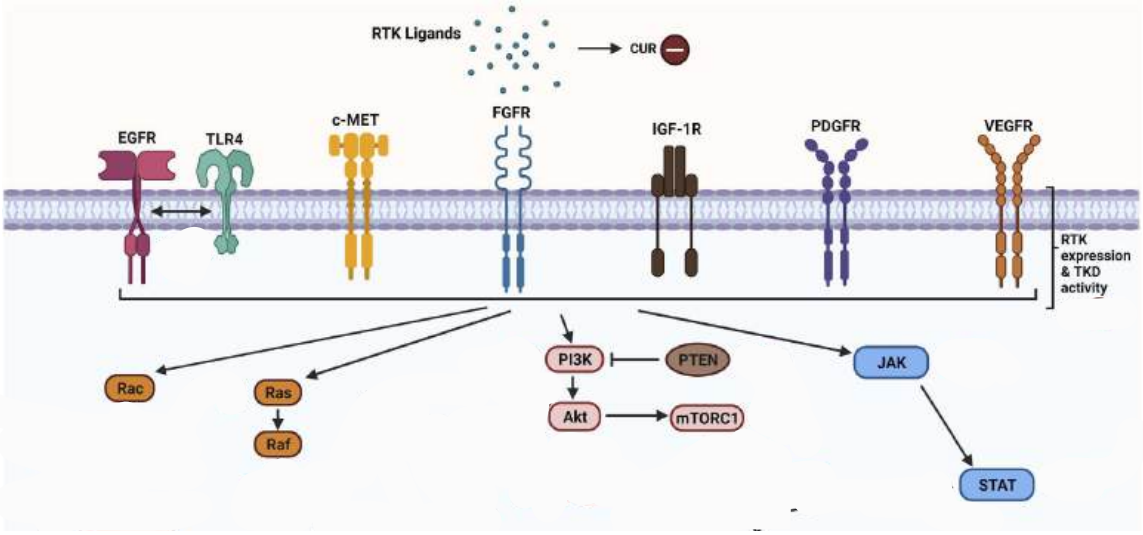


Figure 28.4.8: Some RTKs and downstream signaling events Dev et al. Frontiers in Pharmacology, 12 (2021). https://www.frontiersin.org/article/...ar.2021.772510. DOI=10.3389/fphar.2021.772510. **Creative Commons Attribution License (CC BY).**

1. Downstream signaling from the epidermal growth factor receptor.

Figure 28.4.9 shows events associated with the activation of EGFR.

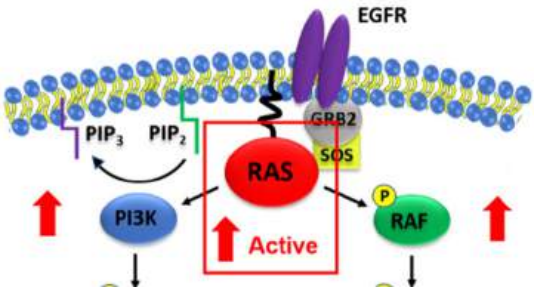


Figure 28.4.9: Near downstream signaling molecules after activation of EGFR

Once bound and activated by binding of growth factors protein signals, the intracellular domains of the now dimeric EGFR autophosphorylates themselves on selected tyrosine side chains. This then recruits a protein called **Growth factor receptor-bound protein 2 (GRB2)**, which has an **SH2 (Sarc Homology 2)** domain that binds phosphotyrosine motifs in proteins. GRB2 acts as an adaptor protein in that in addition to the SH2 domain, it has two **SH3 (Sarc Homology 3)** domains that bind **proline-rich domains** on other signaling proteins, including the protein **Son of sevenless homolog (SOS)**. GRB2 does not have enzymatic activity.

The adaptor protein GRB2 binds through its SH3 domain to the protein SOS, which then interacts with the protein Ras. This last member in the recruited complex is named because it causes **Ra**t **S**arcomas. There are many variants of these proteins but three are key in humans, HRas, KRas, and NRas. Ras is perhaps the key member of the family of intracellular small G proteins that bind GDP/GTP and are activated on the exchange of GTP for GDP. The proteins also have intrinsic GTPase activity as is found in the G_α protein of the heterotrimeric Gαβγ protein.

So what does SOS do? The SOS bound to RAS facilitates the exchange of GTP for GDP on Ras, maintaining it in an active state. SOS is a member of another fascinating class of small proteins that catalyze the exchange of GTP for GDP. The family of GTP/GDP exchange proteins is called **Guanine nucleotide Exchange Factors (GEFs)**. We will explore this in the next section.

The EGFR -GRB2-SOS-Ras complex in the figure above looks somewhat similar to the structure of a GPCR-heterotrimeric G protein $G_\alpha G_\beta G_\gamma$ complex, where G_α is also a GTP/GDP exchange protein with intrinsic GTPase activity. When the ligand binds to the GPCR, a conformational change ensues which facilitates the exchange of GTP for GDP on the G_α protein, activating intracellular signaling.

Once RAS is activated (bound to GTP), it binds and activates key kinases in the cell, including the lipid kinase PI3K and **RAF**, a kinase shown in Figure 28.4.8 Control of RAS activity is critical in signaling. It is one of the most commonly mutated proteins in cancer cells. Mutations that inhibit the intrinsic GTPase activity keep the protein active, leading to unregulated cell growth, proliferation, and differentiation, hallmarks of cancer cells.

The domain structures of GRB2 and SOS are shown in Figure 28.4.10 The proline-rich domain (motif) is not shown in the figure.

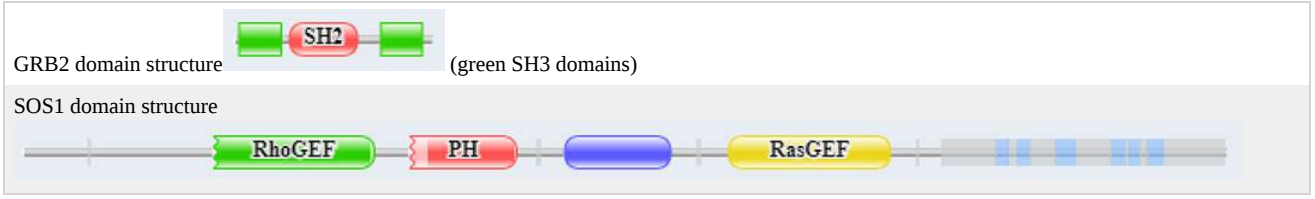


Figure 28.4.10 Domain structure of GRB2 and SOS1 (proline rich domain not shown in SOS)

Note

As we discussed in Chapter xx (domains), new genes encoding proteins with new functionalities can be created by duplicating and adding gene segments for different domains in a preexisting gene. As we have seen with the SH2 and SH3 domains, signaling proteins often share common domains. Table 28.4.1 below, adapted from the excellent book Cell Signaling by Lim, Mayer, and Pawson, shows some common domains found in signaling proteins.

Domains in Signaling Molecules

Domain	Binding Target	Cellular Process	Example protein	Pdb file (examples) Find your own in the PDB

Bromo	Acetyl-Lys	Chromatin reg.	BRD4	2YYN
C1	diacylglycerol	Plasma membrane recruitment	Raf-1	3OMV
C2	Phospholipid (Ca-dependent)	Membrane targeting, vesicle trafficking	PRKCA	3IW4
CARD	Homotypic Interactions	apoptosis	CRADD	3CRD
Chromo	Methyl-Lys	Chromo reg, gene transcription	CBX1	3F2U
Death (DD)	Homotypic inter.	Apoptosis	Fas	3EZQ
DED	Homotypic inter.	Apoptosis	Caspase 8	1F9E
DEP	Memb, GPCRs	Sig trans, protein trafficking	Dsh human disheveled 2	2REY
GRIP	Arf/Art G prot	Golgi traffic	Golgin-97 (Golga5)	1R4A
PDZ	C-term peptide motifs	Diverse, scaffolding	PSD-95 Or discs large homolog 4	1L6O
PH	Phospholipids	Membrane recruirtment	Akt	1O6L 3CQW
PTB	Phospho-Y	Y kinase signaling	Shc 1 SHC-transforming protein 1	1UEF 1irs
RGS	GTP binding pocket of Galpha	Sig trans	RGS4	1EZT
SH2	Phospho-Y	pY-signaling	Src	4U5W
SH3	Pro-rich sequence	Diverse, cytoskeleton	Src	2PTK
TIR	Homo/Heterotypic	Cytokine and immune	TLR4	3VQ2
TRAF	TNF signaling	Cell survival	TRAF-1	3ZJB
VHL	hydroxyPro	ubiquitinylation	VHL	1VCB

Figure 28.4.11 shows an interactive iCn3D model of the GRB2-SH2 domain in complex with a high affinity phosphopeptide KPFpYVNVEF (1BMB)

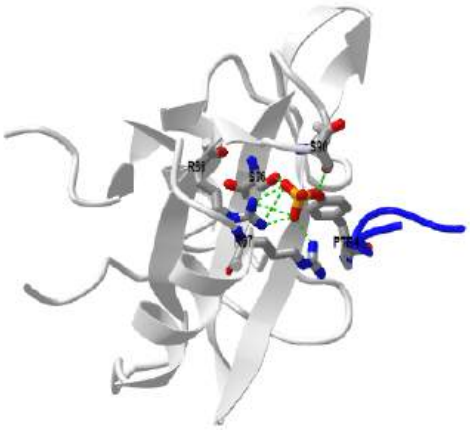


Figure 28.4.11: GRB2-SH2 domain in complex with a high-affinity phosphopeptide KPFpYVNVEF (1BMB) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...gBEpnsN6D8yrR6

Grb2 exists in both a monomeric state which can bind SOS, and a dimeric state which can't. You would expect the equilibrium between the monomer and dimeric form would be highly regulated. When a phospho-tyrosine ligand is bound to Grb2 through its SH2 domain, the dimer dissociated. This also occurs on phosphorylation of tyrosine 160 (Y160) on Grb2, a post-translational modification found in human prostate, colon, and breast cancers.

Figure 28.4.12shows an interactive iCn3D model of the GRB2 N-terminal SH3 domain complexed with a ten-residue proline-rich peptide (1135 Ac-VPPPVPPRRR-NH_2) derived from SOS (1GBQ)

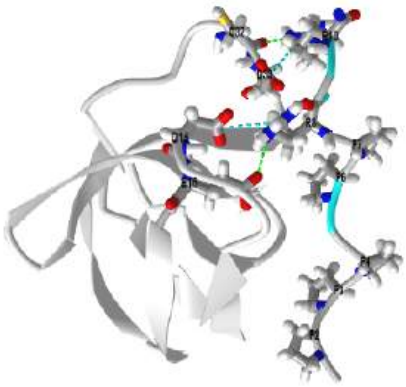


Figure 28.4.12: GRB2 N-terminal SH3 domain complexed with a ten-residue proline-rich peptide (derived from SOS (1GBQ). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...k2jNGaHexaUmT8

We will explore interactive structures of the small G protein Ras with the guanine nucleotide exchange factor SOS in the next chapter section.

After activation of Ras through GTP exchange for GDP, the GTP-Ras protein binds to and activates the kinase Raf. We will continue our exploration of that later in this section.

2. Downstream signaling from the insulin receptor.

When insulin binds to the receptor tyrosine kinase (RTK), it phosphorylates itself, which then leads to the binding of other proteins to the activated receptor and their phosphorylation. These lead to more intracellular signaling and ultimately changes in gene transcription. We'll focus on a specific adaptor protein, the **Insulin Receptor Substrate 1, IRS1,** a "scaffolding protein", which leads to the movement of the glucose transport protein **GLUT4** to the cell surface, allowing glucose uptake. These activities are shown schematically in Figure 28.4.13

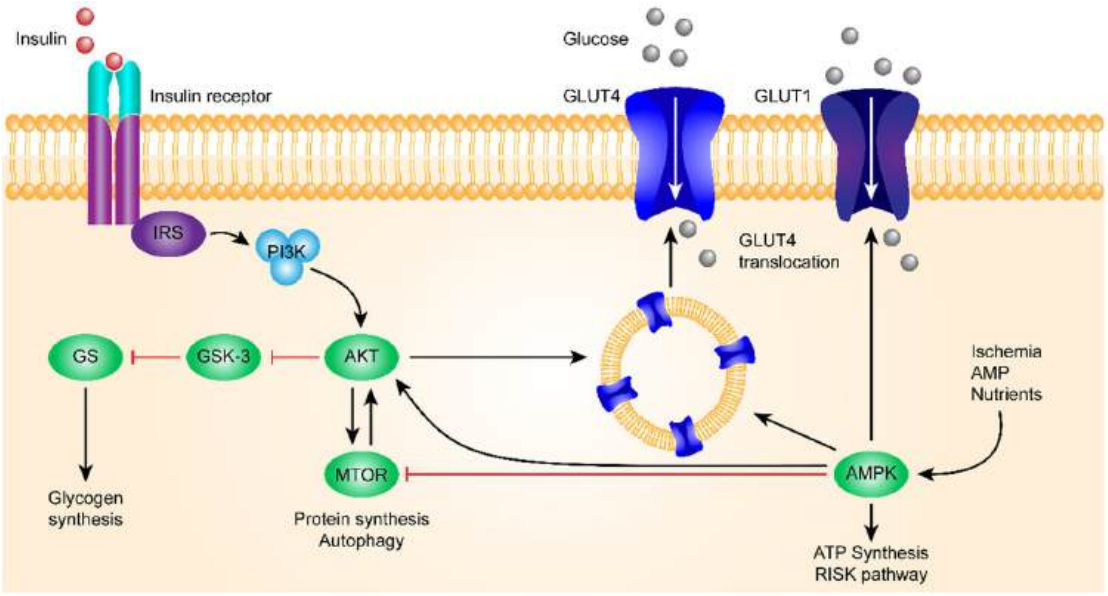


Figure 28.4.13: Binding to insulin stimulates the association of the receptor with downstream mediators, including insulin receptor substrate-1 (IRS-1) and PI3K. The insulin receptor can activate PI3K either directly, by binding to the p85 regulatory subunit, which produces PIP3, or indirectly, which leads to phosphorylation and the activation of AKT. Afterward, AKT phosphorylates the Ser9 site of GSK-3β and inhibits its activity. The PI3K/AKT/GSK-3β signaling pathway is involved in insulin signaling transduction, and GSK- 3β is regulated and controlled by insulin in this signaling pathway. Arneth et al. Int. J. Mol. Sci. 2019, 20, 2467; doi:10.3390/ijms20102467. Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

Human IRS1 has two domains, a PH domain for binding to membranes through phosphorylated IP2 derivatives and an **IRS/PTB** domain which binds phosphotyrosines on proteins. The PTB and the SH2 domains are the most common domains for binding phosphotyrosines on proteins. PTB binds the NPXY, where X is a pTyr.

We show a more detailed view in Figure 28.4.14 in part, to review many of the signaling proteins we have seen before, including PI3K and PDK1.

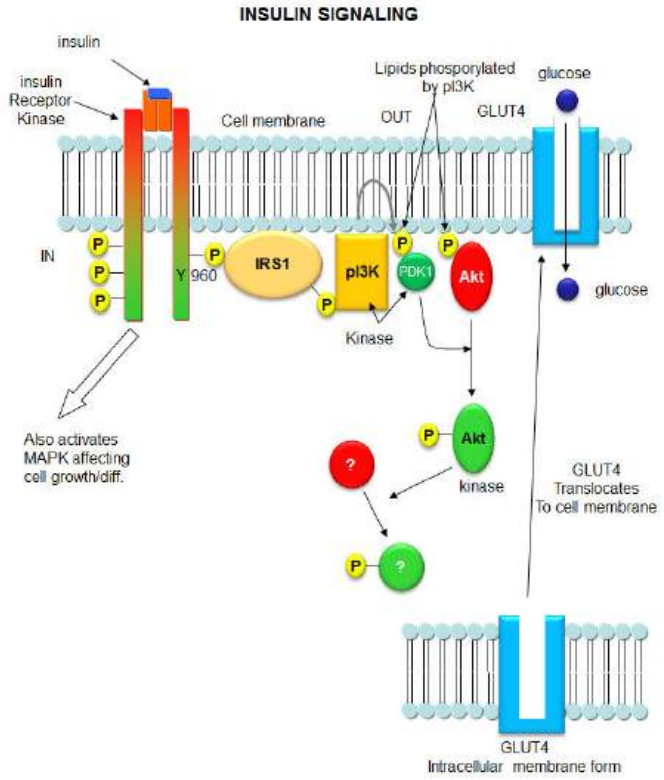


Figure 28.4.14: A more detailed view of the insulin receptor-mediated translocation of GLUT4 to the membrane.

Likewise, another review of PIP2 derivatives is warranted. After phosphorylation by the activated insulin receptor protein tyrosine kinase, IRS-1 binds phosphatidylinositol 3-kinase (PI3K) that causes phosphorylation of the 3'OH on phosphatidyl inositol (PI) in the inner leaflet of the membrane to form PI(3)P. PI3K is a member of a family of kinases that phosphorylates PIP2. The metabolic pathway centered on pI3K is one of the most mutated in human cancers. PI(3)P in turn recruits to the membrane other inactive kinases, phosphoinositide-dependent kinase 1, PDK1, and Akt, also known as PKB. Figure 28.4.15 shows phosphorylated phosphatidylinositol derivatives.

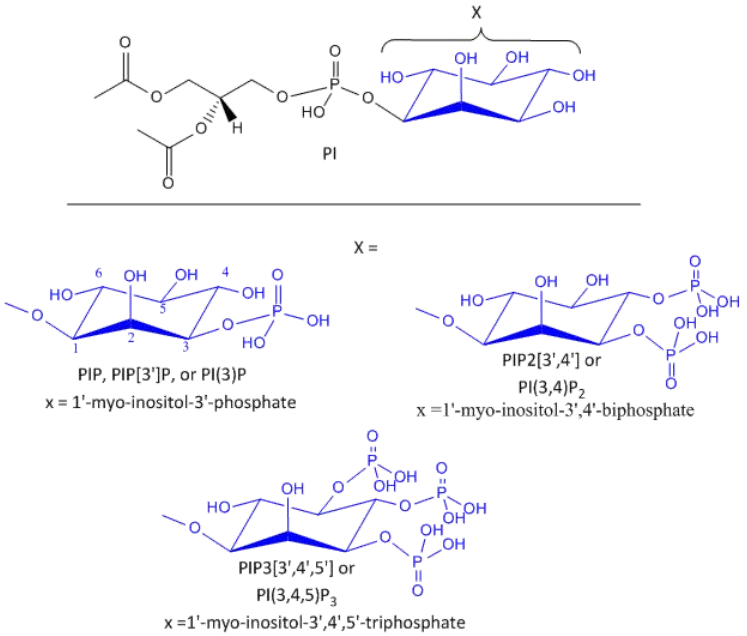


Figure 28.4.15: Phosphorylated phosphatidylinositol derivatives

On binding of PI(3)P, PDK1 becomes an active kinase, which phosphorylates and activates Akt. Akt kinases are major Ser/Thr protein kinase that phosphorylates proteins involved in a host of cell activities, including regulation of glucose transport, cell proliferation, and death. In the insulin signaling pathway, active (phosphorylated) Akt leads to the movement of the GLUT4 protein from intracellular endosomal vesicles to the cell surface, which offers a quicker way to import glucose into the cell than if Akt activated GLUT 4 gene expression. PDK1 phosphorylation of Akt2-T309 is required for insulin-stimulated Glut4 translocation. If T309 is mutated to A309 or if PDK1 is inhibited, GLUT 4 is not translocated to the cell membrane.

Figure 28.4.16 shows an interactive iCn3D model of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog (1IR3))

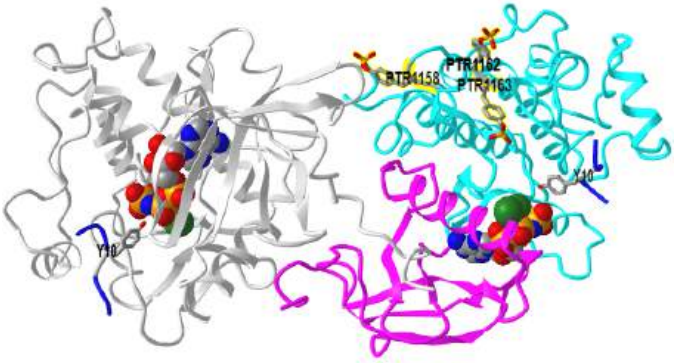


Figure 28.4.16: Activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog (1IR3). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...j11LhWKCUnreV9

The dimeric form of the kinase is shown. ANP is shown in spacefill. The peptide substrate containing the interior tyrosine (stick) for phosphorylation is shown in blue. Three key tyrosines in the activation loop that are autophosphorylated (Tyr1158, Tyr1162, and Tyr1163) are shown and labeled in the right-hand monomer.

As with a protein kinase, the activation loop of the insulin receptor kinase domain is phosphorylated and the resulting conformational change allows ATP and target protein access. The activation loop gets phosphorylated on Tyr1158, Tyr1162, and Tyr1163 with Y1163 being key.

28.4.3: Downstream signaling from Cytokine activated Receptors- The JAK/STAT pathway.

Now we explore two signaling pathways mediated by the adaptor protein JAK and STAT. These are abbreviations for the Janus Kinase (JAK) and the Signal Transducer and Activator of Transcription (STAT). These play key roles in embryonic development, stem cell maintenance, hematopoiesis (formation of blood cells), and immune cell signaling. . This pathway is stimulated by cytokines and interleukins, protein modulators released by immune cells, as well as growth factors.

In general, there are five groups of cytokines:

- tumor necrosis factor alpha (TNF-alpha)
- Interleukin-1 family members (IL-1_
- Transforming growth factor-beta (TGF-bet)
- those that signal through RTKs (such as macrophage colony-stimulating factor (M-CSF)
- Chemokines
- cytokines that signal through JAK/STAT

In contrast to RTKs, which have kinase domains activated on receptor dimerization, cytokine receptors that work through JAK/STAT do **NOT** have kinase domains. On cytokine binding to their target cytokine receptor, the now-activated receptors activate the prebound inactive Janus kinase through conformational changes. The kinase domains autophosphorylate themselves in a trans fashion. The active kinase then phosphorylates the cytoplasmic tails of the cytokine receptors. This trigger further signal transduction reactions mediated by the binding of target signaling proteins to the cytoplasmic region of the phosphorylated cytokine receptor. Just to reiterate, the cytokine receptor is NOT a RTK with latent kinase activity. Instead, it becomes phosphorylated by the bound and cytokine-activated JAK. A portion of the pathway is illustrated in Figure 28.4.17.

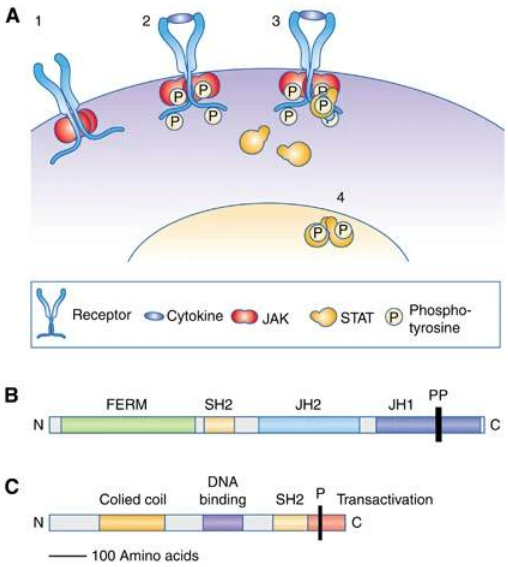


Figure 28.4.17: An overview of the JAK/STAT signaling pathway (A) and the domain structures of JAK (B) and STAT (C). Thomas *et al.* T*Br J Cancer* **113,** 365–371 (2015). https://doi.org/10.1038/bjc.2015.233. Creative Commons Attribution-Non-Commercial-Share Alike 4.0 International License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/4.0/

Panel (A1) shows that the kinase JAK is bound constitutively in an inactive state to the cytokine receptor cytoplasmic region, not through its SH2 domain, but through its FERM domain (Panel B). The figure implies the cytokine receptor is dimeric in the absence of a ligand. On cytokine binding, conformational changes and repositioning of the bound JAK activates its kinase domain, which phosphorylated the C-terminal tails of the cytokine receptor. STAT monomers, through their SH2 domains, bind to the phosphorylated cytokine receptor where they get phosphorylated by the active JAK. The phospho-STAT monomers form a dimer, dissociate from the complex, and translocate to the nucleus where they act as transcription factors. the Janus kinase is named after Janus, the two-faced Roman god of beginnings, endings, and duality since JAK has two nearly identical JH kinase domains. One has kinase activity while the other inhibits the first.

Panels (**B**) shows the domain structure of JAK. The FERM domain facilitates JAK:cytokine receptor binding. The JH2 pseudokinase domain regulates the kinase activity of the JH1 kinase domain. Ps show site on JAK necessary for activation

Panel (**C**) shows the domain structure of STAT. The SH2 domain binds phosphorylated tyrosines. The carboxy terminus transactivation domain is required for full transcriptional activation. P marks the conserved tyrosine residue whose phosphorylation is essential for STAT activation.

The pseudokinase domain of JAK interacts with the kinase domain on the same chain and prevents its activity in the inactive monomer and dimer. Cytokine binding to the cytokine receptor induces a conformational change that promotes the interactions of the pseudokinase domain on one JAK monomer with the same domain on another, promoting dimerization and freeing the kinase domains for activity. Figure 28.4.18shows the pseudokinase domain (orange) interactions in the active JAK dimer.

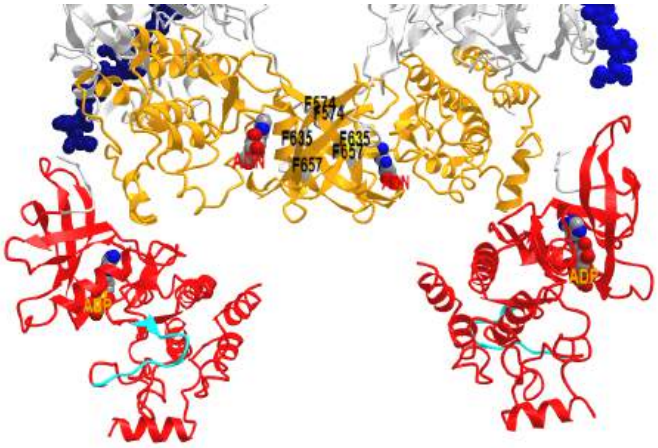


Figure 28.4.18 The pseudokinase domain (orange) interactions in the active JAK dimer.

The kinase domain of each dimer is shown in red and the pseudokinase domain is shown in orange. The activation loop (cyan) and ADP (spacefill, CPK colors) are shown in the red kinase domains. The orange pseudokinase domain has a bound adenosine (not ADP) that is shown spacefill and CPK color. However, it lacks the
DFG motif required for catalysis. The phenylalanine cluster (F635, 657, 574) are labeled. The structure is actually for a mutant that has the V657F mutation that promotes dimerization and JAK activity. Hence the V657F mutation is oncogenic.

JAK activity can be inhibited by the protein suppressor of cytokine signaling **(**SOCS). Transcription of the protein is activated by STAT, and the SOCS protein in a feedback inhibition loop binds to JAKs and also to IFNGR1, which inhibit JAK activity.

Figure 28.4.19 shows an interactive iCn3D model of the structure of the inhibitory protein SOCS1 in complex with JAK1 kinase domain (6C7Y)

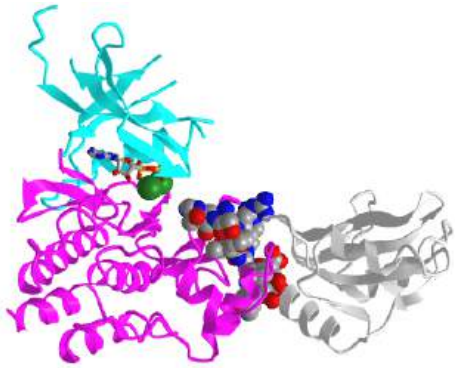


Figure 28.4.19: Structure of the inhibitory protein SOCS1 in complex with JAK1 kinase domain (6C7Y). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...tqtfo2SQ5TxF58

The N-lobe of the JAK1 kinase domain is shown in cyan and the C-lobe in magenta. ADP (sticks) and Mg^{2+} (green) are shown in the interface between the lobes. SOCS1 is shown in gray except for the kinase inhibitor region which is shown as spheres and CPK colors. It binds in the substrate binding grove and prevents substrate access.

Activated JAK activity and signaling do not stop simply with the activation of STAT. In addition to stimulating signaling through phosphorylated dimers of STAT, cytokines also activate other signaling pathways through the same receptors. Examples include the PI3K pathway described in an earlier section and also the MAPK pathways, described in detail later in this section. Both the PI3K and MAPK pathways are activated by binding the cytokine IL6 to its cytokine receptor. The mechanism for PI3K activation is a bit unclear. The MAPK pathway is activated by a phosphatase called SHP2 for Src homology region 2 domain-containing phosphatase 1. This protein binds to pTyr759 on the IL-6 receptor. As binding of the cytokine activates the prebound JAK, it also activates SHP2, ultimately activating signaling through Ras (a small G protein), which activates RAF, a kinase. Figure 28.4.20 shows these three combined signaling pathways for a cytokine receptor: STAT, Ras, and PI3K activations.

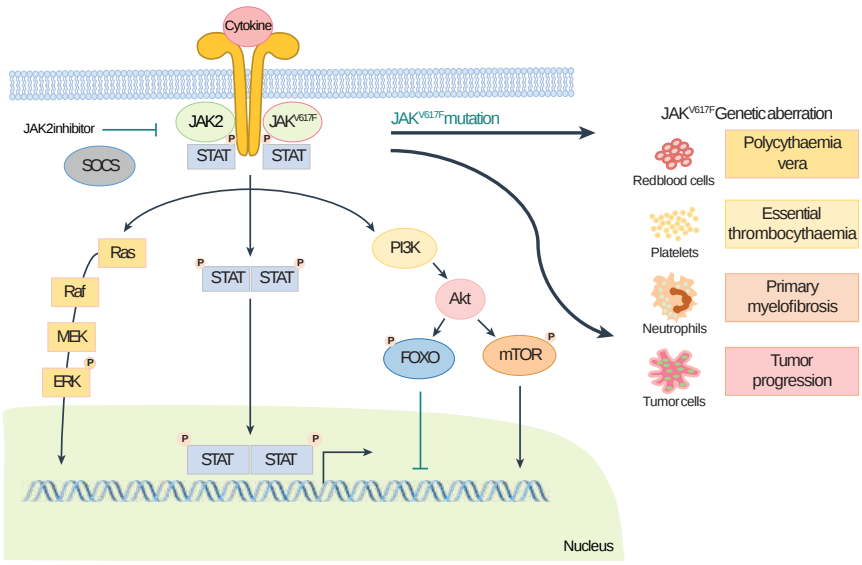


Figure 28.4.20: Schematic representation of the JAK-STAT pathway. Lee et al. Immune Netw. 2017 Aug;17(4):214-227. https://doi.org/10.4110/in.2017.17.4.214. pISSN 1598-2629·eISSN 2092-6685. Creative Commons Attribution Non-Commercial License (https://creativecommons.org/licenses/by-nc/4.0/)

The cytokine receptor induces activation of JAKs after cytokine stimulation following the phosphorylation of STATs. Furthermore, phosphorylated STATs undergo dimerization and translocate to the nucleus to activate target gene transcription. SOCS, suppressors of cytokine signaling; PI3K, phosphatidyl inositol 3 kinase; Akt, protein kinase B; FOXO, Forkhead box protein O; mTOR, mammalian Target Of Rapamycin.

Figure 28.4.21 shows an interactive iCn3D model of the Crystal structure of a tyrosine phosphorylated STAT-1 dimer bound to DNA (1BF5)

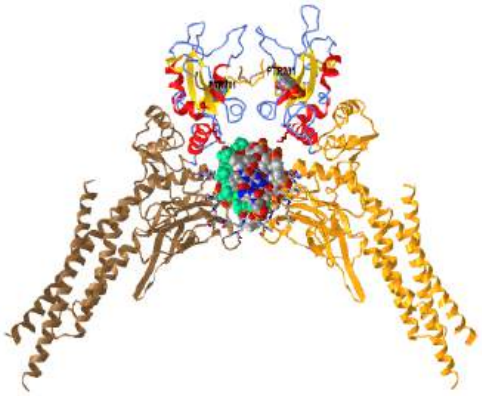


Figure 28.4.21: Crystal structure of a tyrosine phosphorylated STAT-1 dimer bound to DNA (1BF5). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...VEjHZ5t3baKC37

The dimer chains are shown in brown and gold except for those colored-coded by secondary structure (helix red, sheet yellow). The backbone of the nucleotides is shown in spacefill cyan and gray. Zoom in to see noncovalent interactions between the bound DNA and protein monomers. Two phospho-tyrosines labeled pTR701 are shown as well. The DNA binding domain of the STAT dimer has an immunoglobulin fold and forms a "C-shaped clamp" around the DNA. The domains colored by secondary structure are SH2 domains, with each recognizing and binding to the phosphorylated Tyr 701 (labeled pTR701), a very interesting use of the SH2 domain.

Figure 28.4.22 shows an interactive iCn3D model of the active Janus Kinase (JAK) dimer complexed with the intracellular domains (spacefill) of the interferon lambda(a cytokine) receptor (7T6F).

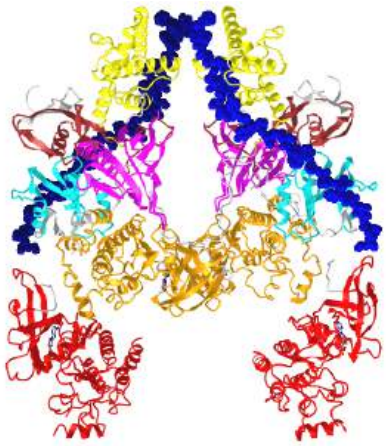


Figure 28.4.22 Active Janus Kinase (JAK) dimer complexed with the intracellular domain of the interferon lambda(a cytokine) receptor (7T6F). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...9s3VBNMxsk7WM7

Domain/protein Coloring

- Red: PTKc-JAK-rpt2 kinase catalytic domain
- Orange: PTK-JAK1-rpt1 pseudokinase domain
- Yellow: FERM F2
- Magenta: FERM C-JAK1
- Cyan: SH2
- Brown: FERM F1
- Navy Blue and Blue spacefill: cytoplasmic domains of the cytokine receptor interferon lambda receptor 1 dimer
- The gold large sphere represents interferon bound to the extracellular domain of the interferon lambda receptor 1 dimer

Figure 28.4.23shows a model combining the active Janus kinase (JAK) containing the intracellular domain of the interferon alpha receptor 1 dimer (navy and blue spacefill) with AlphaFold models of the extracellular and transmembrane domain of the interferon-alpha receptor ((Q8IU57))

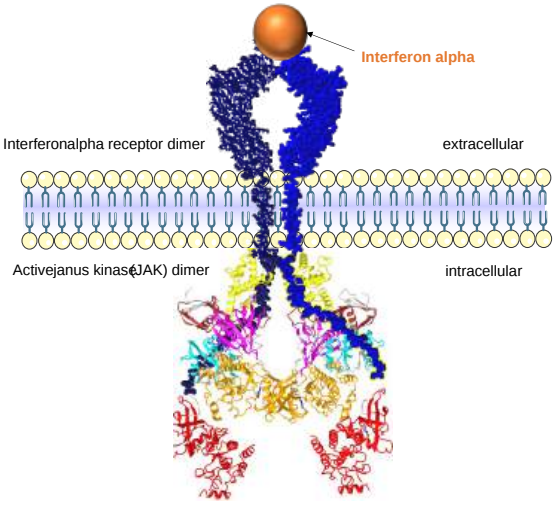


Figure 28.4.23 AlphaFold model of Interferon lambda receptor 1 Extracellular Domain and transmembrane domain (Q8IU57) with Active Janus Kinase (JAK) dimer complexed with cytokine receptor intracellular domain (7T6F)

Again, to reiterate as we did above, the cytokine receptor (in the figure above the interferon lambda receptor) is NOT a RTK with latent kinase activity. Instead, it becomes phosphorylated by the bound and cytokine-activated JAK which is resident in the cytoplasm.

Here is a link to a second iCn3D model which shows the phenylalanine cluster that promotes pseudokinase domain interactions in the active JAK dimer.

28.4.4: Downstream Intracellular signaling through Src Family Kinases - Membrane-associated non-receptor tyrosine protein kinases

Another family of intracellular protein kinases - the Src family - are often activated on upstream activation of many different types of receptors including GPCRs, RTKs, cytokine receptors, as well as integrins and adhesion receptors that we explore in a later chapter section. We have already touched on them when we discussed proteins containing the src homology domains SH2 and SH3. Src, an intracellular Tyr kinase (MW 60,000), is the founding member of this family of protein kinases. Src is activated when it binds through its own SH2 domain to a phosphorylated membrane receptor.

Src has many names including Proto-oncogene tyrosine-protein kinase Sr**c,** proto-oncogene c-Src, pp60c-src, and p60-Sr. These membrane-associated, non-receptor tyrosine kinases regulate cell proliferation, differentiation, apoptosis, migration, metabolism, and cytoskeleton organization. They are one of the major classes of intracellular kinases which are activated after upstream activation of the membrane receptors mentioned above (GPCRs, RTKs, cytokine receptors, as well as integrins and adhesion receptors). They in turn activate further downstream protein kinases by phosphorylation. They even phosphorylate the upstream membranes which led to their activation.

There are 10 members in the Src family: Src, Frk, Lck, Lyn, Blk, Hck, Fyn, Yrk, Fgr, and Yes. They all share the same expanded domain structure shown in Figure 28.4.24

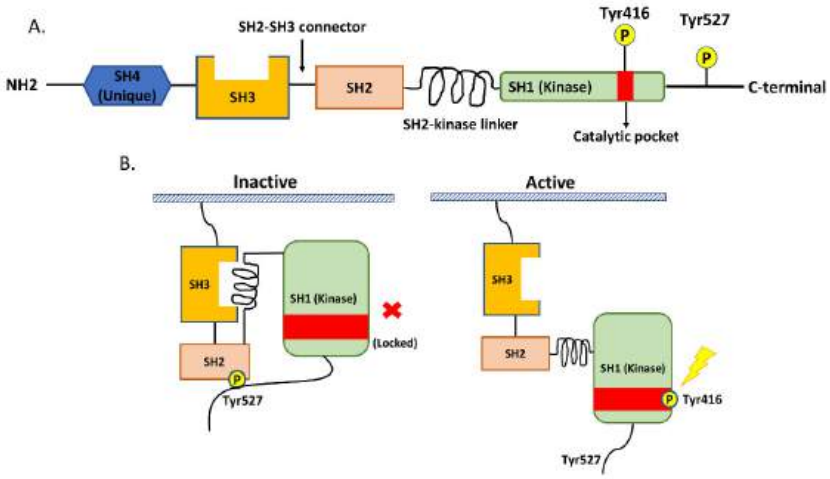


Figure 28.4.24: The domain structure of Src (panel A) and its activation (panel B). Hsu et al, Cancers 12(6):1361. DOI: 10.3390/cancers12061361. CC BY 4.0

They have an SH4 N-terminal region that can be post-translationally modified with fatty acids (myristoylation and palmitoylation), which can anchor it to the membrane. The spacer SH2-SH3 spacer is intrinsically disordered and differs in sequence among members of the Src family. Two key phosphorylation sites (Tyr416 and Tyr527) are important. In inactive Src, Try 527 is phosphorylated, allowing it to bind to the SH2 domain. The SH2-kinase linker also binds to the SH3 domain. This occludes the active site region and prevents the phosphorylation of Try 416 in the "activation loop" of the SH1 kinase domain. When Try 527 is dephosphorylated, a conformation change ensues the opens the binding site allowing autophosphorylation of Tyr 416 and its activation. Hsu et al, Cancers 12(6):1361. DOI: 10.3390/cancers12061361. CC BY 4.0

Figure 28.4.25shows an interactive iCn3D model of the Human tyrosine-protein kinase C-Src in complex with AMP-PNP (2SRC)

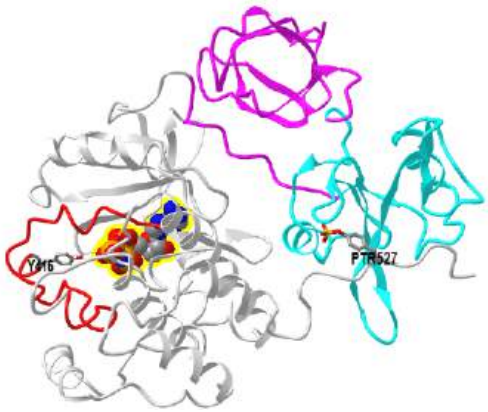


Figure 28.4.25: Human tyrosine-protein kinase C-Src in complex with AMP-PNP (2SRC). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...L45honKx6rrEu7

pY527 is shown in stick and labeled. It binds to the SH2 domain (151-248) shown in cyan. The SH3 domain is in magenta and the kinase domain is shown in gray. The activation loop in the kinase domain is shown in red with the pY416 needed for activation shown in stick and labeled. ANP is shown in spacefill.

We have now seen the structure of many kinase domains. Figure 28.4.26 shows an interactive iCn3D model of the alignment of human c-Src (452 amino acids) (2SRC) and Erk2 (362 amino acids) also called MAPK1 (2Y9Q), a protein kinase will be explored at the end of this section

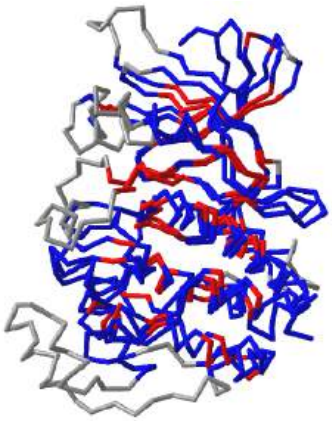


Figure 28.4.26: Aligned structures of human c-Src (2SRC) and Erk2 (2Y9Q). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...ETS1puViPtrQu7

Red shown conserved sequences, blue is aligned (but not conserved), and gray is nonaligned. Use the "a" key to toggle between the states.

After much discussion of the binding of p-Tyrosine side chains to the SH2 domain, we now present Figure 28.4.27, which shows an interactive iCn3D model of a phosphotyrosine peptide bound to the SH2 domain of Fer tyrosine kinase (6KC4)

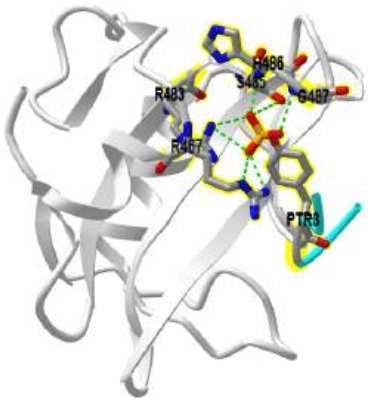


Figure 28.4.27: Phosphotyrosine peptide bound to the SH2 domain of Fer tyrosine kinase (6KC4). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...3eqrcT7NwcHpr8

The SH2 domain is shown in gray. The phosphopeptide (DEpYENVD) is shown in cyan with the labeled pY in sticks. The side chains on the SH2 domains interacting with the pY are shown in stick and labeled.

We are about to explore the last but incredibly important downstream kinases activated in signal transduction cascades, the **mitogen activate protein kinase (MAPK**) cascade. It could also be called the **Erk cascade**. Before we do that, we present parts of three pathways mediated by activated RTKs to refresh your minds and also because the more you see key players in the pathway, the more you start to remember the names, structural features, and function of the signaling molecules.

Figure 28.4.28 offers a quick and abbreviated look at signaling through activated RTKs that proceed through the adaptor protein and the GEF SOS, leading to the activation of Ras, a key small G protein. Ras in turn activates a protein (MAPKKK) in the MAPK cascade.

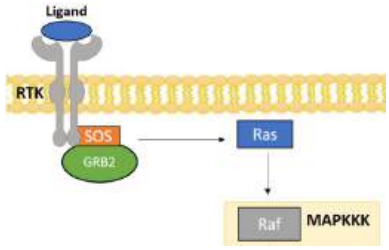


Figure 28.4.28 Simplified cartoon showing the activation of the MAPK cascade protein MAPKKK. Cordover E, Minden A. Signaling pathways downstream to receptor tyrosine kinases: targets for cancer treatment. *J Cancer Metastasis Treat* 2020;6:45. http://dx.doi.org/10.20517/2394-4722.2020.101. Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/),

Figure 28.4.29 shows an abbreviated version of the activation of the lipid kinase PI3K and through the activation of PDK1 and AKT, the activation of two major kinase complexes, mTORC1, and mTORC2, which we will explore in its a separate chapter section.

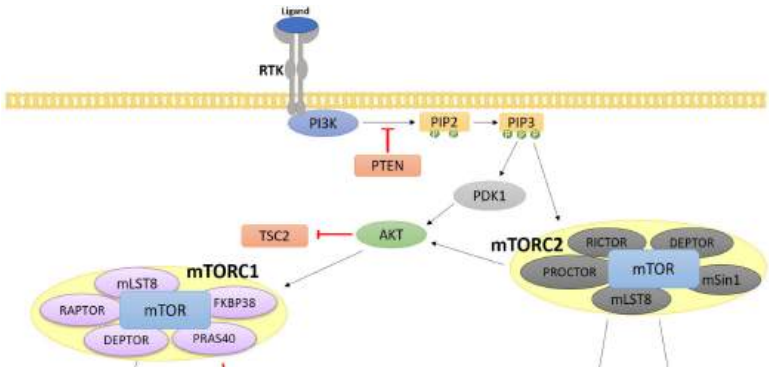


Figure 28.4.29: Activation of mTORC1 and mTORC2, through activation of PI3K, PDK1 and AKT. Cordover E, Minden A. Signaling pathways downstream to receptor tyrosine kinases: targets for cancer treatment. *J Cancer Metastasis Treat* 2020;6:45. http://dx.doi.org/10.20517/2394-4722.2020.101. Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/)

Finally, Figure 28.4.30show the combined activation of both the MAPK (ERK) cascade pathway and the mTOR complex through GPCR signaling using the adaptor protein Grb2.

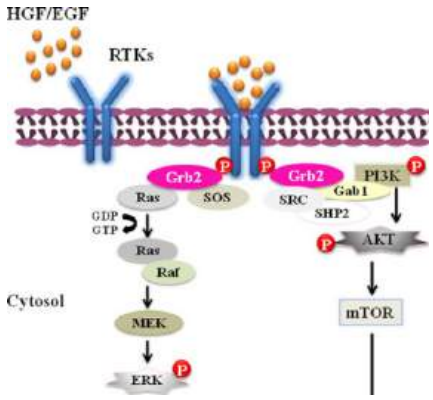


Figure 28.4.30: Combined activation of both the MAPK (ERK) cascade pathway and the mTOR complex through GPCR signaling using the adaptor protein Grb2Yang et al. Journal of Cellular and Molecular Medicine 22(6). DOI: 10.1111/jcmm.13447. License CC BY 4.0

28.4.5: Downstream signaling through the Mitogen-Activated Protein Kinase Cascade

Active upstream kinases like PKA, PKC, and RTKs phosphorylate target proteins and in doing so change their activities. The usual protein targets are kinases, which become active on phosphorylation. They in turn activate other kinases, resulting in a complex cascade and amplification of the original signal. Often the end product of such cascade is a phosphorylated transcription factor that can alter gene expression. Perhaps the most described of these cascades is the **Mitogen-Activated Protein Kinase (MAPK)** pathway. Mitogens are chemical species that lead to mitosis (cell division). The MAPK system has three layers and a confusing nomenclature (until you are used to it). The end (downstream) product of the cascade is the enzyme **mitogen-activated protein kinase**, abbreviated MAPK. It can be phosphorylated several times to produce MAPKP or MAPKPP where the last Ps in the abbreviation signifies the number of added phosphate groups.

The kinase immediately upstream that phosphorylates MAPK is abbreviated **MAPKK** (for mitogen-activated protein kinase kinase) or **MAP2K**. MAPKK (MAP2K) is activated by yet another upstream kinase called **MAPKKK** or **MAP3K**. If these are also targets of another upstream unrelated kinase, they would be abbreviated MAP3KPP for example.

Now, of course, there are families of the MAPK cascade enzyme, each with its own name. Figure 28.4.31shows the names of four different mammalian MAP3Ks leading to the activation of 5 different MAP2Ks which lead to the activation of 4 different MAPKs. Some of these enzymes are so widely discussed in textbooks and journal articles that it is good to remember them specifically with their alternative names. These include the MAP3K enzymes **Raf (**Rapidly Accelerated Fibrosarcoma) and **MEK**, and the MAPK enzymes **ERK** (Extracellular Related Kinase) and **JNK** (c-Jun N-terminal Kinase)

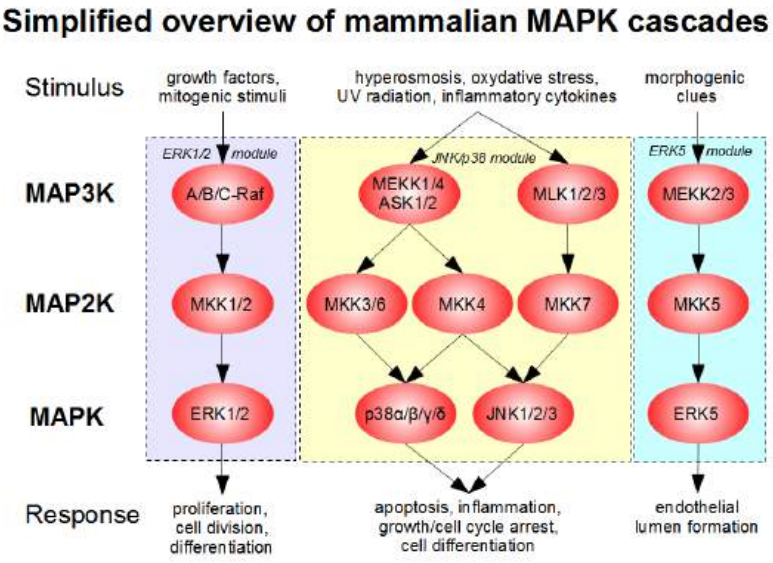


Figure 28.4.31: Simplified overview of the mammalian MAPK cascades https://commons.wikimedia.org/wiki/F...-mammalian.png. Creative Commons Attribution-Share Alike 3.0 Unported

Figure 28.4.32 shows another representation of the MAPK cascade with some different enzyme names and added phosphates shown in circles.

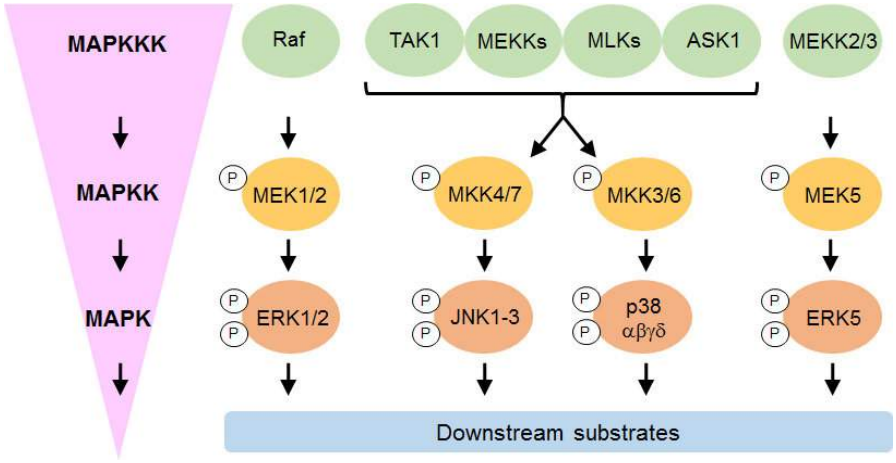


Figure 28.4.32 Another representation of the MAPK cascade. Journal of Cardiovascular Development and Disease 6(3):27. 2019. 10.3390/jcdd6030027. CC BY 4.0

One way to organize a seminar on a complex topic is to use these 3 steps: tell your audience what you will tell them, tell them, and then tell them what you told them. Following that advice, we present in Figure 28.4.33 an integrated view of signaling, starting from the membrane and moving inward to three enzymes in the MAPK kinase cascade, RAF (a MAP3K), MEK (a MAP2K) and ERK (a MAPK). Upstream signaling to the MAPK cascade comes RTKs, GPCRs, and Ca^{2+} signaling, which we will discuss later.

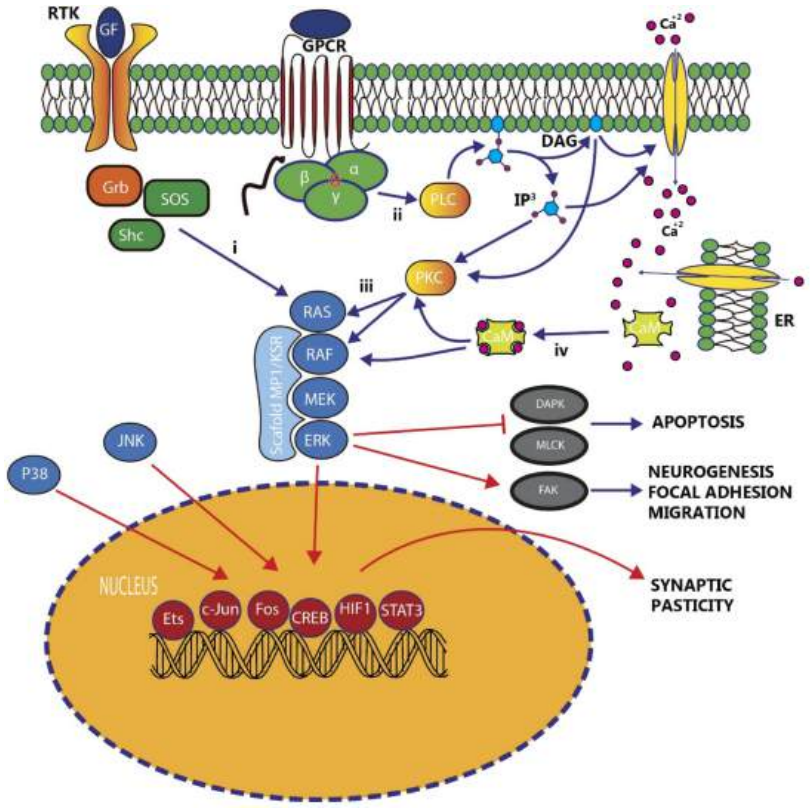


Figure 28.4.33: Upstream signaling to the MAPK cascade from RTKs, GPCRs, and Ca^{2+} signaling Albert-Gasco et al. International Journal of Molecular Sciences 21(12):4471 (2020). DOI: 10.3390/ijms21124471. License CC BY 4.0

The sequence of events is:

- binding of an external signal to membrane receptor and activation of receptor kinase
- phosphorylation of receptor kinase and interaction with an activator GTP binding protein such as Ras
- binding of activated G-protein to and activation of a mitogen-activated protein kinase kinase kinase (MAPKKK)
- MKKK phosphorylates and activates another kinase, MAPKK
- MKK phosphorylates and activates mitogen-activated protein kinase, MAPK
- MAPK phosphorylates inactive transcription factors (or other proteins) and activates them. Unfortunately (from a naming point of view) when the activated proteins are themselves protein kinases, they are called mitogen-activated protein kinase activated protein kinases (MAPKAPK)

There are seven types of MAPKs, four conventional and three atypical. Four typical ones are described in the table below.

Activator GTP binding protein	Ras:GTP			
MAPKKK or MAPK3	Raf-1A/B c-Mos	MEKK1-4 DLK MLK2	MEKK1-4 DLK MLK2	MEKK2/3 Tpl-2
MAPKK or MAPK2	MEK1,2	MEK4,7	MEK3,6	MEK5
MAPK or MAK	ERK1,2	JNK1-3	p38	ERK5
MAPKAPK	RSK 1-4 MNK2 MSK 1,2	MK2,3	MSK1,2 MK2,3	RSK1-4
An eventual Protein Target	c-Jun	c-Jun		

EXTERNAL MAP Kinase System from Cell Signaling

Structural, these proteins are similar in size and domain structure as we have seen for other kinases. Figure 28.4.34 shows an interactive iCn3D model of the alignment of MEK 1 (**4U7Z**) and ERK2 (**5NHJ**)

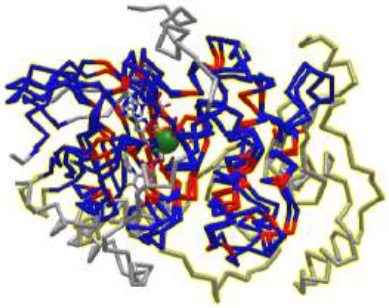
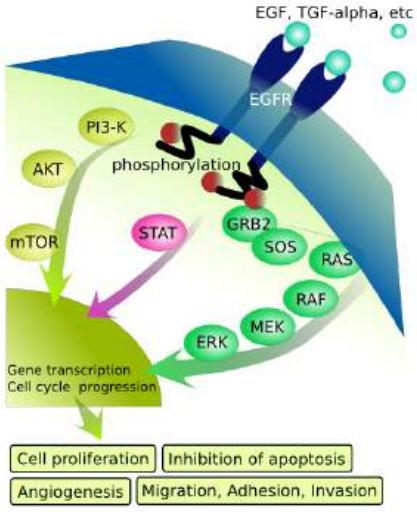


Figure 28.4.34: . (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...Vq6RR7BA6o1Qb7

We now present multiple visual images of the activation of the MAPK cascade. Figure 28.4.35shows two.

Details:



EGFR_signaling_pathway.png. **public domain**

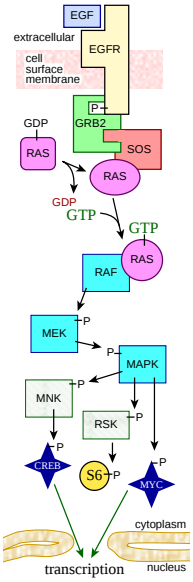


Figure 28.4.x: https://commons.wikimedia.org/wiki/F...ay_diagram.svg. Creative Commons Attribution-Share Alike 4.0 International

Figure 28.4.35 Two visual representations of the activation of the MAPK cascade through to activation of gene transcription.

As with protein kinase A and protein kinase C, signaling efficiency and specificity with minimal phosphorylation of wrong targets occurs when multiple signaling partners are scaffolded. This is also true of the MAPK cascade trio of the kinase. Figure 28.4.36 shows the role of scaffolds KSR and Ste5 in MAPK signaling.

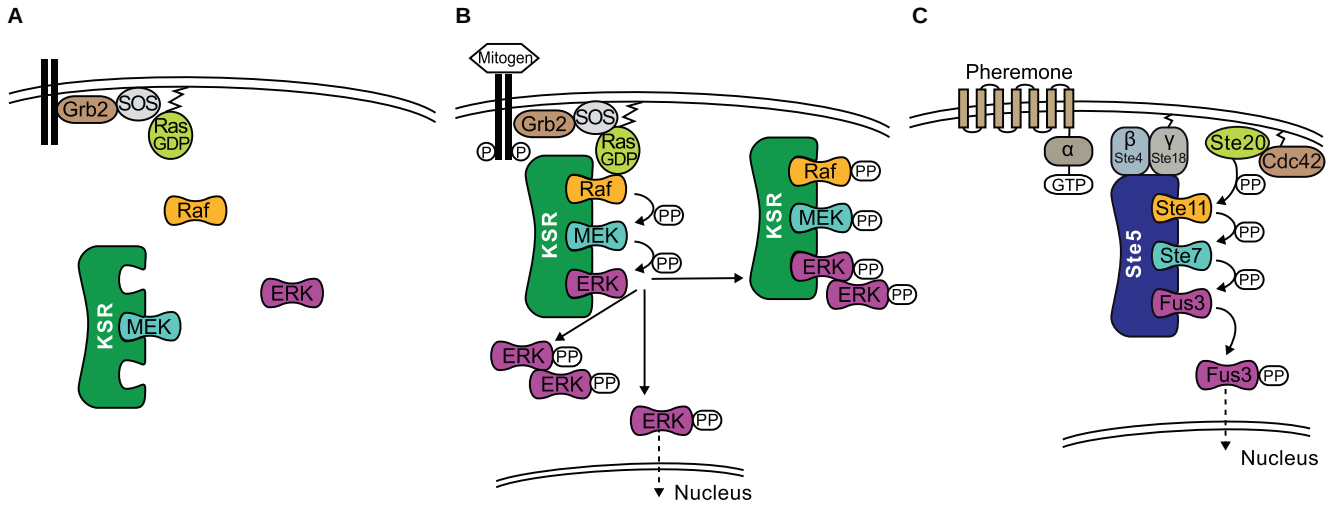


Figure 28.4.36: **The role of scaffolds KSR and Ste5 in MAPK signaling. (A)** In quiescent cells, an inactive KSR/MEK complex exists in the cytosol. **(B)** Upon stimulation of the cell, KSR translocates to the cell membrane and forms an active complex with phosphorylated Raf, MEK, and ERK. Activated ERK detaches from the scaffold with three outcomes; (1) ERK dimerizes in the cytoplasm where the dimer remains or translocates to the nucleus; (2) ERK translocates to the nucleus; (3) KSR acts as a platform where ERK dimers are assembled and the new complex can interact with substrates in the cytoplasm. **(C)** Schematic of the yeast mating pheromone response pathway. Witzel et al. Front. Physiol., 21 December 2012 | https://doi.org/10.3389/fphys.2012.00475. **Creative Commons Attribution License**,

28.4.6: Computational Analyses - MAPK Cascade

You might think that the interconnected reactions of the MAPK cascade are complicated. However, as presented in the figures above, it only consists of 3 enzymes, MAPKKK (MAPK3 or MK3), MAPKK (MAPK2 or MK2), and MAPK. We added complexity by converting the actual enzymes from an inactive state to an active state by phosphorylation. In reality, this pathway is simple compared to the complete signal transduction pathways they are part of, and very simple compared to whole catabolic and anabolic pathways that we will see in Part Two of this book.

We have discussed kinetics in earlier chapters and have shown how it can be used to more fully understand an enzyme and its regulation/control. We need to extend such kinetic analyses to whole pathways as well. We can do that using the VCell. Let's look at two different models of the MAPK cascade using Vcell. One particularly interesting feature is the regulation of the pathway. We will see in the next section on metabolism that pathways are often regulated by the end product of the pathway. This makes sense since if that end product is abundant, it would make biological sense for that product to inhibit the first enzyme in the pathway to avoid making more of the ultimate end product. Of course, that inhibition would be relieved as the concentration of the end product falls. Hence there is a temporal sense to the inhibition.

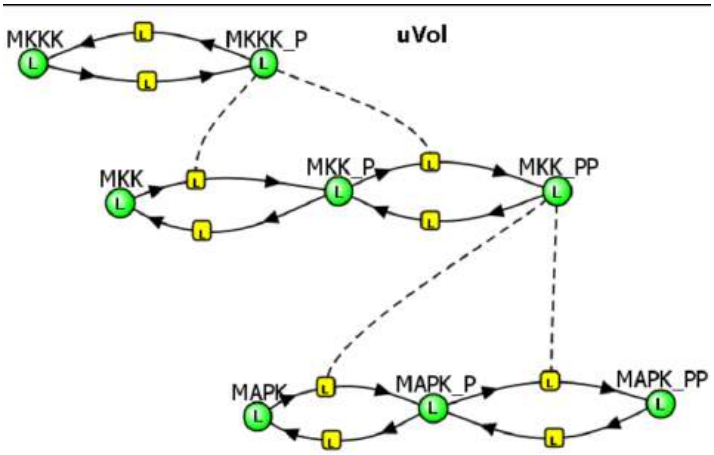
Let's look run two Vcell models for the MAPK Cascade, one with no end product inhibition and one with inhibition of the first step. Since we dealing with linked kinases, the inhibition of the first enzyme (MAP3K=MKKK) and hence the first reaction (MKKK ↔ MKKK_P) is not mediated by a chemical product of the last enzyme (MAPK_PP) but by phosphorylation of the first enzyme (MKKK) by the last (MAPK_PP).

MAPK Cascade - Model 1: **No** feedback inhibition of the MAPK cascade by inhibition of the first step (MKKK ↔MKKK_P) by the "end product" of the cascade (MAPK__PP) try quick

---

VCell MODEL

MAPK Cascade - no feedback inhibition



Initial Condition - See simulation results

Select Load [model name] below

NOFEEDBACK_Load Oscillations_in_MAPK_cascade

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed. For this model, select Vm, Km, Ki and I | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

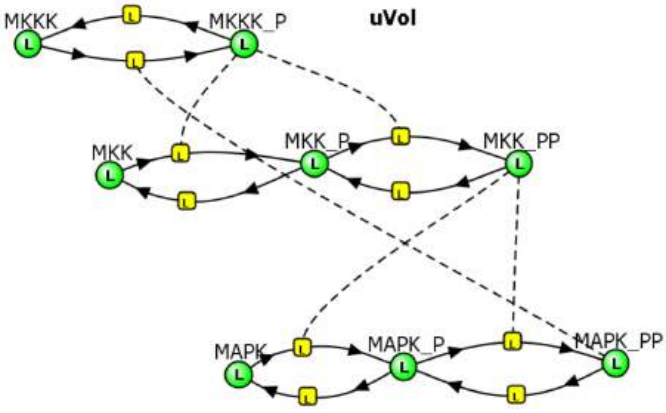
Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

The various concentration vs time curves in the output graph should make "intuitive" sense. There are no surprises!

Now let's add a twist. What if the last active enzyme in the pathway, MAPK that is doubly phosphorylated (MAPK_PP), the "final product" of the cascade, can, in a feedback reaction, inhibit the very first reaction of the cascade, MKKK → MKKK_P through an inhibiting phosphorylation. Run the simulation in Vcell to find out!

VCell MODEL

MAPK Cascade - With feedback inhibition



Initial Conditions: See simulation results

Select Load [model name] below

Load Oscillations_in_MAPK_cascade

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed. For this model, select Vm, Km, Ki and I | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

The various concentration vs time curves in the output graph should make "intuitive" sense. There are no surprises!

Now let's add a twist. What if the last active enzyme in the pathway, MAPK that is doubly phosphorylated (MAPK_PP), the "final product" of the cascade, can, in a feedback reaction, inhibit the very first reaction of the cascade, MKKK → MKKK_P through an inhibiting phosphorylation. Run the simulation in Vcell to find out!

MAPK Cascade - Model 2: Feedback inhibition of the MAPK cascade by inhibition of the first step (MKKK ↔MKKK_P) by the "end product" of the cascade (MAPK__PP)

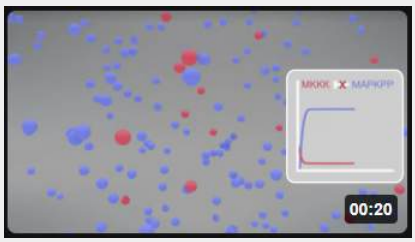
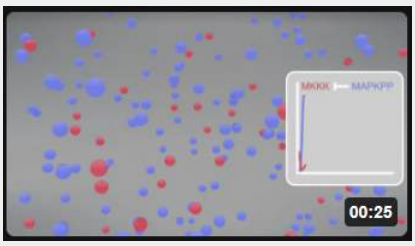
Kholodenko2000 - Ultrasensitivity and negative feedback bring oscillations in the MAPK cascade. https://www.ebi.ac.uk/biomodels/BIOMD0000000010. Based on Kholodenko BN. Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades. Eur J Biochem. 2000 Mar;267(6):1583-8. doi: 10.1046/j.1432-1327.2000.01197.x. PMID: 10712587.

There is utterly no way to have predicted, using intuition or simple mathematical analyses, the oscillatory changes in the concentration of all the enzymes and their phosphorylated variants. Of course, the oscillating concentrations depend on the chosen initial concentrations and on the rate constants.

When enzymes are linked sequentially in signal transduction pathways, their actions are amplified by the preceding enzyme. If the first kinase (MK3) activates a 1000 molecules of the second kinase (MK2), and each of these activates 1000 of the last kinase (MAPK), the net effect of activating the first kinase is a million-fold amplification of the last! This causes this three-kinase pathway to be **ultrasensitive** to small changes to the first enzyme. Add other activating and inhibiting agents and the net activities of the pathway components become even more complicated.

Just the simple feedback inhibition by the last enzyme of the first enzyme in this cascade can bring about the oscillation shown in Vcell model 2. Depending on the concentrations and rate constants used in the model, the oscillations can last from minutes to hours. These oscillations can produce waves of phosphoproteins that propagate through the cytoplasm of the cell.

Here is a simplified animation of the MAPK cascade that shows changes in MKKK (red dots) and MAPK_PP) blue dots with no feedback inhibition (left, Model 1) and feedback inhibition (right, Model 2) in the cascade. (Animations produced by Shraddha Nayak and Hui Lui.)

Model 1: **No** feedback inhibition of the MAPK cascade	Model 2: Feedback inhibition of the MAPK cascade
	

You might expect similar oscillatory behavior in proteins (cyclins and cyclin-dependent protein kinases) controlling the movement of cells through the cell cycle. We will see that in a subsequent section.

28.5: Small G proteins, GAPs and GEFs

28.5.1: G proteins: Cellular Switch for Kinases

In the preceding chapter sections, we discussed two types of small G proteins, Gα, part of the heterotrimeric Gαβγ complex linked to GPCR signaling, and Ras (H, K, and N). Both bind GTP and GDP and have GTPase activity. When bound to GTP they are active while the GDP bound form is inactive. What a perfect molecular switch to turn on signaling and with a built-in off switch (the GTPase activity). It turns out that his simple on/off switch is too simple. For example, a single mutation that inhibits the GTPase activity would leave the protein on continually which could (and does) lead to unregulated growth and tumor formation.

Two new sets of proteins that regulate the on-off activity of small G proteins are found abundantly in cells:

- GTPase activating protein or GAPs: As the name implies they enhance the GTPase activity of the small G proteins, which would decrease G protein signaling;
- Guanine nucleotide exchange proteins or GEFs: These lead to the dissociation of bound GDP and its replacement with GTP, which would increase G protein signaling.

Mammalian cells contain 3 variants of Ras: H, K, and N. They all bind GDP/GTP and have GTPase activity. Ras is targeted to the cell membrane through the post-translational addition of a hydrophobic farnesyl group. When activated by binding to GTP, it can bind to and activate a protein called **Raf**-1, which on binding becomes an active tyrosine kinase. Ras has intrinsic GTPase activity, so eventually, active Ras will deactivate itself.

Ras is just one member of a large superfamily of small G proteins, which all have GTPase activity. However, they are poor GTPases, so they need help to autocatalytically cleave the GTP to GDP. GAPs and GEFs evolved to regulate their activity by modulating the balance of bound GTP (an active form of the protein) and GDP (an inactive form of the small G protein).

Before describing these proteins, we need to have a better understanding of the family of small G proteins.

28.5.2: Small G proteins

Small G proteins in the superfamily have a common 20 K molecular weight catalytic (GTPase) domain with 5 alpha helices, 6 beta strands, and connecting loops. The small G proteins are "active" in the GTP-bound form. Hydrolysis of GTP to GDP causes the protein to become inactive. Figure 28.5.1 shows the domain structures of small G proteins.

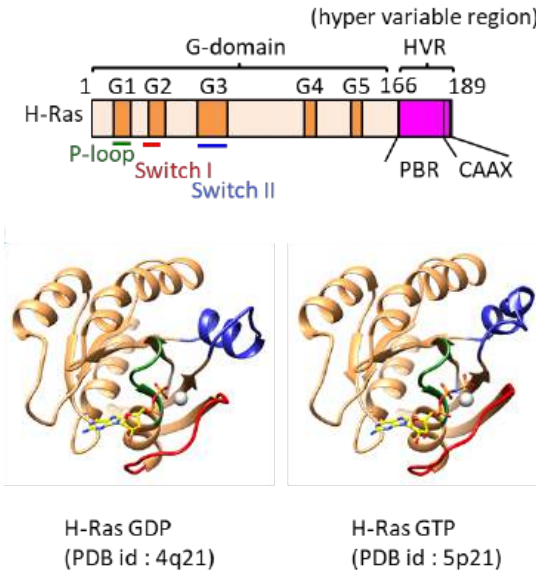


Figure 28.5.1: Domain Structure of small G proteins. Toma-Fukai et al. Molecules 2019, 24, 3308; doi:10.3390/molecules24183308. Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

G boxes of the G domain are highlighted with orange boxes. The hypervariable region, including a polybasic region and a CAAX motif, is highlighted with pink boxes. The P-loop, switch I, and switch II are shown as bars colored green, red, and blue, respectively. The bottom structures show H-Ras bound to GDP and GTP. The P-loop, switch I, and switch II are colored green, red, and blue, respectively.

Figure 28.5.2 shows an interactive iCn3D model of human KRAS G12C mutant covalently bound to AMG 510, a covalent inhibitor (6OIM). This mutation flips the Ras switch so it is permanently on.

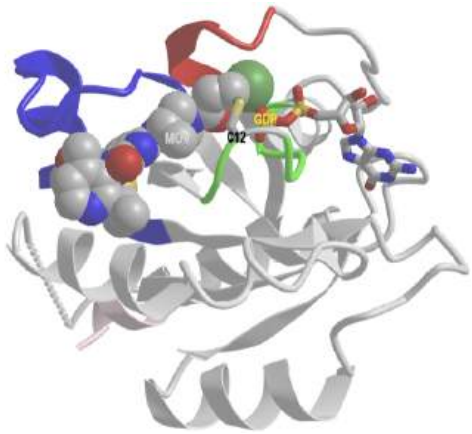


Figure 28.5.2: Human KRAS G12C covalently bound to AMG 510 (6OIM). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...Rx2ksMNmeCB7t6

The coloring coding is the same as in the above Figure 1:

- P-Loop (same as G1) is Green
- Switch 1 is red
- Swithch 2 is blue
- Polybasic region is pink
- The GAAX motif was not observed (low electron density in the crystal structure)

GDP is shown in colored sticks. The covalent inhibitor, AMG 510, is shown in spacefill and labeled MOV. It is covalently attached to the Cys 12 in the mutated version.

The G12C human KRAS mutation is found in about 13% or patients with non-**small cell lung cancer (**NSCLC) with is any type of epithelial lung cancer that is not small cell lung cancer**.** NSCLCs include squamous cell carcinoma, large cell carcinoma, and adenocarcinoma. The G12C mutation causes the protein to be "stuck" in the active conformation, which leads to continued activation of signaling pathways leading to cell proliferation, a trait of cancer cells. Before the discoveryof AMG 510, Ras was considered "undrugable" since it was devote of obvious pockets that could inhibit its activity. The G12C mutation replaces glycine with cysteine, a potent nucleophile that can react with the covalent inhibitor and inhibit the always-switched on G12C mutant in the absence of the inhibitor.

A new generation of inhibitors attempts to affect different aspects of its activity, including its localization to the membrane, the binding of different effectors, and nucletodide exchange.

Figure 28.5.3shows these key structural features of Ras (with different color coding).

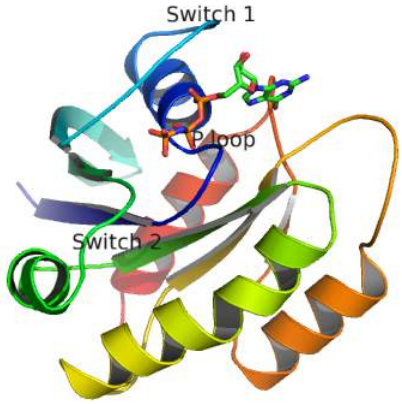


Figure 28.5.3: Key structural features of Ras

Important parts of **Ras** necessary for GTP binding include the phosphate-binding (P loop), residues 10 to 16 (dark blue trace below), switch regions I (30 to 37, light blue trace) and II (60 to 76, green trace), which are flexible loops which sandwich GTP.

Figure 28.5.4is an animation showing structural differences between the GTP bound form (blue, pdb id 5p21) and GDP form (red, pdb id 4q21) of the H-Ras protein. One helix and nearby loops are perturbed.

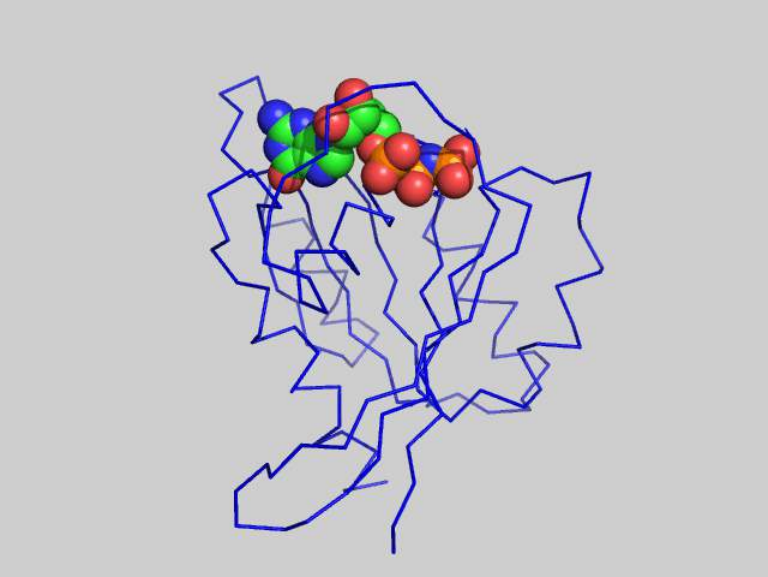


Figure 28.5.4: Structural differences between the GTP bound form (blue, pdb id 5p21) and GDP form (red, pdb id 4q21) of Ras

There are about 150 members of the human Ras superfamily as shown in Figure 28.5.5and Table 1 below.

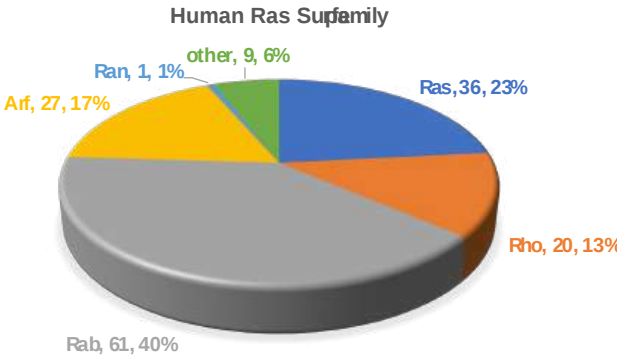


Figure 28.5.5: Members of the human Ras superfamily

Table 28.5.1shows common Ras superfamily functions

Ras	regulation of gene expression, cell proliferation, survival, and differentiation
Rho	regulation of actin cytoskeleton, cell shape, and movement, cell interactions with the extracellular matrix
Rab	vesicle trafficking, endocytosis, secretion
Arf	vesicle trafficking, endocytosis, secretion, microtubule assembly
Ran	nuclear cytoplasm transport, mitotic spindle

We have discussed Ran before as a mediator of protein movement across the nuclear membrane in Chapter 11.5. It's mainly in the GDP-bound form in the cytoplasm and the GTP-bound form in the nucleus. Switches between a cytoplasmic GDP- and a nuclear GTP-bound state by nucleotide exchange and GTP hydrolysis. Nuclear import receptors with bound cargo protein containing a nuclear import signal bind RAN-GTP in the nucleus, leading to the release of the importin and the cargo protein. In contrast, cargo

proteins with a nuclear export signal bind exportins and RAN-GTP in the nucleus and move into the cytoplasm, where the RAN-bound GTP is hydrolyzed on binding a RAN-GAP. This cycle is illustrated in Figure 28.5.6

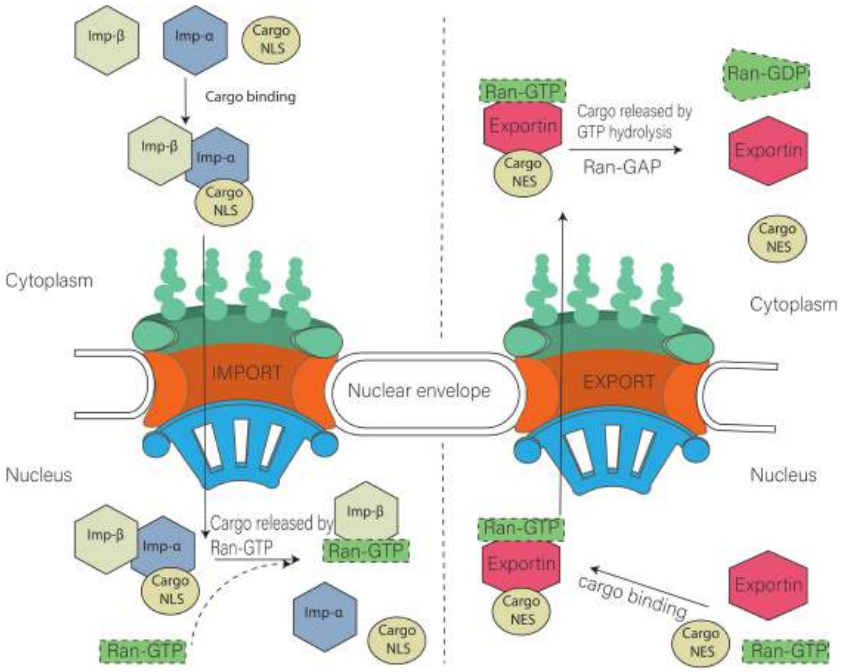


Figure 28.5.6: Model of nuclear import and export. Cargo containing NLS (Nuclear localization signal) is imported with the help of Importin α and importin β heterodimer. Nuclear export of cargo having NES (nuclear export signal) is carried out with the help of exportins. Ran GTP is also required during that process. Khan Asmat Ullah, Qu Rongmei, Ouyang Jun, Dai Jingxing. Front. Physiol., 03 April 2020 | https://doi.org/10.3389/fphys.2020.00239. **Creative Commons Attribution License (CC BY)**.

Small G proteins are a fundamental form of molecular switch. They are simply too important to not be regulated. Probably the most common mutation in human cancer cells involved a single amino acid change in Ras (H, K, and N form). If the GTPase activity is inhibited with mutation, the protein may be constitutively active. Such a Ras mutation is found in almost 90% of pancreatic cancers. Hence researchers have been trying to design drugs that inhibit its GTPase activity. This has proven difficult since it has very few targetable pockets that could bind a drug.

28.5.3: Regulation of small G proteins: GAPs and GEFs

Given the critical importance of small G proteins, it makes biological sense that their on/off activity would be exquisitely regulated. Indeed, they are. Two families of protein have evolved to regulate them by determining whether GTP or GDP is bound to the protein (leading to an active, and inactive small G protein respectively). One family, **GTPase activating proteins (GAPs)** facilitate the hydrolysis of bound GTP, leading to the inhibition of the protein. The other proteins are **GTP exchange proteins (GEFs)**, which facilitate the exchange of GTP for GDP, activating the protein.

The activity of Ras GAPs and GEFs, as well as various proteins interacting with Ras, are depicted in Figure 28.5.7.

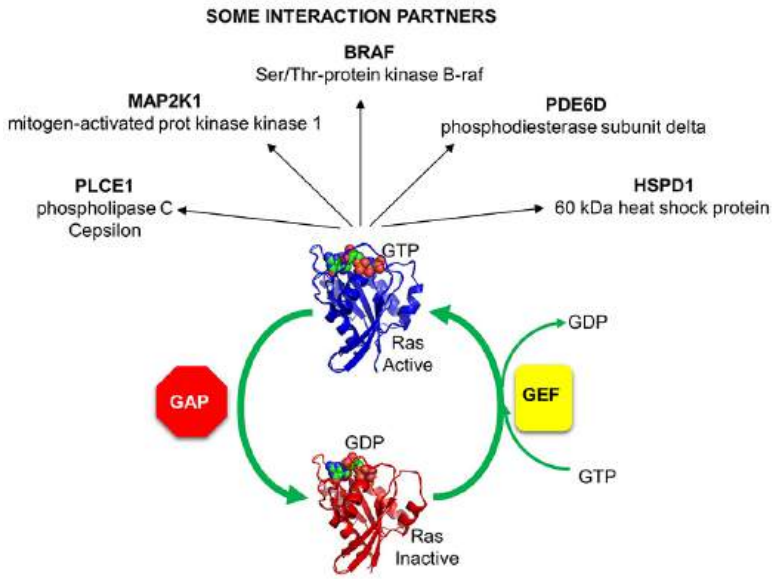


Figure 28.5.7: The activity of Ras GAPs and GEFs, as well as various proteins interacting with Ras

It may seem crazy but the number of GEFs and GAPs is greater than the number of G proteins with which they interact. As shown in Figure 28.5.4, there are 20 Rho G proteins but about 80 GEFs and 70 GAPs for them. This number presumably allows greater control of the specificity of the reactions controlled by the Rho G protein.

GAPs - GTPase Activating Proteins

The hydrolysis of the gamma phosphate of GTP by water in Ras proceeds by a pentavalent transition state with two axial and three equatorial ligands to the P. Developing charge in the transition state would usually be stabilized by catalytic residues in the catalytic domain of Ras. However, Ras is a poor GTPase. That's where GAP comes in. In the Ras/GAP complex, GAP positions its Arg 789 on the GAP in a position to stabilize the transition state for Ras-bound GTP cleavage. This Arg 789 is almost in the same position as Arg 178 in the Galpha inhibitor subunit of a heterotrimeric G protein which inhibits GPCR signaling. Both of these arginines have similar catalytic functions.

Figure 28.5.8shows an interactive iCn3D model of Ras-GAP complex (1WQ1)

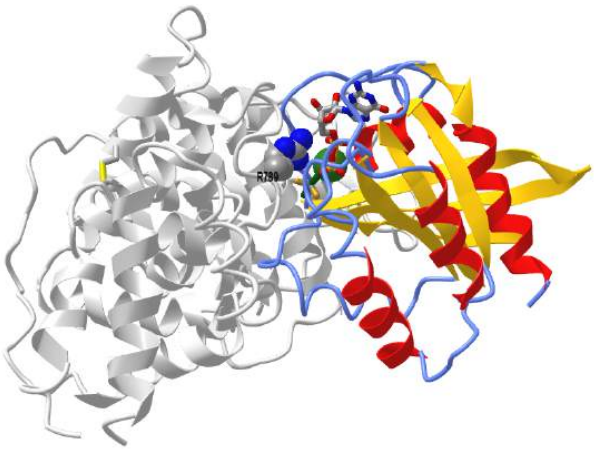


Figure 28.5.8: Ras-GAP complex (1WQ1) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...MJiLGpT4w6kP69

Ras is shown in secondary structure colors, while GAP is shown in gray. GDP-AlF3, a GTP analog, is shown in color spacefill. Arginine 789 in GAP is shown in spacefill with CPK colors and labeled R789. It is a clear position to stabilize the bound GTP in the complex and its cleavage transition state.

GEFs - GTP Exchange Factor

Once bound to Ras, GDP dissociates very slowly. Values of 10^{-5} sec^{-1} have been reported for the first order rate constant of the dissociation of GDP from a small G protein. Assuming that the diffusion is controlled on the rate constant for the complex, the K_D for the G protein:GDP complex would be 0.1 pM and the half-life would be 0.8 days, similar to the lac repressor:DNA operator complex. Hence the protein, when bound to GDP, is essentially locked in the off position. What if it needs to be reactivated quickly? How can the rate at which GDP dissociates be increased so that GTP could replace it? If it were to dissociate, GTP could quickly replace it since from an equilibrium point of view, Ras and other small G proteins would favor GTP binding since its concentration in the cell is higher.

One could envision several ways to change the rate at which GDP dissociates. In organic chemistry, a favorite student response to many questions is to evoke steric effects. In biochemistry, the analog is often conformational changes. How could you change the conformation of Ras such that it might favor GTP binding? That could occur by ligand binding or more likely by a post-translation modification such as phosphorylation as part of a signaling process. It turns out that for the case of small G proteins, another mechanism is evoked: the binding of another protein, a **GTP Exchange Factor** or **GEF**, which promotes GTP exchange for the bound GDP. If the Ras:GEF:GDP complex has a 10,000 increase in k_{off} for GTP, the half-life of the bound GDP is 7 seconds. There are 80 GEFs in the human genome. If you think about it, in GPCR coupled signaling, the ligand-bound GPCR is a GEF for the Gαsubunit of the heterotrimeric Gαβγ protein.

The crystal structure of a Ras GEF, SOS, in complex with Ras allows a detailed understanding of the mechanism. SOS, a cytoplasmic protein, is recruited to the cell membrane where active Ras is found, tethered to the membrane with a hydrophobic farnesyl attachment. Figure 28.5.9 shows an interactive iCn3D model of Ras and SOS (a **GEF**) complex (1bkd).

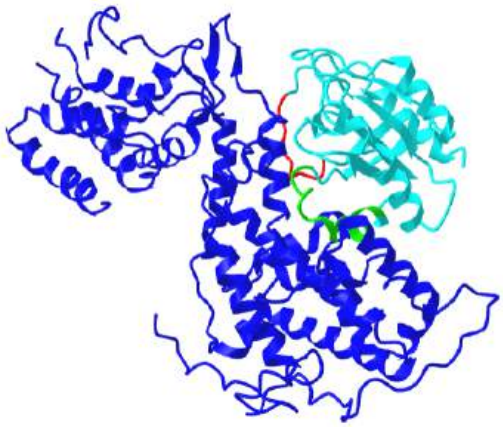


Figure 28.5.9: Ras and SOS (a **GEF**) complex (1bkd) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...XJP4Fajqc9eEL6

The actual biological unit (functional structure) is a hetero 8-mer (A_4B_4) with C_4 symmetry. The iCn3D model shows just a heterodimer for clarity. Ras is shown in cyan and SOS in dark blue. SOS as a GEF affects nucleotide binding to SOS in two essential ways. An alpha helix from SOS displaces Switch 1 (amino acids 30-38, shown in red) in Ras, which opens the binding site for the guanine nucleotides open. Additional conformational changes in Switch II (59–72, green) in Ras and interference from the side chains form the SOS alpha helix interfere with the binding of the phosphates on the bound nucleotide. This promotes the dissociation of the bound nucleotide and Mg^{2+}. Now GTP can preferentially rebind. How?

Before we answer that question, let's explore the conformational differences just in the structure of Ras in the Ras:GDP complex (4q21) and Ras in the Ras:SOS (1bkd) complex. These differences are shown in Figure 28.5.10 The green structure is the Ras:GDP. The cyan structure is Ras without bound GDP but bound to SOS.

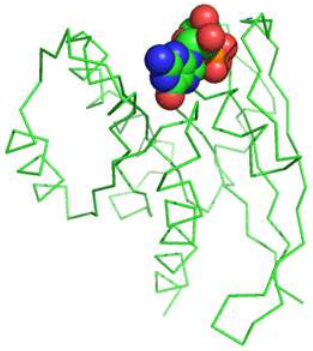


Figure 28.5.10: Structure difference in Ras in the Ras:GDP complex (4q21) and Ras in the Ras:SOS (1bkd) complex

Note the large shift in Switch 1 in the Ras structure from the Ras:SOS complex. This leaves a "gaping" hole from which GDP can "escape". Now how do the opening of the active site and release of bound GDP facilitate GTP binding?

Once bound just to Ras, GDP dissociates very slowly. Values of 10^{-5} sec^{-1} have been reported for the first-order rate constant of the dissociation of GDP from a small G protein. Assuming that the on-rate for complex formation is diffusion-controlled, the K_D for the G protein:GDP complex would be 0.1 pM and the half-life would be 0.8 days, similar to the lac repressor:DNA operator complex. Hence the protein, when bound to GDP, is essentially locked in the off position.

The conformational changes on RAS binding to SOS open up the active site, allowing GDP dissociation. GTP can now replace it since from an equilibrium point of view, Ras and other small G proteins favor GTP binding since the concentration of GTP in the cell is higher than that of GDP. In addition, some additional noncovalent interactions with the extra phosphate on GTP probably help.

Figure 28.5.11shows a cartoon showing the changes in Ras on GEF binding as illustrated in these coupled chemical equilibria:

$$GEF_{loose}:Ras:GDP_{tight} \leftrightarrow GEFtight:Ras:GDP_{loose} \leftrightarrow GEF_{tight}:Ras$$

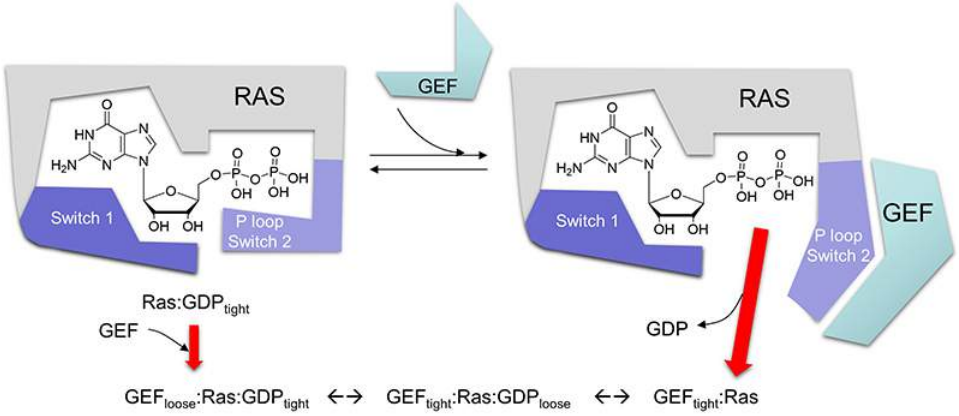


Figure 28.5.11: Chemical equation for Ras:GEF:GDP interactions

28.6: Phosphatases

28.6.1: Introduction

Protein kinases phosphorylate proteins in a process that can either activate or inhibit the target protein's activities. To control signaling processes, the activities altered by protein phosphorylation can be readily reversed by dephosphorylation of the Ser-, Thr- or Tyr-phosphoesters by simple hydrolysis. These reactions are catalyzed by protein phosphatases. Some of these phosphatases also cleave phosphates from lipids as well.

There are three main families of phosphatases, phospho-Tyr phosphatases (PTP), the phospho-Ser/Thr phosphatases, and those that cleave both. Of all phosphorylation sites, most (86%) are on serine, 12% involve threonine and about 2% are on tyrosine. They can also be categorized by molecular sizes, inhibitors, divalent cation requirements, etc. In contrast to kinases which differ in the structure of their catalytic domains, many **p**rotein **p**hosphatases (**PP**s, also abbreviated Ppp for **P**rotein **p**hos**p**hatases) gain specificity by binding protein cofactors which facilitate translocation and binding to specific phosphoproteins. The active phosphatase hence often consists of a complex of the phosphatase catalytic subunit and a regulatory subunit. Regulatory subunits for Tyr phosphatases may contain a SH2 domain allowing binding of the binary complex to autophosphorylated membrane receptor Tyr kinases.

We'll consider examples of all four families. They recognize target proteins through protein:protein interactions and specific binding site motifs. There are over 10,000 pSer and p-Thr sites on target proteins, so targeting specific sites must be quite nuanced.

28.6.2: Serine/Threonine Phosphatases

Important Ser/Thr phosphatases (PPs for Protein Phosphatases) include:

- Protein phosphatase 1 (PP-1 or Ppp1) - This is the most abundant PPP in humans. Different regulatory subunits target this to the liver glycogen particles (GL subunit), striated muscle glycogen, and sarcoplasmic reticulum (GM subunit) or smooth muscle fibers (M subunit). It is also present in the nucleus where it is presumably involved in the regulation of transcription factors. It is also involved in RNA splicing and signaling at neural synapses.
- Protein phosphatase 2A (PP-2A or Ppp2) - is a trimer with catalytic, regulatory, and scaffolding (also regulatory) structural subunits. It is found mainly in the cytoplasm and is involved in a myriad of cellular processes.
- Protein phosphatase 2B (PP-2B or Ppp3) - also called calcineurin or Ca^{2+}/Calmodulin dependent protein phosphatase - It consists of a catalytic subunit (calcineurin A) and a regulatory, calcium-binding subunit (calcineurin B). It is inhibited by the complex of the immunosuppressant cyclosporin and FK506 with immunophilins. PP2B regulates PKA and PKC
- (PP2C) -

The catalytic subunits of PP1, 2A, and 2B share a great deal of amino acid homology, and based on this homology, belong to one family. PP2C belongs to another family. PPs are often categorized into three other families including, phosphoprotein phosphatases (PPPs) and metal-dependent protein phosphatases (PPMs). There are about 30 catalytic PP subunits (many fold fewer than Ser/Thr Kinases). They gain specificity by binding numerous modulatory regulatory subunits.

As with other proteins, the names given to the proteins when discovered often do not reflect an organization scheme that would name different members based on structural similarities. PP-1, 2A, and 2B are better named Ppp1, Ppp2, and Ppp3 which denote members of the Protein PP (PPP) family. PP-2C would be named Ppm1 as the first member of the PPN family. All PPPs have three short sequence motifs that bind divalent cations.

In addition, older names for the given PPs referred to both the catalytic subunit and the dimer with the regulatory subunits. For clarity, the name of the catalytic protein phosphatase 1 is PP1c, and its regulatory subunits as RIPPOs, **r**egulatory **i**nteractors of **p**rotein **p**hosphatase **o**ne.

28.6.2.1: Protein Phosphate-1 (PP-1):

PP-1 is involved in many signaling pathways that control cell division, protein synthesis, etc. It catalyzes most serine–threonine dephosphorylation in cells. It is perhaps best known for its regulation of glycogen mobilization. Insulin signals the well-fed state in healthy people and promotes glucose uptake through the GPCR insulin receptor which we discussed in Chapter 12.4. Under these condition excess glucose is used to elongate glycogen, our main carbohydrate energy storage polymer. In contrast, the starving or low energy state is signaled by the hormone glucagon. You would expect that signaling pathways activated in the presence of insulin would promote glycogen synthesis and inhibit glycogen breakdown. PP-1 is a key factor in the regulation of both processes:

- Insulin activation of glycogen synthesis - PP1 dephosphorylates glycogen synthase (the enzyme that synthesizes glycogen) and in the process ACTIVATES it.
- Insulin inhibition of glycogen breakdown - PPI dephosphorylates two key enzymes involved in glycogen breakdown, phosphorylase kinase and glycogen phosphorylase a (with a pSer14), and in the process INHIBITS them.

PP1c interacts with many different regulatory subunits (RIPPOs) forming unique heterodimers. The regulatory subunits also bind potential substrates for PP1c and help localize the enzyme. at sites. The regulatory subunit involved in the PP1c effect on glycogen is called the glycogen-targeting subunit, G_M. There are 7 such regulatory subunits involved in glycogen metabolism. G_M (R_{GL}) is expressed in muscles and G_L in the liver. All of the regulatory subunits (unfortunately called G-subunits) have a conserved RVxF amino acid sequence which interacts with specific sites on the catalytic subunit typically distal to the active phosphatase site. Binding through the RVxF sequence does not affect the active site of the PP-1c. The binding of the regulatory subunit to the PP-1c can also occur outside of the canonical RVxF sequence. The regulatory subunits also have starch binding domain (SBD), also called the carbohydrate-binding module (CBM21). The subunits are often highly disordered until they are bound.

Figure 28.6.1 shows an interactive iCn3D model of protein phosphatase 1 (PP1) bound to the muscle glycogen-targeting subunit (Gm) and microcystin (6DNO) and the toxin microcystin.

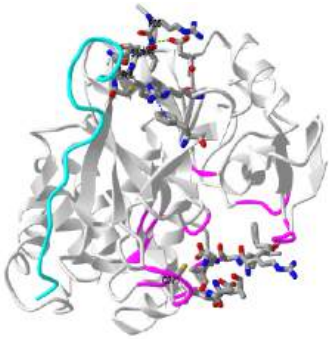


Figure 28.6.1: Protein phosphatase 1 (PP1) bound to the muscle glycogen-targeting subunit (Gm) and microcystin (6DNO). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...6f8f2kBRW4uhv5

PP-1 alpha catalytic chain is shown in gray. Its active site, where it binds Ser/Thr phosphorylated proteins is shown in magenta. Bound in that site is the toxin microcystin. The small chain shown in cyan is protein phosphatase 1 regulatory subunit 3A. PP-1 binds to its regulatory sequence to a 65RVxF68 sequence common on many regulatory PP-1 subunits. This subunit contains a serine 67 which is phosphorylation by protein kinase A. This inhibits the binding of the catalytic and regulatory subunits. (PKA) of the “x” residue in the G_M RVxF motif, Ser^{67}_{GM}, inhibits PP1 binding (*16*). Val79 and Lys80 on GM also form a motif that binds a corresponding pocket in the catalytic subunit.

Note the pi-cation (red) and pi-stack (blue) interactions from Phe 68 of the G_M 65RVSF68 motif of the regulatory subunit to the PPI catalytic subunit. Microcystin is 7-mer peptide ring that has 5 noncanonical amino acids and 2 regular ones. A covalent bond from Cys273 of PP1 (labeled in the above iCn3D model) to the methyl-dehydroalanine (Mdha) of the toxin forms. Microcystins are produced by toxic cyanobacteria and are very toxic and lethal, especially to animals including humans that drink water contaminated with the cyanobacteria. They will pose a greater threat in a warming world from climate change. They also bind to PP-2A.

Another example of a regulatory subunit for PP1 is the PP1 nuclear target subunit (PNUTS). The activity of the complex in the nucleus regulates the phosphorylation state o many proteins involved in the cell cycle including p53, a tumor suppressor in many tumors. It also regulated chromatin structure and RNA processing. As with the muscle glycogen-targeting subunit (Gm), PNUTS is intrinsically disordered when not bound to PP1 and is very extended when bound. The catalytic subunits of PP1c and PP2A have an acidic, hydrophobic, and C-terminal grove. PNUTS binds one of the substrate groves at an arginine subsite, but not the active site.

So far we can say the specificity of the binding of PPP is determined to some degree by the regulatory subunits. We turn to the contributions of the catalytic specificity of PP-1c when we compare it to PP-2Ac below.

Mechanism of PP1-A and PP-2A and B

The catalytical subunits of PP-1 and PP-2 are very homologous with nearly identical key active site side chains. The site contains two Mn^{2+} ions very near each other as shown in Figure 28.6.2

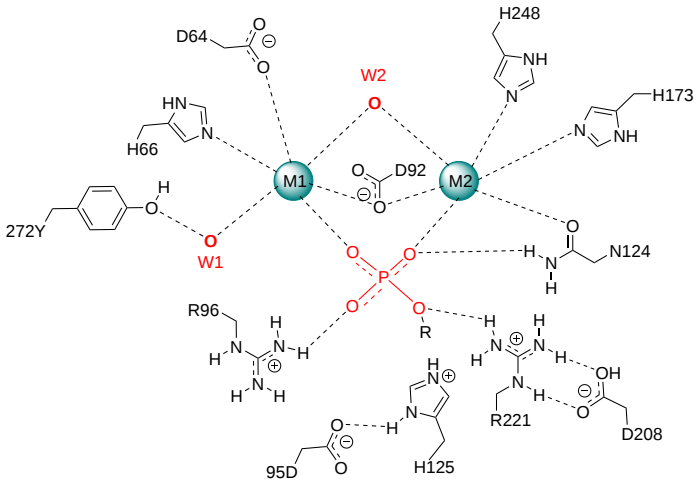


Figure 28.6.2: Interactions with two Mn ions in the active site of protein phosphatase-1 (PP-1). After McWirter et al. *J. Am. Chem. Soc.* 2008, 130, 41, 13673–13682

The figure crudely shows each Mn^{2+} is octahedrally coordinated to side chains and one water oxygen for each metal ion. Water 2 (or more likely OH^-) and an aspartate D92 oxygen bridge the two ions. The phosphate of the target Ser-OPO_3^{2-} or Thr-OPO_3^{2-}, or an inhibitor such as tungstate, also bridges the metal ion.

Figure 28.6.3 shows an interactive iCn3D model of Human Protein Phosphatase 1 Active Site Residue (4mov) which shows the same active site residues

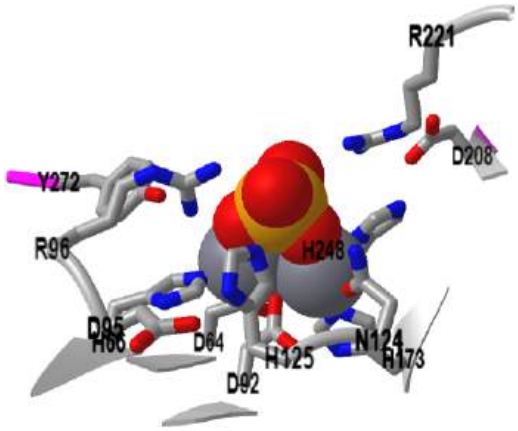


Figure 28.6.3: Human **Protein Phosphatase 1 Active Site Residue** (4mov). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?MYPrWfkX1PBaAggL9

The metals singly or in combination probably reduce the pKa of bound water to produce the deprotonated hydroxide, which engages in an S_N2 attack on the phosphate. Hence the metal ions act as an electrostatic catalyst.

The subtle differences in the active site and the three groups contribute to the specificity of the PPPs.

28.6.2.2: Comparison of the catalytic subunits of PP1A the PP2A

As mentioned above, substrate specificity is altered by subtle changes in the active site and three groves of PPPs. Figure 28.6.4 shows the acidic groves for PP1c and PP2Ac.

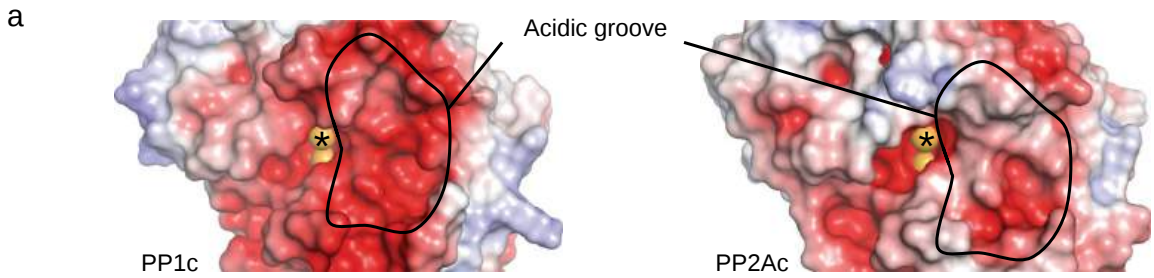


Figure 28.6.4: Acidic groves for PP1c (3EGG, chain A) and PP2Ac (4I5L, chain C)

The color is the electrostatic surface potential with red indicating negative and blue (notably absent) positive. The acidic groove is stronger in PP1c. The asterixis * shows the catalytic cleft containing two Mn^{2+} ions. Hoermann et al. Nature Communication | (2020) 11:3583 | https://doi.org/10.1038/s41467-020-17334-x. Creative Commons Attribution. 4.0 International License, http://creativecommons.org/licenses/by/4.0/.

The negative acidic grove is highly enriched in negatively charged side chains. Figure 28.6.5 shows the actual amino acids contributing to the negative electrostatic potential in aligned PP1c and PP2Ac.

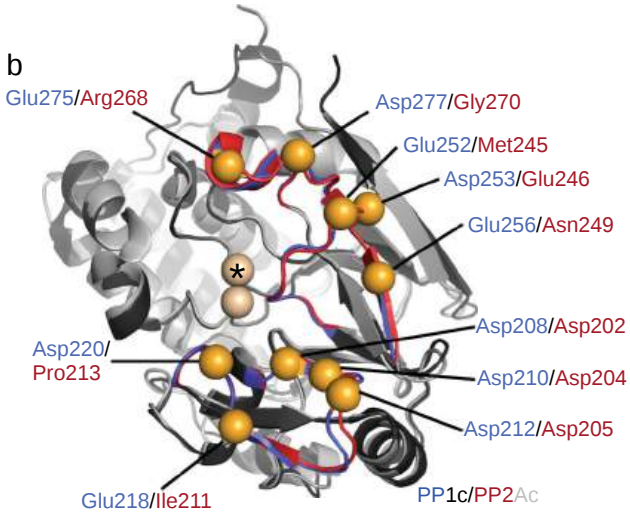


Figure 28.6.5: Amino acids contributing to the negative electrostatic potential in aligned PP1c and PP2Ac. Zhang et al. Hoermann et al. ibid

The orange spheres show the location of acidic side chains. PP1c is shown in blue/black and PP2Ac in red/gray.

The stronger acidic groove in PP1Ac gives it great preferences for pSer/pThr-protein targets with basic motifs than PP2Ac in a fashion that is independent of the bound regulator subunit. In contrast, PP2A needs to interact with regulatory subunits with more acidic composition to target basic motifs in protein targets. These features for PP1c and PP2Ac are compared in Figure 28.6.6

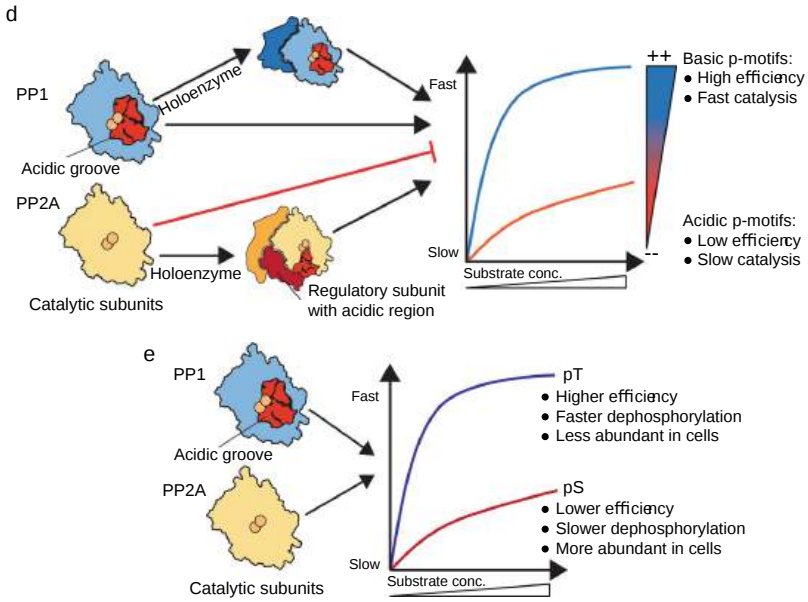


Figure 28.6.6: Comparative activity of PP1 and PP2A to basic target protein substrates.

Panel (d) shows that the holoenzyme for PP1 has a great preference for positive basic motifs than the holoenzyme for PP2A, which needs to associate with a negatively charged regulatory subunit for activity towards target proteins with basic motifs.

Panel (e) also shows that the catalytic subunits of both PP1 and PP2A prefer p-Thr protein targets compared to p-Velocity vs substrate graphs show these effects as well. e Both, PP1 and PP2A holoenzymes have a preference for pT due to higher catalytic efficiency of their respective catalytic subunits towards pT over pS. Hoermann et al. ibid

Figure 28.6.7 shows an interactive iCn3D model of the protein phosphatase 2A catalytic subunit in complex with a larger regulatory subunit and bound to the phosphatase inhibitor and tumor promoter okadaic acid (2IE4). The toxin, found in sponges and shellfish, is produced by dinoflagellates.

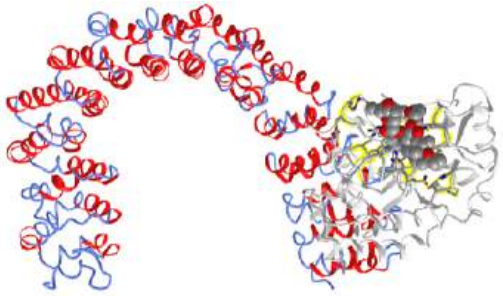


Figure 28.6.7: Protein phosphatase 2A core enzyme (catalytic and regulatory subunit) bound to okadaic acid (2IE4). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...Sb2Vq5iZgbBSz8

The regulatory subunit is shown in secondary structure color. This scaffolding protein is shaped like a horseshoe. The phosphatase inhibitor okadaic acids are shown in spacefill bound to the catalytic subunit shown in gray. The side chains of the catalytic subunit interacting with the 2 Mn^{2+} ions are shown in CPK-colored sticks (zoom in to see them). On binding the catalytic subunit, the scaffolding regulatory subunit is quite flexible and adaptable in interacting with other proteins.

28.6.2.3: Protein Phosphate 2B: Calcineurin (CN)

Calcineurin (CN), or PP2B, is depended on Ca^{2+}. It is involved in the development, immune signaling, and heart function. It consists of a catalytic site (CNA) and a calcium-binding regulatory subunit CNB so it is another example of PPP heterodimers. CNA has a catalytic domain and domains that bind CNB (the regulatory subunit), calmodulin (CAM, a calcium-binding protein that we will explore more in the next chapter section), and an autoinhibitory domain that blocks the active site. On Ca^{2+} release from internal organelles, the ion binds to both CNB and also CAM. These events cause conformational changes that release the bound autoinhibitor.

As with the other phosphatases, much effort has been made to determine how CN interacts with specific pSer- and p-Thr sites on targets. We'll focus on one, the integral membrane Na^+/H exchanger 1 (NHE1). This protein is itself regulated by Ca^{2+} ions and by phosphorylation by kinases we have previously studied, the MAPK ERK1/2 and the JNK kinase. Erk2 phosphorylates NHE1 at 6 Ser/Thr side chains in the recognition sequences name [S/T]P11. Several different phosphatases, including CN, can regulate NHE1 activity through direct dephosphorylation.

CN binds **short linear motifs (SLiMs)** named PxIxIT and the LxVP that are found in interacting partners including regulatory subunits as well as inhibitors and substrates. As we saw above, the regulatory subunits of PP1A and PP2A are highly disordered. Likewise, SLiMs are on intrinsically disordered regions as well as interacting proteins.

- PxIxIT binds to the catalytic domain of CNA22. It also enables interaction between CN and NHE1.
- LxVP binds to a cleft between the CNA and CNB, which is only available in the active form of the protein.

CN doesn't dephosphorylate multiple nearby p-Ser side chains of NHE1 (pS363, pS723, and pS726) since they are close to the NHE1-PxIxIT interaction sit,e which sterically restricts their binding to the active site. However the 3 other phospo S/T sites on NHE1 (pS771, pS785, and the actual target site (pT779) are far enough away from the NHE1's PxIxIT site so they can interact with the CN active site. Making the T779S mutation shows that dephosphorylation of their phosphorylated version shows a faster rate with pT779 and a slower, yet reasonable rate for pS779. Therefore other specificity factors are in play. A newly discovered very short 4-amino acid site motif in NHE1 including pS779 appears to be a source of selectivity. This TxxP motif in NHE1 is 779**T**PAP782. Such short recognition motifs are different than the selection of substrates by PP1 which involved multiple domain binding interactions and steric restrictions imposed by them.

Figure 28.6.8shows a model of the structure of the NHE1 exchanger (left panel a) and the calcineurin CNA/CNB complex.

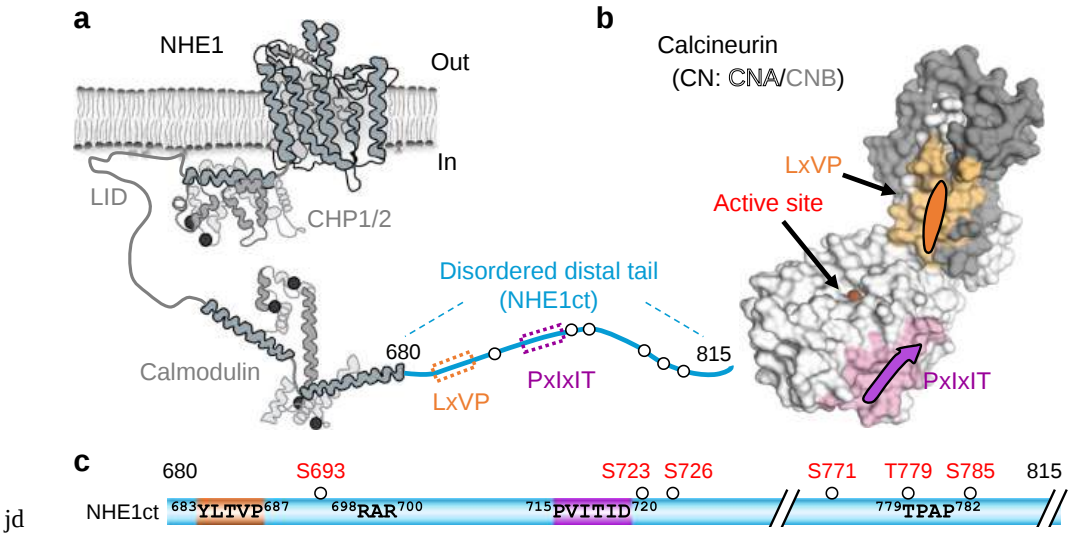


Figure 28.6.8 Docking motifs mediate the interaction of NHE1ct with CN

The motifs present in NHE1 (LxVP and PxIxIT) in the intrinsically disordered tail of NHE are indicated in the left panel and their corresponding binding sites in the CNA dimer are shown in corresponding colors. The calmodulin binding site is also shown. Erk2 phosphorylation sites in NHE1 are shown are listed in panel C along with the consensus motif sequences (PxIxIT in purple, LxVP in orange, and TRAP (uncolored)). Hendus Altenburger et al. Nature Communication (2019) 10:3489 | https://doi.org/10.1038/s41467-019-11391-7Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Figure 28.6.9shows an interactive iCn3D model of calcineurin (PP2B) complex bound to a peptide from the Na^+/H^+-exchanger 1 (6NUC)

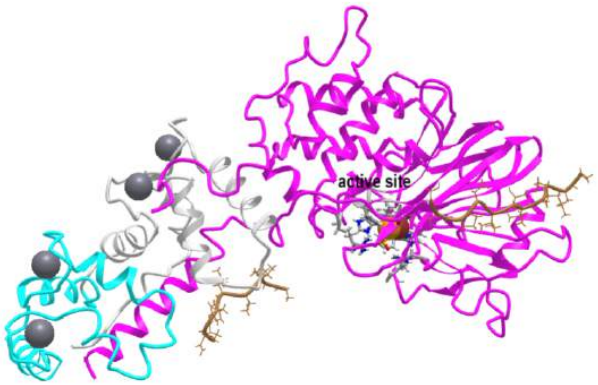


Figure 28.6.9: calcineurin (PP2B) complex bound to a peptide from the Na^+ /H^+ -exchanger 1 (6NUC). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...GgqSWpirpnZGt9

28.6.2.4: Protein Phosphatase 2C

The protein phosphatase 2C (PP2C) is a member of a family of metal-dependent protein phosphatases sometimes abbreviated as PPMs. (Of course, the other phosphatases we discussed above are also metal-dependent.) The required Mg^{2+} or Mn^{2+}. PP2C is a monomeric enzyme with at least four isoforms in humans. In humans, there are at least 17 members. One is unfortunately named protein phosphatase **1D** but also referred to as PPM1D, PP2Cδ, or Wip1). It is involved in heterochromatin silencing and the cycle. Mutations in its gene can accordingly give rise to tumors.

PP2Cs in humans have a binuclear metal cluster which reduces the pKa of water, producing OH^- for S_N2 attack on the phosphorus in the pSer or pThr in target phosphorylated proteins. Figure 28.6.10 describes binding interactions around the binuclear site in PP2Cα.

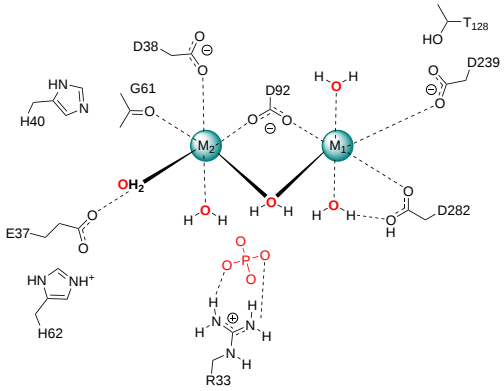


Figure 28.6.10 Binding interactions around the binuclear site in PP2Cα. after Pan et al. *Sci Rep* **5,** 8560 (2015). https://doi.org/10.1038/srep08560htt...cles/srep08560

The metal binding site with many water molecules Interactions with the metals is very different than the PP1 and PP2 active site shown in Figure 28.6.2

If Cd is bound at the M1 site, the activity of the enzyme is blocked so it is required for catalysis. Making the mutations affecting M2 (D38A and D38K) suggests that M2 is involved in binding the phosphate of the substrate, and also stabilizes the transition state and the leaving group in the reaction. H62 probably acts as a general acid catalyst.

Not all PPCs require both metal ions. The plant hormone abscisic acid regulates stress responses in plants. When it binds to a particular receptor called PYL1 (alternative name PYR1-like protein 1), the receptor interacts with a PP-2C called **ABI1** (also

called Absisic acid-insensitive 1). Figure 28.6.11shows an interactive iCn3D model of the ternary complex of Abscisic acid, PYL1 and ABI1 (phospholipase 2C) (3KDJ)

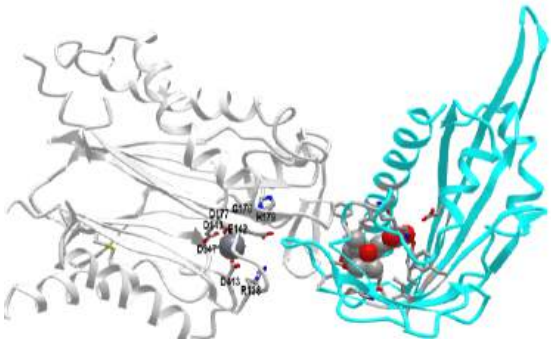


Figure 28.6.11: Ternary complex of Abscisic acid, PYL1, and ABI1 (phospholipase 2C) (3KDJ Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?FbEeZ3n5dWvA6bp56

The catalytic subunit, which in contrast to PP1, PP2A, and PP2B, has only 1 Mn^{2+} ion, is shown in gray with amino acids side chains interacting with the metal ion labeled. The cyan subunit is the receptor of abscisic acid, which is shown in spacefill.

28.6.3: Protein Tyrosine Phosphatases

Protein Tyr phosphatases (PTPs) consist of receptor-like (transmembrane) and intracellular Tyr phosphatases. They more resemble tyrosine kinases in their complexity than the Ser/Thr phosphatases. There are about 100 PTPs in the genome, a number similar to the number of protein tyrosine kinases. PTPs have an active site Cys in a CX5R-(S/T) motif with an active site Cys nucleophile and an Arg in the phosphate binding (P) loop. Some examples we will discuss include:

- Low molecular weight PTPase - These have roles in the metabolism and differentiation of cells. They have a molecular weight of 18,000 and have an active site CX5R-(S/T) motif, where the C (Cys) is an active site nucleophile.
- PTP1B - dephosphorylates many cell surface receptors (insulin, EGF, PDGF) that have been phosphorylated on Tyr residues. Its main activity seems to dephosphorylate nascent receptors in the endoplasmic reticulum before they get to the final cell membrane destination.
- Tyrosine phosphatase nonreceptor type 11, *ptpn11*, commonly called SHP2

Figure 28.6.12shows the protein tyrosine phosphatase (PTP) superfamily.

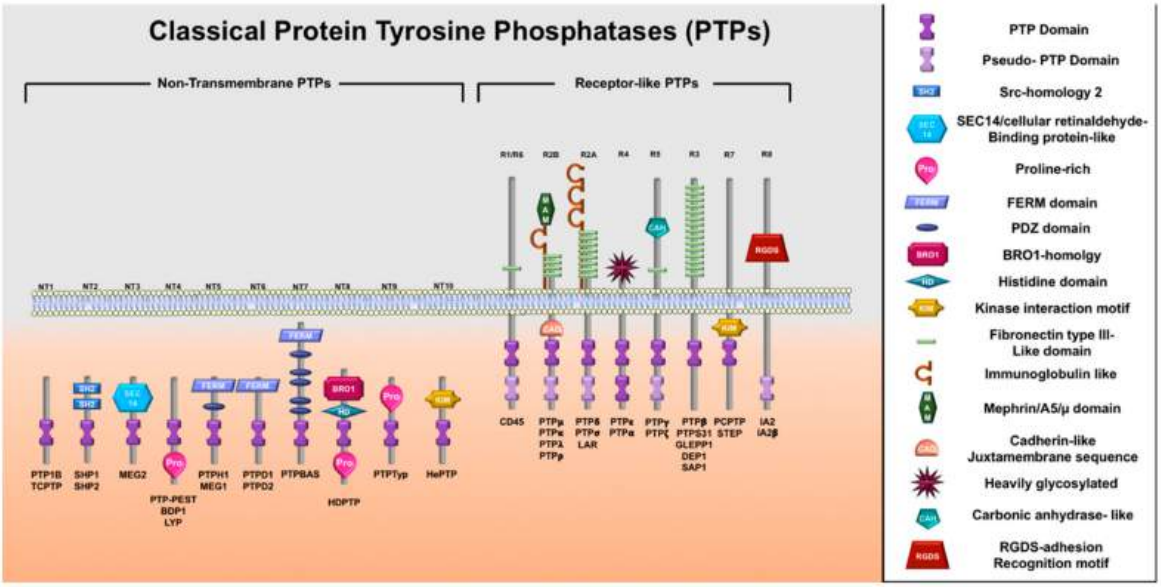


Figure 28.6.12: Protein Tyrosine Phosphatase superfamily Abdelsalam et al. *Biomolecules* **2019**, *9*, 286. https://doi.org/10.3390/biom9070286. Creative Commons Attribution 3.0 Unported License **https://creativecommons.org/licenses/by/3.0/**.

In contrast to the active site of the Ser/Thr phosphatases like PP-1, PP-2A and PP-2B, the active sites of protein tyrosine phosphatases (PTPs) do **not** have a bimetal ion cluster in the active site. Rather they all have an active site cysteine that acts as a nucleophilic catalyst in the hydrolysis of the p-Try phosphoester bond. The active site PTP domain is found in all of the proteins, so all use similar catalytic mechanisms shown in Figure 28.6.13

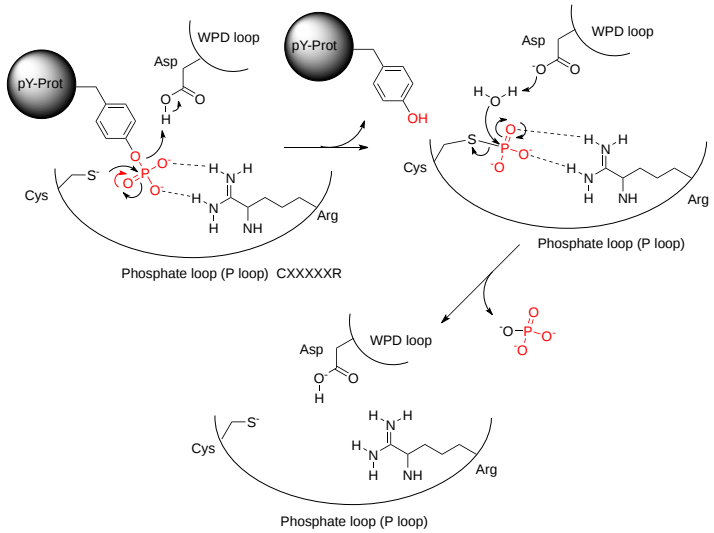


Figure 28.6.13: General catalytic mechanism of protein tyrosine phosphatases

The phosphotyrosine side chain of the target phosphoprotein binds in the phosphate-binding P-loop (H/VCxxxxxRS/T), which contains the nucleophilic Cys 12 and Arg 18 that stabilizes the charge on the phosphate. The Asp in the WPD loop positioned across from the nucleophilic Cys 12 acts as a general acid. Since the active site is nearly identical in the PTPs, it has been hard to design drugs that bind to the active site but that are also selective for specific PTPs.

28.6.3.1: Low molecular weight protein tyrosine phosphatase - LMW-PTP

This protein tyrosine phosphatase is the simplest of all in structure. It has the phosphate-binding P-loop (12CxxxxxR18) with the nucleophilic Cys 12 and Arg 18 that stabilizes the charge on the phosphate. It does not have the conserved WPD loop but deploys Asp 129 across from Cys 12 as a general acid. This enzyme exists as two main isozymes, A and B. Figure 28.6.14 shows an interactive iCn3D model of human low molecular weight protein tyrosine phosphatase bound to sulfate (1xww)

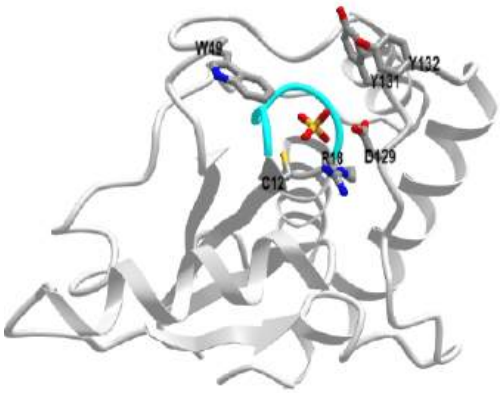


Figure 28.6.14: human low molecular weight protein tyrosine phosphatase bound to sulfate (1xww). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...jaW2syCWWeeLD8

The active site is deep as shown in Figure 28.6.15

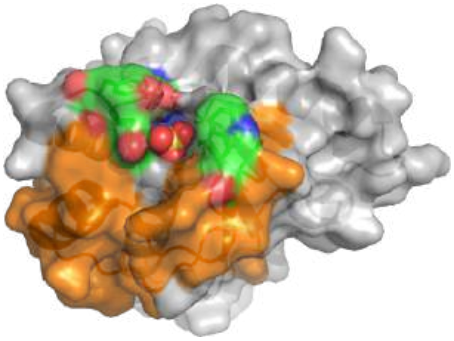


Figure 28.6.15: Active site pocket and surround surface of low molecular weight protein tyrosine phosphatase (1xww)

The sulfate, a mimetic for the phosphate on the p-Try protein target, is deeply buried. The CPK-colored surface (green red, blue) around the sulfate are the side chains of Tyrosine 131 and 132 as well as Trp 49. Y131 and Y132 are part of a loop containing Asp 129, the general acid. This loop is analogous to the WPD loop. The three aromatic amino acids on top of the pocket make it deep enough that pSer and pThr side chains can't reach the active site nucleophile Cys 12. Trp 49 is also in a variable loop (34 amino acids, shown as an orange surface) that differentiates two of the major isozymes, the A and B forms, and contributes to substrate binding specificity.

28.6.3.2: Tyrosine-protein phosphatase non-receptor type 1, also known as PTPN1 or PTP1B PTP1B

PTP-1B regulates the endoplasmic reticulum unfolded protein response and is involved in insulin JAK/STAT and HER2 (ErbB2) signaling. It has a full-length form (MW 50,000) and a C-terminal shortened form (37,000)

Given the difficulty in targeting the active site which is common in all PTPs, efforts have concentrated on the development of allosteric inhibitors that bind to exosites removed from the active site. An example is trodusquemine (MSI-1436) used in the treatment of obesity and type 2 diabetes. It binds much more tightly to the full-length form.

Figure 28.6.16 shows an interactive iCn3D model of the human Protein Tyrosine Phosphatase 1B (1-301) in complex with the inhibitor OTA (5K9W).

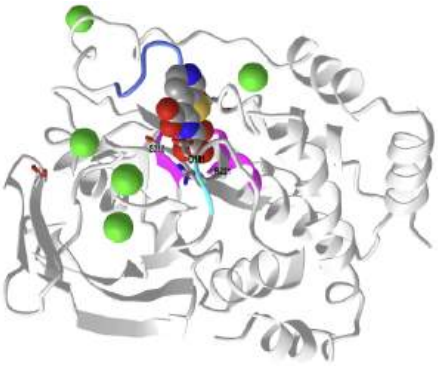


Figure 28.6.16: Human Protein Tyrosine Phosphatase 1B (1-301) in complex with the inhibitor OTA (5K9W). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...fTurfTTFiW2b47

The P-loop is shown in magenta, the WPD loop in cyan, and the substrate binding loop (SBL), which allows entry of pTyr but not pSer and pThr, in blue. The key side chain in the P-loop (Cys 215 and Arg 221) as well as the catalytic general acid (Asp 181) are shown in sticks and labeled. The inhibitor is shown in spacefill, and CPK colors. The movement of the WPD loop is rate-limiting for the hydrolysis of P-Tyr esters. On binding, the WPD starts to close, and in the process Arg 221 moves to form salt bridges with the phosphate. Full closure of the WPD follows, which positions Asp 181 for general acid catalysis. Key interactions of PTP-1B phosphoprotein in insulin and cytokine signaling, are shown in Figure 28.6.17:

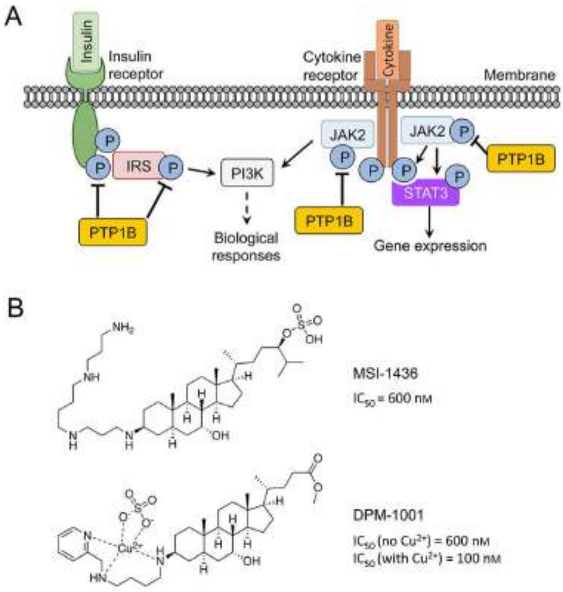


Figure 28.6.17: Protein tyrosine phosphatase 1B (PTP1B) and its effects on signaling. Maja Köhn *ACS Cent. Sci.* 2020, 6, 4, 467–477. Publication Date: March 13, 2020. https://doi.org/10.1021/acscentsci.9b00909. This is an open-access article published under an ACS Author Choice License, which permits copying and redistribution of the article or any adaptations for non-commercial purposes.

Panel (A, left) shows how PTP1B dephosphorylates the insulin receptor and the insulin receptor substrate (IRS), which we have explored in a previous section. Panel (A, right) show its activity in the JAK/STAT pathway, which we have all seen previously. One cytokine receptor that it regulates is the leptin receptor. The hormone leptin, released from fat cells (adipocytes) is a key regulator of lipid metabolism. Pane B shows the structures of key inhibitors of PTP-1B.

28.6.3.3: Protein tyrosine phosphatase nonreceptor type 11 (*ptpn11*) also known as SHP2 (SH2-domain containing phosphatase-2)

This is an example of another phosphatase in which a mutation leads to cancer. It is downstream and activated by most receptor tyrosine kinases (RTKs) involved in the activation of the MAPK pathway with its ultimate links into the nucleus and activation of gene transcription.

Figure 28.6.18shows an interactive iCn3D model of Non-receptor Protein Tyrosine Phosphatase SHP2 in Complex with Allosteric Inhibitor Pyrazolo-pyrimidinone 5 (6MDB)

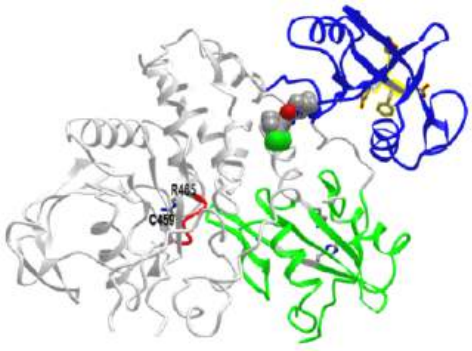


Figure 28.6.18: Non-receptor Protein Tyrosine Phosphatase SHP2 in Complex with Allosteric Inhibitor Pyrazolo-pyrimidinone 5 (6MDB). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?5WCzu8HeY7JrEVyw5

The phosphatase domain is shown in gray. The N- and C-terminal SH2 domains are shown in green and blue, respectively. The allosteric inhibitor is shown in spacefill and CPK colors. The P-loop in the catalytic domain is shown in red with the Cys 459 (active site nucleophile) and R465 (stabilizer of phosphate in the complex) shown in sticks, CPK colors, and labeled. The bound

inhibitor is especially interesting as it binds at an allosteric site. As mentioned above, it is very difficult to design specific inhibitors that target just one PTP given their common active sites and mechanisms. Figure 28.6.19shows multiple features of SHP2.

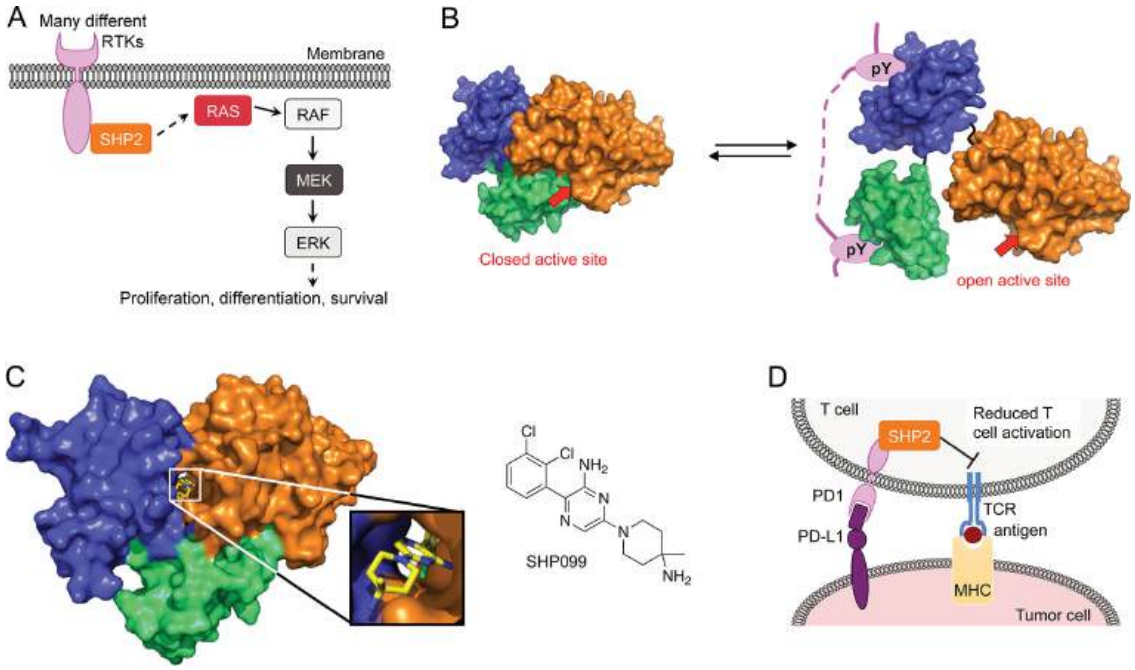


Figure 28.6.19. SH2-domain containing phosphatase-2 SHP2. Köhn ibid.

Panel (A) shows how SHP2 recruited to phosphorylated RTKs activates the MAPK pathway. The dotted line indicates multiple steps. Upon receptor activation, SHP2 is recruited in different ways to activate the MAPK pathway.

Panel (B) shows that in the inactive state, the N-terminal SH2 domain (green) blocks access to the active site. When the N- and C-terminal domains bind pY residues in a single pY-protein, two pY-proteins, or pY residues on its C-terminal tail, a conformational change ensues opening the active site (5EHR).

Panel (C) shows how another allosteric inhibitor keeps the protein in a closed state (5EHR).

Panel (D) shows how SHP2 can decrease T-cell responses through the MHC:Tcell receptor (TCR) complex. Tumor cells express a ligand called PD-L1, which binds to the PD1 receptor on the T cell surface. After binding, SHP2 is recruited to PD1, decreasing T cell activation. This is not a good thing since it inhibits the immune response to the cancer cell.

28.6.4: Dual Specificity Phosphatases (DUSPs)

Another important phosphatase is **phosphatidylinositol 3,4,5-trisphosphate 3-phosphatase and dual-specificity protein phosphatase PTEN (**phosphatase and tensin homolog**).** Dual-specificity protein phosphatase hydrolyzes pTyr- as well as pSer- and pThr- phosphoesters in target proteins. They don't require divalent metal cations and are closer in structure to protein tyrosine phosphatases. They have an active site cysteine in a P-loop also containing arginine. In addition, they are lipid phosphatases, removing phosphate from the inositol ring from phosphatidyl inositol derivatives. These both impact many signaling pathways. Its activity as a lipid phosphatase makes it a tumor suppressor protein as it inhibits the PI3K-AKT/PKB signaling pathway by dephosphorylating phosphoinositides. Hence it modulates both AKT and mTor pathways.

The domain structure of PTEN is shown in Figure 28.6.20

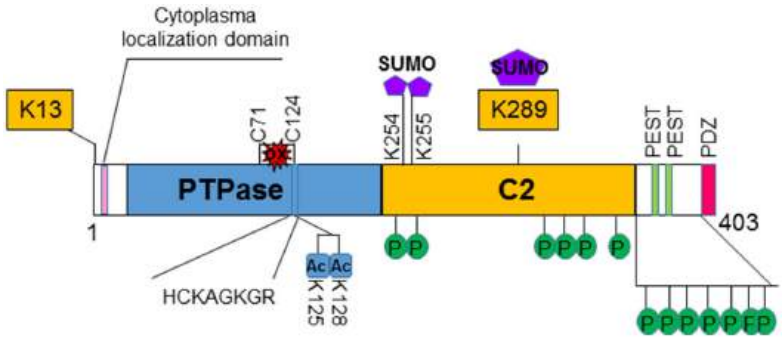


Figure \(\PageIndex{20}\: Structure and regulation of PTEN Chen et al. Endocrinol., 09 July 2018 | https://doi.org/10.3389/fendo.2018.00338. **Creative Commons Attribution License (CC BY)**

The C2 domain enables phospholipid binding. Multiple post-translational modification sites are indicated. The PEST motif is sequence rich in proline (P), glutamic acid (E), serine (S), and threonine (T) and bounded by positively charged amino acids (Lys, Arg, or His) that act as signals for protein degradation. The PDZ domain, often found at the C-terminal of signaling proteins, acts as a scaffolding site for interaction with other signaling proteins. In the next chapter section, we will consider redox signaling, for which PTEN is a great example. Disulfide formation (in a more oxidizing environment) between the nucleophilic Cys 124 and a nearby Cys 71 (figure above) inhibits PTEN phosphatase activity.

Figure 28.6.21 shows an interactive iCn3D model of an AlphaFold computational model of full-length human PTEN (Uniprot P60484).

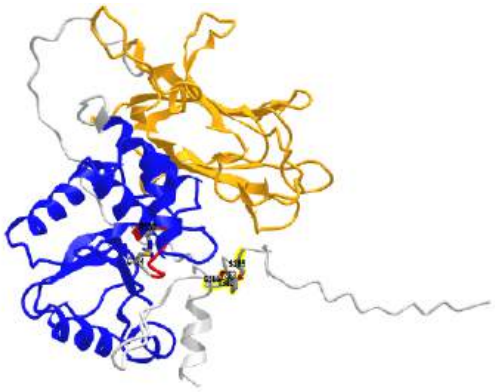


Figure 28.6.21: AlphaFold computational model of full-length human PTEN (Uniprot P60484). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...1nbf3sDfhvMZ19

The phosphatase (PTPase) domain is shown in blue and the C2 domain is in orange. The P-loop is in red with the active site Cys 124 and R130 in colored sticks and labeled. The backbone of the highly extended intrinsically disordered C-terminus region is shown in gray. It contains the clustered residues Ser 380, Thr 382, Thr 383, and Ser 385 (shown in colored sticks and labeled) that are sites for phosphorylation by activated kinases.

iFigure 28.6.22shows key molecules dephosphorylated by PTEN, including the lipid PIP3, and Thr 308 and Ser 473 on AKT.

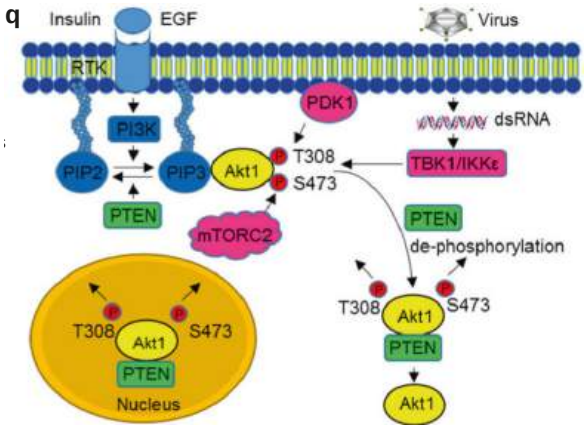


Figure 28.6.22: A model of how PTEN suppresses the activity of Akt. Bu, L., Wang, H., Pan, Ja. *et al*. PTEN suppresses tumorigenesis by directly dephosphorylating Akt. *Sig Transduct Target Ther* **6,** 262 (2021). https://doi.org/10.1038/s41392-021-00571-x. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

When PTEN dephosphorylates Akt1, it inhibits AKT activity and effectively antagonizes the main PTEN-PIP3-PDK1-Akt pathway. PTEN is considered a tumor suppressor for this reason. Mutations in PTEN hence are associated with cancer.

28.7: Calcium Signaling

The following is adapted directly and modified from Sharma et al. Biomedicines **2021**, *9*(9), 1077; **https://doi.org/10.3390/biomedicines9091077.** Creative Commons Attribution (CC BY) license (**https://creativecommons.org/licenses/by/4.0/**).

28.7.1: **Introduction**

Ca^{2+} is central to numerous cellular processes and functions. Its chemical features, including a low hydration energy, high polarizability, relative flexibility of coordination sites and bond length, and large concentration gradient across cellular membranes (100 nM intracellular to 2 mM extracellular) due to low intracellular levels make it the ion of choice at the core of cellular signaling in prokaryotes and eukaryotes alike.

In studying calcium ion signaling we will focus on four key areas:

- Buffering of intracellular Ca^{2+} ion concentrations: Basal low levels must be maintained, which allows transient increases to act as signals. Ca^{2+} ions hence are no different from other second messengers like cAMP, for example. What matters is the rise from a basal level to a threshold concentration level that allows binding to signaling proteins and subsequent signal transmission.
- Storage of intracellular Ca^{2+}: Calcium ions are stored in organelles such as ER, mitochondria, and lysosome. The ions must be released in the presence of specific signals, then returned to the storage organelle to maintain basal Ca^{2+} levels.
- Signaling pathways activated by Ca^{2+} ion: We have seen many pathways simulated by a rise in second messengers and by phosphorylation of lipid and protein molecules in interconnected pathways. We will return to several pathways we have previously studied to see how they integrate with Ca^{2+} in signaling processes.
- Ca^{2+} binding proteins and their binding partners in signaling pathways: We will focus on one key Ca^{2+} binding protein, **calmodulin** (**CAM**), and the kinase that it activates, **Ca^{2+}/CAM protein kinases** or **CAMKs**.

The next two sections are adapted and modified from Sharma et al. Biomedicines **2021**, *9*(9), 1077; **https://doi.org/10.3390/biomedicines9091077.** Creative Commons Attribution (CC BY) license (**https://creativecommons.org/licenses/by/4.0/**).

28.7.2: Buffering of intracellular Ca^{2+} on concentrations

Calcium ions, like the hydronium ion, must be buffered in cells otherwise its potential as a signaling agent would be compromised. The mechanisms adopted by cells for intracellular Ca^{2+} buffering involve sequestration by special proteins as shown in Figure 28.7.1.

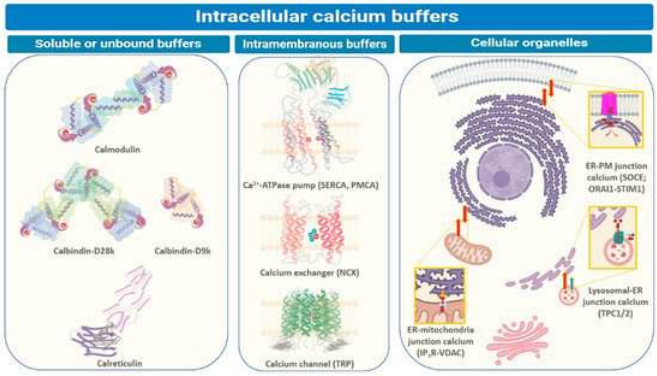


Figure 28.7.1: Types of intracellular Ca^{2+} buffers.

Intracellular Ca^{2+} levels are managed through binding to special proteins or sequestration within different cellular compartments. The three main ways by which intracellular Ca^{2+} is buffered are depicted in Figure 1. These include soluble or unbound proteins that are found in the cytosol or inside organelles, membrane proteins (generally Ca^{2+} channels, ATP-driven pumps (SERCA or PMCA), and ion exchangers (NCX), inside organelles like endoplasmic reticulum (ER), mitochondria, acidic vesicles (mainly lysosomes and Golgi bodies) or organelle junctions (endoplasmic reticulum-plasma membrane (ER-PM), endoplasmic reticulum-mitochondria, or endoplasmic reticulum-lysosomes). The major players regulating inter-organellar Ca^{2+} transfer are IP_3R (inositol-3,4,5-triphosphate receptor), NCX (sodium-Ca^{2+} exchanger), ORAI1/CRACM1 (Ca^{2+} release activated modulator 1), PMCA,

(plasma-membrane Ca^{2+} ATPase), SERCA (sarco-endoplasmic reticulum Ca^{2+} ATPase), STIM1 (stromal interaction molecule 1), SOCE (store-operated Ca^{2+} entry), TPC1/2 (two-pore channel), TRP (transient receptor potential) and VDAC (voltage-dependent anion channel). We will discuss some below.

These proteins are involved in sequestering cytosolic Ca^{2+} upon sensing an increase in its levels and participate in relaying the associated cellular messages. Other proteins that work as intracellular Ca^{2+} buffers exist in the lipid bilayers, plasma membranes, or organelle membranes, like pumps or transporters. Apart from these proteins, intracellular Ca^{2+} is regulated by inter-organellar transport and the influx of Ca^{2+} ions from extracellular space.

28.7.3: Storage of intracellular Ca^{2+} - Proteins

Soluble and Unbound Intracellular Proteins: Calmodulin, Calbindin, and Calretinin

Nonmembrane-associated proteins inside a cell can act as both Ca^{2+} sensors and buffers. Most of these proteins have EF-hand motif(s) that allows Ca^{2+} ions to bind and trigger changes in protein folding, influencing downstream or linked cellular pathways. **Calmodulin (CaM)** is one of the best-studied and ubiquitously expressed Ca^{2+}-sensing proteins known to play a key role in intracellular Ca^{2+} homeostasis. It is a prototype for intracellular Ca^{2+} sensors. It has a 148 amino acid structure with two Ca^{2+}-binding sites in two separate lobes, each with two EF-hand motifs. The lobes are connected in the holo (Ca^{2+}-bound form that binds other proteins. N- and C-termini alpha-helices with a Ca^{2+} coordination loop in between providing affinity for Ca^{2+} ion docking and sequestration. The ability of CaM to transmit a change in free intracellular Ca^{2+} levels into a signal depends on the conformational flexibility of the Ca^{2+}-dependent (apo) form. CaM can exist in a Ca^{2+}-free closed conformational state (Apo-CaM), a semi-open (Ca_2-CaM), or an open state (Holo-CaM or Ca_4-CaM) after Ca^{2+}-binding as shown in Figure 28.7.2

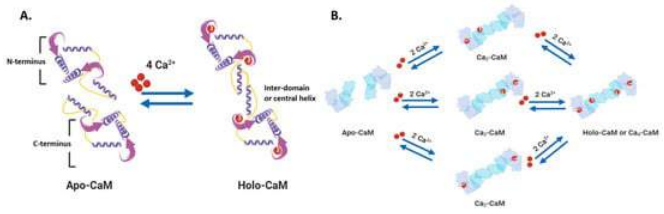


Figure 28.7.2: Overview of calmodulin structure and Ca^{2+} binding conformations.

Figure 28.7.3 shows an interactive iCn3D model of Holo-calmodulin with 4 bound Ca^{2+} ions (1CLL)

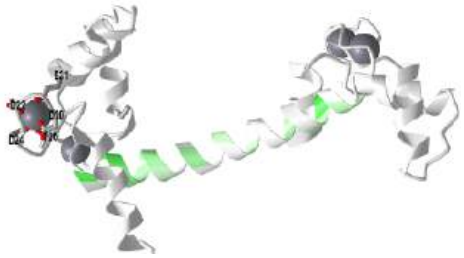


Figure 28.7.3: Holo-calmodulin with 4 bound Ca^{2+} ions (1CLL). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...8nT5HXgNJ4ymc6

Side chains interacting with one Ca^{2+} ion in an EF-hand is shown in the model. The central helix is colored based on hydrophobicity with green indicating more hydrophobic. This amphiphilic helix can bind target proteins in this region through nonpolar interactions.

Figure 28.7.4 shows the change in conformation going from the apo form (without Ca^{2+}) to the holo-form with the fully-formed central helix connecting to the two lobes of the "dumbbell".

Figure 28.7.4: Conversion of apo-CAM to holo-CAM on binding Cal^{2+}.

Differential Ca^{2+} binding to the two lobes of CaM makes fast buffering of a wide range of free intracellular Ca^{2+} possible for this protein. The presence of methionine residues in its lobes and the plasticity of the central linker in its structure also provides CaM with properties to function as an adaptor protein in intracellular Ca^{2+} signaling. CaM can bind to several targets or effector molecules over a variable distance and in multiple orientations to mediate change in intracellular Ca^{2+} signaling. Some major effector proteins that are regulated by CaM binding and are relevant for Ca^{2+} homeostasis include EGFR, PI3K, and connexins.

CaM is required for spatial and temporal regulation of $[Ca^{2+}]$as evident by its role in modulation (activation or inactivation) of Ca^{2+} pumps (such as PMCA and SERCA) and Ca^{2+} channels (such as CaV1.3, TRPV5 and 6, ORAI). CaM also acts via serine/threonine kinases known as Calmodulin-activated Kinases (CaMKs) to influence cellular processes like proliferation (for example, centrosome duplication at G1/S or anaphase to metaphase transition via CaMKII). We will discuss those in detail below.

Integral Membrane protein molecular buffers: SERCA, PMCA, NCX, and TRP

Integral membrane protein Ca^{2+} buffers primarily translocate free Ca^{2+} between domains and organelles. These mainly comprise ion exchangers, channels, and ATP-driven pumps. SERCA, Sarcoendoplasmic Reticulum Ca^{2+} ATPase, is an ATP-dependent ion pump known to significantly maintain free cytosolic Ca^{2+} concentration via actively pumping the ion into the endoplasmic reticulum (or sarcoplasmic reticulum in muscle cell). They share a general structure that includes 10-pass transmembrane helices and three cytoplasmic domain lobes as shown in Figure 28.7.5

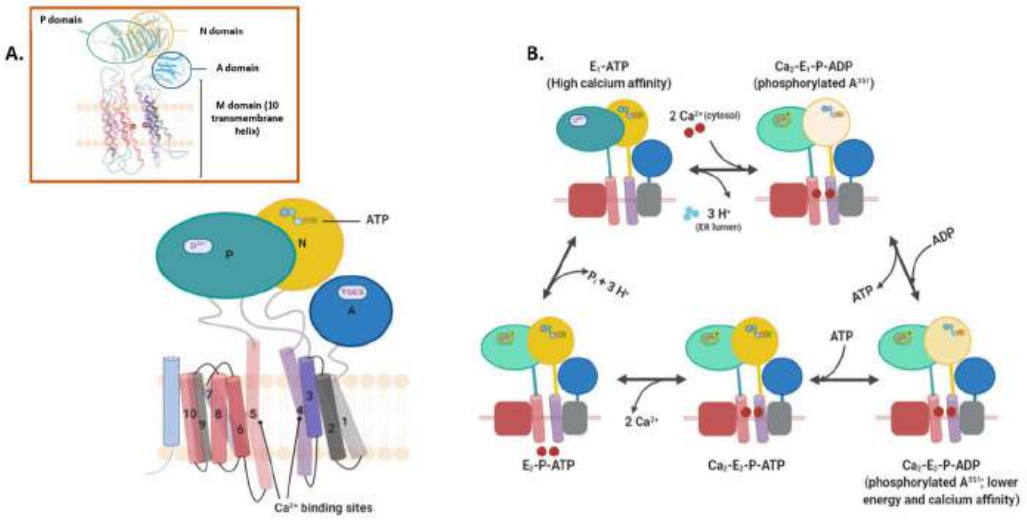


Figure 28.7.5: Sarcoplasmic endoplasmic Ca^{2+}-ATPase structure (Panel A) and mechanism of Ca^{2+} ion transport (Panel B).

P-type Ca^{2+}-ATPases also exist within the plasma membrane and maintain cytosolic Ca^{2+} levels by transferring them into the extracellular space. The Plasma Membrane Ca^{2+} ATPases (PMCAs) can PMCAs transport one Ca^{2+} ion per ATP molecule which differs from the two Ca^{2+} ions per ATP molecule stoichiometry of SERCA. The general structure of such Ca^{2+} transporters comprises 10 transmembrane segments with large cytosolic loops TM 1–2 and TM 3–4, a cytosolic N- and C-termini tails are shown in Figure 28.7.6, panel A below.

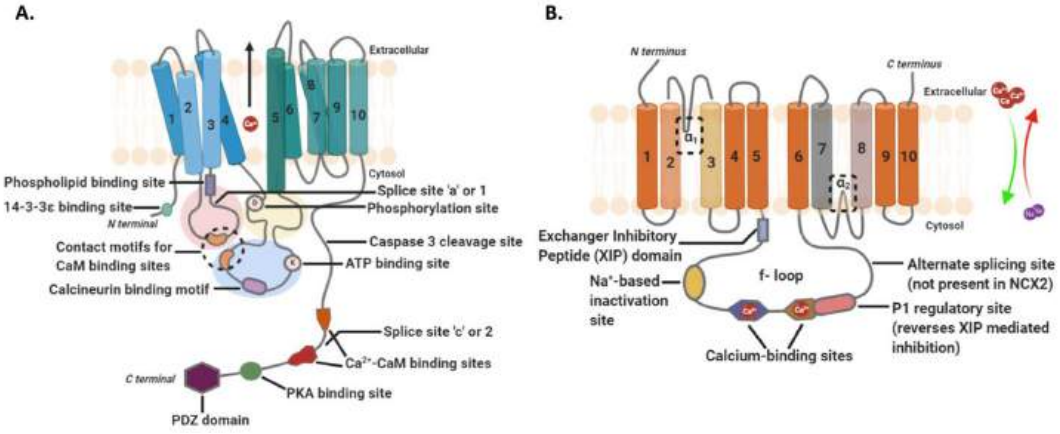


Figure 28.7.6: The general structure of such Ca^{2+} transporters. (A) is the Plasma membrane Ca^{2+}-ATPase (PMCA) and (B) shows the Na^{+}–Ca^{2+} exchanger (NCX)

The cytoplasmic region of PMCA (left) has three loop structures with binding sites for signaling molecules like CAM. They also have phosphorylation sites for additional regulation. The C-terminal tail contains additional CAM sites as well as a PKA site. It has a PDZ domain that can anchor the protein to cytoskeletal components. Differential RNA processing lead to variations in the amino acid sequence in this region and hence binding specificity. The binding of CaM reverses auto-inhibition of the pump due to conformational shifts which displace C-tail from cytosolic loops. Other means of autoinhibition reversal include phosphorylation of C-tail (Ser/Thr residues) by protein kinase A or C, proteolytic cleavage of C-tail, or dimerization via the C-terminus.

Transient Receptor Potential (TRP) channels have a similar function in neurons, epithelial and immune cells. The Mammalian TRP channel superfamily is composed of 28 family members belonging to six subfamilies—TRPC (Canonical), TRPA (Ankyrin), TRPM (Melastatin), TRPV (Vanilloid), TRPP (Polycystin), and TRPML (Mucopilin)—that differ in their sensitivity to various sensory stimulations and affinity for cations (including Ca^{2+} ions) sequestration. Commonly, TRP family members share a structure with six transmembrane domains, intracellular N- and C-termini, and a pore-forming TM 5–6 loop. The cytoplasmic C-terminus of each subunit is a site for protein interaction and post-translational modification. The C-tail of these channels can have PDZ protein binding domains (TRPV and C), sites for interaction with G-proteins (Gq/11)/calmodulin/PLCβ, ADP-ribose binding (NUDIX; TRPM2), or PLC-interacting kinase (PLIK; TRPM6 and 7) domain.

TRP channels act as activators, integrators, as well as downstream effectors of Ca^{2+} signaling at the plasma membrane and in intracellular compartments. Many members of the TRPC subfamily are activated by DAG (diacylglycerol) which is produced by PLC β- or γ-mediated cleavage of phosphatidylinositol 4,5-bisphosphate (PIP_2) after the ligand binding at GPCRs or RTKs. TRPP1/2, TRPA1, TRPM8, and TRPV1-4 are all expressed on the ER membrane. At this site, PLC-independent activation of the TRP channels (such as TRPV1) is suggested to induce ER Ca^{2+} release via inositol triphosphate receptor (IP3R) which further triggers bulk entry of extracellular Ca^{2+} into the cell. On the flip side, cytosolic Ca^{2+} regulates the activity of TRP channels in response to physiological stimuli. This regulatory effect is usually through CaM binding (inhibition of TRPV5, TRPV6, and sensitization of TRPV3) and indirectly through CaM-binding kinase II (CaMKII). These activities are shown in Figure 28.7.7.

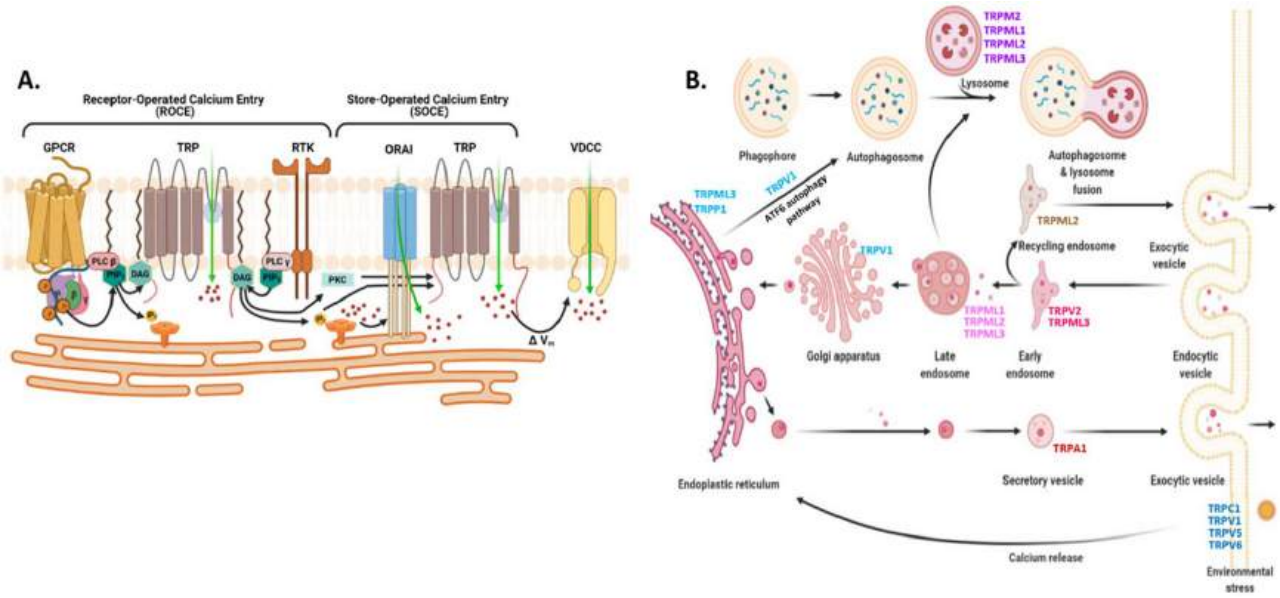


Figure 28.7.7: TRP Channels in Signaling. DAG, Diacylglycerol; GPCRs, G-protein coupled receptors; IP_3, inositol 1,4,5-triphosphate; ORAI (Ca^{2+} release activated Ca^{2+} modulator; RTKs, Receptor-tyrosine kinases; ROCE, Receptor-Operated Ca^{2+} Entry; SOC, store-operated Ca^{2+}; SOCE, Store-Operated Ca^{2+} Entry; STIM, stromal interaction molecule; PIP_2, phosphatidylinositol 4,5-bisphosphate; PKC, protein kinase C; TRP, transient receptor potential; VDCCs, voltage-dependent Ca^{2+} channels.

28.7.4: Storage of intracellular Ca^{2+} ions - organelles

Endoplasmic Reticulum: STIM, ORAI, IP_3Rs, and TRPC1 in SOCE and SOCIC Ca^{2+} Entry Models

The ER serves as the largest and most dynamic organelle reservoir for intracellular Ca^{2+} and is therefore central to many signaling processes for protein synthesis, folding, and post-translational modifications. In contrast to the cytosol, ER Ca^{2+} ion levels can range from 100 uM to 1 mM based on the cell type. ER, and other intracellular organelles buffer excessive cytosolic Ca^{2+} by both housing Ca^{2+}-binding proteins (example: calreticulin in ER) and via active transport (example: SERCA pumps in ER).

- Depletion of Ca^{2+} from the ER lumen actuates an indirect mode of Ca^{2+} entry into the organelle which is termed **Store-Operated Ca^{2+} Entry (SOCE)** or **Ca^{2+} Release Activated Ca^{2+} (CRAC)** entry; it is activated when plasma membrane receptors like PLC-coupled GPCRs (but not voltage-gated channels) trigger Ca^{2+} ion release from the organelle.
- Exhaustion of the intraluminal ER Ca^{2+} ion store following such prolonged release is then sensed by **STIM (Stromal Interaction Molecule**) tethered to the ER membrane and subsequently relayed to the CRAC channels on the plasma membrane.

Figure 28.7.8shows the domain structures of STIM 1/2, ORAI 1-3, and IP3Rs (panel A) and the mechanism of Ca^{2+} influx into the cell (B).

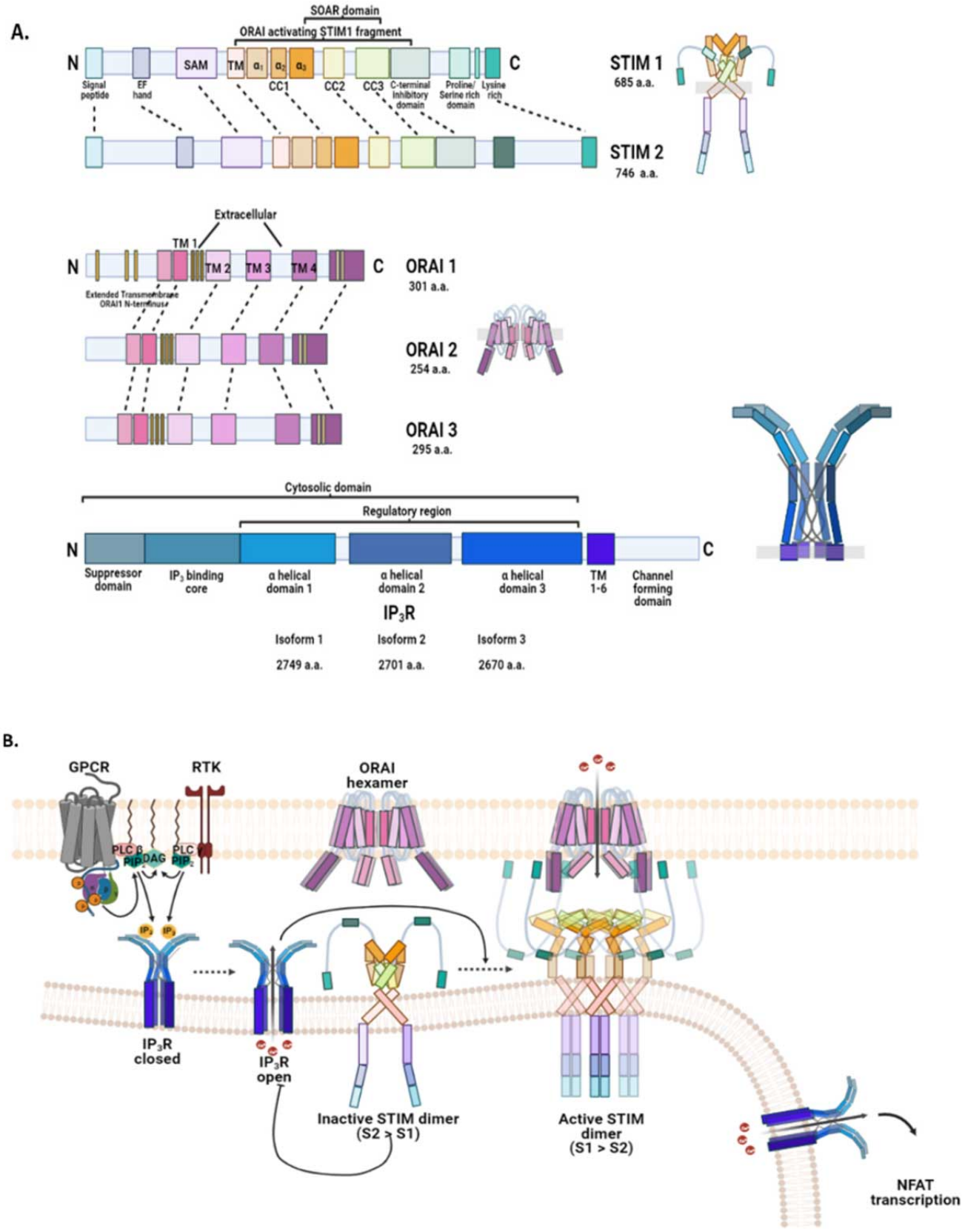


Figure 28.7.8: Role of STIM< ORAI and IP3Rs in Ca^{2+} influx.

IP3 Receptors (IP3Rs), on stimulation by IP3, open and allow Ca^{2+} influx from the organelle lumen into the cytoplasm. After activation of IP3Rs on the ER membrane by ligand IP3 and cytosolic Ca^{2+} from activation of phospholipase C, STIM dimers are activated once the luminal Ca^{2+} concentration drops below basal levels. These receptors provide intracellular Ca^{2+} ions for downstream Ca^{2+} signaling including NFAT-mediated transcription. The red semi-circle in the ER represents high luminal Ca^{2+} levels, the pink semi-circle is for moderately low Ca^{2+} ion concentration, and the pale semi-circle indicates extremely low Ca^{2+} concentration. CC, coiled-coil; NFAT, nuclear factor of activated T-cells; SAM, sterile alpha motif; SOAR, STIM1 Orai1-activating region; TM, transmembrane.

Mitochondria and Acidic Vesicles (Mainly Lysosomes)

Mitochondria also play a critical role in maintaining Ca^{2+} ion levels in the cytosol and endoplasmic reticulum. They are found mostly aggregated around the nucleus and store similar levels of intracellular Ca^{2+} as the cytosol (0.1 μM). Electrochemical proton gradient or membrane potential (Ψmt = −150 to −180 mV) and close association to the ER are the two key factors responsible for Ca^{2+} uptake in mitochondria. The free movement of small molecules (less than 5 kDa) from the outer mitochondrial membrane (OMM) into the inner mitochondrial space and their impermeability across the latter generates a high electrochemical proton gradient for ATP synthesis. This gradient simultaneously draws Ca^{2+} ions from the cytosol.

Transfer of Ca^{2+} ions from ER to mitochondria occurs at specialized microdomains or contact sites known as Mitochondrial Associated Membranes (MAMs). These are characterized by the ER and OMM apposed at 10–25 nm from each other and are strewn with a cluster of channels, transporters, exchangers, and tethering proteins for facilitating Ca^{2+} ion transfer. IP3Rs localized at the ER side of the MAMs release Ca^{2+} ions that gate voltage-dependent anion channels (VDACs) located on the OMM. VDACs (1, 2, and 3) are 30 kDa polypeptides having a 19-strand beta-barrel structure that regulates the flux of metabolites (polyvalent anions like ADP and ATP) across the outer mitochondria membranes. These channels transport cations including Ca^{2+} more readily than anions like chloride. Due to voltage-dependent electrostatic gating, the ion selectivity and flux across VDACs change between open and closed states. Figure 28.7.9 shows couped mitochondrial and lysosomal effects on intracellular Ca^{2+} signaling.

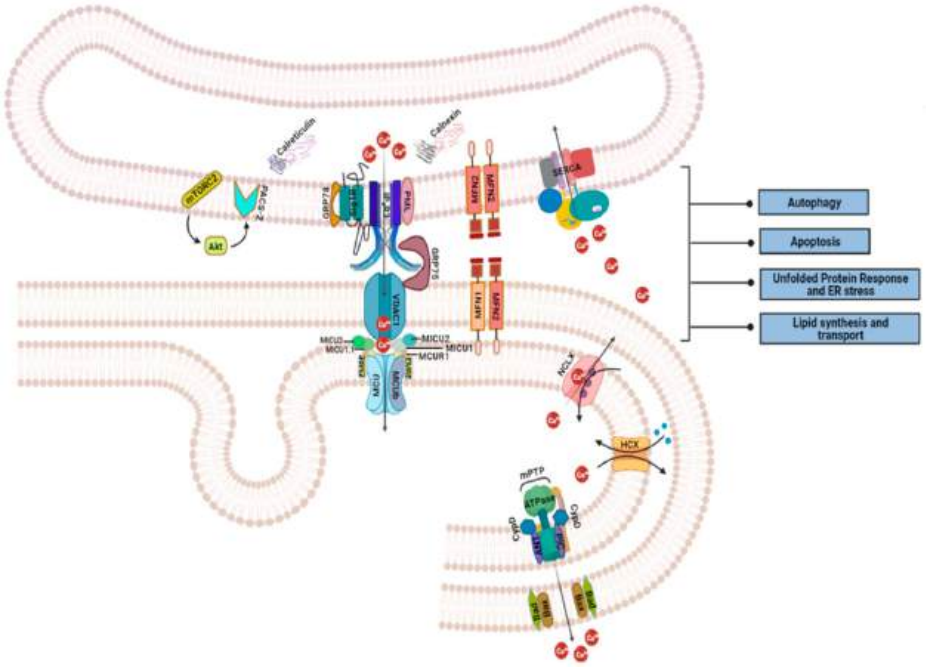


Figure 28.7.9: Mitochondrial and lysosomal impact on intracellular Ca^{2+} signal.

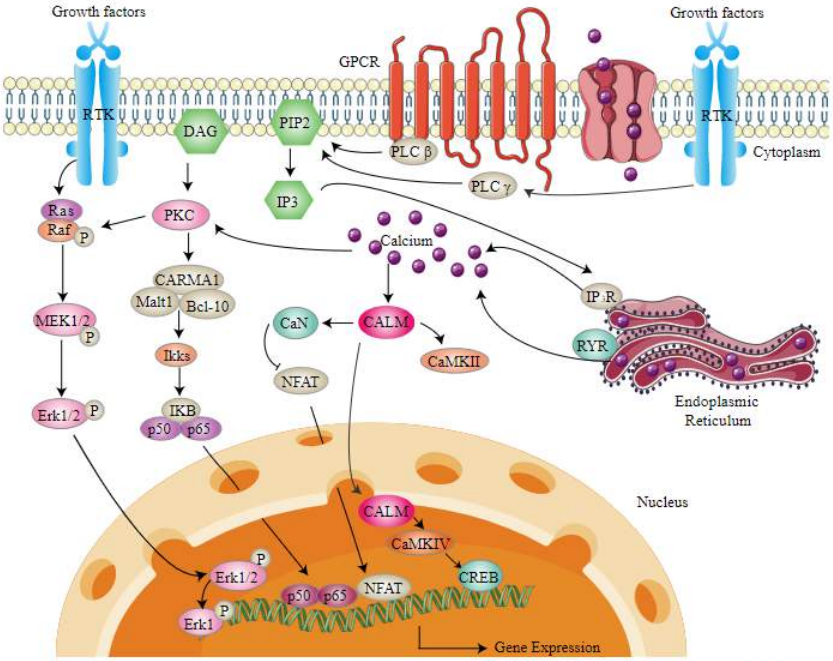
Primary components of Ca^{2+} signaling at the mitochondrial associated membranes (MAMs) include IP_3R3 on the endoplasmic reticulum, VDAC1 on the outer mitochondrial membrane, and MCU complex on the inner mitochondrial membrane [151,154,156,161]. Transport of Ca^{2+} ions from ER to mitochondria plays a crucial role in cellular metabolism (autophagy), cell survival (during unfolded protein response and impinging cell death signals), lipid production, and distribution. F

While IP3 acts as the dominant Ca^{2+}-mobilizing messenger, cADPR (cyclic ADP-ribose) and NAADP (nicotinic acid adenine dinucleotide phosphate) are also known to modulate intracellular Ca^{2+} stores. cADPR evokes Ca^{2+} ion release from ER by acting

on ryanodine receptors (RyR; counterpart of IP3R in myocytes and co-expressed in some other cell types). NAADP releases Ca^{2+} from acidic and/or secretory vesicles such as lysosomes and endosomes. In most mammalian cells, lysosomes comprise ~5 percent of the cell volume and store similar levels of intracellular Ca^{2+} (0.5 mM) as the ER. Due to their relatively smaller size than ER, lysosomes release nearly undetectable amounts of intracellular Ca^{2+} in response to NAADP trigger

28.7.5: Signaling pathways activated by Ca^{2+} ion

It is daunting to both readers and writers to introduce a myriad of new signaling pathways. Instead will show how Ca^{2+} signaling fits into other pathways we have already discussed. A summary showing how Ca^{2+} signaling integrates with other pathways is shown in Figure 28.7.10



Calcium (Ca^{2+}) serves as a ubiquitous second messenger in all eukaryotes. For an ever-increasing number of biological processes, it has been found that temporally and spatially defined changes of Ca^{2+} concentration in the cytoplasm or in defined organelles occur at one point or another. The importance of Ca^{2+} signaling pathway for the implementation of the information provided by Ca^{2+} has been increasingly appreciated, and several distinct families of Ca^{2+} sensing proteins have been identified and characterized.

Figure 28.7.10: https://www.creative-diagnostics.com...ng-pathway.htm

An Overview of Calcium Signaling Pathway

Ca^{2+} signaling, as described above, requires ion buffer, organelle storage, and Ca^{2+} protein pumps and channels. The concentration (amplitude) and frequency of Ca^{2+} release affect signaling. Figure 28.7.10 shows the importance of upstream signaling through GPCRs, phospholipase C, RTKs, and IP3/DAGs. Ca^{2+} also enters the nucleus vs IP3 receptors (IP3Rs) and ryanodine recetors (RYR). An important family of cytoplasmic transcription factors, the **Nuclear factor of activated T-cells** (**NFAT**), which are important in immune responses and development of muscle and nervous systems, are activated in calcium ion signaling.

As we described above, Ca^{2+} release from the ER is sensed by integral ER membrane proteins called STIMs. These bind Ca^{2+} ions as a buffering system, but if most of the ER calcium is depleted, the STIM-bound Ca^{2+} ions are also released. This lead to their self-association and ultimate activation or ORAI1, part of the CRAC complex in the cell membranes which allow extracellular Ca^{2+} ions to flow into the cell in a process called store-operated calcium entry (SOCE). Sufficient calcium now accumulates in the cell to activate the transcription factor NFAT through dephosphoylation by calcineurin (PP2B), also abbreviated CaN in Figure 28.7.10 NFAt then translocates into the nucleus and activates gene transcription. Also, it has been shown that nuclear calcium ions directly can activate the cAMP response element binding protein (CREB), a transcription factor that activates gene transcription. In addition the CAM:CAMKII complex can translocate into the nucleus. Calcium signals also activate ERK1/2-MAPK cascade.

28.7.6: Ca^{2+} binding proteins and their binding partners in signaling pathways

We have already described the key calcium-binding protein, calmodulin. On binding Ca^{2+}, it undergoes a profound conformational change that allows it to interact with a family of key signaling kinases called CAM and Ca^{2+}/CAM-Dependent Protein Kinases (CAMKs).

Ca^{2+}/calmodulin-dependent protein kinase is a key signaling protein activated by Ca^{2+} through the binding of calmodulin to **CAMK**. Activated CAMK is a Ser/Thr kinase. There are many types of CAMKs. We will focus on multifunctional CAMKs that can phosphorylate multiple target proteins. These are important in learning and memory, metabolism, and gene transcription. As with other kinases, they have catalytic and regulatory domains. Some like CAMK II have association domains that allow the formation of CAMKII multimers. In addition, they must have a CAM binding domain. As with all kinases, the CAMKs must be able to switch from an inactive to an active form.

CAMKI has a catalytic, substrate-binding domain and an autoinhibitory domain that blocks the active site. CAMKI is activated by an upstream kinase CAMKK (a naming system similar to the MAPK cascade) on the binding of Ca^{2+} to CAM. It helps regulate transcription, the cell cycle, hormone production, cell differentiation, actin filament organization, and neurite outgrowth. It is found in the cytoplasm and nucleus.

We will focus our attention on **CAMKII**, which has four isoforms (α, β, γ, and δ). It is activated by the binding of Ca^{2+}/CAM which promotes autophosphorylation. After that, it is active in the absence of CAM. It is important in learning and memory and synapse formation in neurons and the regulation of sarcoplasmic reticulum Ca^{2+} transport in skeletal muscles.

They have an N-terminal catalytic domain and a C-terminal association domain that facilitates multimer formation into large holoenzymes with 12 or 14 CAMK monomeric subunits (a homomer or heteromer). These two domains are separated by a linker/regulatory domain that has a CAM binding site, an autoinhibitory region, and key Ser and Thr side chains that are targets for phosphorylation. Figure 28.7.11 shows the domain structure of the CAMKII monomer (a), the overall structure of a homododecamer (b), and the mechanism for activation of kinase activity (c).

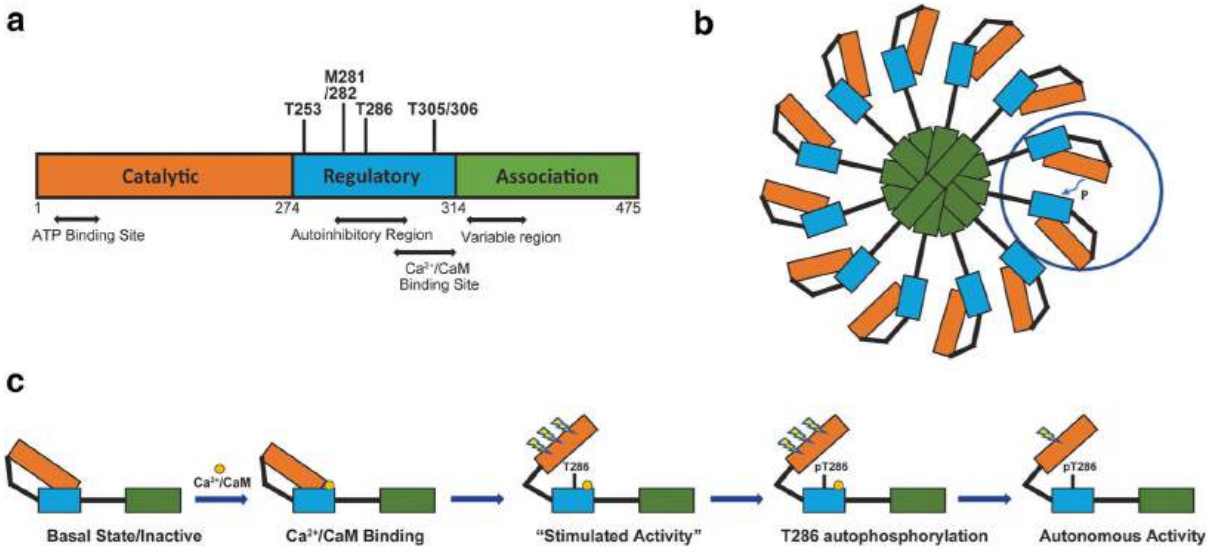


Figure 28.7.11: CamKII - structure and activation. Zhang, X., Connelly, J., Levitan, E.S. *et al.* Calcium/Calmodulin–Dependent Protein Kinase II in Cerebrovascular Diseases. *Transl. Stroke Res.* **12,** 513–529 (2021). https://doi.org/10.1007/s12975-021-00901-9. Creative Commons Attribution 4.0 International License, http://creativecommons.org/licenses/by/4.0/.

T253, T286, and T305/306 are targets of autophosphorylation. M281/282 are also sites for oxidative modification. The C-terminal association domain allows multimer formation. It has a variable region that differentiates CaMKII subtypes. Panel (b) shows the multimer that forms on interactions of multiple association domains on different CAMKIIs. Panel (c) shows binding of CAM promotes the phosphorylation of key residues including T286 (through autophosphorylation).

In the absence of Ca^{2+}/CAM T286 amino acid forms interactions with the catalytic domain to maintain the inactive conformation. The regulatory domain effectively autoinhibits the kinase domain. On the formation of the Ca^{2+}/CAM/CAMKII complex, a conformation change ensues that frees the catalytic domain from autoinhibition and exposes the active site. Each subunit in the dodecamer is activated separately. T286 is now free to be "autophosphorylated" by an adjacent active subunit. Once

phosphorylated, pT286 prevents the rebinding of the autoinhibitory region to the catalytic domain, even when CAM dissociates. At this point, CAMKII is active in the absence of Ca^{2+}/CAM.

Phosphorylation of T286 also regulates its binding to target proteins for their phosphorylation. In addition, CAMKII can autophosphorylate T254 and T306 with further effects on activity. T306 is only autophosphorylated after CAM dissociates and the enzyme is autonomously active. Dephosphorylation by PP1 and PP2A returns the enzyme to an inactive state.

Figure 28.7.12 shows an interactive iCn3D model of a single subunit of human Ca2+ Calmodulin- Dependent Kinase II Holoenzyme (3SOA)

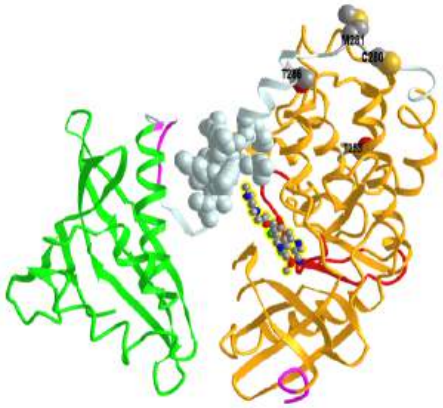


Figure 28.7.12: Single subunit of human Ca2+ Calmodulin- Dependent Kinase II Holoenzyme (3SOA). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...S4XAczeqqPE5Y9

The catalytic domain is shown in orange, the association domain in green, and the linker domain in light cyan. The spacefill light cyan site in the linker domain is the CAM binding site. The side chains of Thr 253 and Thr 286, sites for phosphorylation, are shown in spacefill CPK colors and labeled. The side chains of Cys 280 and Met 281, the site for redox regulation, are shown in spacefill CPK colors and labeled. Bosutinib (shown in ball and stick) is a small molecule BCR-ABL and src tyrosine kinase inhibitor used for the treatment of chronic myelogenous leukemia. It is bound in the ATP binding site of the catalytic subunit. Finally, the activation loop is shown in red.

Figure 28.7.13shows an interactive iCn3D model of Human Ca2+ Calmodulin- Dependent Kinase II Holoenzyme (3SOA)

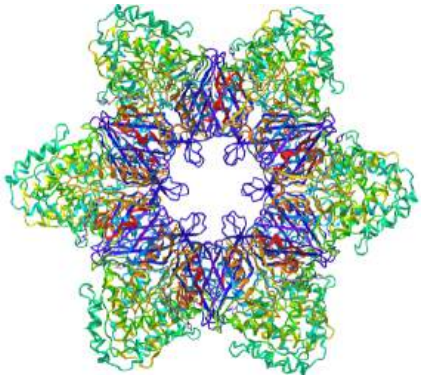


Figure 28.7.13: Human Ca2+ Calmodulin- Dependent Kinase II Holoenzyme (3SOA). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...DPiWLo5VrgiLp9

The response of most protein kinases we have studied depends on the concentration (amplitude) of a binding ligand (like Ca^{2+}). The CAMKII dodecamer also responds to the frequency of Ca^{2+} waves or spikes as it is released from intracellular organelles. This is important in neurons where CAMKII phosphorylates ion channels that control and regulate neuron response. If the frequency of Ca^{2+} spikes reaches a certain threshold level, the enzyme no longer depends on Ca^{2+}.

28.7.7: Modeling of Ca^{2+} signaling

Model for signal-induced Ca^{2+} oscillations and their frequency encoding through protein phosphorylation.

Reference paper: ALBERT GOLDBETER, GENEVIVE DUPONT, AND MICHAEL J. BERRIDGE , Proc. Natl. Acad. Sci. USA, Vol. 87, pp. 1461-1465, February 1990, Biophysics

HVJ: In response to hormones and neurotransmitters, cyclic intracellular oscillations/spikes (as a function of time) and spatial waves of cytoplasmic Ca^{2+} arise. The Ca^{2+} ions derive from the opening of cell membrane channels but also, more importantly, from the release and recapture of the ions from intracellular compartments such as the endoplasmic reticulum (ER). Signaling through membrane GPCRs can lead to the activation of phospholipase C and the formation of inositol 1,4,5-trisphosphate (IP3 or $InsP_3$) as a key second messenger. IP3 can interact with the IP3 receptor on ER membranes, leading to the release of intracellular stores of Ca^{2+} into the cytoplasm in ways that leads to oscillations in its concentration. These processes are shown in Figure 28.7.14below.

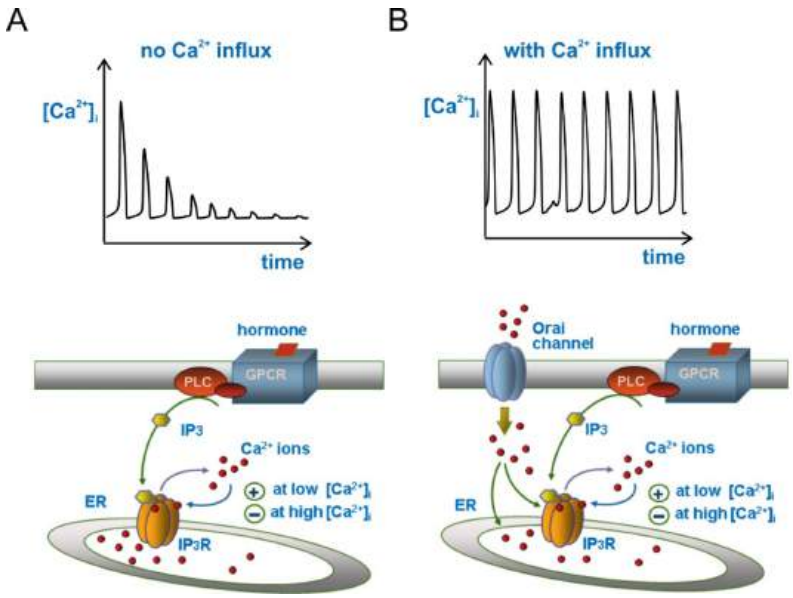


Figure 28.7.14 Ca^{2+} oscillations in response to inositol trisphosphate (IP_3) increase, with and without Ca^{2+} influx from extracellular space. Catacuzzeno L, Franciolini F. Role of KCa3.1 Channels in Modulating Ca^{2+} Oscillations during Glioblastoma Cell Migration and Invasion. Int J Mol Sci. 2018 Sep 29;19(10):2970. doi: 10.3390/ijms19102970. PMID: 30274242; PMCID: PMC6213908. Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

Panel (**A**) bottom shows a drawing illustrating the hormone-based production of IP_3 that activates the IP_3 receptor to release Ca^{2+} from the endoplasmic reticulum (ER). The biphasic effects of cytosolic Ca^{2+} on IP_3 receptor gating (the basic mechanism for Ca^{2+} oscillations), whereby Ca^{2+} modulates positively the receptor at low [Ca^{2+}] but negatively at high [Ca^{2+}], is also illustrated. Top, Ca^{2+} oscillations as produced from the schematics below. Note the decaying trend of Ca^{2+} spikes due to the absence of Ca^{2+} influx from extracellular space;

Panel (**B**) shows Ca^{2+} influx apparatus from extracellular spaces through ER-depletion by activated Orai channels on the plasma membrane (bottom), which generates sustained Ca^{2+} oscillations (top).

Ca^{2+} levels fall through Ca^{2+} pumps on the sarco/endoplasmic reticulum (SERCA), which moves the ion into the SR/ER or the plasma membrane (PMCA) that then moves it to the outside of the cell. The SERCA and PMCA Ca^{2+} pumps as well as the ER STIM protein are not shown in Figure 14 above. The STIM1 protein is involved in what is known as store-operated Ca^{2+} entry (SOCE). This allows the influx of Ca^{2+} influx after intracellular stores are depleted. The STIM1 protein has an EF-hand so it binds and acts as a sensor for Ca^{2+}. When Ca^{2+} is depleted, the STIM1 protein moves from the ER to the cell membrane and activates ORA1, a subunit of the Ca^{2+} release-activated Ca^{2+} (CRAC) channel to promote Ca^{2+} into the cell.

In a variety of cells, hormonal or neurotransmitter signals elicit a train of intracellular Ca^{2+} spikes. The analysis of a minimal model based on Ca^{2+} -induced Ca^{2+} release from intracellular stores shows how sustained oscillations of cytosolic Ca^{2+} may develop as a result of a rise in inositol 1,4,5-trisphosphate ($InsP_3$) triggered by external stimulation. This rise elicits the release of a certain amount of Ca^{2+} from an $InsP_3$-sensitive intracellular store. The subsequent rise in cytosolic Ca^{2+} in turn triggers the release of Ca^{2+}

from a second store insensitive to $InsP_3$. The model shows how signal-induced Ca^{2+} oscillations might be effectively encoded in terms of their frequency through the phosphorylation of a cellular substrate by a protein kinase activated by cytosolic Ca^{2+}.

The release of intracellular Ca^{2+} release by IP3 can be broken down into 3 steps: β

- agonist binding to GPCRs to activate the Phospholipase C pathway to produce the second messengers IP3 and DAG;
- IP3 induces some Ca^{2+} release from intracellular stores in the SR/ER. This can be called the Primer Step characterized by a rate V1 to produce cytosolic Ca^{2+} of concentration [Z] that prime the next step. This step has a rate of v1β.
- released Ca^{2+}induces more Ca^{2+} release, a self-amplification positive feedback step, which causes the Ca^{2+} spike.

A simplified view of this model is shown in Figure 28.7.15below.

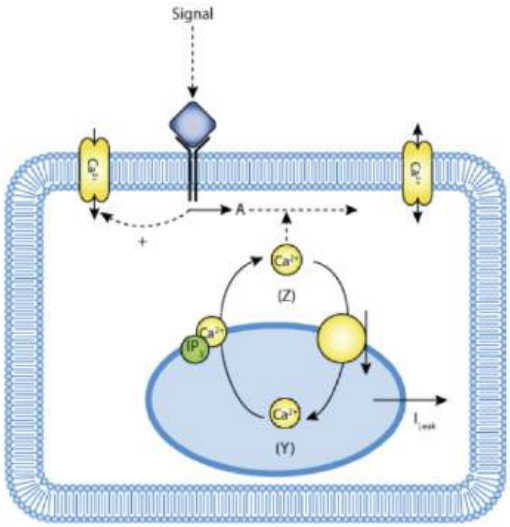


Figure 28.7.15 Schematic representation of the one-pool model based on CICR with Ca2+-stimulated degradation of IP3. Lloyd, C.M., Lawson, J.L., Hunter, P.J. and Nielsen, P.F. The CellML Model Repository. Bioinformatics. 2008 September;24(18):2122-2123 (accessed 4/19/23; 5:40 am EDT). Creative Commons Attribution 3.0 Unported License.

A simple model can be constructed to account for cytosolic Ca^{2+} fluxes. The steps include

- IP3 causes a triggering a constant flux of Ca^{2+} into the cytosol, v1β ,to produce cytosolic [Ca^{2+}] = **Z** (the primer)
- cytosolic Ca^{2+} flows into an IP3-insensitive sequestered pool (concentration **Y**) with rate v2 to keep low levels of cytosolic Ca^{2+}
- Spikes arise when the sequestered pool Y releases Ca^{2+} back into the cytosol at a rate v3 in a process that is activated by cytosolic Ca^{2+}.

Goldbeter at al (Proc. Natl. Acad. Sci. 87, 1461-1465 (1990) constructed a mathematical model to account for the oscillations/spikes In addition to the parameters mentioned above, they included 3 more:

- vo, describing the influx of extracellular Ca^{2+} into the cytosol and k describing the efflux of cytosolic Ca^{2+} from the cell. These are controlled by Ca^{2+} pumps SERCA, PMCA, et al as described above.
- kf, which describes the passive leak of Y into Z.

The Ca^{2+} oscillations then are based on a self-amplified release of Ca^{2+} from intracellular stores.

Here are their equations:

- $dZ/dt = v_0 + v_1 beta - v_2 + v_3 + k_f Y - kZ$
- $dY/dt = v_2 - v_3 + v_3 - k_f Y$
- $v_2 = V_{m2} \frac{Z^n}{K_2^n + Z^n}$
- $v_3 = V_{m3} \frac{Y^m}{K_R^m + Y^m} \cdot \frac{Z^p}{K_A^p + Z^p}$

We'll use Vcell to plot the following species:

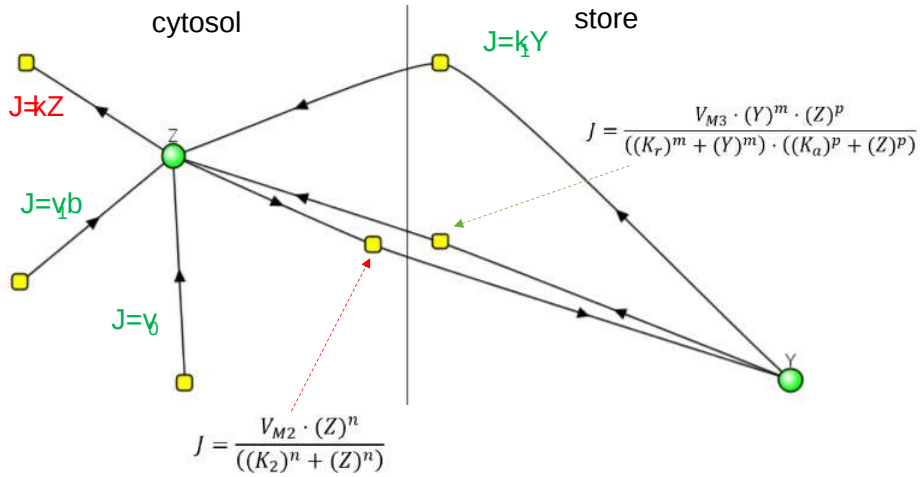
- Species **Z**: Concentration (uM) of cytosolic Ca^{2+}

- Species **Y**: Concentration (uM) of Ca^{2+} stored in the $InsP_3$-insensitive pool.

The model is based on code from EBI-Biomodels: https://www.ebi.ac.uk/biomodels/BIOMD0000000098.

VCell MODEL

Calcium ion oscillation in the cell without oscillations in IP3



Here are the adjustable parameters:

- v0: influx of Ca^{2+} into the cell, uM/s.
- v1: influx of Ca^{2+} into the cell from the $InsP_3$ receptor, uM/s.
- beta: saturation function of the receptor for $InsP_3$, unitless.
- Vm2: maximum rate of Ca^{2+} pumping into the intracellular $InsP_3$-insensitive store, uM/s.
- Vm3: maximum rate of Ca^{2+} released from the intracellular store, uM/s;
- K2: Threshold constant for Ca^{2+} pumping, uM.
- n: Hill function cooperativity coefficient for Ca^{2+} pumping into the store, unitless.
- m: Hill function cooperativity coefficient for Ca^{2+} pumping from the store, unitless.
- Ka: threshold constant for activation, uM.
- kf: efflux of cytosolic Ca^{2+} from the cell, 1/s.
- k: influx of extracellular Ca^{2+} into the cytosol, 1/s.
- p denotes the degree of cooperativity of the activation process, unitless.

Select Load [model name] below

Load Goldbeter_Ca_Oscillations

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real time.

After loading the GoldBeter model, select the '**Start**' button below to simulate the model. Adjust the parameter sliders below the plot to see how they affect Ca^{2+} concentrations (Z, Y). The simulator only displays twelve parameters at a time. To choose others, select the 'Slider' button on the side and chose up to twelve parameters to adjust.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

Adjust the parameter sliders below the plot to see how they affect Ca^{2+} concentrations (Z, Y). The simulator only displays twelve parameters at a time. To choose others, hit the 'Slider' button on the side and chose up to twelve parameters to adjust.

Only two variables, Y and Z, and some intra-connections are required to generate the oscillations, which do NOT required oscillation in the second messenger, IP3. One can imagine that cytosolic Ca^{2+} oscillations might also elicit oscillatory activity of protein kinases activated by it.

Questions:

- Does v0 change the oscillation frequency?
- What other parameters affect the frequency?
- How does Vm2 affect the oscillations?

28.8: Receptor Guanylyl Cyclases, cGMP, and Protein Kinase G

Search Fundamentals of Biochemistry

28.8.1: Introduction

We have considered many signal transduction pathways, starting with an extracellular signal, a primary messenger, that initiates signaling when it binds to a receptor (GPCR, RTK, Cytokine Receptor, etc). It elicits a conformation change in the receptor, which is transmitted to an intracellular domain, from where it can propagate and transmits a signal intracellularly through secondary messengers and phosphorylation of signaling proteins. Some primary messengers, however, actually pass through the cell either passively or through membrane carriers, so there is no need to generate a second messenger. We will consider two signals that translocate through the cell membrane, gases such as nitric oxide (NO), and steroid hormones (which we will discuss in a future section). NO produced within a cell activates the formation of an intracellular second messenger, cyclic GMP (analogous to cAMP). cGMP in turn activates our final member of the AGC Ser/Thr protein kinase family, **protein kinase G (PKG)**. Intracellular NO has an unexpected role on adjacent cells. Given its small size and its nonpolar nature, it can diffuse out of the cell where it is produced and into an adjacent cell, where it can also initiate signaling. Some call this retrograde signaling.

Before we consider cGMP and PKG, let's see how nitric oxide (NO) is produced.

28.8.2: NO formation

NO is synthesized by the enzyme **nitric oxide synthase (NOS)**. There are three isoforms in mammals, neuronal (NOS1 or nNOS), inducible (NOS2 or iNOS), and endothelial (NOS3 or eNO2). Each is a homodimer with a complex domain structure, including a

- N-terminal oxidase (NO_synthase) domain that binds heme
- calmodulin binding site between the N and C terminal domains
- C-terminal reductase domain containing the FMN (Flavodoxin_1) subdomain (which contains an autoinhibitory helix) and FAD/NADPH subdomain.

Ca^{2+} ion activates the enzyme through the binding of Ca^{2+}/CAM. A more detailed description of the domain structures of human NOS is shown in Figure 28.8.1. The dimeric molecular weight of the neuronal NOS1 (321K) is greater than for iNOS (206) and eNOS (266) as it has a N-terminal PDZ domain.

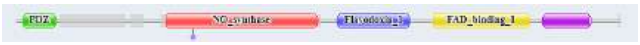
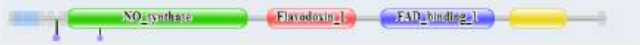
neuronal NOS 1 (MW 321)	 purple NADPH binding
inducible (MW 206K) endothelial (MW 266)	 yellow NADPH binding

Figure 28.8.1: Domain structure of human nitric oxide synthases

NOS catalyzes the conversion of the free amino acid arginine to citrulline and NO, as shown in the chemical equation in Figure 28.8.2

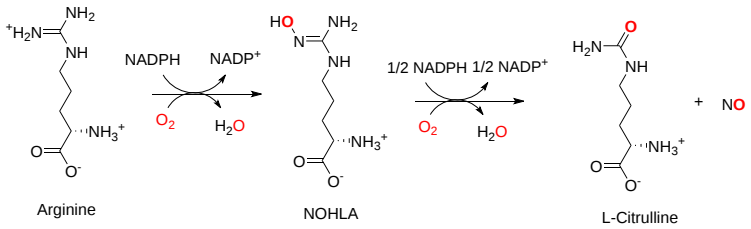


Figure 28.8.2: Synthesis of NO by nitric oxide synthase

The structures of the three enzymes with or without bound CAM are similar. Linkers between the domains and subdomains allow flexibility. Figure 28.8.3 shows the flow of electrons from NADPH into the reductase (NADPH/FAD subdomain to the FMN

subdomain), and on to the NOS synthase domain containing the heme.

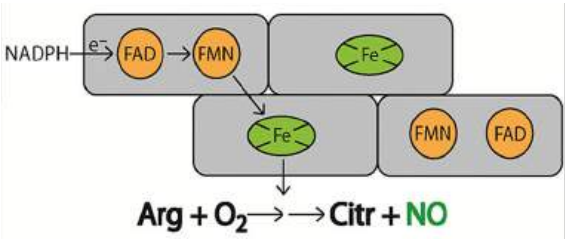


Figure 28.8.3: Electron transfer in nitric oxide synthase. Gephart et al. Nitric Oxide (2019). https://doi.org/10.1016/j.niox.2019.04.007. Creative Commons license

Two monooxygenase reactions occur in NOS synthase (oxidative domain) in which electrons are funneled into the heme and bound dioxygen (O_2) leading to the formation of water and the final product, NO. Electron transfer only occurs within the dimer when calmodulin is bound. However, iNOS is active even at basal Ca^{2+} concentrations.

Figure 28.8.3 shows an interactive iCn3D model of the structure of human neuronal nitric oxide synthase (with its PDZ domain) predicted by AlphaFold (P29475).

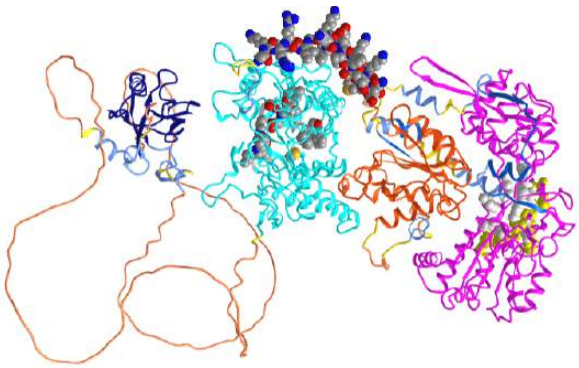


Figure 28.8.10: Human neuronal nitric oxide synthase predicted by AlphaFold (P29475) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...RCgzaSGpc259t6

In the iCn3D model, orient the protein as shown in the figure above. The dark blue (left) is the PDZ domain, and cyan is the oxidase (NOS synthase) domain that contains the heme (which is not shown in AlphaFold models). The spacefill CPK color shown in the cyan domain is active site residues interacting with the heme (not shown). The orange domain is the FMN (Flavodoxin_1) subdomain in the reductase domain. The magenta (far right) shows the FAD/NADPH subdomains of the reductase domain. The yellow spacefill shows the NAD binding pocket and the white spacefill the FAD binding pocket. The structures of amino acids 129-304 between the PDZ domain and the oxidase domain are not predicted with any certainty. Crystal structures are available for the oxidase domain alone. The spacefill CPK-colored helix (amino acid 730-754) represents a helical peptide region that binds to calmodulin.

NO that is synthesized in the cells can signal there or diffuse to another cell and signal there. Figure 28.8.11 shows how NO synthesized in vascular epithelial cells that line blood vessels can move into the nearby muscle cells and initiate signaling through soluble guanylyl cyclase there, leading to vasodilation and a lowering of blood pressure.

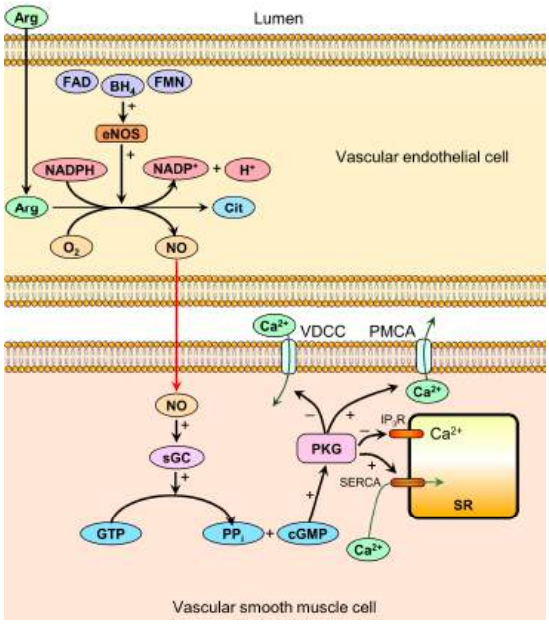


Figure 28.8.11: . Mechanisms of nitric oxide-mediated vasodilation. Khalaf et al. Nutrients 2019, 11, 1679; doi:10.3390/nu11071679 CC BY 4.0

Endothelial nitric oxide synthase (eNOS) located in the vascular endothelium forms NO from plasma arginine. Two substrates, O_2 and NADPH, are required along with the cofactors tetrahydrobiopterin (BH_4), FAD, and flavin mononucleotide (FMN). NO diffuses into smooth muscle cells and activates soluble guanylyl cyclase (sGC), increasing cGMP production. cGMP subsequently activates protein kinase G (PKG), resulting in decreased [Ca^{2+}] by these mechanisms:

- inhibition of voltage-dependent calcium channels (VDCC), reducing calcium influx;
- activation of plasma membrane calcium ATPases (PMCA), increasing ATP-dependent calcium efflux;
- inhibition of inositol triphosphate receptors (IP3R), reducing calcium release from the sarcoplasmic reticulum (SR) to the cytoplasm;
- activation of sarcoplasmic calcium ATPases (SERCA), increasing the ATP-dependent sequestration of calcium from the cytoplasm to the SR.

Decreased [Ca^{2+}] mediates smooth muscle relaxation via the activation of myosin light chain kinase and the inhibition of myosin light chain phosphatase (not shown in the figure), resulting in vasodilation.

Figure 28.8.13 shows retrograde diffusion of NO from an activated post-synaptic neuron back to the presynaptic neuron that excited it on the release of the neurotransmitter glutamate. NO is synthesized by post-synaptic cell nNOS after it's activated by Ca^{2+} inflow and binding to CAM (not shown). NO with the post-synaptic neuron binds to guanylyl cyclase to produce the second messenger cGMP which can directly activate other channels, protein kinase G, or phosphodiesterases (PDEs)

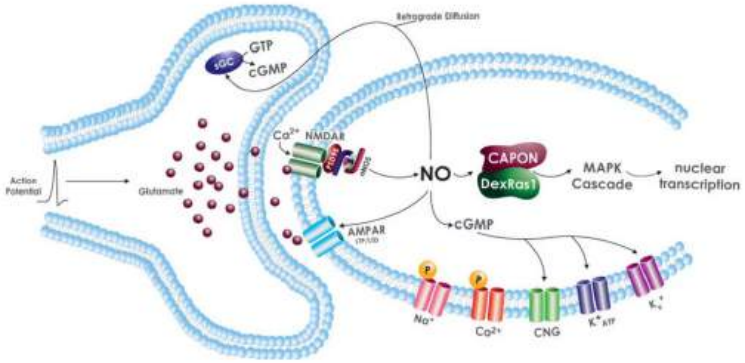


Figure 28.8.13: NO Signaling at a Neuronal Synapse. https://www.caymanchem.com/news/nitr...ion-in-the-cns.

Synaptic glutamate release activates postsynaptic NMDA and AMPA receptors (NMDAR, AMPAR) leading to Ca^{2+}-induced nNOS activation. NO will diffuse back to the presynaptic cell and activate sGC to produce cGMP, which has many signaling roles including affecting presynaptic neurotransmitter release. cGMP directly targets several ion channels in the post-synaptic cell. As we saw in the previous chapter, many ion channels are voltage-gated. However, ion channels can be regulated directly by ions (ex Ca^{2+}, Na^{+}) as well as by cyclic nucleotides such as cAMP and cGMP. The later channels are called cyclic nucleotide-gated (CNG) channels. NO in the post-synaptic cell also associates with CAPON, a nNOS binding protein, leading to downstream MAP kinase cascade.

Carbon Monoxide

Everyone knows that carbon monoxide (CO) in high doses is lethal as it binds to heme Fe^{2+} in hemoglobin and myoglobin with a higher affinity than O_2. Hence it may come as a surprise to you that endogenous CO is a signaling molecule, which now in retrospect might make sense given its similarity in chemical structure to NO.

CO is produced through heme oxygenase (HOs). CO can act as a signaling molecule in neural, cardiovascular, respiratory, gastrointestinal, immune, and reproductive systems. In contrast to the lethal effects of inhaling exogenous CO from incomplete combustion, endogenous CO has anti-inflammatory and antioxidant effects. It can also act to dilate the vasculature system. Other gases like H_2S also are signaling agents.

28.8.3: cGMP formation

The second (or third) messenger cyclic guanosine monophosphate (cGMP) is synthesized after activation of the enzyme **guanylyl cyclase (GC)** by nitric oxide. cGMP has many signaling effects in cells, some of which were outlined above. The cytoplasmic **soluble GC (sGC)** is activated by NO. The membrane-associated "**particulate" GC (pGC)** form is activated on the binding of natriuretic peptides (NPs) to natriuretic peptide receptors, which are NP-activated integral membrane guanylyl cyclase. The peptide hormones (ANP secreted by the atria and BNP secreted by ventricles) decrease blood pressure. The membrane form does not require NO for activation Figure 28.8.14shows the conversion of GTP to cGMP.

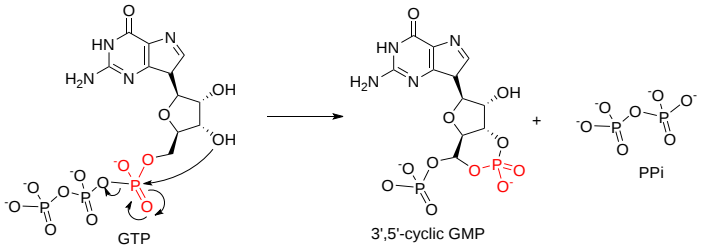


Figure 28.8.14: Conversion of GTP to cGMP

Synthesis of cGMP from soluble GC is activated by NO or molecules like nitrates that can be metabolized to NO. These molecules are called NO donors. Since NO causes vasodilation, they are used to treat angina and hypertension. A class of drugs called **stimulators** (for example riociguat), increases cGMP production from sGC in the absence and synergistically in the presence of NO. They are also used to treat hypertension. Another class of drugs called **activators** can activate sGC even if heme is oxidized or even missing without upstream NO signaling. They are effective even if the heme is oxidized or lost from the NOS catalytic domain.

We saw that cAMP is cleaved to AMP by phosphodiesterase. Likewise, phosphodiesterases (PDEs) cleave cGMP to GMP to attenuate signaling through cGMP. Selective drugs targeting PDE are available. These include sildenafil for the treatment of pulmonary hypertension and erectile dysfunction and tadalafil for benign prostatic hyperplasia (BPH).

Pathways for activation of guanylyl cyclase activity (sGC and pGC) are shown in Figure 28.8.15

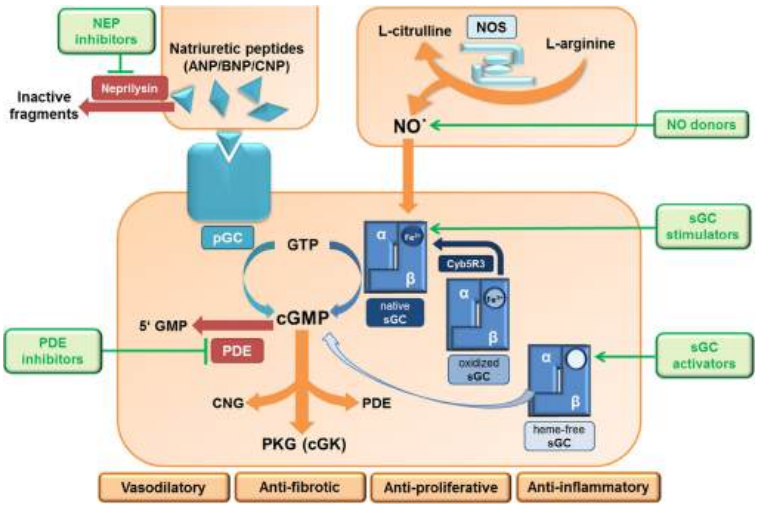


Figure 28.8.15: Pathways for activation of guanylyl cyclase activity (sGC and pGC). NP, atrial natriuretic peptide; BNP, brain natriuretic peptide; cGK, cGMP-dependent protein kinase; cGMP, cyclic guanosine monophosphate; CNG, cyclic nucleotide-gated ion channels; CNP, C-type natriuretic peptide; GTP, guanosine triphosphate; GMP, guanosine monophosphate; NO, nitric oxide; NOS, nitric oxide synthase; NP, natriuretic peptide; PDE, phosphodiesterase; pGC, particulate guanylyl cyclase; PKG, protein kinase G; sGC, soluble guanylyl cyclase. Naunyn Schmiedebergs Arch Pharmacol. 2020; 393(2): 287–302. doi: 10.1007/s00210-019-01779-z. http://creativecommons.org/licenses/by/4.0/.

28.8.4: Soluble guanylyl cyclase (cGC) structure and function

The soluble form of GC is a heterodimer of α and β subunits. The domain structure of guanylyl cyclase is shown in Figure 28.8.16

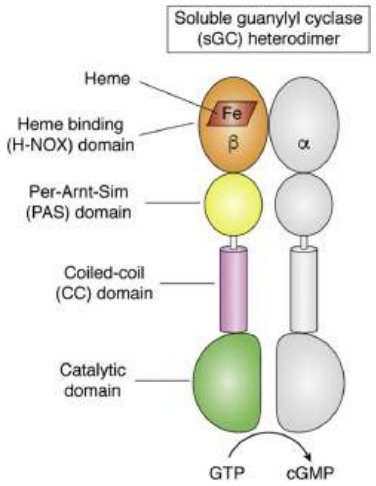


Figure 28.8.16: Domain structure of guanylyl cyclase. Stuehr et al. **JBC REVIEWS**| VOLUME 296, 100336, JANUARY 2021. https://doi.org/10.1016/j.jbc.2021.100336. http://creativecommons.org/licenses/by/4.0/)

It appears that when NO binds to the heme group, a twist in the coiled-coil domain leads to its extension, which leads to the activation of the catalytic domain. Simulators likely cause similar conformational changes initiated by their binding to the top part of the CC domain. Figure 28.8.17 shows an interactive iCn3D model of the human soluble guanylate cyclase in the riociguat (stimulator) and NO-bound state (7D9R)

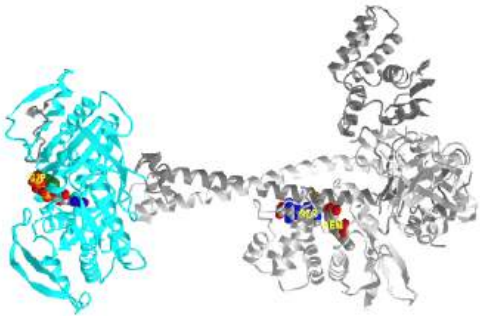


Figure 28.8.17: Human soluble guanylate cyclase in the riociguat and NO-bound state (7D9R) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...vm3dLK1o1odZD6

The A chain is dark gray and the B chain is light gray. The guanylate cyclase domain contributed to by each monomer is shown in cyan. A phosphonate GTP analog (labeled G2P) is shown in spacefill CPK color bound in the guanylyl cylase domain (cyan). The heme (HEM) and the stimulator riociguat (GZO) are shown in the HNOBA (Heme NO Binding Associated) domain and shown in spacefill CPK colors.

The conformation change of A chain of the bent inactive form of guanylyl cyclase (6JT1) to the active extended form (6JT2) is shown in Figure 28.8.18

Figure 28.8.18: The conformation change of A chain of the bent inactive form of guanylyl cyclase (6JT1) to the active extended form (6JT2)

The guanylyl cyclase domain is at the top of the figure. The conformational change is somewhat reminiscent of the change in apo-calmodulin on the binding of Ca^{2+} ions, although the central helix is fully intact in the inactive and active forms of guanylyl cyclase.

Figure 28.8.19 shows how oxidation of the heme iron and S-nitrosation of the protein that occurs in the presence of reactive oxygen species (ROS) and reactive nitrogen species (RNS) leads to heme loss and inactivation of sGC.

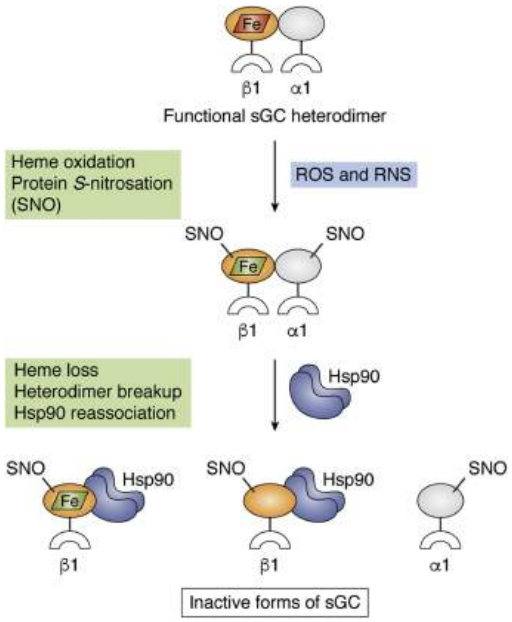


Figure 28.8.19: sGC inactivation pathways and potential structural consequences. sGC inactivation pathways and potential structural consequences. When functional sGC heterodimer (top) is exposed to reactive oxygen and/or reactive nitrogen species (ROS and RNS) it can become unresponsive toward NO by undergoing oxidation of its ferrous heme to ferric (*red* and *green parallelograms*) and protein modifications like Cys S-nitrosation (SNO). These events alone or in combination may lead to the breakup of the sGC heterodimer, heme loss, and/or rebinding of Hsp90 to the freed sGCβ1 subunit. NO, nitric oxide; PAS, Per-Arnt-Sim; sGC, soluble guanylate cyclase. Stuehr et al. ibid.

28.8.5: pGC structure and function

The other source of cytosolic cGMP is **particulate guanylyl cyclase (pGC)**. These are integral membrane protein receptors for natriuretic peptides (NPs), which when bound to the receptor activate the cytoplasmic guanylyl cyclase domain of the receptor. In effect, they are ligand (NPs)-gated receptor enzymes. The peptide hormones (ANP secreted by the atria and BNP secreted by ventricles) decrease blood pressure. There are seven variants of pGC (A-G) found in mammals. GC-A (also called NPR-A or NPR1) and GC-B (NPR-B or NPR2), are both receptors for natriuretic peptides. The domain structure of NPR-A is shown in Figure 28.8.20

Figure 28.8.20: Domain structure of GC-A (also called NPR-A or NPR1)

Figure 28.8.21 shows an interactive iCn3D model of the human atrial natriuretic peptide receptor1 AlphaFold predicted model (P16066)

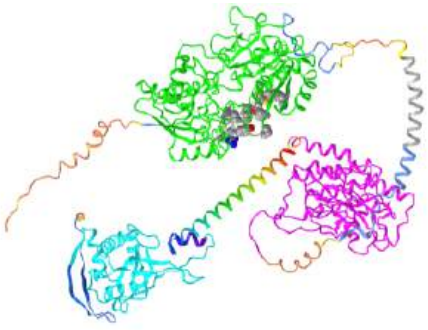


Figure 28.8.21: Human atrial natriuretic peptide receptor1 AlphaFold predicted model (P16066) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...Wd3xvHrY3rEUw6

Orient the model as shown in the figure above. The coloring coding in the model is as follows:

- green: Ligand (ANP)-binding domain of the type A natriuretic peptide receptor (NPR-A);
- magenta: PK-GC pseudokinase domain;
- cyan: cyclase domain;
- rainbow helix: HNOBA domain is found to be associated with the HNOB domain and pfam00211 in soluble cyclases and signaling proteins. The HNOB domain is predicted to function as a heme-dependent sensor for gaseous ligands rainbow;
- gray helix: transmembrane segment amino acids 474-494.

Note the protein is an integral membrane protein that passes through the membrane using a single alpha helix (474-494). The N-terminal domains above it and the C-terminal domains below it would orient like a typical single-pass membrane protein in the presence of a bilayer.

28.8.6: Protein Kinase G (PKG)

To briefly review, NO production leads to the production of cGMP. cGMP can directly bind to and regulate membrane ion channels. In alignment with the basic paradigm of signaling described throughout this chapter, we will now discuss how it activates **Protein Kinase G (PKG)** a member of the Ser/Thr Protein Kinase AGC family.

The are two mammalian genes for PKG1 and PKG2. Both are homodimers. PKG1 acts in the cytoplasm while PKG2 becomes tethered to the membrane by N-terminal myristoylation. Figure 28.8.22 shows the domain structure of PKG-I, which is similar to PKG2.

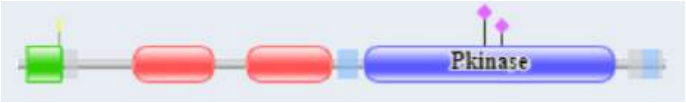


Figure 28.8.22: Domain structure of protein kinase G - I.

Red indicates the two nonidentical cGMP binding domains. Green is the N-terminal coiled-coil dimerization domain, which inhibits kinase activity in the absence of cGMP. On binding of cGMP, autoinhibition of the catalytic domain by the N-terminal domain is relieved. The binding of cGMP to the regulatory domain induces a conformational change which stops the inhibition of the catalytic core by the N-terminus and allows the phosphorylation of self (autophosphorylation) and then substrate proteins. Whereas PKG-I is predominantly localized in the cytoplasm, PKG-II is anchored to the plasma membrane by N-terminal myristoylation.

PKG1 is involved in modulating Ca^{2+} activity, platelet activation, smooth muscle contraction, gene expression as well as neural function. PKG2 helps regulate bone growth, intestinal secretion, and synaptic plasticity. It also regulates gene expression and activates the MAPK cascade in bone cells.

Figure 28.8.23 shows an interactive iCn3D model of the predicted structure of Human cGMP-dependent protein kinase 2 (AlphaFold, Q13237).

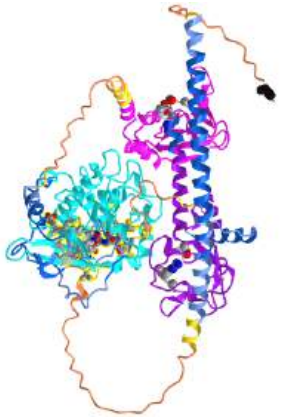


Figure 28.8.23: Predict the structure of Human cGMP-dependent protein kinase 2 (AlphaFold, Q13237). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...A55VorQFGZqpSA

The cyan domain is the STK (Ser/Th Kinase) domain. The magenta and purple domains are the CAP-Ed domains. The spacefill side chain structures within them represent the cGMP (or with lower affinity cAMP) binding sites. The CPK-colored sticks in the cyan kinase domain are the amino acid side chains in the active site where ATP binds. The orange backbone represents the least

confident part of the predicted structure. The black spacefill is the N-terminal Gly (after removal of Met) which is myristoylated, allowing targeting of the modified PKG2 to the cell membrane.

28.9: Gated Ion Channels - Neural Signaling

28.9.1: Introduction to Neural Signaling

Ion channels are membrane proteins that allow the flow of normally impermeant ions across the hydrophobic "sea" of cell or organelle membranes. The channel proteins can be constitutively open or can be "gated" open (or closed) by signals that affect the conformation of the protein. The signals can be ligands (in the form of hormones or neurotransmitters), post-translational modifications (mostly phosphorylation, dephosphorylation), or more difficult-to-understand events (change in transmembrane potential, pressure, or temperature). Gated channels are perhaps best studied through their role in neural signaling.

Neurochemistry is one of the most explosive areas of biological research. Scientists are now starting to unravel the molecular bases for memory, cognition, emotion, and behavior. The next decades will bring a greater understanding of brain chemistry and along with it the potential to alter human mood, and memory, and to treat mental illnesses such as schizophrenia much more effectively. The human brain has about 86 billion neurons. Compare this with some estimates for the number of stars in the Milky Way galaxy (about 100 billion, a number derived from luminosity and mass measurements but some estimates suggest 400 billion). Now imagine that each neuron can form connections - synapses - with 1000 to 10,000 other neurons. Throw in another 86 billion nonneuronal cells like glial cells and you have one of the most complex structures in the universe. This section will explore the biology and chemistry of neurons.

Recent studies suggest that there are around 3300 different types of cells in the brain. We will discuss two kinds of neurons - those that interact with muscles at the neuromuscular junction and those that interact with other neurons in the central nervous system. Neurons consist of a single, nucleated cell body with multiple signal-receiving outgrowths (dendrites) and multiple-signal sending outgrowths (axons) which end in a terminal button. These interact through the synapse with dendrites on other neurons. These characteristics are shown in Figure 28.9.1.

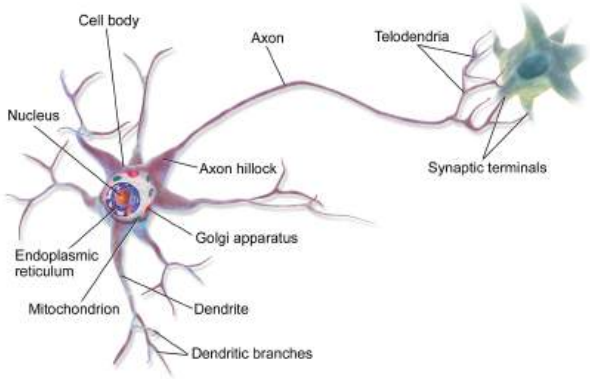


Figure 28.9.1: A neuron. Bruce Blaus - Own work, CC BY 3.0, https://commons.wikimedia.org/w/inde...curid=28761830

A presynaptic neuron can stimulate an adjacent postsynaptic neuron by releasing a neurotransmitter into the synapse between the cells, which binds to a receptor in the membrane of the post-synaptic cell, stimulating the cell, as shown in Figure 28.9.2

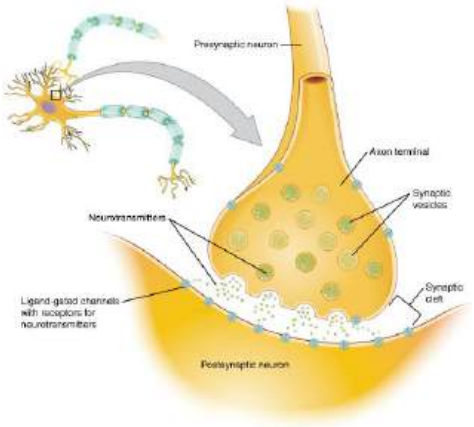


Figure 28.9.2 The synapse. https://cnx.org/contents/FPtK1zmh@8.25:fEI3C8Ot@10/Preface..Creative Commons Attribution 4.0 International license.

We will discuss the events which cause the post-synaptic cell to "fire", but we will not discuss the immediate events which lead to the release of neurotransmitters by the presynaptic neuron.

Neurons (as do all cells) have a transmembrane potential across the membrane. Transient changes in the membrane potential are associated with neuron activation or inhibition. This arises in part due to the imbalance of ions across the membrane which was established by the Na^+-K^+-ATPase in the membrane. This electrogenic antiporter transfers 3 Na ions out of the cytoplasm for every 2 K ions it transports in. Likewise Cl- has a much higher level

outside the cell. Membrane potentials are determined not only by the size of the ion gradients across the membrane but also by the differential permeability of membranes to ions. As we saw previously, synthetic bilayer membranes are not very permeable to ions, as shown in Table 28.9.1 below.

Ion permeability of phosphatidyl serine vesicles

ION	PERMEABILITY (cm/s)
sodium	$<1.6 \times 10^{-13}$ (lowest)
potassium	$<9 \times 10^{-13}$
chloride	$<1.5 \times 10^{-11}$ (highest)

Table 28.9.1: Ion permeability of phosphatidyl serine vesicles

Sodium would be expected to have a lower permeability than potassium since it has a higher charge density. It also is the largest in effective size due to the larger hydration sphere around the ion arising from its higher charge density. Chloride would be expected to have the highest permeability since it has the lowest charge density (due to the repulsion of the electron cloud in the negative ion). Intracellular charged proteins (which are mostly negative) are not permeable and help in creating the negative charge imbalance across the membrane.

Much work has been done on the giant axon of the squid, which has uniquely high intracellular potassium (400 mM compared to 20 mM outside) and high extracellular fluid sodium (440 mM vs, 50 mM inside) and chloride (560 mM vs 52 mM inside). Mammalian cell concentrations are much lower, but the relative size of the gradient is about the same. Typical ion concentrations and permeabilities for mammalian membranes are shown in Table 28.9.2 below.

T

Ion	Cell (mM)	Blood (mM)	Permeability (cm/s)
potassium	140	5	5×10^{-7}
sodium	5-15	145	5×10^{-9}
chloride	4	110	1×10^{-8}
X^- (negative macromolecules)	138	9	0

Table 28.9.2 Typical ion concentrations and permeabilities for mammalian membranes

How can we account for the markedly greater permeabilities of ions (1000x to 1,000,000 x) in mammalian cell membranes compared to synthetic lipid vesicles? Previously, we showed that glucose has a greater permeability through red blood cell membranes than through synthetic liposomes because of a membrane receptor that allows facilitated diffusion across the membrane and down a concentration gradient. The same thing is true of ion permeabilities in intact biological membranes. These membranes have several types of selective ion channels (non-gated - always open, and gated - open only after specific conformational changes). The non-gated channels dramatically increase the permeability of membranes to ions, as the glucose transport protein increases the permeability to glucose. It turns out that this differential permeability contributes to the transmembrane potential. Ion channels in nerves and muscles can move ions across the membrane at a rate of up to 10^9/s, which is comparable to k_{cat} for the best enzymes.

If we envision channels as pores, how can we account for the selectivity of the channel to specific ions? A larger pore should admit any ion less than a maximal size for the pore, so it is hard to imagine the nature of the selectivity filter. Because of this, many people discounted the ideas of channels in favor of a transporter, which would bind the ion selectivity and then, through conformational changes, move the ion across (much like the Na/K ATPase we discussed in the previous guide). This model could not, however, account for the incredible rates of ion flux across the membrane. Selectivity can be accounted for by a channel that contains a narrow opening that acts as an ion sieve. The ion loses most of its hydration sphere and forms specific interactions with amino acid side chains in the pore region. Such an interaction would be transient and not too tight since the ion must pass through the membrane. As we will see later, these ion channels:

- pass ions down a concentration gradient in a thermodynamically favorable process
- are specific for certain ions (although a few are less selective and will pass Na, K, Ca, and Mg ions)
- allow ion flow through either ungated or gated channels.,
- saturate with increasing ion concentration (even though as concentration increases, the ions have a greater thermodynamic drive to pass through the channel). This is consistent with the binding of the ion at a selectivity filter in the narrow part of the pore. The K_D for the interaction is usually in the mM range and indicates weak binding with large dissociation rate constants (k_{off}).

28.9.2: Transmembrane Potentials

Several questions arise about the distribution of ions and the magnitude of the transmembrane potential.

1. How are the ion gradients established?
2. How does the transmembrane ion distribution contribute to the membrane potential?
3. How can the resting electrochemical potential and the ion distribution be maintained?

The answer to these questions will be illustrated using studies on two types of brain cells, glial cells (which function as protectors, scavengers, and feeders for brain neurons) and neurons. Both types of cells have transmembrane potentials.

Glial Cells

The transmembrane ion gradients for ions can be established by different mechanisms. One uses ion-specific ATPases (P-type ion transporters), such as we discussed with the Na/K ATPase. This transporter ejects 3 sodium ions from the inside of the cell for every 2 potassium ions it transports in, all against

a concentration gradient. Since it is an electrogenic antiporter, it helps generate the potential. Additionally, specific ion channels also contribute (as described below) to the transmembrane gradients and potentials.

The harder question is how the ion distribution contributes to the membrane potential. Two things must occur for a membrane potential to exist: First, there must be a concentration gradient of charged ions (for example, sodium, potassium, or chloride) across the membrane. Second, the membrane must be differentially permeable to different ions. If the membrane were completely impermeable to ions, then no movement of ions across the membrane could occur, and no membrane potential would arise. If, however, membranes are differentially permeable to the ions, an electrical potential across the membrane can arise. Remember, synthetic bilayers are quite impermeable to ions, given the hydrophobicity of the internal part of the bilayer. Likewise, it is quite impermeable to glucose. It turns out that glial cells appear to have only a non-gated potassium channel, which allows the outward flow of potassium ions down the concentration gradient. The inside will then have a net negative charge since impermeable anions remain.

The chemical potential gradient causes this outward flow of potassium ions. As more ions leave, the inside gets more negative, and a transmembrane potential develops which resists further efflux of potassium. Eventually, they balance, and the net efflux of potassium stops. The resting transmembrane potential reaches around -75 mV which is exactly the value obtained from the equations we will derive below. Since glial cells appear to only express a non-gated (or leakage) potassium channel, their resting potential is equal to the potassium equilibrium potential. Figure 28.9.3 shows hypothetical transmembrane potentials (Ψ), and the electrical and chemical potentials (ΔG) for K^+ loaded vesicles with and without a non-gated K^+ channel

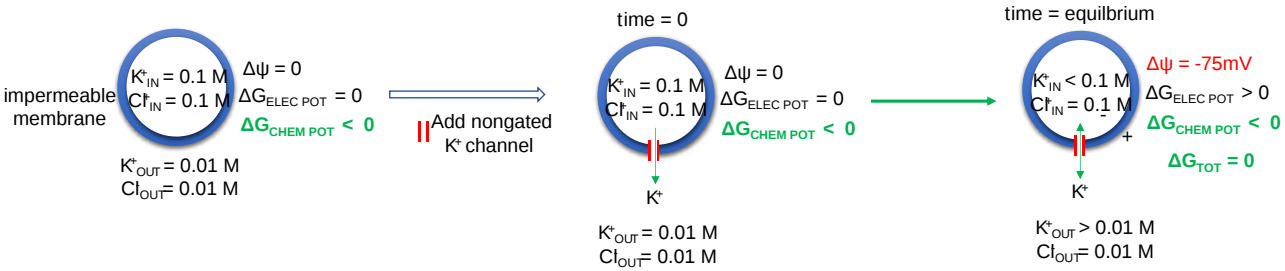


Figure 28.9.3: Visualizing the transmembrane potential in K^+ loaded vesicles with and without a non-gated K^+ channel

As we have shown previously, very little K^+ efflux is required to develop a transmembrane potential. At equilibrium, the K^+_{IN} is simply shown as < 0.1 M while K^+_{OUT} is shown as > 0.1 M

Here is a PhET simulation showing the differences between non-gated (leakage, always open) and gated channels (opened in response to stimuli like ligands, transmembrane potential, mechanical forces, etc).

28.9.3: Neurotransmitter Activation of Neurons

What happens when a neurotransmitter binds to a receptor on the post-synaptic cell? We will study two examples. The first is the simplest: binding of the neurotransmitter acetylcholine, released by a motor neuron, to its receptor on muscle. This region is called the neuromuscular junction. The binding of acetylcholine will lead to a transient depolarization of the muscle cell. Next, we will discuss the interaction of a neurotransmitter with a post-synaptic neuron in the central nervous system. This is a much more complex system. Their differences are described below:

In neurons interacting with muscles:

- Most muscle fibers are innervated by only one neuron - a motor neuron
- Neurotransmitter release at the neuromuscular junction leads only to muscle excitation, not inhibition.
- All fibers are excited by the same neurotransmitter - acetylcholine.

In the central nervous system, life is more complicated:

- Stimuli are received from hundreds to thousands of different neurons.
- Nerves receive both excitatory and inhibitory stimuli from neurotransmitters
- Different kinds of receptors are present to receive stimuli, which control the activity of different kinds of channels.
- The ion channels in neurons are gated by a variety of mechanisms in addition to changes in membrane potential, including gating by heat, cold, stretch, or covalent modification.
- Most nerve cells have a resting potential of about -65 mV compared to -90 mV for a muscle cell.

What happens when a neurotransmitter binds to the receptor on the post-synaptic cell? A depolarization occurs (mediated by conformational changes in the transmitter-receptor complex), raising the membrane potential from the resting equilibrium level. What happens next depends on the identity of the post-synaptic cell. In the muscle cell, the rising potential caused by the binding of acetylcholine ultimately leads to muscle contraction by opening intracellular organelle membrane calcium channels. In a neuron, the rising potential triggers an action potential by opening voltage-gated sodium channels. The potential rises to about + 35 mV but does not reach the Na ion equilibrium potential, because the high positive potential opens a voltage-gated potassium channel. The potential then falls until it reaches the K ion equilibrium potential when the cells are hyperpolarized. It slowly then relaxes

back to the resting potential of -60 mV. This wave of changes in potential sweeps down the post-synaptic cell membrane and is the basis for the "firing" of the neuron. A plot of transmembrane voltage changes vs time for a typical action potential is shown in Figure 28.9.4

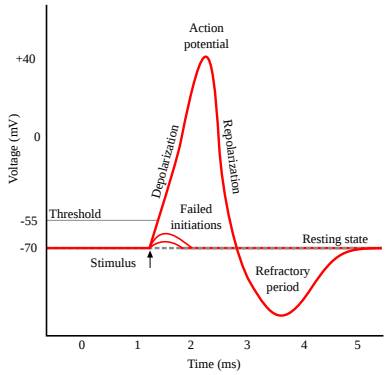
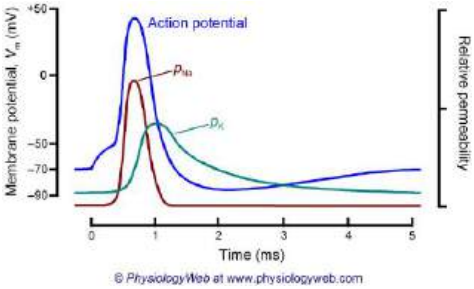


Figure 28.9.4: Action potential of a neuron. Original by en:User:Chris 73, updated by en:User:Diberri, **GNU Free Documentation License**,

Figure 28.9.5shows the actual changes in ion permeability in various phases of the action potential (replace figure, seek permission)



mine

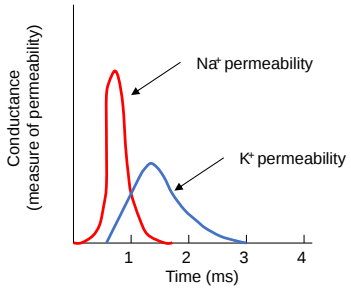
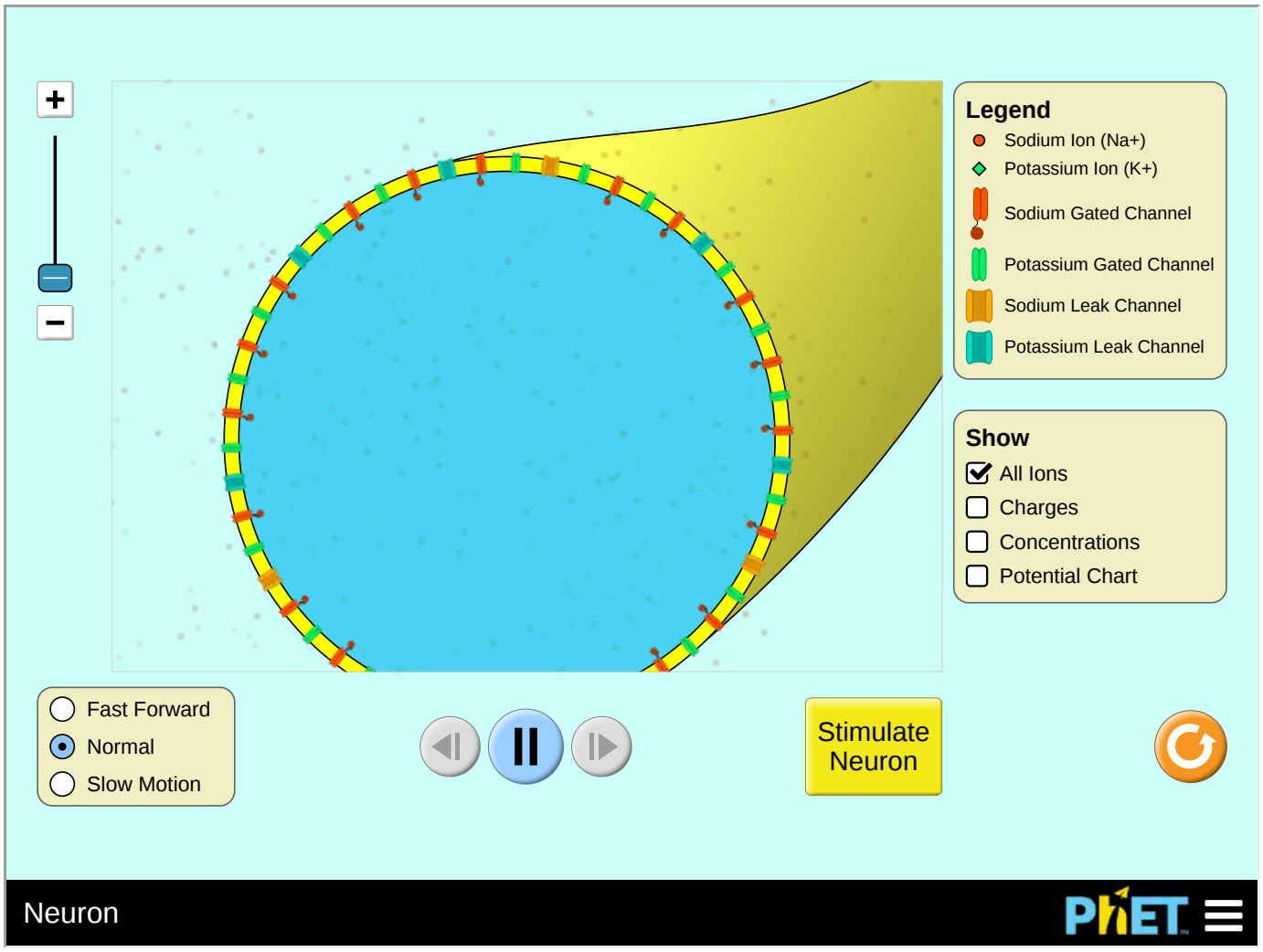


Figure 28.9.5: Na^+ and K^+ permeability during the action potential

Figure 28.9.6shows an animation of a neuron firing showing all the key players. (produced by PhET, University of Colorado, Boulder).



28.9.4: Proteins of the Neural Synapse

We must now account for the rise and fall of the membrane potential to a variety of neurotransmitters, including the cholinergic transmitters (ex. acetylcholine), catecholamines (dopamine, epinephrine, norepinephrine), amino acid derivatives (ex. Glu, Asp, N-methyl-D-Asp, Gly, gamma-amino-butyric acid -GABA), and peptides (endorphins, enkephalins). We will consider five membrane proteins as shown below in Figure 28.9.7.

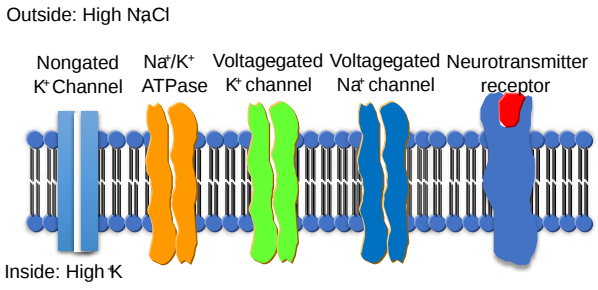


Figure 28.9.7: 5 key membrane proteins involved in the action potential

Many of these we have explored in detail in Chapter 11.3: Diffusion Across a Membrane - Channels, so we will focus on just one new one, the chloride channel, in the following section.

Na^+-K^+-ATPase: It transports both sodium and potassium ions against a concentration gradient using ATP as an energy source. The protein is a sodium-dependent ATPase. Without this protein, the membrane potential could not be maintained since the sodium and potassium gradient would collapse. It also contributes to the potential since it is an electrogenic antiporter. (In addition, we have seen that ungated potassium and sodium channels are also present.)

Neurotransmitter receptor: The receptors we will consider here are typically ligand-gated ion channels. Once the ligand binds, a conformational change occurs in the protein, allowing a flow of ions down a concentration gradient. Depending on the nature of the ion, the channel either initiates depolarization

(when Na^+ enters from the outside and raises ΔΨ) or inhibits depolarization (when Cl^- enters from the outside and lowers ΔΨ). When chloride channels open, they hyperpolarize the transmembrane potential. Stimulatory neurotransmitters (like glutamate) lead to depolarization of the membrane, while inhibitory neurotransmitters (like gamma-aminobutyric acid) lead to hyperpolarization of the membrane (making the potential more negative).

Na^+ channel (voltage-gated): When the ligand-gated channel depolarizes the membrane to some threshold value, sodium channels undergo a conformational change and open allowing Na^+ ions to flood into the cell, raising the potential to a positive approx. 33 mV (a value lower than the equilibrium sodium potential). This membrane protein is a voltage-gated channel, not a ligand-gated one.

Two potent neurotoxins, tetrodotoxin (from Pufferfish) and saxitoxin bind to the channel and act as antagonists (inhibit the activity of the receptor by blocking sodium influx). Their structures are shown in Figure 28.9.8

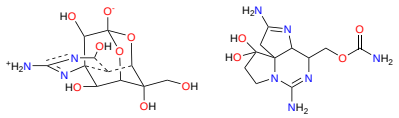


Figure 28.9.8: Structures of tetrodotoxin and saxitoxin

The guanidino group of the tetrodotoxin appears to bind with high affinity to the entrance of the channel that interacts with the hydrated sodium ion. Affinity chromatography using tetrodotoxin beads has been used to purify the protein. Figure 28.9.9 shows the relative sizes of ions used to probe size requirements for the channel. the hydrated K^+ ion can not pass through.

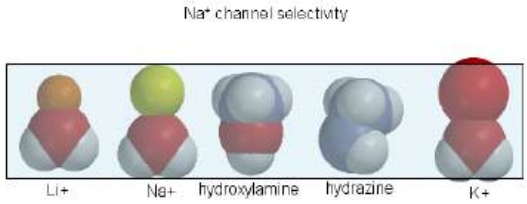


Figure 28.9.9: Na^+ channel ion selectivity

Depolarization of the membrane potential may result in an outward movement and rotation of the positively charged helixes containing Lys and Arg side chains which had presumably formed salt bridges (ion-ion interactions) with negatively charged side chains within the protein. Depolarization of the membrane results in breaking a few salt bridges, and effectively leads to the movement of 1 or 2 charges on the helix through the membrane. Work occurs when charges are moved through an electric field. Work is also related to the ΔG for the system, which is also dependent on the ratio of the open and closed form of the channel. Other voltage-gated ion channels (for potassium and chloride) have a similar membrane topology and an S4 voltage-sensor helix.

K^+ channel *(voltage-gated)*: When the membrane potential reaches around +25 mV or so, the K+ channel, a voltage-gated membrane protein, alters its conformation, allowing K^+ efflux from the cells, lowering the potential until it reaches the potassium equilibrium potential. It slowly relaxes back to the cell resting potential of about -60 mV.

Cl^- channel: If these channels (typically ligand-gated) are open, they will hyperpolarize the cell and make it more difficult to fire.

The selectivity filter is composed of many stacked rings of oxygens that can interact with a dehydrated K ion but not with a dehydrated Na ion which can not approach close enough to form significant interactions. Surrounding the filter are twelve aromatic amino acids which constrain the size of the pore opening. The interactions of the filter O's with the K ion make up for the energetically disfavored dehydration of the ion. The filter contains two K ions which repel each other, assisting in the vectorial discharge of the ions through the membrane. These ions must form weak interactions with the selectivity filter. The actual pore is mostly hydrophobic, which facilitates the flow-through of the ions. The central cavity of the pore can hold some water molecules in addition to the K ions which helps stabilize the ion in the middle of the pore.

28.9.5: Inhibitory Neurotransmitters

The main inhibitory neurotransmitters are GABA (gamma-aminobutyric acid), which is made from glutamic acid through decarboxylation of the α-C group, and glycine. They bind to transmitter-gated chloride channels, which when open hyperpolarize (make more negative) the transmembrane potential. Benzodiazepines (like Valium and Librium - anti-anxiety and muscle-relaxing agents) and barbituates (like phenobarbital-hypnotics) bind at allosteric sites on the GABA receptor and potentiate the binding of each other and GABA. This receptor is also affected by alcohol and anesthetics. Let's focus on the $GABA_A$ channels (also called the GABA(A) receptor or GABA(A)R) since it binds so many interesting pharmaceutical drugs.

Figure 28.9.10 shows an interactive iCn3D model of the human full-length heteromeric α1-β3-γ2L GABA(A)R in complex with picrotoxin, GABA, and megabody Mb38 (6huj) derived from cryo-EM (loads slowly).

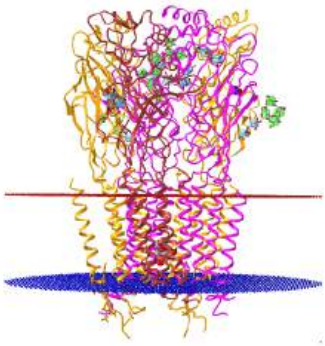


Figure 28.9.10: Human full-length heteromeric alpha1beta3gamma2L GABA(A)R in complex with picrotoxin, GABA and megabody Mb38 (6huj) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...VdExv4QhFPsVX8 (loads slowly)

This structure is a Type-A γ-aminobutyric ($GABA_A$) receptor. It consists of two alpha chains (orange), two beta chains (magenta), and one gamma chain (brown). Two GABAs (spacefill, CPK colors) are bound between the orange and magenta subunits (between the α and β chains). Picrotoxin is shown in spacefill CPK colors in the central pore (bottom), near the cytoplasmic end (blue bilayer dummy atoms). The extracellular domain (above the red) has glycans shown in SNFG cartoon form and shown in Figure 28.9.11.

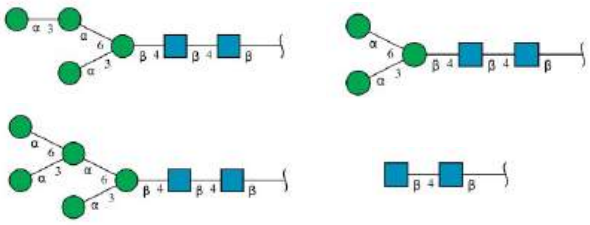
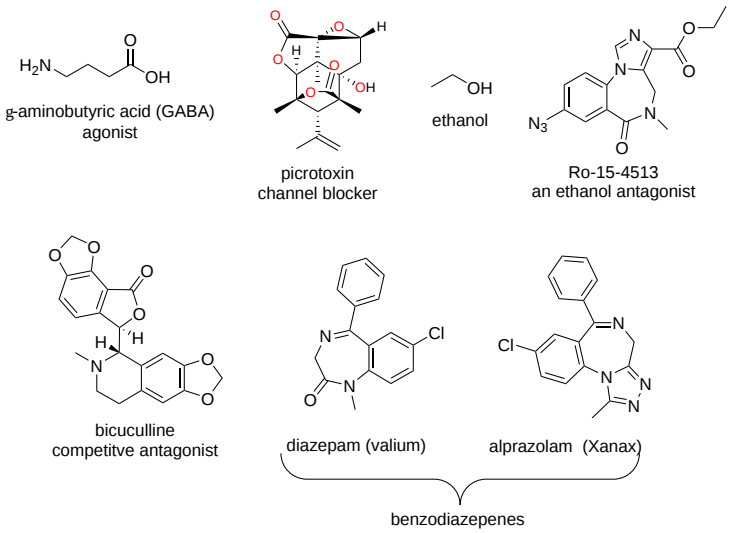


Figure 28.9.11: Glycan structure for the Type-A γ-aminobutyric ($GABA_A$) receptor

The structure of the receptor bound to many pharmacological agents has been solved. The structure of GABA (the agonist), bicuculline (a competitive antagonist), the benzodiazepines alprazolam (Valium) and diazepam (Xanax), a channel blocker (picrotoxin), ethanol, and Ro-15-4513 (an ethanol antagonist) are shown in Figure 28.9.12



(\PageIndex{12}\): Structures of GABA-R binding molecules -GABA (the agonist), bicuculline (a competitive antagonist), the benzodiazepines alprazolam (Valium) and diazepam (Xanax), a channel blocker (picrotoxin), ethanol, and Ro-15-4513 (an ethanol antagonist)

The benzodiazepine diazepam (Xanax) binds at an allosteric site and promotes GABA binding, so it is considered an "**indirect agonist**". It is also an anxiolytic (a drug used to reduce anxiety) and an anticonvulsant. Ethanol activates the inhibitory GABA channel. Too much ethanol leads to cells hyperpolarized cells so neural responses are greatly inhibited, possibly leading to death. Ethanol acts synergistically with benzodiazepine, which makes this combination so lethal.

Figure 28.9.13 shows an interactive iCn3D model of the human full-length heteromeric α1-β3-γ2L GABA(A)R in complex with diazepam (Valium), GABA (6HUP) (loads slowly). https://structure.ncbi.nlm.nih.gov/i...fTZMqePnpdwTy8

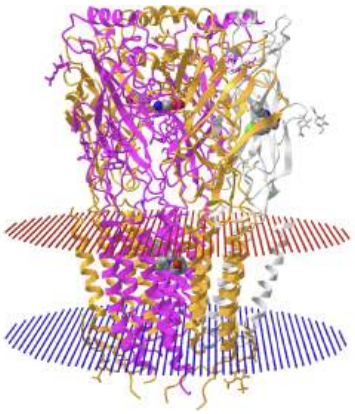


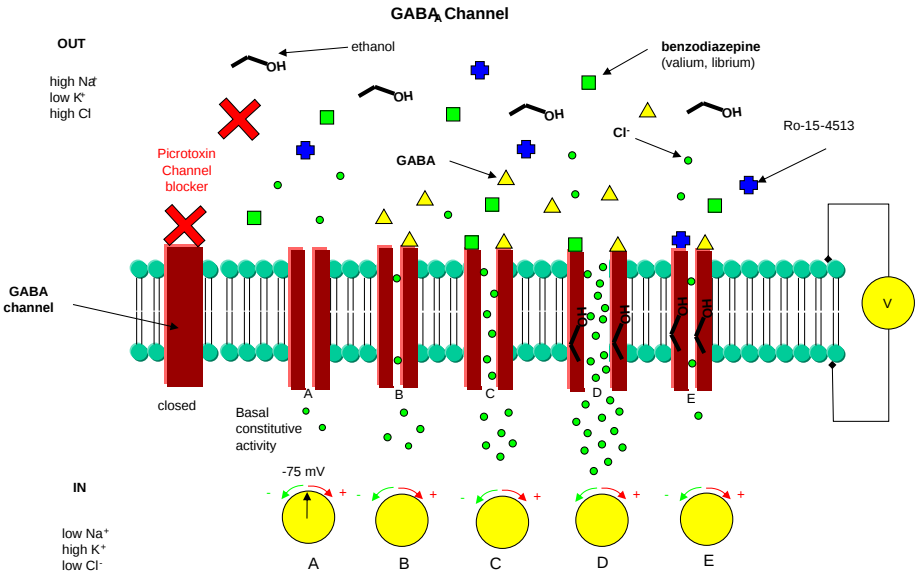
Figure 28.9.13: Human full-length heteromeric α1-β3-γ2L GABA(A)R in complex with diazepam (Valium), GABA (6HUP). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...fTZMqePnpdwTy8

The structure shows two alpha chains (orange), two beta chains (magenta), and one gamma chain (gray). Two GABAs (spacefill, CPK colors) are bound between the orange and magenta subunits (between the α and β chains). Three valiums (spacefill, CPK) are shown in the structure. One is bound to an allosteric site between the orange (alpha) and gray (gamma) chains. Two others are bound within the bilayer, this time between the alpha (orange) and beta chains (magenta) chains. Different benzodiazepines appear to bind in slightly different sites, and binding elicits their specific effects.

A drug, Ro-15-4513, was developed in the 1980s that antagonizes the effect of ethanol. It has a complicated pharmacology. It is considered a benzodiazepine "**partial inverse agonist**". It has no effect by itself. It reverses the anticonvulsant effects of benzodiazepines, and blocks the Cl^- effects of ethanol so it is an inverse agonist of GABA-mediated Cl^- flux. Hence it antagonizes the effect of benzodiazepines and alcohol. If given to intoxicated mice, they act normally! Such drugs have not reached the market as their use poses significant ethical issues.

The effect of pharmacological agents on the GABA receptor

In the "voltmeters" below, draw an arrow indicating if the transmembrane potential becomes more negative or more positive for the conditions given.



Answer

Add texts here. Do not delete this text first.

28.9.6: Metabotropic Neural Receptors

Some signaling molecules, whose effects are regulated by kinases (β-adrenergic and some olfactory signals by PKA and acetylcholine by PKC for example), are neurotransmitters. In all the examples presented previously, the neurotransmitters gate the inactive ion channels directly. These types of membrane receptors are classified as **ionotropic**. Typical examples of neurotransmitter-gated ion channels are the acetylcholine receptor in neuromuscular

junctions and the Glu, Gly, and GABA receptors in the central nervous system. These receptors are multimeric proteins. Receptors with direct gating of ion flow are fast, with activities that last milliseconds, and are used in eliciting behavioral responses.

However, ion channels can also be gated indirectly when the neurotransmitter binds to its receptor and leads to downstream events which subsequently open an ion channel that is distinct from the receptor. In this case, the occupied receptor communicates to an ion channel indirectly through a G protein for example. Examples of this indirect gating of ion channels include the serotonin, adrenergic, and dopamine receptors in the brain. These receptors are classic single-protein serpentine GPR receptors with 7 transmembrane helices and intracellular domains that can interact with G proteins as described above to increase second messenger levels (cAMP, DAG) in the neuron. The receptors are classified as **metabotropic** since they must activate a series of metabolic steps before ion channels are open. The second messengers can either activate kinases in the cell, which phosphorylate ion channels to either open or close them, or can bind directly to the channel and modulate its activity through an allosteric conformational change. In some cases, the G protein directly acts on the ion channel. These different ways are described in Figure 28.9.14

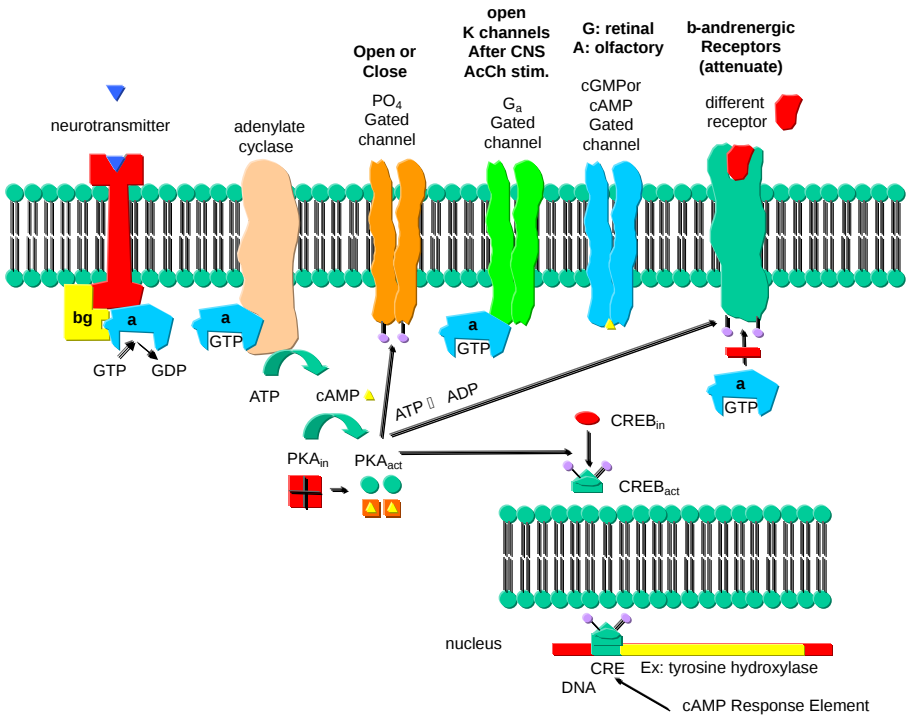


Figure 28.9.14: Neurotransmission: Gating through G Linked Receptors

In contrast to direct gating, receptors that indirectly gate ion channels have activities that are slow and last seconds to minutes. These receptors are usually involved in modulating behavior by changing the excitability of neurons and the strength of neural connections, hence modulating learning and memory. These changes can occur in many ways, summarized below:

Phosphorylating ion channels: Receptors that act through a second messenger system can change ion channel activity by activating kinases that phosphorylate the channels. This may:

- open the channel normally closed at the resting potential and produce an effect like gating.
- close a channel usually opens at the resting potential (such as non-gated K channels which when closed would depolarize the cell and make it more excitable).

Gα interaction with ion channels:

- the Ga subunit of the G protein interacts with K channels after stimulation of the CNS Acetylcholine receptor, opening the channel and hyperpolarizing the cell

Second messenger interaction with ion channels:

- cGMP opens cation channels in retinal cells after activation of the photoreceptor by photons
- cAMP opens cation channels in olfactory cells after activation of the olfactory receptor by odorants.

Second messenger effects on proteins other than ion channels (usually different receptors):

- the β-adrenergic receptors are phosphorylated by PKA and PKC (activated by stimulation of a different neurotransmitter receptor linked through a G protein to produce increased levels of second messengers cAMP and diacylglycerol). When phosphorylated, the β-adrenergic receptor, itself activated through G protein) can't bind Gs. This attenuates the response of the β-adrenergic receptor to its neurotransmitter which leads to desensitization to that signal.

Second messengers regulate gene expression:

- cAMP-activated PKA can phosphorylate an inactive transcription factor in the cell, which then can bind to a section of DNA called the cAMP Response Element (CRE), which is upstream of certain genes, leading to the transcription of the genes. The transcription factor is called CREB for cAMP Response Element Binding protein. Example: tyrosine hydroxylase (a monooxygenase) is involved in the synthesis of epinephrine and norepinephrine. The activity of this protein is increased when it is phosphorylated by PKA. Hence its activity can be increased quickly by this modification of the already present protein. If an animal is subjected to severe or long-term stress (cold or immobilization), presynaptic cells with norepinephrine will be stimulated to release the neurotransmitter. This requires continual synthesis of the neurotransmitter by the presynaptic cell. The increase in the synthesis of this neurotransmitter is caused by the presynaptic cell being stimulated by another neuron, which leads to increased levels of cAMP, and ultimately activation of CREB which increases transcription of the hydroxylase gene.

Caffeine

Caffeine produces a state of arousal in the central nervous system. High levels appear to block the binding of an inhibitory neurotransmitter, adenosine, to the A2A adenosine receptor. In the absence of caffeine, adenosine levels rise during the day, which promotes interaction with its receptor, leading to increased sleepiness and lack of concentration. When adenosine binds normally to its receptors, it activates the adenylate cyclase cascade, which activates PKA, leading to changes in the phosphorylation state of many proteins inside the cell, including protein phosphatase (2A). These changes inhibit neural firing. Caffeine blocks these changes.

Hallucinogenic drugs

Illicit drugs like LSD, psilocybin, and ecstasy can produce hallucinations as they have profound effects on consciousness and perception of self and reality. Recent clinical studies have shown that under tightly controlled conditions and doses, these drugs might have significant therapeutic effects in the treatment of mental health issues such as depression and post-traumatic stress syndrome. Hence their mechanisms of action have been the source of many studies.

All of these drugs bind to the human serotonin 2A receptor (5-HT_{2A}R), a metabotropic GPCR receptor of serotonin (5-hydroxytryptamine)**.** When serotonin binds, it leads to GPCR signaling, partly through beta-arrestins. These are **adapter proteins** that form complexes with ligand-bound and activated GPCRs and GPCR protein kinases.

Structures show that LSD binds to the orthosteric binding site (i.e. the active site, not an allosteric site) for serotonin. The study also found that serotonin and psilocin binding extends into an adjacent site called the extended binding pocket (EBP). These differences suggest that it should be possible to design agonists that have therapeutic, but not hallucinogenic properties. Serotonin binding does not cause hallucinations. A drug, IHCH-7086, in mouse studies seems to not provoke hallucinations but appeared to have antidepressive effects (based on observations of mouse behaviors like twitching and freezing). The structures of serotonin (5HT) and psychotropic drugs that bind the $5HT_{2A}$R are shown in Figure 28.9.15

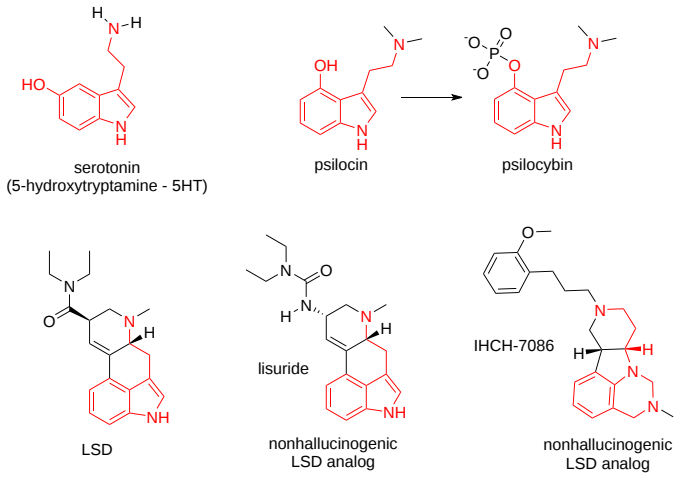


Figure 28.9.15 Structures of serotonin (5HT) and psychotropic drugs that bind the $5HT_{2A}$R

Figure 28.9.16 shows a model of the serotonin receptor (5HT-2A) with each of the bound drugs. Note the different occupancy of the orthosteric and extended binding pockets.

Figure 28.9.16: The serotonin receptor (5HT-2A) with bound serotonin and orthosteric binding analogs.

The GPCRs have been aligned for each of the poses. The ligands are represented in spacefill in a single color, as listed below:

- red: serotonin (5HT), the physiological agonist, 7WC4
- yellow: psilocin, 7WC5 LSD, 7WC6
- magenta: LSD, 7WC6
- cyan: IHCH-7086

An alternative way to explain the fact that serotonin does not cause hallucinations and why the other agonists do is that in addition to binding plasma membrane GPCRs, the agonist might have additional effects that are associated with hallucinations by binding intracellular membrane GPCRs.

The cerebral cortex plays a key role in what it means to be human. It is involved in higher-level processing involving thought, learning, emotion, personality, language, and memory. As such its dysregulation is involved in mental health conditions such as depression et al. in which synaptic plasticity is decreased as evidenced by a reduction in dendritic structure require for synaptic connections. The serotonin (5HT) reuptake inhibitors (SSRIs) used to treat depression appear to increase neural plasticity and connections over time (so they don't give immediate responses). Hallucinogens that target the 5-HT2A receptor (5-HT2AR) also promote neuroplasticity. However, neuron growth in cortical cultures is not affected by serotonin. What is the mechanism that allows various combinations of neuron plasticity and hallucinations all from the same receptor?

One factor that determines these differential effects is that serotonin (5HT) is more polar and cannot readily enter cells, while the HT2AR agonists that are hallucinogenic are less polar (more lipophilic) and could potentially enter cells where they might elicit hallucinogenic effects. Intracellular hallucinogens would bind internal 5-HT2AR receptors, which are found in cortical neurons on the Golgi and other organelles that are more acidic than the cytoplasm or extracellular environment. (In fact, in cortical neurons, the main location of the 5-HT2ARs is intracellular.) This might lead to more prolonged retention of intracellular hallucinogens and longer signaling (LSD has a profound hallucinogenic effect that lasts for 10 hours or more) leading to neuronal growth as well.

If serotonin is given to wild-type mice, no hallucinations occur, as evidenced by the lack of a head twitch response (HTR). However, if given to mutant mice who expressed serotonin transporter (SERT) on cortical neurons, the HTR response was observed. The import of serotonin led to neuroplasticity. It could be that serotonin is not the physiological ligand for intracellular HT2ARs. It might also imply that as with the case with endocannabinoids, we have endogenous psychedelics. The subtle difference in the binding of serotonin and the hallucinogens in Figure 16 might have little to do with their tendency to produce hallucinations.

Recent Updates: LSD binding to the BDNF receptor TrkB 06/09/23

Much is still not known about how antidepressants work and what causes delays in their therapeutic effects. Increased synaptic connections through expanded neuroplasticity appears to be required for their therapeutic effects. The major action of drugs like Prozac, a serotonin reuptake inhibitor, is through blocking the serotonin transporter (SERT), increasing extracellular levels of serotonin in the neural synapse. Tricyclic anti-depressants as well as monoamine oxidase inhibitors (MAOI), also increases monoamine neurotransmitters in the synapse with delayed therapeutic effects (the monomaines act quickly however). Perhaps these agents bind other receptors!

In fact, they do. The binding of both typical and fast antidepressants also occurs in the transmembrane domain of tyrosine kinase receptor 2 (TRKB), the brain-derived neurotrophic factor (BDNF) receptor, which is linked to increases in neuroplasticity. The receptor also binds cholesterol which modulates its activity. The antidepressants binding site is formed on dimerization of their transmembrane domains. Mutations in the transmembrane region block the efficacy of the antidepressants.

LSD and other psychedelics also produce fast and long-lasting antidepressant effects promoted by increases in neuroplasticity. Studies have shown that LSD and psilocin bind to slightly overlapping sites in the transmembrane domain of the BDNF receptor as antidepressants, but with a 1000-fold higher affinity. If the LSD binding site on the serotonin 2A receptor (5-HT_{2A}) is blocked, LSD still has antidepressant and increased neuroplasticity effects.

Figure 28.9.17shows the interaction of LSD and the metabolite of psilocybin, psilocin (PSI), with the TRKB receptor.

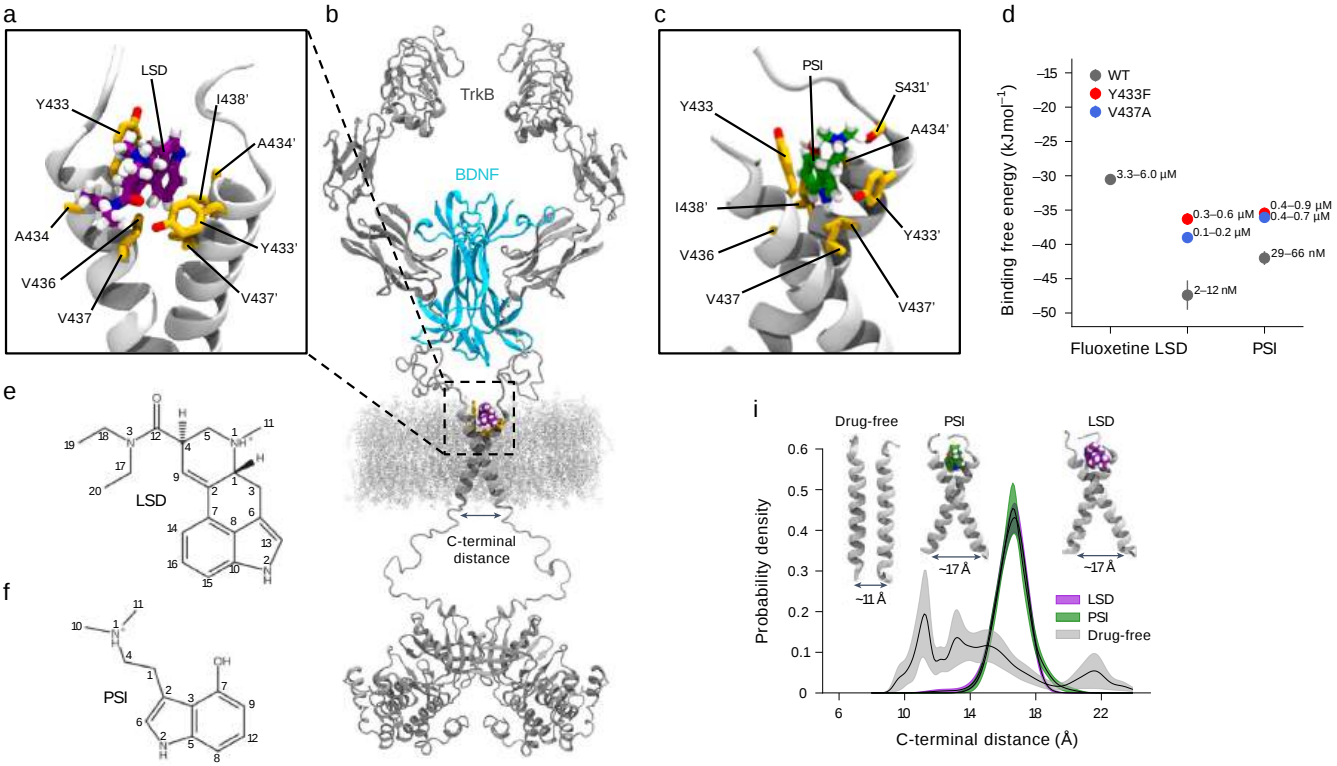


Figure 28.9.17 **Characterization of the psychedelics binding site in the TrkB TMD.** Moliner, R., Girych, M., Brunello, C.A. *et al.* Psychedelics promote plasticity by directly binding to . BDNF receptor TrkB. *Nat Neurosci* **26**, 1032–1041 (2023). https://doi.org/10.1038/s41593-023-01316-5. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Panel a–c show representative MD snapshots showing the binding pocket for LSD (purple) (**a**) and PSI (green) (**c**) in the extracellular-facing crevice of the TrkB TMD dimer (gray). Side chains (yellow) of relevant binding site residues are displayed. A structural model of full-length TrkB dimer (gray) embedded in a lipid membrane is shown with bound BDNF (blue) and LSD (purple) (**b**).

Panel d shows in silico binding free energy estimations for fluoxetine, LSD and PSI. Each free energy estimate (ΔG, circles) and its statistical error (SE, bars) were estimated from a separate set of FEP simulations ($n = 1$). Dissociation constants are given as a range with upper and lower bounds converted from ΔG − SE and ΔG + SE, respectively.

Panel e,**f** show chemical structures of LSD (**e**) and PSI (**f**) with atom numbers annotated.

Panel i shows the distributions of TMD dimer C-terminal distance and shows that LSD and PSI stabilize the cross-shaped conformation of TrkB favorable for receptor activation in a 40 mol% CHOL membrane. Lines represent the mean distribution, and bands represent the standard errors ($n = 10$ independent simulations). TMD conformations corresponding to indicated C-terminal distances and drug-bound states are shown in the inset.

Figure 28.9.18shows the different binding modes of LSD and Fluoxetine (Prozac), a selective serotonin reuptake inhibitor (SSRI) for TrkB. Fluoxetine is used to treat depression, obsessive-compulsive disorder (OCD), bulimia nervosa, and panic disorder.

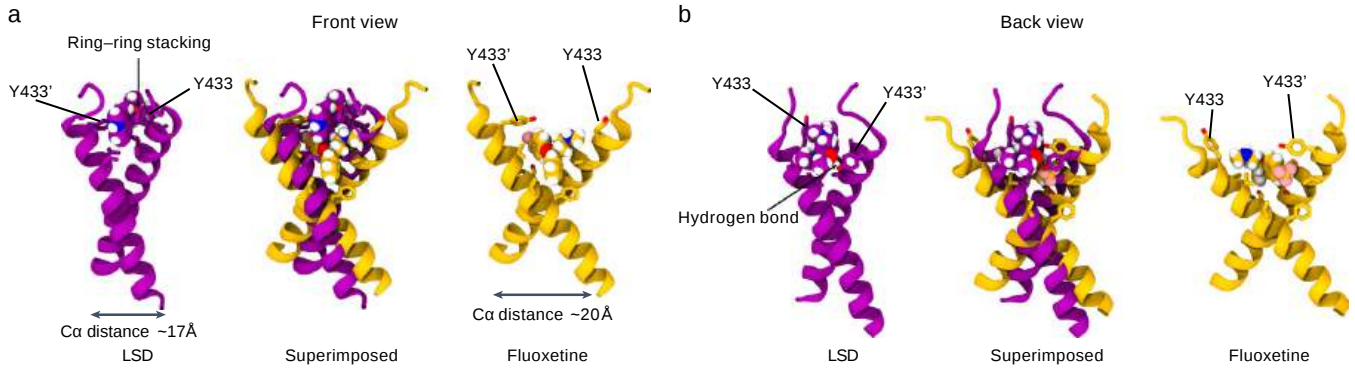


Figure 28.9.18 **Different TrkB binding modes of LSD and fluoxetine.** Moliner et al., ibid.

Panels a,**b**, show representative snapshots of atomistic MD simulations showing the front (**a**) and back (**b**) views of the binding pockets for LSD (purple) and fluoxetine (yellow) in the extracellular-facing crevice of TrkB TMD dimers. Side chains of relevant binding site residues are displayed. Superimposed structures of TrkB optimally bound to LSD or fluoxetine reveal that, while some residues involved in binding are shared (Y433 and V437), the binding modes are different. Fluoxetine binds at a site deeper within the dimer, locking the TMD dimers in a more open cross-shaped conformation (distance between the center of mass L451–L453 Cα atoms of each monomer ~20 Å). In contrast, LSD binds closer to the N-terminus of the TrkB TMD and establishes more stable interactions with the dimer: a hydrogen bond between the oxygen atom of the diethylamide group of LSD and the Y433 residue of one monomer, and pi-stacking of the aromatic backbone of the drug with the Y433 residue of the second monomer, locking the TMD dimer in a tighter cross-shaped conformation (L451–L453 Cα distance ~17 Å) compared with fluoxetine. Drugs are shown in van der Waals representation. Oxygen, nitrogen, and hydrogen atoms are shown in red, blue, and white, respectively.

28.10: Integrins- Bidirectional Cell Adhesion Receptors

This section was derived almost completely, with some modifications and additions, from the following source: Mezu-Ndubuisi, O.J., Maheshwari, A. The role of integrins in inflammation and angiogenesis. *Pediatr Res* **89,** 1619–1626 (2021). https://doi.org/10.1038/s41390-020-01177-9. https://doi.org/10.1038/s41390-020-01177-9. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/. Consult the original article for specific references.

28.10.1: Introduction

Integrins link the extracellular matrix to the intracellular cytoskeleton through a single transmembrane alpha-helical segment. They work with growth factor receptors to regulate cell survival, cell migration, and cell division. They contain a very large extracellular domain comprising most of the protein and a very small intracellular domain.

Integrins are heterodimeric transmembrane cell adhesion molecules made up of alpha (α) and beta (β) subunits arranged in numerous dimeric pairings. These complexes have varying affinities to extracellular ligands. Integrins regulate cellular growth, proliferation, migration, signaling, and cytokine activation and release and thereby play important roles in cell proliferation and migration, apoptosis, and tissue repair, as well as in all processes critical to inflammation, infection, and angiogenesis.

Integrins are a family of ubiquitous αβ heterodimeric receptors that exist in multiple conformations and interact with a diverse group of ligands. These molecules mediate interactions between cells and these cells with the extracellular matrix (ECM) and thereby serve a critical role in signaling and homeostasis. By facilitating dynamic linkages between the intracellular actin cytoskeleton and the ECM, integrins also transduce both external and internal mechanochemical cues and bi-directional signaling across the plasma membrane. Integrins are involved in a diverse range of body processes, including cellular survival, inflammation, immunity, infection, thrombosis, angiogenesis, and malignancy. In this review, we highlight the structure and function of integrins; the mechanisms involved in integrin activation and signaling; their role in inflammation, infection, and angiogenesis; and discuss current advances in integrin-targeted therapies. Understanding the factors that regulate integrin structure, function, and signaling would enable the identification of new therapeutic targets.

28.10.2: Structure of integrins

In mammals, the family of integrins is comprised of 24 αβ pairs of heterodimeric transmembrane adhesion receptors and cell-surface proteins. These pairings are known to involve 18 α and 8 β subunits as shown in Figure 28.10.1

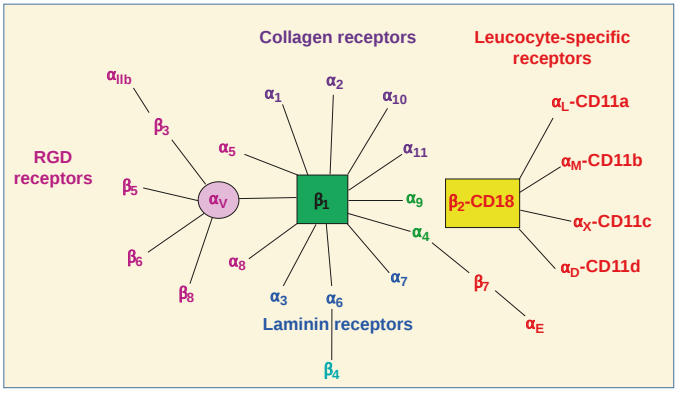


Figure 28.10.1: Integrin heterodimers consist of numerous combinations of α and β subunits.

With respect to ligand specificity, integrins are generally classified as collagen-binding integrins ($\alpha_1\beta_1$, $\alpha_2\beta_1$, $\alpha_{10}\beta_1$, and $\alpha_{11}\beta_1$), Arg-Gly-Asp (RGD)-recognizing integrins ($\alpha_5\beta_1$, $\alpha_V\beta_1$, $\alpha_V\beta_3$, $\alpha_V\beta_5$, $\alpha_V\beta_6$, $\alpha_V\beta_8$, and $\alpha_{IIb}\beta_3$), laminin-binding integrins ($\alpha_3\beta_1$, $\alpha_6\beta_1$, $\alpha_7\beta_1$, and $\alpha_6\beta_4$), and leukocyte integrins ($\alpha_L\beta_2$, $\alpha_M\beta_2$, $\alpha_X\beta_2$, and $\alpha_D\beta_2$).

The right-hand side of Figure 28.10.1 shows that the β_2 integrin subunit (CD18) can pair with one of the four α subunits (α_L-CD11a, α_M-CD11b, α_X-CD11c, and α_D-CD11d), forming leukocyte function-associated antigen-1, Mac1/CR3 (macrophage-1 antigen, complement receptor 3), 150.95/CR4 (complement receptor 4), and CD18/CD11d, respectively.

The structure of the heterodimers and their non-covalent associations are shown in Figure 28.10.2

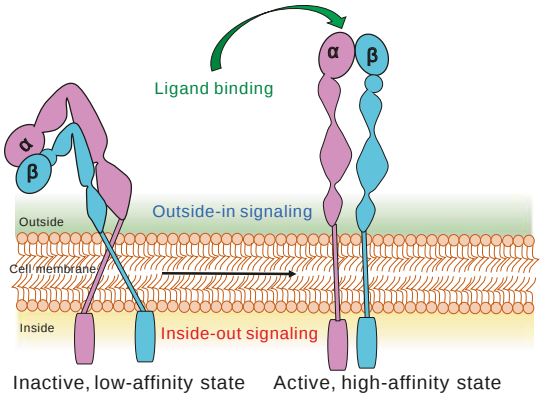


Figure 28.10.2: Structurally, the αβ integrin subunits are type 1 transmembrane proteins.

Each subunit consists of one large multi-domain extracellular segment, one transmembrane helix, and a short cytoplasmic tail. The extracellular region interacts with extracellular matrix (ECM) ligands and is composed of about 1104 residues in the α subunit and 778 residues in the β subunits and shorter cytoplasmic domains with 30–50 residues. The short cytoplasmic tails are composed of 20–70 amino acids and mediate interactions with intracellular cytoskeletal and signaling proteins.

In response to intracellular or extracellular stimuli, integrin activation occurs by ligand binding or by the changes in the cytoplasmic domains, resulting in elongation and separation of the legs. Integrins appear in a closed or “bent” conformation on resting cells and display a low binding affinity for ligands rendering them inactive to ligand binding or signal transduction; while once activated, the integrin shape extends to an open conformation leading to a high affinity. In a closed conformation, integrins show low ligand-binding affinity, partly due to the bend in the center of the α and β subunits, which brings the ligand-binding site within 5 nm of the cell surface. However, when the conformation is open, the two subunits straighten with increased integrin affinity for the ligand. The initial binding of extracellular ligands causes separation of the cytoplasmic domains, allowing interaction with signal transduction and cytoskeletal molecules during outside-in signaling, while separation of the cytoplasmic domains by talin and other activators activates the head to enable ligand binding during inside-out signaling.

The αβ pairings of integrin subunits dictate the specificity of the integrin to a particular ligand, modulate the formation of intracellular adhesion complexes, and regulate downstream signaling. Six α ($\alpha_{1–6}$) and seven β ($\beta_{1–7}$) subunits are known to form several unique αβ subunit associations, as shown in Figure 28.10.1. Interestingly, the earliest discovered integrins, lymphocyte function-associated antigen 1 (integrin $\alpha_L\beta_2$) and macrophage antigen 1 (integrin $\alpha_M\beta_2$), derive their specificity from specific α subunits, but these share the same β subunit.[5]

28.10.3: Integrin α subunit family

The integrin α subunits carry a 200 amino acid “inserted” domain, the I-domain (αI). When present on an integrin, the αI domain is an exclusive ligand-binding site. αI integrins have 13 extracellular domains in 2 subunits, which interact with a variety of ligands. The I-domains are seen in 6 out of the 15 integrin α subunits.

Integrin alpha-1/beta-1 is a receptor for laminin and collagen. It recognizes the proline-hydroxylated sequence G-F-P-G-E-R in collagen. the human α_1 subunit (1179 amino acids) has the domain structure shown in Figure 28.10.3

Figure 28.10.3: Domain structure of human integrin α_1

The green represents the Van Willibrand Factor Type A domain. The middle pinkish domains are FG-GAP extracellular domains. They are repeated up to 7 times in alpha integrins. The reddish domain at the very C-terminus is the transmembrane helix domain (1142-1164). This membrane protein is very different than those we have seen before as it has just a 15 amino acid C-terminal tail exposed in the cytoplasm.

Figure 28.10.4shows an interactive iCn3D model of the predicted structure of human Integrin alpha-1 (AlphaFold, P56199).

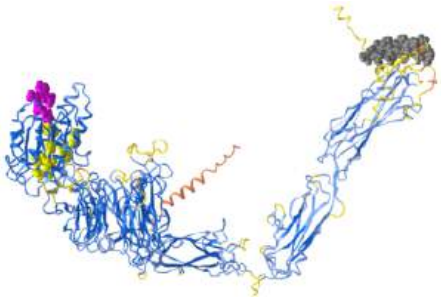


Figure 28.10.4: Predicted structure of human integrin alpha-1 (AlphaFold, P56199) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...9r15PAYYbNbit8

The color coding is gray spacefill for the C-terminal transmembrane helix, yellow spacefill for the inhibitor binding pocket and magenta for the collagen-binding site

28.10.4: Integrin β subunit family

In humans, integrin β subunits have a cytoplasmic tail that has <75 amino acids in length, except the β_4 tail which is about 1000 amino acids long (includes four fibronectin type III repeats). The integrin β tails have one or two NPxY/F motifs (x refers to any amino acid) that recognize protein modules, phosphotyrosine-binding domains, that are involved in several signaling and cytoskeletal proteins at the cytoplasmic face of the plasma membrane through phosphorylation of the tyrosine (Y) in the NPxY/F motif. The integrin β subunit family includes $\beta_{1–7}$, which bind the α subunits in different combinations. The most frequently seen β subunit integrin heterodimers are β_1.

Although β_2 integrins show functional overlap, the corresponding α subunit defines its individual functional properties. The β_2/CD18 chain has also received attention because of its involvement in several inflammatory receptors such as $\alpha_L\beta_2$, lymphocyte function-associated antigen-1 (LFA-1), and the $\alpha_M\beta_2$, Mac-1, complement receptor 3 (CR3). In these β_2 integrins, the α subunits bind specific ligands such as the intercellular adhesion molecules (ICAMs). The non-I-domain α subunits in other integrins, such as the laminin-binding α_3, α_6, and α_7, and others that recognize the arginine (R), glycine (G), aspartic acid (D) (RGD) motif (α_V, α_8, α_5, and α_{IIb}), are also closely related to each other.

- The α subunit of each integrin is the primary determinant of its extracellular ligand specificity.
- The β chain binds acidic residues in ICAMs and cytoplasmic adapters such as paxillin, talins, and kindlins to facilitate cellular adhesions with the ECM. Integrins interact with the actin cytoskeleton through the talin- and kindlin-binding motifs present in the cytoplasmic domains of their β subunits.

28.10.5: Characteristics of specific integrin heterodimers

Integrin αβ heterodimers are divided into four classes (leukocyte, collagen-binding, Arg-Gly-Asp (RGD)-binding, and laminin-binding integrins (as shown in Figure 28.10.1, based on evolutionary associations, ligand specificity, and restricted expression on white blood cells (β_2 and β_7 integrins).

- *Leucocyte integrins* have a common β_2 chain that is linked to CD-18 and binds receptors such as ICAM and plasma proteins such as complement components C3b and C4b.
- *Collagen-binding integrins* have a common β_1 chain that binds various α chains in integrins $\alpha_1\beta_1$, $\alpha_2\beta_1$, $\alpha_{10}\beta_1$, and $\alpha_{11}\beta_1$. The $\alpha_2\beta_1$ integrin binds its primary ligand, collagen, and chondroadherin, a matrix protein.
- The *RGD-binding integrins* have a common α_V chain or β_1 chain. The RGD peptide motif was first discovered in fibronectin but was later found in several other ECM proteins, such as fibronectin, osteopontin, vitronectin, von Willebrand factor (VWF), and laminin.

Among the 24 human integrin subtypes known to date, eight integrin dimers recognize the tripeptide RGD motif within ECM proteins, namely: $\alpha_V\beta_1$, $\alpha_V\beta_3$, $\alpha_V\beta_5$, $\alpha_V\beta_6$, $\alpha_V\beta_8$, $\alpha_5\beta_1$, $\alpha_8\beta_1$, and $\alpha_{IIb}\beta_3$. Laminin-binding integrins ($\alpha_3\beta_1$, $\boldsymbol{\alpha_6\beta_1}$, $\alpha_7\beta_1$, and $\alpha_6\beta_4$) mediate the adhesion of cells to basement membranes in various tissues. The $\alpha_4\beta_1$, $\alpha_9\beta_1$, and $\alpha_4\beta_7$ integrin family binds fibronectin in an RGD-independent manner.

Figure 28.10.5shows an interactive iCn3D model of the structure of **$\alpha_6\beta_1$** integrin in complex with laminin-511 (7CEC)

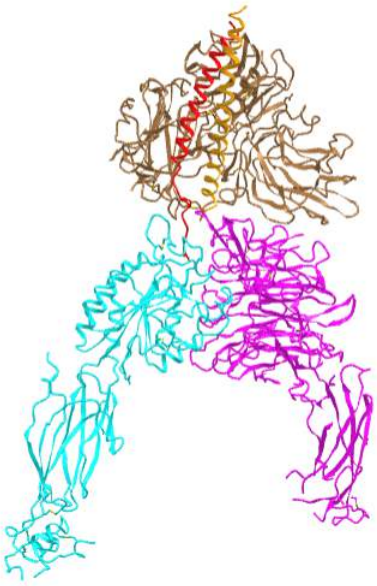


Figure 28.10.5: Structure of $\alpha_6\beta_1$ integrin in complex with laminin-511 (7CEC) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...CxCTt6caaWrPd6

The cyan chain is integrin α_6 and the magenta chain is integrin β_1. The laminin chains are α (brown), β (orange), and γ (red). Note that each of the laminin subunits interacts with the $\alpha_6\beta_1$ integrin. Two carboxylates on the C-terminal region of the laminin γ chain interact with the metal ion-dependent binding sites on the integrin beta subunit and an Asp 189 in the alpha subunit.

28.10.6: Integrin–ligand binding and consequent activation

The structure and function of integrins are complex. Integrins bind numerous extracellular ligands, intracellular signaling molecules, and the cytoskeleton in a bivalent-cation-dependent manner with varying specificities. Integrins also have many states with multiple conformations and affinities.

28.10.6.1: Mechanism of integrin-ligand binding and conformational states

Integrins bind cell-surface ligands to promote cellular interactions with the ECM and with other cells in the transduction of complex signals that modulate many cellular processes, such as adhesion, migration, and differentiation. These soluble, ECM, or cell surface-bound ligands may include growth factors, structural constituents of the ECM, proteases, cytokines, plasma proteins, microbial pathogens, or receptors specific to immune cells. The affinity and avidity of a ligand may change actively by inside-out signaling in specific pathways. Ligand affinity may vary with the strength of interaction and dissociation of a monovalent protein and its ligand, where ligand avidity refers to its ability to form multiple combinations of bonds.

Integrins exist primarily in three conformational states: bent–closed (inactive; the predominant state), extended–closed (active; low affinity or intermediate state), and extended–open (active; high affinity). The affinity of integrins to various inhibitory and stimulatory ligands is modulated by bivalent cations, which induce a range of conformational changes in integrins ranging from a folded, inactive, and low-affinity state to a high-affinity conformation as shown in Figure 28.10.2 These conformational changes in the extracellular domains of integrins modulate both ligand binding and downstream cellular signaling.

28.10.6.2: Integrin activation

The activation of integrins increases the affinity of these molecules to extracellular ligands. Integrin tail domains play a critical role in these steps, and any genetic mutations in these parts of integrins can disrupt downstream intracellular signaling. Integrin-mediated signaling across cell membranes is typically bi-directional and termed “outside-in” and “inside-out” signaling. When integrins interact with ECM ligands, a conformational change allows adherence to downstream adaptor molecules in the cell-membrane plane. Once clustered, integrins can recruit and activate kinases such as Src family kinases, focal adhesion, and scaffold molecules such as the adaptor protein p130CRK-associated substrate/breast cancer anti-estrogen resistance 1 (p130CAS/BCAR1). These integrin-associated complexes include discrete active and inactive integrin organizations, which can activate unique signaling pathways.

The extracellular domains of integrins are known to undergo a diverse range of conformational changes that alter the ligand-binding domains. In the cytoplasmic tails of integrins, α-helices are seen as heterodimers, and the β-strands often bind intracellular proteins, such as talin or filamin. The cytoplasmic tail may undergo several specific conformational changes to bind a range of other signal transducers.

This section is derived from Mechanobiology. https://www.mechanobio.info/what-is-mechanosignaling/what-is-the-extracellular-matrix-and-the-basal-lamina/what-is-integrin/how-is-integrin-activated/. Creative Commons Attribution-NonCommercial 4.0 International License.

Integrin can be activated from two directions, from the inside by the regulated binding of proteins to the cytoplasmic tails, and from the outside by multivalent ligand binding. In either case, **talin** binding to the integrin β tails is an essential and the final common step. Though the two processes are conceptually separate, they are mutually cooperative i.e one can lead to the other. Some structural studies done with force application to mimic ligand/intracellular protein suggested that the combined action of these two events favors the transition from the closed, low affinity to an open, high-affinity conformation of integrin. Activation leads to bidirectional signaling crucial in a variety of anchorage-dependent events such as adhesion, cell spreading, migration, polarity, and organization of the ECM leading to physiological changes. Figure 28.10.6shows the different states of the integrin dimer in inside-out and outside-in signaling.

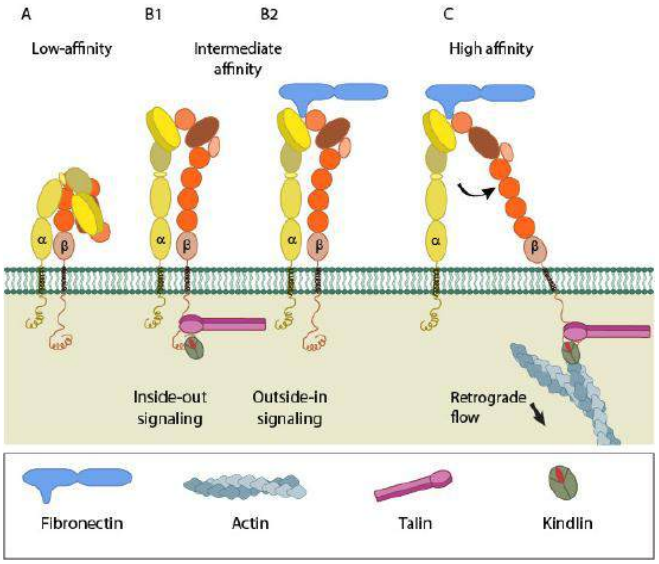


Figure 28.10.6 Integrin dimer in inside-out and outside-in signaling.

(A) shows integrin in a low affinity, inactive, bent, conformation. (B1) and (B2) show inside-out integrin activation by cytoplasmic proteins or outside-in integrin activation via extracellular matrix (ECM) ligands, both of which lead to the complete extension of the extracellular domains. (C) shows high affinity and active integrin characterized by separation of the cytoplasmic leg domains.

28.10.7: Inside-out signaling

Signals received by other receptors foster the binding of **talin** and **kindlin** to the cytoplasmic end of the **integrin β** subunit, at sites of actin polymerization. Substantial information on signaling pathway leading activation is available for integrin αIIbβ3.

Talin binds to integrin β-tail via the F3 phospho-tyrosine binding (PTB) domain, a unique interaction with the membrane-proximal (MP) region of the integrin (NPxY motif). This permits competition between a conserved lysine on talin and an aspartic acid on integrin α essential for α/β salt bridge disruption and sufficient for integrin activation. Addition interactions through the basic patches in the FERM subdomain F2 help to orient the β-subunit to promote spatial separation of the cytoplasmic domains.

Kindlin is also an essential co-activator of integrin and binds to a membrane distal NxxY motif on β-integrin via its FERM F3 subdomain. A preceding threonine patch on integrins β1 and β3 that gets phosphorylated and a tryptophan on kindlin F3 are also required for binding. However, kindlins are not known to activate integrins on their own but may render integrin-specific effects.

28.10.8: Outside-in signaling

Ligand binding to the external domain causes conformational changes that increase ligand affinity, modify protein-interaction sites in the cytoplasmic domains, and thence the resulting signals.

Besides conformational changes that extend integrin dimers, multivalent ligand binding leads to the clustering of integrins, which in turn activates the **Src family of kinases (SFKs)** by autophosphorylation. SFKs phosphorylate tyrosines of the integrin cytoplasmic domain (NPxY motifs) and other proteins leading to

1. control of ligand binding strength
2. alteration of binding with signaling molecules (kinases, GTPases, and adaptors), that constitute dynamic adhesion structures such as focal adhesions and podosomes

Nevertheless, whether clustering triggers outside-in signaling to facilitate integrin activation or occurs after integrin activation is uncertain.

David G. Menter, Raymond N. DuBois, "Prostaglandins in Cancer Cell Adhesion, Migration, and Invasion", *International Journal of Cell Biology*, vol. 2012, Article ID 723419, 21 pages, 2012. https://doi.org/10.1155/2012/723419. Creative Commons Attribution License,

Let's look at another more detailed representation of integrin states. Each αβ dimer recognizes a different intercellular adhesion molecule (ICAM), ligand, or protein substrate in the basement membrane or extracellular matrix. The α subunit dictates the ligand specificity by a seven-bladed β-propeller head domain connected to a leg support structure made of a "thigh", a "calf"-1, a "calf"-2, a transmembrane, and a cytoplasmic domain. The β subunit interacts with the cell cytoskeleton and contains an N-terminal plexin-semaphorin-integrin (PSI) domain, a hybrid domain, a βI domain, four cysteine-rich epidermal growth factor (EGF) repeats, a transmembrane, and a cytoplasmic domain.

In many cases, the N-terminal β-I domain of a β subunit inserts into the 7-bladed β-propeller domain of an α subunit (α1, α2, α10, α11, αL, αM, αX, and αD) to form a bulbous-binding headpiece complex. The formation of integrin receptor complexes depends on divalent cation (i.e., Ca^{2+}, Mn^{2+}, Mg^{2+}) that bind to metal-ion-dependent adhesion site (MIDAS) motifs in the α subunits and adjacent to MIDAS (ADMIDAS) motifs in β subunits found in the N-terminus of these receptors. Together they joined α and β subunit termini form an N-terminal **headpiece**. These detailed features of the integrin dimer structure are shown in Figure 28.10.7.

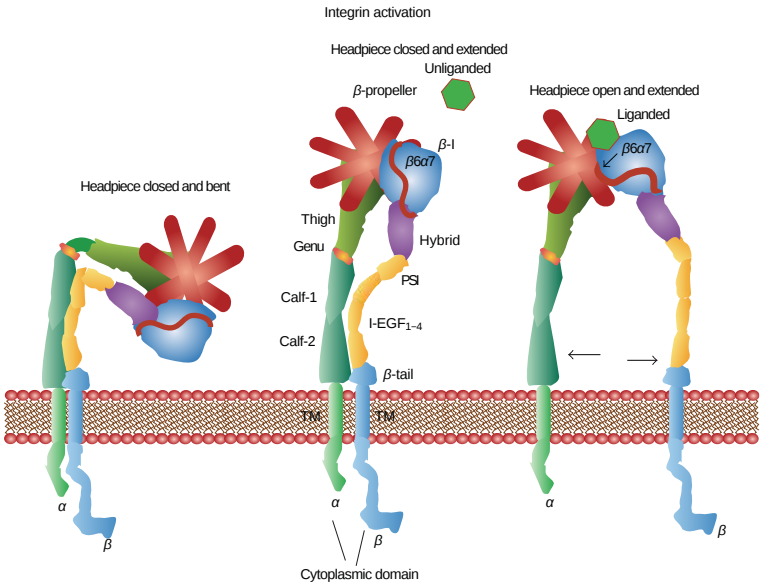


Figure 28.10.7: Detailed domain/motif structures in integrin dimers. Menter and DuBois. https://doi.org/10.1155/2012/723419. Creative Commons Attribution License,

Three conformation states exist for α and β subunit complexes. (1) The unliganded conformation has a closed headpiece and a bent receptor structure with the EGF domains of the β-subunit touching the calf-1-calf-2 domains of the α-subunit. (2) The headpiece remains closed, but structural changes in the β-subunit EGF domains cause a separation from the calf-1-calf-2 domains of the α-subunits causing an extended structure. (3) Conformational changes in the β_6-α_7 loops expose the ligand-binding site along with a

complete separation of the β-subunit from the calf-1-calf-2 domains in the α-subunit. These conformational changes engage the specific integrin headpiece with its ligand"

Figure 28.10.8 shows an interactive iCn3D model of the headpiece of integrin $\alpha_{IIb}\beta_3$ in the headpiece extended and open conformation (3FCU)

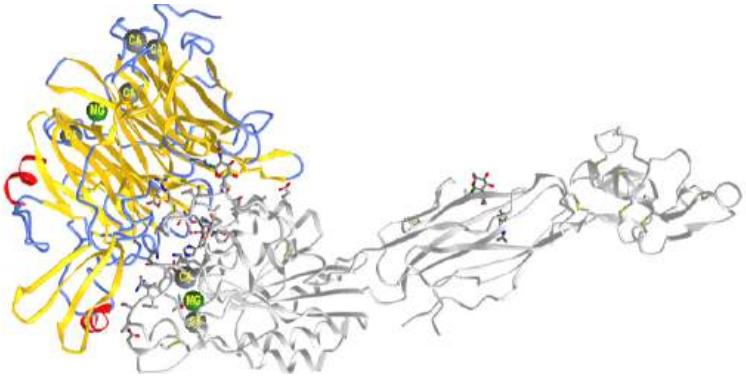


Figure 28.10.8: Headpiece of integrin $\alpha_{IIb}\beta_3$ in the headpiece extended and open conformation (3FCU) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...sG7piAXGhipg4A

The α_{IIb} part of the headpiece is shown colored based on the secondary structure with the yellow sheets comprising the beta propellor secondary structure motif. The gray chain is the β_3 chain. The side chains of the β_3 chain forming the binding interface between the α_{IIb} and β_3 chains are displayed as colored sticks. The three close metal ions (2 Ca^{2+} and 1 Mg^{2+}) are important for ligand binding with the Mg^{2+} involved in coordination to acidic side chains of integrin ligands. These metal ions are present before ligand binding. The RGD binding motif of some integrin ligands binds through their aspartate to the Mg^{2+}. Without bound Mg^{2+}, acidic side chains around the site would interfere with binding.

It appears that lateral forces are most important in activating integrins. This is in contrast to tensile forces which act along the length of the receptor. Tensile forces appear to stabilize the closed, extended low-affinity form, while lateral forces at the beta subunit, a mimic for moving cytoskeletal filaments, stabilized the open, extended, high-affinity form. This links conformational allosteric changes to cytoskeletal changes. The mechanism for activation is hence mechanochemical.

Similar to conventional cell surface signal transducing receptors, integrins bind ligands and transmit information in an “outside-in signaling” as shown in the top panel of Figure 28.10.9 “Outside-in signaling” behavior typically involves the engagement of integrins with the extracellular matrix or ICAM surface receptors. When external factors bind to exposed ligand binding sites on integrins this results in conformational changes described in the previous section. Most ECM proteins exhibit multivalent or recurrent molecular patterns, which trigger integrin clustering. As cells engage the repetitive patterns in the ECM, these events occur simultaneously thereby activating intracellular signals. The myriad of different extracellular signals that cells encounter in their microenvironment mediates cell polarity, cytoskeletal structure, adhesion, migration, invasion, gene expression, cell survival, and proliferation.

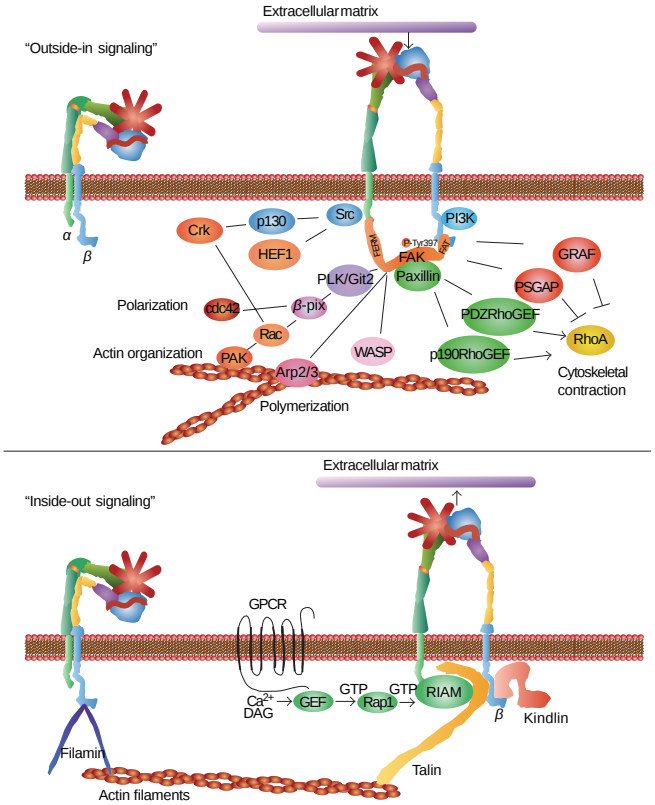


Figure 28.10.9: Detailed view of "outside-in" and "inside-out" integrin signaling. Menter and DuBois, ibid

The "outside-in" binding of ECM ligands to cell surface integrins stimulates conformational changes that activate focal adhesion kinase (FAK). FAK then is autophosphorylated on Tyrosine 397 near the catalytic domain, which binds Src. FAK contains a central kinase domain bordered by FERM (protein 4.1, ezrin, radixin, and moesin homology) domain at the N-terminus and a focal adhesion targeting (FAT) sequence at the C-terminus. Activated Src interacts with human enhancer of filamentation1 (HEF1) and p130 CRK-associated substrate (p130CAS) scaffold proteins that help to positively regulate Src-FAK-Crk interactions with Rac. FAK also activates (PKL/Git2)-β-Pix complexes and β-pix then serves either as an exchange factor for Cdc42 or a scaffold protein to promote signaling via Rac and p21-activated protein kinases (PAK). FAK also interacts with actin-related proteins (ARP2 and ARP3) which are regulated by the Wiskott-Aldrich Syndrome Protein (WASP). ARP2/ARP3 initiates the polymerization of new actin filaments. FAK also influences actin contraction and polarization through another GTPase protein, Rho. The regulation of Rho GTPase hydrolysis of GTP (active) to GDP (inactive) form occurs through the opposing activities of guanine nucleotide exchange factor (GEFs). The GTPase regulator associated with FAK (GRAF) and p190RhoGAP blocks actin cytoskeleton changes. In contrast, PDZRhoGEF and p190RhoGEF both serve to activate Rho. "Outside-in signaling" transfers integrin-mediated external signals to the inside of cells."Inside-out signaling" depends on talin and kindlin. Both talin and kindlin contain FERM (4.1/ezrin/radixin/moesin) domains and a highly conserved C-terminal F3 domains. Talins bind β integrin, actin through the C-terminus, and also vinculin. Kindlins bind integrins, the cell membrane, and various actin adaptor proteins like migfilin, or integrin-linked kinase (ILK). Talin activation occurs through G-protein-coupled receptors that increase cytoplasmic Ca^{2+} and diacylglycerol. This activates GEF function in conjunction with Ras-proximate-1/Ras-related-protein-1-(Rap1-) GTPase. Rap1 then binds to the Rap1-GTP-interacting adaptor molecule (RIAM). RIAM recruits talin to the membrane and the α and β integrin cytoplasmic domains. Kindlin interacts with β integrin cytoplasmic domain stabilizing the activated state of the integrin complex. "Inside-out signaling" strengthens adhesive contacts and the appropriate force necessary for integrin-mediated cell migration, invasion, ECM remodeling, and matrix assembly.

In the case of "outside-in signaling" initiated by ECM proteins, a single ligand-binding event can trigger integrin activation, but repetitive regularly spaced molecular patterns provide a more effective stimulus [122, 123]. This type of mechanoreception has been explored using nanopatterned molecular printing techniques that form regular cRGDfK patch spacings on a polyethylene glycol background matrix [122–125]. These adhesion-dependent sensory mechanisms lead to signal transduction inside the cell by

the activation of multiple pathways. Focal adhesions are often formed as a result of cell interactions with the ECM substrata, which initiate signal transduction via kinase cascades and other mechanisms.

28.10.9: Integrins in inflammation and infection

In the resting state, β_2 integrins are expressed specifically on leucocyte receptors. During inflammation, the inflammatory cytokines activate these integrins and promote cellular adherence to the counter-receptors such as **ICAMs** and promote phagocytosis and cytotoxic killing. Integrin receptors on leukocytes, such as the macrophage-1 antigen (Mac-1, also known as CR3, $\alpha_M\beta_2$, CD11b/CD18) interact with platelet antigens such as the glycoprotein Ibα (GPIbα) during inflammation. Integrins bind to the pro-domain of transforming growth factor (TGF)-β_1 to activate it and promote its secretion. The pro-TGF-βs are biosynthesized and stored in tissues in latent forms, and integrins $\alpha_V\beta_6$ and $\alpha_V\beta_8$ can uniquely bind and activate pro-TGF-β_1 and pro-TGF-β_3. The $\alpha_V\beta_6$ integrin is known to specifically bind the RGDLXXL/I motif in TGF-β_1 and TGF-β_3.

β_2 integrins promote the recruitment of leukocytes to the sites of inflammation by promoting the adhesion of circulating leukocytes to vascular endothelium, transendothelial migration, the formation of immunological synapses in leucocytes, and inflammatory signaling in involved cells. β integrins on the leukocyte surface are also involved in the tethering, rolling, and adhesion of leukocytes to activated endothelial cells. β_2 integrins can also initiate intracellular signaling pathways in macrophages and neutrophils and stimulate cytokine secretion from these cells either directly or in synergy with Toll-like receptors (TLRs). Integrins may also integrate the impact of the epidermal growth factor receptor, platelet-derived growth factor receptor, insulin receptor, met receptor superfamily (hepatocyte growth factor receptor), and the vascular endothelial growth factor receptor (VEGFR) in inflammatory cells.

β_2 integrins are important regulators of adhesion, leukocyte recruitment, and immunological signaling. These integrins mediate adhesive interactions between myeloid cells, endothelial cells, antigen-presenting cells, T cells, and the ECM. **L-selectin**, the CCR7 chemokine receptor, interacts with specific carbohydrate epitopes on the endothelium and promotes leukocyte rolling and transmigration through the vascular endothelium. Leukocyte rolling induces a rapid, although a transient, increase in the affinity of the β1 and β2 integrins to the endothelial ligands. Conformational changes in the structure of the inserted (I) domain of the α_L subunit of LFA-1 enhance firm leukocyte adhesion under shear flow.

28.11: Signaling by Steroid Hormones

28.11.1: Introduction

We will now consider signaling by steroid hormones. These hormones are derivatives of cholesterol, which is found in membrane bilayers and lipoproteins. They are mostly nonpolar. Steroid hormones can affect signaling in two major ways:

- through binding to membrane receptors, which when occupied affect signaling through the myriad of ways we discussed throughout this chapter. These effects would be rapid.
- through binding to cytoplasmic receptors after they diffuse into the cell through passive and active means. This signaling is hence similar to retrograde signaling by nitric oxide, which can passively diffuse out of a cell and enter an adjacent cell to effect signaling in that cell. If the steroid primary messenger is in the cell, it most often can enter the nucleus and regulate gene transcription. Binding to cytoplasmic receptors account for most of the biological effects of steroid hormones. Since transcription is involved, the pathways elicit a slower response.

We will briefly discuss the first type of signaling (binding to membrane receptors) mostly by presenting figures which describe their signaling pathways. Then, we will focus on steroid hormone activation of gene transcription.

There are many classes of steroid hormones. These are illustrated in Figure 28.11.1along with their overall synthetic pathway.

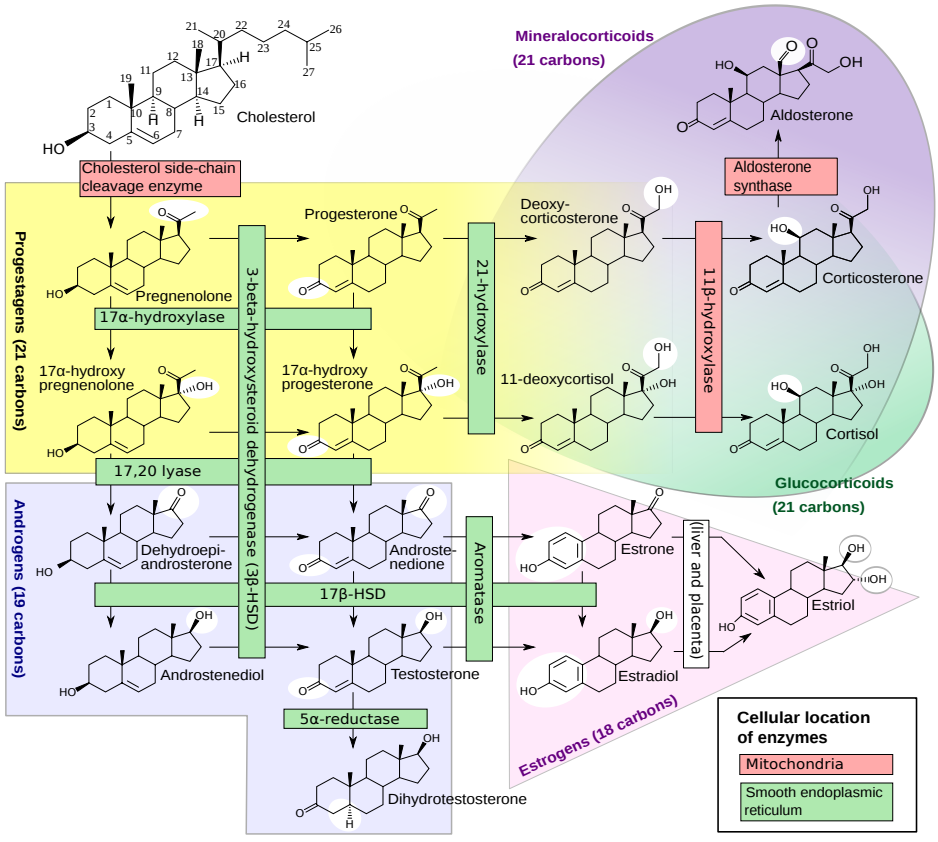


Figure 28.11.1: Diagram of the pathways of human steroidogenesis. *WikiJournal of Medicine* **1** (1). DOI:10.15347/wjm/2014.005. ISSN 20018762.

The overall families include progestogens and estrogens (female sex hormones), androgens (male sex hormones), glucocorticoids (like cortisol which affects immune and metabolic systems), and mineral corticoids (which affect salt/water balance).

We will mostly use the estrogens in the section as a prototypical example because of their involvement in breast cancer and epidermal growth factor receptors.

28.11.2: Steroid hormone signaling through binding to membrane receptors

Most of this subsection is taken from and adapted from the following source: *Masi et al. Cells* **2021**, *10*(11), 2999; **https://doi.org/10.3390/cells10112999.** Creative Commons Attribution (CC BY) license (**https://creativecommons.org/licenses/by/4.0/**).

Most people think of steroid signaling, it is through their nuclear effects on gene transcription (the predominant signaling effect). Signaling at the cell membrane is an emerging area of interest. We will present steroid signaling at the cell nucleus in four figures with their associated captions. We present these lesser-known effects of steroid signaling first since they utilize many of the pathways we have already studied. The figures will also give you a short review of some canonical pathways, which is always good in a field so complicated as signal transduction. The figures focus on the signaling effects in cancer.

Figure 28.11.2shows membrane signaling through androgen receptors.

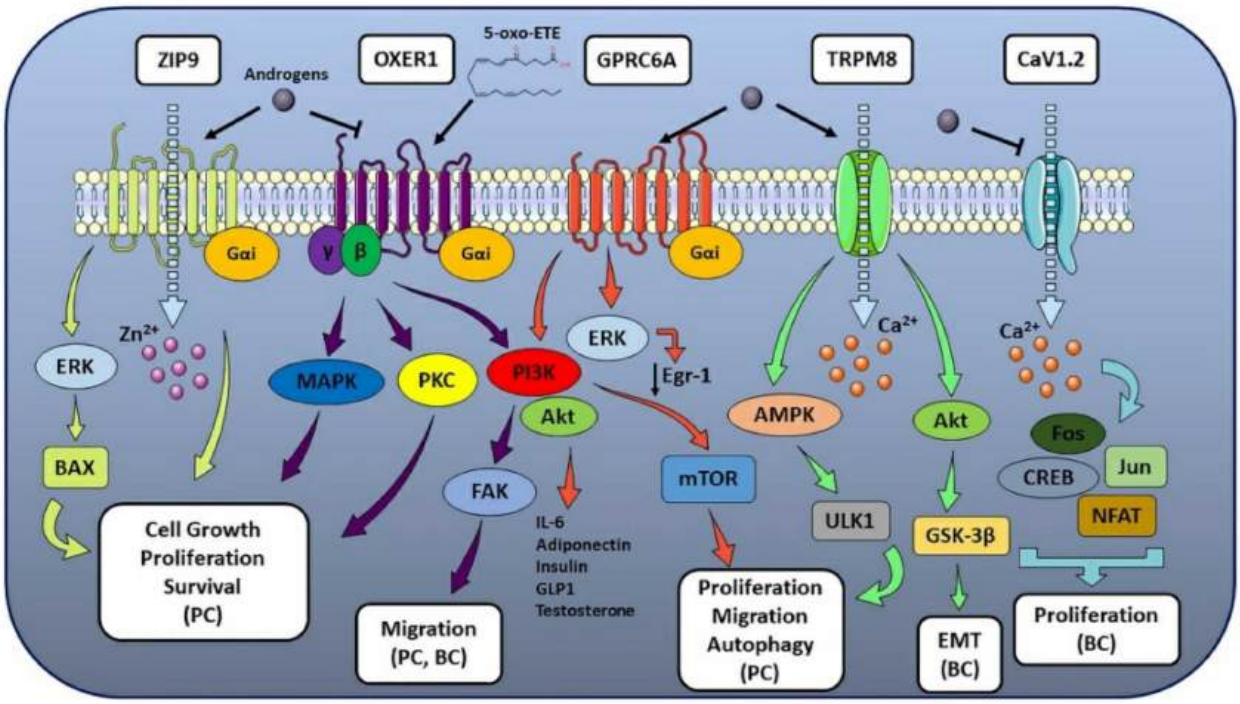


Figure 28.11.2: Membrane androgen receptor-associated pathways and their effect on hormone-sensitive cancer progression (breast cancer as BC, prostate cancer as PC). Receptors are ZIP9 (Zinc transporter), OXER1 (Oxoeicosanoid receptor 1), TRPM8 (Transient receptor potential cation channel subfamily M member 8), CAV1.3 (Voltage-dependent L-type calcium channel subunit alpha-1C). *Masi et al. ibid*

Figure 28.11.3shows membrane signaling through estrogen receptors

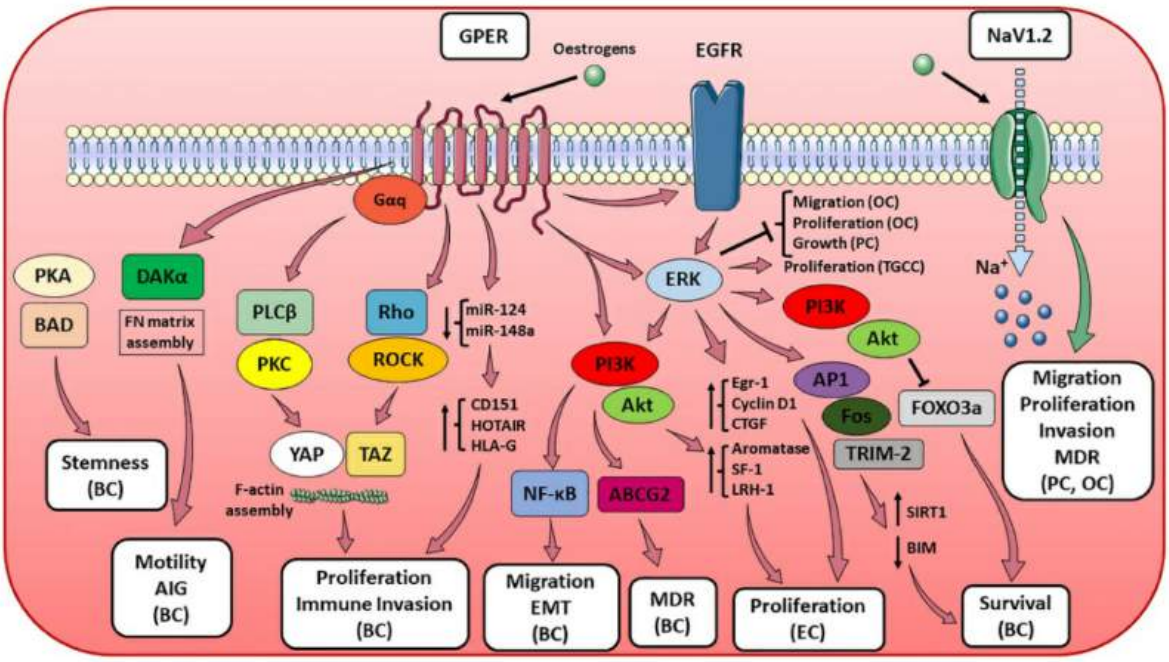


Figure 28.11.3: membrane estrogen receptor-associated pathways and their effect on hormone-sensitive cancer progression (breast cancer as BC, prostate cancer as PC, endometrial cancer as EC, ovarian cancer as OC and testicular germ cell cancer as TGCC). Other abbreviations: multi-drug resistance (MDR), anchorage-independent growth (AIG). Receptors are GPER (G-protein coupled estrogen receptor), NaV1.2 (Sodium channel protein type 2 subunit alpha). Masi et al. Ibid

A tamoxifen (a selective estrogen receptor modulator used in breast cancer therapy) binding site has been found in NavMs voltage-gated sodium channels.

Figure 28.11.4show membrane signaling through membrane progesterone receptors

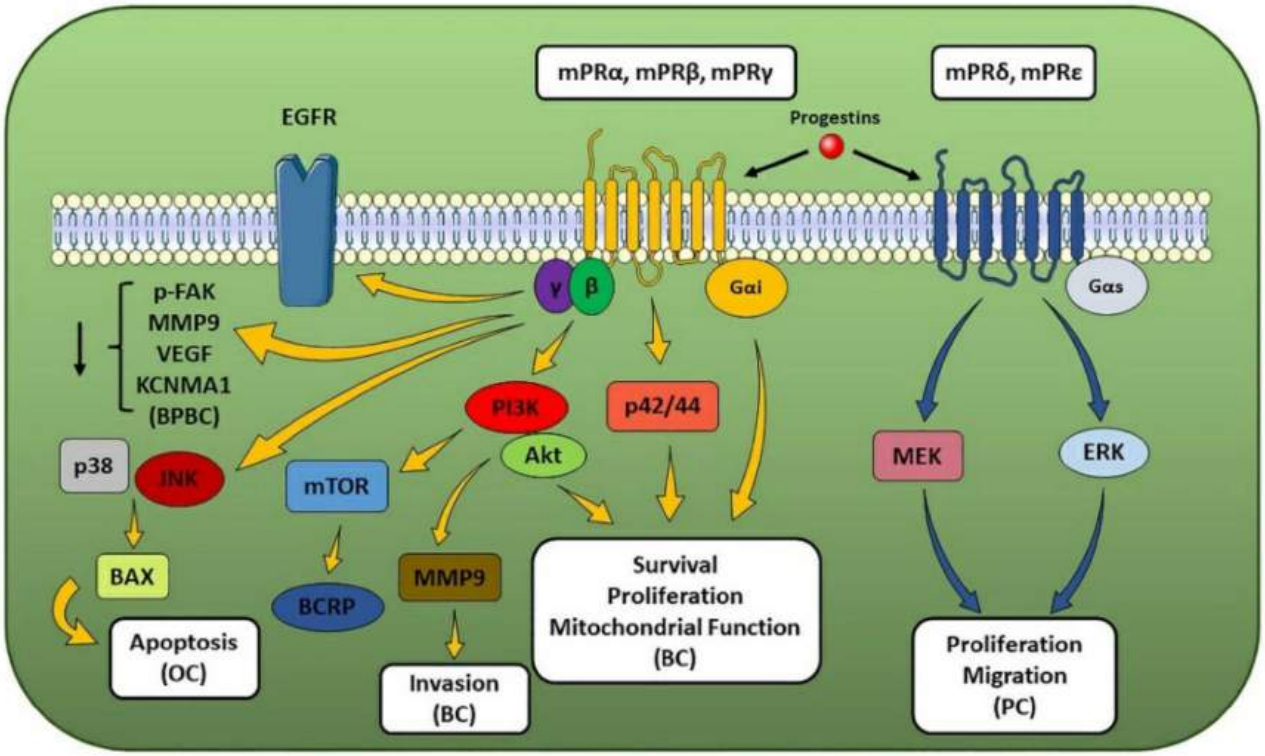


Figure 28.11.4: Membrane progesterone receptor-associated pathways and their effect on hormone-sensitive cancer progression (breast cancer as BC, prostate cancer as PC, ovarian cancer as OC, and basal phenotype breast cancer as BPBC). Receptors are mPR (Membrane progestin receptor). Masi et al. ibid

Figure 28.11.5shows membrane-associated progesterone receptor effects from progesterone signaling.

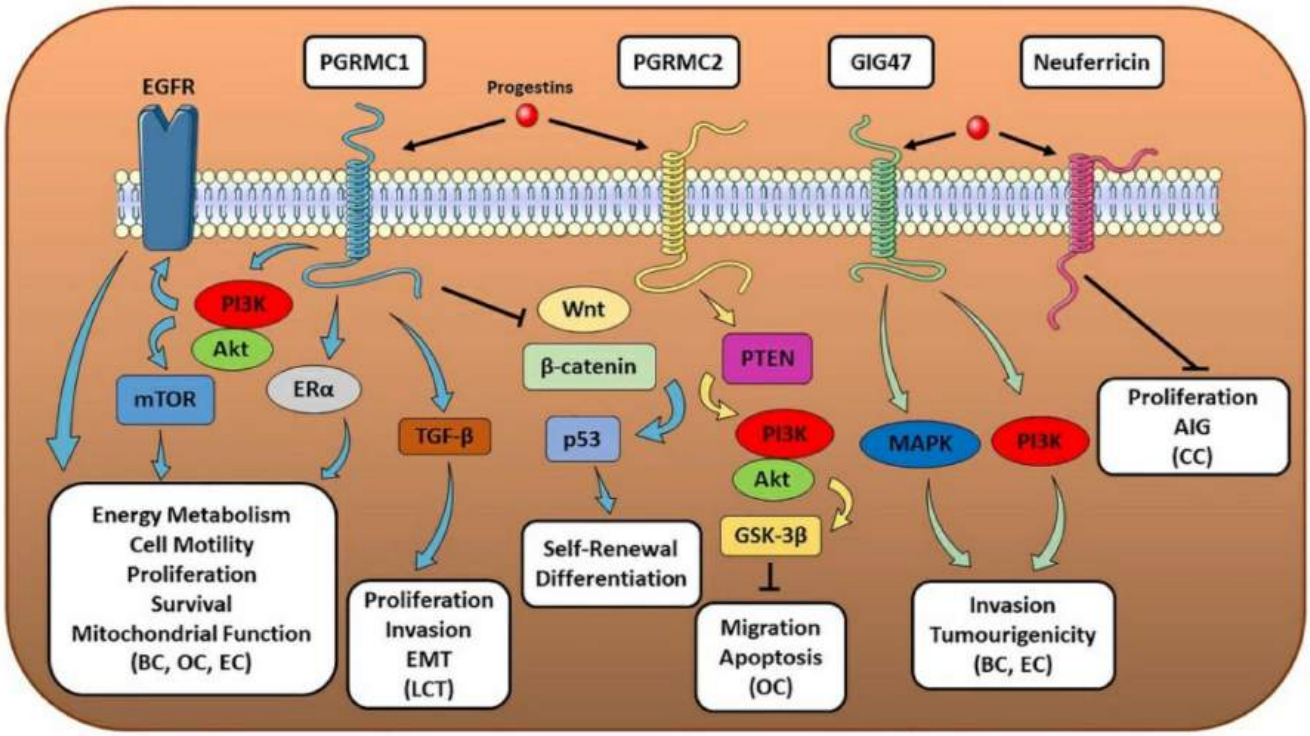


Figure 28.11.5: MAPK-associated pathways breast cancer as BC, endometrial cancer as EC, ovarian cancer as OC, Leydig cell tumor as LCT, and cervical cancer as CC). Other abbreviations: anchorage-independent growth (AIG). Receptors are PGRMV (Membrane-associated progesterone receptor component 1), GIG47 (Neudesin). Masi et al. ibid.

28.11.3: Steroid hormone signaling through binding to cytoplasmic receptors and activation of gene transcription

Now we will explore the major effects of steroid hormones on signaling, in which steroids enter the cell, bind to cytoplasmic receptors, and then translocate into the nucleus where they regulate gene expression. The naming of steroid receptors can be confusing since it is important to differentiate steroid receptors that are resident in the cell membrane and those that move to the nucleus. An added complexity arises from the factor that some nuclear receptors can be covalently modified with a fatty acid (palmitoylated) and targeted to the cell membrane. Estrogen receptors targeted to the membrane can then act independently of their nuclear transcriptional activity.

We will primarily focus on estrogen effects in breast cancer (most diagnosed cancer) that are mediated through the nuclear steroid receptors, which belong to the nuclear receptor superfamily. Tamoxifen, a commonly used drug in the treatment of breast cancer is an estrogen receptor antagonist (also called an estrogen receptor modulator—SERM).

28.11.4: Molecular Function of Steroid Receptors—Common Features

Steroid receptors (SR) consist of four main domains, the C-terminal ligand-binding domain (LBD), the DNA-binding domain (DBD), the hinge region, and amino-terminal domain (NTD). Each SR contains also two motifs called activation function 1 and 2 (AF1 and AF2) within NTD and LBD, respectively, and are crucial for the regulation of gene transcription. Two zinc fingers are located in the DBD. Figure 28.11.6

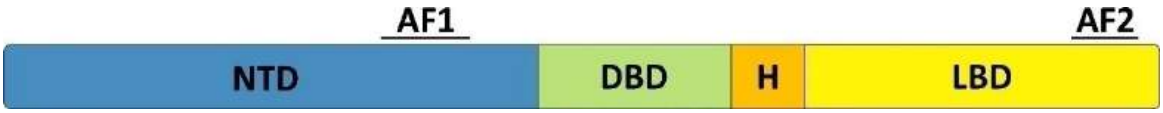


Figure 28.11.6: Schematic illustration of steroid receptor structure. NTD—amino-terminal domain, DBD—DNA-binding domain, H—hinge region, LBD—ligand-binding domain, AF1—activation function 1, AF2—activation function 2. The diagram does not show the exact length proportion of the domains because it differs between distinct SRs. Kowalczyk, W.; Waliszczak, G.; Jach, R.; Dulińska-Litewka, J. Steroid Receptors in Breast Cancer: Understanding of Molecular Function as a Basis for Effective Therapy Development. *Cancers* **2021**, *13*, 4779. https://doi.org/10.3390/cancers13194779. Creative Commons Attribution (CC BY) license (**https://creativecommons.org/licenses/by/4.0/**).

The domains are often labeled A-F. The N-terminal domain (NTD) is also called the A/B domain. It can also bind DNA and can weakly activate transcription in the absence of hormones. The C domain is the DNA binding domain (DBD) containing Zn fingers, which bind to the steroid response element in promoters of key genes. The D domain is the hinge domain, and The E domain binds hormones (like estrogen) and protein regulators and when bound can activate gene transcription. The last domain (F) varies in length and its function is not completely clear.

The specific DNA structure for the estrogen receptor from Pfam is shown in Figure 28.11.7

Figure 28.11.7: Estrogen receptor domain structure.

The green is the N-terminal Oest_recep domain (NTD, A/B). The red zf-C4 is the DNA binding domain which has two Zn fingers that bind DNA (DBD, C). The blue Hormone_Receptor is the ligand (estrogen) binding domain (LBD, domain E). The yellow is the C-terminal domain (F).

These two genes for the estrogen receptor encoding ERα and ERβ. The transcriptionally active form is a dimer that forms on the binding of estrogens. The dimer then translocates to the nucleus and activates transcription at ERE sites. The ERα dimer promotes estrogen-dependent growth while the ERβ dimer inhibits it. Heterodimers can form which seem to reduce the proliferative effects of ERα. Both ERα and ERβ can be expressed in

Figure 28.11.8shows an interactive iCn3D model of the human estrogen receptor computational model (P03372, AlphaFold)

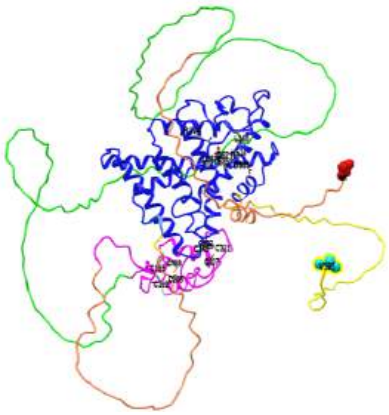


Figure 28.11.8: Full human estrogen receptor computational model (P03372, AlphaFold) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...XAk1jUHred87c9

- red spacefill: N-terminal Met
- green backbone trace: N-terminal domain disordered, which can bind DNA
- magenta backbone trace: DNA binding domain with Zn fingers. The Cys side chains of the Zn fingers are shown in sticks, CPK colors, and labeled.
- blue: Hormone_Receptor is the ligand (estrogen) binding domain (LBD, domain E)
- yellow: C-terminal tail (domain)
- Cyan spacefill: C terminal Val

There are no full-length crystal structures of ER dimer. Most available structures are for the estrogen-binding domain. It's really useful to see the full predicted structure to see how all the domains are connected, but perhaps more interestingly, the extended regions of disordered structures, which you should imagine adopting specific structures on interaction with key signaling partners.

Figure 28.11.9 shows an interactive iCn3D model of the ligand binding domain of human estrogen receptor ERα bound to the antagonist tamoxifen (3ERT).

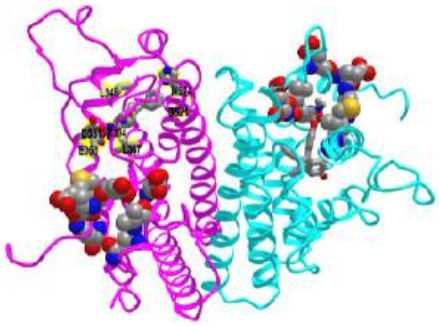


Figure 28.11.9: Ligand binding domain of human estrogen receptor ERα bound to the antagonist tamoxifen (3ERT) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?hhKXF19sVHS5RPys9

The ligand binding domains (one for each of the ERα monomers) are shown in magenta and cyan. The antagonist Tamoxifen, one bound in each of the ligand binding domains, is shown in sticks, CPK color. The amino acids comprising the binding site for tamoxifen, are shown in stick, and CPK colors and labeled in the magenta subunit. The CPK-colored spheres show the binding site on the ligand binding domain for other binding proteins called **corepressors** or **coactivators** (not shown, discussed below).

Figure 28.11.10 shows the structures of estrogens and selective estrogen receptor modulators (SERMs).

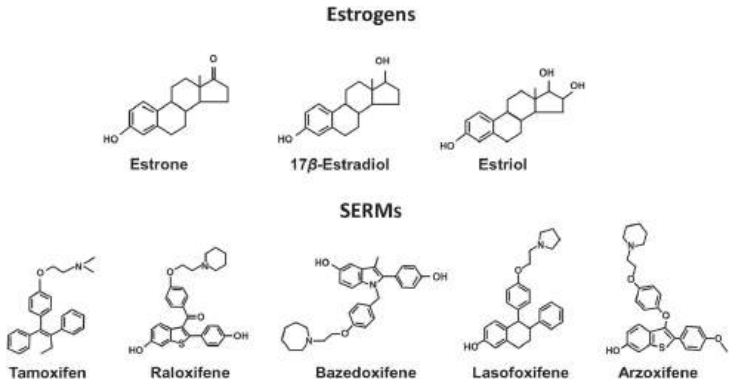


Figure 28.11.10: Structures of estrogens and selective estrogen receptor modulators (SERMs). Smith et al. Breast Cancer Research. 2014, 16:212. http://breast-cancer-research.com/content/16/3/212. CC BY 2.0

The iCn3D model for the human estrogen receptor ERα bound to the antagonist tamoxifen (3ERT) showed binding sites for other proteins called coactivators or corepressors. The ER-estrogen complex, after binding to DNA, can also bind a protein **coactivator**, which activates transcription. Likewise binding of a **corepressor** to the DNA-bound complex inhibits transcriptional activity. Tamoxifen binding to the ERα monomer leads to dimerization and DNA binding. The DNA-bound dimer can then bind either a corepressor (the usual case for tamoxifen binding to ER in the breast), leading to inhibition of DNA transcription (i.e. tamoxifen antagonizes ER transcriptional function), or a coactivator which stimulates gene transcription.

A cartoon diagram illustrating the role of ER coactivators and corepressors is shown in Figure 28.11.11.

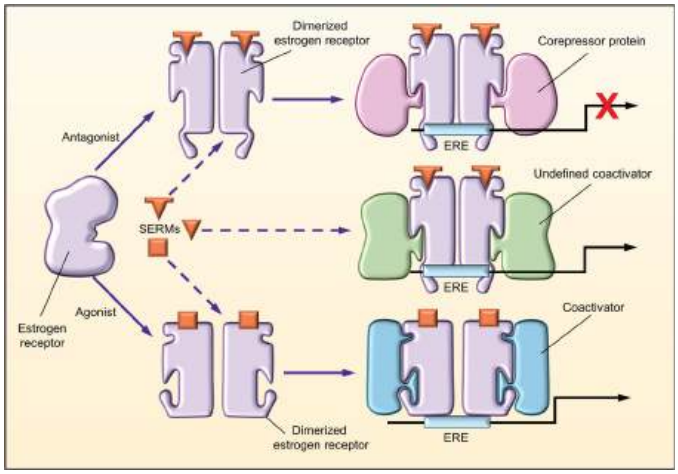


Figure 28.11.11: Molecular activity of selective estrogen receptor modulators (SERMs) at estrogen receptors. Smith et al. ibid

When a SERM binds to the estrogen receptor, the receptor adopts a unique conformation that allows dimerization and interaction with estrogen response elements (EREs) of the target genes. The unique conformational change induced by the binding of the SERM may result in a distinct pattern of cofactor recruitment.

Before steroid binding, most steroid receptors are found in the cytoplasm bound to heat shock proteins like Hsp90. Phosphorylation of the Hsp:SR complex leads to dissociation of the Hsp, followed by dimerization and translocation into the nucleus. In some cases, hormone binding occurs in the nucleus.

Figure 28.11.12shows an interactive iCn3D model of the estrogen receptor DNA-binding domain bound to DNA (1HCQ).

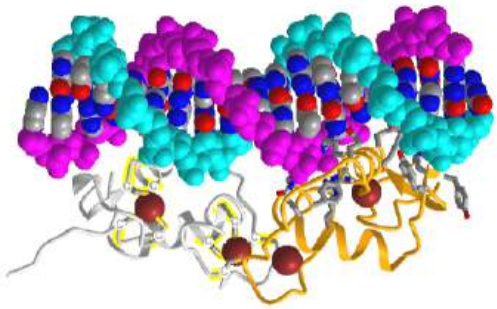


Figure 28.11.12: Estrogen receptor DNA-binding domain bound to DNA (1HCQ) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...rUUhnx8J8ezh8A

The backbones of the dsDNA are shown in spacefill cyan and magenta. The DNA bases are shown in CPK colors. The two chains of the DNA binding domain of the estrogen receptor are shown in gray and gold. Zn^{2+} ions are shown as brown spheres. The coordinating Cys side chains in the gray DNA binding domain are shown in sticks, CPK colors, and labeled "C". The amino acid side chains from the gold DNA binding domain that interact with DNA are shown in sticks CPK colors.

There are two types of ways that steroid hormones activate gene transcription, a direct and an indirect method

Direct (classical): The DNA binding domain (containing the Zn fingers binds) of the dimer bind to the target hormone response element (HRE) or for steroids the steroid response element (SRE) sequences in the promoter site of specific genes under steroid hormone control. As seen in the iCn3D above, one of the two Zn fingers on each of the hormone receptors binds to the target site in the major groove of DNA. The other Zn fingers are involved in the dimerization of the hormone receptor. The SRE contains two 6-base pair repeats separated by 3 base pairs. The DNA sequence shown in the iCn3D above is CCAPGGTCA. The general consensus sequence for steroid hormones is 5′-GGTACAnnnTGTTCT-3′. The ER binds to 5′-GGTCAnnnTGACC-3′. Note that the complementary strand sequence is 5'-GGTCAnnTGACC so the sequence is a palindrome (the complementary strand has the same sequence going in the opposite direction. After binding the hormone:receptor dimer to DNA, coregulators bind. These modify histones and remodel the DNA to facilitate or inhibit transcription.

Indirect: In this method, the steroid receptors bridge other DNA-bound transcription factors without the steroid hormone binding to its response element.

The direct and indirect methods for steroid hormone effects on transcription are shown in Figure 28.11.13

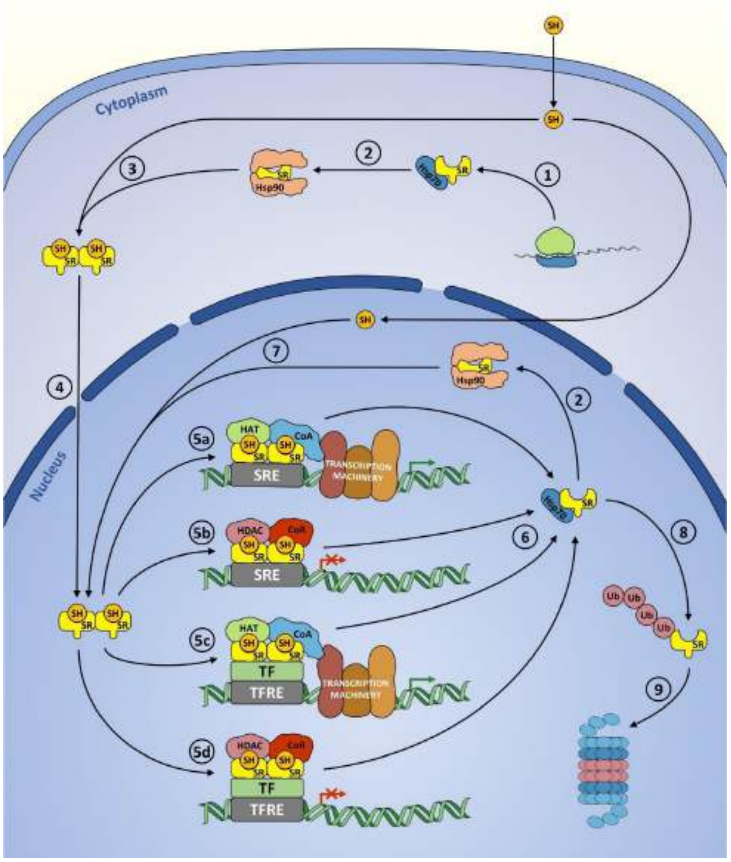


Figure 28.11.13: SR synthesis, genomic signaling, and degradation.

(1) Translation of a SR and binding of Hsp70. (2) Hsp70 to Hsp90 transition. (3) Ligand binding, Hsp90 dissociation, and dimerization. (4) Nuclear translocation. (5) Transcriptional action: induction (5a, 5c) or inhibition (5b, 5d) of target gene expression, performed either in the classical mechanism involving SRE-binding (5a, 5b) or by tethering other TFs (5c, 5d). (6) Ligand dissociation followed by disassembly of the transcriptional complex and SR binding to a molecular chaperone. (7) Rebinding of the ligand. (8) Ubiquitination. (9) Proteasomal degradation. SR—steroid receptor, SH—steroid hormone, Hsp 70—heat shock protein 70, Hsp90—heat shock protein 90, SRE—steroid response element, CoA—coactivators, CoR—corepressors, HAT—histone acetyltransferase, HDAC—histone deacetylase, TF—transcription factor, TFRE—transcription factor response element, Ub—ubiquitin. Although histone acetyltransferases (HATs) and histone deacetylases (HDACs) are classified as coregulators, here they are shown separately to emphasize their role. Illustration created using elements from Servier Medical Art **https://smart.servier.com/**, reproduced under Creative Commons Attribution 3.0 Unported License **https://creativecommons.org/licenses/by/3.0/**.

Hormone dissociation leads to steroid hormone dissociation from DNA by chaperone proteins. The Hsp90-SR complex can reenter the cycle. The steroid hormone receptor can be targeted by proteolysis by the proteasome by ubiquitination in either the nucleus or the cytoplasm.

28.11.5: Modulation of ER function by phosphorylation

No pathways stand in isolation, so it should be no surprise the estrogen receptor (again used as an example) is regulated by post-translational modification, especially by phosphorylation. Phosphorylation can be ligand dependent or independent. Multiple kinases are involved in the phosphorylation of the N-terminal region. An especially important one occurs at Ser 118 at a cyclin-dependent kinase (involved in cell cycling) when the receptor is bound to estradiol. Ser 118 also gets phosphorylation through epidermal growth factor signaling by MAPK. This may increase cell proliferation in breast cancers even in the absence of estrogen. Figure 28.11.14shows the activation of ER by phosphorylation through growth factor and cytokine signaling pathways.

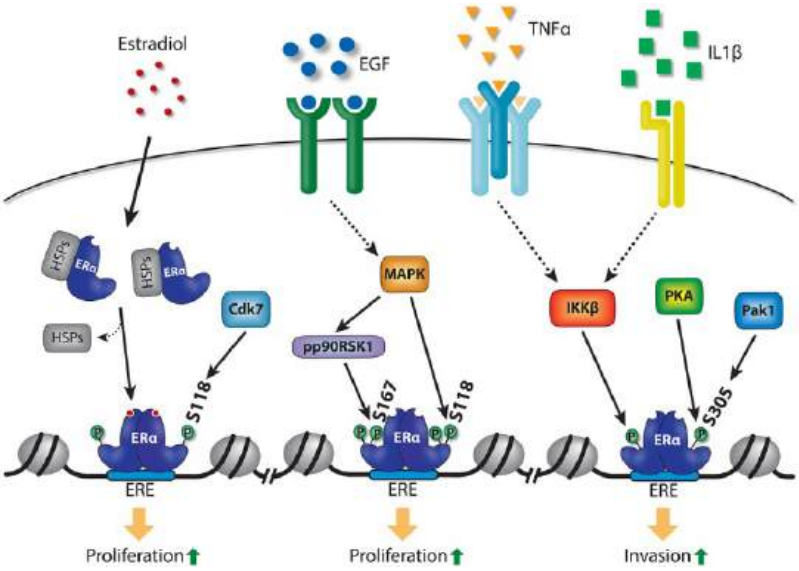


Figure 28.11.14: Activation of ER by phosphorylation induced by growth factor and cytokine signaling pathways. Siersbæk et al ,Genes Dev. 2018 Sep 1; 32(17-18): 1141–1154. doi:10.1101/gad.316646.118. Creative Commons License (Attribution-NonCommercial 4.0 International), as described at http://creativecommons.org/licenses/by-nc/4.0/.

28.12: mTOR and Nutrient Signaling

28.12.1: Introduction: mTOR and AMPK

Imagine when you were in high school, you weighed 135 pounds (61 kg). Let's say it's 50 years later and you still weighed 135 pounds. In the intervening years, think of how much food and liquid you consumed. In 2011, the USA Food and Drug Administration indicated that the average American consumed about 2000 pounds (907 kg) each year, including liquids other than drinking water. (Compare this to 365 kg for someone in Somalia!) So over 50 years you would have consumed 100,000 lbs (45,400 kg) of food or 740 times your body weight. These numbers are essentially unchanged even if you gained one pound a year for a total of 50 pounds.

These figures suggest that we have an elaborate system that regulates how much we eat and how much weight we gain or lose. There are obvious times in our lives when we are growing and actively gaining body mass. Incoming food is not only processed into energy but also into net protein, lipid, carbohydrate, and nucleic acid synthesis. This system has become dysregulated in an ever-increasing number of people with type II diabetes and obesity throughout the world. Obvious candidates for regulators in general of net body weight and in specific of protein and lipid synthesis are the nutrients we consume and store. Many systemic hormones and neurotransmitters are involved in hunger, satiety, and eating behavior. This chapter will not focus on those but rather on mechanisms of nutrient signaling pathways in growth which requires new protein, lipid, and nucleic acid synthesis for cell growth and division. Likewise, it will not focus on nutrient signaling through hexosamines and UDP-GlcNAc.

A key player in these signaling pathways is mTORC (mammalian or mechanistic Target Of Rapamycin Complex). A key protein in this multiprotein complex is mTOR, a Ser/Thr kinase that regulates cell growth, division, protein synthesis, RNA synthesis (transcription), and even autophagy (the major process whereby cells die and their contents are recycled for use). There are two physiologically relevant complexes of mTOR, mTORC1, and mTORC2. These two complexes have been called the master regulators of metabolic and growth processes.

- mTORC1 activates protein, lipid, and nucleotide synthesis, all required for cell growth and division; it is inhibited by rapamycin. For activation, it needs two obvious conditions: energy and growth factors. In addition, it needs amino acids.
- mTORC2 activates many processes through phosphorylation; it is not inhibited by rapamycin.

What is Rapamycin? It sounds like an antibiotic but it is an antifungal agent produced by certain bacteria as a defense against fungal (eukaryotic) pathogens. It blocks cell division in fungi by stopping cell growth. The cell cycle consists of the following general sequential steps: (Go-G1)→S →G2→M→ G1. Gap 0 (Go) is a quiescent phase outside of the cycle. In G1, cells are growing and preparing for DNA synthesis which occurs in the S phase. After DNA replication/synthesis (S phase), cells grow again and prepare for mitotic cell division (M phase). Rapamycin traps fungal cells in the G1 phase. It also traps mammalian cells, and in particular immune lymphocytes in G1 as well, preventing lymphocytes from dividing. Hence rapamycin has been used to prevent rejection of transplanted tissue as it suppresses the immune system. Rapamycin inhibits mTORC1. That it inhibits mTOR is consistent with its immunosuppressive (antigrowth and antiproliferative) effects

Another key player regulated by mTOR is the energy sensor AMP Protein Kinase (AMPK). We will discuss both below.

28.12.2: mTOR Inhibition by Rapamycin

The structure of the mTORC1 complex is complex itself, in part, since some of its components were studied and named before their roles in the mTORC1 complex were elucidated. Investigators were interested in the molecular target(s) of rapamycin. Yeast (a fungus) was an easy organism to study using genetic techniques. Three genes were found that when mutated inhibited the effect of the inhibitor rapamycin (i.e. so that rapamycin did not inhibit mTOR). Two of these were the Ser/Thr kinases mTOR1 and mTOR2. The other was an analog of a protein found in humans (FK Binding Protein 12 - FKBP12). The family of FK binding proteins act as protein chaperones and have Pro-X peptidyl-prolyl isomerase activity (PPI). FKBP12 specifically binds a drug (FK506 also called Tacrolimus) that is an immunosuppressant (do you see the general link with immune cell division and growth?). The binary complex of FK506 and FKBP12 binds a third protein, a phosphatase called calcineurin, and blocks its phosphatase activity and signal transduction required for T cell activation and proliferation. This activity of FKBP12 strangely does not require its PPI activity.

Now back to mTORC1. Here are the known components of the core complex:

- mTOR, a Ser/Thr kinase;
- Raptor, a Regulatory-associated protein of mTOR; modulates the specificity of the kinase;

- mLST8, the mammalian lethal with SEC13 protein 8 (also called Mammalian Lethal With SEC13 Protein 8 e

In addition, other proteins associate with the core complex

- PRAS40, a proline-rich Akt substrate of Akt
- DEPTOR, DEP Domain Containing MTOR-Interacting Protein (where DEP is Disheveled, Egl-10 and Pleckstrin domain found in these 3 proteins and others involved in G-protein signaling)
- FKBP, which binds rapamycin.

mTORC2 is not sensitive to rapamycin. Instead of RAPTOR, it has a protein called RICTOR (rapamycin-insensitive companion of TOR).

Figure 28.12.1 shows an interactive iCn3D model of the human mTORC1 containing mTOR, Raptor and mLST8 bound to FK506 binding protein (FKBP)-rapamycin complex (5FLC)

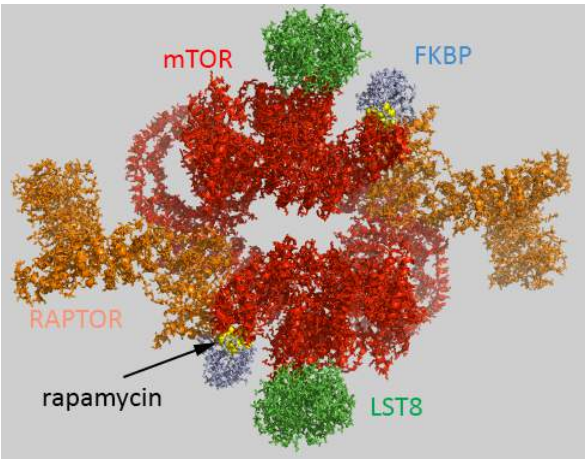


Figure 28.12.1: Human mTORC1 containing mTOR, Raptor and mLST8 bound to FK506 binding protein (FKBP)-rapamycin complex (5FLC). Click the image for a popup or use this external link https://structure.ncbi.nlm.nih.gov/i...iSuYwaymutw867

The complex acts physiologically as a dimer with rotational C2 symmetry. The static image in Figure 28.12.1 is color coded as below. The iCn3D image is similarly.

- chains B ,F ,1-4: pieces of mTOR (mol ID 1-3) as a homodimer - Red;
- chains A and E: Raptor (mol id 4) orange;
- chains D and H: LST8 (mol id 5) green;
- chains C and G: FKBP (mol id 6) light blue;
- rapamycin in C and G, yellow spacefill.

Raptor has been likened to tape as it interacts with the two mTOR subunits, holding them together into a larger, donut-like structure, and stabilizing the dimer.

How does rapamycin inhibit mTOR? You have studied many kinds of inhibitors (competitive, uncompetitive, mixed, or noncompetitive inhibition) in which the inhibitor binds to either free E or the ES complex. The figure above suggests that rapamycin binds at the interface of the mTOR kinase (red) and FKBP. Presumably the RAP:FKBP complex binds to mTOR.

Figure 28.12.2 shows how rapamycin (shown in sticks) is sandwiched between mTOR (red) and FKBP (light blue). This offers clues as to how it inhibits mTOR. The rapamycin:FKBP complex is an allosteric inhibitor with its effect dependent on both substrate and post-translational phosphorylation.

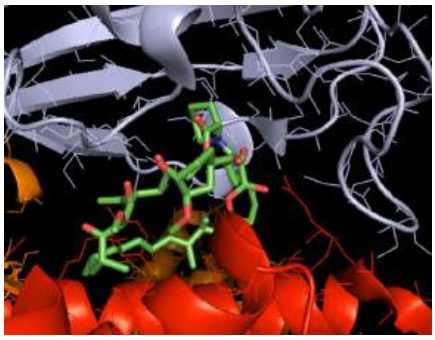


Figure 28.12.2: Rapamycin (shown in sticks) interactions with mTOR (red) and FKBP (light blue)

Right across from the light blue FKBP protein is the green LST8 protein (see above). In between these is a cleft which is the active site. One could imagine that RAP:FKBP binding to mTOR might interfere with substrate binding in mTORC1 but not in mTORC2. RAPTOR in mTORC1 probably helps recruit substrates and hence is involved in the determination of substrate specificity. RAPTOR interacts with a short section called the TOR signaling (TOS) motif in mTORC1 substrates. The part of RAPTOR that binds TOS in substrates is at the base of the mTORC1 active site, probably narrowing it further as it provides exquisite substrate selectivity.

28.12.3: mTOR Protein Kinase Structure and Activity

How does the structure of such an important kinase (mTOR) differ from other kinases? Remember that there are 388 S/T Kinases, 90 Y Kinases, and 40 atypical protein kinases in the human genome. The generic structure is shown below with ATP and substrate binding between N- and C-terminal lobes. The C lobe has a catalytic lobe which contains an Asp side chain acting as a general base in nucleophilic attack on the gamma P of ATP. A disordered activation loop in the C lobe often prevents substrate binding to the enzyme and keeps the kinase in an inactive state. On phosphorylation of the activation loop or elsewhere, or substrate binding, conformational changes lead to movement of the activation loop away from the active site, activating the kinase activity. The structure of generic kinases is shown in Figure 28.12.3

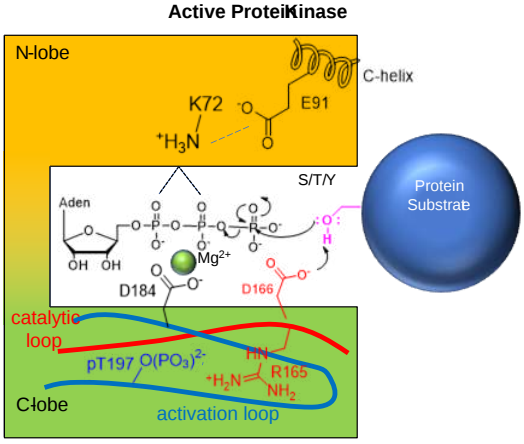


Figure 28.12.3: (Adapted from the superlative text, Cell Signaling by Lim, Mayer, and Pawson)

Compared to generic protein kinases, mTOR has several insertions (about 200 amino acids) into the protein sequence and these must be involved in the determination of its specificity toward protein substrates. The C-terminal domain structure of the mTOR kinase is shown in Figure 28.12.4

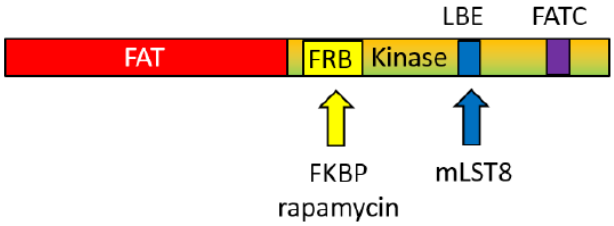


Figure 28.12.4: Domain structure of mTor

The FAT domain precedes the mTOR kinase domain. The FRB is inserted into the N lobe of the kinase domain whereas the LBE and FATC are inserted into the C lobe of the kinase domain.

Figure 28.12.5shows an interactive iCn3D model of the mTOR with the LST8 protein bound through the kinase LBE domain and with bound AGS, a nonhydrolyzable ATP analog (4JSP).

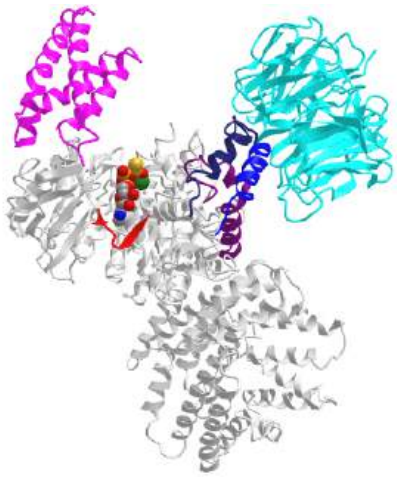


Figure 28.12.5: mTOR - LST8 complex with bound AGS, a nonhydrolyzable ATP analog (4JSP) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...TfzH75NecRodj7

Key regions are:

- FAT domain: 1385-2000
- N Lobe: 2003-2240
- C Lobe: 2241-2549
- FRB domain (magenta) inserted into N lobe: 2021-2118
- LBE domain (medium blue) inserted into C lobe: 2259 (ILL) to 2296 (TAG)
- Catalytic Loop (red): 2337 (GDR) to 2344 (SNL)
- Activation loop (dark blue): 2357 (DFG) to 2379 (FRL)
- FATC domain (purple) inserted in C lobe: 2519 (LDV) to 2549 (PFW)

Now imagine the FKBP:RAMP complex binding to the FRB domain in the figure above and you can easily imagine how RAMP could inhibit a large protein substrate from binding.

A close-up of the active site of the kinase showing bound AGS, the catalytic loop (red) containing the general base Asp 2388, and the activation loop is shown in Figure 28.12.6

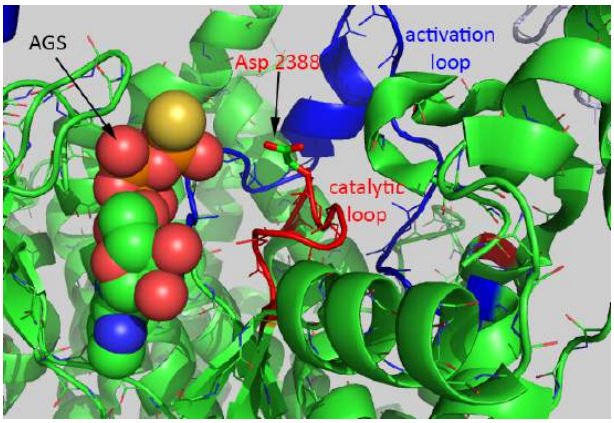


Figure 28.12.6: mTOR with bound AGS, the catalytic loop (red) containing the general base Asp 2388, and the activation loop

The mTOR kinase seems to be always primed for catalysis since the Asp 2388 is in a position to act as a general base. The FRB domain seems to be involved in directing substrate access (such as S6K1) and hence in controlling substrate specificity. Binding of FKBP:RAMP to the FRB domain would prevent substrate binding. Important substrates for each complex are shown below.

mTORC1:

Eukaryotic Translation Initiation Factor 4E Binding Protein 1 (EIF4EBP1): This protein inhibits translation by binding eukaryotic translation initiation factor 4EIF (eIF4E). In the absence of active IF4EBP1, eIF4E is part of a complex that recruits 40S ribosomal subunits to the 5' end of mRNAs, which allows the initiation of translation. The binding protein inhibits complex assembly and represses translation. Active mTORC1 phosphorylates the binding protein in a variety of conditions (UV irradiation and insulin) which leads to dissociation of the binding protein which allows eIF4E to initiate translation.

Ribosomal S6 kinase 1 (RPS6KB1 aka S6K1): This is a Ser/Thr kinase involved in proliferation, protein synthesis, cell growth, and cell proliferation. It phosphorylates eIF4B. In nutrient depletion (non-growth conditions), it forms a complex with the EIF3 translation initiation complex which inhibits translation. Under growth conditions, it is phosphorylated by mTORC1, causing its dissociation from the EIF3 complex and activation of translation. The active form then phosphorylates and activates several substrates in the pre-initiation complex, including the EIF2B complex and the cap-binding complex component EIF4B. In the presence of amino acids, both *EIF4EBP1 and S6K1* are phosphorylated. If amino acids are depleted, they are dephosphorylated.

Lipin 1 (LPIN1): This is a phosphatase that converts phosphatidic acid to diacylglycerol in triglyceride synthesis. Interestingly, it is also a transcriptional coactivator with PPARs (peroxisome proliferator-activated receptors) to modulate genes involved in lipid synthesis.

A summary figure of mTORC1 signaling is shown in Figure 28.12.7.

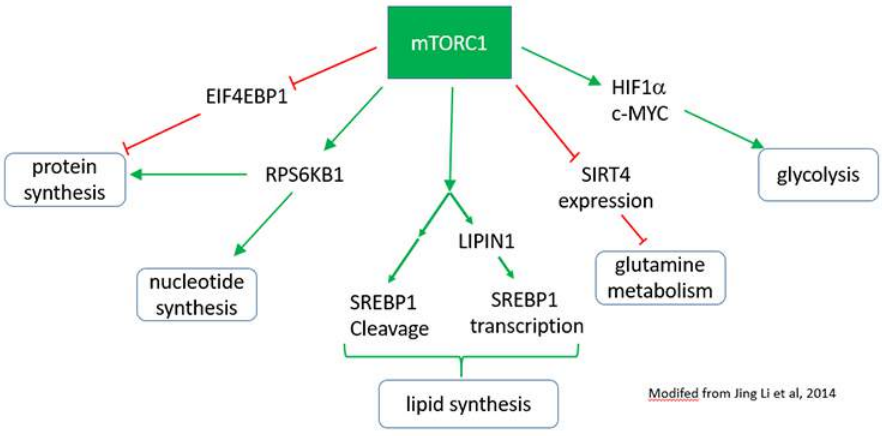


Figure 28.12.7: Summary figuring showing mTORC1 signaling

mTORC2:

Akt (also known as Protein Kinase B): This is a Ser/Thr kinase that is involved in the regulation of metabolism, proliferation, cell survival, growth, and angiogenesis. It has a notable role in the movement of the glucose GLUT4 transporter to the cell membrane in response to insulin signaling. Akt also interacts with mTORC1.

Serum/Glucocorticoid Regulated Kinase 1 (SGK1): This serine/threonine protein kinase is involved in cellular stress response. It activates certain potassium, sodium, and chloride channels. It also activates membrane transporters, enzymes, and transcription factors. Its effects regulate neuronal activity, cell growth, proliferation, survival, migration, and apoptosis.

Protein Kinase C alpha (PRKCA): This is a Ser/Thr kinase involved in cell adhesion, proliferation, differentiation, and migration.

Rho and Rac: These are small G-protein involved in cytoskeletal structure and cell cycle.

Figure 28.12.8 shows a more complete pathway of activation, regulation, and activity of both mTORC1 and mTORC2. Figure 28.12.x below is used with courtesy of Cell Signaling Technologies (www.cellsignal.com). This chapter section will mostly focus on mTORC1.

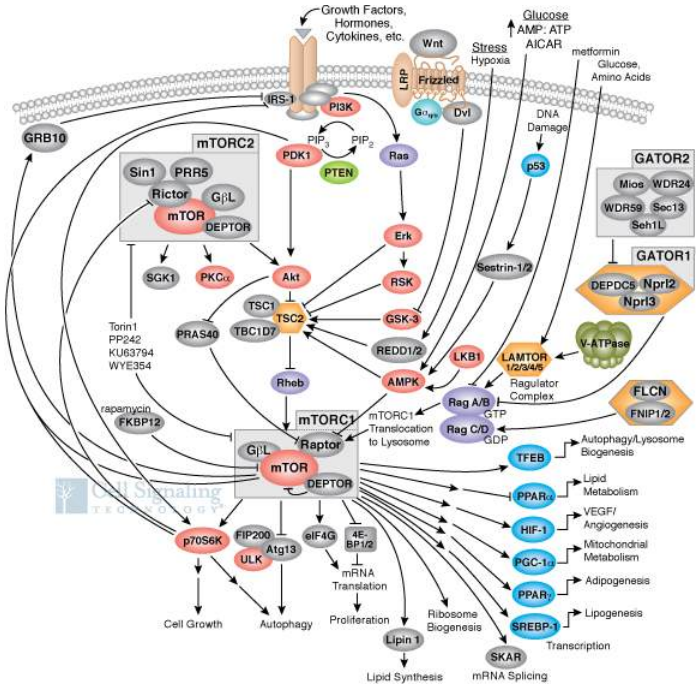


Figure 28.12.8: mTOR1 and mTOR2 signaling.

28.12.4: Regulation of mTORC1 by Leucine

mTORC1 is a key regulator of protein synthesis but that begs the question as to how it determines that protein synthesis is required. How does it sense that? Rregulators of mTORC1 might be amino acids in cells, but who would have thought that the master regulator would be **leucine**, a simple branched chain hydrophobic amino acid.

It would be nice if free leucine bound directly to mTORC1, but it's not that simple. Rather, it binds to a "leucine" receptor, sestrin 2 (SESN2). Figure 28.12.9shows the binding interactions between Leu (spacefill) and key side chains in sestrin 2 (5dj4).

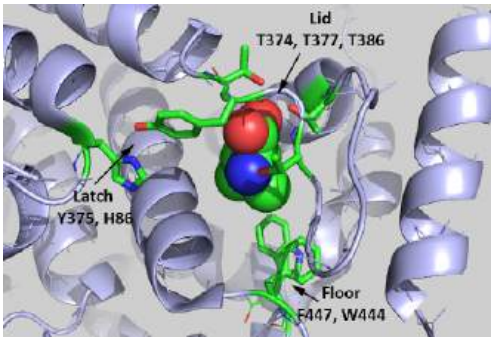


Figure 28.12.9: Binding interactions between Leu (spacefill) and key side chains in sestrin 2 (5dj4)

The Leu is rather buried, which suggests a conformational change ensues on binding to the protein. Saxton et al (2016) describe three types of sestrin2 side chains involved in the interaction:

Lid: Thr374, Thr377, and Thr386 form H bonds with the Leu amine and carboxyl group. Leucine is represented as a stick model (orange).

Latch: Tyr375 and His86 form hydrogen bonds to the Leu. Note that these residues are distal in the chain and are probably pulled together during the conformational changes which occur after binding to form a latch to sequester the bound Leu.

Floor: F447 and W444 which interact with the nonpolar side chain of Leu.

Figure 28.12.10shows an interactive iCn3D model of Leucine-bound Sestrin2 (5DJ4).

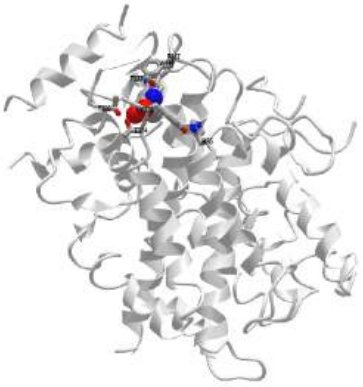


Figure 28.12.10: Leucine-bound Sestrin2 complex (5DJ4) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...nPyYVRUr9oQfv8

What happens after leucine binds? It's a complicated but understandable process described below in words and images. But first a quick review. Kinases must be regulated to be turned on and off at the right time. They are often regulated by phosphorylation, as mTOR is. In addition, they can be regulated by binding proteins as mTOR is (by Raptor, FKBP, etc). They can also be regulated by small G proteins (like Ras) which are active when bound to GTP and inactive when bound to GDP. Of course, whether small G proteins have GTP bound depends in part if they interact with **GAPs** (GTPase activating protein which inactivates small G proteins) or **GEFs** (which facilitate the exchange of GDP for GTP and activate them). Such a master regulator of growth as mTORC1 is regulated by all of these, in addition to the presence of abundant leucine.

In the absence of leucine, sestrin 2 is bound to a protein called GATOR2 (GTPase-activating protein - GAP - activity toward Rags 2). The binding of leucine to sestrin 2 causes the dissociation of GATOR2. This is shown in Figure 28.12.11

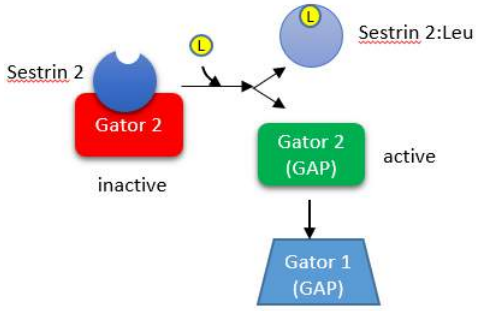


Figure 28.12.11: Sestrin 2 - GATOR interactions

Free GATOR2 is a GAP that regulated mTORC1. Specifically, it regulates the activity of a heterodimer of small GTP binding proteins, RagA/B:RagC/D (see pathway above) which are associated with the outer leaflet of the lysosome. There they interact with a membrane protein, SLC38A9, and a protein that regulates the Rag proteins, which of course is named Ragulator. Active RagA/B:RagC/D recruits mTORC1, presumably through the Raptor subunit) from the cytoplasm to the lysosome membrane. Small G proteins like Ras, when activated by exchanging bound GDP for GTP, can interact with and activate kinases (like the Raf kinase for Ras). When mTORC1 binds to active RagA/B:RagC/D, it becomes activated.

We often think of activating a protein by ligand binding, which promotes a conformational change, or by post-translational modification, which can provide a binding interaction or conformational change to activate the protein. Another way is to inhibit an inhibitor of a protein, as shown in Figure 28.12.12: Y inhibits Z as denoted by the blunt blunt-ended. If X inhibits Y, then Y can't inhibit Z, which is now active. This is analogous to the quote that "the enemy of my enemy is my friend", which has been attributed to Kautilya (from India) in the 4th century BCE.

Figure 28.12.12: An inhibitor of an inhibitor is an activator

Leucine binding to sestrin 2 leads to free GATOR, which activates mTORC1 by blocking downstream inhibitors. Figure 28.12.13 (after Buel and Blenis, 2016) shows the interactions from an activation (arrow) or inhibition (blunt arrow) perspective.

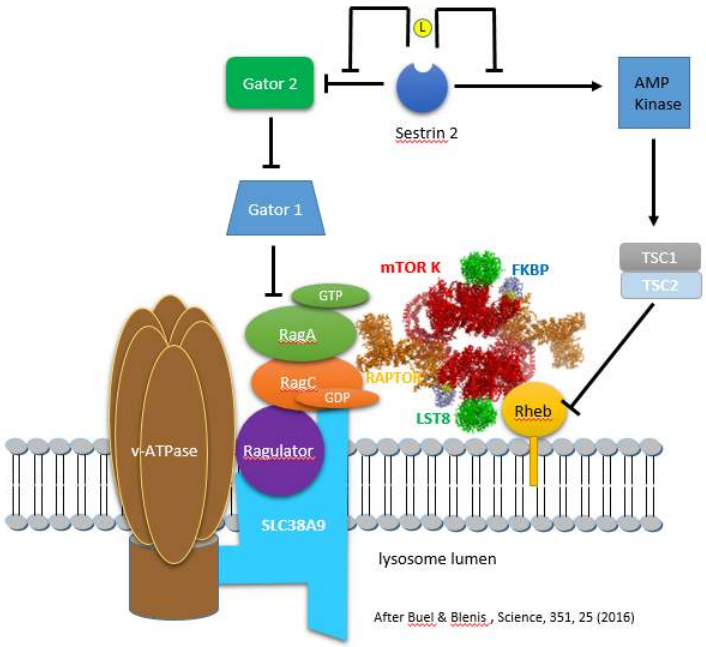


Figure 28.12.13: Leucine binding to sestrin 2 leads to free GATOR, which activates mTORC1 by blocking downstream inhibitors

The figure above shows the involvement of multiple proteins in the lysosome membrane that are involved in mTORC1 activation. There is yet another way that the RagA/B and RagC/D proteins are regulated (other than by the GATOR GAP activity. The main one appears to be **Ragulator**, which is a **GEF** for the Rag proteins. Here is a summary of the components of this lysosomal membrane recruitment center for mTORC1.

- Ragulator (what a great name) binds and recruits the small G proteins Rag to the lysosome membrane where Ragulator acts as a GEF for RagA/B
- SLC38A9 is a weak amino acid transporter in the lysosome membrane, with a preference towards polar amino acids. More likely it is yet another sensor of amino acids, particularly of arginine, which has a high concentration in the lysosome. The protein has a high Km for the transport of Arg. It has a Ragulator binding domain and is hence part of the complex that recruits mTORC1 to the lysosome
- vacuolar adenosine triphosphatase (v-ATPase): function unclear

These interactions, which involve multiple activations and inhibitions, are difficult to follow even with a diagram. The actions of small G proteins can be especially difficult to understand since the G protein is biologically **INACTIVE** in its GDP-bound form towards its target binding protein. This occurs when the GTPase activity of the G protein is **ACTIVE**. The arrows and blunt end arrows in the figure above represent the activity of the protein toward its target protein.

Here are two alternative ways to make sense of the interactions:

- Stepping backward from Rag A/B, Gator 1 (a GAP) inhibits the ACTIVITY of the protein Rag A/B as it acts as a GAP to leave Rag A/B in the inactive GDP-bound state. Paradoxically this occurs as the inherent GTPase activity of the protein is activated as

described above). Free Gator 2 (also a GAP) appears to inhibit the GAP activity of Gator 1 (through an unknown mechanism), thereby increasing the amount of GTP-bound Rag A/B, which then can activate mTORC1. Free Gator 2 does this only if Sestrin 2 is bound to Leu which allows the Gator 2 to dissociate from the inactive sestrin 2:Gator 2 complex.

- The diagram above shows that in the absence of leucine, three blunt end (inhibition) arrows occur between Sestrin 2 and Rag A/B. One blunt arrow denotes inhibition, two activation (inhibition of inhibition), and hence three net inhibition Hence in the absence of Leu (when Sestrin is bound to Gator 2, Rag A/B is **inhibited in its ability to activate mTORC1** as Rag A/B is in the GDP-bound state. However, free leucine unblocks the inhibitor action of sestrin 2 as Gator 2 is now free and active on its own.

Amino acids (especially arginine, which is abundant) in the lumen of the lysosome activate, through the v-ATPase and SLC38A9, the GEF activity of Ragulator. When Rag A/B has sufficient GTP, some conformational changes must ensue to allow mTORC1 recruitment to the lysosomal membrane.

28.12.5: Regulation of mTORC1 by Energy Availability - AMP Kinase

Believe it or not, another small G protein with GTPase activity, Rheb (Ras homolog enriched in the brain), is involved in both mTORC1 recruitment to the lysosomal membrane and activation of mTOR. This interaction is also shown in the figure above. Mostly, Rheb is involved in the activation of the kinase activity of the mTORC1 complex and specifically the phosphorylation by mTOR of the substrates S6K1 and EIF4EBP1. In the presence of growth factors, Rheb is localized to the membrane by a lipid anchor (a farnesyl group). The mTORC1 kinase-activating activity of Rheb stands in contrast to the role of the Rag G proteins which appears to be chiefly recruitment.

How is the small G protein Rheb regulated? Of course, by its interaction with yet another GAP, named the tuberous sclerosis complex (TSC). In the absence of growth factors, TSC binds to Rheb and, acting as a GAP, promotes GTP hydrolysis. This inactivates Rheb, inhibiting mTOR kinase activity.

How then is Rheb regulated? One way is through phosphorylation by **AMP Kinase (AMPK)**, an enzyme that is itself regulated by the energy level of the cells. AMPK phosphorylates and activates the TSC, which, acting as a GAP, inactivates the small G protein Rheb complex (TSC complex). Sestrins 1 and 2 may also regulate AMPK. Let's look at the energy sensor of the cell in more detail.

AMP Kinase is one of the cell's major fuel sensors and also in mammals responds systemically to hormone and nutrient levels. The enzyme is a heterotrimeric protein consisting of an alpha (catalytic), beta (regulatory), and gamma (regulatory) subunit that binds AMP, ADP, and ATP. Cellular ATP levels are determined in part by the enzyme **adenylate kinase** which helps interconvert adenine nucleotide (AXPs) as shown in the following equilibrium:

Adenylate Kinase: ADP + ADP ↔ ATP + AMP, Keq = 0.44

In red blood cells, the concentrations of ATP, ADP, and AMP are approximately 1850 uM, 145 uM, and 5 uM. Even in cells that use lots of ATP (muscle for example), ATP never falls by much. Using the values above and simple general chemistry, an 8% drop in ATP would lead, through the action of adenylate kinase, to an ATP concentration of about 1710 uM and an AMP concentration of 20 uM. This value for AMP is still much lower than ADP and ATP. However, this change represents a 4 fold increase in AMP which, even with the low actual concentration of AMP, leads to the activation of AMPK.

Another "normalized" indicator of cell energy status (or "charge") is the Energy Charge, EC. It is defined by an equation that gives a value from 0-1 where 0 indicates that all AXPs are in the AMP form and 1 where only ATP is present. The numerator of the equation of EC below represents the number of moles of phosphoanhydride linkages in the AXP pool (two for each ATP and one for ADP) and the denominator is the number of moles of AXPs (mass balance). The 1/2 term allows the bracketed term to equal 1 when only ATP exists and 0 when only AMP exists. The EC values of cells are regulated to remain around 0.85.

$$\mathrm{EC} = \frac{1}{2}\left[\frac{2\mathrm{ATP} + \mathrm{ADP}}{\mathrm{ATP} + \mathrm{ADP} + \mathrm{AMP}}\right] \tag{28.12.1}$$

Before we explore the mechanism of energy sensing by AMPK, let's look at the domain structure of the three subunits of AMPK. They are shown in Figure 28.12.14

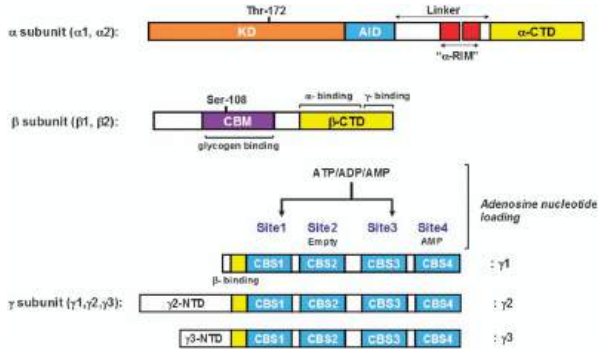


Figure 28.12.14 Domain structure of the alpha, beta, and gamma subunits of AMPK. Kim et al. Experimental & Molecular Medicine (2016) 48, e224. https://www.nature.com/articles/emm201616.pdf. a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. http:// creativecommons.org/licenses/by-nc-sa/4.0/

The mammalian α1/α2 and β1/β2 isoforms are very similar, and their characteristic features are shown. AMPKα subunits: KD, kinase domain containing Thr-172 for the activation by upstream kinases; AID, autoinhibitory domain; two α-RIM, regulatory subunit interacting motifs triggering the conformational changes in response to AMP binding to the AMPKγ subunit; α-CTD, C-terminal domain binding to the β-subunit. AMPKβ subunit: CBM, carbohydrate-binding module, in which Ser108 is important for the action of some direct AMPK activators, such as thienopyridone (A-769662) and salicylate; β-CTD, C-terminal domain containing α-subunit-binding site and immediately followed by the domain for γ-subunit interaction. AMPKγ subunit: three γ-subunit isoforms have variable N-terminal domains (NTDs); four CBS, cystathione-β-synthase domain, which forms two Bateman domains that create four adenosine nucleotide-binding sites (Sites 1–4). Site 2 appears to be always empty and Site 4 has a tightly bound AMP, whereas Sites 1 and 3 represent the regulatory sites that bind AMP, ADP, or ATP in competition.

How AMPK detects this exponential but still small molar increase in AMP is interesting, especially given the much higher concentrations of ADP and ATP. AMPK contains four binding sites that can bind AMP, ADP, and ATP (AXPs) in the regulatory subunit (gamma). This is in addition to the binding of ATP and ADP at the active site in the catalytic subunit (alpha). What binds depends on the Kd for binding of different AXPs as well as their concentrations. Bound AMP seems to have three effects on AMPK. When bound to the gamma subunit, AMP

- increases phosphorylation of Thr 172 in an "activation" loop in the catalytic alpha subunit by an upstream kinase which increases the kinase activity of AMPK by 100-200 fold. Phosphorylation is essential for the activity of the enzyme;
- inhibits dephosphorylation of Thr 172 which is perhaps the major way that AMP enhances the kinase activity of the catalytic subunit. ADP binding also inhibits dephosphorylation as shown by studies that show that the binding of ADP and the dephosphorylation of the phospho-AMPK have the same ADP concentration dependency;
- allosterically activates ten-fold the kinase activity of the catalytic alpha subunit (a secondary effect). ADP has no such effect.

These effects are altered by the markedly higher concentrations of ATP which counteracts all these effects, enhancing the Energy Charge sensor activity of this enzyme.

The gamma regulatory subunit has 4 binding sites for AXP. Crystal structures show site 2 is empty, site 4 is always bound to AMP, and sites 1 and 3 can bind AMP, ADP, or ATP. Site 1, which mediates the allosteric effects on AMPK binds all AXPs with similar affinity. This appears paradoxical since given the high energy charge, one would expect ATP and ADP to out-compete AMP for binding. However, it was found that the Mg^{2+} -ATP complex has marked lower affinity for the site, allowing both AMP and ADP, which under cellular conditions are mostly not bound to Mg^{2+} while ATP is, to out-compete Mg^{2+} -ATP for binding. Site 3 binds AMP and ADP with a 30-fold lower affinity but on binding protects p-AMPK from dephosphorylation of Thr 172.

Figure 28.12.15 shows an interactive iCn3D model of human AMPK (a2b2g1) in complex with a small molecule activator SC4 (6B2E)

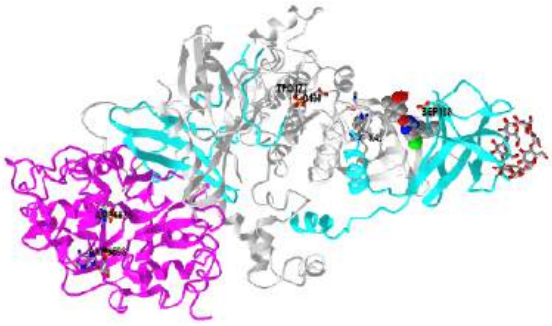


Figure 28.12.15: Human AMPK (a2b2g1) in complex with a small molecule activator SC4 (6B2E). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...HpqpDGibsRqQt7

SC4 or similar molecules may be important drugs to target AMPK and be useful in the treatment of insulin resistance and Type II diabetes. SC4 activates α2 complexes and glucose uptake into muscle cells. Some would call this type of drug an importagog, as it increases the uptake of important metabolites into cells.

The alpha catalytic kinase subunit is shown in gray with key catalytic residues and phosphorylated Thr and Ser shown in sticks, CPK coloring, and labeled. The beta subunit is shown in cyan. It appears to be two chains since the connecting section is not ordered in the crystal structure. The gamma subunit with bound AMP (spacefill, CPK colors, labeled) is shown in magenta.

What effect does activated AMPK have on the cell? Active AMPK has an amazing number of effects (see figure below). It activates liver glycolysis (by phosphorylating phosphofructokinase 2 which forms F2,6-BP, an activator of PFK) and inhibits by phosphorylation enzymes involved in fatty acid synthesis (acetyl-CoA carboxylase), glycogen synthesis (glycogen synthase) and cholesterol synthesis (HMG-CoA reductase). Yeast AMPK has recently been shown to be also controlled by acetylation of the equivalent beta subunit (Sip2). Acetylation increases its interaction with the alpha catalytic subunit (Snf1) which decreases its kinase activity. This decreases the phosphorylation of downstream kinases (including an analog of Akt1 called Sch9) which slows the growth and increases longevity. Normal aging is associated with decreased acetylation of Sip2.

Figure 28.12.16shows how the many signaling pathways we have studied interact with AMPK

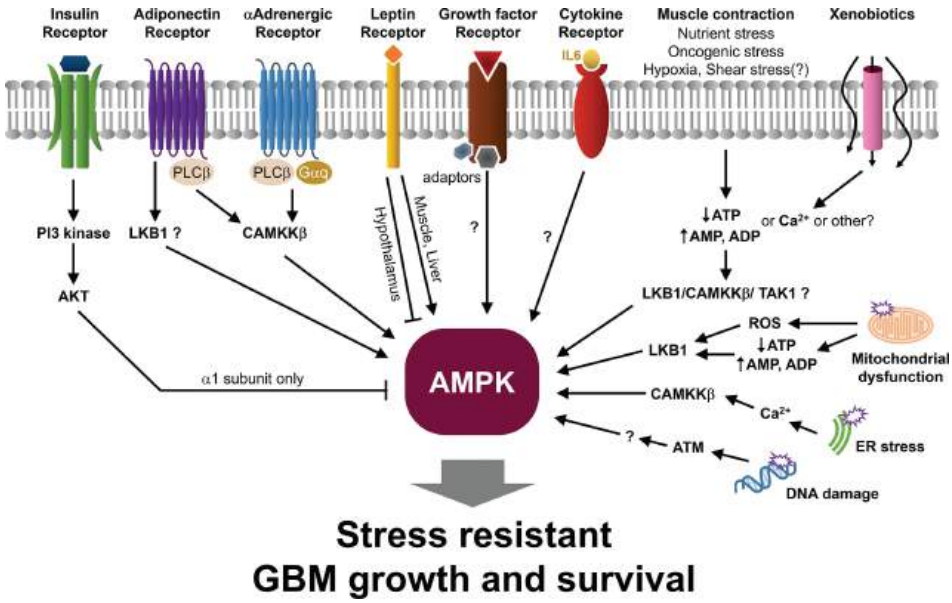


Figure 28.12.16: Signaling from membrane proteins to AMPK. Chapter 7, Targeting Energy Metabolism to Overcome Therapeutic Resistance of Glioblastoma and Tumor-associated Edema. Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) https://creativecommons.org/licenses/by-nc/4.0/

Figure 28.12.17 shows a more complete pathway of activation, regulation, and activity of AMPK. The illustration is used with courtesy of Cell Signaling Technologies (www.cellsignal.com).

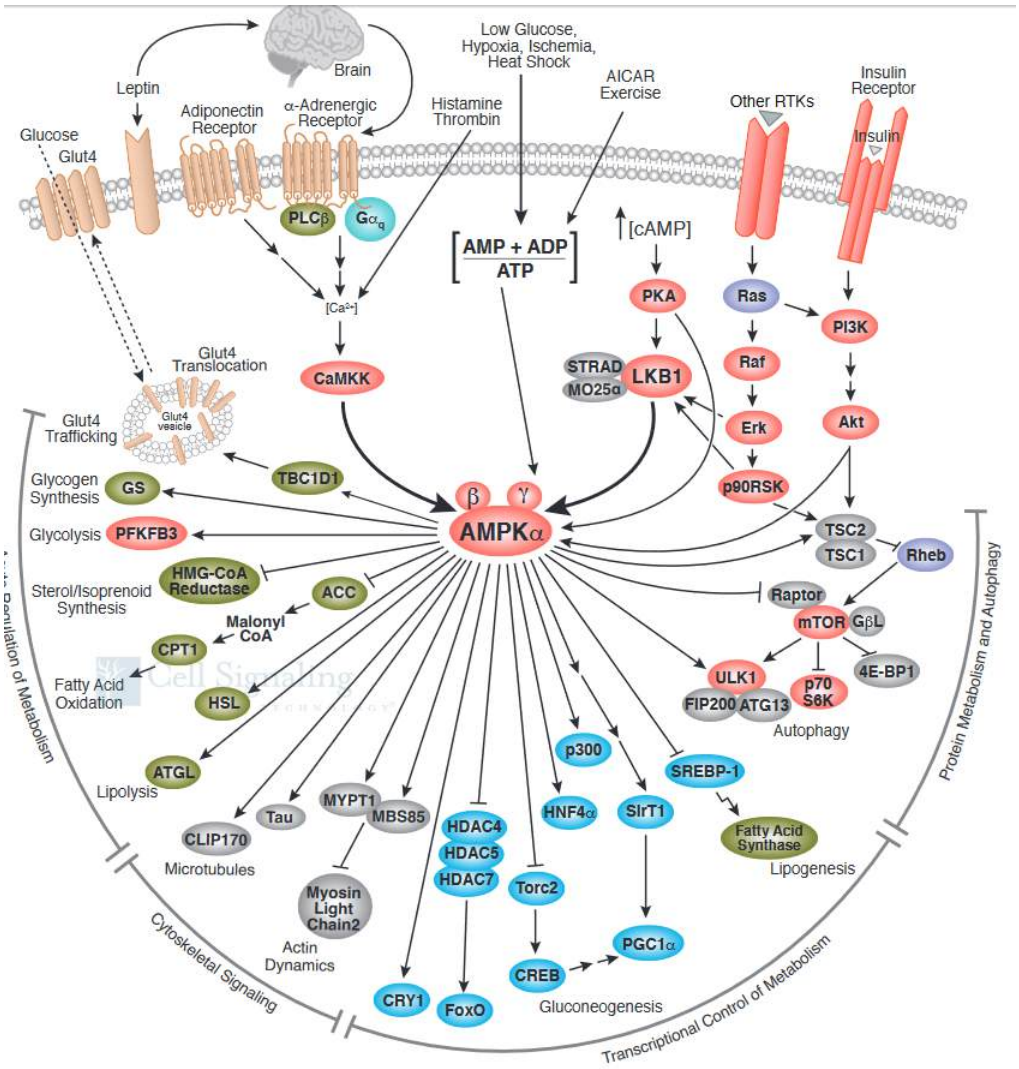


Figure 28.12.17: Signaling through AMPK

It also shows the effect of AMPK on the master regulator of protein, lipid, and nucleic acids synthesis, mTOR. Synthesis of these molecules is necessary for cell growth and proliferation, two activities that cells do not engage in when AMP levels are high, which signifies an energy-depleted state.

28.12.6: Regulation of mTORC1 by Insulin and Growth Factors

mTORC1 is regulated by local factors (amino acids, energy state) and systems factors (growth factors). This list is growing daily. The following have been shown to lead to mTORC1 activation including small molecules such as amino acids, ATP (through AMP kinase), oxygen, and glucose, and larger ones such as insulin, other growth factors, cytokines (immune growth factors and regulators), oncogenes (which promote cell proliferation) and some infectious agents. Other molecules or processes inhibit mTORC1, including tumor suppressors and stress.

mTORC1 promotes mRNA and protein synthesis as described above but also nucleotide and lipid synthesis, which is not described in detail above. In addition, it promotes aerobic glycolysis (Warburg effect), to supply not energy but intermediates for biosynthesis, as well as the pentose pathway, which forms NADPH for reductive biosynthesis and ribose for nucleic acid synthesis.

Let's look at 2 specific external hormones, insulin and epidermal growth factor (EGF), and how they affect mTORC1 activity;

Insulin:

Insulin binding to its receptor leads to the activation through phosphorylation of the kinase Akt (aka Protein Kinase B) after upstream phosphorylation of membrane phosphoinositides in the membrane and activation of phosphoinositide-dependent kinase 1, PDK1. Atk, as shown in the signaling figure for AMPK, phosphorylates TSC2, the GAP for Rheb. The arrows on the AMPK

kinase figure above **are not** consistent with our previous use of arrows. In the figure from CST, arrows show that both AMPK and Akt phosphorylate TSC2. The phosphorylated TSC is shown to inhibit Rheb, the small G protein. This would not make physiological sense. Phosphorylation of TSC2 by AMPK (signaling energy depletion) activates the TSC2 GAP protein which would inhibit RheB (the G protein) and hence inhibit mTORC1. In contrast, phosphorylation of TSC2 by Akt (signaling the abundance of glucose) leads to the **inhibition** of the GAP activity of TSC2. That would keep Rheb in the active, GTP-bound form, which leads to the activation of the bound mTORC1. A more complete description of the pathway where insulin binds to its receptor (an insulin-gated receptor tyrosine kinase) and leads to activation of mTORC1 through Akt is shown in Figure 28.12.18

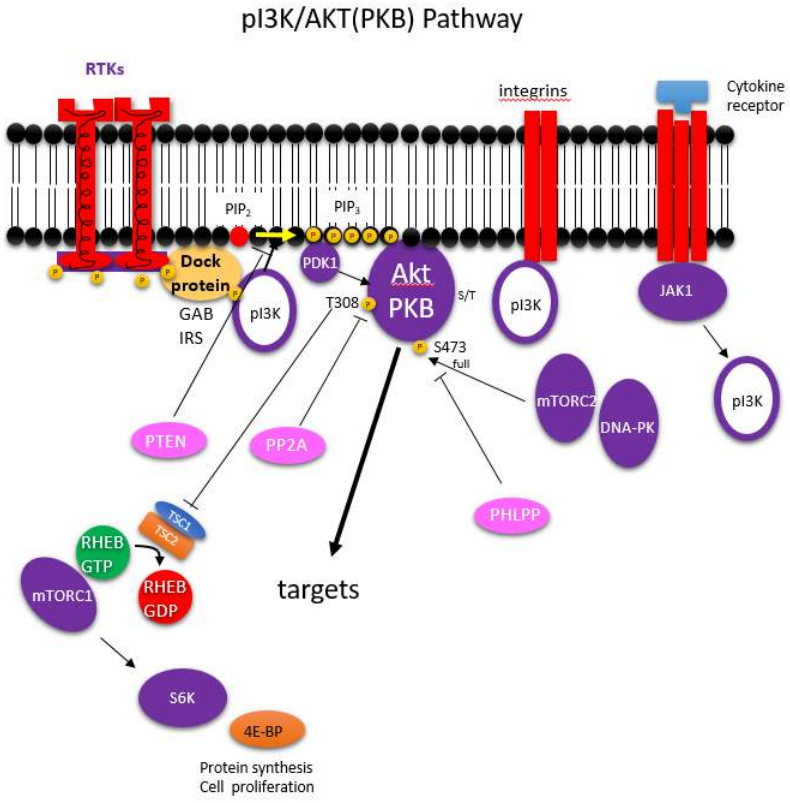


Figure 28.12.18: Insulin Receptor and mTOR signaling

MAPK Cascade:

EGF binds its receptor, activating it as a receptor tyrosine kinase. Typical of other receptor kinases, it activates the mitogen activate protein kinase system. This process is mediated by Ras (a small G protein) activator of Raf (a mitogen-activated protein kinase kinase kinase or MAP3K). Active Raf phosphorylates and activates MEK (a MAPK2) which activates ERK (a MAPK). Erk phosphorylates mTORC1 directly, which activates it. It also phosphorylates TSC2/TSC1, which inhibits this GAP protein, leading indirectly to the activation of the small G protein Rheb, which also activates mTORC1. These steps are shown in Figure 28.12.19

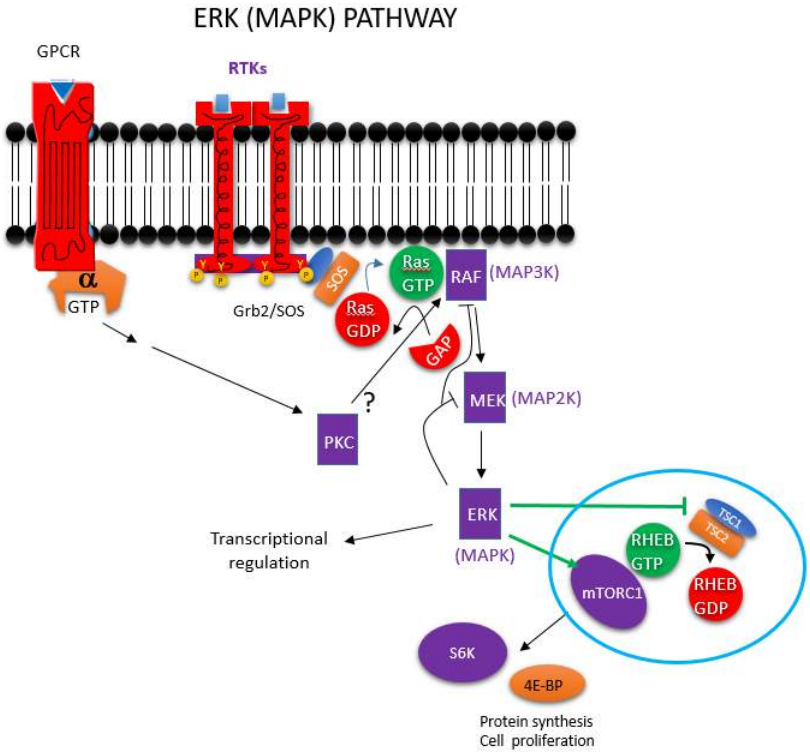


Figure 28.12.19: MAPK and mTOR signaling

A summary showing the kinases that activate or inhibit TSC2/1 is shown in Figure 28.12.20 We tend to concentrate on our favorite protein and confer it with special status as critically important in a pathway. One could pick the GAP protein TSC2 as especially important in regulating the activity of mTORC1.

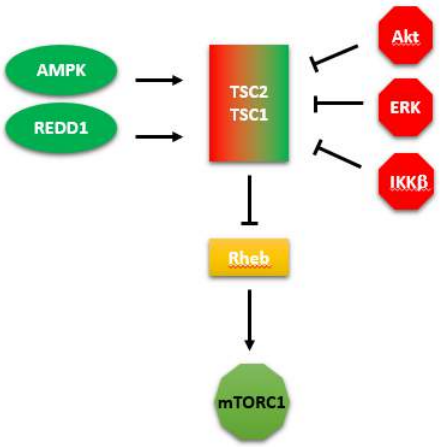


Figure 28.12.20 kinases that activate or inhibit TSC2/1

The figure above shows two additional proteins. One is REDD1 (not a kinase), which activates the GAP protein TSC2, leading to the inhibition of the small G protein Rheb, and hence the inhibition of mTORC1.

REDD1 (regulated in development and DNA damage responses 1) is also called DDIT4 (DNA-Damage-Inducible Transcript 4). It is a gene whose expression is activated during hypoxia by hypoxia-inducible factor-1 and also during DNA damage. Hypoxia alters metabolism very quickly. The protein is degraded by the proteasome after it is targeted for degradation by the post-translational addition of ubiquitin. This suggests yet another way to regulate the activity of mTORC1.

The other is IKK beta, also known as IKBKB (Inhibitor Of Kappa Light Polypeptide Gene Enhancer In B-Cells, Kinase Beta). It is a kinase that phosphorylates and inhibits TSC2 which inhibits Rheb, leading to the activation of mTORC1.

This kinase is activated by many stimuli including inflammation (mediated by cytokines), bacterial or viral infections, and DNA damage. It phosphorylates a bound inhibitor of NF-kappa beta. This allows ubiquitinylation of the inhibitor, targeting it for proteasomal degradation. The free NFKB can then enter the nuclease and alter the transcription of genes involved in the immune response and hence promote proliferation. Under these conditions, one would expect the activation of mTORC1.

A final note: Less is known about how lipids regulate mTORC1. Two possible lipid signaling molecules, phosphatidic acid, and phosphatidyl inositol -3-phosphate are probably involved. The enzyme that makes them, phospholipase D and phosphoinositide 3-kinase (from the PIK3C3 gene), also known as VPS34 (for vacuolar protein sorting from yeast), are found in phagosomal and lysosomal vesicles and are involved in their processing, seem to be involved in mTOR signaling. Obesity and people with high-fat diets have elevated mTOR activity.

28.13: Regulation of the Cell Cycle by Protein Kinases

This section is an integration of materials as referenced, with significant modifications and additions.

Aleem and Arceci. Targeting cell cycle regulators in hematologic malignancies. Article in Frontiers in Cell and Developmental Biology 2015. DOI: 10.3389/fcell.2015.00016.

28.13.1: Introduction

For a cell to undergo successful division, it has to perform four key tasks in a highly ordered fashion. First, there is a preparatory synthetic phase (G1) that results in increased cell size in anticipation of DNA replication (S phase). Cells then proceed through (G2-phase) to prepare to equally segregate duplicated DNA (M phase) and finally divide into two equal daughter cells. From G1 a cell can also exit the cell cycle and enter a state of quiescence (G0), undergo differentiation, or re-enter the cell cycle to proliferate in response to mitogenic signals.

The core molecular machinery controlling the mammalian cell cycle consists of a family of serine/threonine protein kinases called **cyclin-dependent kinases (CDKs**). These are catalytic subunits, which are activated in most cases by association with **cyclin regulatory subunits**. The activity of CDK/cyclin complexes is further regulated by CDK inhibitors **(CKIs)**, phosphorylation and dephosphorylation, ubiquitin-mediated degradation, transcriptional regulation, substrate recognition, and subcellular localization. The family of CDKs/cyclins/CKIs contains more than 30 members. They are implicated in essential cellular functions such as transcription, DNA damage repair, epigenetic regulation, metabolism, proteolytic degradation, stem cell self-renewal, neuronal functions, and spermatogenesis. Figure 28.13.1shows the cell cycle and the involvement of CDKs/cyclins at key points.

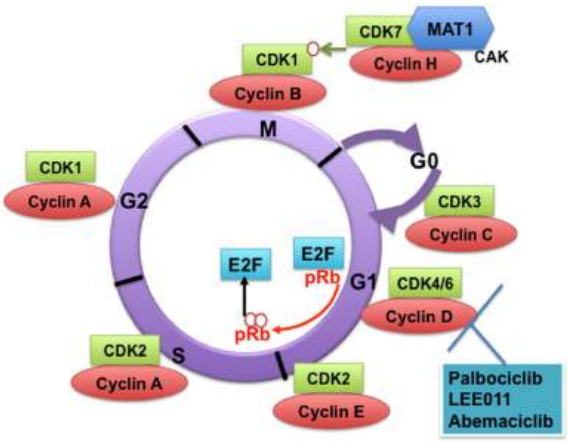


Figure 28.13.1: Cyclin-dependent kinases (CDKs) and their cyclin regulatory subunits. CDK-cyclin complexes with direct functions in regulating the cell cycle. Aleem and Arceci. Frontiers in Cell and Developmental Biology 2015. DOI: 10.3389/fcell.2015.00016.

CDK3/cyclin C drives cell cycle entry from G0. CDK4/6/cyclin D complexes initiate phosphorylation of the retinoblastoma protein (pRb) and they sequester $p21^{Cip1}$ and $p27^{kip1}$ (not shown), which are both inhibitors of CDK2, thus promoting the activation of CDK2/cyclin E complex. In late G1, the CDK2/cyclin E complex completes phosphorylation and inactivation of pRb, which releases the E2F transcription factors and the G1/S transition takes place. DNA replication takes place in the S phase. CDK2/cyclin A complex regulates progression through the S phase and CDK1/cyclin A complex through the G2 phase in preparation for mitosis (M). Mitosis is initiated by **CDK1/cyclin B** complex (which will model at the end of this section). The activity of CDK1/cyclin B is tightly regulated by activating phosphorylation by the CDK-activating kinase CAK (a heterodimer of cyclin H-CDK7-MAT1) and inhibitory phosphorylations by Wee1 and Myt1 on Tyr15 and Thr14 (not shown). Some specific CDK4/CDK6 pharmacological inhibitors are also shown

28.13.2: CDKs with Direct Functions in Cell Cycle Regulation

The classical CDKs that directly regulate the mammalian cell cycle in complexes with cyclin subunits include CDK3, CDK4, CDK6, CDK2, and CDK1. CDK3 promotes cell cycle entry from quiescence in association with cyclin C. CDK8 has also been suggested to play a role in cell cycle entry from G0 and in the G1/S transition. In its simplest model, the mammalian cell cycle proceeds as follows:

- In early G1, CDK4/CDK6 in complex with cyclin D receive mitogenic signals that result in activation of cell cycle entry, as shown in Figure 28.13.1 Key signaling events include the initiation of retinoblastoma protein (pRb) phosphorylation and the sequestration of $p21^{Cip1}$ and $p27^{kip1}$, which are both inhibitors of CDK2, thus promoting the activation of CDK2/ cyclin E complex. In late G1, CDK2 in complex with cyclin E completes the phosphorylation and hence inactivation of pRb, which in turn releases the E2F transcription factors. E2F promotes transcription of cyclin E which is necessary for the G1/S transition.
- Progression through the S phase is mediated by CDK2/cyclin A complex.
- Mitosis is then initiated by **CDK1/cyclin B complexes**. We will model the regulation of CDK1 later in this section. CDK1/cyclin A complexes contribute to the preparation for mitosis in the G2 phase. The activity of CDK1/cyclin B is tightly regulated by activating phosphorylation by the CDK-activating kinase (CAK) (a heterodimer of cyclin H and CDK7) and inhibitory phosphorylations by WEE-1 and Myt1 on Tyr15 and Thr14. Mitosis starts after WEE-1 is degraded and CDC25C phosphatase releases the inhibitory phosphorylation on CDK1/cyclin B.

The cyclins are also expressed in a coordinated fashion throughout the cell cycle. The cyclin expression cycle is shown in Figure 28.13.2 The timing of expression is consistent with the explanations above.

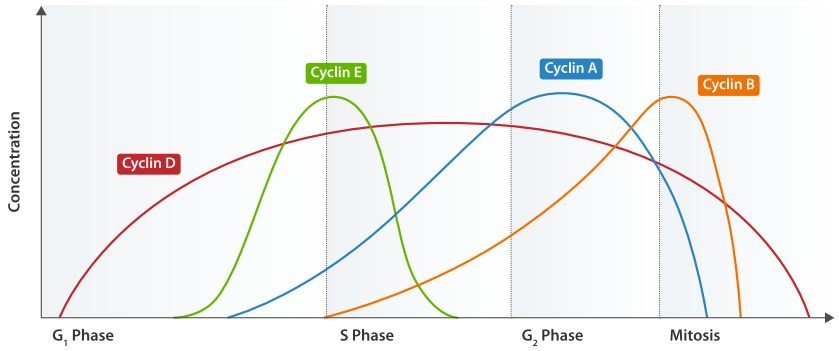


Figure 28.13.2: Cyclin expression cycle. https://commons.wikimedia.org/wiki/F...Expression.svg

A graph showing multiple progressions through the cell cycle is shown in Figure 28.13.3

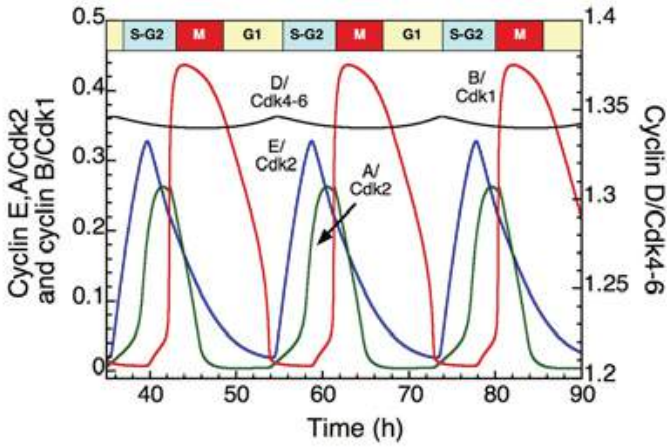


Figure 28.13.3 Sustained oscillations of the CDK network in mammalian cells. Gérard and Goldbeter, Front. Physiol., 02 November 2012
Sec. Systems Biology Archive. https://doi.org/10.3389/fphys.2012.00413. Creative Commons Attribution License

The time evolution of cyclin D/Cdk4–6 (in black), cyclin E/Cdk2 (in blue), cyclin A/Cdk2 (in green), and cyclin B/Cdk1 (in red) is shown in the presence of a suprathreshold level of growth factor. Cyclin D/Cdk4–6 is the total active form of the kinases, which is composed of cyclin D/Cdk4–6 and also the complex formed by cyclin D/Cdk4–6 and p21/p27.

Note again the oscillatory nature of the cyclin B/Cdk1 complex, which will explore at the end of this section.

28.13.3: CDKs with Transcriptional and Other Functions

In addition to their direct role in the mitotic cell cycle regulation, some classical CDK/cyclin complexes have essential functions in meiosis, such as CDK2, in transcription and/or DNA repair. Other CDKs act by activating the classical CDKs, such as CDK7/cyclin H (CAK) and the related CDK20, also known as cell cycle-related kinase (CCRK). Some CDKs function mainly in influencing transcription by phosphorylating the carboxy-terminal domain (CTD) of ribonucleic acid (RNA) polymerase II (RNA pol II). This phosphorylation also serves as a platform for RNA processing and chromatin regulation.

CDKs that have important transcriptional roles include CDK7/cyclin H/MAT1 complex, a component of the basal transcription factor, TFIIH, and facilitate transcriptional initiation. CDK8/cyclin C, in addition to its role in transcription, is also involved in the Wnt/β-catenin pathway and inhibition of lipogenesis. Cyclin C can recruit CDK8 or CDK19 to the CDK8 module of the Mediator complex, which can function as a positive or negative regulator of transcription by RNA pol II. CDK3/cyclin C also plays a role in NHEJ-mediated DNA damage repair. While CDK9 in complex with cyclin T forms the phospho-transcription elongation factor b (p-TEFb) and promotes transcriptional elongation, CDK9 also functions in the DNA damage response when complexed with cyclin K. CDK10/cyclin M phosphorylates the Ets2 transcription factor and positively controls its degradation by the proteasome. Ets2 plays a key role in cancer and development. CDK11/cyclin L controls the assembly of the RNA pol II mediator complex. CDK12 and CDK13 in complex with cyclin K control RNA pol II transcription, and CDK12/cyclin K controls DNA damage response.

28.13.4: Structure of CDKs–cyclins

The structures of inactive cyclin-free kinases are very similar but vary at the N-terminal and C-terminal ends. Figure 28.13.4shows an interactive iCn3D model of the prototypical active human cyclin-dependent kinase 2 with a bound ATP (1HCK).

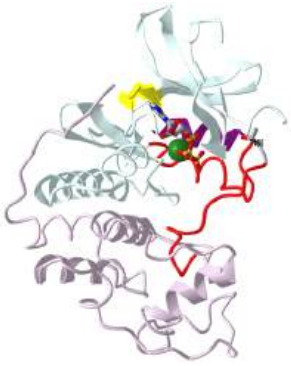


Figure 28.13.4: Human cyclin-dependent kinase 2 with a bound ATP analog (1HCK) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...3Sf81pwLKQZx9A

The model shows that CDK2 has structural features shown in all the kinases we have studied previously:

- a smaller N-terminal lobe (light cyan) and larger C-terminal lobe (light magenta) in between which ATP binds (along with Mg^{2+}).
- the C-helix (residues 45 – 55, purple), which contains a conserved Glu. It forms an interaction with and helps position a key Lys in the active, which facilitates ATP binding and transition state stabilization;
- hinge (residues 80 – 84, yellow),
- activation loop (residues 145– 172, red), which contains T160 (sticks, CPK colors, labeled) that becomes phosphorylated on activation by yet another kinase called **CDK-activation kinase (CAK).** When T160 is phosphorylated, the kinase binds to cyclin A. The loop starts and ends with the conserved residue DFG and APE, respectively.
- Not highlighted in the model is a conserved conformationally flexible glycine-rich region (residue 12-16) with the motif GXGXXG

In the inactive form, the N-terminal end of the activation loop has a short alpha helix that prevents the C-helix from adopting the correct position for catalysis. Activation requires movement of the C-helix allowing the Glu in the C-helix to position the active site Lys.

28.13.5: CDK2–cyclin A activation

The binding of cyclin A to CDK2 activates it through the repositioning of the C-helix and the activation loop. When CDK2 is phosphorylated and bound to cyclin A, there is a large shift in the C-helix allowing the interaction of the C-helix Glu with the active site Lys.

First, let's look at the structure of cyclins. Each cyclin has a unique sequence and structural features that allow them to interact with specific CDKs and associated proteins. However they all have a conserved "cyclin box" structure containing about 100 amino acids.

Figure 28.13.5shows an interactive iCn3D model of bovine cyclin A (1VIN).

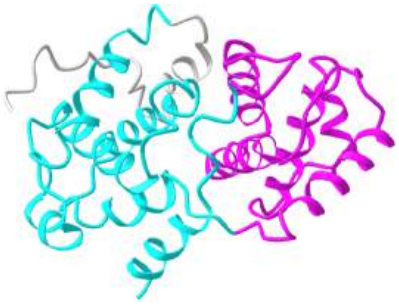


Figure 28.13.5: Bovine cyclin A (1VIN) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...6j7kLcPzpRTAA9

Cyclin A has two linked cyclin box folds, each containing around 100 amino acids and comprised of five helices. They interact with the more disordered parts of unphosphorylated CDK2, which result in some low levels of activity. Cyclin binding causes a large movement of the C-helix enabling the Glu -- Lys interaction. Phosphorylation of T160 leads to the repositioning of the activation loop.

Figure 28.13.6shows an interactive iCn3D model of Phosphorylated cyclin-dependent kinase 2 bound to cyclin A (1JST)

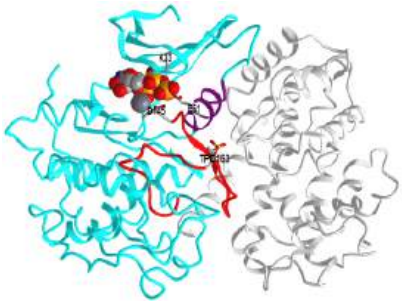


Figure 28.13.6: Phosphorylated cyclin-dependent kinase 2 bound to cyclin A (1JST) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...cP6YsNcgUtGHW8

CDK2 is shown in cyan and cyclin A in gray. Here are some structural features represented in the model.

- the C-helix (residues 45 – 55, purple),
- activation loop (residues 145– 172, red), which contains pT160 (sticks, CPK colors, labeled)
- the catalytic "triad" Lys^{33}, Glu^{51}, and Asp^{145}

Figure 28.13.7shows an animation of structural changes in just CDK2 when "apo"-CDK2 (without bound cyclin A, pdbID 1HCK) binds cyclin A (1JST).

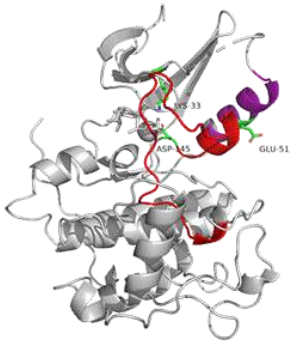


Figure 28.13.7: Structural changes CDK2 when it binds cyclin A

Gray represents the structure of CDK2 in the absence of cyclin A. The structure of just CDK2 in the presence of cyclin A is shown in magenta. Note that large shifts in the C-helix (purple) and activation loop (red) on binding cyclin A.

28.13.6: Cyclin partners of CDK1 and CDK2

CDK1 is the closest member of the CDK family to CDK2 and for which structures of the cyclin-free and authentic cyclin-bound forms can also be compared.

Depending on cyclin availability and concentration, CDK2 can bind cyclin A, B (if CDK1 expression is knocked down), and E (see Figure 28.13.1. The binding interface between CDK2 and the cyclins is quite large compared to the interface between CDK1 and cyclins. Three large aromatic side cyclin side chains (Y170, Y177, and Y258) are conserved in the binding interface. In cyclin E, the corresponding amino acids are smaller (N112, I119, and L202).

The binding interface between CDK1 and the cyclins is smaller so it appears that it might preferentially interact with cyclins A and B to gain binding affinity through the more robust interactions with the aromatic groups in the interface in the CDK2: Cyclin A and CDK2:cyclin B complexes.

A comparison of the CDK1–cyclin B and CDK2–cyclin A/B/E structures also highlights the potential for these closely related CDKs to be differentially regulated by reversible phosphorylation. The antagonistic activities of Wee1/Myt1 kinases and Cdc25 phosphatases regulate the phosphorylation status of the CDK glycine-rich loop (defined by the GXGXXG motif, residues 11–16 in CDK2). The structure of CDK2–cyclin A phosphorylated on Y15 illustrates how phosphorylation promotes a glycine loop structure that antagonizes both peptide substrate binding and the ATP conformation required for catalysis. The flexibility of the glycine-rich loop is compatible with a model in which the phosphorylated Y15 side chain is solvent exposed and accessible to both kinases and phosphatases. CDK1 is also regulated by active-site phosphorylation, and the conserved nature of the structure in this region suggests that the mechanism of inhibition is also conserved.

28.13.7: CDK substrate recognition

Local and distal sequence motifs must be used to confer specificity to the binding of specific cyclins and other proteins to specific CDKs. One interesting example is provided by examining the structure of a phospho-CDK2 Cyclin A in complex with a peptide substrate derived from the protein CDC6. Figure 28.13.8shows an interactive iCn3D model of Phospho-CDK2:Cyclin A complex with a peptide containing both the substrate and recruitment sites of CDC6 (2CCI)

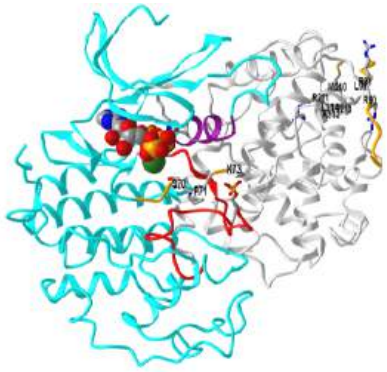


Figure 28.13.8: Phospho-CDK2 Cyclin A in complex with a peptide containing both the substrate and recruitment sites of CDC6 (2CCI) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...Ge8kdSBjxkaNd8

The color coding is the same as the models above:

- CDK2 is shown in cyan and cyclin A in gray.
- the C-helix (residues 45 – 55, purple),
- activation loop (residues 145– 172, red), which contains pT160 (sticks, CPK colors, labeled)
- the catalytic "triad" Lys^{33}, Glu^{51}, and Asp^{145}

The 30 amino acids peptide (numbers 67-96) shown in gold is a substrate for phosphorylation by the CDK2:cyclin A complex. It derives from an actual biological substrate in the protein cell division control protein 6 homolog, also called CDC6-related protein. It is involved in a checkpoint control of the cell cycle that "checks" that DNA replication is completed before mitosis. It is discontinuous in the model since part of the bound peptide is intrinsically disordered and not observed in the crystal structure.

- The 1st fragment of the CDC6 peptide (67-73) contains the binding motif sequence S/T)P*X*(K/R) (the CDC6 substrate has the sequence 70Ser-Pro-Arg-Lys). Ser 70 is the target amino for phosphorylation by CDK2:cyclin A.
- The second fragment seen in the model (amino acid 85-96) contains another binding motif, RXL (in this peptide RRL), which acts to recruit cyclin A. This binds to the sequence MRAIL (210-214) in cyclin A.

What is so interesting is that this second binding site on cyclin A for its target protein is so far away from the active site of the CDK2:cyclin A. These kinds of interactions work to determine the specificity for CDKs and the binding cyclin partners.

CDKs 7, 9, 12, and 13 phosphorylate the RNA polymerase carboxy-terminal domain (CTD). The sequence of the CTD is unusual, composed of 52 heptad repeats in humans, with the consensus sequence YSPTSPS. Extracted from cells, CTD residues S2 and S5 are the most abundantly phosphorylated serine residues, while S7 is phosphorylated to a lesser extent. CDK7 has been shown to predominantly phosphorylate S5 and S7, CDK9 to have activity towards all three serines, and CDK12 and CDK13 to predominantly phosphorylate S2.

CDKs form complexes not only with target protein substrates but other proteins which can serve as scaffolding anchors that bind both the CDK and the cyclin. Figure 28.13.9 shows an interactive iCn3D model of human CDK-activating kinase (CAK), a complex composed of cyclin-dependent kinase (CDK) 7, cyclin H, and the scaffolding protein MAT1 (6xbz).

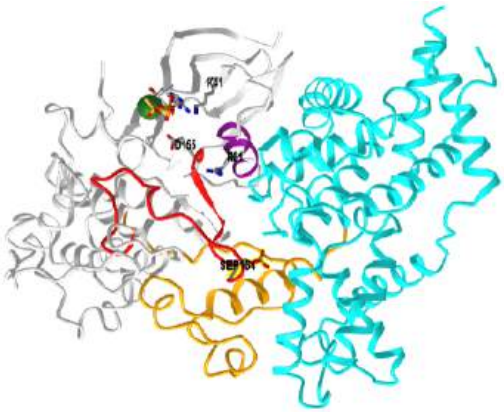


Figure 28.13.9: Human CDK-activating kinase (CAK), a complex composed of cyclin-dependent kinase (CDK) 7, cyclin H, and MAT1 (6xbz)(Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...3bdn39ARgLijB7

The gray protein is CDK7, the cyan is cyclin H and the orange MAT1. The purple again represents the C-helix of the CDK, and the red is the activation loop. The catalytic triad side chains in the active site of the CDK are shown in CPK-colored sticks. Also shown is phospho-Ser in the activation loop.

The CDK activating kinase (CAK) shown above phosphorylates the target S/T in the activation loop (which is also called the T-loop) in CDKs, activating the kinase. In addition, it regulates the initiation of transcription by phosphorylating the YSPTSPS repeats in the C-terminus of RNA polymerase II subunit RPB1. There are 15 consecutive repeats in the sequence as well as others dispersed in the C-terminal domain.

We already mentioned the motif RXL found in cyclin binding proteins that recruit them to cyclins (through, for example, their interaction with MRAIL (210-214) in cyclin A. Likewise short motifs in cyclins are used to bind to proteins that increase CDK activity or decrease it.

A number of cyclin-encoded protein-binding sites or short peptide motifs have been structurally characterized. A well-characterized example is the recycling of the cyclin RXL recruitment site that is exploited to either enhance or inhibit CDK activity. Alternatively, short motifs encoded within the cyclin sequence can be used both to dock cyclins to substrates to enhance CDK activity and alternatively to localize them to CDK regulators frequently resulting in a loss of CDK activity. Members of the p27KIP1/p21CIP1 cyclin-dependent kinase inhibitor (CKI) family share an RXL motif with RXL-containing substrates and compete with them for CDK–cyclin association. The INK (inhibitors of CDK) family of CKIs selectively inhibits CDK4 or CDK6 and, through an allosteric mechanism, disfavors CDK–cyclin binding [15]. Their tandem ankyrin repeat structures exemplified bCy CDK6–p19INK4d bind in the vicinity of the CDK hinge on the interface opposite to the surface remodeled upon cyclin association.

28.13.8: Modeling the Cell Cycle - Oscillations

Lastly, we will focus on mathematical models that show how the oscillatory behavior of CDK1/Cyclin B arises (remember than CDKs become active on binding to a cyclin). When CDK1 is activated, the cell is driven into mitosis. It is driven out of mitosis by the activation of the anaphase-promoting (APC) complex, which contains APC-Cdc20, an E3 ubiquityl ligase. Yeast Cdc20 is an activator protein that regulates the ubiquitin ligase activity of APC by binding at the right time in the cell cycle to B cyclins that contain a D box motif. This recruits the B cyclin:CDK1 to APC which ubiquitinates the cyclin B, leading to its degradation by the proteasome.

Figure 28.13.10shows the levels of both cyclin B/CDK1 (red) and cyclin A/CDK2 (green) with time. Now in your mind, imagine another curve on the graph showing similar oscillations of activated APC, only frameshifted a bit in time so that the active APC trails that of active cyclin B/CDK1. When cyclin B/CDK1 is at iitsactive peak, active APC is already beginning its rise.

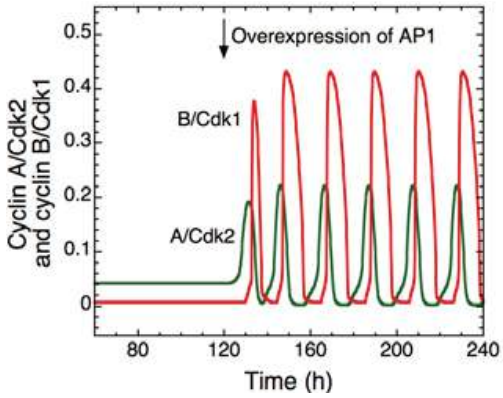


Figure 28.13.10 The switch to sustained oscillations of cyclin A/Cdk2 (in green) and cyclin B/Cdk1 (in red) is shown following the overexpression of AP1. Gérard and Goldbeter, ibid.

The dissociation constant, K_D, for CDK1 and cyclin B is about 28 nM, which represents high affinity binding. Ubiquitinylation of cyclin B and its degradation allows for the freeing of CDK1 and inhibition of its activity. The oscillatory behavior in activity occurs only on the overexpression of yet another protein, AP1.

Let's explore in detail a model proposed by Ferrell et al that accounts for the 25 min oscillatory behavior of CDK1 in Xenopus (frog) eggs. The eggs are very large and perhaps because of their size have different constraints on their cell cle cycle. For example, cells can enter mitosis before the completion of DNA synthesis. The players which regulate its activity are shown below in Figure 28.13.11

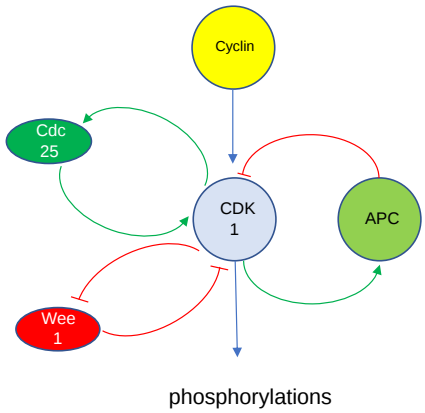


Figure 28.13.11: Proteins that help regulate the activity of CDK1 in Xenopus eggs.

- Cyclin: binds to and activates CDKs; active CDK1 drives cells into mitosis.
- APC: anaphase-promoting complex; active APC drives cell out of mitosis with its E3 ubiquitin ligase activity, modifying cyclin and targeting it for proteolysis;
- Wee1: nuclear Ser/Thr kinase
- Cdc25: cell cycle division phosphatase that activates cyclin-CDK1.

Our goal in this discussion is not only to model the actual oscillatory behavior of CDK1 but also to show you how models are built and tested. When modeling enzyme inhibition data, it is important to fit the data to many models (reversible competitive, uncompetitive, mixed, noncompetitive) to find out which best fits the data. Typically one starts with the simplest possible model and then advances to more complicated models until the data with the best statistical fit to the data is found. Other examples, which are more relevant here, involve fitting binding and kinetic data using equations that give hyperbolic (for saturation binding and simple Michaelis-Menten kinetics) and sigmoidal fits (for cooperative binding and allosteric enzymes).

So let's start with the simplest model that might lead to oscillatory activity of CDK1. In this section, active CDK1 will be designed as CDK1*.

Model 1: One-step process with negative feedback - CDK1* inhibits its own activation (for example by activating APC and hence ubiquitinylation of cyclin)

We saw in the Vcell models for the MAP kinase cascade that feedback phosphorylation of the first enzyme in the cascade (MAPKKK) by the last enzyme in the cascade, MAPKPP, leads to oscillations in enzyme activity. Could it explain the oscillations in the activity of CPK1 in Xenopus eggs? Let's make the following set of assumptions, which all be understandable from the material presented in previous chapters:

1. CDK1* is inhibited by APC* (active APC), and to make it simpler, APC* can be expressed as a simple function of CDK1* so there is just one species that vary in the flux equation;
2. CDK1 activation occurs on rapid high-affinity binding of cyclin, which is synthesized at a constant rate a1;
3. The rate of CDK1 activation to produce CDK1* is given by mass action = rate activation - rate inactivation;
4. APC is activated to APC "instantaneously" by CDK1* so APC* is a very sensitive “cooperative” function of CDK1* which can replace the APC*. For this type of "instantaneous (or highly cooperative effect), we will use the Hill equation which gives a sharp, sensitive, cooperative rise in complex instead of a simple formation of a complex between CDK1* and APC. We explored this property of the Hill equation in Chapter section 5.3.5 on Mathematical Analysis of Cooperative Binding.

Here is the Hill expression commonly use to **empirically fit** the fractional saturation of a species

$$Y = \frac{L^n}{K_D^n + L^n} \tag{28.13.1}$$

It's a bit different than the Hill equation we saw for modeling the cooperative binding of O_2 to hemoglobin (Chapter 5.3.5) since the K_D term is also raised to the power n (which is not in the actual Hill equation). However, we did see that for oxygen binding to hemoglobin,

$$\mathrm{K_D = P_{50}^n} \tag{28.13.2}$$

Hence the empirical expression used in fitting Model 1 is completely in accord with the Hill treatment of cooperativity. Again we use the Hill equation when modeling binding and kinetic data that show significant sensitivity to conditions. It gives yet another parameter to help fit the data and to test models.

Two different representations of a reaction diagram showing Model 1 are shown in Figure 28.13.12

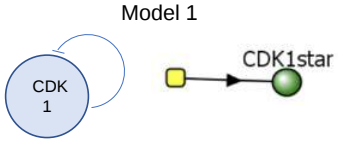


Figure 28.13.12 Two different representations of a reaction diagram for Model 1 - Activity of CDK1.

The representation on the right is from Vcell. The one on the left shows that CDK1 can inhibit itself.

Model 1 and its associated assumptions lead to the following differential equation that can easily be solved numerically in Vcell. All of the VCell outputs shown below were obtained from Vcell models kindly provided by Leslie Loew.

$$v = \frac{dCDK1^*}{dt} = a_1 - b_1\,(CDK1^*)\,(APC^*) = a_1 - b_1\,(CDK1^*)\left(\frac{CDK1^{*n1}}{K_1^{n1} + CDK1^{*n1}}\right) \tag{28.13.3}$$

This rate equation has two terms (assumption 3). The first is the rate that CDK1* forms (a constant a1 defined by the rate of cyclin synthesis) and the rate at which it is degraded by APC*. The constant b1 in the second term can be thought of as a second-order rate constant for the interaction of CDK1* and APC*, a process that inactivates CDK1*.

The [APC*] in the middle equation is replaced with the Hill equation for the effective fractional saturation concentration of APC (see assumption 4 of Model 1 described above) in the right-hand side.

$$(APC^*) = \left(\frac{CDK1^{*n1}}{K_1^{n1} + CDK1^{*n1}}\right) \tag{28.13.4}$$

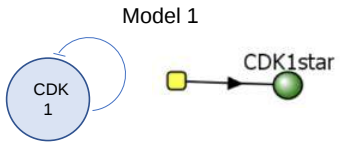
We will define the activity of the system as the rate at which CDK1* forms.

$$\text{Activity} = \frac{dCDK1^*}{dt} \tag{28.13.5}$$

Now let's see if changing the Hill coefficient n1 can cause oscillations in CDK1*.

VCell MODEL

CDK - Model 1



Initial Conditions

J	reaction rate	$\left(a1 - \frac{b1 \cdot CDK1star \cdot (CDK1star)^{n1}}{(K1)^{n1} + (CDK1star)^{n1}}\right)$
a1	user defined	0.1
n1	user defined	8.0
K1	user defined	0.5
b1	user defined	1.0

Select Load [model name] below

Load Model1CDK

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

Note that all the graphs plateau quickly at which point the CDK1* activity is constant. The graph (gray) of the curve with the highest value of the Hill coefficient (n1=24) is linear and then abruptly plateaus. The slope of the velocity curve over the entire linear part of the n1=24 graph curve is 0.1, which is the value set for the rate of activation of CDK1. Then suddenly at around 4 seconds, an "almost infinitely cooperative" shift to a constant rate of formation occurs arising from an abruptly reached rate of inactivation of CDK1* by the APC complex. These graphs do not show oscillations.

Now let's see the graph with no feedback inhibition, much as we did with the MAPK cascade in Chapter 12.4. The easiest way to do that is to set b1, the "second" order rate constant for the interaction of CDK1* and APC* in the model to 0. The graph is shown in Figure 28.13.13

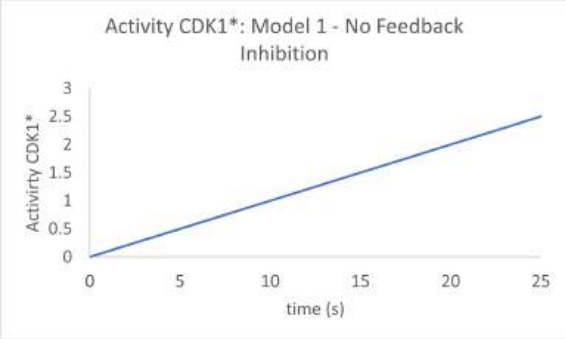


Figure 28.13.13 Activity of CDK1* in Model 1 in the absence of feedback inhibition.

The activity of CPK1* continually increases. When feedback inhibition is added, the curve "bends" to a plateau, but it does **not** start to decrease and shows no signs of oscillations. Time to move on to a more complex model!

Model 2: Two-species model with activation and inhibition-

This model is more complicated and shows 2 species (CDK1 and APC) both of which are activated and inhibited. We need 2 mass action differential equations, one for each. Figure 28.13.14

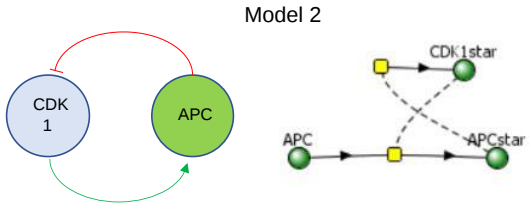


Figure 28.13.14 Two different representations of a reaction diagram of Model 2

The equation for dCDK1*/dt is the same as in Model 1, as is repeated below.

$$v = \frac{dCDK1^*}{dt} = a_1 - b_1 \left(CDK1^*\right)\left(APC^*\right) = a_1 - b_1 \left(CDK1^*\right)\left(\frac{CDK1^{*n1}}{K_1^{n1} + CDK1^{*n1}}\right) \tag{28.13.6}$$

Likewise, the equation for dAPC*/dt consists of two terms, one for its activation and one for its inhibition.

$$v = \frac{dAPC^*}{dt} = k_f APC - k_r APC^* \tag{28.13.7}$$

Assume that k_f, the rate constant of the activation of APC, is equal to a constant a_2 times a Hill function of CDK1*, and k_r, the rate constant for the inactivation of APC*, is simply b_2. Then the equation becomes

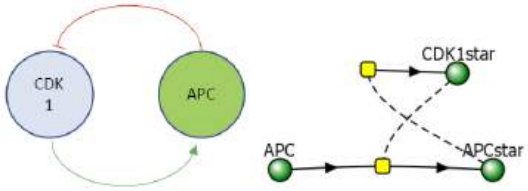
$$v = \frac{dAPC^*}{dt} = a_2 \left(\frac{CDK1^{*n2}}{K_2^{n2} + CDK1^{*n2}}\right) APC - b_2 APC^* \tag{28.13.8}$$

Let's look at the output graphs for the following initial condition:

- CDK1* = 0 uM
- APC = 1 uM
- APC* = 0 uM

VCell MODEL

CDK Model 2



See equations in text.

Select Load [model name] below

Load Ferrell_Tsai_Model2CDK

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

Figure 28.13.15shows statics graphs of just CDK1* activity vs time for n1 values of 1, 4, 8, and 24 for Model 2.

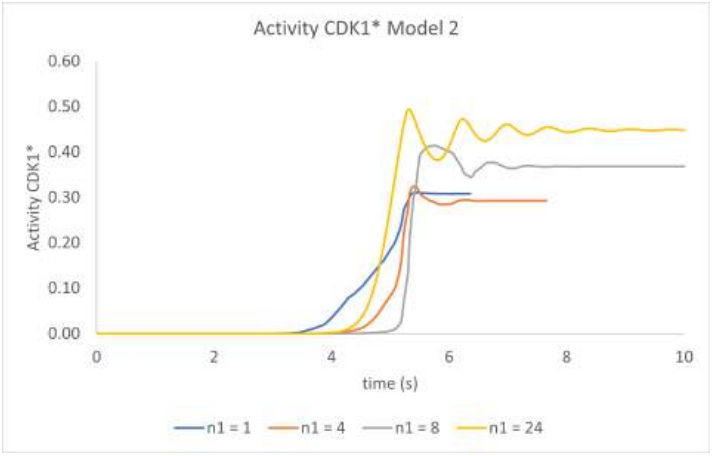


Figure 28.13.15 Graphs of just CDK1* activity vs time in Model 2 for n1 values of 1, 4, 8, and 24.

Wow! By simply adding an additional species to the model and a second differential equation for it, we see the first signs of oscillatory behavior in the activity of CDK1*. The output at higher n1 values is best described as damped oscillations. Now let's try a final third model.

Model 3: Three species model with activation and inhibition-

This model contains the enzyme Plk1 (Polo-like kinase 1, also called serine/threonine-protein kinase 10-A), along with APC and CDK1. These three species are all activated and inhibited. Assume that Plk1 is activated by CDK1 and that it also helps activate APC. Following the arrows in the left part of the figure below shows that it acts as an "intermediary" between CDK1 and APC. Two different representations of a reaction diagram of Model 3 are shown in Figure 28.13.16

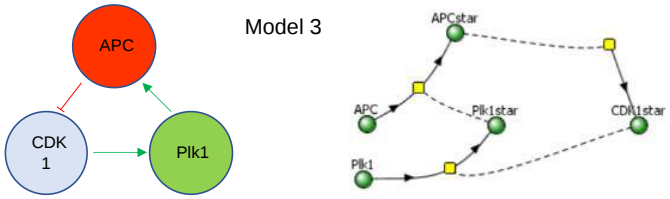


Figure 28.13.16 Two different representations of a reaction diagram of Model 3

We have 3 species, so we need three differential equations, as shown below.

$$v = \frac{dCDK1^*}{dt} = a_1 - b_1 \left(CDK1^*\right)\left(\frac{APC^{*n1}}{K_1^{n1} + APC^{*n1}}\right) \tag{28.13.9}$$

$$v = \frac{dPlk1^*}{dt} = a_2 \left(1 - Plk1^*\right)\left(\frac{CDK1^{*n2}}{K_2^{n2} + CDK1^{*n2}} - b_2 CDK1^*\right. \tag{28.13.10}$$

and

$$v = \frac{dAPC^*}{dt} = a_3 \left(1 - APC^*\right)\left(\frac{Plk1^{*n3}}{K_3^{n3} + Plk1^{*n3}} - b_3 APC^*\right. \tag{28.13.11}$$

The equation for activation of CDK1 is the same as in Models 1 and 2.

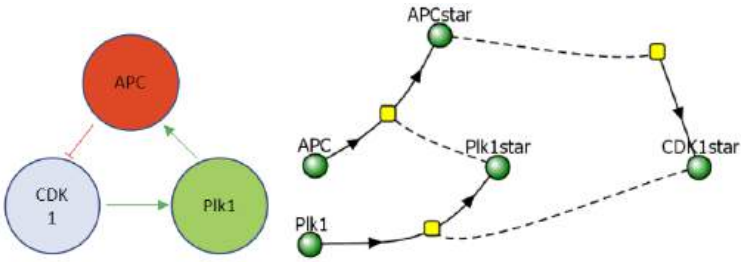
The equation for the activation of APC is similar to Model 2 with k_f modeled as a Hill function of Plk1star, which activates APC

The equation for the activation of Plk1 is similar to Models 2 and 3 with k_f modeled as a Hill function of CDK1star which activates Plk1

Although you probably can't write these differential equations by yourself, hopefully, you can see that they make sense.

VCell MODEL

CDK Model 3



Initial values

J	reaction rate	$(a1 - \frac{b1 \cdot CDK1star \cdot (APCstar)^{n1}}{(K1)^{n1} + (APCstar)^{n1}})$
a1	user defined	0.1
n1	user defined	8.0
K1	user defined	0.5
b1	user defined	3.0
J	reaction rate	$(Kf \cdot APC - Kr \cdot APCstar)$
Kf	forward rate constant	$a3 \cdot \frac{(Plk1star)^{n3}}{(K3)^{n3} + (Plk1star)^{n3}}$
Kr	reverse rate constant	b3
a3	user defined	3.0
n3	user defined	8.0
K3	user defined	0.5
b3	user defined	1.0
J	reaction rate	$(Kf \cdot Plk1 - Kr \cdot Plk1star)$
Kf	forward rate constant	$a2 \cdot \frac{(CDK1star)^{n2}}{(K2)^{n2} + (CDK1star)^{n2}}$
Kr	reverse rate constant	b2
a2	user defined	3.0
n2	user defined	8.0
K2	user defined	0.5
b2	user defined	1.0

Select Load [model name] below

Load Ferrell_Tsai_Model3CDK

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

Figure 28.13.17shows graphs of just CDK1* activity vs time for n1 values of 1, 4 and 8 for Model 3

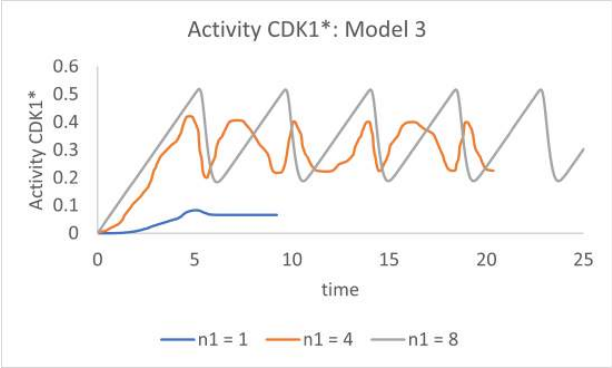


Figure 28.13.17: Graphs of CDK1* activity vs time for n1 values of 1, 4 and 8

Finally, we observe oscillatory behavior in the activity of CDK1*, but only for higher values of the Hill coefficient (n1 = 4 and 8).

Other models can produce oscillation, but this one seems perhaps most comprehensible to students who have studied mass action equations along with Hill binding and kinetic equations. The three-component system described in Model 3 is of course embedded

in a large pathway of inputs and outputs so other factors most likely affect the oscillatory behavior.

28.14: Programmed Cell Death

28.14.1: Introduction

We have discussed often how cell signaling might go awry and lead to cancer. However, there are signaling systems that lead to cell death. There are many ways in which cells can die. We'll discuss not "accidental" cell death but one, **apoptosis**, that is programmed into the genome and highly regulated. Figure 28.14.1 shows how normal cell proliferation and growth can be modulated by two classes of genes, oncogenes that cause proliferation, and tumor suppressor genes that inhibit it.

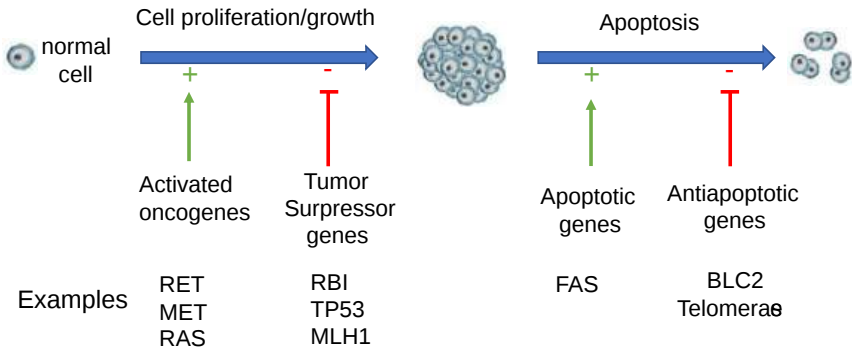


Figure 28.14.1: Overview of oncogenes, tumor suppressor genes, apoptotic genes, and antiapoptotic genes. Adapted from https://www.researchgate.net/figure/...fig1_332774933. Creative Commons Attribution 3.0 Unported

Apoptosis involves chromatin aggregation and cleavage, the concentration of cell material, and apoptosis body formation. Mutations to aberrantly activate oncogenes or inhibit the expression of tumor suppressor genes lead to cancer. These cells would ideally undergo programmed cell death or **apoptosis**. As with control of proliferation, some genes promote apoptosis as well as anti-apoptotic genes which inhibit programmed cell death. Dysregulation of these can also cause cancer. Apoptosis is an important mechanism to kill viral-infected cells. However, this can go too far. For example, T helper cells (T_H) infected with the HIV virus die. However, the collapse of the population of these cells is in part attributed to apoptosis.

As we learn more about programmed cell death, it is clear that apoptosis is not the only way the genome is programmed to cause cell death. These other ways include:

autophagy - This is a catabolic pathway in which intracellular proteins, protein complexes, and organelles are collected into large autophagosomes in which incorporated lysosomes and their degradative enzymes reprocess damaged or unneeded cell material. It is a highly programmed process, which if dysregulated, could lead to cell death.

necroptosis: Infections and toxins are known to cause necrosis, which is a "passive" form of cell death. In contrast, programmed necrosis is called necroptosis.

28.14.2: Overview of apoptosis.

Apoptosis consists of 4 steps:

- the decision to activate the pathway;
- the actual "suicide" of the cell;
- engulfment of the cell remains by specialized immune cells called phagocytes;
- degradation of the engulfed cell.

The actual steps in cell death require:

- condensing the cell nucleus and breaking it into pieces
- condensing and fragmenting of cytoplasm into membrane-bound apoptotic bodies; and
- breaking chromosomes into fragments containing multiple numbers of nucleosomes (a nucleosome ladder)

To commit suicide must be an extremely important cellular decision. Hence you would expect this process to be regulated and highly complicated. When would it be advantageous to the organism to want a cell to kill itself (or be told

to kill itself)? Cell death would be used to:

- "sculpt" an organism during development such as during embryo development, metamorphosis, and tissue atrophy
- regulate the total number of cells.
- defend and remove unwanted or dangerous cells like tumor cells, virally infected cells, or immune cells that recognize self (which could lead to autoimmune disease).

Unregulated apoptosis could exacerbate or cause diseases such as:

- AIDS, in which T helper cell numbers plummet. Part of the dramatic decline in these cells might be caused by health T helper cells being tricked into committing suicide;
- neurodegenerative diseases like Alzheimer's;
- ischemic stroke, when restricted blood flow to certain regions of the brain can lead to neural death through increased apoptosis'
- cancer, in which tumor cells lose their ability to undergo apoptosis;
- autoimmune disease, in which self-reactive immune cells trick normal body cells to kill themselves;
- viral disease;

Apoptosis does not require new transcription or translation, suggesting that the molecular machinery required for cell death lay dormant in the cell, and just requires appropriate activation. What "signals" induce apoptosis?

Signals can be extracellular:

- a hormone (such as thyroxine that causes apoptosis in tadpole tails
- a lack of a "survival" signal (which inhibits apoptosis) such as a growth factor
- a cell:cell contact from an adjacent cell

Signals can be intracellular:

- ionizing radiation
- virus infection
- oxidative damage from free radicals

28.14.3: Apoptosis

Much of this section was derived directly from the following reference, with modifications and additions.

There are two apoptotic pathways in cells:

- The **extr**insic pathway: **extr**acellular apoptotic ligands bind to membrane death receptors, leading to the assembly of the **death-inducing signaling complex (DISC)**. Similar to the inflammatory response we have already discussed in Chapter 5, two specific cysteine-aspartic proteases (caspases) are activated, caspases 8 and 10. These activate other caspases in an amplification of the process.
- The **intr**insic pathway: intracellular signals such as damaged DNA or proteins are sensed by Bcl-2 proteins on the outer membrane of mitochondria. The BcL-2 (B-cell lymphoma-***2***) family of proteins all have Bcl homology domains. Their functions are carried out at the outer mitochondrial membrane. Some members of this family are antiapoptotic (Bcl-2, Bcl-xL, Mcl-1, Bcl-w, A1/Bfl-1, and Bcl-B/Bcl2L10), while others are proapoptotic (Bid, Bim, Puma, Noxa, Bad, Bmf, Hrk, Bik Bax, Bak, and Bok/Mtd). Apoptosis leads to activation of Bax/Bak, which initiate mitochondria degradation, starting with damage to the outer membrane and release of pro-apoptotic proteins like the inner membrane space protein cytochrome C into the cytoplasm. This leads to the assembly of the apoptosome and activation of caspases 9 and 13. Again this is very similar to the formation and activity of the inflammasome which we saw in Chapter 5.

An overview of the extrinsic and intrinsic apoptotic pathways is shown in Figure 28.14.2 We will explore some of the proteins involved in the section below.

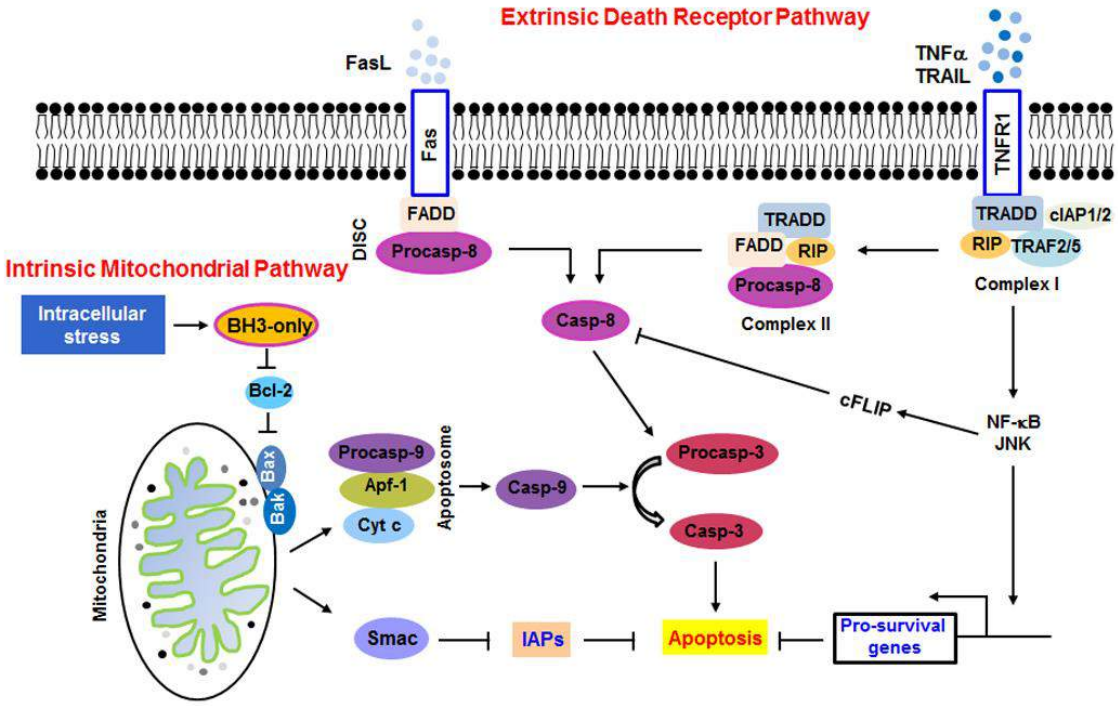


Figure 28.14.2: Overview of the extrinsic and intrinsic apoptotic pathways

The extrinsic death receptor pathway is activated by death receptor ligands, including FasL, TNF-α, DR3, DR4, and DR5 or TRAIL, etc. FasL is an integral membrane protein found in cells. In addition, there are soluble versions of it. The binding of FasL to Fas, an integral membrane protein, initiates the recruitment of FADD, TRADD, and caspase-8 to form the DISC complex, which in turn activates caspase-8 and downstream caspases. The binding of tissue necrosis factor alpha (TNF-α) to its receptor, TNFR1 (a Fas protein), initiates the recruitment of TRADD, RIP, TRAF2/5, and cIAP1/2 to form complex I, which activates NF-κB and JNK pathways and increases the transcription of pro-survival genes. However, the modification of RIP or degradation of cIAP1/2 can lead to the disassociation of complex I. TRADD and RIP then associate with FADD and caspase-8 to form complex II, the so-called death complex.

The intrinsic death receptor pathway is initiated by the BH3-only protein, BCL-2 homology 3 (BH3-only), under intracellular stress such as DNA damage. The BH3-only proteins activate **apoptosis by binding and neutralizing the pro-survival proteins**, allowing Bax/Bak to homo-oligomerize and permeabilize the mitochondria. For example, BH3-only protein can inactivate Bcl-2 and prevent Bcl-2 from effectively neutralizing Bax and Bak, leading to the activation of Bax and Bak. The activated Bax and Bak on the mitochondrial membrane alter its permeability, depolarizes the membrane, and leads to the release of cytochrome c and Smac, normally found in the inner membrane space, from mitochondria. Figure 28.14.3shows how monomeric BAK can form an altered dimeric form in the presence of detergent.

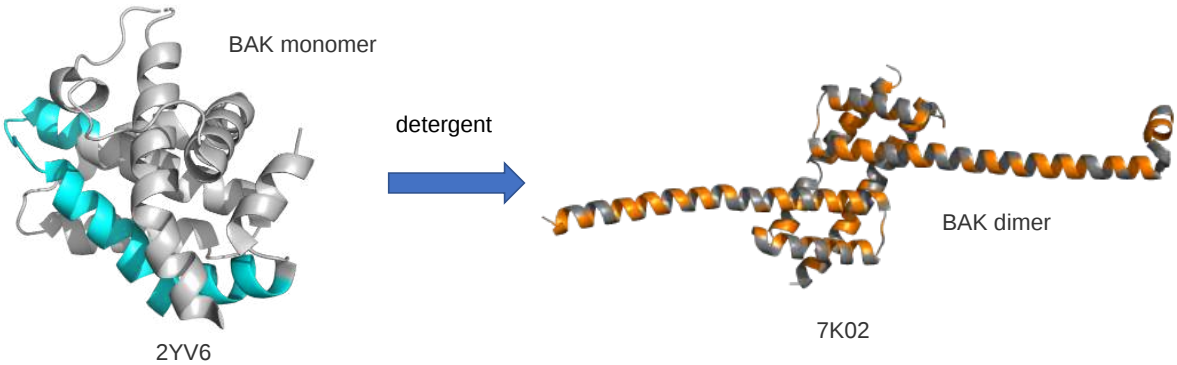


Figure 28.14.3: Monomeric and membrane associated dimeric BAK

The extended left-hand helix on the right-hand side is color-coded to show nonpolar residue (orange) and the polar/charged amino acids in gray. That same section of the protein is shown in cyan in the monomeric protein to the right. One can easily imagine how the apparent amphiphilic helices of the BAK dimer could bind to the outer mitochondrial membrane and alter its structure.

Cytoplasmic cytochrome c associates with Apaf-1 and caspase-9 to form the apoptosome, which activates caspase-9 and downstream executing caspases. Smac can regulate apoptosis by inhibiting the inhibitor of apoptosis proteins (IAPs)." Zhou and Li. Chapter 9, Apoptosis in Polycystic Kidney Disease: From Pathogenesis to Treatment.

Another diagram of the extrinsic and intrinsic apoptotic pathways that show more detail on the domain structures of some key protein and the "executioner" caspases is shown in Figure 28.14.4

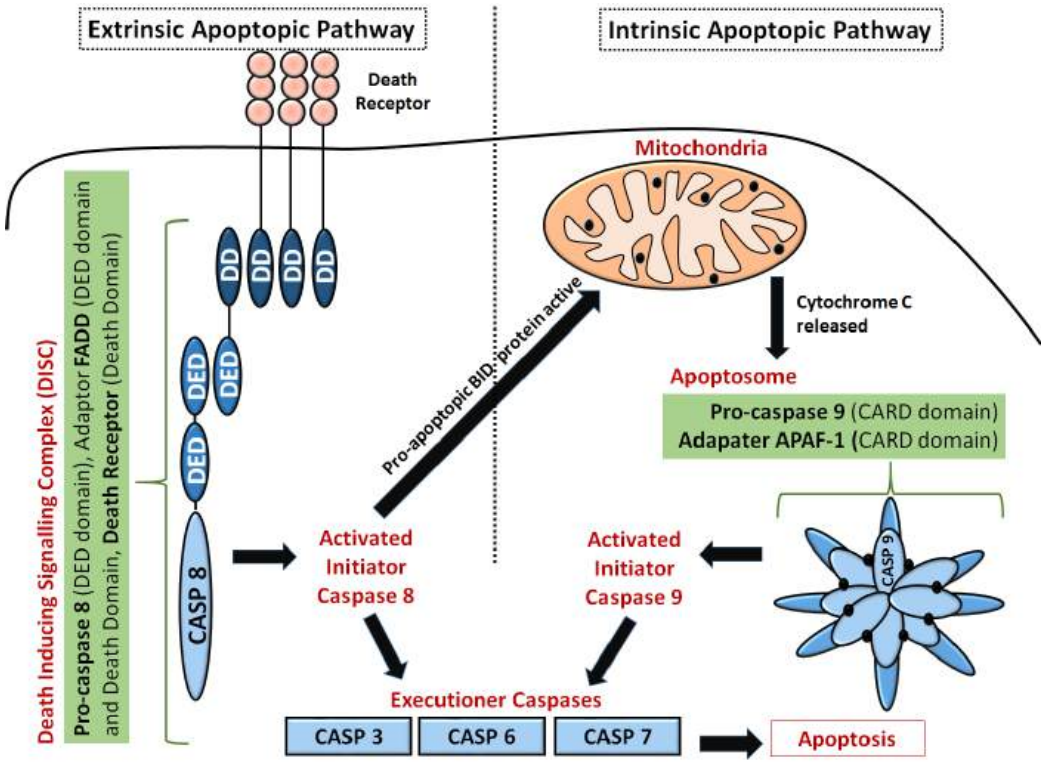


Figure 28.14.4: Extrinsic and intrinsic apoptotic pathways https://commons.wikimedia.org/wiki/File:Apop1.png

Extrinsic pathway: The first step is an association of death receptors with their cognate ligands, which leads to the recruitment of adaptor molecules, including FAS-associated death domain protein (FADD), and then caspase 8. Caspase 8 cleaves and activate caspase 3 and caspase 7 and can proteolytically activate BH3-only protein BH3-interacting domain death agonist (BID). Proteolytically activated BID (tBID) promotes mitochondrial membrane permeabilization through the activation of the assembly of BAX-BAK channels and represents the main link between the extrinsic and intrinsic pathways.

Now let's look more closely at the ligands that activate the extrinsic pathway, as shown in Figure 28.14.5

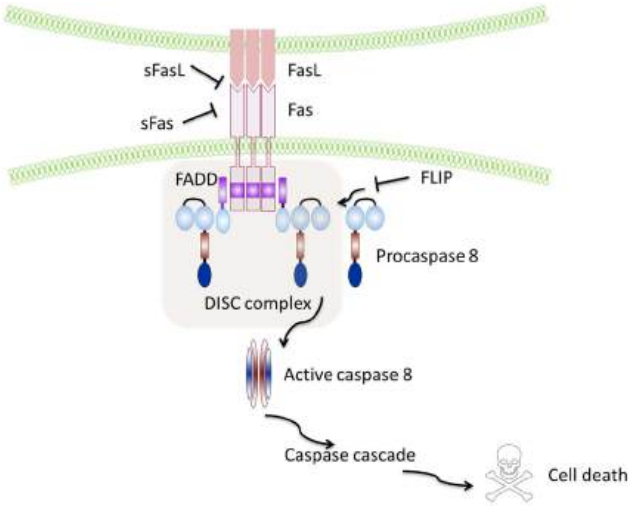


Figure 28.14.5: Fas/FasL activation pathway

Soluble Fas and soluble FasL bind to the respective ligands inhibiting activation of the pathway. FLIP inhibits the activation of caspase-8 and is thus a major anti-apoptotic protein. Volpe E et al. (2016) Fas–Fas Ligand: Checkpoint of T Cell Functions in Multiple Sclerosis. Front. Immunol. 7:382. doi: 10.3389/fimmu.2016.00382. Creative Commons Attribution License (CCBY).

Now we are in a position to examine the actual structure of some key components of the extrinsic pathway.

28.14.4: Active human apoptosome with procaspase-9 (5JUY)

Figure 28.14.6shows an interactive iCn3D model of the active human apoptosome with procaspase-9 (5JUY)

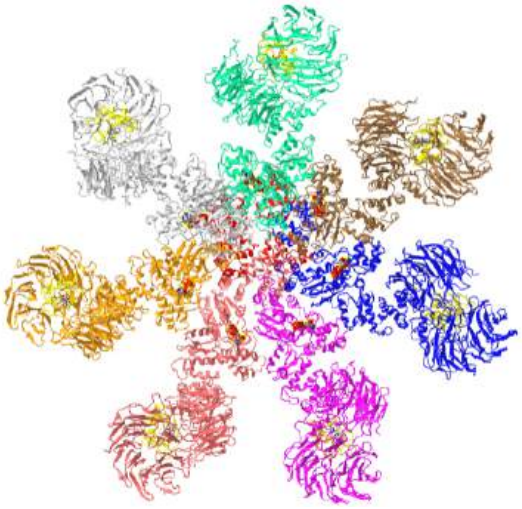


Figure 28.14.6: Active human apoptosome with procaspase-9 (5JUY) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...ek6wzp36TGUpp6 (long load time)

Each of the 7 different subunits of the apoptotic protease-activating factors (Apaf-1) is shown in a different color. The seven yellow subunits are cytochrome Cs. The 4 red subunits underneath the disk plane of the other subunits are the zymogen procaspase 9s. The small spacefill CPK color ligands are 2'-deoxyadenosine 5'-triphosphate. The Apaf-1:pc9 pairs, interacting through their CARD domains, form a spiral underneath the disk.

Apaf-1 can be considered an adaptor protein with an N-terminal **c**aspase **a**ctivation and **r**ecruitment **d**omain (**CARD**), followed by a **nucleotide-binding and oligomerization domain (NOD**, also known as **NB-ARC**).

Figure 28.14.7shows the domain structure of caspase 9 and Apaf-1.

Figure 28.14.7: domain structure of caspase 9 and Apaf-1.

The presence of CARD domains in both allows their mutual binding and the assembly of the full apoptosome.

28.14.5: An AlphaFold model of the Cas 9 zymogen

Figure 28.14.8shows an interactive iCn3D model of human Cas 9 AlphaFold model (P55211)

Figure 28.14.8: Human Cas 9 AlphaFold model (P55211) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...74TiVhNfiX4376

The green is the CARD domain and the salmon is the caspase (peptidase_C14) domain. Procaspase 9 is cleaved at Asp 315 (sticks, CPK colors, labeled) into two chains for activation. The activated Cas 9 has two key active site residues, His 237 and the catalytic nucleophile C287 (sticks, CPK colors, labeled). Phosphorylation at Thr-125 by MAPK1/ERK2 blocks procaspase activation by proteolysis. to block caspase-9 processing

28.14.6: Apaf-1

Oligomeric Apaf-1 mediates the cytochrome c-dependent autocatalytic activation of pro-caspase-9 (Apaf-3), leading to the activation of caspase-3 and apoptosis

Figure 28.14.9shows an interactive iCn3D model of Human Apaf-1 AlphaFold model (O14727)

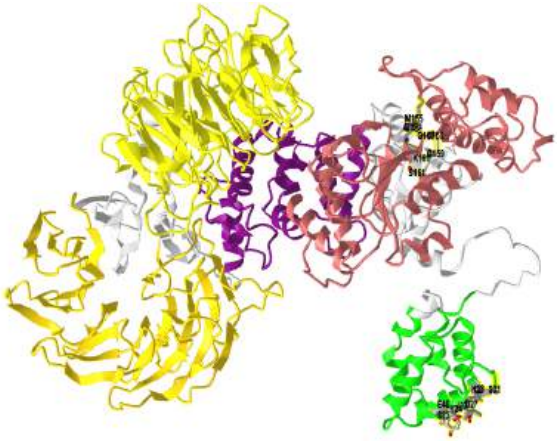


Figure 28.14.9: Human Apaf-1 AlphaFold model (O14727) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...9ZrDFxEpdHR9K6

Domain colors:

- The green is the N-terminal CARD domain
- light red NB-ARC (nucleotide-binding and oligomerization domain - NOD)
- purple is Apaf
- yellow is WD40, gold the C-terminal WD40.

Again the model above does not show the actual structure since the intrinsically disordered regions are not more structured.

The CARD domain of Cas 9 inhibits the catalytic domain of Cas 9. When the CARD domain of Cas 9 interacts with the CARD domain of Apaf-1, the autoinhibition is removed. In addition, the Apaf-1 stimulates the catalytic activity of the protease.

Before assembly into the apoptosome, Apaf-1 is monomeric and in an inactive dATP or ATP conformation. When cytochrome C is released into the cytoplasm, it binds to the WD domains, facilitating a dATP/ATP-cleavage associated conformation change in the Apaf-1, which in the presence of heat shock protein 70 (Hsp) folds to form which leads to the assembly of the active apoptosome.

28.14.7: Fas - **Tumor necrosis factor receptor superfamily member 6 -** P25445

Figure 28.14.10shows an interactive iCn3D model of Fas-Tumor necrosis factor receptor superfamily member 6 AlphaFold model (P25445)

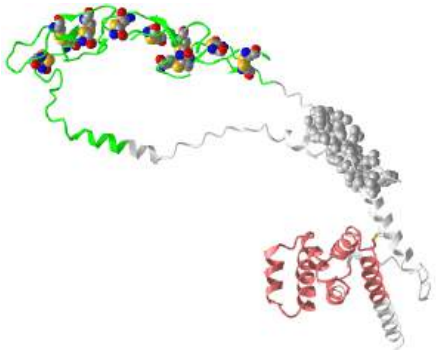


Figure 28.14.10: Fas-Tumor necrosis factor receptor superfamily member 6 AlphaFold model (P25445) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...Dpm3neKspeCrP8

The green is the N-terminal TNFR/NGFR domain that is highly enriched in Cys (spacefill, color CPK) in disulfide bonds. The gray spheres are the transmembrane helix. The Red shows the Death Domain.

The death domains are common protein:protein binding domains that serve as adaptors or scaffolds. They can form homo- or heterodimers with other proteins containing the domain which is a part of the CARD domain, DED (Death Effector Domain), and PYRIN.

28.14.8: Human FasL and a soluble Fas Receptor DcR

Figure 28.14.11shows an interactive iCn3D model of the complex of Human FasL and Its Decoy Fas Receptor DcR (4MSV)

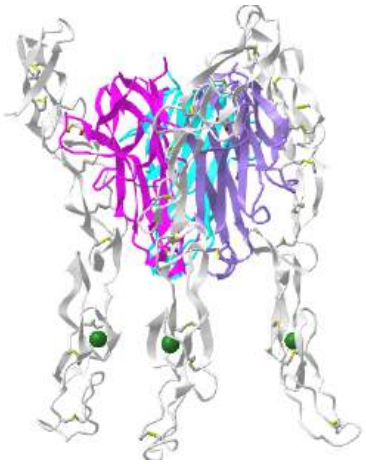


Figure 28.14.11: Complex of Human FasL and Its Decoy Fas Receptor DcR (4MSV). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...bHLh8fBNZ7E5Q9

The three gray subunits are soluble decoy receptor (DcR) versions of the Fas TNFR/NGFR domain, which again is highly enriched in Cys. It is structurally very similar to its typical membrane receptor ligand Fas (**tumor necrosis factor receptor superfamily member 6** - P25445)

DcR is a secreted member of the TNF family and disrupts apoptosis, which can allow tumors to survive.

28.14.9: Fas and FADD death domain interactions

Figure 28.14.12shows an interactive iCn3D model of two Fas death domains bound to two FADD death domains (3EZQ)

tetrameric arrangement of four FADD death domains bound to four Fas death domain

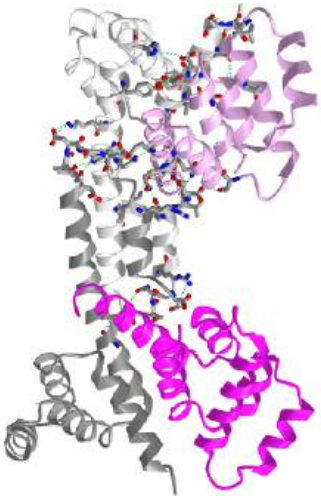


Figure 28.14.12: two Fas death domains bound to two FADD death domains (3EZQ). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...S9Uj45N4mqT5G8

Two Fas death domains are shown in a different shade of gray. The two FADD death domains are shown in a different shade of magenta. Each domain consists of six alpha-helical bundles. Interaction between the dark and light grays Fas death domains and between the light gray Fas and light magenta FADD death domains are shown with side chains in stick with CPK colors.

The Fas-FADD-procaspase 8 complex is collectively called the **Death Inducing Signaling Complex (DISC)**. The Fas-FADD interactions lead to the binding of capase 8 and the completion of the DISC. The actual disc appears to contain 4 FADD death domains bound to 4 Fas death domains. Conformational changes allow both FADD:Fas and Fas:Fas interactions, some of which are weak but when formed switch on the activity of the building complex. The need for 4 monomers probably prevents accidental assembly which would be deadly to the cell.

28.14.10: **Mechanism and regulation of apoptosis**

Caspases

Characterization of apoptotic mechanisms and cellular players started with the study of *C. elegans*, a roundworm. The mature worm has about 1000 cells. During development, 131 cells die. Two mutations were found in which the 131 cells did not die. These mutations were called ced3 and ced4 (**ced** stands for **ce**ll **d**eath). The sequence of ced 3 was very homologous to a protein called **i**nterleukin **c**onverting **e**nzyme (**ICE**) which is required for proteolytic activation of the precursor to interleukin 1, a protein hormone released by certain immune cells during activation and which can promote inflammation. This suggested that proteolysis was required for apoptosis. Subsequent studies show that a whole family of proteases (about 10 in humans) called caspases (ICE has been renamed caspase 1) are required for programmed cell death. These proteases are found in the cell in an inactive form which must undergo limited proteolysis for activation. These caspases form a cascade of proteases which are activated in this process. They are endoproteases that have an active site Cys (**C**) and cleave at the C-terminal side of Asp residues (**asp**) and hence are known as **casp**ases - **c**ys containing-**asp** specific prote**ases**).

ICE is not normally involved in apoptosis, but its artificial activation in cultured mammalian cells can lead to it. Each caspase had the same sequence as they are designed to cleave, so it became evident that they probably cleave each other in an activation cascade mechanism, similar to the coagulation protease cascade of activation of precursors (zymogens) of serine proteases which activate the next in the series. Two series of caspases seem to be involved. One set initiates the process of caspase activation. Just as in the clotting system, the question of what activates the first caspase appeared problematic until investigators found that the initiator caspase can be activated if they aggregate to a critical concentration. This could occur by binding a suicide signal molecule at the cell surface. Conformational

changes in the receptor can lead to aggregation of surface receptor molecules with concomitant aggregation of intracellular caspases which interact with the aggregated receptors.

Intracellular signals

How might intracellular activators of apoptosis (like radiation or reactive oxygen species) work? Research indicated the involvement of mitochondria in the apoptotic pathway. Believe it or not, **cytochrome C**, the heme protein which acts as a water-soluble mobile carrier of electrons in mitochondrial oxidative phosphorylation, shuttling electrons through cytochrome C oxidase or complex IV, leaks out of the intermembrane space and binds to a cytoplasmic protein called **Apaf-1** for **a**poptotic **p**rotease **a**ctivating **f**actor-**1**. This then activates an initiator caspase-9 in the cytoplasm.

These proteins seem to leak out of mitochondria after a collapse of the electrochemical potential across the inner membrane. The potential collapses as a consequence of the opening of a channel called a nonspecific inner membrane **permeability transition pore**, composed of both an inner membrane protein (**a**denine **n**ucleotide translocator - **ant**) and an outer membrane protein (porin, the **v**oltage-gate**d** **a**nion **c**hannel - **VDAC**). These proteins act together, probably at sites where the inner and outer membranes are in contact. This channel passes anything smaller than molecular weight 1500. Collapsing the proton gradient uncouples oxidation and phosphorylation in the mitochondria. Changes in ionic strength cause a swelling of the matrix. Since the inner membrane is highly convoluted and has a much greater surface area than the outer membrane, swelling of the matrix leads to a rupture of the outer membrane, spilling the inner membrane space proteins (cytochrome C and Apaf-1) into the cytoplasm.

What causes all these changes in the mitochondria? Several interrelated events appear to be involved:

1. disruption of ox-phos. and electron transport, caused by irradiation and certain second messengers such as ceramide.
2. changes in cell redox potential and generation of **r**eactive **o**xygen **s**pecies (**ROS**).
3. DNA damage (caused by radiation, ROS, etc). A protein called p53 is often expressed in cells with DNA damage. Expression of this protein results in inhibition of cell division, or apoptosis, both of which would keep the damaged cell from becoming a tumor cell. Hence the p53 gene is a tumor suppressor gene. It is inactivated by mutation in approximately 50% of all human tumor cells studied. p53 can induce gene expression. Of the 14 different genes whose expression is significantly altered by p53, many seem to be used by cells to generate or respond to oxidative stress. Cells undergo p53 apoptosis through oxidative damage.
4. increases in intracellular calcium ions through signal transduction.

Caspase targets:

Apoptosis involves:

1. condensing of the cell nucleus and breaking into pieces
2. condensing and fragmenting of cytoplasm into membrane-bound apoptotic bodies
3. breaking chromosomes into fragments containing multiple numbers of nucleosomes (a nucleosome ladder)

How does caspase activation lead to these events? A protein has been uncovered that when cleaved by a caspase leads to a nuclear breakup. The target protein is usually bound to another protein, a DNA endonuclease. When the target protein is cleaved, the DNase is free to migrate to the nucleus and begin the execution. Membrane changes in apoptosis occur when caspase 3 cleaves gelsolin, a protein involved in maintaining cell morphology. The cleaved gelsolin cleaves actin filaments inside the cell. Another protein is necessary to form apoptotic bodies: a kinase named **p21-activated kinase 2 (PAK-2)**. This kinase is activated by caspase-3 by limited proteolysis. Caspases also cleave beta-amyloid precursor protein which might generate more beta-amyloid protein, causing neural cell death in Alzheimer patients.

Controlling Apoptosis

It should be clear that cells keep tight control of the caspases. Two players which appear to inhibit apoptosis are the mitochondrial proteins **Bcl-2** and **Bcl-X**, which can block the release of cytochrome C from the mitochondria. The Bcl family of proteins has a hydrophobic tail and binds to the outside surface of mitochondria and other organelles like the

nucleus and endoplasmic reticulum. These proteins seem to be able to form ion channels in liposomes. So far 15 members of this family (related to ced-9 of C. elegans) have been discovered in humans. Bcl-2 can also bind to Apaf-1 (mentioned above) and inhibit its activation of initiator caspase-9. Bcl-2 is regulated by changes in the expression of the Bcl-2 gene, by post-translational phosphorylation by kinases, or by cleavage by caspases. Overexpression of Bcl-2 can cause a cell to become a tumor cell. Other members of the family, **BAX** and **BAD** bind to mitochondria and facilitate apoptosis by stimulating cytochrome C release.

In addition, other proteins called **IAP**s (**i**nhibitors of **ap**optosis) can inhibit caspase or other apoptotic proteins. Some viruses make the protease to keep their host cells viable.

Cell Membrane Events

Cells can be instructed to undergo apoptosis through cell surface interactions with other cells which are often immune cells. One of the jobs of the immune cell is to destroy an altered cell (for example a virally-infected cell or a tumor cell). Immune cells themselves must also die after they are activated in an immune response. Activated lymphocytes (like cytotoxic T cells or natural killer cells) can target and kill cells using several ways which can involve apoptosis. In one, an activated lymphocyte binds to a target cell (like a virally infected cell) and secretes perforin, a protein that assembles in the target cell membrane to form a transmembrane channel. Other proteins released by the activated lymphocyte can enter the target cell through the pore and initiate apoptosis. One such protein that enters, granzyme B is a protease that activates caspases in the target cell.

Target cells that express a specific membrane protein called **CD95 (**also called **Fas**) are also targeted for apoptosis. This protein receptor, a member of the tumor necrosis factor receptor (TNFR) binds to a membrane protein-ligand on the surface of an activated lymphocyte called CD95 Ligand - CD95L- (also called the **Fas ligand**). On binding, the CD95 (Fas) receptors on the target membrane aggregate after conformation changes. An adapter protein in the cell, FADD (Fas-associated death domain) binds to the aggregated cytoplasmic domain (the death domain) of CD95 (Fas) and recruits inactive caspase-8 to the site, where their concentration increases. This leads to the activation of the caspases.

This mechanism is used to get rid of activated lymphocytes after they have finished their work. Activated immune cells start expressing Fas a few days after activation, targeting them for elimination. Some cells which have been stressed express both Fas and Fas ligands and kill themselves. Various cells express CD95 (Fas), but CD95L (Fas-Ligand) is expressed predominately by activated T cells.

Cell surface events also can inhibit apoptosis. The binding of "survival" factors (like growth factors) to cell surface receptors can shut off apoptotic pathways in the cells. Some survival factor receptors are coupled to **PI-3-kinase (phosphoinositol-3-kinase**) through the G protein **ras (p21)** which is targeted to the cell membrane by post-translational addition of a hydrophobic anchor. The activated kinase produces PI-3,4-P2 and PI-3,4,5-P3, which activates **Akt**, a Ser/Thr protein kinase. This activated kinase phosphorylates the proapoptotic-protein BAD, which then becomes inactive. In addition, active Akt phosphorylates procaspase, which in its phosphorylated form will not interact with cytochrome C, hence inhibiting apoptosis.

The endpoint of apoptosis is the engulfment of the fragmented cell by a phagocytic cell (such as a macrophage). In a recent article (Nature, 405, pg 85, 2000), it was shown that the activity of phagocytes could be inhibited stereospecifically by the addition of phosphatidyl serine (PS) to the mixture, but not by other negative phospholipids. If you remember from our description of lipids, PS is found exclusively in the inner leaflet of red blood cells). The investigators cloned a gene from the phagocytic cell for a receptor that recognizes PS. When added to ordinary T and B lymphocytes (immune cells), these cells could also take up apoptotic cells. The gene is homologous to genes in Drosophila (fruit fly) and *C. elegans* (roundworm) suggesting that it is conserved in nature. The message: when cells undergo apoptosis, PS, normally found only in the inner leaflet, is exposed to the outside. It can then bind to receptors on phagocytic cells to complete the process of apoptosis.

28.14.11: Therapeutics

Figure 28.14.13shows points of therapeutic intervention in the intrinsic and extrinsic apoptotic signaling pathways.

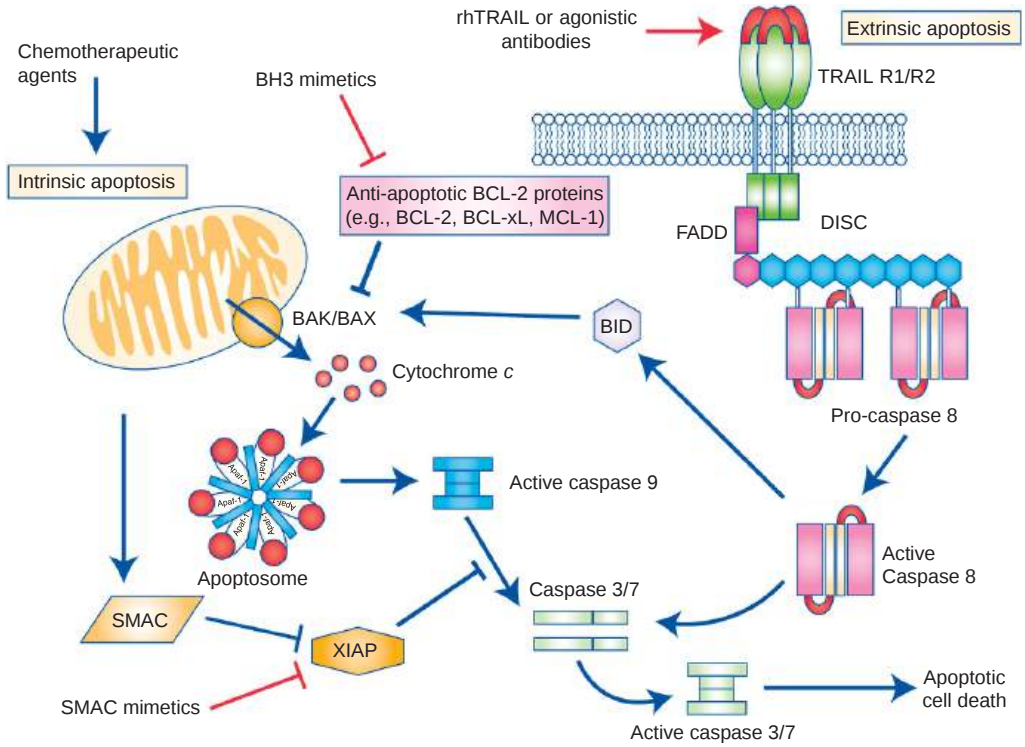


Figure 28.14.13: **Intrinsic and extrinsic apoptotic signaling pathways and points of therapeutic intervention.** Fox, J., MacFarlane, M. Targeting cell death signaling in cancer: minimizing 'Collateral damage'. *Br J Cancer* **115,** 5–11 (2016). https://doi.org/10.1038/bjc.2016.111. Creative Commons Attribution 4.0 International License. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/

Intrinsic and extrinsic apoptotic signaling pathways and points of therapeutic intervention. Apoptosis can be initiated by signals originating from either the plasma membrane via death receptor ligation (extrinsic pathway) or at the mitochondria (intrinsic pathway). Stimulation of the extrinsic pathway by TRAIL results in TRAIL receptor (TRAIL-R) aggregation and formation of the DISC, in which pro-caspase 8 becomes activated and initiates apoptosis by direct cleavage of downstream effector caspases. The addition of either agonistic TRAIL-R1/R2 antibodies or recombinant human TRAIL (rhTRAIL) has been used to trigger the extrinsic pathway for therapy. The intrinsic pathway is regulated by the BCL-2 family of proteins, which regulate pore formation in the outer mitochondrial membrane and the release of apoptogenic factors such as cytochrome *c* or SMAC from the mitochondria. The release of cytochrome *c* into the cytosol triggers caspase 9 activations through the formation of the cytochrome *c*/Apaf-1/caspase 9-containing apoptosome complex. SMAC promotes caspase activation through the neutralizing the inhibitory effect of IAPs. The intrinsic pathway has been targeted for therapy either by blocking the inhibitory action of the pro-survival BCL-2 family proteins with BH3 mimetics or by inhibiting the anti-apoptotic action of IAPs with SMAC mimetics. The extrinsic and intrinsic pathways are interconnected, for example, by BID, a BH3 domain-containing protein of the BCL-2 family, which upon cleavage by caspase 8 triggers intrinsic apoptosis, thereby further amplifying the signal from the extrinsic pathway.

28.14.12: The entire pathway

Now we can present detailed pathways showing apoptosis in its complexity. Trace the interconnections in the different views.

Figure 28.14.14 View 1

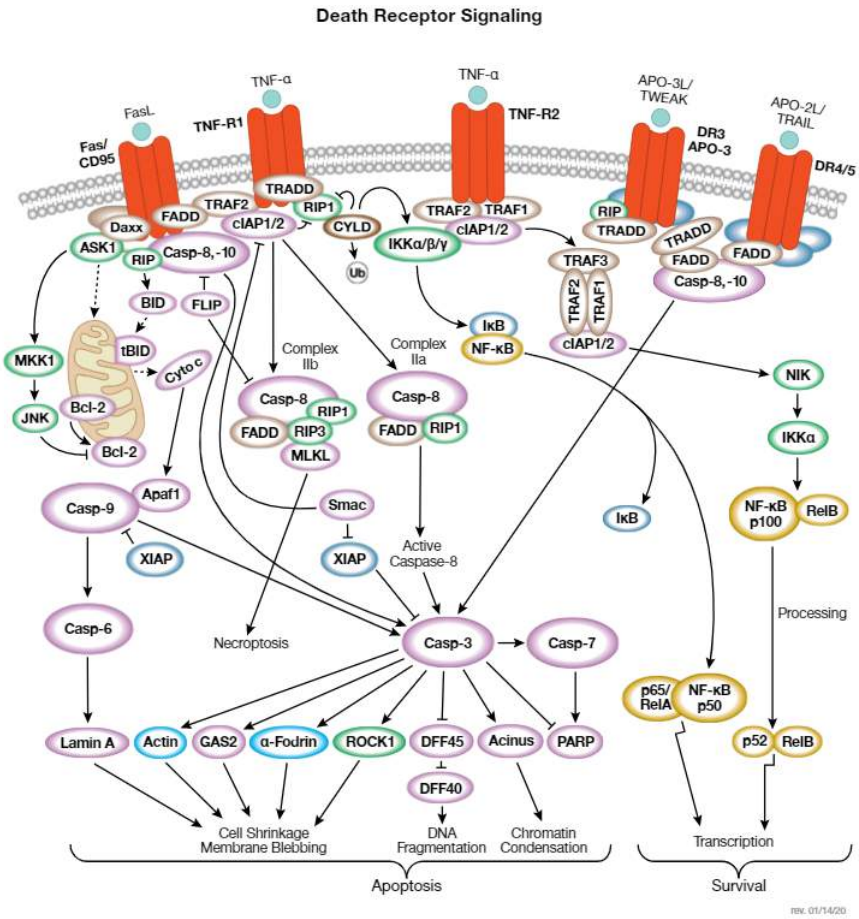


Figure 28.14.14: View 1 from Cell Signaling Technologies (with permission) https://www.cellsignal.com/pathways/...ptor-signaling

Figure 28.14.15presents a second view

Pathway 2 https://www.sinobiological.com/pathw...ceptor-pathway

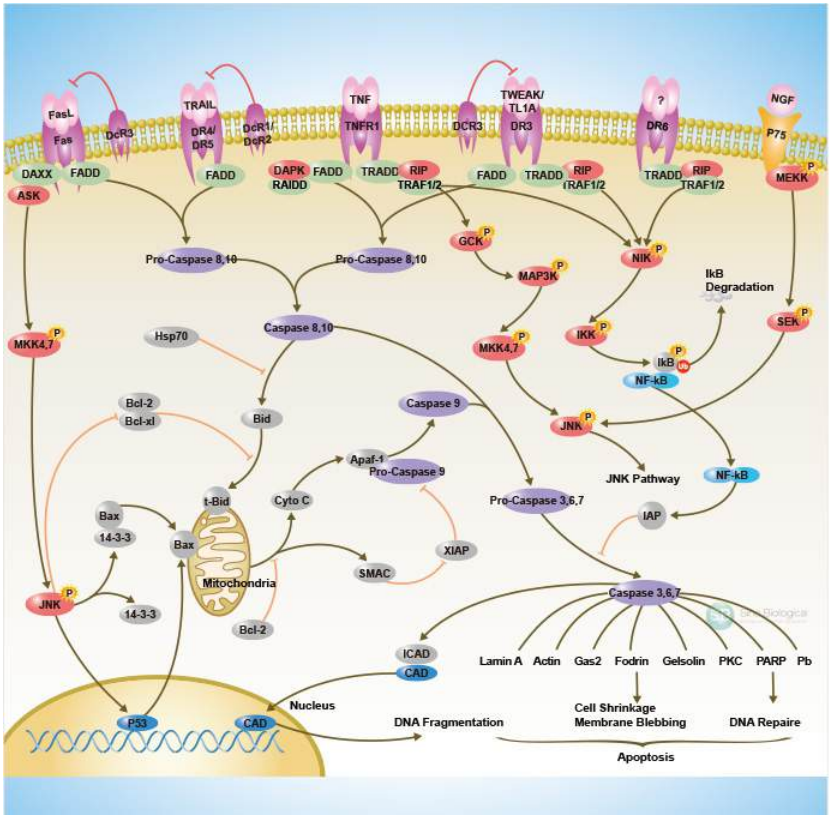


Figure 28.14.15: View 2 from Sinobiological https://www.sinobiological.com/pathw...ceptor-pathway

28.15: Signaling in Microorganisms

The main organization of this section derives from **Bacterial transmembrane signaling systems and their engineering for biosensing. Jung et al :25 April 2018https://doi.org/10.1098/rsob.180023.** Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/. Significant content from the source has been integrated into the section.

28.15.1: Introduction

Bacteria constantly interact with their surroundings. They identify and actively acquire nutrient resources, sense and respond to environmental stresses, and exchange information with other cells, while commensals and pathogens adapt their lifestyles for survival in their hosts. The cytoplasmic (inner) membrane of bacterial cells separates the cytoplasm from the outer world. Therefore, all information from the outside must be transferred across this interface, which contains various sensors that carry out this function.

Bacteria use three major types of signaling systems: membrane-integrated one-component systems (for example -ToxR-like receptors), two-component systems consisting of a **receptor histidine kinase** and a **response regulator**, and extracytoplasmic function (ECF) sigma factors. These are shown in Figure 28.15.1.

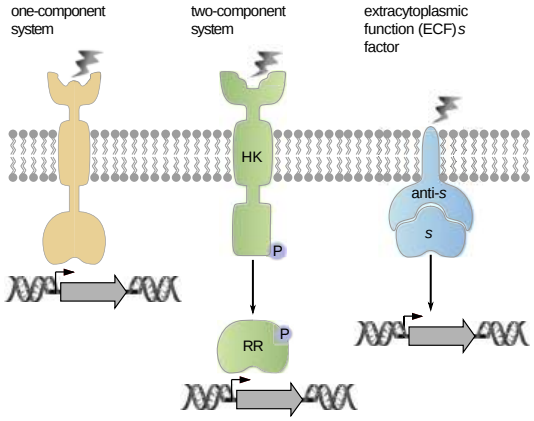


Figure 28.15.1: Schematic presentation of the major types of transmembrane signaling systems in bacteria. One-component signaling systems, consisting of sensor and DNA-binding domain (yellow), two-component systems with a membrane-integrated histidine kinase (HK) and a response regulator (RR) (green), and extracytoplasmic function (ECF) sigma factors (σ) that will be released from the anti-sigma factor (anti-σ) upon stimulus perception (blue).

The one-component signaling family ToxR (named after the main regulator of virulence in *Vibrio cholerae*) is the simplest. They have a periplasmic sensor domain, a single transmembrane helix, and an intracellular winged helix-turn-helix DNA-binding domain. The family is named after the main regulator of virulence in *Vibrio cholerae*, ToxR.

In two-component systems, the membrane-integrated histidine kinase generally acts as a sensor for various stimuli and is also responsible for information transfer across the membrane. This process usually results in the autophosphorylation of the protein and the phosphoryl group is subsequently transferred to a specific soluble response regulator which usually acts as a transcription factor (see Figure 28.15.1). The number of histidine kinase/response regulator systems varies widely between bacterial species, ranging from 30 in *Escherichia coli* and 36 in *Bacillus subtilis* to 132 in *Myxococcus xanthus*. In chemotactic systems, a soluble histidine kinase perceives the signal(s) conveyed by membrane-integrated chemoreceptors and transduces this information via phosphorylation/protein–proteins interaction to the flagellar motor.

The ECF sigma factors are small regulatory proteins that bind to RNA polymerase and stimulate the transcription of specific genes. Many bacteria, particularly those with more complex genomes, contain multiple ECF sigma factors, and these regulators often outnumber all other types of sigma factors. Little is known about the roles or the regulatory mechanisms employed by the majority of ECF sigma factors. Most of them are co-expressed with one or more negative regulators. Often, these regulators include a transmembrane protein that functions as an anti-sigma factor, which binds and inhibits the cognate sigma factor.

Let's look at three examples.

28.15.2: One-component system: pH sensor CadC

pH in E. Coli is regulated by a series of Cad proteins. CadA is a cytoplasmic decarboxylase, which converts lysine to cadaverine, while CadB is a membrane-integrated lysine/cadaverine antiporter. CadC acts as a homodimeric one-component regulator. Together, their activities lead to an increase in both internal and external pH, which favors the survival of *E. coli* under moderate acid stress and helps to maintain pH homeostasis. Their activities are shown in Figure 28.15.2

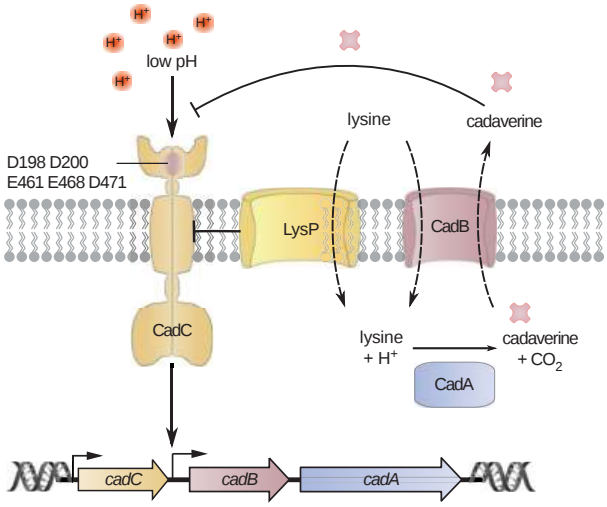


Figure 28.15.2: The regulation of CadC, a one-component system representative.

CadC is the regulator of the *cadBA* operon encoding the lysine decarboxylase CadA and the lysine/cadaverine antiporter CadB. Under non-inducing conditions, the lysine-specific transporter LysP inhibits CadC. When cells are exposed to low pH in the presence of lysine, the interaction between LysP and CadC is weakened, rendering CadC susceptible to protonation and transcriptional activation. The end-product of decarboxylation, cadaverine, binds to CadC and thereby inactivates this receptor.

CadC is activated by two stimuli, low pH (less than 6.8) and the presence of external lysine (greater than 0.5 mM), which are perceived by different mechanisms. The periplasmic domain of CadC directly senses a decrease in pH. It has two distinct subdomains: the N-terminal subdomain comprises a mixture of β-sheets and α-helices, and the C-terminal subdomain consists of a bundle of 11 α-helices. A patch of acidic amino acids (D198, D200, E461, E468, D471) detect changes in the external pH through protonation changes, altering their charges and noncovalent interactions between the subdomains/monomers. This in turn leads to dimer formation of the periplasmic domain, triggering receptor activation.

Figure 28.15.3shows an interactive iCn3D model of a Transcriptional activator CadC One Component Model - AlphaFold Model (P23890)

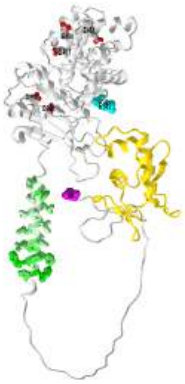


Figure 28.15.3: Transcriptional activator CadC One Component Model - AlphaFold Model (P23890) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...UfZh5LVeoaToj6

The gold represents the N-terminal CadC DNA binding domain. The green (hydrophobic) is the transmembrane helix and the gray is the outside periplasmic domain containing the acidic side chains (D198, D200, E461, E468, D471). The N term Met is magenta spacefill and the C-terminal Ser is cyan spacefill. Again, the long intrinsically disordered region between the DNA binding domain and the transmembrane domain would have a more defined structure in the actual membrane-bound form and probably form

additional interactions with other molecules. It also presumably leads to a conformation change in the DNA binding domain which leads to its interaction with the target DNA.

CadC senses external lysine only in interaction with the lysine-specific permease LysP. In addition, the products of lysine decarboxylation, CO_2, and cadaverine, act as feedback inhibitors on CadC. Cadaverine binds to the periplasmic domain of CadC, thereby switching off *cadBA* transcription.

28.15.3: Two Component System

Two-component Histine kinase systems have (in general) two main components

- A HK **sensor protein** binds a ligand in a receptor binding domain leading to the transfer of a gamma-phosphate from ATP to a His by the kinase domain. This component is often called a **transmitter**. The conserved His located in the **H box**.
- a separate **response regulator** (or effector) protein containing a reactive Asp which receives the phosphate from p-Histidine. The conserved Asp (D) is located in the **D box**. This activates the response regulator protein. This component is often called the **receiver**. It may also transfer it to another His in a phospho-relay system.

Histidine kinase/response regulator systems are the most commons in bacterial signaling across the membrane. In contrast to the myriad of serine/threonine (S/T) and less abundant tyrosine (Y) kinases that dominate signaling in mammals, the histidine kinase predominates in bacterial signaling.

Before we present more detail on the two-component system, let's looks at protein kinases in general and see what is different about histidine kinases. ATP is a donor of a gamma-phosphate in both S/T/Y and H kinases. However, their products are very different energetically. pS, pT, and pY of the O-phosphoproteome are all phosphoesters, which are not high energy compared to their hydrolysis products. (**Remember, there is no such thing as a "high energy" bond.**) In contrast, pHis, a member of the N-phosphoproteome (along with pLys and pArg), is not a phosphoester but more analogous to the mixed anhydride of a carboxylic acid as in the case of phosphorylated aspartic acid. In pHis and pAsp, there is an electronegative N (in pHis) and O (in pAsp) bridging two atoms which are each connected to another atom by double bonds. This type of structure, which allows for bridging resonance between the center N (in pHis) and O (in pAsp) is also high energy compared to its hydrolysis products. Hence the phosphorylation of His by ATP is not as energetically favored as the phosphorylation of Ser, Thr, or Tyr since it produces another high-energy molecule (with respect to its hydrolysis products.

Since the pHis is also considered high energy compared to its hydrolysis products, it can act as a phosphate donor to another receiving group. That could be water in a simple hydrolysis reaction or, if sheltered from water in an active site of an enzyme or receptor, to a carboxylate receiver like Asp to form another high energy mixed anhydride which is isoenergetic with the pHis. This is the process that occurs in the two-component His kinase signaling pathways in bacteria. Figure 28.15.4compares Ser, Thr, Tyr, and His kinase reaction and their products.

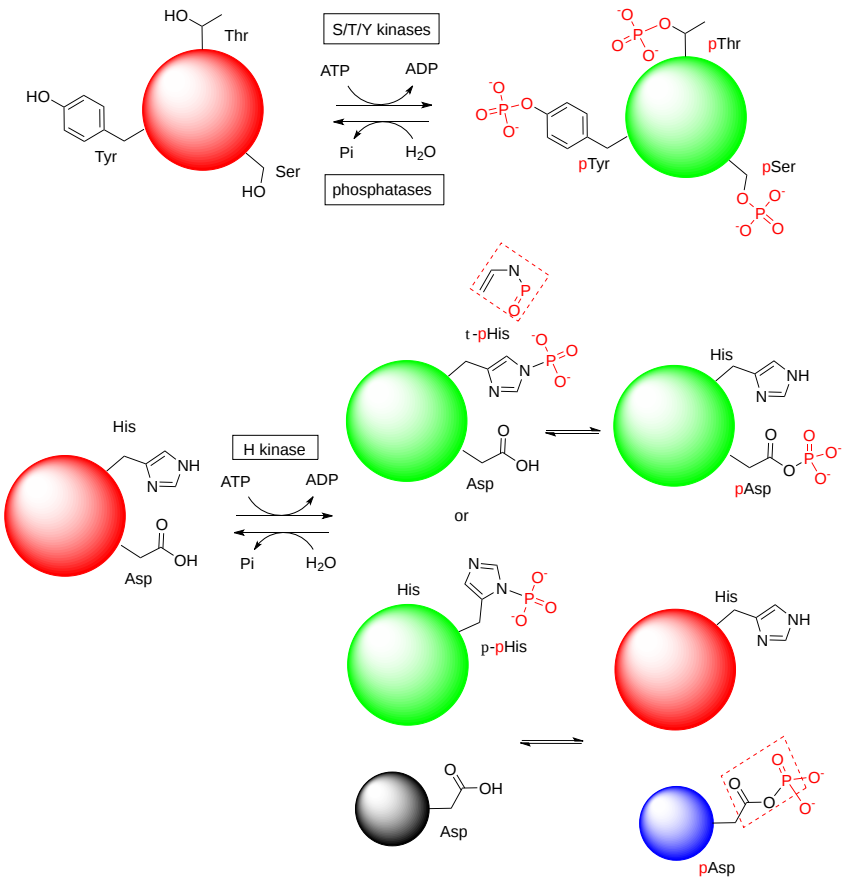


Figure 28.15.4: A comparison of Ser, Thr, Tyr, and His kinases

The more common mammalian S/T/Y kinases are shown at the top and the histidine kinase at the bottom. Note that phosphorylation of His can occur at either nitrogen to produce either τ- or π-pHis. In the two-component system, instead of water being the receiver of the phosphate from the pHis (hydrolysis), the receiver is an Asp or another His in the same His Kinase receptor or in another receiver protein (shown in gray and its phosphorylated blue form in Figure 28.15.9 You can imagine the phosphate on the original pHis jumping to a receiver, which then donates it to another receiver in the signaling process in a relay process.

Proteins like the receptor His-kinase proteins in two-component systems have multiple domains with different functions. It's really helpful to present domain structure diagrams to help in understanding the protein's structure and activities. At the same time, the domain structures determine by various bioinformatic programs vary. Nevertheless is it useful to see multiple representations of domain structure, especially if they are shown in conjunction with actual structures.

The first component of the two-component signaling system is the **receptor His-kinase** which can be viewed as a stimulus-activated kinase (much like receptor tyrosine kinases - RTKs). The second component is the **response regulator protein,** which is typically a second protein. Each of these in turn has its own domain structure. For example, the periplasmic sensing domain regulates the kinase domain of the receptor His-kinase (component one). The phosphate from the p-His in the first component is transferred to an Asp in the second component. The domain structures and phospho-transfer are illustrated in Figure \ (\PageIndex{5\) below.

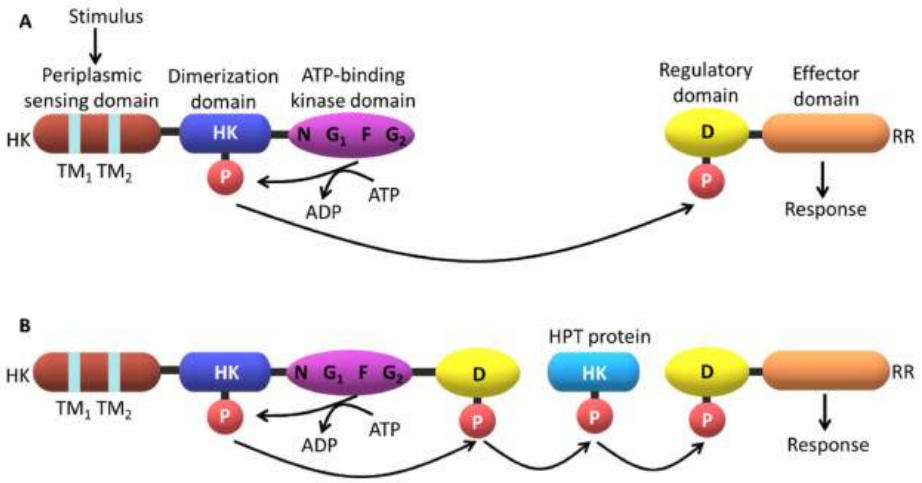


Figure 28.15.5: Domain structures and phosphotransfer mechanisms in two-component systems. Silva et al. *Int. J. Mol. Sci.* **2021**, *22*(16), 9034; **https://doi.org/10.3390/ijms22169034.** Creative Commons Attribution (CC BY) license (**https://creativecommons.org/licenses/by/4.0/**).

Panel (**A**) shows a prototypical two-component pathway. The transmembrane sensor HK (component one) and a cytoplasmic response regulator (RR) protein (component two) are shown. (Note: the actual protein is a **dimer** in the membrane.) The transmembrane segments are labeled TM_1 and TM_2. N, G_1, F, and G_2 are conserved sequence motifs in the ATP-binding domain. HK catalyzes ATP-dependent autophosphorylation of a specific conserved His residue within the HK dimerization domain. The phosphoryl group (P) is then transferred to a specific aspartate residue (D) at the conserved RR domain (component two). Phosphorylation of this domain usually triggers an associated (or downstream) effector domain, which ultimately produces a specific cellular response.

Panel(**B**) shows a multi-component phospho-relay system that often involves a variant of HK with an additional internal C-terminal RR domain. In these complex systems, at least two His–Asp phosphoryl transfer events occur, typically involving a His-containing phosphotransfer protein (HPT) operating as a His-phosphorylated intermediate.

In most prokaryotic systems, the response is directly carried by the RRD which functions as a transcription factor. Two-component systems also exist in some eukaryotes. They often interact with other downstream signaling pathways such as the MAPK system. However, in eukaryotic systems, the TCS are placed at the start of the pathways and establish an interface with more conventional signaling strategies such as mitogen-activated protein (MAP) kinase and cyclic nucleotide cascades

The domain structure shown in Figure 28.15.10 doesn't show the actual orientation of the proteins in a membrane sytem. Orientation is important since the sensing domain of the HK receptor must be in the environment of the stimuli. Stimuli can be encountered in the periplasmic (equivalent to extracellular) region, in the transmembrane region and in the cytoplasm. Variants of the HisK receptors exist that recognize stimuli in each of these locations as shown in Figure 28.15.6

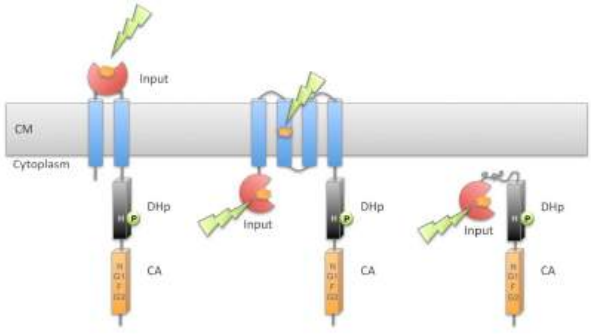


Figure 28.15.6: Hannah Schramke. https://edoc.ub.uni-muenchen.de/2103...mke_Hannah.pdf

Most histidine kinases sense extracellular signals (left-hand structure). All have their cytoplasmic transmitter domains which contain the pHis. As mentioned previously, the histidine kinases are dimeric, so when activate they phosphorylate a His on the other monomer (transphosphorylation). In addition to the H box with the reactive histidine, they also have N, G1, F, and G2 boxes. The H box is also involved in dimerization The transmitter domain can be further divided into two parts: the H-box is involved in

dimerization and obviously in phosphotransfer. The figure also shows the CA domain (HK-type ATPase catalytic or HATPase_c), also known as the catalytic and ATP-binding (CA) domain (Figure 1.3-3)

The histidine kinase senses a variety of stimuli in its sensory domain. The stimuli can be generally grouped into organic (e.g. dicarboxylates, citrate, etc), ions (e.g. Mg^{2+}, H^+, K^+), gaseous ligands (e.g. O_2, N_2), and physical changes (osmolarity/turgor, light, and temperature). Stimuli are "sensed" by a variety of different characterized folds. Some common sensing domains are PAS (Per-ARNT-Sim), CHASE (cyclase/histidine kinase-associated sensing extracellular), four-helix bundle (4HB), and NIT (nitrate and nitrite-sensing) classes. We will focus on one particular His kinase system, histidine kinase KdpD.

28.15.3.1: Histidine kinase KdpD and the regulation of K^+ ion concentration

It is often difficult to identify the primary stimulus for a receptor, as exemplified by the histidine kinase KdpD which, together with the response regulator KdpE, controls the expression of a high-affinity K^+-uptake system in many bacteria. K^+ is the most abundant cation in all living cells, especially in bacteria it is crucial for the regulation of cell turgor and intracellular pH and the activation of several enzymes. To ensure a sufficient supply of K^+, most bacteria have more than one K^+-uptake system. For example, *E. coli* has at least three such systems, the constitutively expressed systems Trk and Kup, and the inducible high-affinity K^+-uptake system KdpFABC. The genes *kdpF*, *kdpA*, *kdpB*, and *kdpC* form an operon that codes for four inner membrane proteins. The *kdp* operon is induced when *E. coli* is grown under K^+ limitation, lacks the major K^+ transporter Trk or has an increased need for K^+ when under hyperosmotic stress. Under all these conditions, the membrane-integrated histidine kinase KdpD autophosphorylates and transfers the phosphoryl group to the cytoplasmic transcriptional (response) regulator KdpE, resulting in the induction of the *kdp* operon, as shown in Figure 28.15.7.

What is amazing is that KdpD has not only kinase activity but also phosphatase activity towards phosphorylated KdpE, which switches the signaling cascade off. It is a bifunctional enzyme/receptor. A single substitution (T677A) in the C-terminal domain results in no phosphatase activity.

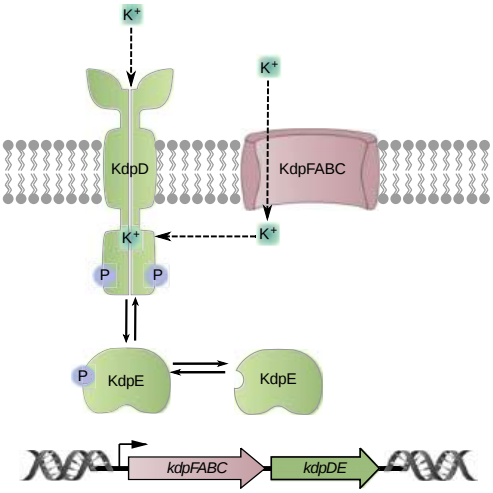


Figure 28.15.7: Schematic of the KdpD regulation system.

Hence KdpD can be thought of as a bifunctional receptor acting as both kinase and phosphatase to regulate gene expression. The bifunctional receptor histidine kinase KdpD acts as both an autokinase (including phosphotransferase) and phosphatase for the response regulator KdpE. Phosphorylated KdpE activates the expression of the genes encoding the high-affinity K^+ transporter KdpFABC. KdpD autokinase activity depends on the external K^+ concentration, and the phosphatase activity is influenced by the internal K^+ concentration. K^+ ions don't move through a channel in KdpD but through the KdpFABC.

The cartoon in Figure 28.15.7, as with all cartoons, can be misleading with respect to scale. Figure 28.15.8 shows an interactive iCn3D model of the actual membrane domain of E. coli histidine kinase receptor KdpD (2KSF)

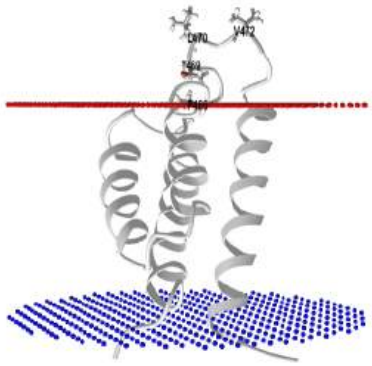
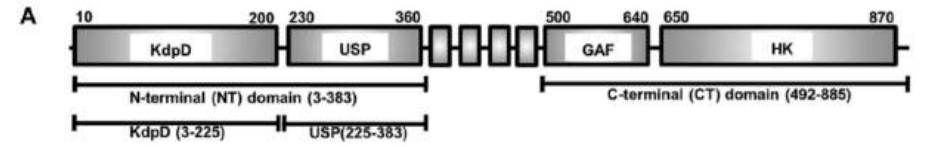


Figure 28.15.8: Membrane domain of E. coli histidine kinase receptor KdpD (2KSF)Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...yMgkv2fxn9RKo8

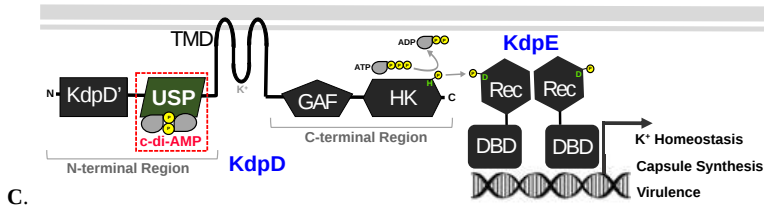
4 helices space the membrane. Hence both the N-terminal and C-terminal domains are actually in the cytoplasm, not one in the periplasmic space and one in the cytoplasm as you would infer from Figure 28.15.7. The actual periplasmic (outsithe de of cell" domain) consists of only about 6 amino acids. Hence it most clearly is represented by the middle model in Figure 28.15.6 Mutations of key residues in the periplasmic loop region (P466A, T469A,, L470A and V472A) drastically affect K^+ recognition. Actually how it "senses" periplasmic K^+ ions is not clear.

Domain representation

We present three different domain diagrams for KdpD in Figure 28.15.9 not to confuse readers, but to show the utility of mulrepresentationstation they are likely to encounter in reading the literature.



Moscoso et al. Journal of Bacteriology. 198 (2016) http://dx.doi.org/10.1128 /JB.00480-15. Creative Commons Attribution 3.0 Unported license



Dutta et al. JBC. 296, 100771 (2021). DOI:https://doi.org/10.1016/j.jbc.2021.100771. Creative Commons Attribution (CC BY 4.0)

Figure 28.15.9 Multiple representations of the domain structure of the histidine kinase KdpD

In panel A, the 4 boxes btw 360 and 500 are 4 transmembrane helices, which would not be evident in a simple cartoon as in Figure 28.15.7. In panel B, the domains are shown as follows:

- Green: K+ channel His kinase sensor domain 21-230;
- Red: 4 transmembrane helices 402-508; the helices are lumped together in the red domain representation;
- Blue: GAF domain 527-644;
- Yellow and Purple combo: HK domain;
  - Yellow: HisKinase A phosphoacceptor domain 663-730 which contains the His acceptor, in effect the substrate of the kinase domain;
  - Purple: His Kinase/HSP 90 is like ATPase 773-883.

The domain structure in panel C is the most detailed and also has the domain structure of the receiver (response regulator). Again the periplasmic domain which senses K^+ consists of only a few amino acids, 424-427 and 467-474. Neither representations A nor B show that the major N-terminal and C-terminal halves of the protein are in the cytoplasm. Panel C shows more information about the domains, their function, and their orientation in the intracellular environment. It turns out there is also a sensor for intracellular K^+ions, which is depicted in Figure 28.15.7.

A central question is how histidine kinase KdpD responds to changes in K^+ concentrations. Both the kinase and phosphatase activities are regulated by K^+.

- When periplasmic (extracellular) K^+ is > 5 mM (high), the ion appears to bind to the small extracellular loops (see Figure 28.15.8), which inhibits the autokinase activity. Under the same conditions the intracellular C-terminal tail senses K^+ and activates the phosphatase activity, which cleaves its pHis. These combined effects inhibit high-affinity K^+ transport.
- When periplasmic K^+ becomes low, kinase activity is activated and the protein is autophosphorylated, ultimately leading to the activation of the gene for the high-affinity K^+ transporter KdpFABC. As long as intracellular K^+ levels are high, the phosphatase is active. When intracellular K^+ levels drop sufficiently, the phosphatase becomes inhibited, which further simulations the transcription of both high-affinity K^+ transporter KdpFABC.

Hence the histidine kinase KdpD system is regulated by **both** periplasmic and cytoplasmic K^+ ions.

Yet another signal regulates the KdpD His Kinase receptor two-component signal. What has been conspicuously absent from this discussion about signaling in bacteria is the involvement of second messengers like cAMP (which activates Protein Kinase A and some membrane proteins). There does appear to be one major second messenger in bacteria - cytoplasmic di-AMP (c-di-AMP), whose structure is shown in Figure 28.15.10

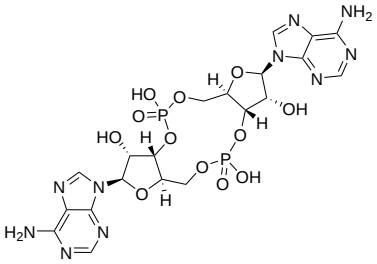


Figure 28.15.10: The bacterial second messenger cytoplasmic di-AMP (c-di-AMP)

It binds in the N-terminal region of KdpD His Kinase receptor protein "sensor" domain region to a specific domain called the Universal stress protein (USP) domain as shown in panels A and C of Figure 28.15.9 Figure 28.15.11shows an interactive iCn3D model of the Staphylococcus aureus universal stress protein (USP) domain of KdpD histidine kinase in complex with second messenger cyclic diadenosine phosphate (c-di-AMP) (**7JI4)**

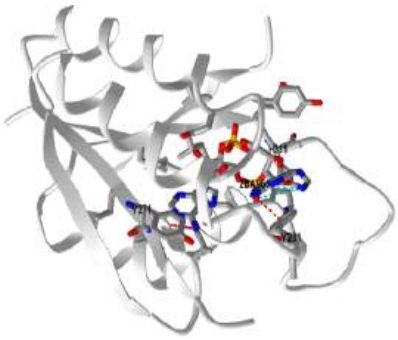


Figure 28.15.11: Staphylococcus aureus universal stress protein (USP) domain of KdpD histidine kinase in complex with second messenger c-di-AMP (**7JI4).** (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...ij5DqgazLFvV5A

"Dual sensing thus emerges as a highly optimized regulation strategy. The key advantage of this strategy is that it confers on cells the ability to directly sense changes in both the supply of and demand for the limiting resource. It is, in fact, analogous to strategies that are widely used in control engineering, e.g. modern heating systems work with both exterior and interior thermometers to ensure constant room temperature."

28.15.3.2: *Escherichia coli* nitrate/nitrite sensor kinase NarQ

Let's examine another TCS protein, the *Escherichia coli* nitrate/nitrite sensor kinase NarQ, to see how the binding of a ligand to the periplasmic domain might transmit a signal so far into the cell through the plasma membrane. We won't discuss the His Kinases that lack transmembrane regions (about 1/4). The sensor domain hence is mostly in the periplasm, followed by the transmembrane domain (see Figure 28.15.8), followed often by a cytoplasmic HAMP domain, with a four-helical parallel coiled-coil. The HAMP domain transmits the signal to downstream signaling domains like Dhp in the protein.

Nitrate/nitrite is sensed by two different two-component systems, sensing systems NarX-NarL and NarQ-NarP, which regulate anaerobic respiration. NarQ phosphorylates two different proteins, NarL and NarP in the presence of nitrate or nitrite and dephosphorylates both proteins in the absence of ligands. Both NarQ (and NarX) have seven domains: a four-helical periplasmic sensor domain, TM bundle, HAMP domain, so-called signaling helix, (S-helix), GAF-like domain, dimerization and histidine phosphotransfer domain, and, finally, catalytic kinase domain, as shown in Figure 28.15.12

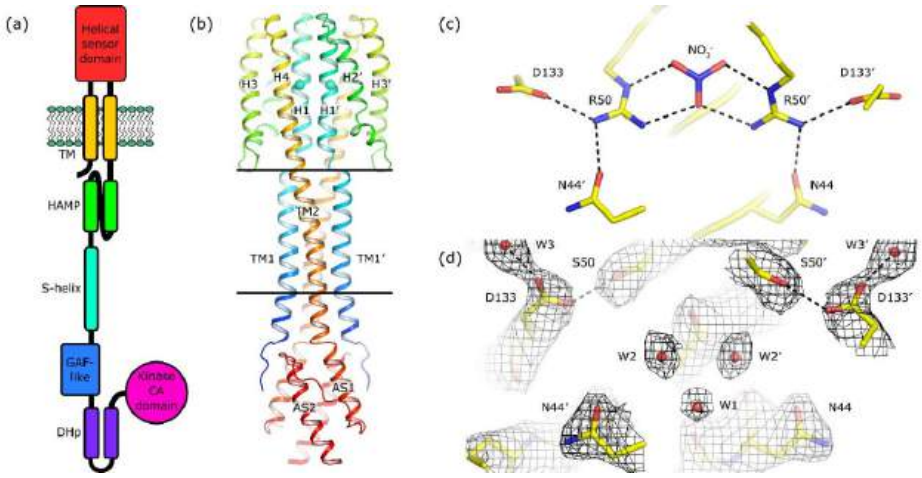


Figure 28.15.12: Architecture of NarQ and structure of the R50S mutant. Elements belonging to the second protomer in the dimer are marked with primes. Gushchin et al. Int. J. Mol. Sci. 2020, 21, 3110; doi:10.3390/ijms21093110. Creative Commons Attribution. (CC BY) license. (http://creativecommons.org/licenses/by/4.0/).

Pane (**a**) shows the architecture of NarQ. Note that the functional protein is homodimeric. Approximate domain boundaries, according to InterPro [17], are TM1, residues 14–34; sensor, 39–146; TM2, 147–167; HAMP, 172–227; S-helix, 228–246; GAF-like, 247–360; DHp, 361–425; CA, 424–560.

Panel (**b**) shows the overall structure of the sensor-TM-HAMP fragment of the R50S mutant (which allowed crystallization). The position of Ser50 is highlighted with spheres. The backbone structure is identical to that of the WT protein.

Panel (**c**) shows the structure of the ligand (nitrate) -binding site in the WT protein.

Panel (**d**) shows the structure of the ligand-binding pocket in the R50S mutant. Asp133 is reoriented towards Ser50. $2F_o–F_c$ electron density maps are contoured at the level of 1.2 × r.m.s. Putative water molecules are shown as red spheres. Gushchin et al. Int. J. Mol. Sci. 2020, 21, 3110; doi:10.3390/ijms21093110. Creative Commons Attribution. (CC BY) license. (http://creativecommons.org/licenses/by/4.0/).

Although the sequence identity between NarQ and NarX is ~28%, their ligand binding sites—membrane-proximal parts of the sensor domain's helices H1 called P boxes—are very well conserved: 14 out of 15 amino acids (residues 42–56 in NarQ) are identical, and the differing ones, Ile 45 in NarQ and Lys49 in NarX, are responsible for the differentiation between nitrate and nitrite.

It appears that the ligand-induced conformational change in the ligand-binding site is helical rotation, which results in diagonal scissoring of the sensor domain helices, leading to the change in the secondary structure of the sensor-TM linker and, eventually, piston-like shifts of the transmembrane α-helices.

Nitrate causes changes in the transmembrane region when the apo (nitrate free) and holo (nitrate bound) state structures of NarQ are compared. On binding of nitrate, the induced conformation changes in NarQ have been described as a "combination of changes in the lateral arrangement of the TM helices and piston-like shifts of the helices in the direction perpendicular to the membrane plane." This results in either symmetric or asymmetric changes and scissoring of the transmembrane helices (based on two different crystal structures of the holo-form. A "piston-like" movement of the helices is observed on both holo-forms.

Results show that the binding of ligand to NarQ causes a piston-like displacement of the TM helices, which is accompanied by extensive symmetric or asymmetric rearrangements and scissoring of the TM helices. The rearrangements are different in the two presented holo-state structures, but the piston-like displacement is perfectly conserved. Thus, the latter appears to be a more robust mechanism of TM signal transduction.

Figure 28.15.13 shows an interactive iCn3D model of a fragment of nitrate/nitrite sensor histidine kinase NarQ (mutant R50K) in the symmetric holo state (5IJI)

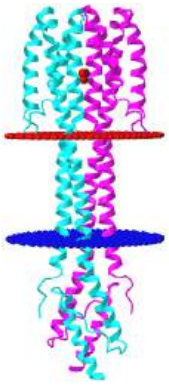


Figure 28.15.13: Fragment of the nitrate/nitrite sensor histidine kinase NarQ (mutant R50K) in the symmetric holo state (5IJI) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...gsRXjmKBbqbis8

A homodimer (cyan and magenta) is shown with the N-terminus and C-terminus on the cytoplasmic side. Each monomer passes through the membrane with an alpha-helical domain twice. Nitrate is shown in spacefill in the extracellular domain outside of the outer leaflet (red spheres) of the membrane.

Figure 28.15.14 shows the conformational transition going from the symmetric apo state (5JEQ) magenta to the symmetric holo state (IJI) (cyan) with bound nitrate (not shown).

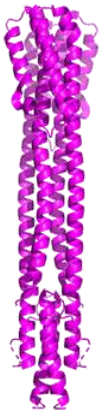


Figure 28.15.14: Conformational transition in the nitrate/nitrite sensor histidine kinase fragment going from the symmetric apo state (5JEQ) magenta to the symmetric holo state (IJI) (cyan)

Conformational changes in the HAMP domain seem to amplify and convert the piston-like conformational changes in the transmembrane domain. Note the splaying out to the helices at the bottom (cytoplasmic end) in the apo form.

28.15.3.3: Phototaxic Photoreceptors

We have discussed two-component systems that have a His-Kinase receptor transmitter protein. There are two other major types of bacterial receptors, chemoreceptors (involved in chemotaxis) and photoreceptors, involved in phototaxis. These often have similar modular domain structures. Chemotaxis receptors are, like the His Kinase receptor, dimers with extracellular domains that bind the chemotactic signal. The photoreceptors appear to be active as a trimer of dimers.

The basic dimeric structure contains the microbial light sensor rhodopsin which contains the chromophore opsin, and its transducer Htr. In the halobacteria *N. pharaonis*, *(*archaeal, not a bacterial cell) the proteins are sensory rhodopsin II (NpSRII) with its transducer (NpHtrII), mediates negative phototaxis in halobacteria *N. pharaonis*.

Microbial **rhodopsins** are phototransducing proteins with a conjugated chromophore retinal, covalently attached to the protein opsin through a Schiff base (imine) linkage. The holoprotein (opsin with the attached retinal) is called rhodopsin. Retinal is derived from beta-carotene. The structures of animal and microbial retinals are shown in Figure 28.15.15

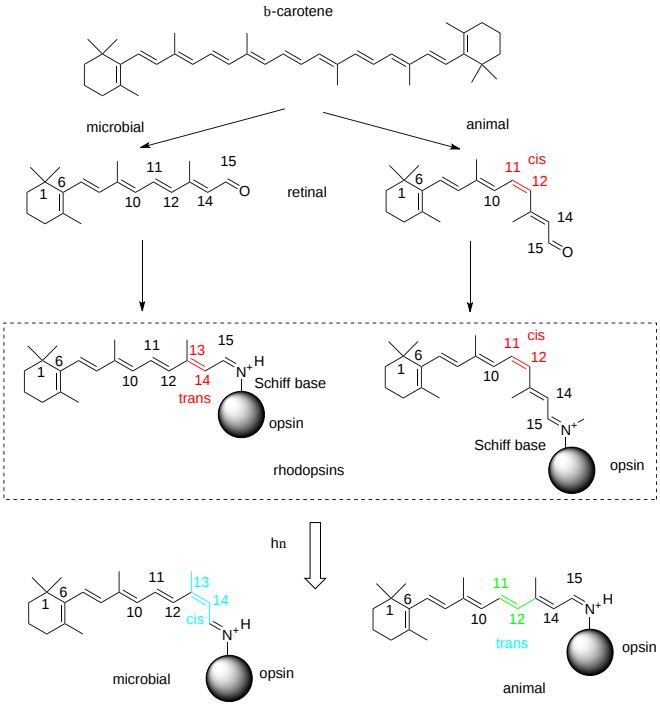


Figure 28.15.15: Structures of animal and microbial retinals

When light of the correct wavelength is absorbed, an electron in a pi molecular orbital in retinal is promoted to a pi antibonding molecular orbital, breaking a 2 electron pi bond in the structure at a certain site in the isoprenoid chain, allowing rotation around the now single bond. The final result after the electrons return to the ground state is photoisomerization of the trans 13-14 and cis 11-12 bonds in microbial and animal retinal, respectively, to their respective cis 13-14 and trans 11-12 configuration. This conformational change in the bound retinal induces a conformational change in the protein opsin, leading to signaling.

Figure 28.15.16 shows an interactive iCn3D model of one monomeric of bacteria rhodopsin (1C3W) containing retinal attached through a Schiff base to Lys 216.

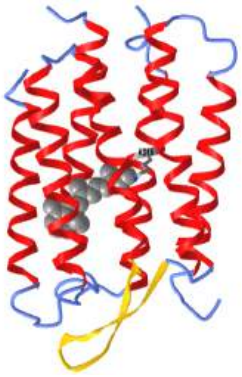


Figure 28.15.16: Bacterial rhodopsin (1C3W). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...9HzvfdFvb9kcc6

The covalently attached retinal is shown in gray spacefill. The protein opsin is a membrane protein that spans it with seven helices. Hence it is very similar to a GPCR.

In the phototaxic receptor, when light is absorbed by rhodopsin (NpSRII), the resulting conformational change in the protein causes conformational changes in the transducer protein NpHtrII associated with it, leading to signaling through a two-component signal system. (The chemotaxis response to chemical signals occurs through a similar but ligand-induced process.) Figure 28.15.17 shows the structure of the Archaeal photoreceptor complex.

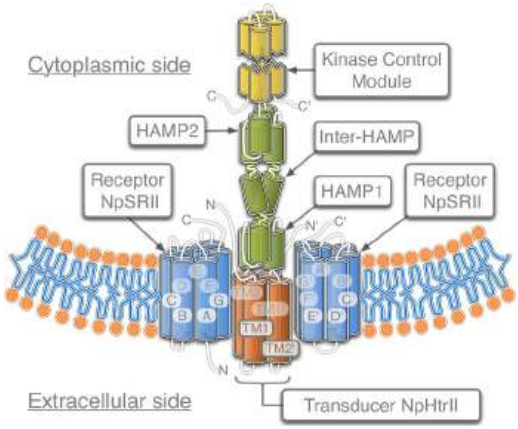


Figure 28.15.16: The dimer of two NpHtrII proteins is flanked by two NpSRII proteins.

A-G, TM1, and TM2 are the transmembrane helices. The cytoplasmic part of NpHtrII consists of two HAMP domains (HAMP1 and HAMP2) connected by an α-helical linker (Inter-HAMP) and the kinase control module. Primes denote symmetry mates of the complex. Ishchenko, A., Round, E., Borshchevskiy, V. *et al.* New Insights on Signal Propagation by Sensory Rhodopsin II/Transducer Complex. *Sci Rep* **7,** 41811 (2017). https://doi.org/10.1038/srep41811. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/

Figure 28.15.17 shows a more details representation of both the transmitter and receiver in the Archeal two-component phototransduction system and how it leads to phototaxis.

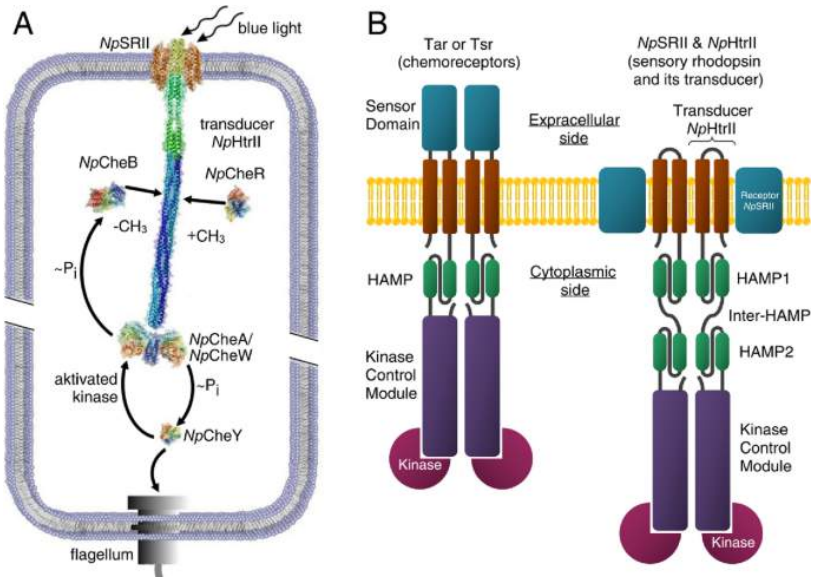


Figure 28.15.17: Signal transduction pathway in case of the two-component phototaxis system of *Natronomonas pharaonis*[5] and domain architecture of membrane chemo- and photoreceptors of TCS. Ryzhykau, Y.L., Orekhov, P.S., Rulev, M.I. *et al.* Molecular model of a sensor of the two-component signaling system. *Sci Rep* **11,** 10774 (2021). https://doi.org/10.1038/s41598-021-89613-6. Creative Commons Attribution 4.0 International License, http://creativecommons.org/licenses/by/4.0/.

Pane (**A**) shows Light activated sensory rhodopsin II (*Np*SRII) induces conformational and/or dynamical changes in the transducer (*Np*HtrII), which are converted by two HAMP domains and conveyed along the 200 Å long transducer to the tip region. Activated by the transducer histidine kinase CheA (bound to the adapter protein CheW) undergoes auto-phosphorylation and further transfers the phosphate group to the response regulators CheY or CheB. CheY affects the rotational bias of the flagellar motor, while the methylesterase CheB along with the methyltransferase CheR controls the adaptation mechanism.

Panel (**B**) shows cartoon representations of the chemoreceptor dimer (Tar and Tsr in complex with kinases) from *E. coli* and of the photosensor dimer of the complex of the sensory rhodopsin II with its cognate transducer *Np*HtrII and kinases from *N. pharaonis*.

28.16: Signaling in Plants

28.16.1: Introduction

Plants are obviously comprised of cells. Hence they must engage in cell signaling within and between cells. It is beyond the scope of this book to give a detailed description of cell signaling in plants. Instead, we focus on 5 key classic plant hormones, **auxins, cytokinins, ethylene, gibberellins, and abscisic acid**, which are produced by leaves, flowers, shoots, roots, or fruit, and see how they initiate signaling in plants. Finally, we would be remiss if we didn't include the profound signaling in plants initiated by light. Most of this section comes directly from a series of sources, with modifications and additions (mostly molecule models).

Auxins (3-indolebutyric acids derivatives) are regulators of growth and development and are found in actively growing parts of the plant (root, shoot, leaves) but mostly in the cell stem. Auxins facilitate the bending of plants toward the light, for example. They work in conjunction with other hormones like **cytokinins**. When auxins are higher than cytokinins, roots will form, while the opposite produces shoots. Auxins facilitate the elongation of cells, while cytokinins promote cell division and growth as well as wound repair. **Gibberellins** also are plant growth regulators and facilitate cell elongation. They also help in germination, elongation of the stem, fruit ripening, and flowering. **Abscisic acid** affects **seed development and maturation** and helps plants tolerate environmental or biotic stresses. It also inhibits growth and metabolism. **Ethylene** affects fruit ripening, organ abscission, and growth by restricting cell elongation.

We will focus on the hormones, their protein receptors, and how the hormone:receptor complex initiates some key events in the cell.

28.16.2: Auxin

Auxins, the first plant hormones discovered, regulate plant growth and development. The most common auxin is 3-indole acetic acid. Figure 28.16.1shows the structures of naturally occurring auxins.

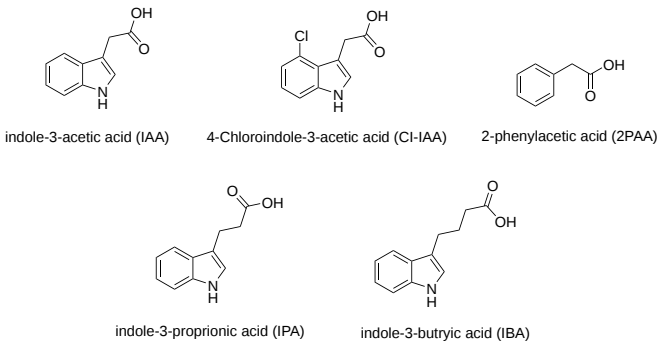


Figure 28.16.1: Structures of naturally occurring auxins.

Figure 28.16.2shows an interactive iCn3D model of auxin bound to its receptor, TIR1 ubiquitin ligase (2P1Q)

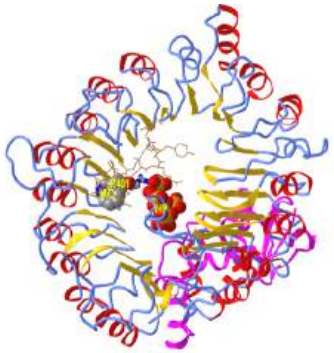


Figure 28.16.2 Auxin bound to its receptor TIR1 ubiquitin ligase (2P1Q) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...6xSkEZ3aJA4gT6

Auxin (IAA) is shown in spacefill CPK colors along with an unexpected binding cofactor, inositol hexakisphosphate (IHP), shown in spacefill CPK colors. The peptide shown in light brown sticks is part of the protein Auxin-responsive protein IAA7, a member of a class of proteins called **AUX/IAAs**. These are short-lived transcriptional factors that function as repressors of early auxin response genes at low auxin concentrations.

The magenta subunit, TIR1 (transport inhibitor response 1), is part of the larger TIR1 complex, the SCF(TIR1) E3 ubiquitin ligase, of which only TIR1 is shown. Its mere name suggests that it is involved in the ubiquitinylation of a key protein involved in auxin activity, which will be targeted for proteolysis. That protein is the repressor protein IAA7 (an AUX/IAA protein).

Auxin binds in a hydrophobic pocket, which accounts for the binding of the other largely hydrophobic auxins shown in Figure 28.16.1 Note however that Arg 401 that forms a salt bridge (ion-ion interaction) with the carboxylate of 3-indole acetic acid

Figure 28.16.3shows an interactive iCn3D model of auxin bound in the hydrophobic pocket of its receptor, TIR1 ubiquitin ligase (2P1Q).

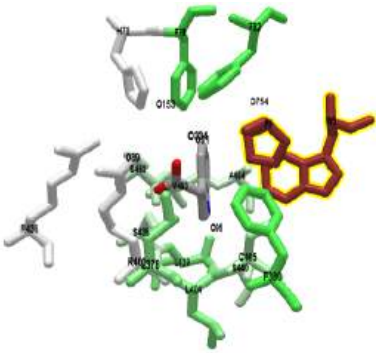


Figure 28.16.3: Auxin bound in the hydrophobic pocket of its receptor TIR1 ubiquitin ligase (2P1Q) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...TFkf6HdMYMiZT9

The green color represents nonpolar side chains. The brown side chains, Trp and Pro, are from the auxin-responsive protein IAA7 peptide. To reiterate, the protein IAA7 is a member of a class of proteins called AUX/IAAs which repress auxin activity. The IAA7 peptide packs over the auxin in the binding pocket.

Now we can see how auxin function to regulate gene transcription. First, we must introduce another protein family, the **auxin response factors (ARFs)**. These are transcription factors that bind to a key DNA sequence, the auxin response element (AuxRE) in promoter sequences of auxin-activated genes. Once bound they can either activate or repress transcription from target genes. Auxin binds its receptor TIR1 enabling the binding of an AUX/IAAs (like IAA7) repressor and the binding of the complex to ARF. The TIR1 ubiquitin ligase activity of the complex ubiquitinylates the bound AUX/IAAs (like IAA7) repressor, targeting it for degradation, freeing the ARF to become active in the regulation of gene transcription in the nucleus.

ARFs are structurally similar, with most members containing three regions: DBD (DNA-binding domain), MR (middle region), and PB1 (Phox and Bem 1). Figure 28.16.4shows a model of how auxin affects ARF transcriptional activity.

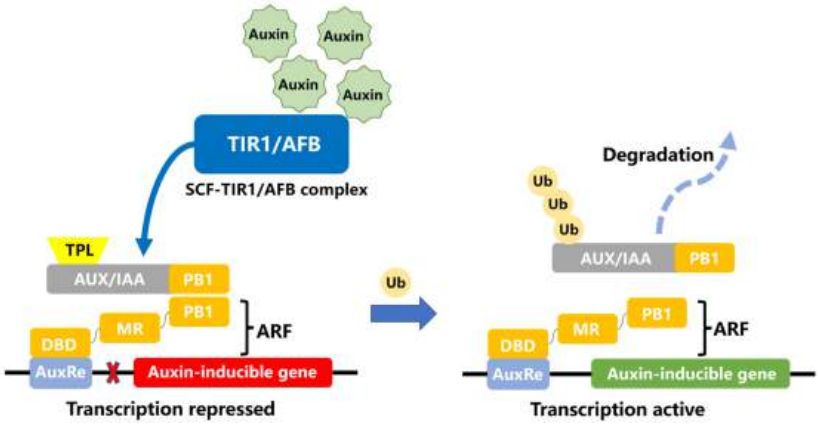


Figure 28.16.4: Model of ARF involvement in auxin response. DBD, DNA-binding domain; MR, middle region; PB1, Phox, and Bem 1. *Kou et al. Appl. Sci.* **2022**, *12*(3), 1360; **https://doi.org/10.3390/app12031360.** Creative Commons Attribution (CC BY) license (**https://creativecommons.org/licenses/by/4.0/**).

At low concentrations of auxin, the AUX/IAA repressor binds to the ARF transcription factor through their PB1 domains. The PB1 (Phox and Bem1) domain is about 80 amino acids in length. It acts as a protein binding module allowing heterodimerization or homo-oligomerization with proteins have also contained the PB1 domain.

The dimer of AUX/IAA and ARF recruits the co-repressor TPL (TOPLESS) to inhibit the ARF activity and the expression of auxin-responsive genes. When the concentration of auxin is increased, Aux/IAA binds to the SCF TIR1/AFB complex and is ubiquitinated and then degraded by 26S protease. The ARF transcription factors are released to activate the transcription of downstream genes. DBD, DNA-binding domain; MR, middle region; PB1, Phox, and Bem 1.

- At low concentrations of auxin, the AUX/IAA repressor binds to the ARF transcription factor and forms a dimer that recruits the co-repressor TPL (TOPLESS) to inhibit the ARF activity and the expression of auxin-responsive genes;
- When the concentration of auxin is increased, Aux/IAA binds to the auxin:SCF TIR1/AFB complex (remember that the auxin receptor is the TIR1 component of the complex) and is ubiquitinated by the TIR1, which is also a ubiquitin ligase;
- The ubiquitinated AUX/IAA protein is degraded by proteolysis by the 26S protease, allowing the ARF to become an active transcription factors

It appears that the MR domain of ARF determines whether it activates or inhibits transcription. If it is rich in proline, serine, and threonine, it acts as an inhibitor. If it is enriched in glutamine and leucine it acts as an activator. Some reports show that Aux/IAA and ARFs can form not only dimers but also larger complexes (oligomers), noting that oligomerization of Aux/IAA proteins may be essential for the inhibition of ARF proteins and only sufficient amounts of Aux/IAA proteins can exert the inhibitory effect of ARF proteins.

Figure 28.16.5 shows an interactive iCn3D model of the DNA binding domain of arabidopsis thaliana auxin response factor 1 (ARF1) in complex with auxin response element-like sequence ER7 (4LDX)

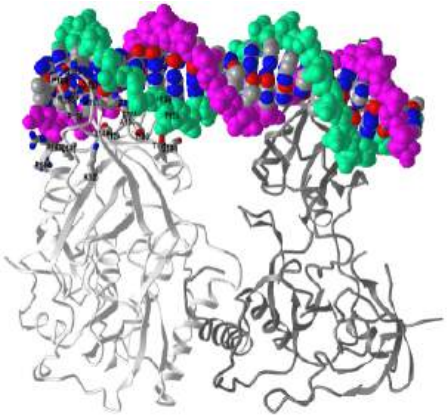


Figure 28.16.5: DNA binding domain of arabidopsis thaliana auxin response factor 1 (ARF1) in complex with promoter-like sequence ER7 (4LDX) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...9fNHEiFaVANAk6

The ARF protein must translocate to the nucleus to regulate gene transcription. ARF7 and ARF19 have been shown to form micron-sized aggregates in the cytoplasm. These have low responses to auxin. Aggregation occurs through PB1 domain interactions between ARFs as well as through intrinsically disordered regions. Mutation of a single lysine in the PB1 prevents aggregation and leads to morphological changes in the plant. This shows the importance of regulating not only transcription but also the translocation of proteins to the nucleus.

Figure 28.16.6reviews the activation of ARFs and some of the genes affected by ARF.

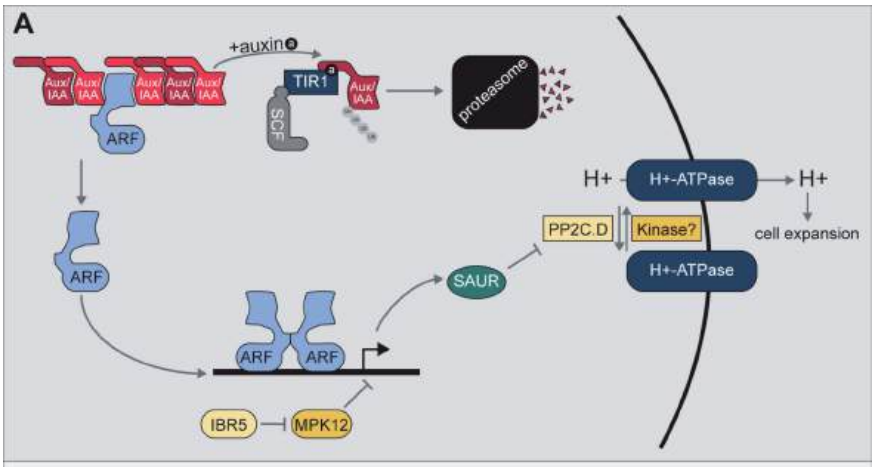


Figure 28.16.6: Auxin signal transduction pathways. Powers and Strader. Up in the air: Untethered Factors of Auxin Response (doi:10.12688/f1000research.7492.1). License CC BY 4.0

Auxin promotes the formation of the TIR1/AFB Auxin/Indole-3-acetic acid inducible (Aux/IAA) co-receptor to promote the ubiquitylation and subsequent degradation of the Aux/IAA repressor. Aux/IAA degradation relieves repression of auxin response factor (ARF) transcription factors, allowing for auxin-responsive gene expression. One of the transcript families upregulated by auxin is the SAUR family. The **s**mall **a**uxin **u**p **R**NA (SAUR) proteins encoded by these transcripts have been suggested to play roles in multiple processes, one of which is interaction with and inhibition of members of the PP2C.D family of phosphatases, which act to regulate H^+-ATPase activity. Further, indole-3-butyric acid response 5 (IBR5) and mitogen-activated protein kinase 12 (MPK12) have been implicated in regulating auxin-responsive gene transcription; this regulation is not through destabilization of the Aux/IAA repressors, suggesting a yet-to-be-discovered mechanism of regulating auxin-responsive gene expression. F

28.16.3: Cytokinins (CKs) and Ethylene (ET)

Much of this material derives from Bidon et al. Cells 2020, 9, 2526; doi:10.3390/cells9112526 . Creative Commons Attribution (CC BY) license. (http://creativecommons.org/licenses/by/4.0/).

Cytokinins (CKs) and ethylene (ET) are among the most ancient organic chemicals on Earth. The structure of a representative cytokinin (kinetin) and ethylene are shown in Figure 28.16.7.

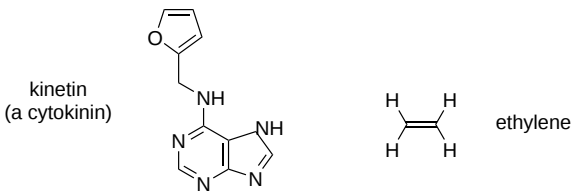


Figure 28.16.7: Structure of kinetin (a cytokinin) and ethylene

A wide range of organisms including plants, algae, fungi, amoebae, and bacteria use these substances as signaling molecules to regulate cellular processes. Because of their ancestral origin and ubiquitous occurrence, CKs and ET are also considered to be ideal molecules for inter-kingdom communication. Their signal transduction pathways were first determined in plants and are related to the two-component systems of bacteria (which we explored in a previous section), using histidine kinases as primary sensors.

CKs share a common structure of N^6-substituted adenine (see Figure 28.16.7), with biological activities defined by the N^6-substituents (isoprenoids or aromatic groups). They were originally described as the major hormones regulating cell division but are also implicated in the control of morphogenesis and embryogenesis and inhibition of senescence. Conversely, ET is a simple gas, often referred to as the senescence hormone in plants, acting to stimulate the senescence of leaves and petals as well as the ripening of fruits. Both CK and ET are also well known to orchestrate plant responses to many types of biotic and abiotic stresses.

Signaling pathways in plants are related to the two-component systems typically described in prokaryotes. CKs and ET are perceived by two types of membrane-bound **histidine kinase receptors, CRE1** and **ETR1** as shown in Figure 28.16.8

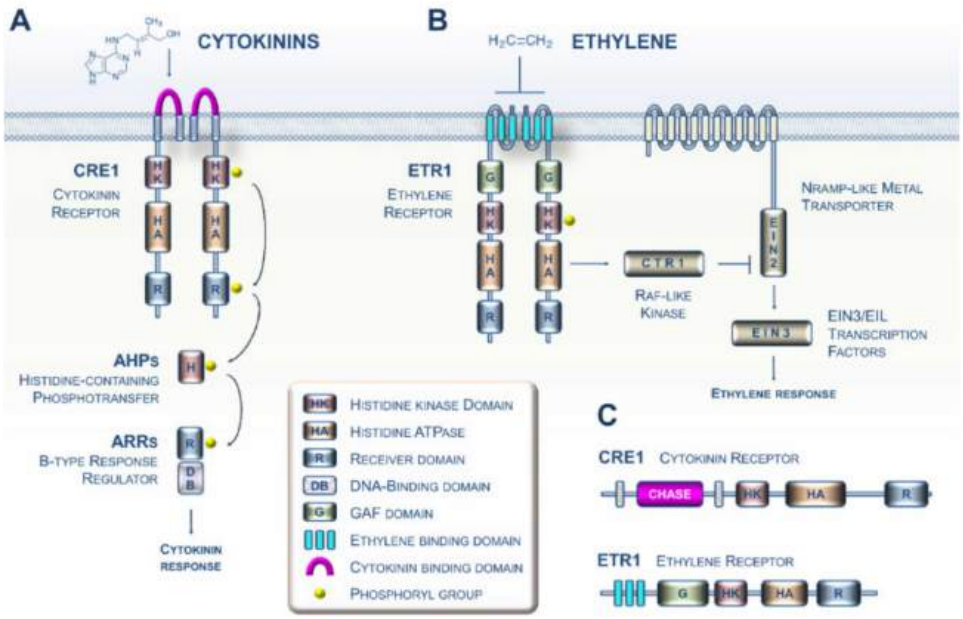


Figure 28.16.8: Perception and transduction of cytokinin (CK) and ethylene (ET) signals in the model plant *Arabidopsis*.

Panel **A** shows the cytokinin signaling pathway. CKs in *Arabidopsis* primarily are recognized by dimerized receptors such as the CRE1 receptor via the **cyclase/histidine kinase-associated sensing extracellular (CHASE) domain**. CRE1 then auto-phosphorylates (histidine kinase (HK) activity) and immediately transfers its phosphate group to the conserved histidine of a protein belonging to the histidine-containing phosphotransfer (HPt) family. This small protein then acts as a cytoplasm-to-nucleus shuttle and in turn phosphorylates a type B response regulator, which, when activated, positively regulates the transcription of response genes to the CK signal.

Panel (**B**) shows the ET signaling pathway. Ethylene molecules are detected by ethylene receptors (labeled ETR1) with ethylene binding to the three transmembrane helices (shown in sky blue). The binding of ET to the dimerized ETR1 receptor downregulates its activity. In the absence of ET, ETR1 activates the serine/threonine kinase CTR1. The CTR1 protein then phosphorylates the EIN2 protein located in the ER membrane, leading to the proteolysis of EIN2. In the presence of ET, ETR1 activity is reduced, leading to less CTR1 activity; this leads to lower phosphorylation and accumulation of EIN2 protein and subsequent activation of the EIN3 and related transcription factors. EIN3 then positively regulates the transcription of ET signal response genes.

Panel (**C**) shows the domain structure of the *Arabidopsis* ET (ETR1) and CK (CRE1) receptors.

Mechanistically, the two pathways use fundamentally different families of downstream modules.

It is now known that bacteria also use CK and ET signaling, as described in Figure 28.16.9

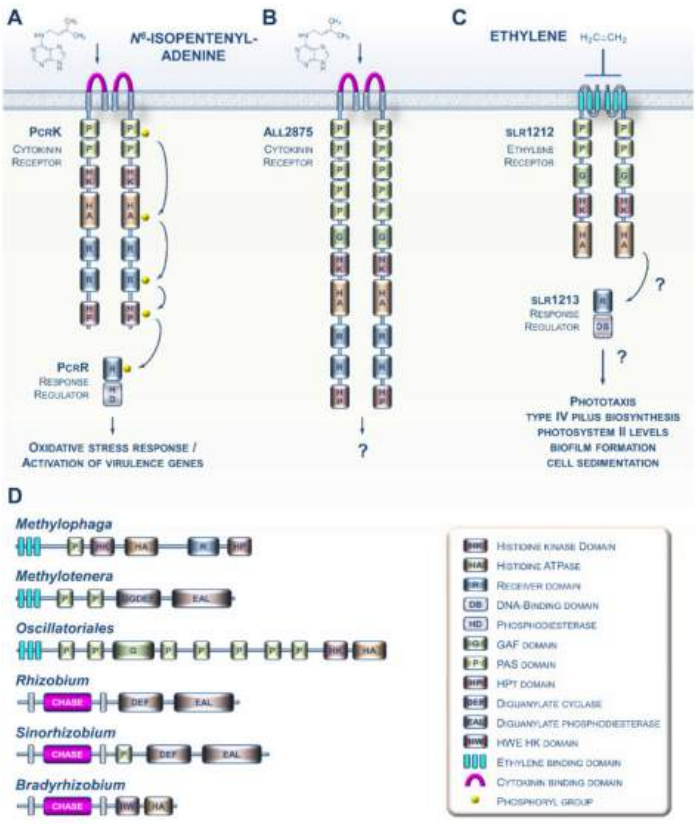


Figure 2.

Let's look in more detail at CRE1, cytokinin response 1, the main cytokinin receptor in plants. Different computational programs often show different domain structures. Figure 28.16.10shows the domain structure determined by Pfam.

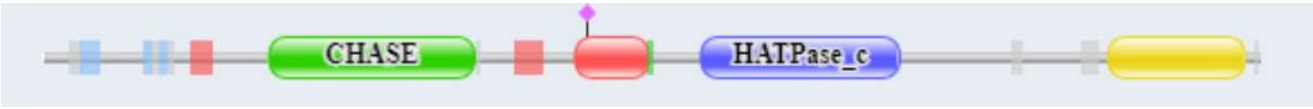


Figure 28.16.10: Domain structure of cytokinin response 1, the main cytokinin receptor in plants.

Uniprot describes this domain structure, color-coded as in Figure 28.16.11

- 131-149: transmembrane
- 200-382: Green Chase (**C**yclases/**H**istidine kinases **A**ssociated **S**ensory **E**xtracellular)
- 420-443 transmembrane
- 472-537: Red His Kinase A Phosphoaccepter domain
- 584-760 Blue HK kinase, DNAgyrase, HSP-like ATPase
- 786-920: Yellow Reg REsp 1
- 946-1071: Yellow Reg Regulator receiver domain

Figure 28.16.11shows an interactive iCn3D model of Histidine kinase 4 - Cytokinin receptor 1 (CRE) from Arabidopsis thaliana (AlphaFold model - Q9C5U0). The coloring matches the Pfam domains shown in Figure 28.16.10

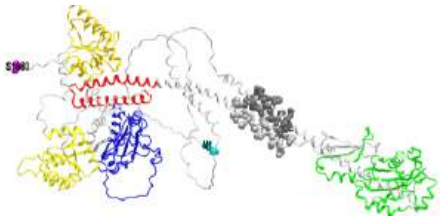


Figure 28.16.11: Histidine kinase 4 - Cytokinin receptor 1 (CRE) from Arabidopsis thaliana (AlphaFold model) - Pfam Domain organization (Q9C5U0) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...jq7BTpshpS3d46

Figure 28.16.12 shows an interactive iCn3D model of Histidine kinase 4 - Cytokinin receptor 1 (CRE) from Arabidopsis thaliana (AlphaFold model) - Domain organization (Q9C5U0) that clearly shows the extracellular and intracellular domains.

Figure 28.16.12: Histidine kinase 4 - Cytokinin receptor 1 (CRE) from Arabidopsis thaliana (AlphaFold model) - Domain organization (Q9C5U0) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...2Q9WsRkZk75ZZ8

The N-terminal methionine is in cyan spacefill and the C-terminal Ser is in spacefill. Two transmembrane helices are shown in dark gray spacefill (125-145) and light gray spacefill (430-450). These connect the extracellular domain (cyan, 146-429) and the two cytoplasmic domains (magenta 1-124, which is mostly disordered in the model, and 451-1080). This model does not reflect the relative disposition of the protein in the actual structure, but clearly shows the extracellular and cytoplasmic domains. The extracellular domain (cyan, 146-429) is the CHASE domain.

Here are the step involved in cytokinin signaling through its receptor (shown in Figure 28.16.8:

- the cytokinin binds to the CHASE domain
- the receptor autophosphorylates a His in the HK domain
- a phosphotransfer from the pHis to an Asp in the Yellow Reg (Regulator) Receiver domain
- a phosphotransfer from pAsp to the His in the histidine-containing phosphotransfer protein (HPt)
- a final transfer from pHIs to an Asp in a response regulator (RR)

The MAPK cascade is activated in the cytokinin signaling pathway. Phosphorylated pRR can also regulate target gene transcription. Type-A RRs are negative regulators of cytokinin signaling. It also acts with phytochromes (discussed at the end of this section) to regulate red light signaling. Cytokinin receptors can bind synthetical chemicals that act as defoliants and herbicides.

Intermolecular interactions in the cytokinin signaling pathway leading to transcriptional effects are illustrated in Figure 28.16.13

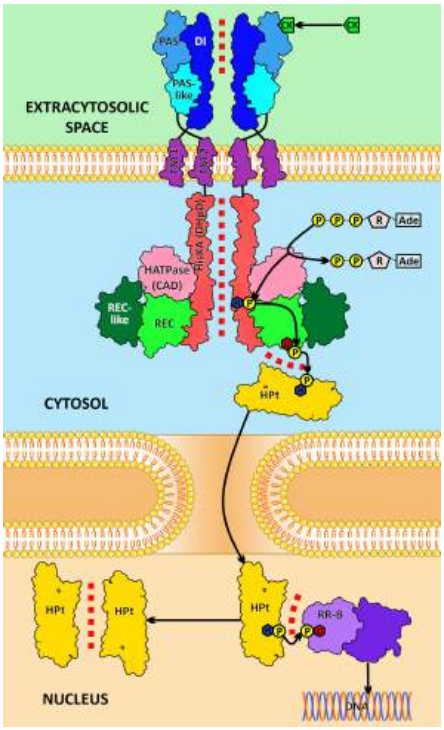
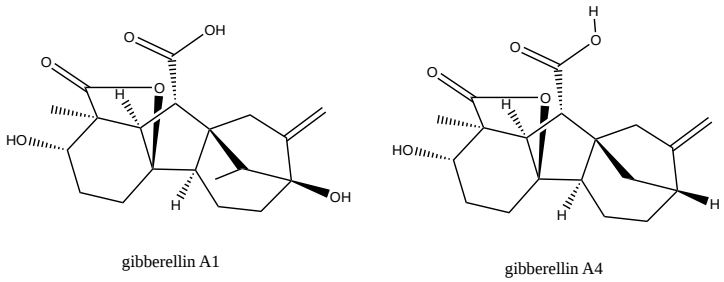


Figure 28.16.13: Intermolecular interactions in the cytokinin signaling pathway. CK, cytokinin; Ade, adenine; R, ribose; P, phosphate; D, conserved aspartate; H, conserved histidine; DI, dimerization interface domain of the sensor module; PAS and PAS-like are subdomains of the CHASE domain of the sensor module; TM1 and TM2, transmembrane domains; HisKA (DHpD), histidine kinase A domain (dimerization and histidine phosphotransfer domain); HATPase (CAD), adenosine triphosphatase domain (catalytic and ATP-binding domain); REC-like, receiver-like domain; REC, receiver domain; HPt, histidine-containing phosphotransfer protein (phosphotransmitter); RR-B, type B response regulator (transcription factor). Protein–protein interactions (PPI) are indicated by a red dotted line. *Arkhipov et al. Int. J. Mol. Sci.* **2019**, *20*(9), 2096; **https://doi.org/10.3390/ijms20092096**Creative Commons Attribution (CC BY) license (**http://creativecommons.org/licenses/by/4.0/**).

28.16.4: Gibberellin

Much of this material derives from Hedden, P., Sponsel, V. A Century of Gibberellin Research. *J Plant Growth Regul* **34,** 740–760 (2015). https://doi.org/10.1007/s00344-015-9546-1. https://doi.org/10.1007/s00344-015-9546-1. Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/)

Gibberellin controls growth and development pathways in plants and fungi. They act in plants by removing growth limitation by promoting the degradation of the growth-inhibiting DELLA proteins which contain the Asp-Glu-Leu-Leu-Ala (DELLA) motif. The name gibberellin derives from the fungus *Gibberella fujikuroi*. There are many types of gibberellins, which are all diterpenoids. The structures of the main bioactive GAs in plants, GA_1 and GA_4, are shown below in Figure 28.16.14



gibberellin A1 gibberellin A4

Figure 28.16.14: Structures of gibberellin A1 and A4

Gibberellin initiates signaling by binding to the nuclear gibberellin receptor. One such receptor is the Gibberellin Insensitive Dwarf1 (GID1). When bound, it leads to the proteolysis of another protein bound to it called a DELLA protein (an example is GAI), a transcriptional regulator that inhibits growth. The control of transcriptional activity by gibberellins is hence reminiscent of that of auxins.

Figure 28.16.15 shows an interactive iCn3D model of the gibberellin(GA3)- active gibberellin receptor GID1L1 bound to the DELLA domain of GAI (2ZSH).

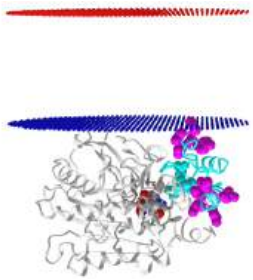


Figure 28.16.15: gibberellin(GA3)- Active gibberellin receptor GID1L1 bound to the DELLA domain of GAI (2ZSH). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...jVpzjnbUoQMhz5

The gibberellin receptor is gray and associated with the inner leaflet (blue) of the membrane. The plant hormone gibberellin A3 is bound to the receptor. Amino acid side chains in the receptor involved in the interactions with gibberellin A3 are shown in sticks, colored CPK. The DELLA protein GAI RAG protein is shown in cyan. The 3 amino acid motifs within it (DELLA, cyan spacefill; and VHYNP and LExLE, both magenta spacefill) are also shown. As with the auxin receptor, GID11A binds gibberellin in a deep pocket, which is covered by an N-terminal helix of the receptor. That helix recognizes and binds to the DELLA sequence in the DELLA transcription regulator protein.

Figure 28.16.16shows the effects of a mutation that leads to deficiencies in gibberellin 1-3 (right-hand side).

Figure 28.16.16: Physiological action of GA as illustrated by a comparison of the Landsberg *erecta Arabidopsis* plant with a GA-deficient mutant (ga1-3). In the absence of a GA, response stem elongation, leaf enlargement, floral development, seed set, and fruit development do not occur

The mechanism by which GAs promote growth is shown in Figure 28.16.17.

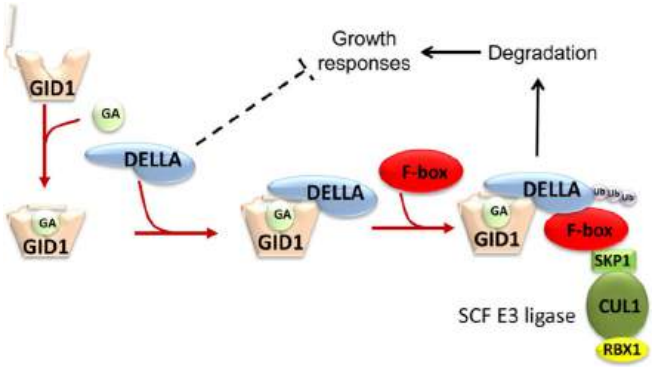


Figure 28.16.17: Representation of GA perception and signal transduction.

Binding of bioactive GA results in a conformational change in the GID1 receptor that promotes interaction with DELLA proteins. Recruitment of an F-box protein initiates ubiquitination of DELLA by an SCF E3 ubiquitin ligase targeting the DELLA for proteasomal degradation. Loss of DELLA relieves growth repression and suppresses other DELLA-mediated responses

28.16.5: Abscisic acid

Much of this material derives from Hewage et al. (2020). Advanced Science. https://doi.org/10.1002/advs.202001265. This is an open-access article under the terms of the Creative Commons Attribution License.

The phytohormone abscisic acid (ABA) is the best-known stress signaling molecule in plants As such it will be key as plants struggle to adapt to climate change. Its structure is shown in Figure 28.16.18

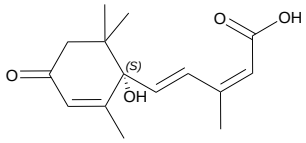


Figure 28.16.18: Structure of abscisic (ABA)

ABA protects land plants from biotic and abiotic stresses. ABA receptors proteins (**PYLs**) contain a conserved pyrabactin resistance/pyrabactin resistance-like/regulatory domain (**PYR/PYL/RCAR**) that binds ABA and triggers a cascade of signaling events.

ABA has significant roles throughout a plant's life cycle. From the single-celled zygotic stage to the mature multicellular plant, plant developmental stages involve ABA. ABA allows germination only under optimum conditions and inhibits growth under stress conditions. The adult plant as well as the seedling experience biotic and abiotic stressors that vary in severity and persistence. ABA allows that plant to survive by inducing both short-term and long-term stress responses, including rapid and reversible stomatal closure, long-term growth inhibition, dormancy, senescence, and abscission. ABA is therefore both a developmental and a stress-signaling molecule with diverse roles, as shown in Figure 28.16.19

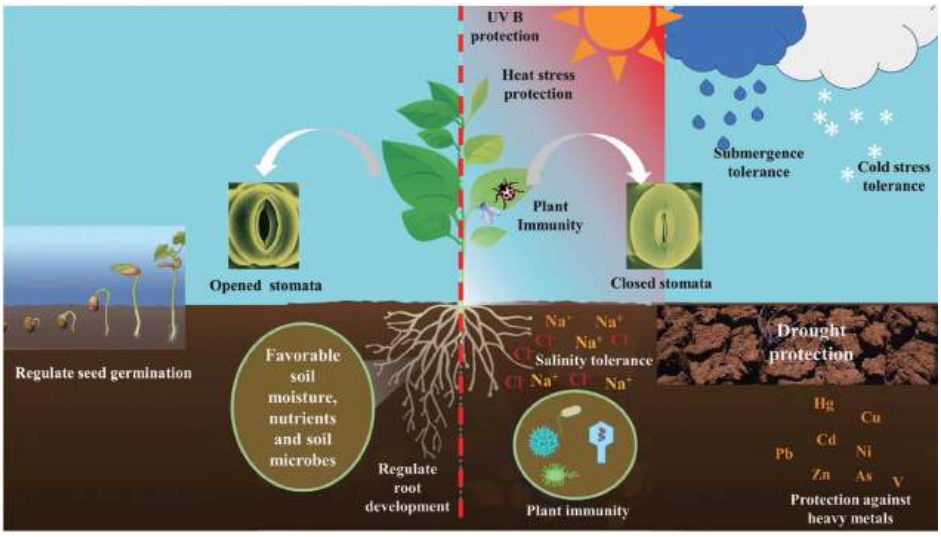


Figure 28.16.19: Summary of ABA involvement in the development and stress tolerance.

ABA signaling in drought

Let's look at a specific example of ABA signaling in the presence of drought, stress that will expand as the world's climate changes due to the combustion of fossil fuels.

Insufficient levels of soil water can result in an imbalance of water between the cells and the outer environment. A resulting change in the cellular electrolyte content affects metabolism, resulting in an osmotic imbalance or stress. The osmotic stress thus leads to the accumulation of ABA in cells and triggers ABA signaling. The cellular pool of ABA is dramatically increased during drought. Biosynthesis, catabolism, conjugation, and transportation of ABA are coordinated to increase ABA levels. ABA rapidly regulates plant water levels by controlling stomata. Stomata, which are microscopic pores controlled by two highly differentiated epidermal cells (guard cells), have the primary role in regulating gas exchange between the air and plant. Open stomata allow CO_2 to diffuse into the leaf mesophyll and reach the sites of photosynthesis. They also allow water vapor to exit from the plant interior to the atmosphere.

By allowing transpirational water loss, stomata allow the cooling of the plant and the managing of the interior water levels. Thus, stomata are essential regulators that connected the plant interior to the outside environment. Increases in osmotic pressure in guard cells lead to water uptake and then to cell expansion; as the cells expand, the pore opens because of differential thickenings of guard cell walls. Stomatal movements are regulated by numerous environmental signals such as light, plant growth regulators, pathogens, drought, cold, and nutrient status. Stomatal movement is the quickest response to ABA signaling. Therefore, the core

ABA signaling is essential for guard cell function. The involvement of ABA signaling events in guard cell function is summarized in Figure 28.16.20

Figure 28.16.20: Simplified representation of ABA-mediated core signaling events in guard cells.

In the absence of ABA, (left) ABA receptors (PYLs) are in ligand-free form. H^+ATPase action pumps H^+ ions outside of the plasma membrane. The SnRK2 protein kinases and the S-type anion channel SLAC1 are kept dephosphorylated by PP2Cs. The dephosphorylation state of SLAC1 prevents the nonspecific activation of S-type anion channels.

In the presence of ABA (right) PYLs bind to and inhibit PP2Cs. ABA inhibits H^+ATPase activity, blocking the H^+ pumping outside. The Ca^{2+}-independent protein kinases (SnRK2s) are released from PP2C inhibition and activated by auto-phosphorylation. Ca^{2+}-permeable cation (I_{Ca}) channels are released from PP2C-mediated inhibition, causing increases of ABA-responsive Ca^{2+} in cytosol leading to activate CPKs. The activated SnRK2s and CPKs phosphorylate SLAC1 The SnRK2.6/OST1 protein kinase phosphorylates and activates the R-type anion channel ALMT12/QUAC1. The K^+ ions are effluxed via voltage-dependent outward K^+ (K^+_{out}) channel GORK, causing a guard cell turgor decrease leading to stomatal closure. PYLs: ABA receptors; ABA: abscisic acid; PP2C: protein phosphatase 2C proteins; OST1: open stomata 1/SnRK 2.6 protein kinase; Ca/CPK: $Ca^{2+/}$calcium dependent protein kinases; I_{Ca2+}: plasma membrane nonselective cation channel permeable to Ca^{2+} SLAC1: slow anion channel-associated 1 (SLAC1); QUAC: aluminum-activated malate transporter 12/quickly activating anion channel 1 (ALMT12/QUAC1); GORK: guard cell outward rectifying K^+ channel (GORK); KAT1: K^+ activated 1 potassium ion channel; A−: anions; K^+: potassium ions.

The ABA receptor core complex

There are three main phases of ABA signaling: ABA synthesis/metabolism, long-distance transport, and ABA binding to its receptor. Downstream signaling ensues. through transcriptional activators/repressors and plasma membrane-located channel proteins. The main components of the core ABA signaling pathway are shown in Figure 28.16.21

Figure 28.16.21: Components of the Core ABA signaling pathway.

In the absence of ABA (A above), SnRK2 kinases are dephosphorylated by protein phosphatase 2C (PP2Cs). In the presence of ABA (B above) PP2Cs are inhibited by the complexes PYLs-ABA. Thus, the SnRK2 kinases are released and make a cascade of downstream transcription factors, NADPH transporters, and ion channels phosphorylate the transcription factors that induce ABA-responsive gene transcription, and ion channels act on the guard cells to bring about transpirational control.

ABA receptors (PYLs) bind ABA, PP2C, and protein kinases. The ABA:PYL complex binds PP2Cs, leading to conformational changes in the active sites of PP2Cs that inhibits the phosphatase. This in turn leads to the release of downstream protein kinases (SnRK2s) from PP2C-mediated inhibition. The SnRK2s undergo autophosphorylation to activate a series of ion channels, NADPH oxidases, and transcription factors via phosphorylation. This activates both short-term and long-term ABA responses such as stomatal closure and upregulation of ABA-dependent gene expression. MAPKKKs (MAPK3s) also activate SnRK2.6 by phosphorylating a specific site during salinity stress.

ABA binding regulates a double-negative regulatory system, in which the ABA receptor (PYLs) act as ABA receptors, PP2Cs as negative-regulatory coreceptors, and SnRK2s as negative regulators. In addition to the regulation by SnRK2 and PP2Cs, several post-translational modifications also regulate ABA signaling. Phosphorylation, dephosphorylation, ubiquitination, farnesylation, and sumoylation have been found to modulate ABA signaling by targeting core components (PYLs or PP2Cs) or other interacting proteins downstream.

ABA Receptors (PYLs)

PYLs are soluble proteins, and among the 14 PYLs in *Arabidopsis*, 13 functions as ABA receptors. All PYLs are known to share a dominant helix-grip structure. This characteristic motif consists of a seven-stranded antiparallel β-pleated sheet, which is flanked by two α helices. The β-pleated sheets enfold a long carboxy-terminal α-helix of PYLs. The apo-PYLs contain a sufficiently large hydrophobic pocket of 543˚A between the C-terminal helix and β sheet. The size of this pocket is estimated to be 480˚A in the ABA-bound state. The 23 pocket residues are highly conserved and are more hydrophobic than the other parts of PYLs. The interactions of ABA and PYL2 are shown in Figure 28.16.22

Figure 28.16.22 The binding mode of ABA and PYL2 (PDBID: 3KDI) in A) 3D and B) 2D (redrawn).[211] In the 3D structure, the cartoon of PYL2 is colored in white (A). The important residues and ABA are shown in sticks with blue and yellow colors

respectively. The H-bonds are marked with red dotted lines.

Figure 28.16.23shows an interactive iCn3D model of Abscisic acid bound to the Abscisic Acid Receptor (PYL2) (pdbid: 3KDI)

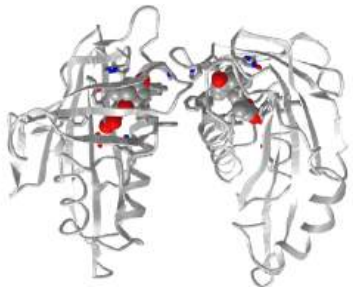


Figure 28.16.23: Abscisic acid bound to the Abscisic Acid Receptor (PYL2) (pdbid: 3KDI) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...uToGsSPpnZfPeA

In the absence of ABA, the apo-PYL2 has a pocket surrounded by four surface loops. When ABA binds, one loop (CL2) closes onto the pocket, forming a PP2C binding site for the phosphatase ABI1 and ABI2. This blocks the active site of the phosphatase.

Figure 28.16.24shows an interactive iCn3D model of ABA-bound PYL1 and the Protein Phosphatase 2C ABI1 (pdbid 3kdj)

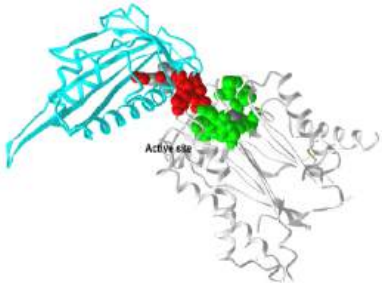


Figure 28.16.24: ABA-bound PYL1 and the Protein Phosphatase 2C ABI1 (pdbid 3kdj) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...tQgyC7NVXk1UVA

The PP2C phosphatase (ABI1) is shown in gray with its active site highlighted in green spacefill. The ABA receptor PYL1 is shown in cyan with ABA shown in sticks, colored CPK. The CL2 loop of the ABA-bound PYL1 receptor is shown in red spacefill. It projects into the PP2C active site, inhibiting its activity.

28.16.6: Light Signaling through Phytochromes

Much of this material derives from Liu, Y., Jafari, F. & Wang, H. Integration of light and hormone signaling pathways in the regulation of plant shade avoidance syndrome. *aBIOTECH* **2,** 131–145 (2021). https://doi.org/10.1007/s42994-021-00038-1. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/. get rid of this

Plants deal with competing plants in regions of high plant density by sensing changes in the intensity and wavelengths of light. Signaling leaves to responses (stem elongation, reduced branching, early flowering, etc) called shade avoidance syndrome (SAS). A photosensory system initiates signaling that alters gene transcription. In the SAS in plants in a large canopy, the upper leaves used red and blue light for photosynthesis. Multiple photoreceptors are used. Some transcription factors are also sensitive to light. For example, PIF3, a transcription factor, binds to light-responsive genes only when it binds to another transcription factor called Pr. Pr is resident in the cytoplasm but moves to the nucleus after altering conformation on absorbing red light.

Legris, M., Ince, Y.Ç. & Fankhauser, C. Molecular mechanisms underlying phytochrome-controlled morphogenesis in plants. *Nat Commun* **10,** 5219 (2019). https://doi.org/10.1038/s41467-019-13045-0. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/. do

Phytochromes are present in bacteria, cyanobacteria, fungi, algae, and land plants. We will focus mostly on phytochromes in Arabidopsis. In land plants, phytochromes are red and far-red light receptors that exist in two forms. They are synthesized in the inactive Pr state, which upon light absorption converts to the active Pfr conformation. Pfr is inactivated upon far-red (FR) light absorption or through thermal relaxation, which depends on temperature. Phytochromes act as dimers, resulting in three possible phytochrome species: Pr–Pr, Pfr–Pr, and Pfr–Pfr. Pr and Pfr have different absorption maxima, but due to overlapping spectra both conformers are always present in the light while only prolonged darkness returns all phytochrome to Pr. Given that phytochrome

responses depend on the proportion of Pfr conformers, signaling is influenced by a combination of light quantity, color, and temperature. These features of phytochromes are summarized in Figure 28.16.25

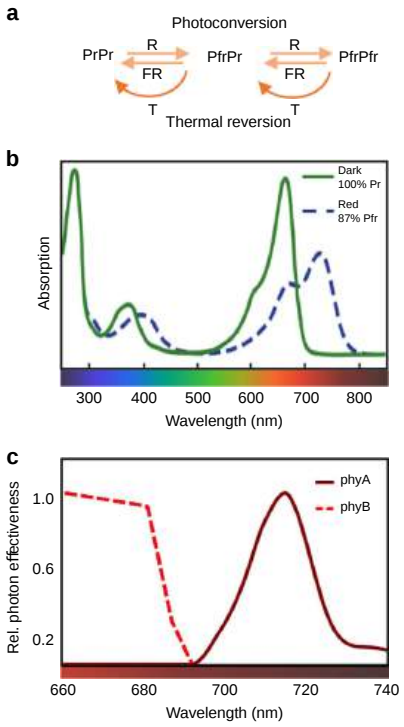


Figure 28.16.25: Control of phytochrome activity.

Panel a shows factors that control phytochrome activity. Phytochromes exist in two conformations, Pr and Pfr, the latter being the active form. They exist as dimers so three species can be found. Each monomer can be activated by red light (R) and inactivated by far-red light (FR) or by thermal reversion, a process that depends on temperature (*T*). At least in the case of phyB, Pfr in heterodimers reverts much faster than that in homodimers, allowing phyB to perceive temperature both during the day and during the night.

Panel b shows plant phytochrome absorption spectra of the Pr and Pfr conformations. In dark-adapted seedlings, phytochromes are in the Pr form. Upon a saturating R pulse, due to overlapping absorption spectra of Pr and Pfr, only 87% of Pfr is achieved.

Panel c shows action spectra for phyA and phyB in the control of hypocotyl elongation. Fluence rate (number of particles passing per unit time) response curves are measured at different wavelengths and fluence rate that leads to 40% inhibition compared with dark control is determined. To specifically determine action spectra for phyA and phyB, for phyB the curve was performed with phyB-GFP/*phyAphyB* seedlings, and for phyA using *phyB-5* seedlings. Values are relative to the response obtained at the most efficient wavelength in each case

Plant phytochrome structure

Plant phytochromes are dimeric, each monomer consisting of ~1150 amino acids. The chromophore, a linear tetrapyrrole named phytochromobilin (PΦB), whose structure is shown in Figure 28.16.26, is attached to the protein.

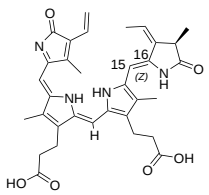


Figure 28.16.26: Structure of phytochromobilin (PΦB)

The domain structure of phytochromes is shown in Figure 28.16.27.

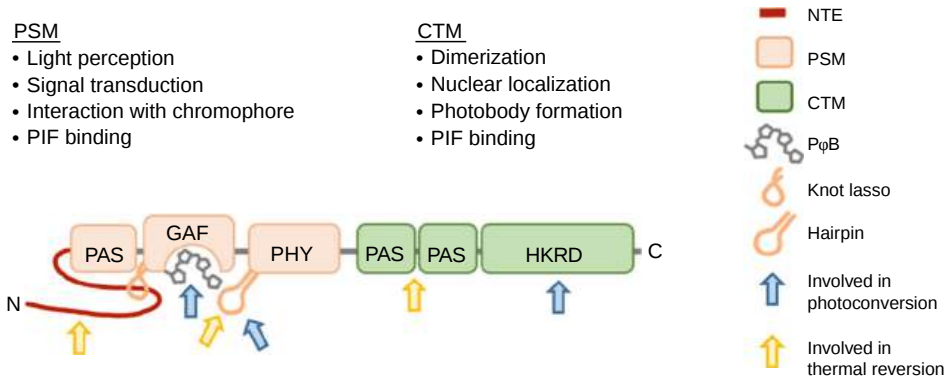


Figure 28.16.27: Structural domains of phytochromes and their role in the perception of environmental signals and downstream signaling. NTE N-terminal extension, PSM photosensory module, CTM C-terminal module.

The apoprotein can be divided into the N-terminal PSM, which consists of the N-terminal extension (NTE), for which structural information remains scarce, and three structurally related domains Period/Arnt/SIM (PAS), cGMP phosphodiesterase/adenylyl cyclase/FhlA (GAF), and a phytochrome-specific domain (PHY) and a C-terminal module (CTM) comprising two PAS domains and a histidine kinase-related domain (HKRD). The chromophore is bound covalently to a conserved cysteine in the GAF domain, which has intrinsic chromophore lyase activity. Light perception triggers a Z to E isomerization around the C15–C16 double bond of PΦB, which leads to a cascade of structural modifications in the protein. Figure 28.16.28shows an interactive iCn3D model of phytochrome (Deinococcus) Pfr form in the Photoactivated State (5C5K)

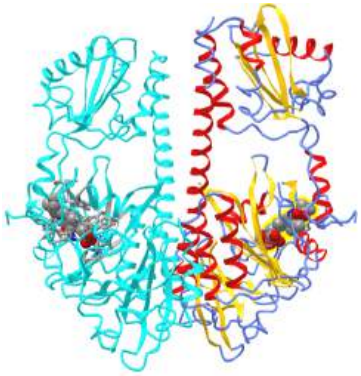


Figure 28.16.28: Phytochrome (Deinococcus) Pfr form in the Photoactivated State (5C5K) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...VzE7AqtgejbfeA

The protein is shown in its active dimeric state. One chain is shown in secondary structure colors and the other in cyan. The chromophore (heme derivatives) are shown in spacefill in both subunits. Side chains surrounding the chromophore are shown in colored sticks in the cyan chain.

At first glance, the presence of the histidine kinase-related domain (HKRD) would seem to suggest that phytochromes transduce their signal through the C-terminal module (CTM). Although many bacterial and cyanobacterial phytochromes have a C-terminal histidine kinase domain and act as light-regulated histidine kinases, plant phytochromes are not histidine kinases, and their role as Ser/Thr kinases remains contentious. The photosensory module (PSM) fused to a nuclear localization signal and a dimerization sequence is sufficient to restore most phyB functions, pointing to key signaling functions of the PSM.

Major The major of the plant phytochrome CTM are dimerization, nuclear import, and localization to sub-nuclear structures known as photobodies. However, it was recently shown that the C-terminal part of phyB also engages in light-regulated interactions and regulation of PIF activity. Moreover, the activity of the CTM is controlled by post-translational modification with SUMOylation limiting the ability of active phyB to interact with downstream signaling targets thereby limiting light responses. In addition, the CTM modulates active (Pfr) phytochrome levels with the HKRD inhibiting the Pr–Pfr photoconversion while the PAS–PAS promotes thermal reversion. Hence, while the division of plant phytochromes into PSM and a CTM helps describe the molecule, both parts of the photoreceptor contribute to the regulation of active Pfr levels and downstream signaling activities.

Figure 28.16.29shows a simplified mechanism for phytochrome control of transcription factors in different light environments.

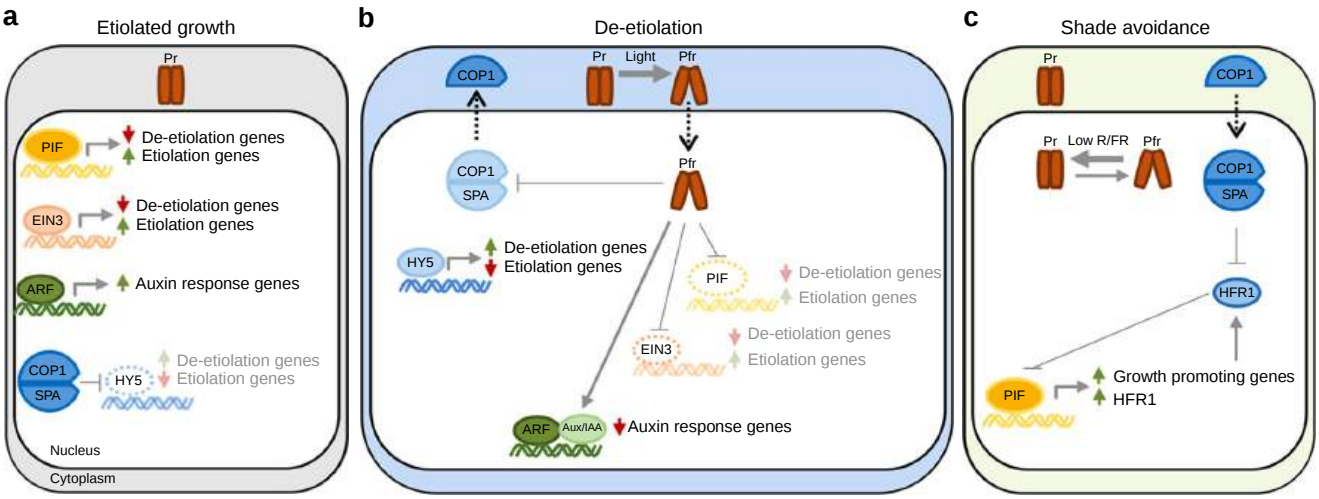


Figure 28.16.29: Simplified mechanism for phytochrome control of transcription factors in different light environments.

Panel a shows the response below the soil surface during growth in partial or a complete absence of light (called etiolated growth). For simplicity, we consider that phytochromes remain inactive (Pr) below the soil surface, which results in the accumulation of transcription factors PIFs, EIN3, and ARFs and subsequent induction of etiolation and auxin response genes. The COP1/SPA ubiquitin E3 ligase accumulates in dark and leads to proteasome-mediated degradation of HY5, a transcription factor that suppresses the expression of genes required for etiolation and induces expression of genes required for de-etiolation.

Pane b shows changes that occur when light intensity increases (de-etiolation). Light perception activates phytochromes (Pfr) which promote de-etiolation by directly inhibiting PIFs and EIN3, and indirectly inhibiting ARFs by stabilizing Aux/IAA proteins. The Pfr form of either phyA or phyB interacts with SPA proteins, resulting in the inhibition of COP1/SPA. This results in the stabilization of HY5 leading to the induction of de-etiolation-related gene expression and repression of etiolation genes.

Pane c shows de-etiolated plant in response to shade (reduced R/FR). Low R/FR in shade reduces the fraction of active phytochrome (Pfr/Ptot). PIFs accumulate and induce growth-promoting gene expression. In addition, PIFs induce a negative feedback loop exemplified by *HFR1* expression. HFR1 (and other HLH proteins) binds to PIFs forming non-DNA-binding heterodimers. COP1/SPA is also involved in this loop by leading HFR1 to proteasome-mediated degradation. Arrows indicate positive regulation, blunt-ended arrows indicate negative regulation and dotted-lined arrows indicate nucleo-cytoplasmatic relocalization

Figure 28.16.30summarizes how phytochromes affect transcription.

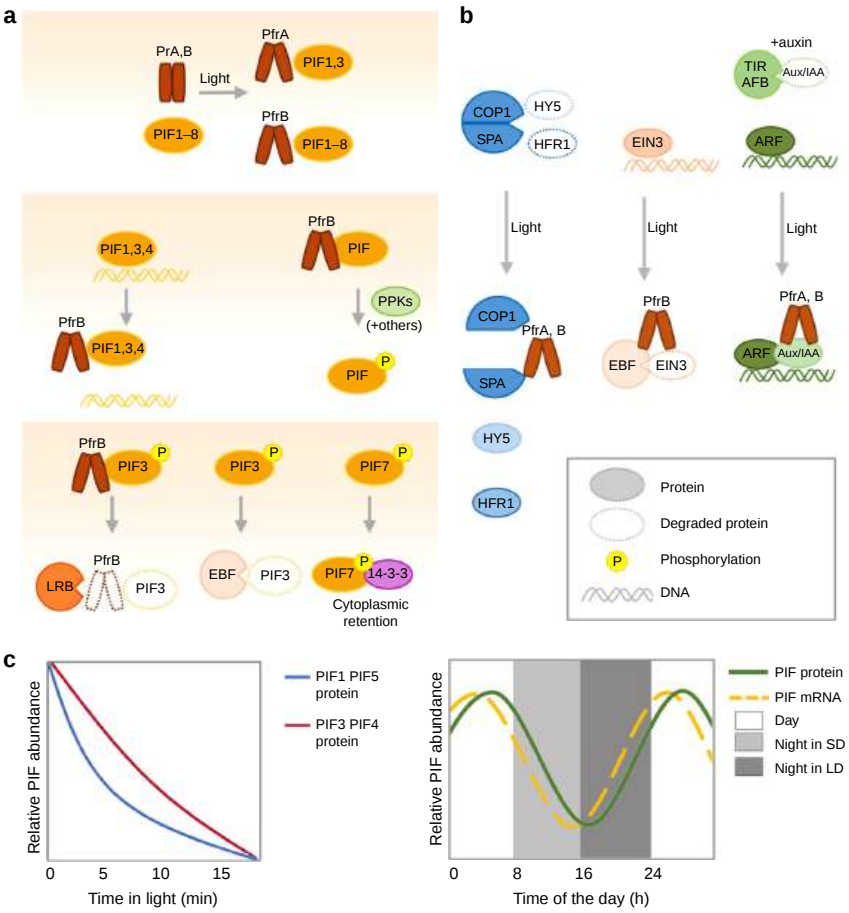


Figure 28.16.30: Mechanisms of phytochrome-mediated regulation of transcription factors.

Panel a (top to bottom) shows sequential steps by which Pfr inhibits PIFs. Top: PfrA interacts with PIF1 and PIF3 while PfrB interacts with PIF1–PIF8. Middle left: for PIF1, 3, and 4 phytochrome inhibits DNA binding. Middle right: Interaction with Pfr leads to the phosphorylation of PIFs. Many kinases have been found to phosphorylate PIFs (see text) with PPKs phosphorylating PIFs in response to light. Bottom: after light-induced phosphorylation, PIF3 is degraded by LRBs and EBFs with phyB co-degradation occurring in the LRB-mediated process (left, center), phosphorylated PIF7 interacts with 14-3-3 proteins and remains in the cytoplasm (right).

Panel b shows other mechanisms of transcriptional control by phytochromes. Left: PfrA and PfrB interact with SPA and inhibit the COP1/SPA complex. Center: PfrB interacts with EIN3 to promote ERF-mediated EIN3 degradation. Right: PfrA and PfrB interact with Aux/IAA to prevent their degradation by $SCF^{TIR1/AFB}$.

Panel c shows Patterns of PIF abundance depending on the developmental state and growth conditions. In etiolated seedlings PIFs accumulate to high levels, promoting etiolated growth. Upon light exposure, PIFs are rapidly degraded in a phytochrome-dependent manner, with half-lives of ~5 min for PIF1 and PIF5, and ~10 min for PIF3 and PIF4 (left). In contrast, in light-grown seedlings PIFs are under strong transcriptional control, allowing accumulation of the protein even in conditions when phytochrome activity is predicted to be high (right), SD (short days), LD (long days)

28.17: Signal Transduction - Vision and Olfaction

28.17.1: Sensory Transduction in Photoreceptors and Olfactory Sensory Neurons

Photoreceptors and olfactory sensory neurons (OSNs) have highly specialized structures that enable them to capture their respective stimuli of light and odorant ligands. Both photoreceptors and OSNs have evolved highly specific abilities to detect and discriminate light wavelengths or odors. They use intricate transduction mechanisms to convert sensory stimuli into electrical signals. Their transduction cascades not only can greatly amplify the signal but also to enhance the signal to noise, enabling these cells to detect and distinguish minute stimuli within very noisy background conditions. Such transduction mechanisms provide for modulation at multiple steps to adapt the sensory neurons to different background stimulation and optimize the capture of useful information about the surrounding world.

In this review, we summarize some of the key structural and functional features of vertebrate rod and cone photoreceptors and of OSNs, and the molecular mechanisms that underlie their function. While describing the features of both cell types, we emphasize the similarities and differences between photoreceptors and OSNs and the unique features of each cell type that make them perfectly suited to perform their function.

28.17.2: Signal Detection in Photoreceptors and Olfactory Sensory Neurons' Specialized Cilia

Vertebrate rod and cone photoreceptors as well as OSNs are ciliary neurons (Figure 1) with specialized cilia where the initial detection of the sensory stimulus takes place to activate a sensory transduction cascade. Rods and cones have a single cilium that has evolved to accommodate a stack of ~1,000 membrane disks where the visual pigment is expressed at a very high 3–5 mM concentration (Figure 1A; Palczewski, 2006). In the case of rods, the disks are enveloped by the plasma membrane, whereas in cones the disks are formed by invaginations of the plasma membrane. As light enters the eye and reaches the retina, it travels along the length of the rod and cone outer segments. The orientation of the elongated outer segments along the light path, together with the high density of visual pigment in their disks results in ~50% probability that an incident photon is absorbed by a visual pigment molecule (Bowmaker and Dartnall, 1980). In the case of OSNs (Figure 1B), odorant ligands are detected in the ~20 cilia protruding from each dendritic knob which are immersed in the mucus layer covering the olfactory epithelium. The olfactory cilia, which are motile in amphibians but not in rodents, are only about 0.1–0.2 μm thin but can reach up to 100 μm in length depending on the species (Kleene and Gesteland, 1981; Ukhanov et al., 2021). While this greatly increases the surface membrane area available to incorporate olfactory receptor (OR) proteins to detect odorants, it also greatly reduces the ciliary volume with potentially detrimental effects (see below).

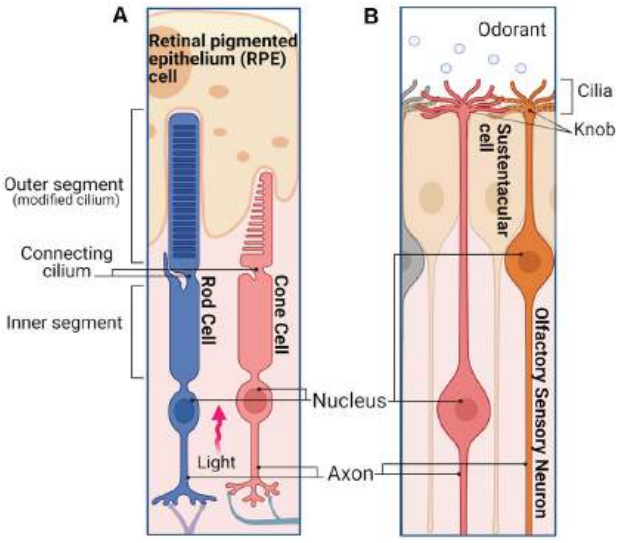
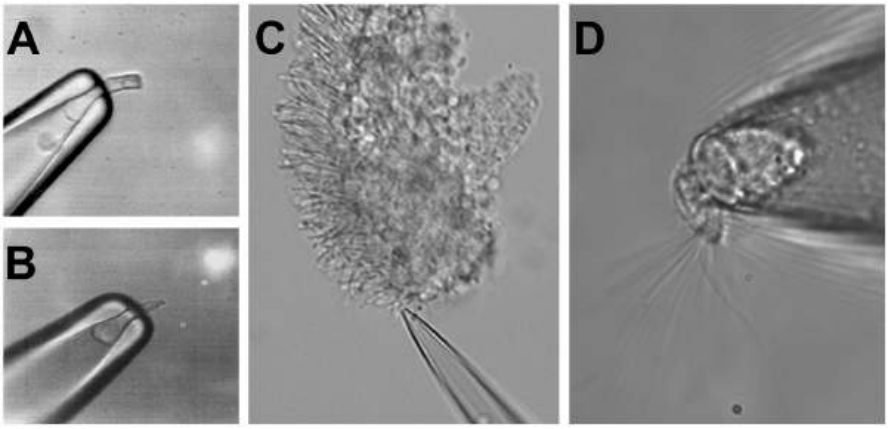


Figure 1. Photoreceptors and olfactory sensory neurons (OSNs). **(A)** Simplified schematic representation of a rod and a cone in the retina. Photoreceptors are polarized neurons with a specialized morphology optimized to detect light stimuli. The outer segments of

both rods and cones are modified sensory cilia, containing membrane disks organized in a stack. In the case of rods, the outer segment has a slim rod-like structure in which the disks are enclosed by the plasma membrane. The outer segment of the cones has a stocky conical-shaped structure, in which the disks are constituted by invaginations of the plasma membrane. The outer segment does not contain any proteins of the cell translation machinery, which are mostly localized in the inner segment, including the endoplasmic reticulum, Golgi, and mitochondria. Outer and inner segments are connected by the connecting cilium, while distal to the inner segment is the cell body containing the nucleus, followed by the axon and synaptic termini that extend into the outer plexiform layer where they synapse with the second-order neurons. When the light enters the eye, after reaching the retina, it travels along the length of the rod and cone inner segment until finally reaching the outer segments. **(B)** Simplified schematic of an OSN in the olfactory epithelium. OSNs are ciliated bipolar neurons, their apical dendrites extend to the surface of the epithelium terminating with a spherical structure called a dendritic knob, from which the sensory cilia enter the mucus layer. The ciliary membrane contains the olfactory receptors (ORs) necessary to detect different odorants. Distal from the knob is the cell body of the OSN with its nucleus, followed by a long axon that projects to the olfactory bulb, where it synapses with the second-order neurons. Images created with BioRender.com.

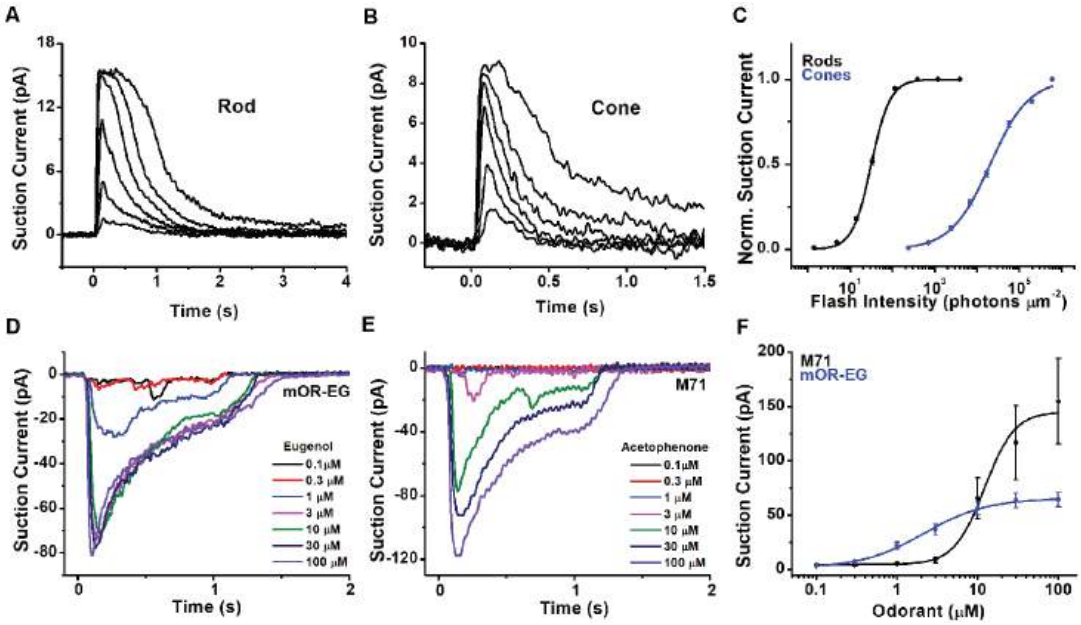
28.17.3: Electrophysiological Approaches to Record Light- and Odorant-Induced Responses

The similar morphological structure of rods, cones, and OSNs, with a ciliary part able to detect the respective stimuli and an adjacent cell body, allows similar electrophysiological approaches to recording stimulus-induced responses in these cell types. The cell body of a photoreceptor or an OSN can be sucked into the tip of a recording pipette by using a loose-patch (or suction pipette) recording configuration (Baylor et al., 1979; Lowe and Gold, 1991). This leaves the outer segment of photoreceptors or the olfactory cilia exposed and accessible to bath solution changes, e.g., the application of pharmacological agents or odorants, in the case of OSNs. Suction pipette recordings can be performed from isolated sensory neurons, as shown in Figures 2A,B,D (respectively, a salamander rod, salamander cone, and salamander OSN) but also from dissected retina tissue, as in the case of the outer segment of a mouse rod drawn in the recording electrode from a piece of the retina (Figure 2C). This recording configuration measures the transduction current entering the photoreceptor outer segment or olfactory cilia, and leaving *via* the cell body.



Figures modified from Reisert and Matthews (2001), Kefalov et al. (2010), and Kefalov (2012) with permission.

A fundamental difference between photoreceptors' and OSNs' responses to stimuli lies in their polarity. In the absence of light, rods, and cones are kept depolarized by a standing inward current of approximately 20–40 pA for amphibian cones and rods, and 7–15 pA for mouse photoreceptors. This depolarizing current is gradually suppressed upon light stimulation until, for sufficiently high light intensities, it is reduced to zero (Figures 3A, B, mouse rod, and cone responses, respectively), leading to photoreceptor hyperpolarization. Similar to rods, but differently from cones, the OSNs show comparatively little spontaneous activity in absence of stimuli (Reisert, 2010; Connelly et al., 2013). Different OSNs show varying levels of spontaneous basal activity determined by the constitutive activity of their ORs (Reisert, 2010; Connelly et al., 2013).



Figures modified from Kefalov et al. (2010), Sakurai et al. (2011), Deng et al. (2013), and Dibattista and Reisert (2016) with permission.

In the presence of odorants, OSNs generate an inward receptor current which leads to depolarization, and the generation of action potentials (Firestein and Werblin, 1989; Kurahashi, 1989; Reisert and Matthews, 1999). This receptor current is odorant concentration-depend and increases progressively with increasing stimulation until it eventually saturates at high odorant concentrations. Responses recorded from OSNs expressing different olfactory receptors can generate fairly different response amplitudes when stimulated with their respective agonists (Figures 3D, E: responses recorded from mouse OSNs that express the mOR-EG or the M71 olfactory receptor, which are activated by the ligands eugenol and acetophenone, respectively).

The hyperpolarization and signals carried by graded potentials in photoreceptors vs. depolarization and signals carried by action potentials in OSNs represent another fundamental difference between these two types of sensory neurons. These topics and the differences in synaptic structure and transmission between photoreceptors and OSNs go beyond the focus of this review and are discussed in an excellent recent review on this topic (Lankford et al., 2020).

28.17.4: Sensitivity of Photoreceptors and Olfactory Sensory Neurons

In part due to their unique structure, photoreceptors, and, to a lesser extent, OSNs have achieved exquisite sensitivity that optimizes the detection of stimuli within the respective sensory organs. In addition, both sensory receptors use a transduction cascade to amplify the signal (see below). As a result, rod photoreceptors can reliably detect single photons (Baylor et al., 1979), enabling humans to perceive light with as few as six photons detected by adjacent rods (Hecht et al., 1942). This renders rods perfectly suited for dim light vision, with a dynamic range spanning lights from a dark cloudy night to sunrise (Fain et al., 2010). Cones are ~100-fold less sensitive than rods, making them suited for daytime light conditions. Figure 3C compares the intensity-response function of mouse rods and cones, demonstrating the much lower sensitivity of cones compared to rods.

Most OSNs respond to odor concentrations in the low micromolar range (Bozza et al., 2002; Grosmaitre et al., 2009; Saito et al., 2009; Lee et al., 2011; Dibattista and Reisert, 2016), but they can also reach exquisite sensitivity and are capable of detecting odors at the nanomolar concentration range. Picomolar sensitivity is reached by a subset of OSNs that express receptors specialized in detecting amines, the trace-amine-associated receptors (Zhang et al., 2013). In comparison to rods, OSNs do not reach such high sensitivity, and cannot be activated by a single odorant molecule but instead require around 30 odorant binding events to begin firing action potentials reliably (Bhandawat et al., 2010). The detection of odorants in the olfactory epithelium can be further enhanced by the expression of a wider number of different OR genes, more than 350 in humans and 1,000 in mice (Malnic, 2007), with overlapping response profiles to odorants. A larger number of OSNs, particularly in species relying heavily on their sense of smell, may enhance further the detection of odorants. For instance, the human olfactory epithelium covers ~3–4 cm^2 and contains approximately 5–6 million OSNs while in the case of dogs, the area of the olfactory epithelium is 18–150 cm^2 and contains 150–300 million OSNs (Lippi and Heaney, 2020).

28.17.5: Detection of Stimuli

In both photoreceptors and OSNs, the detection of stimuli is mediated by G protein-coupled receptors. In photoreceptors, this function is achieved by rod and cone visual pigments, which consist of a protein, opsin, covalently attached to the visual chromophore, typically 11-cis-retinal (Ebrey and Koutalos, 2001). The chromophore serves as a reverse agonist, keeping the receptor molecule in the inactive ground state (Crouch et al., 1996). Absorption of a photon by 11-cis-retinal triggers its conformational change to all-trans-retinal, which, in turn, results in the rearrangement of the opsin transmembrane helices and switch of the visual pigment molecule into its active state.

We discussed bacterial rhodopsin in an earlier section. We present it here again to show and note its similarity to animal light-sensing proteins.

Review: Retinal, Opsin and Light Transduction in Bacteria

Microbial **rhodopsins** are phototransducing proteins with a conjugated chromophore retinal, covalently attached to the protein opsin through a Schiff base (imine) linkage. The holoprotein (opsin with the attached retinal) is called rhodopsin. Retinal is derived from beta-carotene. The structures of animal and microbial retinals are shown below.

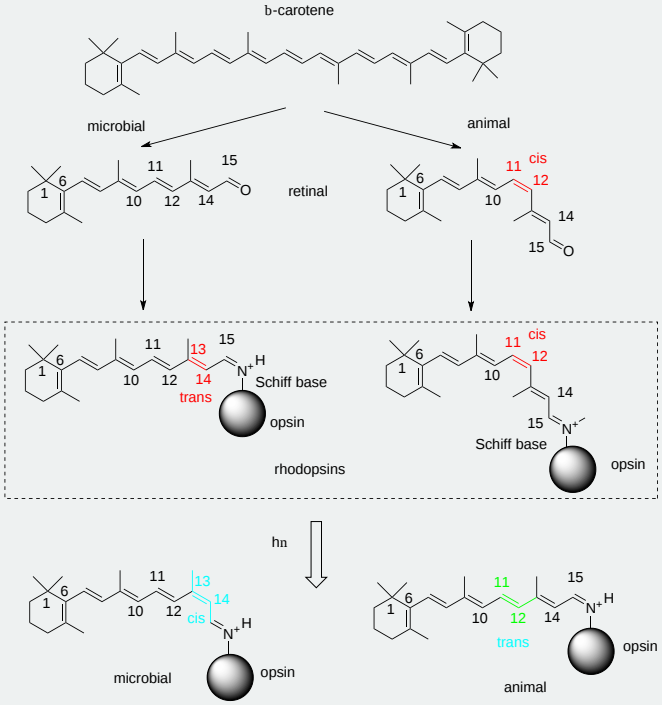
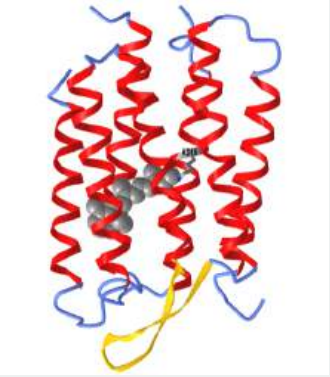


Figure 28.17.15: Structures of animal and microbial retinals

When light of the correct wavelength is absorbed, an electron in a pi molecular orbital in retinal is promoted to a pi antibonding molecular orbital, breaking a 2 electron pi bond in the structure at a certain site in the isoprenoid chain, allowing rotation around the now single bond. The final result after the electrons return to the ground state is photoisomerization of the trans 13-14 and cis 11-12 bonds in microbial and animal retinal, respectively, to their respective cis 13-14 and trans 11-12 configuration. This conformational change in the bound retinal induces a conformational change in the protein opsin, leading to signaling.

The figure below shows an interactive iCn3D model of one monomeric of bacteria rhodopsin (1C3W) containing retinal attached through a Schiff base to Lys 216.



Bacterial rhodopsin (1C3W). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...9HzvfdFvb9kcc6

The covalently attached retinal is shown in gray spacefill. The protein opsin is a membrane protein that spans it with seven helices. Hence it has very similar to a GPCR.

The activated visual pigment then binds to a G protein, transducin, activating it. The activation of transducin triggers the transduction cascade that ultimately generates the cellular response (Pugh and Lamb, 1993). Eventually, the all-trans-retinal chromophore is released from opsin after the covalent Schiff base between them is hydrolyzed, leaving behind chromophore-free opsin (Saari, 2016). Notably, without chromophore, opsin has residual activity, and in sufficient quantities can produce steady activation of the photoreceptors, similar to steady background light, thus modulating the sensitivity of photoreceptors (Fain et al., 1996). This process is known as bleaching adaptation, indicating the production of free opsin after the photoactivation of the visual pigment and dissociation of the visual chromophore.

Unlike in photoreceptors, where the ligand, a light-sensitive reverse agonist, is covalently attached to opsin, in olfaction, the ligands are dissolved in the mucus covering the surface of the olfactory epithelium and come into direct contact with the OR proteins expressed in the OSN ciliary membrane. This results in the activation of the receptor protein that, in turn, is transduced downstream to a G protein to trigger a transduction cascade resulting in the cellular response. The binding of the ligand to the receptor protein is noncovalent and rapidly reversible. ORs, like other G protein-coupled receptors, do display antagonism, inverse, and partial agonism, leading to suppressed responses to their agonists, a reduction in basal activity in the absence of stimulation or suppression of the maximal response (Firestein et al., 1993; Oka et al., 2004; Reisert, 2010).

28.17.6: Discrimination Between Stimuli

The spectral sensitivity of individual rod and cone photoreceptors is dictated by the absorption properties of their visual pigments. Typically, each photoreceptor type expresses only one type of opsin; in the case of the human retina, rods express rod opsin, whereas cones express long wavelength (LW, red), middle wavelength (MW, green), or short wavelength (SW, blue) opsin (Nathans, 1987). When bound to the chromophore, the amino acid structure of each opsin determines the optical properties of the resulting visual pigment and the spectral sensitivity of the photoreceptors expressing it. As a result, species existing in environments with characteristic light distribution, such as deep-sea fish, have visual pigments that have evolved to optimize their spectral sensitivity (Hope et al., 1997). A second factor controlling the optical properties of the visual pigment is the structure of the visual chromophore. Most species, including mice and humans, use 11-cis-retinal, a derivative of Vitamin A, also known as A1. However, some amphibians and fish also use 3,4-dehydro 11-cis retinal, also known as A2. This chromophore has an extra conjugated double bond in its structure, which shifts the absorption spectrum of A2 visual pigments to longer wavelengths compared to the A1 visual pigment embedded in the same opsin molecule (Corbo, 2021). Some aquatic and amphibian species use the A1/A2 chromophores to shift their spectral sensitivity from murky waters dominated by longer wavelengths of light to seawater and air, dominated by shorter wavelength lights (Bridges, 1964). One notable example includes the toad, where the retina is populated by A1 visual pigment in its ventral section, receiving light from above the surface of the water, and by A2 visual pigment in its dorsal section, receiving light from below the surface of the water (Reuter et al., 1971). A shift in the chromophore can also occur during the lifetime of the animal as its environment changes, such as the A2 to A1 shift in salamanders as they metamorphose from the larval (aquatic) to the adult (terrestrial) stage (Ala-Laurila et al., 2007), or the A2 to A1 shift in Atlantic salmon during migration from sea to freshwater (Beatty, 1966).

Similar to photoreceptors, the ligand specificity of the OSNs is also dictated by the expression of OR genes in their cilia. As photoreceptors, each OSN expresses generally only one receptor gene so that its ligand specificity is determined by the structure of the OR expressed in that particular cell. However, photoreceptors typically use no more than five opsin genes to cover the visible spectrum, while OSNs can use hundreds, in the case of humans, to thousand and more, for rodents and dogs, OR genes to cover the odor space (Malnic, 2007). The same OR can be activated by multiple odorants with different sensitivities, and a given odorant can activate different ORs with different half-maximal concentrations (Buck, 1996; Ache, 2020). This generates a complex mosaic of ORs and odorants response pairs. Figure 3F compares the dose responses of OSNs expressing either the mOR-EG or the M71 OR to eugenol and acetophenone, respectively. In this case, mOR-EG OSNs display higher sensitivity to their agonist compared to M71 OSNs. However, this does not preclude the possibility that the M71 OR is more sensitive to another ligand resulting in a more left-shifted dose-response relation than the one seen with acetophenone. Conversely, the dose-response relation of M71 OSNs to benzaldehyde is approximately 10-fold right-shifted compared to acetophenone (Bozza et al., 2002).

Determining the ligand specificity of ORs is an ongoing endeavor (Abaffy et al., 2006; Saito et al., 2009; Kurian et al., 2021). Due to the large number and diversity of OR genes, as well as the nearly endless number of odorant molecules, understanding the overall mechanisms that control their ligand binding affinity and specificity remains a challenge. Receptor modeling approaches to understand and predict OR–odorant molecule interactions can provide valuable insights but are somewhat hampered by the lack of a crystal structure of any vertebrate OR. The rhodopsin structure is often used as a guide and homology model to predict the structure of ORs (Katada et al., 2005; Bavan et al., 2014).

28.17.7: Sensory Transduction Activation

In both photoreceptors and OSNs, the detection of stimuli by their respective G protein-coupled receptors is converted into electrical signals *via* the activation of a G protein-coupled to a second messenger transduction cascade. The two pathways, though distinct, share an amazing level of similarity (Figure 4). Thus, in both cases, the second messenger is a cyclic nucleotide, cGMP in photoreceptors (Pugh and Lamb, 1990) and cAMP in OSNs (Sklar et al., 1987; Bakalyar and Reed, 1990). As a result, the activation of both transduction cascades results in a rapid shift in the equilibrium between synthesis and hydrolysis of the respective cyclic nucleotide, which is then sensed by the cyclic nucleotide-gated (CNG) transduction channels in the plasma membrane of the photoreceptor outer segment or olfactory cilium.

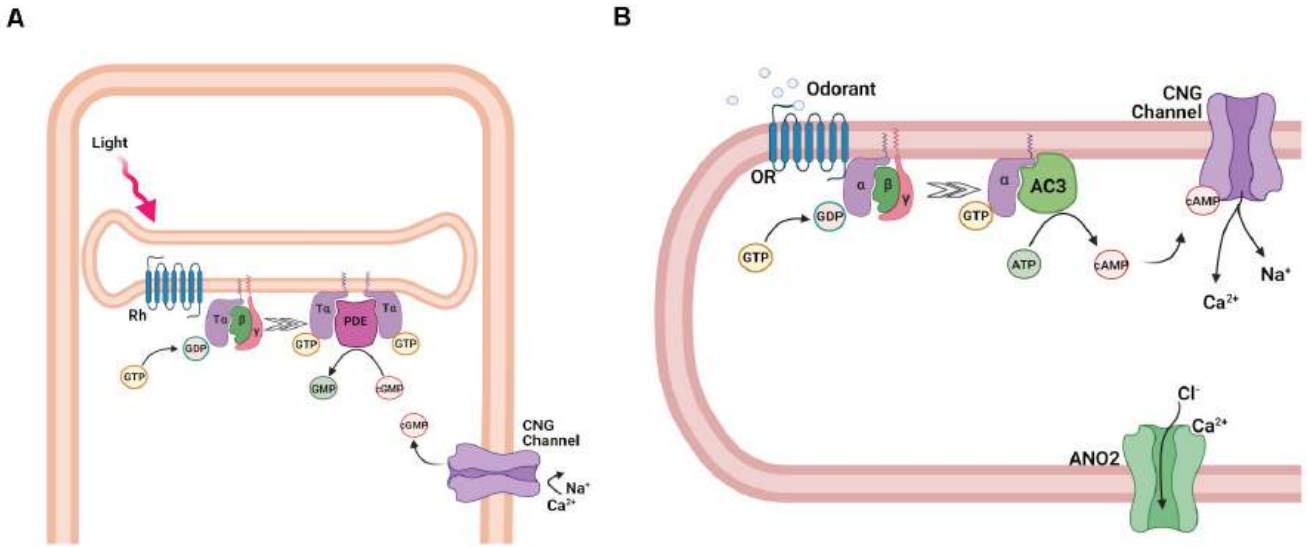


Figure 4. Activation of the transduction cascade in rod photoreceptors and OSNs. **(A)** Schematic representation of phototransduction cascade in rods. Abbreviations: rhodopsin (Rh), Tα, β, and γ subunits of the retinal G protein, transducin (T), guanosine-5′-triphosphate (GTP), guanosine-5′-diphosphate (GDP), phosphodiesterase (PDE), guanosine monophosphate (GMP), and cyclic guanosine monophosphate (cGMP), and cyclic nucleotide-gated (CNG) channel. **(B)** Schematic representation of the olfactory transduction cascade in OSNs. Abbreviations: Olfactory receptor (OR), guanosine-5′-triphosphate (GTP), guanosine-5′-diphosphate (GDP), $G\alpha_{olf}$, *β*, and *γ*, subunits of the olfactory G protein; adenylyl cyclase 3 (AC3), adenosine-5′-triphosphate (ATP), cyclic adenosine monophosphate (cAMP), cyclic nucleotide-gated (CNG) channel; Ca^{2+}-activated Cl^- channel anoctamin 2 (ANO2). Images created with BioRender.com.

In the case of photoreceptors, the photoactivated visual pigment binds to and activates the trimeric G protein transducin (T) (Figure 4A), causing the exchange of GDP for GTP on its α-subunit, which is part of the $G\alpha_t$ protein family. Following the subsequent dissociation of the α-subunit (Tα) from its β/γ complex (Tβγ), Tα then binds to the cGMP phosphodiesterase (PDE) complex, relieving the inhibition of its catalytic α- and β-subunits by its inhibitory γ-subunits (Ebrey and Koutalos, 2001; Burns and Arshavsky, 2005). All these transduction proteins are embedded in or tethered to the disc membranes inside rods or are contained in the cell membrane of cones. As a result of their activation, the hydrolysis of free cGMP in the outer segment is upregulated, causing its rapid decline, and partial or complete closure of the cGMP-gated channels expressed in the rod and cone cell membrane (Luo et al., 2008). The closure of the CNG channels leads to the reduction of the inward transduction current, followed by the hyperpolarization of the cells, and a reduction of neurotransmitter release to second-order neurons within the retina. Inversely, in the absence of light, the opening of CNG channels and the resulting inward transduction current is sustained by the continuous cGMP production by guanylyl cyclase (GC).

Similarly, in OSNs (Figure 4B), the ligand-activated OR proteins bind to the G protein G_{olf}, causing its dissociation into active $G\alpha_{olf}$ and olfactory β- and γ-subunit, $G\beta\gamma_{olf}$. In contrast to transducin, however, $G\alpha_{olf}$ is part of the $G\alpha_s$ protein family and binds to adenylyl cyclase 3 (AC3), activating it. As a result, the synthesis of cAMP in the olfactory cilia is upregulated, causing its rapid increase and the opening of cAMP-gated channels (Kleene, 2008; Su et al., 2009; Boccaccio et al., 2021).

While both photoreceptors and OSNs use CNG transduction channels, their respective channels have different subunit compositions (Bradley et al., 2005). Rods and cones express heterotetramers consisting of the main A1 and A3 and the modulatory B1a and B3 subunits in 3:1 and 2:2 stoichiometries respectively. The olfactory CNG channel is a heterotetramer consisting of two units of the main A2 subunits and one each of the modulatory A4 and B1b subunits. Interestingly, the rod and the olfactory CNG channels express different splice variants of the same B1 subunit. In OSNs, the initial inward Na^+ and Ca^{2+} current generated by the opening of the CNG channel raises ciliary Ca^{2+} and opens a secondary ion channel, the Ca^{2+}-activated Cl^- channel Anoctamin 2. A high intraciliary Cl^- maintained by the $Na^+/K^+/2Cl^-$ cotransporter 1 ensures a Cl^- efflux which further depolarizes the OSNs (Dibattista et al., 2017; Boccaccio et al., 2021). This depolarization triggers the generation of action potentials which further propagate along the axons, inducing glutamate release at synapses with the second-order neurons in the olfactory bulb (Murphy et al., 2004). In photoreceptors, the transduction cascade upon stimulation does not ultimately generate action potentials in the receptor cell, but only a graded receptor potential that directly causes a change in neurotransmitter release.

28.17.8: Amplification

As for any other sensory modality, proper amplification of the signal is required to detect small stimuli and the resulting high sensory sensitivity is critical for the survival and propagation of the species. Nature has reached the highest physically possible sensitivity in the case of rod photoreceptors that can produce a detectable electrical response to the absorption of a single photon. This impressive feat is achieved by employing a transduction cascade that allows tremendous amplification of the signal. During the ~50 ms active lifetime, a single photoactivated rhodopsin molecule activates ~20 transducins, producing an immediate 20-fold amplification (Burns and Pugh, 2010). The following activation of PDE by transducin does not directly produce amplification as each transducin has to bind to a PDE molecule to activate it. However, once activated, each PDE enzyme can hydrolyze thousands of cGMP molecules. Lastly, as the binding of cGMP to the CNG transduction channels is cooperative, a slight change in cGMP levels can reduce the number of cGMP molecules bound to the channel from 3 to 2. This results in channel closure and a sharp reduction in the transduction current, further enhancing the detection of photostimulation. Despite the similarities in the transduction cascades of rods and cones, the amplification in cone photoreceptors is substantially lower as a result of fine-tuning at several of the phototransduction steps (Yau, 1994; Kawamura and Tachibanaki, 2008). Interestingly, even though rod and cone visual pigments activate transducin with similar efficiencies, the lower thermal stability of the cone visual pigment results in higher intrinsic activity in cones compared to rods in darkness (Kefalov et al., 2003), effectively desensitizing the cones and shifting their function towards brighter daytime light conditions (see Figure 3C).

Curiously, the activation of G_{olf} by the OR molecule does not result in amplification. Indeed, the dwell time of the odorant ligand on the OR appears to be very short and on a millisecond timescale (Bhandawat et al., 2005). As a consequence, on average, this results in the activation of less than one G protein per activated receptor. As such, in contrast to phototransduction, where the lifetime of the activated rhodopsin greatly influences the response size and kinetics, in OSNs the response depends more prominently on the coupling efficacy of downstream transduction components while the odorant presence keeps the OR activated. To compensate for the lack of initial amplification at the G protein level, OSNs employ a secondary amplification step on top of the cAMP transduction cascade. The activation of AC3 by G_{olf} results in the synthesis of most likely hundreds of cAMP molecules, the opening of the CNG channels which is followed by a unique secondary amplification based on excitatory Ca^{2+}-activated Cl^-

channels in the cilia (Figure 4B). The Cl^- current carries up to 80% of the overall transduction current (Dibattista et al., 2017). Physiological experiments with pharmacological and genetic modulation of the Cl^- conductance indicates that the Cl^- channels serve to set the length of the action potential train generated in response to odorant stimulation (Pietra et al., 2016) and to promote recognition of novel odorants (Pietra et al., 2016; Neureither et al., 2017).

A puzzling aspect of the secondary amplification step is why Cl^- is the charge-carrying ion and not Na^+, which could be achieved easily by increasing the expression level and/or the ion permeation and conductance of the olfactory CNG channel. Recent theoretical work hinted at two main advantages of Cl^-, instead of Na^+, as the charge carrier. As the external environment of cilia is the nasal mucus, currents will depend on the ion concentration in the mucus, which can be unstable. A current that depends on the intracellular ion concentration, as is the case for Cl^- but not for Na^+, is much less dependent on the mucosal ion concentration. For instance, this could become an issue in the case of a cold with a runny nose or during swimming, when the mucus becomes diluted. The second advantage results from the "compromise" to increase the ciliary surface area, at the expense of having a very small ciliary volume, in the femtoliter range. In such small volumes, even small ionic currents can lead to large changes in ion concentration and osmotic pressures. If the main charge carrier was Na^+ this would lead to a large increase (tens of mM). This would cause a large increase in osmotic pressure and also would prevent Ca^{2+} clearance *via* the olfactory Na^+/Ca^{2+}, K^+ exchanger (see below) with greatly deleterious effects. In contrast, high intracellular Cl^- is maintained throughout the OSN so that its local depletion in the cilia upon ligand activation is rapidly reversed by diffusion from the cell soma. Both these issues do not exist for photoreceptors as they are embedded in the interstitial fluid of the eye and photoreceptors are sufficiently large and their transduction currents are sufficiently small that ion concentration changes due to changes in transduction currents are relatively small (Reisert and Reingruber, 2019). Nevertheless, rod photoreceptors undergo osmotically-driven length changes upon light activation, an effect that is mitigated by the translocation of G protein subunits into the cytosol (Zhang et al., 2017).

28.17.9: Receptor and G Protein Inactivation

Timely and effective transduction inactivation is critical for allowing sensory neurons to continue to detect stimuli with high temporal resolution. Equally important is to extract behaviorally relevant information from the presented stimuli. In both photoreceptors and OSNs, all active transduction components need to be turned off and the level of cyclic nucleotides within the cells needs to be restored to the resting level before the sensory cell can be reset to the inactive state and become ready for subsequent activation (Figure 5). In the case of photoreceptors, the identity of the step determining the overall kinetics of the photoresponse inactivation was the subject of intense research and debate over several decades. As the visual chromophore ligand is covalently attached to opsin, inactivation of the visual pigment could potentially be extremely slow. Indeed, if left on its own, the active state of rhodopsin decays with a time constant of ~50 s (Imai et al., 2007). Its inactivation in photoreceptors is a two-step process, involving phosphorylation of the rhodopsin C-terminus by rhodopsin kinase (GRK1) which partially quenches its activity, followed by the binding of arrestin1, which completely inactivates the visual pigment (Figure 5A). Though the decay of the active state of cone pigment is significantly faster at ~2 s (Fu et al., 2008), this is still clearly too slow to enable the timely termination and reset of phototransduction. Thus, in both rods and cones, the visual pigments are inactivated by phosphorylation by rhodopsin kinase and the subsequent binding of arrestin long before they would decay spontaneously (Makino et al., 2003). The effective time constant of rod visual pigment inactivation is ~50 ms (Krispel et al., 2006). The slowest step in the inactivation of rod phototransduction turned out to be the hydrolysis of GTP which shuts off Tα, a reaction driven by the transducin GTPase activity and enhanced by a GTPase (GTPase activating protein, GAP) complex consisting of Gβ5 and the membrane anchoring protein R9AP (Arshavsky and Wensel, 2013). Inactivation of transducin results in its release from PDE, allowing the two PDE γ inhibitory subunits to resume their inhibition on the two catalytic subunits (α and β) of this enzyme. The kinetics of this reaction determines the overall kinetics of response inactivation in rod photoreceptors. In contrast, work from amphibian cones suggests that in cones the photoresponse duration is Ca^{2+}-dependent and involves the quenching of the cone visual pigment (Matthews and Sampath, 2010).

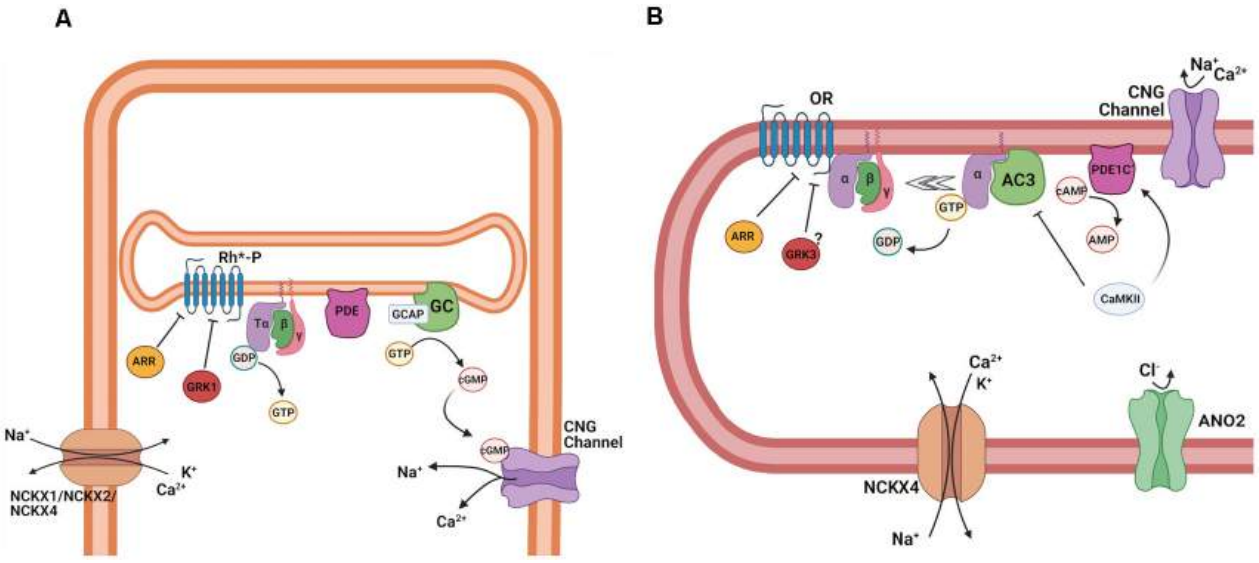


Figure 5. Termination of transduction cascade in rod photoreceptors and OSNs. **(A)** Schematic representation of the termination of phototransduction in rods. Abbreviations: Phosphorylated light-activated rhodopsin (Rh*-P), arrestin (ARR), G protein-coupled receptor kinase 1 (GRK1), Tα, β, and γ subunits of the retinal G protein, transducin (T), guanosine-5′-triphosphate (GTP), guanosine-5′-diphosphate (GDP), phosphodiesterase (PDE), guanosine monophosphate (GMP) and cyclic guanosine monophosphate (cGMP), and cyclic nucleotide-gated (CNG) channel, guanylate cyclase (GC), guanylate cyclase-activating protein (GCAP), cyclic nucleotide-gated (CNG) channels, K^+ dependent Na^+/Ca^{2+} exchanger 1, 2 and 4 (NCKX1, NCKX2, NCKX4). **(B)** Schematic representation of the termination of the olfactory transduction cascade. Abbreviations: Olfactory receptor (OR), arrestin (ARR), G protein-coupled receptor kinase 3 (GRK3) $G\alpha_{olf}$, *β*, and *γ*, subunits of the olfactory G protein, guanosine-5′-triphosphate (GTP), guanosine-5′-diphosphate (GDP), adenylyl cyclase 3 (AC3), activated phosphodiesterase 1C (PDE1C), cyclic adenosine monophosphate (cAMP), adenosine monophosphate (AMP), Ca^{2+}/calmodulin-dependent protein kinase II (CaMKII), cyclic nucleotide-gated channel (CNG); Ca^{2+}-activated Cl^-channel anoctamin 2 (ANO2), K^+-dependent Na^+/Ca^{2+} exchanger 4 (NCKX4). Images created with BioRender.com.

In OSNs, the inactivation by phosphorylation and arrestin are potentially not needed for the timely shutoff of the olfactory transduction cascade, due to the extremely short lifetime of the active ligand-bound receptor molecule. Early biochemical experiments suggested that OR phosphorylation does control cAMP kinetics (Dawson et al., 1993; Schleicher et al., 1993; Peppel et al., 1997), but it seems to play little, if any, role in the control of odorant-response kinetics for one particular OR, mOR-EG (Kato et al., 2014). It remains to be established whether this applies to all ORs, or whether a subset of ORs is subject to phosphorylation and inactivation. β-arrestin interacts with ORs, mediating internalization during prolonged stimulation and altering adaptation to repetitive odor stimuli (Mashukova et al., 2006). Experiments on isolated human and rat OSNs suggested a role for protein kinases A (PKA) and C (PKC) in the termination of the olfactory response. Ca^{2+} imaging showed that the inhibition of PKA and PKC increases intracellular Ca^{2+} responses in the presence of odorant mixtures, and blocks their termination after odorant stimulation ceases. While the inhibition of both PKA and PKC modulated the odor-induced intracellular Ca^{2+} increase in the human OSNs, only PKC and not PKA affected the Ca^{2+} response to odorants in rat OSNs, suggesting differences among species in the termination of the olfactory response (Gomez et al., 2000).

The control of the lifetime of the olfactory G protein seems to be more complex and less well-understood compared to phototransduction. Ric-8B (resistant to inhibitors of cholinesterase-8B) has been identified as a GTP exchange factor (GEF) expressed in OSNs, which facilitates the exchange of GDP for GTP on $G\alpha_{olf}$ and its activation. Unusually, Ric-8B not only interacts with the G protein α-subunit but also with γ13, the olfactory γ-subunit. In a heterologous system, Ric-8B co-expression with olfactory transduction components can greatly increase cAMP production, suggesting that it could indeed modulate olfactory transduction (Von Dannecker et al., 2005; Kerr et al., 2008). A knockout mouse for Ric-8B displays impaired olfactory behavior, and, surprisingly, greatly reduced odorant responses. Ric-8B is localized primarily in the cell body and the dendritic knob of OSNs. Ric-8B knockout OSNs are devoid of $G\alpha_{olf}$ (Machado et al., 2017), suggesting that this gene is needed for the stable expression of $G\alpha_{olf}$, and excludes addressing its potential role as a GEF in the odorant response. The Ric-8B knockout mice also display higher OSN cell death. Regulators of G protein signaling (RGS) are GAPs that modulate the lifetime of an activated G protein as described above. RGS2, instead of functioning as a GAP, directly inhibits AC3 to control the size of the odorant response (Sinnarajah et al., 2001). However, inconsistent and contradictory data on RGS2 and RGS3 expression and their roles in OSNs suggest that more research is needed (Norlin and Berghard, 2001; Kanageswaran et al., 2015; Saraiva et al., 2015).

28.17.10: Adaptation

Adaptation plays a critical role in the capacity of our sensory neurons to remain able to detect stimuli above the background in a complex and rapidly changing environment. For instance, in constant light conditions, the dynamic range for both rods and cones is only 100-fold, spanning a range from threshold stimulation to saturation (Figure 3C). However, as a result of light adaptation, photoreceptors can shift their functional range over a very wide span of light conditions, ranging from cloudy night to sunrise for rods, and starry night to bright sunny day for cones (Weale, 1961). Thus, using the adaptation of individual photoreceptors, the visual system can remain responsive to stimuli over a wide range of light conditions. In contrast, the ability of OSNs to adapt is rather limited even at modest levels of background odorant (Reisert and Matthews, 1999). Nevertheless, increasing concentrations of the same odorant can recruit less sensitive ORs, and therefore less sensitive OSNs, preserving its perception at higher concentrations and ensuring to report the presence of that odorant to the brain.

In both types of sensory neurons, adaptation is mediated by a change in Ca^{2+} upon stimulation. This change is sensed by several Ca^{2+}-binding proteins that trigger a negative feedback on the vision and olfaction transduction cascades by modulating several of their steps. In the outer segments of rods and cones and olfactory cilia, Ca^{2+} levels are controlled by the balance of influx *via* the CNG channels, whose current is carried in part by Ca^{2+}, and efflux *via* Na^+/Ca^{2+}, K^+ exchangers (NCKXs) that use the electrochemical gradient for Na^+ and K^+ to extrude Ca^{2+} (Figure 5; Yau and Nakatani, 1984). In rods (Figure 5A), this task is accomplished by rod-specific NCKX1, whereas cones employ two separate exchangers, NCKX2 and NCKX4 (Vinberg et al., 2017). At rest, both in darkness and in steady-state light, the influx of Ca^{2+} is matched with its extrusion and, as a result, the level of free Ca^{2+} in the outer segments is maintained constant. Upon photostimulation, the transduction cascade is activated, resulting in the depletion of cGMP, closure of CNG channels, and reduction in the influx of Ca^{2+} into the outer segments. However, Ca^{2+} extrusion by the Na^+/Ca^{2+}, and K^+ exchangers carry on for at least a while and, as a result, the level of Ca^{2+} in the outer segments declines. Direct Ca^{2+} measurements in amphibian photoreceptors indicate a dynamic range from 670 to 30 nM in rods (Sampath et al., 1998) and 400–5 nM in cones (Sampath et al., 1999), in darkness and bright light, respectively.

The light-driven decline in Ca^{2+} causes its release from several Ca^{2+}-binding proteins. The dominant Ca^{2+}-dependent feedback mechanism in both rods and cones controls the synthesis of cGMP by membrane-bound GC *via* a pair of GC-activating proteins (GCAPs)—GCAP1 and GCAP2. When Ca^{2+} in the outer segments is high, Ca^{2+}-bound GCAPs bind to and partially inhibit the activity of GC. Upon photoactivation and the decline in Ca^{2+}, GCAPs become Ca^{2+}-free and release from GC, resulting in the upregulation of cGMP synthesis which restores the dark current after photostimulation and modulates the activation of the transduction cascade in the presence of background light (Dizhoor, 2000; Sakurai et al., 2011). Another mechanism by which Ca^{2+} modulates phototransduction involves the Ca^{2+}-binding protein recoverin. As GCAPs, recoverin is a member of the EF-hands protein family, and when bound to Ca^{2+} in darkness, it inhibits rhodopsin kinase, thus slowing down the inactivation of the visual pigment (Makino et al., 2004; Sakurai et al., 2015). When the photoreceptors are activated and Ca^{2+} declines, it is released from recoverin, which in turn dissociates from rhodopsin kinase and relieves its inhibition. This enhances the phosphorylation of visual pigments and accelerates their inactivation, effectively reducing the activation of the transduction cascade by the background light. Finally, direct modulation of the CNG channels has also been suggested. However, in the case of rods, such modulation appears to play a marginal, at best, role (Koutalos and Yau, 1996) and is not mediated by the Ca^{2+}-binding protein calmodulin (Chen et al., 2010). In zebrafish cones, the modulation of the CNG channels appears to play a more substantial role and is mediated by the Ca^{2+}-binding protein CNG-modulin (Korenbrot et al., 2013). It is still unclear whether the mammalian homolog of CNG modulin, EML1 plays a similar role in mammalian cones.

Adaptation in OSNs is less well understood compared to phototransduction. Early data, mostly of biochemical nature or obtained from heterologously-expressed proteins of interest, suggested three main molecular targets for adaptation. All three of them are mediated by the Ca^{2+} influx during the odorant response: Ca^{2+}/calmodulin-mediated desensitization of the olfactory CNG channel to close the channel even in the presence of high cAMP (Chen and Yau, 1994); phosphorylation *via* CaM-kinase 2 of AC3 to reduce the rate of cAMP production (Wei et al., 1996, 1998); and Ca^{2+}-mediated upregulation of phosphodiesterase 1C, which is expressed in olfactory cilia, and is assumed to degrade cAMP to AMP to terminate the response (Borisy et al., 1992). Follow-up experiments using recordings from OSNs all seem to indicate that none of these mechanisms plays as prominently or as originally thought role in transduction (Reisert and Zhao, 2011). A mouse with a mutation in the CNGB1b channel subunit that entirely prevents desensitization by Ca^{2+} surprisingly displays normal olfactory adaptation but instead shows a delayed response termination, suggesting that Ca^{2+}/calmodulin-mediated desensitization of the CNG channel speeds up response termination (Song et al., 2008). A mouse model that carries a mutation in AC3 that prevents phosphorylation does not show a discernable phenotype of the olfactory response (Cygnar et al., 2012), although it might be possible that other, unknown phosphorylation sites in AC3 might be important. Finally, a knockout mouse for PCE1C has no deficits in response termination but instead shows much-reduced

response amplitudes for unclear reasons (Cygnar and Zhao, 2009). This begs the obvious question as to what the role of PDE1C might be and what might actual to cAMP that is generated during the odorant response. For the latter, an interesting option is that cAMP diffuses out of the cilia into the cell body as a means to reduce ciliary cAMP, allowing OSNs to recover from stimulation (Cygnar and Zhao, 2009). One aspect that is reasonably understood is NCKX4, the Ca^{2+} exchanger in OSNs that is required to lower intraciliary Ca^{2+} during and after odorant stimulation, allowing the transduction cascade to recover from adaptation (Reisert and Matthews, 1998; Stephan et al., 2012).

28.17.11: Diseases

Disorders affecting photoreceptors are among the leading causes of blindness in the human population. One of the prevalent visual disorders, called retinitis pigmentosa, is a complex disease caused by a wide range of mutations in photoreceptors. Many of these mutations affect the expression, structure, and function of the rod's visual pigment (Athanasiou et al., 2018). Because of the very high expression of opsin in the outer segments of rods, this protein plays not only a functional role but is also critical for the proper formation of the outer segment itself. As a result, mutations affecting the expression, folding, or targeting of opsin to the rod outer segments, cause gradual degeneration of the rods. Other genes implicated in rod dysfunction and degeneration include those for phosphodiesterase (e.g., rd1, rd10; Chang et al., 2002), the CNG channels A and B subunits (channelopathies; Michalakis et al., 2018), GC, and GCAPs (Olshevskaya et al., 2002). Another diverse set of visual disorders is caused by abnormal chromophore production or supply to photoreceptors, which limits the ability to detect light and can also lead to degeneration (Ku and Pennesi, 2020). Notably, the efficiency of the visual system to produce chromophore seems to decline with age, which may result in poor rod function in dim light even in normally aging adults. It is also an early indicator of age-related macular degeneration, a devastating blinding disorder that affects the function of cones in the central retina responsible for acute vision and color discrimination (Jackson et al., 2002). Interestingly, rods and cones seem to coexist synergistically in the retina, and diseases caused by rod-specific mutations that result in rod degeneration, eventually lead to the loss of cones and central vision as well. Thus, considerable efforts are currently focused on developing methods for preserving rods even when they are not functional, as a way of protecting daytime cone-driven vision. Because the eye is a relatively accessible organ, novel therapeutic approaches for vision protection and restoration have led the field, with successful examples of gene therapy and stem cell therapy in experimental and clinical trial phases (Ovando-Roche et al., 2017).

Compared to vision, in olfactory transduction, very few mutations in transduction components are known that lead to deleterious effects. Several aspects might account for this. Mutations causing a partial reduction of olfaction might go unnoticed in the human population as very little systematic olfactory testing is done. OSNs regenerate throughout life and only have a lifespan of a few weeks. Hence any slow degeneration as those seen in photoreceptors might not manifest in that time window. In a screen of families with congenital anosmia, no potentially causative mutations were found in three main transduction proteins ($G_{\alpha olf}$, CNGA2, AC3), with these genes also being under purifying selection (Feldmesser et al., 2007). An interesting exception is patients suffering from retinitis pigmentosa, which is caused by mutations in the gene encoding the CNGB1 subunit expressed in both rods and OSNs. Those patients, identified because of their visual function decline, were found to be hyposmic or anosmic when tested for their olfactory ability (Charbel Issa et al., 2018). If congenital anosmia is considered to be a relatively rare and little-understood condition, more known and frequently detected are specific anosmias, which manifest in the inability to detect certain odorants (Keller et al., 2007; Trimmer et al., 2019). Broadly speaking, this is the olfactory equivalent of color blindness and is caused by known OR mutations.

Arguably, the most common causes of smell loss are events that lead to the destruction of the olfactory epithelium and/or the olfactory nerves connecting it to the central nervous system (CNS). These events include head or face trauma, inhalation of toxic chemicals, or viral infection (such as SARS-CoV2), and, neurodegenerative diseases such as Alzheimer's and Parkinson's disease (Attems et al., 2015; Cooper et al., 2020). In the former, the origin of the smell disorder can be tracked down to the periphery, the olfactory epithelium. In the case of neurodegenerative diseases, it has been thought that olfactory dysfunction originates centrally in the CNS, but it is becoming clearer now that peripheral olfaction can be affected in these cases as well, although the respective mechanisms have not been fully elucidated.

28.18: Signal Transduction - Taste (Gustation)

28.18.1: G Protein-Coupled Receptors in Taste Physiology and Pharmacology

Heterotrimeric G protein-coupled receptors (GPCRs) comprise the largest receptor family in mammals and are responsible for the regulation of most physiological functions. Besides mediating the sensory modalities of olfaction and vision, GPCRs also transduce signals for three basic taste qualities of sweet, umami (savory taste), and bitter, as well as the flavor sensation kokumi. Taste GPCRs reside in specialized taste receptor cells (TRCs) within taste buds. Type I taste GPCRs (TAS1R) form heterodimeric complexes that function as sweet (TAS1R2/TAS1R3) or umami (TAS1R1/TAS1R3) taste receptors, whereas Type II are monomeric bitter taste receptors or kokumi/calcium-sensing receptors. Sweet, umami, and kokumi receptors share structural similarities in containing multiple agonist binding sites with pronounced selectivity while most bitter receptors contain a single binding site that is broadly tuned to a diverse array of bitter ligands in a non-selective manner. Tastant binding to the receptor activates downstream secondary messenger pathways leading to depolarization and increased intracellular calcium in TRCs, that in turn innervates the gustatory cortex in the brain. Despite recent advances in our understanding of the relationship between agonist binding and the conformational changes required for receptor activation, several major challenges and questions remain in taste GPCR biology that is discussed in the present review. In recent years, intensive integrative approaches combining heterologous expression, mutagenesis, and homology modeling have together provided insight regarding agonist binding site locations and molecular mechanisms of orthosteric and allosteric modulation. In addition, studies based on transgenic mice, utilizing either global or conditional knock-out strategies have provided insights to taste receptor signal transduction mechanisms and their roles in physiology. However, the need for more functional studies in a physiological context is apparent and would be enhanced by a crystallized structure of taste receptors for a more complete picture of their pharmacological mechanisms.

28.18.2: Introduction

G protein-coupled receptors (GPCRs) are the largest and the most diverse group of membrane receptors in eukaryotes. They are activated by a wide variety of ligands in the form of light energy, lipids, sugars, peptides, and proteins (Billington and Penn, 2003; Schoneberg et al., 2004; Lundstrom, 2009) which convey information from the outside environment into the cell to mediate their corresponding functional responses. The conformational changes of GPCRs upon ligand binding initiate a series of biochemical reactions within the cell. These intracellular reactions regulate sensory functions of smell, taste, and vision, and a wide variety of physiological processes such as secretion, neurotransmission, metabolism, cellular differentiation, inflammation, and immune responses (Lagerström and Schiöth, 2008; Rosenbaum et al., 2009; Venkatakrishnan et al., 2013; Ahmad et al., 2015). Taste is one of the most important sensations in human life, enabling us to perceive different tastes from the diverse range of food available in nature, and is a major determinant of our ingestion decisions.

The anatomical units of taste detection are taste receptor cells (TRCs) that are assembled into taste buds distributed across different papillae of the tongue and palate epithelium. Taste processing is first achieved at the level of TRCs that are activated by specific tastants. They transmit information *via* sensory afferent fibers to the gustatory cortex in the brain for taste perception (Figure 1). Three different morphologic subtypes of TRCs in taste buds sense the different tastes we perceive. Type I glial-like cells detect salty taste while type II cells expressing GPCRs detect sweet, umami, and bitter tastes. Type III cells sense sour stimuli (Janssen and Depoortere, 2013).

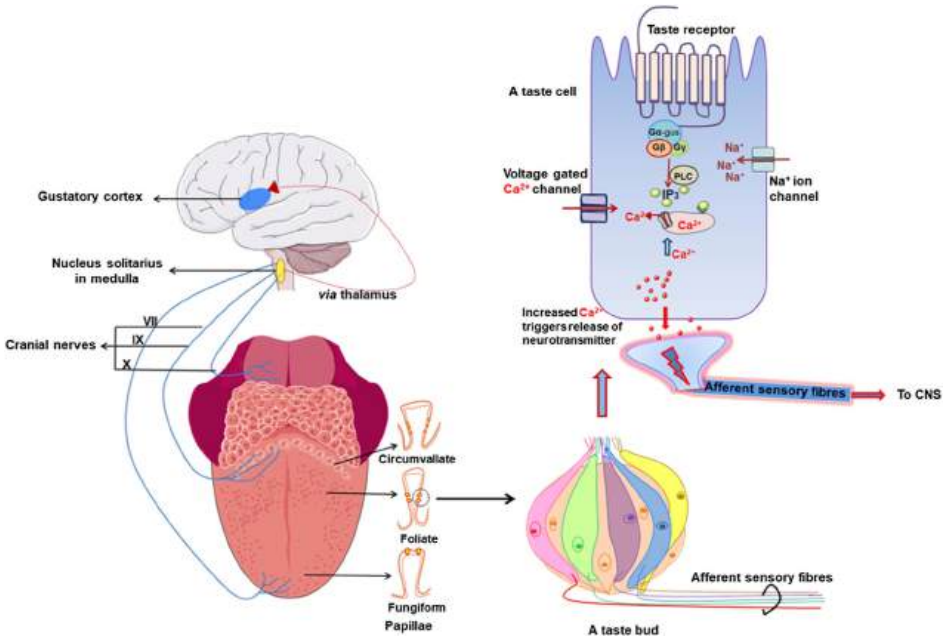
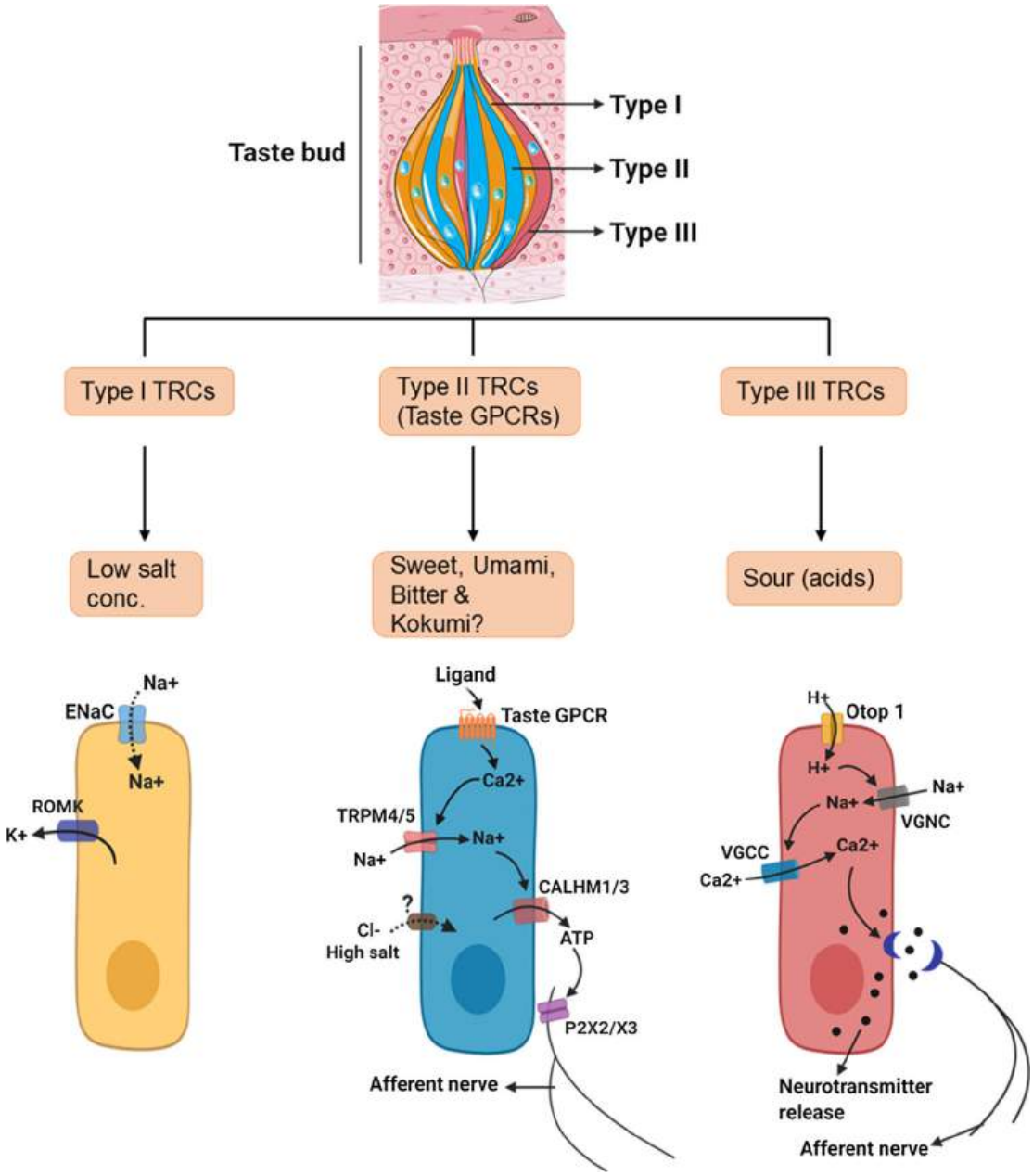


FIGURE 1. A schematic diagram shows taste signal transmission between the tongue and brain. Taste buds present in different papillae in the tongue and palate contain taste receptor cells (TRC) which contain taste G protein-coupled receptors (GPCRs). The left side shows how afferent nerves transmit a signal to the gustatory cortex in the brain via cranial/glossopharyngeal nerves. The right side shows taste buds with taste TRCs and a simplified signal transduction pathway of taste receptors where taste GPCRs are activated by a tastant that in turn recruits a specific G protein that further induces intracellular calcium release (created with BioRender.com).

Sweet and umami stimuli are transduced by Type 1 taste GPCRs while the bitter taste is sensed by Type 2 taste GPCRs (Figure 2; Table 1). The more recently described kokumi sensation is mediated by another GPCR, the calcium-sensing receptor (CaSR) (Figure 2; Table 1). Taste GPCRs are activated by specific taste ligands present in foods and recruit G proteins to activate downstream signaling effectors (Figure 3).



Figures 1, 3) (created with BioRender.com).

TABLE 1 | Taste GPCRs classification and their downstream signaling regulators.

Taste/flavor sensation	Sweet	Umami	Bitter	Kokumi
GPCRs	TAS1R2 + TAS1R3	TAS1R1 + TAS1R3	TAS2Rs	Calcium sensing receptor
Class of GPCR	Class C	Class C	Class A	Class C
Type	Type 1 taste GPCRs	Type 1 taste GPCRs	Type 2 taste GPCRs	ND
Oligomerization	Heterodimer	Heterodimer	Monomer	Homodimer
GαSubunit	Gα-gustducin	Gα-gustducin	Gα-gustducin	Gq/11, Gi/o, G12/13, Gs
Signaling pathway	Gβγ-PLCβ2-TRPM5 pathway-semaphorin 7A	Gβγ-PLCβ2-TRPM5 pathway	Gβγ-PLCβ2-TRPM5 pathway-semaphorin 3A	Gβγ-PLCβ2-TRPM5 pathway
Ligand binding sites	VFT; TMD	VFT, TMD	ND	VFT, TMD
Potent ligands	Sucrose, aspartame, sucralose, cyclamate	L-amino acids, Gultamate	Peptides, alkaloids, flavonoids	Ca^{2+}, divalent and trivalent cations, Glutathione, *γ*-glutamyl peptides, polyamines, amino acids
Antagonists/NAMs	Lactisole, gymnemic acid	Lactisole, clofibric acid	GIV3727, probenecid	NPS2143
Food source	Sugar/sweets, dairy products	Savoury food, fermented dairy/meat/fish/chicken broth, mushroom, tomato	Wine, tea, coffee, cheese, broccoli	Garlic/onion meat/fish/broth/mushroom, soy sauce

TAS2R, bitter taste receptors; GPCR, G protein-coupled receptors; VFT, extracellular Venus flytrap domain; TMD, seven transmembrane domain; ND, not defined.

TABLE 1. Taste GPCRs classification and their downstream signaling regulators.

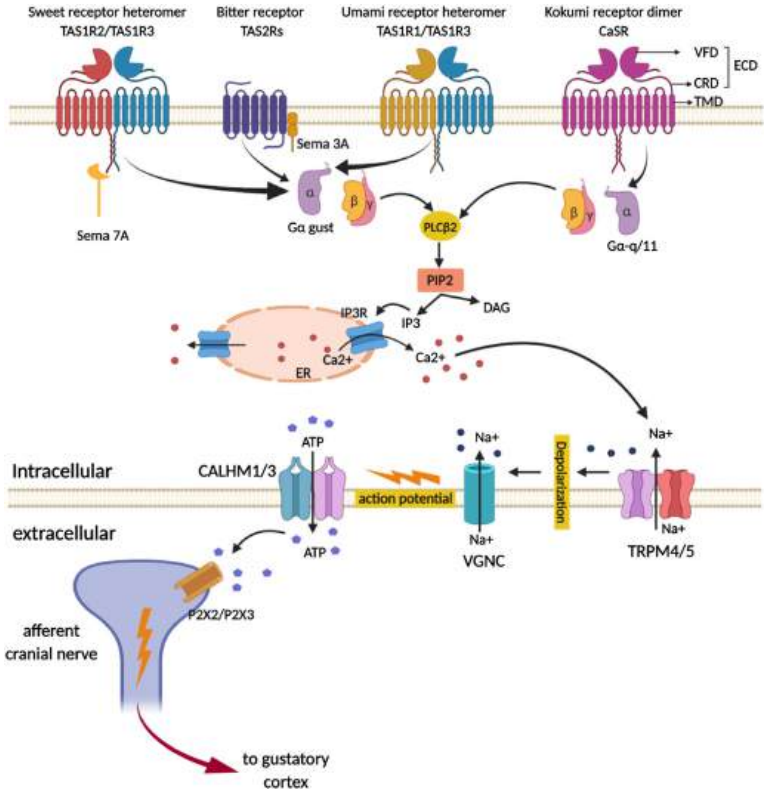


FIGURE 3. Schematic representation of signal transduction pathway of sweet, umami, bitter, and kokumi-calcium sensing receptors (CaSR) in taste receptor cells on the tongue. Ligand-induced stimulation of the sweet (TAS1R2/TAS1R3), umami (TAS1R1/TAS1R3), bitter receptors (TAS2Rs) and kokumi sensation expressed in type II taste cells within taste bud activate a trimeric G protein composed of α-gustducin (Gα-gust) in sweet, umami, bitter and Gα-q/11 in kokumi-receptor and a complex consisting of Gβγ proteins. The released Gβγ-complex activates phospholipase C isoform β2 (PLCβ2) which then induces the production of inositol 1,4,5-trisphosphate (IP3) and diacylglycerol (DAG); the second messenger IP3, in turn, activates the IP3 receptor (IP3R), an intracellular ion channel that allows Ca^{2+} release from the intracellular endoplasmic reticulum (ER store). An increase in intracellular Ca^{2+} then activates the complex of transient receptor potential cation channel subfamily M member 4 and 5 (TRPM4/5) that are plasma membrane-localized sodium-selective channels which leads to depolarization and subsequent activation of voltage-gated sodium channels (VGSC). The combined action of increased Ca^{2+} and membrane depolarization activates the complex of calcium homeostasis modulator 1 and 3(CALHM1/3) channel and pannexin1 channels, thus resulting in the release of

the neurotransmitter ATP. Increased ATP, in turn, activates P2X ionotropic purinergic receptors 2 and 3 (P2X2/P2X3) on the afferent cranial nerve generating an action potential that subsequently signals to the gustatory cortex for sensory perception. Besides well-known taste GPCR pathways, connecting proteins semaphorin 7A (Sem 7A) and 3A (Sem 3A) are depicted in close contact with sweet and bitter receptors as they provide instructive signals that fine-tune to sweet or bitter ganglion neurons, respectively. VFT, venus flytrap domain; CRD, cystine-rich domain; ECD, extracellular domain. (created with BioRender.com).

In this review, we will first explore the basic architecture of the gustatory sensory system and its peripheral signal transmission. Then we will discuss taste GPCR signal transduction mechanisms for the different taste modalities, their molecular structure, and the conformational changes that occur following orthosteric/allosteric binding of endogenous and food-derived ligands.

28.18.3: Taste Buds and Neural Transmission

In mammals, taste buds on the tongue comprise 50–100 elongated epithelial cells and a small number of proliferative basal cells (Sullivan et al., 2010). Ultrastructural studies and patterns of gene expression with cell function reveal three distinct anatomical types of TRCs within each taste bud: Type I, Type II and Type III (Murray, 1986) (refer to Figure 2; Table 2).

TABLE 2 | Summary of taste receptor cell characteristics.

Taste receptor cells	Type I	Type II	Type III
Cell type	50% of the total population	20–30%	5–10%
Taste responses	Low salt taste amiloride sensitive response	Sweet, biiter, umami and probably kokumi flavor response (?) via GPCRs	Sour response via otopterin 1
Morphology	Spindle shaped with long brush like microvilli (1–2 mm), no synapse	Long and slender short microvilli, no synapse	Single large microvillus, have synaptic contact with afferent nerve
Other functions	Support function, ion redistribution, neurotransmitter clearance	High salt taste amiloride insensitive response?	Probable salt response?
Marker proteins	GLAST, K^+ channel (ROMK)	CALHM1, CALHM3, pannexin, connexins	Kir2.1, PDK2L1, SNAP25, synapsinII, NCAM

GLAST, glutamate aspartate transporter; ROMK, renal outer medullary potassium channel; PDK2L1, polycystin 2 like 1, transient receptor potential cation channel; CALHM, calcium homeostasis modulator; Kir2.1, inward rectifier K(+) channel; NCAM, neural cell adhesion molecule; SNAP25, synaptosome associated protein 25 kDa.

TABLE 2. Summary of taste receptor cell characteristics.

Type II TRCs express either sweet, umami, or bitter taste receptors at their cell surface. These receptors share some commonality with their signal transduction mechanisms that are intrinsic to TRCs. Taste GPCRs (sweet, umami, and bitter) couple to heterotrimeric G proteins that include Gα-gustducin, Gβ3, and Gγ13 (Huang et al., 1999) and initiate a series of signal transduction cascades involving activation of phospholipase C-β2 (PLCB2), production of inositol-1,4,5-triophosphate (IP3), and IP3-dependent Ca^{2+} release from the endoplasmic reticulum (ER) via the IP3 receptor (IP3R). The increased intracellular $[Ca^{2+}]_i$ then activates the transient receptor potential cation channel subfamily M member 4 and 5 (TRPM4/5) in the basolateral plasma membrane, leading to membrane depolarization that triggers Na^+ action potential firing, and depolarization-induced release of ATP. In turn, ATP acts as the primary neurotransmitter stimulating purinergic receptors 2 and 3 (P2X2 and P2X3) on afferent cranial nerves whose activation triggers an action potential that subsequently activates the gustatory cortex in the brain (McLaughlin et al., 1992; Wong et al., 1996; Margolskee, 2002). α-gustducin is a distinct G protein selectively expressed in ∼30% of type II TRCs and shares 80% identity with retinal protein α-transducin (McLaughlin et al., 1992) and is a key contributor to signal transduction for sweet and bitter taste receptors (McLaughlin et al., 1992; Wong et al., 1996; Margolskee, 2002).

An important aspect of taste transduction is how ATP signaling is conducted. Recent studies have discovered that calcium homeostasis modulators 1 and 3 (CALHM1/3) are enriched in type II TRCs where they interact and form a functional complex. Their genetic deletion abolishes responses to sweet, bitter, and umami tastes, supporting the requirement of the CALHM1/3 complex as an ATP release channel for the GPCRs mediated tastes (Taruno et al., 2013; Ma et al., 2018).

New information has provided insight into how specific taste qualities are fine-tuned to recognize their partner ganglionic neurons in the brain. Lee et al. (2017) discovered semaphorin proteins, 7A and 3A as the physical links between sweet and bitter TRCs, respectively, and their partner ganglion neurons in the brain. It remains to be determined what physical links exist between umami TRCs and their corresponding neurons in the brain. Delineating the underlying molecular basis for this interaction would provide a further understanding of purinergic transmission in the taste system. In addition, whether these mechanisms are relevant for kokumi sensation has not yet been investigated, despite CaSR having distinct expression in TRCs and significant functional synergy with other prominent taste qualities (sweet, umami, and salty). Moreover, there is still debate regarding the recognition of kokumi as a

sixth taste entity, consequently, the calcium-sensing receptor (CaSR) is not yet included in the nomenclature for any subtypes of taste GPCRs, although it would best fit with Type 1 taste receptors.

28.18.4: Type 1 Taste G Protein-Coupled Receptors (Sweet and Umami)

The type 1 taste receptors (TAS1Rs) belong to the class C GPCRs, which possess a large N-terminal extracellular domain (ECD) fused to the heptahelical seven transmembrane domain (TMD). The ECD is further divided into two ligand-binding domains (LBD1 and LBD2), having a bi-lobed structure called a Venus flytrap domain (VFT) due to its resemblance to this shape (Hoon et al., 1999). Except for $GABA_B$ receptors, a cysteine-rich domain (CRD) connects the VFT to the TMD (Leach and Gregory, 2017).

In contrast to other receptors from this class C of GPCRs, such as the metabotropic glutamate receptor (mGluR) or γ-aminobutyric acid type B receptors (GABABRs) which function as homo- or heterodimers, respectively (Jones et al., 1998; Kaupmann et al., 1998; White et al., 1998; Kunishima et al., 2000), the TAS1Rs function as obligatory heterodimers. The distinct expression pattern of TAS1R1 and TAS1R2 in different subsets of murine cells led to the idea that they could detect two different taste profiles. However, following the discovery of the TAS1R3 subtype, it was clear that when TAS1R1 heterodimerizes with TAS1R2, the receptor detects sweet taste substances (Nelson et al., 2001; Ohkuri et al., 2009; Kim et al., 2017). On the other hand, if heterodimerized with TAS1R3 (TAS1R1/TAS1R3), it is responsible for umami or amino acid taste detection (Li et al., 2002; Nelson et al., 2002). Please refer to figure 4A for the basic structure of sweet and umami receptors.

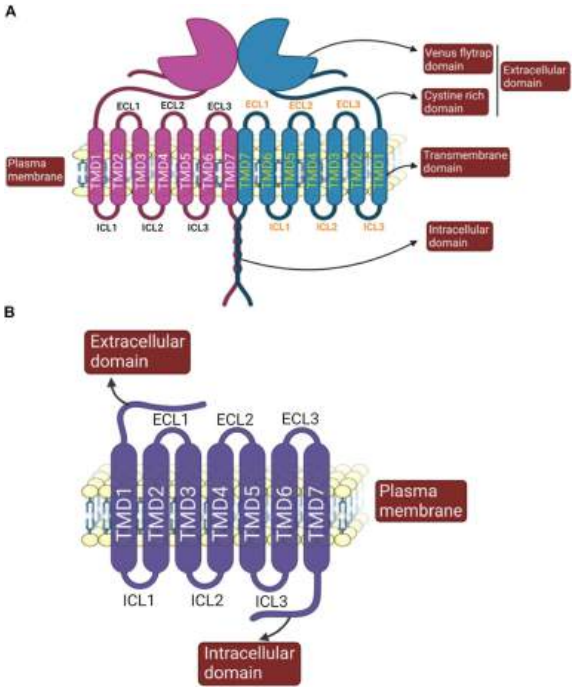


Figure depicting basic structural features of sweet/umami/kokumi dimeric receptor **(A)** and monomeric bitter receptor **(B)** (created with BioRender.com).

Figure 28.18.x shows an interactive iCn3D model of Taste receptor type 1 member 2 (TAS1R2) AlphaFold model (uniprot Q8TE23)

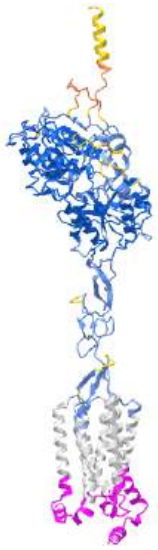


Figure 28.18.x: Taste receptor type 1 member 2 (TAS1R2) AlphaFold model (uniprot Q8TE23). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...mkhYb2uKoR68b9

The gray is the predicted transmembrane helices. The cyan is the intracellular domain. The blue is the extracellular domain. The predicted model of the structure has high confidence except for the yellow/orange at the distal end of the extracellular domain. Key residues in the ligand binding domain are shown as sticks, CPK colors, and labeled.

28.18.5: Sweet Taste Signal Transduction Mechanisms

The TAS1R2/TAS1R3 receptor recognizes a wide variety of sweet substances including natural sugars, artificial sweeteners, amino acids, and proteins (Li et al., 2002; Xu et al., 2004; Jiang et al., 2005a; Jiang et al., 2005c) (Table 3). This was demonstrated in studies using heterologous expression systems as well as knockout mice for TAS1R2 and/or TAS1R3 subtypes that showed a blunted response to sugars, sweeteners, and D-amino acids, confirming the TAS1R2/TAS1R3 heterodimer as the main sweet taste receptor *in vivo* (Li et al., 2002; Zhao et al., 2003; Xu et al., 2004).

TABLE 3 | Agonists of sweet taste receptors along with their EC50 values.

Agonists	Nature	Binding pocket	EC50 (mM)	References
Sucrose	Natural carbohydrate	VFT (TAS1R2 and TAS1R3)	62	(Li et al., 2002; Servant et al., 2010; Zhang et al., 2010; Zhang
				et al., 2003)
Aspertame	Peptide	VFT (TAS1R2)	0.75	(Li et al., 2002; Liu et al., 2011; Masuda et al., 2012)
Neotame	Peptide	VFT (TAS1R2)	5	(Li et al., 2002; Masuda et al., 2012)
Cyclamate	Sulfamate	TMD (TAS1R3)	3.1	(Xu et al., 2004; Jiang et al., 2005c)
Brazzein	Protein	CRD (TAS1R3)	0.08	(Li et al., 2002; Jiang et al., 2004; Ide, et al., 2009; Masuda et al.,
				2012)
Thaumatin	Protein	CRD (TAS1R3)	0.005	Masuda et al., 2012; Jiang et al., 2004
Monellin	Protein	VFT (TAS1R3), VFT (TAS1R2)	0.01	Koizumi et al., 2007; Jiang et al., 2004
Neoculin	Protein	VFT (TAS1R2)	0.001	(Jiang et al., 2004; Koizumi et al., 2007)

Agonists	Nature	Binding pocket	EC50 (mM)	References
Saccharin	N sulfonyl amide	VFT (TAS1R2)	0.19	(Li et al., 2002; Masuda et al., 2012; DuBois, 2016)
Suosan, cyanosuasan	Arylurea	VFT (TAS1R2)	ND	(Tinti and Nofre, 1991; Du Bois, 2016)
SC-45647	Guanidinoacetic acid	VFT (TAS1R2)	0.3	(DuBois, 1995; Sanematsu et al., 2014)
Sucralose	Halogenated carbohydrate	VFT (TAS1R2 and TAS1R3)	0.06	(Li et al., 2002; Masuda et al., 2012)
Acesulfame K	Sulfamate ester	VFT (TAS1R2)	0.54	(Li et al., 2002; Masuda et al., 2012)
Perillartine	Oxime, ethoxyphenyl urea, alkoxyaryl urea,	TMD (TAS1R2)	15	(Li et al., 2002; Servant et al., 2010)
Dulcin	Ethoxyphenyl urea	TMD (TAS1R2)	0.01	(Servant et al., 2010)
S819	Alkoxyaryl urea	TMD (TAS1R2)	0.025	(Zhang et al., 2008)
D-tryptophan	Amino acid	VFT (TAS1R2)	2.09	(Li et al., 2002; Masuda et al., 2012)
Xylitol, sorbitol	Polyols	VFT (TAS1R2)	ND	(Mahalapbutr et al., 2019)
Maltotriose, acarbose	Oligosaccharide, pseudotetrasaccharide	ND	ND	(Pullicin et al., 2017; Pullicin et al., 2019)

Where VFT, venus flytrap domain; TMD, transmembrane domain; ND, not determined.

TABLE 3. Agonists of sweet taste receptors along with their EC50 values.

The sweet receptor couples to heterotrimeric Gα-gustducin which include Gβ3 and Gγ13, as mice lacking Gα-gustducin, showed a reduced response to sweet substances either natural or artificial (McLaughlin et al., 1992; Wong et al., 1996; Margolskee, 2002). Moreover, a point mutation in the C-terminal region of gustducin (G352P) (critical for its receptor interaction) results in the loss of its ability to activate taste GPCRs while keeping other functions intact. Further, G352P acts as a dominant negative to block heterotrimeric G protein interaction with taste receptors and disrupts the responses to sweet and bitter compounds in both wild-type (WT) and null mice (Ruiz-Avila et al., 2001). In addition, the G352 mutant further reduces any residual sweet/bitter taste responses of the null mice by acting as a "βγ sink" to bind all unbound βγ-subunits and remove them from the viable pool of G protein heterotrimers available to the receptor (Ruiz-Avila et al., 2001). These observations confirm the essential requirement of Gα-gustducin in sweet and bitter taste transduction.

In addition to the Gα-gustducin pathway, sweet taste transduction occurs *via* two additional signaling pathways involving different secondary messengers. The first one involves cAMP and the second one involves IP3. Normally, sugars elevate the level of cAMP, while sweeteners stimulate IP3 production (Tonosaki and Funakoshi, 1988; Uchida and Sato, 1997). Sucrose or other sugars bind to either TAS1R2 or TAS1R3 and recruit Gαs protein that leads to increased cAMP levels which initiate the influx of cations through ion channels. Alternatively, cAMP activates protein kinase A that leads to TRC cell depolarization resulting in an influx of calcium ions and neurotransmitter release (Avenet et al., 1988; Tonosaki and Funakoshi, 1988; Margolskee, 2002). Sweetener binding to the TAS1R2/TAS1R3 heterodimer recruits Gα-gustducin proteins that stimulate PLCβ2 which in turn hydrolyzes phosphatidylinositol 4,5-bisphosphate (PIP2) to diacylglycerol (DAG) and IP3 (Margolskee, 2002; Chandrashekar et al., 2006). IP3R3 (Hisatsune et al., 2007) induced Ca^{2+} release from ER stores (Figure 3) activates TRPM5 (Zhao et al., 2003; Hisatsune et al., 2007; Dutta Banik et al., 2018) that leads to an action potential (Yoshida et al., 2005; Yoshida et al., 2006) and subsequent release of neurotransmitters.

Interestingly, Dutta Banik et al. (2018) confirmed that TRPM4 also mediates taste signaling independent of TRPM5, and knocking out both channel proteins (TRPM4/5) abolishes the sweet, umami, and bitter taste response completely. This revealed another layer of complexity to sweet signal transmission. This in-depth mechanistic research has increased our understanding of sweet and bitter receptors and presents a challenge to dissect the taste signal transmission pathways for umami and kokumi as well.

28.18.5.1: Structural, Molecular, and Conformational Changes of Sweet Receptor

Since the sweet taste receptor has not yet been crystallized, determining the structure of the sweetener binding site and mechanism of activation has been a challenge. Based on homology with other class C GPCRs (mGluRs and $GABA_BRs$), multiple studies propose similar activation mechanisms for the sweet receptor (Kunishima et al., 2000; Tsuchiya et al., 2002; Jingami et al., 2003; Muto et al., 2007; Perez-Aguilar et al., 2019). The many different sweet agonists and their diverse binding sites across receptor domains (VFT, TMD, and CRD) (Table 3) may explain its complex yet broadly tuned nature. For example, a single residue in VFT (I60) of TAS1R3 of the TAS1R2/TAS1R3 heteromer is required for a saccharin preference in inbred mouse strains (Max et al., 2001; Reed et al., 2004).

Several studies utilizing homology and computational modeling based on the crystal structure of mGluR and $GABA_BRs$ have predicted structural and functional aspects of orthosteric and allosteric binding sites for the sweet receptor (Kim et al., 2017; Cheron et al., 2019; Park et al., 2019). They reported that both VFT regions undergo ligand-dependent conformational changes and intersubunit interactions between ECDs that further stabilize heterodimer formation for subsequent downstream signaling (Perez-Aguilar et al., 2019). The binding of orthosteric agonists to VFT of TAS1R2 leads to major conformational changes that form a TMD6/TMD6 interface between TMDs of TAS1R2 and TAS1R3, which is consistent with the activation process observed biophysically on the mGluR2 homodimer. The initial role of the bound agonist is to pull the bottom part of VFT3 (VFT of TAS1R3) toward the bottom part of VFT2 (VFT of TAS1R2) to transmit this movement from VFT2 (where agonists bind) through the VFT3 and the CRD3 (VFT and CRD of TAS1R3) to the TMD3 (TMD of TAS1R3). This facilitates G protein coupling and downstream signaling. The CRDs are crucial in this streamlined relay of structural changes where disulfide bonds provide rigidity to the CRD and amplify the mechanical constraints that help in attaining an active conformation (Cheron et al., 2019). This is empirically supported by a study in which a single mutation (A537P) in the CRD of TAS1R3 abolished the response to all sweeteners, indicating that the CRD3 must couple ligand binding in VFT2 to the conformational changes required in TMD3 for receptor activation.

Trafficking and cell surface expression are also crucial factors for sweet taste transduction. Molecular modeling with mutagenesis scanning revealed specific regions consisting of hydrophobic residues in ECD (site II; at the tip of CRD) and TMD regions (site IV; includes TMD6 and the cytoplasmic base of TMD5) of the TAS1R2 subunit to be important for dimerization with TAS1R3. Moreover, the CRD region and ECL2 domain of the transmembrane region seems to be important for surface co-expression of the TAS1R2/TAS1R3 dimer. In particular, the cytosolic C-terminus portion of the CRD region of TAS1R2 needs to be properly folded for coexpression and trafficking (Park et al., 2019). This reflects the difficulty in expressing these receptors at consistent levels in mammalian cell lines (Li et al., 2002; Shimizu et al., 2014).

28.18.5.2: Positive Allosteric Modulation of Sweet Receptor

Class C GPCRs pose an ideal target for allosteric modulation either positive (PAM) or negative (NAM). PAMs show little or no agonist activity on their own but significantly enhance agonist activity. Sweet taste is a major target of the food industry globally and non-caloric sweeteners are highly sought to exploit a huge commercial market. In a first comprehensive high throughput screen by Servant et al. (2010), novel PAMs (SE1, SE2, SE3; Table 4) for the sweet heteromer were reported that were not sweet on their own but significantly enhanced the sweetness of sucralose or sucrose. Agonist binding to the VFT region of TAS1R2 facilitates a closed conformation which constitutes an active state of the sweet receptor, while its open conformation represents an inactive state. Molecular modeling and mutagenesis studies revealed that these PAMs follow a similar mode of binding as that reported for umami PAMs (IMP and GMP). They bind near the opening of the binding pocket of the VFT region adjacent to their agonists, through Van der Waals and hydrogen bonding interactions, and utilize several critical residues for their activity. Although these residues are not in direct contact with any receptor-bound sweetener, mutation of some of them (K65, Y103, L279, D307, and R383) diminishes the response to sweeteners suggesting that these residues normally stabilize the closed conformation. The initial closing of the VFT region by agonist binding and further stabilization of the closed conformation by subsequent binding of SE modulators occurs in two steps. First, by interacting with the ECD region of TAS1R2, and second, by strengthening the hydrophobic interactions between the two lobes of ECD and lowering the free energy needed for their closure (Zhang et al., 2010).

TABLE 4 | Sweet taste receptor's positive allosteric regulators with concentration (used in cell-based assays in studies) and negative allosteric modulators with their IC50 values.

Positive allosteric modulators (PAMs)	Nature	Binding pocket	Conc. (mM)	References
SE1, SE2, SE3	Undisclosed	VFT (TAS1R2)	0.05	(Servant et al., 2010; Zhang et al., 2010)

Positive allosteric modulators (PAMs)	Nature	Binding pocket	Conc. (mM)	References
Neohesperidin dihydrochalcone (NHDC)	Flavonoid	TMD (TAS1R3)	0.25	(Jiang et al., 2005c; Winnig et al., 2007)
Unnatural tripeptides (several)	Biaryl derivative tripeptides	ND	2 – 20	Yamada et al., 2019
Sodium, cholesterol	Cation, lipid	TMD (TAS1R2)	ND	Perez-Aguilar et al., 2019
NAMs	—	—	IC50 (mM)	—
Lactisole	Carboxylic acid salt	TMD (TAS1R3)	0.041	(Jiang et al., 2005c)
(2-(2,4-dichlorophenoxy)propionic acid)	Carboxylic acid salt	TMD (TAS1R3)	0.006	(Nakagita et al., 2019)
Gymnemic acid	Triterpenoid glycoside	TMD (TAS1R3)	6.9	(Sanematsu et al., 2014)
Clofibric acid	Herbicide	TMD (TAS1R3)	1.4	(Maillet et al., 2009; Kochem and Breslin, 2017)
Amiloride	Diuretic	TMD (TAS1R2)	0.87	(Imada et al., 2010; Zhao et al., 2018)
Umami compounds: MSG, Glu-Glu, Glu-Asp	Peptides	VFT (TAS1R2)	ND	(Shim et al., 2015)

Where VFT, Venus flytrap domain; TMD, transmembrane domain; ND, not determined.

TABLE 4. Sweet taste receptor's positive allosteric regulators with concentration (used in cell-based assays in studies) and negative allosteric modulators with their IC50 values.

Using a high throughput chemical screening approach and heterologous expression of the TAS1R2/TAS1R3 heteromer, several unnatural tripeptides with a novel core biaryl structure were found as potential sweet enhancers (Yamada et al., 2019). This study divided the potential molecule into three parts namely, "head and linker" which together are essential for its sweet enhancer activity, while the "tail" determines the level of activity. This approach provided some useful inputs toward the synthesis of potent PAMs. Firstly, an amine incorporated at the α-position of carbonyl moiety in the tail structure interacts with the TAS1R2 subunit thereby increasing allosteric activity. Secondly, additional hydrophobic substitutions in the tail structure provided an increased allosteric activity to the molecule. Lastly, the distance between the head and linker and the insertion of an amide bond is crucial for its synthesis. Although their binding characteristics and allosteric mechanisms are not yet known, these observations provide a starting point to identify and synthesize new sweet PAMs in the future.

Small molecule PAMs can also bind to the transmembrane domain in class C GPCRs, in contrast to agonist which binds to the extracellular domain (Urwyler, 2011). For example, the flavonoid sweetener, neohesperidin dihydrochalcone (NHDC) binds to TMD regions to enhance the agonist-induced sweet response. It interacts with a receptor binding pocket in the TMD of TAS1R3 and requires seventeen critical residues in TMDs and extracellular loop 2 for its allosteric activity (Winnig et al., 2007). These residues also contribute to cyclamate and lactisole binding sites. Among seventeen residues, eight alter receptor activation by NHDC ($Q637^{3.29}$, $S640^{3.32}$, $H641^{3.33}$, $Y699^{4.60}$, $W775^{6.48}$, $F778^{6.51}$, $L782^{6.55}$, and $C801^{7.39}$) and influence lactisole mediated inhibition. Similarly, nine of the seventeen residues ($Q637^{3.29}$, $H641^{3.33}$, $H721^{ex2}$, $S726^{5.39}$, $F730^{5.43}$, $W775^{6.48}$, $F778^{6.51}$, $L782^{6.55}$, and $C801^{7.39}$) mediate activation by cyclamate, while six ($Q637^{3.29}$, $H641^{3.33}$, $W775^{6.48}$, $F778^{6.51}$, $L782^{6.55}$, and $C801^{7.39}$) influence receptor inhibition by lactisole as well as receptor activation by cyclamate [superscript refers to the nomenclature suggested for class C GPCRs by Pin et al. (2003) where first number denotes TMD region and the second number denotes residue position from the most conserved residue].

Notably, three critical residues in TMD6 ($W775^{6.48}$, $F778^{6.51}$, $L782^{6.55}$) and one in TMD7 ($C801^{7.39}$) of TAS1R3 were found crucial for allosteric binding, as their mutation to alanine altered the receptor's sensitivity to NHDC and cyclamate, as well as to the inhibitor lactisole (Winnig et al., 2005). Therefore, TMD6 and TMD7α helices of TAS1R3 are integral to allosteric modulation of the sweet receptor, implicating them in TAS1R2 and TAS1R3 subunit interactions and indicating an important role for this

structural region in the conformational changes involved in receptor activation. Furthermore, these residues are conserved across mammalian species (Cheron et al., 2019).

28.18.5.3: Negative Allosteric Modulation of Sweet Receptor

Like PAMs, negative allosteric modulators (NAM) such as lactisole and gymnemic acid bind to the TMD region of TAS1R3 and inhibit sweet substance-induced responses. Lactisole, an aralkyl carboxylic acid not only inhibits sweet but also the umami receptor response in humans and presents a rare opportunity to study the structural cross-talk between these two taste qualities. Using heterologous expression and mutagenesis, Jiang et al. (2005b) reported that lactisole's sweet inhibition might be mediated by its binding to TMD3, TMD5, and TMD6 of TAS1R3 and induce a conformation change which restricts the movement required to stabilize the active state. Residues A733$^{5.46}$ in TMD5, L798$^{7.36}$ in TMD7, and R790^{ex3} in extracellular loop 3 were found to be crucially important for sensitivity to lactisole in humans (Jiang et al., 2005b). These observations were confirmed in a recent study where 2-(2,4-dichlorophenoxy)propionic acid (2,4-DP) was found to be a more potent antagonist and utilize the same residues as well as four additional ones (H641$^{3.37}$, H734$^{5.43}$, F778$^{6.53}$, and Q794$^{7.32}$) in binding to TAS1R3. Moreover, the (S)- isomer of both compounds was found to be more strongly bound to the TMD of TAS1R3 and be a more effective inhibitor [lactisole; (S)-lactisole IC50, 20 μM while (R)- lactisole exerted no inhibition at this concentration.; 2,4-DP: (S)-isomer was 10-fold more effective than (R)-2,4DP]. The (S)- lactisole isomer interacts with the TMD via its carboxyl group and stabilizes in only one orientation in the binding pocket that does not allow for very strong binding. In contrast, (S)-2,4- DP binds through two moieties simultaneously, a carboxyl group and an aromatic ring with two Cl^- groups and stabilizes in several different orientations through hydrophobic interactions that allow stronger binding, resulting in stronger negative allosteric modulation (Nakagita et al., 2019).

These observations provide information about the relevance of structural modification in NAM compounds that could affect their interaction with the receptor. Although TMDs of TAS1R3 are the most likely regions responsible for allosteric modulation, TMDs and VFT regions of TAS1R2 cannot be ruled out completely. For example, the diuretic amiloride binds to TAS1R2 (TMD3, TMD5, TMD7) and inhibits the sweet response in a species-dependent manner (Zhao et al., 2018). Further, the umami compound [monosodium glutamate (MSG)] and peptides (Glu-Asp, Glu-Glu) bind to the VFT region of TAS1R2 and inhibit the sweet-induced response (Shim et al., 2015). These observations suggest that both subunits are important for the allosteric activity of TAS1R2/TAS1R3 and further structural studies are required to design novel sweet allosteric modulators.

28.18.6: Umami Taste Signal Transduction Mechanisms

In contrast to four well-known basic human tastes (sweet, bitter, salty, and sour), umami or ‘savory taste’ is relatively recent and was introduced in early 2000 by Kikuna Ikeda (Ikeda, 2002) as a new seasoning element in food. The main stimulus for the umami taste is the amino acid, L-glutamate present in the diet mainly in the form of MSG (Roper, 2007). Glutamate was first extracted from konbu/kombu (dried kelp of *Fucus vesiculosus*) and described as having a “unique taste” and “very different from other tastes”. The terminology “umami” comes from the Japanese word “umai” meaning “delicious.” Moreover, the taste of umami is also produced by food such as mushrooms and soy sauce that contain amino acids (L-aspartate), peptides, and synthetic ingredients similar to glutamate and some organic acids (Roper, 2007; Kinnamon, 2009) (Table 5).

TABLE 5 | Umami receptor agonists with their EC50 values and other pharmacological properties.

Agonist	Nature	EC50 (mM)	Binding pocket	References
L-amino acids (glutamate, aspartate, alanine, serine, asparagine, arginine, histidine, threonine, glutamine)	Amino acids	3 (glutamate), ND for others	VFT (TAS1R1)	(Li et al., 2002; Nelson et al., 2002; Zhang et al., 2008; Toda et al., 2013)
L-theanine	Amino acid (plant origin)	ND	VFT (TAS1R1)	(Narukawa et al., 2014)
VFT, venus flytrap domain; ND, not determined.				

TABLE 5. Umami receptor agonists with their EC50 values and other pharmacological properties.

The umami receptor (TAS1R1/TAS1R3) is a heteromeric member of the class C GPCRs, whereas most other receptors of this class exist as homodimers (Nelson et al., 2002; Temussi, 2009; Leach and Gregory, 2017). TAS1R1/TAS1R3 is the predominant umami taste receptor (Zhao et al., 2003; Behrens and Meyerhof, 2011) and the TAS1R1 subtype is critical for sensing umami taste as its

deletion abolished the response to umami taste stimuli (Mouritsen and Khandelia, 2012). However, TAS1R1/TAS1R3 is not the only receptor capable of detecting umami ligands (Chaudhari et al., 2000; Kunishima et al., 2000; Li et al., 2002; Nelson et al., 2002). Studies using heterologous expression, afferent nerve recordings, and behavioral experiments have confirmed that metabotropic glutamate receptors 1, and 4 (taste-mGluR1 and taste-mGluR4) also sense umami stimuli (Chaudhari et al., 2000; Kunishima et al., 2000; Li et al., 2002; Nelson et al., 2002). Notably, TAS1R3 knock-out mice show a strongly diminished response to glutamate and sweet stimuli (Damak et al., 2003) and taste cells isolated from these mice respond to IMP and glutamate which is abolished in presence of mGluR antagonists (Pal Chaudhry et al., 2016). TAS1R1/TAS1R3 is not only activated by glutamate but this activation is strongly enhanced in the presence of 5′-ribonucleotides, (inosine 5′ monophosphate; IMP) a response that is a hallmark of umami taste (Rifkin and Bartoshuk, 1980).

The main transduction components following the activation of TAS1R1/TAS1R3 are similar to those for sweet taste (Zhang et al., 2003), i.e., α-gustducin (and γ13/β1 or β3), PLCβ2, IP3R, and TRPM4/5. Cyclic nucleotides may also contribute to the transduction of umami taste in TRCs. When taste tissue is stimulated with umami, its cyclic AMP level is decreased (Abaffy et al., 2003). However, the consequence of decreased cAMP in TRCs has not yet been fully elucidated. Both α-transducin and α-gustducin are involved in umami taste signal transduction, as mice lacking the gene for one of these proteins showed a reduced response to this taste (He et al., 2004; Leach and Gregory, 2017). In the taste palate fungiform papillae, α-gustducin and α-transducin activate PDE that reduces cAMP levels. Ligand binding to the TAS1R1/TAS1R3 heterodimer releases Gβγ subunits to stimulate PLCβ2, which hydrolyzes PIP2 to DAG and IP3 (Kinnamon, 2009). IP3 then activates IP3R3 which results in the release of calcium ions from intracellular compartments (Clapp et al., 2001; Leach and Gregory, 2017) (Figure 3). Calcium ions activate TRPM5 and TRPM4 channels, leading to an influx of sodium ions, subsequent cell membrane depolarization, and finally release of ATP, which activates ionotropic purinergic receptors located in sensory fibers (Perez et al., 2002; Sugita, 2006). This pathway was confirmed when mice devoid of TRPM5, TRPM4, PLCβ2, and IP3R3 showed a reduced response to umami taste perception following glutamate stimuli (Damak et al., 2006; Kinnamon, 2009; Eddy et al., 2012).

28.18.6.1: Structural, Molecular, and Conformational Changes of Umami Receptor

In the last decade, several in-depth modeling and mutagenesis approaches have improved structural and molecular understanding of the umami receptor. The VFT regions of both subunits of TAS1R1/TAS1R3 comprise orthosteric and allosteric ligand binding sites for umami stimuli.

Mutagenesis and molecular modeling studies reveal that the cognate agonist glutamate binds in the VFT region of the TAS1R1 subunit of TAS1R1/TAS1R3 and stabilizes the closed active receptor conformation. Moreover, four residues in the TAS1R1 VFT region (S172, D192, Y220, and E301) showed no detectable response to glutamate when they were mutated to alanine suggesting that they are critical for glutamate binding. The glutamate binding and stabilization of the closed conformation of TAS1R1, activates the downstream signaling pathway, while TAS1R3 remains in an open (inactive) conformation. Therefore, closure of the VFT is the key event that sensitizes umami taste receptor signal transduction (Lopez Cascales et al., 2010). Apart from glutamate, other L amino acids were also found to elicit functional responses by binding to the corresponding VFT region of TAS1R1. Six residues that contributed to the acidic amino acid agonist (L-glutamate and L-alanine) responses have been identified (S148, R151, A170, E174, A302, and D435).

28.18.6.2: Allosteric Modulation of Umami Receptor

Because of significant advancements in understanding and food industry application of umami taste, its allosteric modulators are sought after. Several allosteric umami ligands have been discovered with varying potency, only a few of which have been characterized at the molecular level. The best-characterized umami PAMs, the 5′-ribonucleotides: inosine 5′-monophosphate (IMP) and guanosine 5′-monophosphate (GMP), interact with the VFT region of the TAS1R1 subunit to enhance the glutamate-induced response that is the hallmark of umami taste (Table 6). IMP and GMP binding sites in the VFT are adjacent to that for glutamate binding. The mutation of four residues (H71, R277, S306, and H308) abolished the IMP/GMP-induced glutamate response suggesting their involvement in the allosteric binding of these nucleotides. Structurally, IMP and GMP stabilize the closed form of the TAS1R1 VFT region through electrostatic interactions and coordinate the positively charged residues that act as pincers. The ability of IMP and GMP to interact with the VFT region (as opposed to the TMD region) represents a unique mechanism of positive allosteric regulation within class C GPCRs (Urwyler, 2011).

TABLE 6 | Umami receptor allosteric modulators with concentrations used in cell-based assays and other pharmacological properties.

Allosteric modulators	Nature	Conc. (mM)	Binding pocket	References

Allosteric modulators	Nature	Conc. (mM)	Binding pocket	References
IMP/GMP	Nucleotide	1	VFT (TAS1R1)	(Li et al., 2002; Nelson et al., 2002; Zhang et al.,
				2008)
Cyclamate	Sodium	8	TMD (TAS1R3)	(Xu et al., 2004)
	cyclohexylsulfamate			
Methional (3-methylsulfanylpropanal)	—	0.12	TMD (TAS1R3)	(Toda et al., 2018)
Lactisole (2-4-methoxyphenoxy propionic acid)	Carboxylic acid salt	5	TMD (TAS1R3)	(Xu et al., 2004)
Clofibric acid (4-chlorophenoxy)-2-methylpropanoic	Herbicide acid	4	TMD (TAS1R3)	(Maillet et al., 2009; Kochem and Breslin, 2017)
acid				

Where VFT, Venus flytrap domain; TMD, transmembrane domain.

TABLE 6. Umami receptor allosteric modulators with concentrations used in cell-based assays and other pharmacological properties.

In contrast to IMP and GMP which bind to the TAS1R1 extracellular domain, the well-known flavor compound methional and its analogs bind to the TMD region and allosterically regulate the umami receptor in a species-dependent manner (Toda et al., 2018). Importantly, methional utilizes several distinct residues in different TAS1R1 transmembrane domains (TMD2-7) to act as a PAM in the human umami receptor, yet it behaves as a NAM in the mouse counterpart. This unusual phenomenon provided an opportunity to study the mechanisms of both positive and negative modulation in TAS1R1 simultaneously (Toda et al., 2018).

Construction of chimeric receptors between human (h) and mouse (m) and their functional analysis demonstrated that the TMD of TAS1R1 is the key domain for switching the PAM/NAM activities of methional. Point mutation substitutions between these species identified four residues (h/m; F768/L769, N769/H770, S799/T800, and S802/G803) that are collectively required to switch PAM/NAM activities. A similar mode of allosteric regulation and PAM/NAM mode switching has been reported for mGluR5 (Gregory et al., 2013) suggesting this is an unusual and distinct phenomenon of the class C GPCRs. Further, alanine scanning mutagenesis in TAS1R1 of the corresponding residues vital for the activity of other taste inhibitors (sweetener inhibitors; NHDC and cyclamate; sweet and umami taste inhibitor; lactisole) revealed three residues required for PAM (W697$^{4.50}$ F728$^{5.40}$ and F732$^{5.44}$) and a single residue (F642$^{3.40}$) for NAM. These results suggest that both the PAM and NAM activities of methional are conferred by residues that are distinct from those required for the PAM/NAM switch. Knowing that methional is an important part of food seasoning globally, these observations could help in maximizing its use in enhancing flavors along with amino acids and nucleotides.

Despite PAMs being a central focus for umami allosteric modulation, there has also been considerable research on negative allosteric modulation where lactisole emerged as a prominent NAM of the umami receptor, TAS1R2/TAS1R3. Because umami and sweet receptors share the TAS1R3 subunit, findings from studies on sweet receptor lactisole binding are relevant. A comprehensive study on the sweet receptor identified critical residues within the TMD regions (S640$^{3.32}$, H641$^{3.33}$ in TMD3 and F778$^{6.51}$, L782$^{6.55}$ in TMD6) of TAS1R3 required for lactisole binding pocket and showed a large effect on sensitivity to lactisole (Xu et al., 2004; Jiang et al., 2005b). Because lactisole shares structural similarities with two other classes of compounds: fibrates and phenoxy-herbicides, researchers studied them to search for novel sweet/umami inhibitors (Maillet et al., 2009). The lipid-lowering drug, clofibric acid inhibits the TAS1R3 umami receptor-mediated response both *in vitro* and *in vivo* (Table 6). Like lactisole, clofibrate inhibits the umami taste from glutamate by binding with a similar affinity to TAS1R1/TAS1R3. However, its specificity against the umami receptor still needs to be validated alongside other umami taste receptors (mGluR1, mGluR4, or NMDA).

28.18.7: TYPE 2 TASTE G PROTEIN-COUPLED RECEPTORS (BITTER RECEPTORS)

Type 2 taste GPCRs are represented by bitter taste receptors that have a distinct subset of bitter sensing cells in type II TRCs and notably, 25 bitter taste receptors (TAS2Rs) are reported to be expressed in humans (Chandrashekar et al., 2000; Devillier et al.,

2015; Behrens and Meyerhof, 2018). A significant amount of work has been done to explore the diversity among TAS2Rs and their agonists in taste biology (Adler et al., 2000; Behrens and Meyerhof, 2009; Behrens and Meyerhof, 2018). Some TAS2Rs (TAS2R3, TAS2R5, TAS2R13, TAS2R50) are narrowly tuned to structurally similar bitter compounds, whereas others are broadly tuned (TAS2R10, TAS214, TAS2R46), responding to several bitter compounds. Initially, it was believed that each bitter-sensitive type II TRC expressed every TAS2R isoform (Adler et al., 2000) but other studies suggest that TAS2Rs can be expressed differentially, allowing for possible discrimination among bitter compounds (Caicedo and Roper, 2001; Behrens and Meyerhof, 2009; Behrens et al., 2009). Please refer to figure 4B for the basic structure of the bitter receptor.

28.18.7.1: Bitter Taste Signal Transduction Mechanisms

The bitter taste is the most complex of all the five basic tastes and protects against the ingestion of toxic substances by eliciting an innate aversive response across species (Chandrashekar et al., 2006; Behrens and Meyerhof, 2018). The TAS2Rs that mediate bitter taste perception are among ∼50 TAS2Rs identified in mammals, and 25 are known to be expressed in humans (Adler et al., 2000; Devillier et al., 2015; Yoshida et al., 2018). The TAS2R family is the most diverse and binds to a wide range of agonists compared with the other taste GPCRs (Jaggupilli et al., 2016) (Supplementary Table 1).

TAS2Rs are distinctive among class A GPCRs in that many of them bind agonists with low apparent affinity in the micromolar range, rather than the nanomolar range (Di Pizio et al., 2016). The activation of TAS2Rs by harmless, minute amounts of bitter compounds such as those contained in most vegetables would limit the availability of food resources appearing safe for consumption and therefore could negatively affect survival. Hence, the concentration ranges at which bitter taste receptors are activated are well-balanced to allow species to maintain a healthy diet yet avoid ingestion of spoiled food containing strongly bitter ligands.

Hundreds of bitter compounds have been reported to evoke bitterness and activate human bitter receptors in different cell-based assays. These bitter agonists include plant-derived and synthetic compounds such as peptides, alkaloids, and many other substances (Supplementary Table 1). (Pronin et al., 2004; Meyerhof et al., 2010; Iwata et al., 2014). Some TAS2Rs are activated by a wide range of compounds, whereas others show strict specificity for a single bitter compound (Behrens et al., 2009; Sakurai et al., 2010a; Born et al., 2013). Interestingly, TAS2R31, TAS2R43, and TAS2R46 have around 85% sequence homology, but they bind to different agonists (Brockhoff et al., 2010; Jaggupilli et al., 2016), reinforcing the idea that each TAS2R might have a unique ligand-binding pocket.

The canonical TAS2R signal transduction cascade signaling molecules shared among bittersweet and umami receptors (Wong et al., 1996; Huang et al., 1999; Mueller et al., 2005), includes the heterotrimeric G protein subunits (Gα-gustducin, Gβ3, and Gγ13), (Ishimaru, 2009; Shi and Zhang, 2009), a phospholipase C (PLCβ2), an inositol trisphosphate receptor (InsP3R), and the TRPM5 ion channel. Upon receptor activation by bitter ligands, the G protein α-gustducin dissociates from its βγ subunits. The latter activates PLCβ2, leading to a release of Ca^{2+} from IP3-sensitive Ca^{2+} stores, resulting in Na^{+} influx through TRPM5 channels. This Na^{+} influx depolarizes the cells and causes the release of neurotransmitter ATP through gap junction hemichannels or CALHM1 ion channels (Finger et al., 2005; Chaudhari and Roper, 2010; Taruno et al., 2013) (Figure 3).

28.18.7.2: Structural, Molecular, and Conformational Changes of Bitter Receptors

Classification of TAS2Rs has always been ambiguous because they were originally considered to be a distinct family (Horn et al., 2003) or grouped with the frizzled receptors (Fredriksson et al., 2003; Jaggupilli et al., 2016), but most recent analyses (Di Pizio et al., 2016) support their classification with Class A GPCRs. The ability of bitter taste receptors to interact with numerous structurally diverse substances compared to other GPCRs is remarkable and includes a wide range of drugs/antibiotics, polyphenols, bacterial metabolites, salts, and metal ions (Supplementary Table 1). Therefore, exploring the criteria for the identification of highly heterogeneous bitter compounds with pronounced selectivity has become a major research area. Some of these studies rely solely on in silico homology/computational modeling (Dai et al., 2011; Tan et al., 2012; Di Pizio et al., 2020; Dunkel et al., 2020) and others on *in vitro* genetic modification and functional assay systems (Pronin et al., 2004; Nowak et al., 2018; Jaggupilli et al., 2019).

As a group of over ∼50 receptor subtypes, TAS2Rs recognize structurally diverse agonists where some are broadly tuned (TAS2R46, TAS2R14, TAS2R10, and TAS2R43) recognize diverse agonists, while others (TAS2R1, TAS2R4, TAS2R7) show strong selectivity and narrow tuning (Liu et al., 2018; Wang et al., 2019). The agonist binding cavity in most bitter GPCRs is located deep within their transmembrane domain (TMD), except TAS2R7 in which it resides on the extracellular surface (Liu et al., 2018). TAS2Rs are also distinct in containing highly conserved TMD regions, with thirteen key residues and two motifs (LXXXR in TMD2 and LXXSL in TMD5) that are absent in class A GPCRs, and may reflect their different activation mechanisms (Singh et

al., 2011). LXXSL plays a structural role by stabilizing the helical conformation of TMD5 at the cytoplasmic end and a functional role by interacting with residues in intracellular loop 3 (ICL3) which is important for proper receptor folding and function (Singh et al., 2011). Moreover, mutation of the conserved residues in LXXSL and LXXXR motifs results in protein misfolding and poor surface expression (Singh et al., 2011; Pydi et al., 2014a).

The initial study highlighting the structure–activity relationship of bitter taste receptors was performed with receptors belonging to a subfamily of closely related TAS2Rs (Pronin et al., 2004). By physically swapping the extracellular loop 1 (ECL1) between TAS2R43 and TAS2R31, chimeric TAS2R31/TAS2R43 (ECL) gained responsiveness to the compound n-isopropyl-2methyl-5-nitrobenzenesulfonamide (IMNB), whereas the reverse chimera TAS2R31 (ECL)/TAS2R43 lost responsiveness for IMNB. Although this report supports an important contribution of residues located within the transmembrane region of the investigated receptors, the extracellular loops appear to be of importance for agonist selectivity. This empirical finding contrasts with earlier computational studies which predicted the agonist binding site to lie within the helical bundle of TAS2Rs without particular contacts between extracellular loops and docked agonists (Floriano et al., 2006; Miguet et al., 2006).

28.18.7.3: Bitter Receptor Ligand Binding Pocket

The emergence of TAS2Rs as the most broadly tuned taste receptors might give the impression that their specific interaction with numerous agonists is because of several binding pockets that accommodate subgroups of bitter compounds. However, structure–function analysis of TAS2Rs (except for TAS2R7) has demonstrated the presence of only a single agonist binding pocket comprising the upper parts of TMD2, TMD3, TMD5, TMD6, and TMD7. The reason for their broad tuning and recognition of such a broad spectrum of agonists might most likely be attributed to the presence of an additional extracellular binding site called a “vestibular site,” in addition to the orthosteric selecting as reported for TAS2R46 (Sandal et al., 2015). This two-site architecture offers more ligand recognition points than a single one and thus might help in selecting the appropriate agonists. Moreover, the presence of the vestibular site may also help to discriminate among the wide spectrum of bitter ligands.

Although broadly tuned receptors (TAS2R46, TAS2R31, and TAS2R43) have high homology in amino acid sequence, their agonist profiles only slightly overlap (Kuhn et al., 2004; Brockhoff et al., 2007; Di Pizio and Niv, 2015) which suggests the involvement of key residues at different positions in agonist specificity. Consequently, when strychnine interacting positions in TAS2R46 (residues differ at this position in TAS2R31, TAS2R43) were exchanged between these two receptors not only was the strychnine responsiveness transferred to the recipient receptor (TAS2R31, TAS2R43), but also sensitivity to additional TAS2R46 agonists (absinthin and denatonium). Sensitivity to activation by aristolochic acid was lost in the mutant receptors (Brockhoff et al., 2010). This experimental evidence supports the presence of a common agonist binding pocket and agrees with other studies on TAS2R16, TAS2R14, and TAS2R7 receptors (Sakurai et al., 2010a; Sakurai et al., 2010b; Thomas et al., 2017; Liu et al., 2018; Nowak et al., 2018).

Recent studies used homology modeling and mutagenesis to elucidate the nature of the ligand-binding pocket in TAS2R7, TAS2R14, and TAS2R16 receptors (Thomas et al., 2017; Liu et al., 2018; Nowak et al., 2018). They reported that the binding pocket is flexible and wide open to accommodate molecules of diverse sizes and shapes, and thus permits chemical modifications among agonists as well (Thomas et al., 2017; Liu et al., 2018; Nowak et al., 2018). Although the molecular basis for the promiscuity of bitter receptors is attributed to their apparent flexible spacious binding site, future work elucidating the contact points between TAS2Rs binding site residues and its agonists in terms of additional binding locations is required.

28.18.7.4: Bitter Receptors Ligand Binding Domain and Amino Acid Residues

A majority of the TAS2R studies are based on molecular modeling, mutagenesis, and heterologous expression systems (Biarnes et al., 2010; Brockhoff et al., 2010; Tan et al., 2012; Nowak et al., 2018; Shaik et al., 2019) suggest that the ligand binding pocket is formed by several key residues in most TMDs (TMD1, TMD2, TMD3, TMD5, TMD6, and TMD7), except for TMD4.

Studies show similarities as well as differences regarding residues and positions involved in agonist-receptor interactions. However, most of them agree that besides position $N^{3.36}$ in TMD3 (superscript as per Ballestros-Weinstein nomenclature for class A GPCRs) (Ballesteros and Weinstein, 1995) and other residues ($L^{3.32}$, $L^{3.33}$, and $E^{3.37}$) in its close proximity, play a role in agonist activation of several broadly tuned TAS2Rs (TAS2R1, TAS2R16, TAS2R30, TAS2R38, TAS2R46) (Pronin et al., 2004; Biarnes et al., 2010; Brockhoff et al., 2010; Sakurai et al., 2010b; Dai et al., 2011). In contrast, for the narrowly tuned TAS2R7, one position in TMD3 ($H94^{3.37}$) and another in TMD7 ($E264^{7.32}$) were found crucial for metal ion binding (Wang et al., 2019). Mutagenesis and molecular modeling revealed that these two residues contribute to the metal ion binding pocket in TAS2R7. Moreover, metal ions bind distinctively to residues lining the binding pocket and interestingly, the presence of calcium in the assay solution appears to affect the TAS2R7 response to metal ions. It is not clear how calcium affects metal ion binding to TAS2R7, but it might work

cooperatively with certain ions and not others. Future studies focusing on structural interactions between the receptor and metal ions will provide further insights into how they activate the receptor.

In TMD2, two studies suggest that position $N^{2.61}$ is critical for binding in TAS2R1 (Singh et al., 2011) and TAS2R46 (Brockhoff et al., 2010). Likewise, in TMD7, position $265^{7.39}$ is implicated in binding to TAS2R46 (E265) and TAS2R1 (I263) (Dai et al., 2011). In TMD5, position $H^{5.43}$ is implicated in binding in TAS2R16 and $E^{5.46}$ in TAS2R1 (Dai et al., 2011) while, in TMD7, position $E^{7.32}$ was crucial for metal ion binding (Wang et al., 2019). These residues represent putative contact points for agonist interaction and form a pattern of being spaced one helical turn from each other.

Recent mutagenesis studies (Nowak et al., 2018; Di Pizio et al., 2020) performed in broadly tuned TAS2R14 with agonists (aristolochic acid, picrotoxinin, thujone) found several residues in TMDs to be involved in agonist binding. However, in contrast to TAS2R10 (Born et al., 2013) and TAS2R46 (Brockhoff et al., 2007), mutation of TAS2R14 did not result in a complete loss of function for all agonists but a varied reduction in responsiveness or selectivity toward agonists. Among several mutants, only mutation of W89A resulted in a complete loss of responsiveness against picrotoxinin while others showed more subtle agonist selective changes. This indicates that TAS2R14 is not streamlined for the most sensitive detection of selected agonists, but rather tailored to detect numerous diverse agonists, with comparatively lower apparent affinity.

The binding characteristics of bacterial acyl homoserine lactones (AHLs) on TAS2Rs (TAS2R4, TAS2R14, and TAS2R20) suggest the presence of a single orthosteric site situated close to the extracellular surface and reinforce the significant role of the extracellular loop structure (ECL2) in TAS2R ligand binding and activation (Jaggupilli et al., 2018). The crucial AHL binding residues in TAS2R4 and TAS2R14 are predominantly located in the ECL2, while in TAS2R20 they are present in TMD3 and TMD7 helices. The ECL2 residues, N165 in TAS2R4, and R160 and K163 in TAS2R14 were found crucial for lactone binding. In contrast, TAS2R20 residues W88 (TMD3) and Q265 (TMD7) are essential for agonist binding (Pydi et al., 2014c; Zhang et al., 2017; Jaggupilli et al., 2018). In addition, the hydrophobic amino acids in the three TAS2Rs are considered important in directing the orientation of the hydrophobic acyl chains of lactones that facilitate receptor activation.

The transmembrane domain in GPCRs is composed mainly of hydrophobic amino acids accommodated in the plasma membrane. Therefore, hydrophobic properties of the receptor binding pocket are important for any membrane-accessible agonist. Hydrophobic residues in TMD3 and TMD7 of TAS2R16 are important in forming a wide ligand-binding pocket (Thomas et al., 2017) that accommodates larger ligands like the β-glycosides. By using salicin analogs as TAS2R16 novel agonists (differ structurally from salicin in β-glucoside core constituents), several critical residues were identified that are required for signaling. Interestingly, these were identical to the residues critical for salicin signaling, except for W261, which was not required for activation by the analog 4-NP-β-mannoside. Importantly, all these residues are in the TMD helices or intracellular face of the receptor, consistent with classical GPCR signal transduction. These results suggest that larger ligands bind to the wide binding pocket of TAS2R16 on the extracellular side, and then their signal is transduced via conserved residues on the intracellular side. This can account for the broad spectrum of ligand recognition conferred by TAS2R16.

Unlike broadly tuned receptors, narrowly tuned ones like TAS2R7 show two different types of critical residue in ligand binding. The first type includes D86, W170, and S181 which are agonist independent and their mutation significantly reduces the ability of TAS2R7 to bind agonist, while a second group consisting of D65 and W89 are selective for quinine and enhance binding to a specific category of ligand (Liu et al., 2018).

Despite the variation in the amino acid type and location important for agonist binding among receptors of the bitter family, for the most part, ligand binding pockets are present on the extracellular surface of TMDs or ECL2. The function of the residues at these binding pockets is dictated by multiple factors that include the type of ligand, the movements in TMDs, and the associated movement of ECL2 to accommodate the ligand. Structure–function studies have identified a conserved KLK/R motif in the intracellular carboxyl-terminal domain of 19 TAS2Rs that is critical for cell surface expression, trafficking, and receptor activation (Upadhyaya et al., 2015; Jaggupilli et al., 2016).

28.18.7.5: Agonist, Antagonist Binding and Modulation of Bitter Receptors

In simple pharmacological terms, an antagonist is a ligand that inhibits the biological response induced by an agonist and does not induce any response of its own, while a ligand that reduces the constitutive/basal activity of a GPCR is considered an inverse agonist. An antagonist acts as a competitive inhibitor to block receptor activity. Large numbers of agonists have been identified for bitter receptors, but few antagonists have been found so far (Table 7). Finding an antagonist/inhibitor for bitter taste would not only help in understanding the TAS2R mechanism of signal transduction but have potential use in foods to overcome unwanted bitterness in consumer products. Such bitter blockers have been proposed to increase the palatability of bitter-tasting food and

beverages, increase compliance in taking bitter-tasting drugs, especially children's formulations and reduce or prevent off-target drug effects in extra-oral tissues (Clark et al., 2012)

TABLE 7 | Bitter taste receptor inhibitors with their IC50 values and other pharmacological properties.

Antagonist	Mode of action	Bitter receptors	Tested agonists	IC50 (μM)	References
GIV3727or 4-(2,2,3-trimethylcyclopentyl) butanoic acid	Competitive orthosteric inhibitor	31	acesulfameK	6.4	(Slack et al., 2010)
		43	Aristolochic acid	11.33	
		4	Colchicine	108	
		40	Cohumulone	6.24	
Gamma-aminobutyric acid (GABA)	Orthosteric inhibitor	4	Quinine	3.2	(Pydi et al., 2014b)
3β-hydroxydihydrocostunolide (3HDC)	ND	46	Absinthin	14.1	(Slack et al., 2010; Brockhoff et al., 2011)
			Andrographolide	4.9	
			Denatonium	6.8	
			Picrotoxinin	4.7	
			Strychnine	15.3	
3-hydroxypelenolide(3HP)	ND		Absinthin	57.8	(Brockhoff et al., 2011)
			Andrographolide	44.5	
			Denatonium	51.4	
			Picrotoxinin	22.9	
			Strychnine	84.9	
Probenecid	Allosteric inhibitor	16	Salicin	292	(Greene et al., 2011)
Sakuranetin	ND	31	Saccharin	5.5	(Fletcher et al., 2011)
6-Methoxysakuranetin	ND	31	Saccharin	10.2	(Fletcher et al., 2011)
Jaceosidin	ND	31	Saccharin	11.7	(Fletcher et al., 2011)
6,3′-dimethoxyflavanone	ND	39	Epicatechin gallate (ECG)	4075	(Roland et al., 2014)
			Denatonium	240	
6-Methoxyflavanone	ND	39	Epicatechin gallate (ECG)	479	(Roland et al., 2014
N,N-bis(carboxymethyl)-l-lysine(BCML)	ND	4	Quinine	0.059	(Pydi et al., 2014b)
(±) abscisic acid (ABA)	ND	4	Quinine	34.4	(Pydi et al., 2015)
ND, not determined.					

TABLE 7. Bitter taste receptor inhibitors with their IC50 values and other pharmacological properties.

To date, ~12 bitter inhibitors have been reported to interact with only 10 TAS2Rs subtypes (Table 5) by binding to transmembrane domains in a similar manner to agonists. GIV3727 (4-(2,2,3-trimethylcyclopentyl) butanoic acid) was the first TAS2R antagonist discovered and to be well characterized structurally (Slack et al., 2010) that acts as an orthosteric competitive antagonist for TAS2R31. It competes with the acesulfame K agonist both *in vitro* and *in vivo*. GIV3727 is moderately selective because it inhibits multiple bitter receptors including, TAS2R4, TAS2R40, and TAS2R43. Homology modeling revealed that the -COOH group in GIV3727 is important for ligand-receptor interactions as its replacement with an ester or the corresponding alcohol abolished its antagonist activity. Moreover, a mutagenesis study in TAS2R31 and TAS2R43 revealed residues $K265^{7.39}$ and $R268^{7.39}$ in TMD7 to be crucial for its antagonistic activity (Slack et al., 2010). Similarly, another non-selective inhibitor, probenecid (p-(dipropylsulfamoyl) benzoic acid) was found to act as NAM of TAS2R16 activity and inhibits TAS2R38 and TAS2R43 as well (Greene et al., 2011). Two point mutations, P44T and N96T in TMD3 of hTAS2R16 were found to significantly suppress the ability of probenecid to inhibit salicin activity. Hydrophobicity seems important for their pharmacological activity as observed for both probenecid and GIV3727. The sesquiterpene lactone, 3β-hydroxydihydrocostunolide (3HDC) is an interesting bitter blocker as it acts as a competitive antagonist of TAS2R46, TAS2R30, TAS2R40, yet activates TAS2R4, TAS2R10, TAS2R14, and TAS2R31 as an agonist (Brockhoff et al., 2011).

Similarly, various flavanones were also noted as antagonists for TAS2R31, and TAS2R39 with varying efficacy. Taken together most of the currently known antagonists are non-selective and there is an urgent need for studies that focus on selective antagonists of major broadly tuned TAS2Rs (such as TAS2R10, TAS2R14, TAS2R16, and TAS2R46). To target bitterness in terms of food industry needs, potential peptide inhibitors from different protein sources such as hen protein hydrolysates (inhibits TAS2R4, TAS2R7, TAS2R14) and beef proteins (inhibits TAS2R4) (Zhang et al., 2018; Xu et al., 2019) are reported to be effective. Several umami glutamyl peptides isolated from soybeans have been found to act as non-competitive allosteric inhibitors of TAS2R16 against the salicin-induced response (Kim et al., 2015).

28.18.7.6: Constitutive Activity of Bitter Receptors

A phenomenon in GPCR activity is that of constitutive activity, essentially an active state occurring in the absence of an agonist which has been demonstrated in more than 60 GPCRs (Seifert and Wenzel-Seifert, 2002). It is the production of a second messenger or downstream signaling by a receptor in a ligand-independent manner. The constitutive activity provides another possibility for taste inhibitor discovery using inverse agonists. Inverse agonists can inhibit both agonist-dependent and agonist-independent activity, while antagonists can inhibit only agonist-dependent activity (Chalmers and Behan, 2002). Interestingly, some mutations in GPCRs can lead to constitutive activity and receptors with this characteristic (including constitutively active mutants or CAM) are important tools to investigate new bitter inhibitors. Although constitutive activity has not been observed naturally in TAS2Rs, when induced by mutation these receptors provide a useful means to investigate the relationship between an active receptor conformation and inverse agonist pharmacology.

Molecular modeling and functional assays report five CAMs critical residues for TAS2Rs, one in TMD7 ($S285^{7.47}$) and four others in intracellular loop 3 (H214A, Q216A, V234A, and M237A) (Pydi et al., 2014a; Pydi et al., 2014b). Of the five CAMs, only the TAS2R4 with H214A mutation shows a 10-fold increase in constitutive activity. This histidine residue is highly conserved in most TAS2Rs. Mutation of H214 (H214A) helped in finding two new inverse agonists (GABA and ABA; Table 7) (Pydi et al., 2015). Similar pharmacological approaches can be used to generate mutants of all TAS2Rs to screen for their inverse agonist/bitter taste blockers. However, for better characterization and interpretation of TAS2Rs, future *in vivo* studies should be performed to understand the functional relevance of these CAMs. At the same time, it is worth noting that the potential presence of endogenous agonists makes it difficult to determine the true constitutive activity of GPCRs including TAS2Rs (Devillier et al., 2015).

28.18.8: Kokumi Sensation Signal Transduction

In addition to the five basic tastes, sensations beyond these add another dimension to taste perception. One such example is “kokumi” which is distinct from the other five tastes in that it does not have a taste as such but rather induces a sensation of “mouthfulness,” depth, thickness, and aftertaste in the flavors. Although this flavor has been used historically and is well recognized in Japanese cuisine, it was first characterized by Ueda et al. (1990) who isolated a kokumi taste substance from water extracts of garlic and onion and identified, γ-glutamylcysteinylglycine or glutathione (GSH) as the main active ingredient of kokumi flavor (Ueda et al., 1990; Ueda et al., 1997; Dunkel et al., 2007). GSH is abundantly present in food-grade yeast extract and has been used to make foods more flavorsome.

Kokumi signal transduction was unknown until CaSR expression was reported in a subpopulation of taste cells in mice and rats that suggested it could function as a taste receptor for calcium and amino acids (San Gabriel et al., 2009; Bystrova et al., 2010).

However, its apparent role in kokumi stimuli detection was not confirmed. Ohsu et al. (2010) for the first time reported that kokumi peptides (GSH, γ-Glu-Val-Gly, and various γ-glutamyl peptides; Table 8) signal through CaSR and can synergize with sweet, salty, and umami taste qualities to impart an augmented kokumi sensation, i.e., increased depth of flavor which was further complemented by later studies (Maruyama et al., 2012; Kuroda and Miyamura, 2015). By using heterologous expression systems and human sensory analysis these studies demonstrated that kokumi peptides impart kokumi sensation to sweet, salty, and umami taste via CaSR as the kokumi component was specifically suppressed in the presence of the CaSR-specific NAM NPS-2143. To further validate this idea, Maruyama et al. (2012) identified a distinct population of taste cells expressing CaSR in mouse lingual tissue which did not express either sweet or umami receptors. Notably, these cells are specifically responsive to kokumi substances and elicit a Ca^{2+} response to focally applied kokumi stimuli in mouse lingual slices. Moreover, this response was inhibited in the presence of NPS-2143. These findings support the idea that CaSR mediates kokumi sensation effects in TRCs

TABLE 8 | Kokumi sensation receptor agonists, allosteric modulators with concentrations used in cell-based assays.

Ca^{2+}	Orthosteric agonist/cation	1[a]	VFT	
Mg^{2+}	Orthosteric agonist/cation	10[a]	VFT	
Gd_{2+}	Orthosteric agonist/cation	0.02[a]	VFT	
Al^{2+}	Orthosteric agonist/cation	0.5[a]	VFT	
Sr^{2+}	Orthosteric agonist/cation	0.5[a]	VFT	
Mn^{2+}	Orthosteric agonist/cation	0.5[a]	VFT	
Ni^{2+}	Orthosteric agonist/cation	0.5[a]	VFT	
Ba^{2+}	Orthosteric agonist/cation	0.2[a]	VFT	
Ca^{2+}	Orthosteric agonist/cation	1[a]	VFT	Ca^{2+}
Spermidine	Orthosteric agonist/polyamine	0.002[a]	VFT	(Nemeth et al., 2018)
Neomycin	Orthosteric agonist/aminoglycoside antibiotic	0.06[a]	VFT	(Katz et al., 1992)
Gentamicin	Orthosteric agonist/aminoglycoside antibiotic	0.15[a]	VFT	Katz et al., 1992)
Kanamycin	Orthosteric agonist/aminoglycoside antibiotic	0.1	VFT	(Katz et al., 1992)
Amyloid β-peptides	Orthosteric agonist/Peptide	0.001–0.04	—	(Ye et al., 1997)
Poly-Lysine	Orthosteric agonist/peptide	0.03 μM[a]	VFT	(Brown et al., 1991; Nemeth et al., 2018)
Poly L-arginine	Orthosteric agonist/peptide	0.004 μM[a]	VFT	Brown et al., 1991; Nemeth et al., 2018)
Lysozyme	Agonist/protein	0.59[a]	ND	(Yamamoto et al., 2020)
Thaumatin	Agonist/protein	0.07[a]	ND	(Yamamoto et al., 2020)
Aromatic L-amino acids (Trp, Phe, His, Ala, Ser)	PAMs	10	VFT	(Conigrave et al., 2000; Mun et al., 2004; Geng et al., 2016)
Anions (SO_4^{2-})	NAM	10	VFT	(Geng et al., 2016)
Cinacalcet	PAM/phenylalkylamine	0.051 μM[a]	TMD	(Miedlich et al., 2002; Petrel et al., 2004; Nemeth et al., 2004)

Calindol	PAM/phenylalkylamine	0.31 μM[a]	TMD	Miedlich et al., 2002; Petrel et al., 2004)
NPS R-568	PAM/phenylalkylamine	0.5 μM[a]	TMD	(Miedlich et al., 2002; Petrel et al., 2004)
NPS R-467	PAM/phenylalkylamine	0.01	TMD	(Miedlich et al., 2002; Petrel et al., 2004)
γ-Glu-Val-Gly	PAM/Peptide	0.041 μM[a]	—	(Ohsu et al., 2010)
γ-Glu-Cys-Gly (Glutathione)	PAM/Peptide	76.5 μM[a]	VFT	(Ohsu et al., 2010; Wang et al., 2006
γ-Glu-Ala	PAM/Peptide	3.65 μM[a]	ND	(Wang et al., 2006; Ohsu et al., 2010)
γ -Glu-Val	PAM/Peptide	1.34 μM[a]	ND	(Wang et al., 2006; Ohsu et al., 2010)
γ -Glu-Cys	PAM/Peptide	0.45 μM[a]	VFT	(Ohsu et al., 2010; Wang et al., 2006)
γ -Glu-α-aminobutyryl-Gly (Opthalmic acid)	PAM/Peptide	0.018 μM[a]	ND	(Ohsu et al., 2010)
NPS2143	NAM	0.0003 (IC50)	TMD	(Gowen et al., 2000; Petrel et al., 2004)
Calhex 231	Mixed PAM/NAM	0.1–1 μM (PAM); 3–10 μM (NAM)	TMD	(Petrel et al., 2003; Petrel et al., 2004; Gregory et al., 2018)

Where VFT, venus flytrap domain; TMD, transmembrane domain; ND, not determined. a shows EC50 value.

TABLE 8. Kokumi sensation receptor agonists, allosteric modulators with concentrations used in cell-based assays.

More recently, kokumi peptides have been found to have an extraoral physiological role in the gastrointestinal tract where they stimulate the secretion of hormones (cholecystokinin and glucagon-like peptide1 by activating CaSR (Depoortere, 2014; Yang et al., 2019). However, future studies with tissue-specific deletion of CaSR in taste buds would help delineate its role in taste physiology.

CaSR involvement in taste is a relatively recent discovery, but its central role in extracellular calcium homeostasis in mammals is well recognized (Brown et al., 1993; Brown, 2013). Diverse ligands activate CaSR, including cations (Ca^{2+} and Gd^{3+}), peptides, polyamines (Brown and MacLeod, 2001), and amino acids (Conigrave et al., 2000; Conigrave and Hampson, 2006) (Table 8). Unlike other taste modalities (sweet, bitter, and umami), CaSR–ligand binding and recruitment of G protein results in the activation of an intricate, amplifying signaling network that initiates numerous intracellular functions. The functional diversity of CaSR results from its ability to activate multiple Gα proteins (Gq/11, Gi/o, G12/13, and Gs) (Magno et al., 2011; Conigrave and Ward, 2013) which subsequently affect multiple signaling pathways related to the pathophysiology of parathyroid hormone secretion, cancer, and metastasis (Kelly et al., 2007; Wettschureck et al., 2007; Mamillapalli et al., 2008).

Kokumi substrates activate CaSR and transmit their signal through Gαq/11 proteins which further activate PLCβ that results in the release of intracellular Ca^{2+} stored through activation of IP3 receptor channels in the ER. Whether the kokumi pathway strictly relies on Gαq/11 protein or can also use Gα-gustducin, like other taste modalities for downstream signaling, is still unknown (Figure 3). The growing number of reports on kokumi flavor signal transduction are shedding light on its potential use as a flavor enhancer.

28.18.8.1: Structural, Molecular, and Conformational Changes of Kokumi Receptor

CaSR belongs to the class C GPCR. Within this class, CaSR and metabotropic glutamate receptors (mGluRs) are known to function as disulfide-linked homodimers (Bai et al., 1998; Ward et al., 1998; Pidasheva et al., 2006) (Figure 4A). Structurally, the human CaSR is similar to sweet and umami taste receptors but differs in being a homodimer instead of a heterodimer (Hendy et al., 2013). The ECD of CaSR not only senses nutrients (Ca^{2+}, L-Phe, and polypeptides; Table 8) and allows ligands to modulate CaSR cooperatively, but is also required for its dimerization (Ray et al., 1999; Zhang et al., 2014). The binding of Ca^{2+} and other ligands

to the ECD changes the conformation of the seven transmembrane domains, causing alterations in the intracellular loops and the intracellular domain (ICD), which further trigger downstream signaling pathways (Brown et al., 1975). The ICD is relatively diverse among species and participates in controlling CaSR signaling in multiple ways by modulating receptor expression, trafficking, and desensitization (Gama and Breitwieser, 1998; Ward, 2004; Huang et al., 2006).

Homology modeling, mutagenesis, and heterologous expression revealed distinct and closely located binding sites for Ca^{2+} and aromatic L-amino acids, in VFT and the cleft of the VFT, respectively (Silve et al., Conigrave et al.,2000; Huang et al., 2009). Notably, four putative Ca^{2+} binding sites of varying affinity have been predicted in the VFT of the CaSR and in which the interaction between site 1 and the other three sites plays a central role in positive cooperativity in sensing Ca^{2+} (Zhang et al., 2014). Besides Ca^{2+}, aromatic L amino acids (L-Trp, L-Phe) also activate the CaSR by binding adjacent to the VFT region through three serine and one threonine residue (S169/S170/S171/T145). Interestingly, the double mutation T145/S170 was found to selectively impair L amino acid (Phe, Trp, His) sensing of CaSR, while Ca^{2+} sensing remained intact (Mun et al., 2004; Mun et al., 2005).

The recent crystal structure of the entire extracellular domain of CaSR (Geng et al., 2016) identified four novel Ca^{2+} binding sites in each protomer of the homodimer including one at the homodimer interface which does not correspond to any of the sites reported previously by Huang et al., (2007). It is unclear why these additional calcium-binding sites were not found in earlier studies. This might be due to the different expression systems used, crystallization conditions, and methods of analysis. The conditions of the more recent studies may have stabilized an active conformational state in which these calcium sites become available (Geng et al., 2016). Among these four Ca^{2+}-binding sites, site 4 seems most relevant to receptor activation as it directly participates in the active CaSR conformation. Moreover, a previously reported natural mutation G557E (Hendy et al., 2009) reduced the potency of Ca^{2+} possibly by affecting backbone conformation, thereby weakening the affinity of Ca^{2+} for this site. This confirms that a Ca^{2+} ion at site 4 stabilizes the active conformation of the receptor by facilitating homodimer interactions between the membrane-proximal LBD2 region and CRD of CaSR.

The most interesting aspect of Ca^{2+} and L-amino acid interplay was reported by Zhang et al. (2014) who studied L-Phe binding characteristics by monitoring intracellular $[Ca^{2+}]_i$ oscillations in living cells and performing molecular dynamic simulations. Their findings supported a previous observation that the L-Phe binding pocket is adjacent to the Ca^{2+} binding site 1. Importantly, by binding to this site, L-Phe influences all Ca^{2+} binding sites in the VFT region and enhances CaSR functional cooperativity through positive heterotropic cooperativity to Ca^{2+}. Moreover, the dynamic communication of L-Phe at its predicted binding site in the hinge region with the Ca^{2+} binding sites not only influences the adjacent Ca^{2+} binding site 1 but also globally enhances cooperative activation of the receptor in response to alterations in extracellular Ca^{2+}.

The crystal structures (Geng et al., 2016) of the entire ECD region of CaSR in the resting and active conformations have provided additional information about the dynamics between calcium and L-amino acid binding (Geng et al., 2016). Most importantly, by using L-Trp, the study provided direct evidence that L-amino acids are CaSR co-agonists, and they act concertedly with Ca^{2+} to achieve full receptor activation. Several lines of evidence support this contention: 1) L-Trp binds at the interdomain cleft of the VFT, which is a canonical agonist-binding site for class C GPCRs (Kunishima et al., 2000; Muto et al., 2007; Geng et al., 2016) and shares a common receptor-binding mode with the endogenous agonists (amino acids or their analogs) of mGluR and $GABA_B$ receptors, (Kunishima et al., 2000; Tsuchiya et al., 2002; Muto et al., 2007; Geng et al., 2016). 2) L-Trp interacts with both LBD1 and LBD2 in ECD to facilitate its closure, a crucial first step during CaSR activation. In contrast, no Ca^{2+} ion is found at the putative orthosteric agonist-binding site to induce domain closure. 3) Mutations of L-Trp-binding residues (S147A, S170A, Y218A, and E297K) severely reduced Ca^{2+} induced IP accumulation and intracellular Ca^{2+} mobilization (Zhang et al., 2002; Silve et al., 2005), indicating that L-Trp is required for a Ca^{2+} induced receptor response. Notably, the presence of extracellular Ca^{2+} above a threshold level is required for amino-acid-mediated CaSR activation, amino acids increase the sensitivity of the receptor toward Ca^{2+}. Taken together, amino acids and Ca^{2+} ions act jointly to trigger CaSR activation.

Knowing that aromatic L-amino acids (Trp, Phe, His) are important tastants in kokumi flavor, CaSR becomes more relevant for taste biology. Moreover, the kokumi tripeptide, glutathione (GSH), and glutamyl peptide are suggested to bind allosterically to CaSR at the same site as L-amino acids (Wang et al., 2006; Broadhead et al., 2011) and enhance its activity in the presence of 0.5–1 mM free calcium, thereby acting as a positive allosteric modulator. In addition, an ECD crystal structure might help to explain the structural and molecular details of the GSH binding pocket such as the nature of critical residues and their binding characteristics. Given recent reports of calcium emerging as a taste modifier, it would be worth investigating how GSH and Ca^{2+} operate in kokumi human perception.

28.18.8.1.1: Allosteric Modulation of Calcium-Sensing Receptor

Classically CaSR is known to be involved in the pathophysiology of parathyroid and renal-related diseases by sensing calcium ions in the extracellular fluid (Brown, 2007). Research on related therapeutic applications has identified several classes of PAMs and NAMs that modulate CaSR agonist sensitivity. More recently this has been applied to kokumi taste signal transduction.

28.18.8.2: Endogenous Modulators (L-amino Acids, Anions, and Glutathione Analogs)

Several studies based on molecular modeling and mutagenesis report L-amino acids (L-Phe, L-Tyr, L-His, and L -Trp) as PAMs because they enhance the Ca^{2+}-induced response of CaSR. Aromatic L-amino acids bind in the VFT domain (Mun et al., 2004) and require a highly conserved five residue binding motif (S147, S170, D190, Y218, and E297) (Conigrave and Hampson, 2006; Geng et al., 2016). Among these residues, E297 was identified through the natural mutation E297K as essential for structural and functional activity (Table 8) (Pollak et al., 1993; Bai et al., 1998; Conigrave et al., 2000; Zhang et al., 2002; Mun et al., 2004).

As recently identified NAMs, anions SO_4^{2-} and PO_4^{3-} are important modulators of the Ca^{2+}-induced response. They bind in the VFT region and act as moderate NAMs for CaSR activity (Geng et al., 2016; Centeno et al., 2019). Based on anomalous difference maps, four anion-binding sites were identified in the inactive and active CaSR ECD structures. Sites 1 and 3 are located above the interdomain cleft in LBD1, while site 4 lies in the LBD2 region. Sites 1 and 3 appear to stabilize the inactive conformation while site 2, which is present in both active and inactive conformations appears important for receptor function as mutations in its residues (R66H, R69E, and S417L) abolished the Ca^{2+}-induced response. In addition, each protomer structure contains one Ca^{2+} ion and three SO_4^{2-} ions which together contribute to the structural integrity of the receptor (Geng et al., 2016). Taken together, anions along with Ca^{2+} and amino acids are involved in an intricate interplay for CaSR activation to maintain conformational equilibrium between inactive and active states.

As positive allosteric modulators, γ glutamyl peptides including glutathione (γGlu-Cys-Gly) and its analogs (Table 8) are predicted to have overlapping binding sites with L-amino acids in the VFT region (Wang et al., 2006; Ohsu et al., 2010; Broadhead et al., 2011). Kokumi peptides that activate CaSR resemble amino acids in having free α-amino and free α-carboxylate groups because they contain both amide bond formation between the γ-carboxylate group of L-glutamate and the α-amino group of its neighboring Cys residue. However, compared to amino acids, glutathione analogs have much larger side chains and are more potent activators of CaSR (Wang et al., 2006). Nonetheless, free sulfhydryl is not required for CaSR activation (Ohsu et al., 2010; Maruyama et al., 2012).

The crystal structure of ECD enables mapping of the GSH binding site and investigation into how GSH binding works in synergy with Ca^{2+} to modulate the kokumi sensation. NPS2143, the sole kokumi NAM identified to date has been reported to inhibit kokumi taste sensation to GSH and its analogs which provides an opportunity to screen for novel kokumi-enhancing molecules in a cell-based assay.

28.18.8.3: Synthetic Drugs as Allosteric Ligands of Calcium-Sensing Receptor

Because of its pathophysiological importance, various synthetic PAMs and NAMs of CaSR have been identified and are in clinical use. The allosteric modulation of CaSR by synthetic drugs has been recently reviewed (Hannan et al., 2016; Chavez-Abiega et al., 2020; Leach et al., 2020). Since the 1990's the terms calcimimetics and calcilytics, have been used for drugs that mimic or antagonize the effect of extracellular Ca^{2+} on CaSR activity, respectively. Pharmacologically, a calcimimetic activates the CaSR and includes agonists (type I) and allosteric ligands (type II). Most type I calcimimetics are either inorganic or organic polycations (e.g., Mg^{2+}, Gd^{3+}, neomycin), whereas type II calcimimetics are small naturally occurring molecules (aromatic amino acids or GSH) or synthetic drugs and peptides (NPS R-568, cinacalcet). Type II calcimimetics (like aromatic amino acids) bind in the ECD while others (e.g., NPS R-568, NPS R-467) bind in the TMD of the CaSR. Calcilytics are thus small organic molecules that appear to act as NAMs and bind in the TMD of the receptor (Widler, 2011; Nemeth, 2013).

Homology modeling and mutational studies show that both PAMs and NAMs have overlapping but non-identical binding sites in TMD and can partially allosterically modulate CaSR activity in the complete absence of the ECD, but their potencies vary among structurally different compounds (Collins et al., 1998; Ma et al., 2011) (Table 8). Several residues reportedly critical for allosteric modulation, $W818^{6.48}$, $F821^{6.51}$ (TMD6) and $E837^{7.39}$, $I841^{7.43}$ (TMD7), $R680^{3.28}$, $F684^{3.32}$, $F688^{3.36}$ (TMD3) impair calcimimetic and calcilytic induced CaSR signaling (Miedlich et al., 2004; Petrel et al., 2004; Leach et al., 2016). Nevertheless, subtle differences in ligand–receptor interactions drive negative vs. positive modulation of CaSR signaling, by NPS2143 or cinacalcet and NPSR-568, respectively (Miedlich et al., 2004; Leach et al., 2016; Keller et al., 2018). The details of CaSR allosteric modulation by synthetic drugs is out of the scope of the current review, for a comprehensive explanation refers to these studies (Chaves-López et al., 2014; Hannan et al., 2016; Leach et al., 2020).

28.18.8.4: Conclusion

Taste GPCR research has advanced rapidly over the past two decades providing a more thorough understanding of receptor molecular pharmacology and signal transduction pathways. Except for the kokumi receptor ECD, high-resolution crystal structures for any taste receptor would be a major step toward designing novel and potent surrogate taste receptor ligands and selective antagonists. This has been a challenge due to low taste GPCR functional heterologous expression, appropriate post-translational modifications, high conformational flexibility, and low detergent stability. However, significant advancements in structural biology technologies of serial femtosecond crystallography using X-ray free-electron lasers and high-resolution cryo-electron microscopy provide promising tools for understanding conformational dynamics and visualizing the process of receptor activation with high spatial and temporal resolution. The physiological relevance of taste GPCRs will be further advanced through *in vivo* studies to help provide information on potential synergies in taste signal transduction mechanisms, particularly among bitter, umami, sweet, and kokumi receptors.

28.19: Signal Transduction - Temperature

This chapter section is taken in entirety from: **Ion Channels and Thermosensitivity: TRP, TREK, or Both?** Lamas et al. *Int. J. Mol. Sci.* **2019**, *20*(10), 2371; **https://doi.org/10.3390/ijms20102371.** Creative Commons Attribution (CC BY) license (**http://creativecommons.org/licenses/by/4.0/**).

28.19.1: Introduction

Mammals and other animals spend large amounts of energy in maintaining a nearly constant body temperature, irrespective of the temperature of the environment. The mechanisms controlling thermal regulation are complex and often rely on negative feedback, where it is first necessary to determine the body and ambient temperature. The temperature of the environment can be sensed by external receptor cells, mainly located in the skin, whereas body temperature is sensed by internal receptors expressed by cells located in several internal organs. Traditionally, only the skin and core thermoreceptors (spinal cord, hypothalamus) have attracted the attention of researchers, but more recently, some very interesting information has emerged regarding visceral thermal receptors, even in humans. Although a hypothesis conceived many years ago, the terminals of receptor neurons are thought to contain branches of nerve fibers without any apparent structural specialization. Indeed, only recently have we begun to understand the molecular basis of thermoreception by cells.

Many biochemical processes like chemical reactions, and physical processes like conformational changes, are extraordinarily dependent on temperature, and although these processes generally occur faster at higher temperatures, the relationships can be very complex [3]. If we consider the nervous system (NS), the effects of temperature on the resting membrane potential (RMP) were the first to be studied, as were its effects on the kinetics and speed of compound and single action potentials, long before the existence of ion channels was demonstrated.

All neurons and ion channels are affected by changes in temperature, not least because channel gating is generally a temperature-dependent process. However, only some neurons can be called thermoreceptors and very few ion channel types can be designated as thermosensors. In general, only channels with a temperature coefficient (Q10) ≥2–5 are considered temperature dependent. (Q10 is the ratio of a reaction at two different temperatures that differ by 10^{o} C. See Chapter 32.11 for more details.) Thermoreceptors are sensitive to changes in temperature rather than to the value of the temperature itself, probably due to their characteristic strong adaptation. These receptors are classified into two groups depending on whether their discharge frequency increases when they are heated or cooled (**Figure 1**). Based on this classification, it is common to speak of four thermal sensations (cold −10 to 15 °C, cool 16–30 °C, warm 31–42 °C and hot 43–60 °C), whereby cold and hot are potentially noxious and/or painful [11,12].

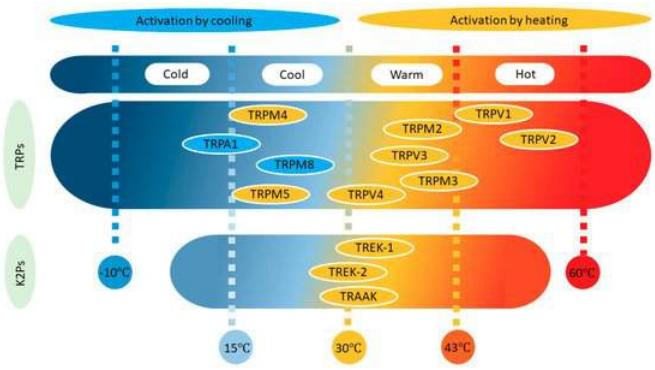


Figure 1. Distribution of transient receptor potential (TRP) and TWIK-related potassium (TREK) channels as a function of their temperature threshold. Note that while TREK channels are activated by increases in temperature (orange), TRP channels may also be activated by lowering the temperature (blue).

The modulation of TWIK-related potassium (TREK) channels by temperature has been touched on in several reviews, yet very few have dealt exclusively with this exciting topic. Conversely, after transient receptor potential (TRP) channels sensitive to temperature were discovered, they were studied extensively to understand how thermal stimuli were transduced. Such interest led to the appearance of good reviews covering this issue. In this review, we will focus on the less well-known role of TREK channels in thermosensation, and we will compare the behavior of these channels to that of TRP channels. Other thermosensitive proteins have also been described, like the Na/K ATPase and ENaC channels, or P2X receptors, and while these should also receive attention, we consider this to fall beyond the scope of this review. Indeed, cell thermosensitivity seems to be governed by the interplay of a number of channel types, as reported in hypothalamic neurons.

28.19.2: TREK Channels

The TWIK-related potassium channel (TREK) subfamily belongs to the two-pore domain potassium channels family (K2P) and is comprised of three members: TREK1, TREK2, and TRAAK (TWIK-related arachidonic acid-activated potassium channel). These are background potassium channels characteristically modulated by several physical and chemical stimuli, such as membrane stretch, pH, unsaturated fatty acids, general anesthetics, and temperature. In general, TREK channels display very weak activity at room temperature and normal pressure, even when overexpressed in heterologous systems. However, their activity increases strongly when a number of different stimuli are applied, including an increase in temperature. From a physiological point of view, it is important to note that at 37 °C, all three members of the TREK subfamily respond to stimuli (pH, membrane stretch, or arachidonic acid), much like they do at room temperature. TREK channels may fulfill a dual role in the transmission of thermal pain. Thus, their strong activation by noxious heat results in an outward current that provokes membrane hyperpolarization and a reduction of thermoreceptor firing, provoking heat-pain relief. Conversely, inhibition of TREK channels by noxious cold should depolarize thermoreceptors and increase their excitability, cooperating in the transduction of noxious cold sensations (see **Figure 2**).

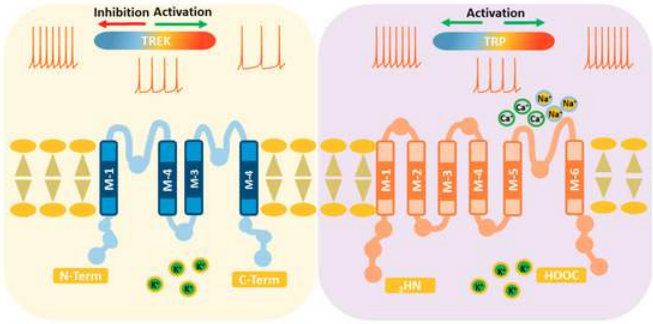


Figure 2. Proteins forming TREK channels have four transmembrane segments and two pore domains so that they assemble as dimers. Activation of TREK channels by increasing the temperature reduces the excitability of thermoreceptors (left). In contrast, TRP channels possess six transmembrane segments and only one pore domain, so they are tetramers. Activation of TRPs increases the excitability of thermoreceptors, but some are activated by an increase and others by a decrease in temperature (right).

28.19.2.1: TREK1

Soon after their discovery, it was shown that TREK1 channels are strongly and reversibly activated by an increase in temperature when expressed in heterologous systems (cell lines derived from kidney (COS) cells, oocytes). If we consider that these are mostly voltage-independent channels open at resting potentials, TREK1 channels should function as cold sensors because low temperatures would dampen their activity and depolarize these thermoreceptors (see **Figure 2**). Many authors have demonstrated that macroscopic TREK1 currents are strongly outwardly rectifying at room temperature. While the outward current is not evident at 12 °C, it is strongly enhanced at 37 °C, and the current progressively increases as the temperature gradually augments. Indeed, the current increases around 7-fold with an increase of 10 °C in the range of 14 to 42 degrees, and importantly, maximal sensitivity (0.9-fold per degree) was reached at nearly physiological temperatures, between 32 and 37 °C. The current induced by heating is also outwardly rectifying, and it reverses at potentials close to the equilibrium potential for potassium (E_K). In heterologous systems, the activation of TREK1 by temperature may be reversibly inhibited by cAMP, and this inhibition is suppressed by mutation of the C-terminal region that harbors a phosphorylation site for protein kinase A (PKA). Moreover, chicken embryonic atrial myocytes express TREK-like currents, and they have a resting membrane potential of around −20 mV in culture, which increases to −70 mV when the temperature rises to 35 °C, a change that was ascribed to the activation of TREK1/2 channels. In voltage-clamp, the outward current recorded at +60 mV increased 9-fold. Both TREK1 and TWIK-related arachidonic acid-activated potassium (TRAAK) channels have been proposed to shut down the firing of hippocampal neurons when the temperature rises too high.

Figure 28.19.*a* shows an interactive iCn3D model of the mouse temperature sensitive K2P2.1 (TREK-1) potassium channels (6W84)

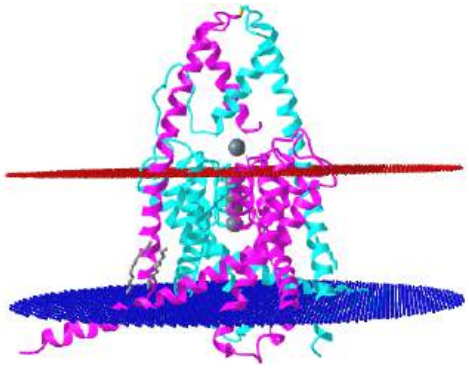


Figure 28.19.*a*: Mouse temperature sensitive K2P2.1 (TREK-1) potassium channels (6W84). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...U28PvTKakHpHW9

The red dots represent the outer leaflet, The gray spheres are potassium ions.

The threshold for the activation of slowly conducting C-fibers by noxious heat (30–50 °C) recorded in a skin-nerve preparation decreases in TREK1 KO mice, and the range of activation of these fibers by heat corresponds closely to the range in which TREK1 is activated (30–45 °C). The number of action potentials in response to a heating ramp (30–50 °C) was higher in the KO mice, although the response of C-fibers to a cooling ramp (32–10 °C) was similar in native and KO mice. Indeed, TREK1 KO mice were hypersensitive to thermal pain up to 50 °C but not at higher temperatures (52–56 °C), indicating that TREK1 channels may be important for the perception of low-threshold but not high-threshold thermal stimuli, for which TRP channels may be more important. Accordingly, the proportion of small diameter, cultured, dorsal root ganglia (DRG) neurons that respond to noxious heat (34%) increases in TREK1 KO (64%) and TREK1/TRAAK KO (74%) mice, as does the proportion of heat-responsive C-fibers in nerve-skin preparations from the single and double KO. TREK1 and TRAAK channels may counteract the stimulatory effect of heat-activated TRP channels in pain-transducing fibers when temperatures increase, such that the overall response may reflect a balance of the activity of these two functionally contrasting channel types. The threshold of thermoreceptors should certainly increase in the presence of TREK channels as temperatures increase.

Cooling of DRG neurons in culture from 32 to 20 °C induces a depolarization of about 10 mV and often the firing of action potentials, an effect shown to be due to the inhibition of a background potassium current. Accordingly, the inhibition of a native TREK1-like current may underlie the excitation (depolarization and firing) produced by cold in small, cultured, trigeminal ganglion (TG) neurons (see **Figure 2**). Interestingly, cold induces subthreshold oscillations in cold-sensitive DRG neurons. Transduction seems to be rather complex, involving the dampening of a hyperpolarization-activated cationic current and a permissive role for a slowly inactivating potassium current. Interestingly, the TREK1/TRAAK double KO mutant shows a consistent cool allodynia (pain due to a stimulus that does not normally provoke pain), and oxaliplatin, a cancer therapy that causes peripheral nerve neuropathy, exacerbates cold sensitivity in many patients and animals, inducing allodynia to cool temperatures.

Neither the deletion of TREK1 nor TRAAK increases the fraction of small DRG neurons sensitive to noxious cold stimuli (below 20 °C and down to about 10 °C), although the TREK1/TRAAK KO and the triple TREK1/TREK2/TRAAK KO showed a significant increase in such neurons. Similar results were obtained when recording C-fibers in a skin-nerve preparation, in which case the double KO C-fibers fired more strongly than the single TRAAK KO and wild-type fibers. Oxaliplatin also induced hypersensitivity to noxious cold temperatures, while double and triple KO mice but not the TREK1 KO mice are hypersensitive to cold, which is not further affected by oxaliplatin. Hence, the deletion of two of the three TREK channels appears to be sufficient to reach maximal hypersensitization. The neuroprotective agent riluzole induces an analgesic effect against painful cold in normal animals, but also in oxaliplatin-pretreated TREK2 KO and TRAAK KO animals. However, riluzole did not affect pain sensitivity in TREK1 KO animals treated with oxaliplatin, in animals treated with the TREK1 inhibitor spadin, or in untreated TREK1 KO mice or triple KO animals. Similarly, a presumed TREK leak outward current recorded in DRG neurons was inhibited by riluzole and fluoxetine at 22 and 30 °C but not at 14 °C, probably because the current was already inhibited at low temperatures. Together, these

experiments suggest that TREK1 channels fulfill an essential role in the perception of noxious cold and that TREK1 and TRAAK channels work together in sensing cold.

Cell-attached patches demonstrated that the basal activity of expressed TREK1 channels is insignificant at room temperature, gradually increasing as the temperature rises (17-fold for an increase of 20 °C) and with a threshold around 25 °C. The current activated by temperature also displays outward rectification and reverses around the equilibrium potential for K^+ [31], although the single-channel conductance remains unaffected TREK1-like channels naturally expressed in cardiac ventricular myocytes and DRGs and recorded in cell-attached patches, do not open at 24 °C, yet they are very active at 37 °C Surprisingly, temperature increases fail to modulate TREK1 activity in outside-out and inside-out patches, but under the same conditions, TREK1 is still strongly activated by arachidonic acid

TREK1 channels are ideally positioned to act as thermosensors because they are expressed in structures related to thermosensitivity and thermoregulation such as DRGs, the TG, nodose ganglia (NG), or the anterior and preoptic hypothalamus .

28.19.2.2: TREK2

Heterologously expressed TREK2 channels also produce strong outward rectification when recorded in whole-cell configuration at room temperature, which increases greatly at temperatures around 37 °C in several heterologous systems In COS cells, a small TREK2 current was observed at 0 mV that augmented progressively with a gradual rise in temperature to about 40 °C. Notwithstanding, the response of TREK2 to abrupt changes in temperature was rapid. Importantly, the IVs of the TREK2 current at different temperatures (24 and 37 °C) showed that the effect of temperature was not voltage-dependent: both inward and outward currents increased to the same degree. In this range of temperatures, the current increased 14-fold per 10 °C, indicating a very strong temperature dependence that was even bigger than that of TREK1. Much like TREK1, TREK2 responds to temperature changes around the physiological range, with current activated reasonably well at 37 °C and at resting membrane potential (RMP). Most experiments on TREK channels have been carried out at room temperature and at 0 mV. However, in the future these currents should be investigated using more physiological parameters, around a resting potential and 37 °C, providing a more precise idea of their role in the behavior of central neurons

Figure 28.19.*b* shows an interactive iCn3D model of the human two-pore domain temperature-sensitive potassium ion channel TREK2 (K2P10.1)(4BW5)

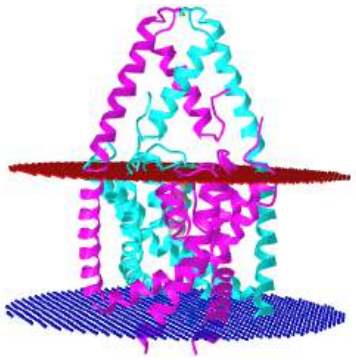


Figure 28.19.*b*: Human two-pore domain temperature-sensitive potassium ion channel TREK2 (K2P10.1)(4BW5). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...wSt9GE2pbDwfq5

Cerebellar granule and DRG neurons expressed native TREK2-like channels with weak activity at 24 °C in cell-attached patches, yet when the temperature increased to 37 or 41 °C they became very active at all voltages (−80 to +80 mV). Moreover, cultured cortical astrocytes have TREK2-like whole-cell outward currents that are strongly enhanced in the temperature range of 23–40 °C. Interestingly, ischemia significantly augmented the outward current provoked by an increase in temperature in these astrocytes In addition, it was recently reported that TREK2 channels contribute about 10 mV to the RMP of DRG neurons at about 30 °C Furthermore, single TREK2 and triple TREK1/TREK2/TRAAK KO mice were more sensitive to warm temperatures (40–42 °C) when tested with the tail-flick reflex.

Using a skin-nerve preparation, it was demonstrated that the proportion of heat-sensitive C-fibers and their activity (the number of action potentials) increased in the TREK2 and triple KO mice when temperatures rose to noxious heat levels (ramped from 30 to 50 °C), whereas the temperature threshold for firing decreased At high temperatures (between 40 and 50 °C), the triple but not the single KO fibers were more active than their wild-type counterparts, indicating that TREK2 regulates C-fiber responses at

temperatures below 40 °C, while at higher temperatures other TREK channels participate in these responses. Both KOs suffered hyperalgesia at temperatures around 45 °C, but only the triple KO showed the same behavior above this temperature.

The withdrawal latency in the tail immersion test was reduced in both the TREK2 KO and the triple KO mice when innocuous cooling temperatures were tested (20–25 °C). As such, the KOs show enhanced sensitivity to temperatures in the normal range, and similar results were obtained in a temperature preference test. The percentage of C-fibers responding to moderate cold (30–21 °C) was higher in single and triple KOs when compared to those recorded from the nerve-skin preparation of wild-type mice. Interestingly, the cold threshold for C-fiber firing (21 °C) was lower in the triple KO (24 °C) but not in the TREK2 KO (23 °C) mice [63]. Moreover, oxaliplatin induces mice to spend more time on a hot plate (30 °C) than on a cold plate (20–25 °C) when compared to untreated animals, indicating that neuropathic mice have enhanced sensitivity to moderate cold [63]. It has been suggested that TREK2 is implicated in the neuropathic hypersensitivity induced by this drug and indeed, oxaliplatin almost halved the TREK2 mRNA detected in DRG neurons. Generally, the data suggest that TREK2 channels may be essential for the control of the C-fiber response to cold at moderate temperatures. The tail immersion test showed that triple KO mice were hypersensitive to noxious cold temperatures (15–5 °C), while the single TREK2 KO mice behaved much like the wild-type mice. Moreover, very similar results were obtained in the nocifensive dynamic cold plate test. Accordingly, it was suggested that TREK2 may not be important in noxious cold sensitivity but that it might be essential for thermoreception at moderately cool temperatures (25–20 °C).

A clear, fast, and reversible increase in activity was also reported for single TREK2 channels in cell-attached patches held at −40 mV when the temperature increased (24 to 37 °C), with a threshold for this increase at 25 °C (from 24 °C) and not affecting the conductance. It should be noted that in these circumstances, the activity of TREK2 single channels was very low at 24 °C. Significantly, neither TASK3 nor TRESK2 showed such dependence on temperature. However, like TREK1, the activity of TREK2 in inside-out patches was not modified by changes in temperature (24 to 42 °C). Finally, it is important to consider that TREK2 channels are expressed strongly in the DRG, TG, and hypothalamus, yet less than TREK1 in the NG.

28.19.2.3: TRAAK

Much like the other members of the family, TRAAK currents showed a strong open-channel outward rectification when recorded in whole-cell configuration, and these currents increase strongly when the temperature rises (24 to 42 °C). Moreover, the percentage of small-diameter DRG neurons responding to noxious heat, in culture, is increased in TRAAK and TRAAK/TREK1 KO mice. Consistently in skin-nerve preparations, the percentage of fibers responding to heating (30–50 °C) and the number of action potentials in response to a heating ramp also clearly increases, while the firing threshold is reduced. Notably, TRAAK and TRAAK/TREK1 KO mice suffer heat hyperalgesia when evaluated in the tail immersion test in the 46–50 °C range. Moreover, the double but not the single KO also shows hypersensitivity at higher temperatures (52–56 °C) in the hot plate test.

Knock-out of TRAAK did not modify the percentage of DRG neurons in culture that respond to noxious (12 °C) cold. Moreover, in the cold plate assay, TRAAK KO mice behave like wild-type mice, whereas TREK1/TRAAK KO mice are more sensitive to cooling in the 10 to 20 °C range. The activity of single TRAAK channels heterologously expressed in COS cells and recorded in cell-attached patches at −40 mV was very low at 24 °C, yet it increased progressively as the temperature rose from 24 to 37 °C. The threshold for activity was around 30 °C, slightly higher than that reported for TREK1 and TREK2. However, the behavior of TRAAK channels in inside-out patches mimics that of TREK1 and TREK2 such that their activity was not affected by changes in temperature (from 24 to 42 °C). Native TRAAK-like channels in DRG neurons displayed little activity at room temperature, but there was clear activity in all cell-attached patches at 37 °C [36, 48]. Finally, TRAAK channels are clearly expressed in the hypothalamus, TG, and DRG , yet they are only weakly expressed in the NG.

28.19.2.4: Molecular Origin of Thermosensitivity

When first discovered, mouse TREK1 was reported to have four transmembrane segments, two pore domains, and a sequence of 370 aa [37]. The activation of heterologously expressed TREK1 currents by increasing temperature is unaffected by the deletion of the cytoplasmic N-terminal region. By contrast, partial deletion of the C-terminal region (Δ103) or replacement of this region with that of TASK1 strongly dampens the activation of TREK1 by heat. The sensitivity of the TREK1 channels to temperature can be eliminated by mutating helix 1 of the pore (G137I), suggesting that temperature affects the TREK1 and TREK2 channels by manipulating the C-type gate. It was suggested that functional coupling between the C-terminal domain and the C-type gate through the M4 segment is crucial for the heat sensitivity of the TREK1 channel. Thus, it is tempting to speculate that increasing the affinity of the C-terminal domain for phospholipids of the inner leaflet would increase the activity of TREK1 by heat, as proposed for other stimuli like stretch, PUFAs, phospholipids, or pH. Conversely, dissociation of this domain from the membrane would result in TREK1 inhibition. Surprisingly, the replacement of the C-terminus of TREK2 with that of TASK3 did not reduce

the sensitivity of the channel to changes in temperature in the range of 24 to 37 °C under similar conditions, although it became insensitive to pH and arachidonic acid. Heat enhances the activity of TREK1, TREK2, and TRAAK in whole-cell and cell-attached recordings, yet not in outside- and inside-out patches, indicating that the integrity of the cell, and probably also a second messenger, are necessary for this modulation. The contribution of TREK channels to maintaining the RMP has often been questioned; however, this assertion is mostly based on experiments carried out at room temperature. Thus, new experiments should be performed at physiological temperatures to ascertain the role of these channels on both the RMP and neuronal excitability.

28.19.3: TRP Channels

Six transient receptor potential (TRP) channels are considered thermosensors, four of them responding to heat and two to cool. Temperature-sensitive TRP channels (Thermo-TRP) are extremely dependent on temperature, showing very high Q10 values (>20).

28.19.3.1: Heat-Sensitive TRP Channels

Four TRP subtypes are activated by an increase in temperature (**Figure 1**). Two of them respond to warm stimuli (TRPV4 Warm >27 °C and TRPV3 Warm >34 °C), and the other two to hot-painful stimuli (TRPV1 Hot >43 °C and TRPV2 Hot >52 °C).

TRPV1s are voltage- and temperature-dependent channels that display outward rectification when expressed in human embryonic kidney (HEK) cells and that is strongly enhanced by heating (to 48 °C) and by capsaicin.

Figure 28.19.*c*shows an interactive iCn3D model of human TRPV1 with capsaicin at 48 degrees Celsius in an open state (7LPE)

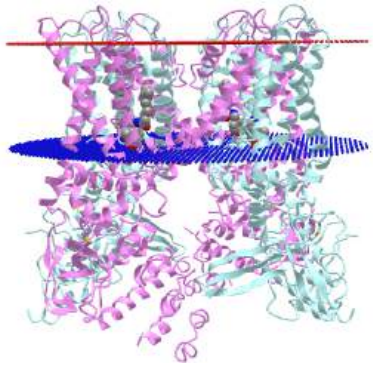


Figure 28.19.1: Human TRPV1 with capsaicin at 48 degrees Celsius in an open state (7LPE). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...7Hsf5zpDQoQu99

At room temperature, the current passing through these channels is negligible below 0 mV, but at 42 °C the channel activates more or less between −100 and +50 mV. These cationic channels are ten times more permeable to Ca^{2+} than to Na^{+} (P_{Ca}/P_{Na} ~10) and are thought to be sensors for noxious heat but not activated by innocuous heat. Indeed, the response to noxious heat in mice lacking TRPV1 (KO) or DRG neurons was clearly weaker, although other channels may also contribute to the perception of noxious thermal stimuli because heat still evokes receptor activation in several preparations. The NG sensory neurons that innervate the lungs produce an inward current in response to an elevation in temperature (from 23 to 41 °C, with a threshold around 35 °C and a Q10 of about 30 in the range of 35–41 °C) as well as membrane depolarization and action potential firing. This response was ascribed to the presence of TRPV1 channels, even though the participation of TRPV2-4 could not be ruled out. We obtained similar results with NG neurons in culture, although these were slightly more complex because a hyperpolarization was observed before the depolarization and firing (unpublished data). It is interesting to note that inflammatory mediators like ATP and bradykinin strongly reduce the threshold of TRPV1 activation (30 °C) such that warm temperatures become painful. TRPV1 is strongly expressed in small-diameter sensory neurons of the DRG, TG, and NG, but also in the hypothalamus, sites where they may exert an important role in thermoreception.

TRPV2 is activated at extremely high temperatures (52 °C), although it is not affected by capsaicin and shows an outwardly rectifying IV curve and a P_{Ca}/P_{Na} ~3 . This channel has a Q10 of around 100, and it is thought that the temperatures that activate TRPV2 are more harmful than those that activate TRPV1. These channels are strongly expressed by myelinated medium-large diameter DRG neurons (Aδ and Aβ), as well as in the hypothalamus and the NG.

TRPV3 channels are activated at warm, close to hot, temperatures (around 34–39 °C, with a Q10 around 6), generating currents with pronounced outward rectification and a P_{Ca}/P_{Na} ~12. They are capsaicin-insensitive channels but stimulated by camphor, and they are thought to be involved in thermosensation and thermal nociception. Indeed, it has been suggested that TRPV3 channels

contribute more to the speed with which mice select a more comfortable temperature than to the choice of the value of the temperature itself. By contrast, TRPV4 channels are more likely to be involved in choosing the preferred temperature from a non-painful range. Interestingly, it was proposed that TRPV3 channels transmit thermal stimuli through skin keratinocytes, which in turn will transmit this information to sensory endings. TRPV3 channels are expressed in sensory DRG and NG neurons but also in the hypothalamus. Interestingly, they co-localize with TRPV1 in DRG neurons.

TRPV4 are cationic (P_{Ca}/P_{Na} ~6) channels activated at even lower warm temperatures (around 27 °C, with a Q10 of about 10), generating outwardly rectifying currents and responding dynamically to temperature changes in the physiological range. These channels were proposed to play a role in thermosensation and thermoregulation although some authors were unable to activate these channels by increasing the temperature. Similarly, some behavioral studies reported a reduced response to temperature changes in TRPV4 KO mice, a behavior that was less clear in other studies. Much like TRPV3, the expression of these channels in keratinocytes was proposed to play an important role in the transmission of thermal information, which probably contributed to the controversy generated. The sensitivity of this channel to temperature is lost in excised patches, suggesting that it requires a soluble intracellular factor TRPV4 channels are expressed in DRG, TG, NG, and preoptic/anterior hypothalamic neurons, although in the hypothalamus they seem to be expressed in terminals rather than in the soma, such that their role in body thermoregulation is unclear.

TRPM2 (>35 °C), TRPM3 (>40 °C), TRPM4 (>15 °C), and TRPM5 (>15 °C) are channels that can also be activated by warming (**Figure 1**), yet they have received less attention, probably because it was initially thought that they were not expressed by somatosensory neurons or keratinocytes. TRPM2 is voltage-insensitive, shows a P_{Ca}/P_{Na} ~1, activates at 35 °C, and has a Q10 of around 15. TRPM3 is expressed broadly, generating an outwardly rectifying current, having a P_{Ca}/P_{Na} between 0.1 and 10, and activating at >40 °C with a Q10 of 7. It is important to say that TRPM3 has been described as part of a triad of TRPs, together with TRPV1 and TRPA1, involved in the transduction of acute noxious heat in mice. The combined ablation of these channels (triple KO) was necessary for the complete reduction of acute noxious sensing; single or double KO combinations resulted in deficits in heat responsiveness, but mice still conserved vigorous withdrawal responses to noxious heat. Heat activation of TRPM2 and TRPM5 was obtained in inside-out patches, suggesting a membrane-delimited mechanism. Interestingly, TRPM2 activation seems to result from the increase in the IV slope while that of TRPM4 and TRPM5 results from a shift of the activation curve to negative potentials. These last two channels are essentially not permeable to calcium.

28.19.3.2: Cold-Sensitive TRP Channels

Two TRP channels are activated by decreases in temperature (**Figure 1**), TRPM8 (<25 °C) activates in the cool range while TRPA1 (<18 °C) senses cold-painful temperatures. Similarly, cool fibers (Aδ and C) have activation thresholds at about 30 °C, and cold fibers (C) have activation thresholds <20 °C. Accordingly, two populations of TG neurons were described in terms of their activation threshold when temperatures decrease: 30 and 20 °C for a low and high threshold, respectively. In general, cold fibers fire continuously at normal skin temperatures and they increase their firing frequency when the skin is cooled down, or they shut down when the skin is warmed. In addition, cold fibers can adapt to small decreases in temperature.

TRPM8 channels are voltage-dependent cationic channels that are permeable to Na^+, K^+, Cs^-, and Ca^{2+} (P_{Ca}/P_{Na} ~3). When expressed in HEK cells and recorded in the whole-cell configuration, they show a voltage-dependent outwardly rectifying current that strongly increases upon cooling from 30 to 15 °C or through the application of menthol.

Figure 28.19.*d* shows an interactive iCn3D model of TRPM8 ion channel in complex with the menthol analog WS-12 and PI(4,5)P2 (6NR2)

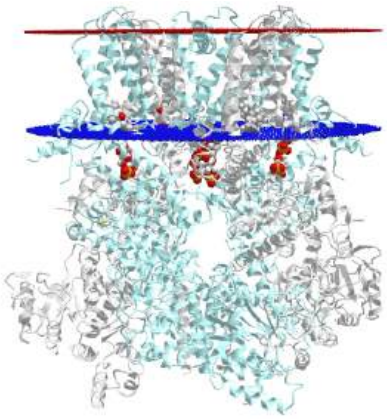


Figure 28.19.*d*: TRPM8 ion channel in complex with the menthol analog WS-12 and PI(4,5)P2 (6NR2). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...94muQHwYr7eij6

The methanol analog is shown in spacefill and CPK colors in the membrane bilayers. PI(4,5)P2 is shown in spacefill just below the lower cytoplasmic leaflet.

The TRPM8 receptor is a Ca^{2+} cation channel. Cooling compounds like menthol and WS-12 depend on allosteric interactions and membrane phosphatidylinositol 4,5-bisphosphate (PIP_2).

Importantly, both basal and cold-stimulated currents reverse around 0 mV and were almost negligible below this potential. Cooling CHO cells expressing TRPM8 (in the range of 25 to 15 °C) also induces an increase in intracellular calcium, and the Q10 in the range of 25 to 18 °C is around 24. The effect of temperature is due to an increase in the open probability and a shift in the conductance–voltage relationships along the voltage axis [9]. Similar results were obtained in inside-out macropatch recordings, although the stimulation occurred at lower temperatures, suggesting that the integrity of the cell is important but not indispensable. The role of this channel as a detector of painful cold has been questioned in experiments on KO mice, but it is accepted that it is an important cold sensor in vagal, TG, and especially DRG afferents. It was predicted that cold transduction may require the activation and inhibition of several different ion channels (see **Figure 3**), such as TRP, TREK, and ENaC channels. If this were the case, TRP channels would probably be more important in the noxious-cold range, whereas TREK channels might participate more strongly in the cool range of temperatures (**Figure 1**). TRPM8 is expressed in small-diameter DRG and TG neurons, presumably thermoreceptors, yet it seems not to co-localize with TRPV1.

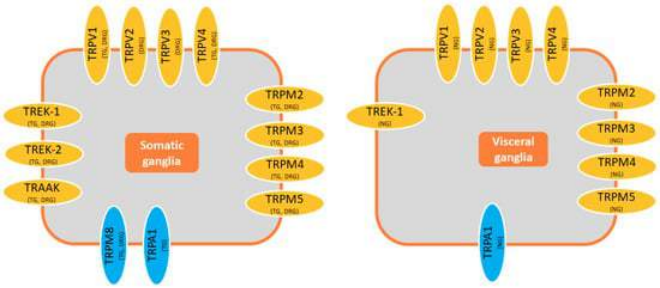


Figure 3. Most temperature-modulated channels are expressed in both somatic and visceral afferents. The expression of TREK2, TRAAK, and TRPM8 in visceral ganglia seems to be low when compared with other channels; however, colocalization of all these channels in the same neuron should be investigated to have a real picture of their relative importance in temperature sensing.

TRPA1 is activated by lower temperatures than TRPM8 (<18 °C), and while it would be expected to be involved in cold nociception, this is not that clear. TRPA1 generates an outward rectifying cationic current, both in control conditions and when cold activated (about 10 °C), with similar permeability for Ca^{2+} and Na^{+} (P_{Ca}/P_{Na} ~1). Cinnamaldehyde can selectively activate currents through this channel in native DRG neurons, as can bradykinin (when co-expressed with BK receptors), strongly suggesting a role in sensing nociceptive stimuli. However, TRPA1 KO mice do not seem to have difficulties in sensing cold stimuli through the skin, while the response of TRPM KO mice to cold is significantly dampened. By contrast, about 50% of NG neurons in culture were activated by cooling (<24 °C), mainly through TRPA1 channel activation (increase in $[Ca]_i$, depolarization, and AP firing). Interestingly, about 10% of the NG neurons responded to cold through a TRPA1- and Ca-independent pathway. TRPA1 often co-localizes with TRPV1, and in fact, this could explain the paradoxical hot sensation experienced with an extremely cold stimulus. Interestingly, most NG neurons sensitive to cold are also sensitive to heat. TRPA1 is expressed in DRG and TGs, while TRPA1 and

TRPM8 are not co-expressed in DRG neurons, but they are in TG neurons. In summary, the data available suggest that TRPA1 is the principal ion channel involved in cold sensation in visceral (NG) neurons, while TRPM8 would fulfill the same role in somatic neurons.

28.19.3.3: Molecular Origin of Thermosensitivity

The mechanism by which temperature modulates TRP channels is still unclear, yet several hypotheses have been proposed: (1) changes in temperature could produce a ligand that binds to a receptor and affects the channel; (2) changes in temperature could produce a structural change in the channel that provokes its opening; (3) temperature changes could affect the structure of the membrane, causing changes in tension that would, in turn, affect ion channels. Because capsaicin induces burning pain, it has been hypothesized that both capsaicin and heat may use a common mechanism to activate TRPV1 and produce pain. Both stimuli affect excised patches, and in general, it is accepted that TRPV1 is directly activated by noxious heat, so that it can be considered a true heat sensor.

The fact that TRPM8 can be activated by cooling in inside-out patches suggests that the mechanism is membrane delimited, also arguing against the participation of a second messenger pathway. Notwithstanding, inhibition of phospholipase C strongly dampened the increase in calcium provoked by cold stimuli in TRPM8-expressing CHO cells. Cooling activates TRPM8-expressed channels by causing a shift in the voltage dependence of activation to negative values, and the same mechanism is responsible for the activation of TRPV1 by heatbut not the activation of TRPM2. It has been proposed that temperature induces large rearrangements of the protein and thus, the existence of a temperature-sensing domain or “temperature sensor” in the structure of TRPM8 channels. Much like for TRPM8, inhibition of phospholipase C strongly reduces the increase in calcium provoked by cold stimuli in TRPA1-expressing CHO cells.

28.19.4: Conclusions

There are several important differences between the two main types of thermosensor channels that have been reviewed in this article (see **Figure 2**). First, TREKs are potassium channels with negative reversal potentials, such that their activation would result in a reduction in thermoreceptor activity. By contrast, as the reversal potential of TRPs (cationic channels) is close to 0 mV, their activation will result in increased thermoreceptor excitability. Second, the three TREK channels appear to increase their open probability as temperatures increase, while there are two possible situations in the case of TRPs: one in which its activity increases by increasing the temperature; and another in which activity increases when the temperature decreases (**Figure 2**). Although the activation of these channels generates opposing effects on thermoreceptors (depolarization versus hyperpolarization), the information available to date regarding the participation of TRP and TREK channels in thermosensitivity strongly suggests that both types of channels collaborate and complement each other to generate the sensations of heat, cold, and thermal pain. In support of this hypothesis, TREK, and TRP channels are very often co-expressed in thermoreceptors and other sensory neurons TRP channels are generally accepted as the primary thermosensors; however, several lines of evidence indicate that other channels are necessary to explain the full plethora of mechanisms involved in thermosensation. TREK2 channels appear to be important in thermoreception at moderate temperatures and sensing innocuous cold but not aversive cold, while TREK1 and TRAAK acting together may be important in sensing painful cold.

28.20: Signal Transduction - Pressure

This chapter section is taken in entirety from Fang, XZ., Zhou, T., Xu, JQ. *et al.* Structure, kinetic properties, and biological function of mechanosensitive Piezo channels. *Cell Biosci* **11,** 13 (2021). https://doi.org/10.1186/s13578-020-00522-z Creative Commons Attribution 4.0 International License, http://creativecommons.org/licenses/by/4.0/. We added iCn3D molecular models.

28.20.1: Introduction

Mechanotransduction, the process by which mechanical stimuli are converted into electrochemical signals, is essential for various biological processes, including neuronal cell development, pain sensation, and red blood cell volume regulation. As pivotal mechanosensors of in the mechanotransduction process, mechanosensitive (MS) ion channels have been found in organisms from bacteria to mammals. Extensive studies have revealed a variety of ion channels in eukaryotic cells that can sense various forms of mechanical forces (Table 1). These ion channels include transient receptor potential (TRP) channels and voltage-gated Na^+, $K^{+,}$ and Ca^{2+} channels, whose dysfunction may be associated with human genetic diseases. Notably, the MS candidates identified in invertebrates either have no homologs (e.g., TRPN) or no functional conservation (e.g., DEG/ENaC/ASIC) in mammals. Furthermore, most MS candidates (the TRP channel in particular) are activated not only by mechanical stimuli but also by chemicals, temperature, osmolarity, and heat (> 27–34 °C). Defining the molecular details of MS cation channels in mammals is therefore of paramount importance to understand the mechanotransduction process and find potentially novel therapeutic strategies for mechanosensitivity disorders.

Channel family	Channel isoforms	Ref.
TRP channels	TRPA1	[6]
	TRPC1	[7]
	TRPC6	[8]
	TRPV1	[9]
	TRPV4	[10]
	TRPM4	[11]
	TRPM7	[12]
	TRPN	[13]
	TRPP2	[14]
K^+ channels	Shaker (Kv1.1)	[15]
	Ca^{2+}-activated K^+ (BK)	[16]
	TREK1/2	[17]
	TRAAK	[18]
	HCN2	[19]
Na^+ channels	Nav1.5	[20]
Ca^{2+} channels	L-type	[21]
	N-type	[22]
	T-type	[23]
Cl^- channels	CFTR	[24]
OSCA protein family	*Sc*CSC1, *Hs*CSC1	[25]
DEG/ENaC superfamily	C.elegans MEC (MEC-4, MEC-10)	[26]
	ASIC	[27]
Other channels	TMC1/2	[28]

In 2010, Coste et al. revealed a novel family of mechanically activated (MA) cation channels in eukaryotes consisting of Piezo1 and Piezo2 channels, which have been proposed as the long-sought-after MS ion channels in mammals. The Piezo1 channel is present in nonsensory tissues, with particularly high expression in the lung, bladder, and skin; by contrast, the Piezo2 channel is predominantly present in sensory tissues, such as dorsal root ganglia (DRG) sensory neurons and Merkel cells. Since their discovery, tremendous effort has been made to reveal the structures and biological functions of Piezo 1 and 2. The partial molecular structure of a Piezo channel was determined by cryo-electron microscopy (cryo-EM). Furthermore, Piezo channels have been linked to various pathological and physiological processes, including erythrocyte volume regulation, cell division, and innate immunity. Moreover, Piezo channel mutations are associated with multiple hereditary human diseases, such as autosomal recessive congenital lymphatic dysplasia, hereditary xerocytosis (a rare, autosomal dominant congenital hemolytic anemia characterized by macrocytic stomatocytosis, and decreased red cell osmotic fragility due to a defect in cation permeability), and an autosomal recessive syndrome of muscular atrophy with perinatal respiratory distress. Considerable progress has been made toward characterizing the structural features, physiological significance, and biophysical properties of Piezo proteins. Given the importance of Piezo channels in understanding mechanotransduction processes, this review focuses on their structural details, kinetic properties, and potential functions as mechanosensors. We also briefly review the hereditary diseases caused by mutations in the Piezo genes, which is key to understanding their functions.

28.20.2: Structure of Piezo channels

Piezo proteins have an uncommonly large predicted size of approximately 2500 amino acids and encompass numerous transmembrane (TM) regions. Subsequent research has revealed that the mouse Piezo1 (mPiezo1) channel is an evolutionarily conserved pore-forming ion channel directly gated by membrane stretch. Several published cryo-EM studies have revealed that mPiezo1 exhibits a three-bladed, propeller-shaped homotrimeric structure that includes a central cap, three peripheral blade-like structures on the extracellular side, three long beams on the intracellular side that bridge the blades to the cap, and a TM region between these features (Fig. 1).

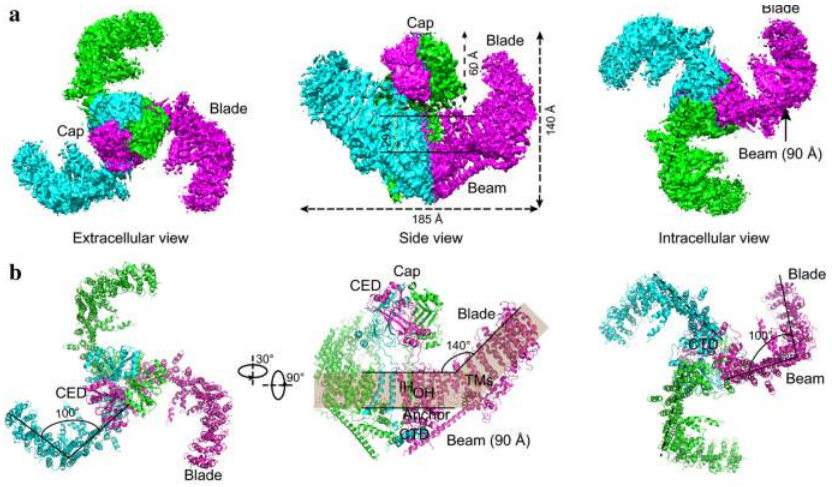


Figure 1: Cryo-EM structure of the mPiezo1 channel (adapted from Zhao et al.). **a** Multiple views of the sharpened map of the trimeric channel with the major domains labeled, with the three subunits colored red, green, and blue. **b** Cartoon model in which the three subunits are colored red, green, and blue. In the middle panel, the front subunit has been omitted to provide a better view of the curvature of the TMs

28.20.3: **Structure of the Piezo1 channel**

28.20.3.1: Unprecedented 38-TM topology

Piezo channels are predicted to possess an unusually large number of TM regions, ranging from 10 to 40. Zhao et al. recently produced high-resolution structures of mouse Piezo1 (mPiezo1), revealing a unique 38-TM topology in each subunit (Fig. 2a, b). The two TM regions (TM37 and TM38) closest to the center of the protein are designated as the inner helix (IH) and outer helix (OH), respectively, and enclose the transmembrane pore of the central pore module. The other 36 TM regions (TM1-36) form a curved blade-like structure with nine repetitive folds containing 4 TM regions each, named transmembrane helical units (THUs)

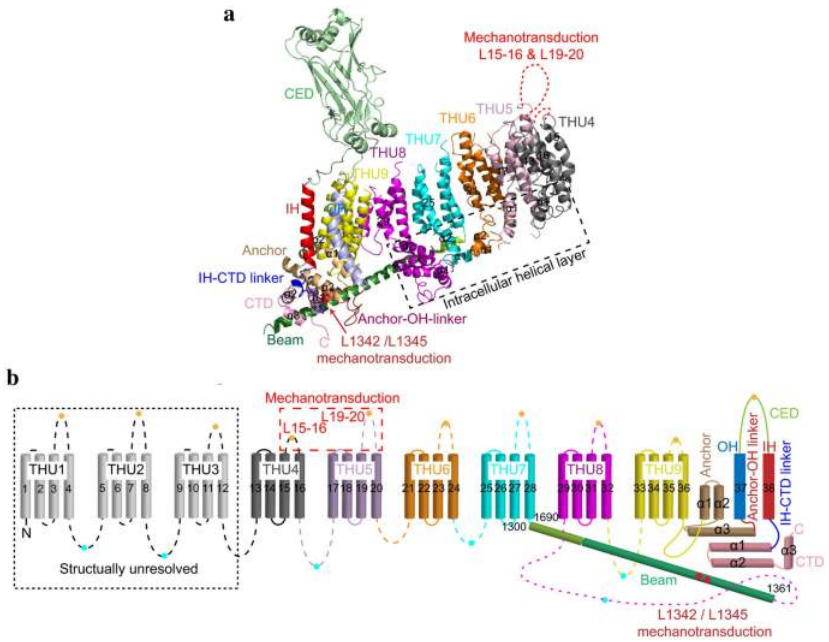


Figure 2: A 38-TM topology model and key functional sites in mPiezo1(adapted from Zhao et al.). **a** A model showing one subunit with individual THUs and featured structural components. Residues L1342 and L1345 in the beam are indicated by red spheres. **b** A 38-TM topology model color-coded to match the cartoon model in A

Figure 28.20.a shows an interactive iCn3D model of the mouse mechanosensitive Piezo1 channel (5Z10) (long load time)

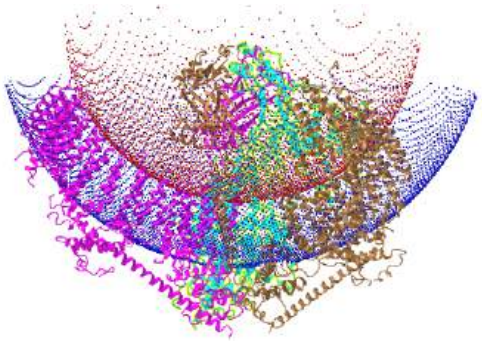


Figure 28.20.a: Mouse mechanosensitive Piezo1 channel (5Z10). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...ZcVpJczxvaaG79 (long load time)

28.20.3.1: Central cap

Kamajaya and colleagues [48] employed topological prediction modeling and found that residues 2210 to 2457 in Piezo1 form an extracellular loop following the last TM region from the C-terminus, defined as the C-terminal extracellular domain (CED) **(**Fig. 1**)**. The deletion of residues 2218 to 2453 from the Piezo1 protein abolished expression of the central cap, suggesting that this region trimerizes to form the central cap (Figs. 1 and 3). Further analysis revealed that the central cap consists of the CED in the form of a trimeric complex that encloses an extracellular vestibule (EV) with openings (Fig. 3).

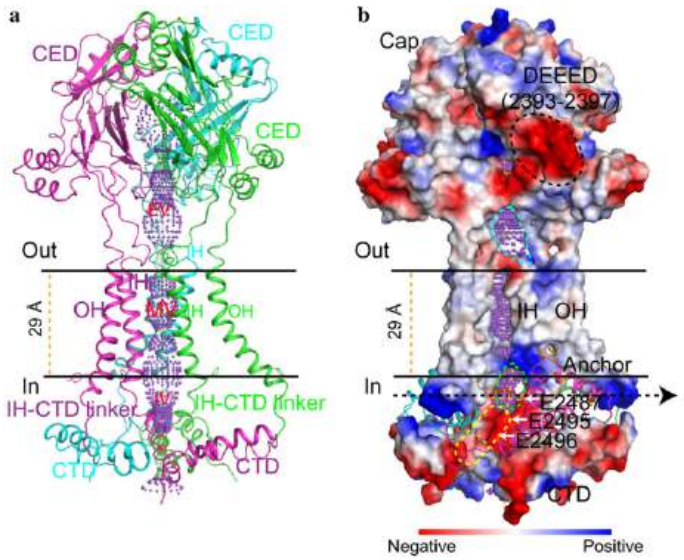


Figure 3. Structure of the central pore module (adapted from Zhao et al. [35]). **a** Ribbon diagram of the ion-conduction pore comprising the OH, CED, IH, and CTD from three color-coded subunits. The central solvent-accessible pathway is marked with a dotted mesh generated by the program HOLE. **b** Pore module presenting the surface electrostatic potentials showing negative (red) and positive (blue) potential. Extracellular and intracellular fenestrations are marked by cyan and green dashed lines, respectively. The lateral portal is marked by yellow dashed lines

28.20.3.1: Anchor

A hairpin structure, referred to as the anchor, connects the OH-IH pair to the C-terminal domain (CTD) plane, which moves the OH-CED-IH-containing region of one subunit into the neighboring subunit in a clockwise direction (Figs. 1 and 2). The anchor is made up of three helices (α1, α2, and α3). Helices α1 and α2 were found to organize into an inverted V-shaped structure, which maintains the integrity of the ion-conducting pore (Fig. 2b). In parallel with the membrane plane, the long α3 helix links to the OH via a lysine-rich anchor-OH linker that interacts with the polar residue-rich α2–3 turn in the anchor and the glutamate-rich region of the CTD. A few mutations in Piezo1 at locations including KKKK (2182-K2185), T2143, T2142 (T2127 in human Piezo1), R2514, E2523, and E2522, which are located in α3 in the anchor, have been reported to cause severe disease. Additionally, SERCA2, a Piezo-interacting protein, suppresses Piezo1 by acting on the anchor-OH linker. These findings support the structural and functional importance of the anchor region.

28.20.3.1: The long intracellular beam

On the intracellular surface, Piezo1 contains three beam-like structures 90 nm in length that are organized at a 30° angle relative to the membrane plane (Figs. 1 and 2). Residues H1300-S1362 form the beam structure. The large intracellular THU7-8 loop, which contains approximately 390 residues, might provide the beam with the structural basis for mechanical transmission. Functionally, the three long intracellular beams not only support the whole TM skeleton but also physically bridge the distal THUs to the central ion-conducting pore. When residues 1280 to 1360 (which form this beam structure) were deleted, the resulting mutant protein was absent, suggesting the structural importance of the beam.

28.20.3.1: Highly curved blades

The nine peripheral THUs in each subunit form blade-like structures, with each blade twisted clockwise (Fig. 1b). The proximal TM25–TM36 and peripheral TM13-24 interact at a 100° angle, as viewed from 90º relative to the plasma membrane plane, and a 140° angle, as viewed from a line parallel to the plasma membrane plane. Another important feature of the blades is the L-shaped helical structures formed by TM13, TM17, TM21, TM25, and TM29. Both identifiable structural features appear to be ideal not only for mechanosensation but also for the induction of local membrane curvature. Intriguingly, the peripheral TM13-24 appears to be within a highly curved membrane plane, indicating that the Piezo1 channel can curve the membrane in which it resides. This is consistent with past studies implying that Piezo1 ion channels can be regulated by cellular membrane curvature and tension.

28.20.3.1: The ion-conducting pathway

As pore-forming ion channels, Piezo proteins contain a trimeric ion-conducting channel made up of residues 2,189 to 2,547, which contain the last two TMs (Fig. 3). The continuous central channel consists of three parts, an EV within the cap region, a transmembrane vestibule (MV) within the membrane, and an intracellular vestibule (IV) underneath the membrane. Both the EV

and IV possess an opening that connects to the MVs, which are positioned above and below the membrane. Importantly, DEEED (2393–2397), a patch of negatively charged residues residing in the opening of the extracellular “cap” structure consisting of the CED, is required to ensure efficient ion conduction and determine the selection of cations over anions. Additionally, two critical acidic residues, E2495 and E2496, located at the CTD-constituted IV, may be responsible for divalent calcium ion selectivity, unitary conductance, and pore blockage.

28.20.4: **Structure of the Piezo2 channel**

Similar to Piezo1 channels, Piezo2 channels are large membrane proteins consisting of over 2,800 residues. However, the Piezo2 channel and Piezo1 channel share approximately only 42% sequence homology. Recent studies have shown that the overall structure of the Piezo2 channel is very similar to that of Piezo1 in that it forms a three-bladed, propeller-like homotrimeric structure comprising a central ion-conducting pore module and three peripheral blades with 38 TMs.

Figure 28.20.*b*shows an interactive iCn3D model of the mammalian tactile channel PIEZO2 (6KG7) (long load time)

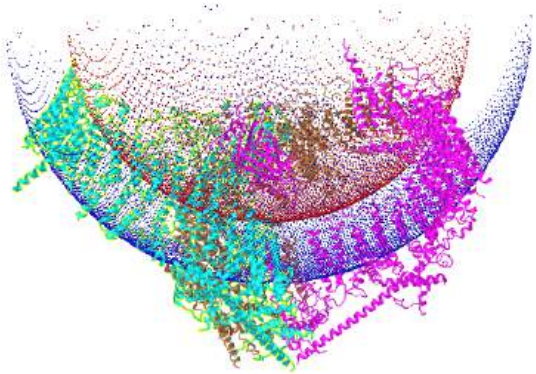
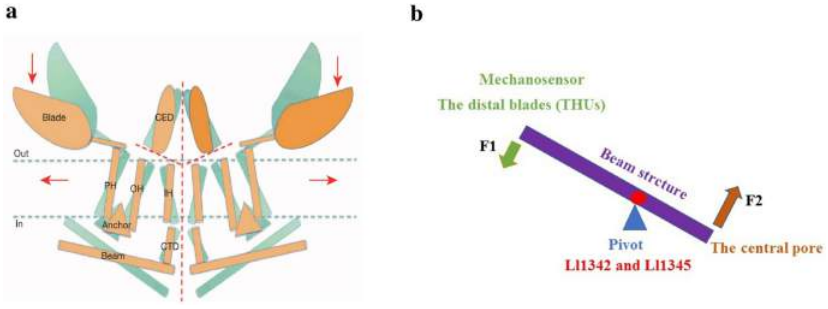


Figure 28.20.*b*: Mammalian tactile channel PIEZO2 (6KG7). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...4FaL71AL23KiA9 (long load time)

In the Piezo2 channel, the charged residues at the interface between the beam and the CTD are required to ensure the normal mechanosensitivity of the channel. Moreover, single-channel recordings indicated that a previously unrecognized intrinsically disordered domain adjacent to the beam acts as a cytosolic plug that limits ion permeation, possibly by clogging the inner vestibule in both Piezo1 and Piezo2. Furthermore, by structurally comparing the Piezo1 and Piezo2 channels, Wang et al. found that the Piezo2 channel has additional constriction sites at L2743, F2754, and E2757 that might serve as a transmembrane gate controlled by the cap domain.

28.20.5: Lever-like mechanotransduction mechanism

Based on the unique topological features of the mPiezo1 channel, a lever-like mechanotransduction mechanism to explain its extraordinary mechanosensitivity was proposed (Fig. 4). In the mPiezo1 channel, the curved blades composed of THUs can act as a mechanosensor, while the beam structure, with the residues Ll1342 and Ll1345 acting as a pivot, can act as a lever-like apparatus. Coupling the distal blades and the central pore through the lever-like apparatus converts mechanical force into a force used for cation conduction.



Adapted from Ge et al. **b** A lever-like mechano-gating model in Piezo1. The blue and red dashed arrows indicate input and output forces, respectively

Figure 4: Model of the lever-like mechanotransduction model. The curved blades can act as a mechanosensor, while the beam structure, with residues Ll1342 and Ll1345 acting as a pivot, can act as a lever-like apparatus. Coupling of the distal blades and the central pore through the lever-like apparatus converts mechanical force into cation conduction. **a** Proposed model of the force-induced gating of Piezo channels. The blue and orange models represent the channel in its closed and open states, respectively. Red dashed lines indicate possible ion-conduction pathways.

Because the pivot of the lever is positioned closer to the central pore than to the distal blades, the input force is effectively amplified through the lever-like apparatus. Additionally, a large conformational change in the distal blades is converted into a relatively slight opening of the central pore, allowing cation-selective permeation.

28.20.6: Kinetics properties of Piezo channels

28.20.6.1: Activation mechanisms of Piezo channels

Normal Piezo channel kinetics can be modeled with three states: open, closed, and inactivated; these states have emerged, collectively, as an important mechanism in the Piezo channel function. Studies have proposed that the Piezo1 channel is gated directly by bilayer tension that can be modified by cytoskeletal proteins and linkages to the extracellular matrix (ECM). For example, in overhydrated red blood cells (RBCs), Piezo1-mediated Ca^{2+} influx activates K^{+} efflux through the Gardos channel (KCa3.1), which in turn leads to water loss and RBC dehydration.

Piezo1 and Piezo2 channels not only exhibit a three-bladed, propeller-shaped trimeric architecture but also can locally deform lipid membranes into a dome-like shape. In addition, changes in the projection area of Piezo channels from closed to open are essential for their mechanosensitivity; this was investigated by calculating the available free energy. Based on these findings, the membrane dome mechanism was proposed and experimentally proved to explain the activation mechanisms of Piezo channels (Fig. 5). Essentially, the dome shape created by Piezo channels in their closed conformation acts as a potential energy source for MS gating. Under tension, lateral membrane tension flattens the Piezo dome, which increases the energy of the membrane-channel system in proportion to the expansion of the projected area of the dome. Piezo channels then open due to the relative energy difference. This mechanism can account for the highly sensitive mechanical gating of Piezo channels with a cation-selective pore. Although the membrane dome mechanism explains the exquisite mechanosensitivity of Piezo channels, it does not consider the shape of the surrounding membrane. Haselwandter et al. [57] proposed the membrane footprint hypothesis, which states that the Piezo1 channel deforms the shape of the membrane outside the perimeter of the channel such that it exhibits a curved membrane footprint, which amplifies the sensitivity of Piezo1 to changes in the membrane tension. Nevertheless, further experiments are needed to test and refine these hypotheses.

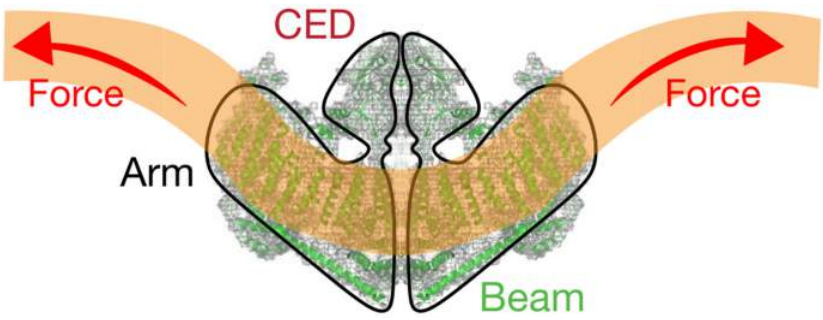


Figure 5: Model of the membrane doming mechanisms. Changes in membrane curvature lead to a gating force applied to the Piezo1 channel

28.20.6.1: Inactivation kinetics of Piezo channels

Various types of mechanical stimulation trigger Piezo channel activation and sequentially elicit an MA current with rapid decay, even in the presence of continued stimulation, due to rapid inactivation. Coste et al. first described detailed information about the voltage-dependent inactivation kinetics of Piezo channels, characterized as fast at rather negative membrane potentials and slow at rather positive membrane potentials. Additionally, Piezo1 channel inactivation is relatively slow compared with Piezo2 channel inactivation. Several point mutations in Piezo channels have been reported to slow down the inactivation process, which produces larger cation fluxes and results in various human diseases. Given its demonstrated key role in normal channel function, we next review what is known about the inactivation kinetics of Piezo channels with a focus on the inactivation mechanism.

The available information regarding the structures (residues/domains) and human disease-related point mutations have helped to clarify the mechanisms of ion channel inactivation. Currently, six gain-of-function mutations associated with dehydrated hereditary xerocytosis (DHS) have been found to slow the inactivation rate of Piezo channels (Table 2), most of which are clustered at the central core region of the Piezo channel structure. This implies that the pore region, which contains an OH, an IH, an extracellular cap domain, and an intracellular CTD, determines the kinetics of inactivation. Further detailed links between structural domains and inactivation kinetics have been investigated. Wu et al. identified that the distinct inactivation kinetics of Piezo1 and Piezo2 channels and characteristic voltage-dependent inactivation appears to be determined by the C-terminal extracellular domains (cap domain). Two potential inactivation gates within the IH and CTD have been thought to be sufficient for the normal inactivation of the Piezo1 and Piezo2 channels. Recently, three small subdomains within the extracellular cap were shown to individually confer Piezo channel inactivation. These results support the idea that the ion-conducting pore region of Piezo channels is essential for their inactivation properties.

Table 2 Mutations in Piezo1 and Piezo2 Associated with Human Diseases: Full-size table

Interestingly, a slowly inactivating MS current in mouse embryonic stem cells (mESs) has been described that is dependent on the Piezo1 channel. However, heterologous expression of Piezo1 cDNA from mES cells displays fast inactivation kinetics, indicating that additional regulatory mechanisms other than the amino acid sequence determine the slow kinetics of the Piezo1 channel in mES cells [70]. Recently, sphingomyelinase activity has been revealed to be a crucial determinant of Piezo1 inactivation. Various modulators, such as pH, temperature, divalent ion concentrations, alternative splicing, osmotic swelling, membrane lipid composition, co-expression of other membrane proteins, and G-protein-coupled pathways have also been reported to regulate the Piezo channel kinetics; however, we still know very little about the relationships among these factors and pivotal structural domains.

28.20.6.1: Pharmacological modulators of Piezo channels

Despite the relatively recent discovery of Piezos, there has been progress regarding small-molecule modulators of Piezo1. Piezo1 chemical activators, including Yoda1 and Jedi1/2, were able to open Piezo1 ion channels in the absence of mechanical stimulation. Jedi1/2, a novel hydrophilic Piezo1 chemical activator, acts through the peripheral blades and utilizes a peripheral lever-like apparatus consisting of the blades and a beam to gate the central ion-conducting pore whereas Yoda1 acts as a molecular wedge, facilitating force-induced conformational changes, effectively lowering the channel's mechanical threshold for activation However, the reason why Yoda1 does not efficiently activate the Piezo2 channel is unclear. Specific inhibitors of Piezo1 are in their infancy. As nonspecific inhibitors of the ion pore in stretch-activated ion channels, gadolinium, and ruthenium red have also been shown to block mouse Piezo1 channels with IC50 values of approximately 5 mM. The commonly used toxin inhibitor of mechanosensitive channels, GsMTx4, was also found to inhibit the Piezo1channel, but it might not bind Piezo1 directly, rather acting via modulating local membrane tension near the channel. Dooku1, an analog of Yoda1 without a stimulatory effect, antagonizes Yoda1-evoked activation of Piezo1 and aortic relaxation.

28.20.7: The function of Piezo channels

Piezo channels are expressed in a wide range of mechanically sensitive cells and allow Ca^{2+} influx in response to different types of external forces, such as fluid flow, pulling, and ultrasonic forces. The biological function of Piezo channels was recently investigated in a number of studies (Fig. 6). The results of these studies verified the pivotal roles of Piezo channels in mechanotransduction under physiological and pathophysiological conditions. Here, we focus on reviewing the biological function of Piezo channels in several different types of MS tissues and cells.

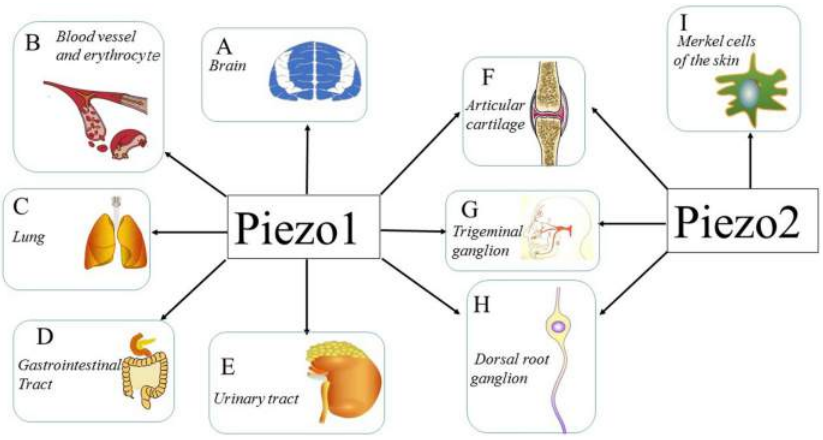


Figure 6: Expression and function of Piezo channels Multiple tissues and cells express Piezo channels, and each of those shown is discussed in this review. **a–e** demonstrate the vital role of the Piezo1 channel in the CNS, blood vessels, erythrocytes, lungs, gastrointestinal tract, and urinary tract. **f–h** illustrate the expression of both the Piezo1 channel and Piezo2 channel in articular cartilage, trigeminal ganglia, and dorsal root ganglia. **i** shows that the Piezo2 channel is expressed in Merkel cells, which are involved in sensing light touch

Consult the original article for details on the roles of Piezo channels on the systems above.

CHAPTER OVERVIEW

29: INTEGRATION OF MAMMALIAN METABOLISM - CAPSTONE VOLUME II

29.1: Overview of Metabolism
29.2: Regulation of glucose metabolism from a liver-centric perspective
29.3: Emerging role of the brain in the homeostatic regulation of energy and glucose metabolism
29.4: Skeletal Muscle Regulates Metabolism
29.5: Intestinal Fructose and Glucose Metabolism in Health and Disease
29.6: Metabolic consequences of obesity and type 2 diabetes- Balancing genes and environment for personalized care
29.7: Lipid-Induced Mechanisms of Metabolic Syndrome
29.8: Fundamentals of cancer metabolism
29.9: Immunometabolism- Cellular Metabolism Turns Immune Regulator
29.10: Ethanol metabolism- The good, the bad, and the ugly
29.11: Gut microbiota-derived metabolites as central regulators in metabolic disorders

29.1: OVERVIEW OF METABOLISM

29.1.1: AN OVERVIEW OF METABOLIC PATHWAYS - CATABOLISM

Biological cells have a daunting task. They must carry out 1000s of different chemical reactions required to carry out cell function. These reactions can include opposing goals such as energy production and energy storage, macromolecule degradation and synthesis, and breakdown and synthesis of small molecules. All of these reactions are catalyzed by proteins and RNAs enzymes whose activities must be regulated, again through chemical reactions, to avoid a futile and energy wasting scenario of having opposing pathways functioning simultaneously in a cell.

Metabolism can be divided into two main parts, catabolism, the degradation of molecules, usually to produce energy or small molecules useful for cell function, and anabolism, the synthesis of larger biomolecules from small precursors.

CATBOLISM: Catabolic reactions involve the breakdown of carbohydrates, lipids, proteins, and nucleic acids to produce smaller molecules and biological energy in the form of heat or small thermodynamically reactive molecules like ATP whose further degradation can drive endergonic process such as biosynthesis. Our whole world is reliant on the oxidation of organic hydrocarbons to water and carbon dioxide to produce energy (at the expense of releasing a potent greenhouse gas, CO2). In the biological world, reduced molecules like fatty acids and partially oxidized molecules such as glucose polymers (glycogen, starch), as well as simple sugars, can be partially or fully oxidized to ultimately produce CO2 as well. Energy released from oxidative reactions is used to produce molecules like ATP as well as heat. Oxidative pathways include glycolysis, the tricarboxylic acid cycle (aka Kreb's cycle) and mitochondrial oxidative phosphorylation/electron transport. To fully oxidize carbon in glucose and fatty acids to carbon dioxide requires splitting C-C bonds and the availability of series of oxidizing agents that can perform controlled, step-wise oxidation reactions, analogous to the sequential oxidation of methane, CH4 to methanol (CH3OH), formaldehyde (CH2O) and carbon dixoxide.

- Glycolysis: This most primitive of metabolic pathways is found in perhaps all organisms. In glycolysis, glucose (C6H12O6), a 6C molecule, is split (or lysed) into two, 3C carbon molecules, glyceraldehyde-3-phosphate, which are then partially oxidized under anaerobic conditions (without O2) to form two molecules of pyruvate (CH3COCO2-). Instead of the very strong oxidizing agent, O2, a weaker one, NAD+ is used, which is reduced in the process to form NADH. Since none of the carbon atoms is oxidized to the state of CO2, little energy is released compared to the complete oxidation to CO2. This pathway comes to a screeching halt if all cellular NAD+ is converted to NADH as NAD+ is not replenished by the simple act of breathing as is the case with O2 in aerobic oxidation. To prevent the depletion of NAD+ from inhibiting the cycle and to allow the cycle to continue under anaerobic conditions, excess NADH is reconverted to NAD+ when the other product of glycolysis, pyruvate is converted to lactate by the enzyme lactate dehydrogenase. Glycolysis occurs in the cytoplasm of the cell.

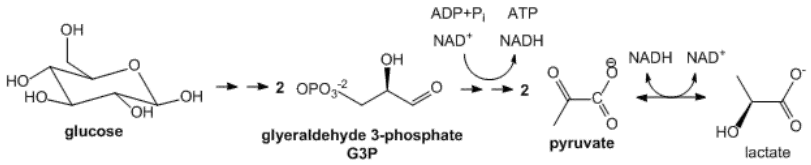


Figure 29.1.1: Summary of Glycolysis (Copyright; author via source)

- Tricarboxylic Acid (Kreb's) Cycle: The TCA cycle is an aerobic pathway which takes place in an intracellular organelle called the mitochondria. It takes pyruvate, the incompletely oxidized product from glycolysis, and finishes the job of oxidizing the 3C atoms all the way to CO2. First the pyruvate moves into the mitochondria where is is oxidized to the 2C molecule acetylCoA with the release of one CO2 by the enzyme pyruvate dehydrogenase. The acetyl-CoA then enters the TCA cycle where two more CO2 are released. As in glycolysis, C-C bonds are cleaved and C is oxidized by NAD+ and another related oxidizing agent, FAD. What is very different about this pathway is that instead of being a series of linear, sequential reactions with one reactant (glucose) and one product (two pryuvates), it is a cyclic pathway. This has significant consequences since if any of the reactants within the pathways becomes depleted, the whole cyclic pathway can slow down and stop. To see how this happens consider the molecule oxaloacetate (OAA) which condenses with acetyl-CoA to form citrate (see diagram below). In this reaction, one OAA is consumed. However, when the cycle returns, one malate is converted to OAA so there is no net loss of OAA, unless OAA is pulled out of the TCA cycle for other reactions, which happens.

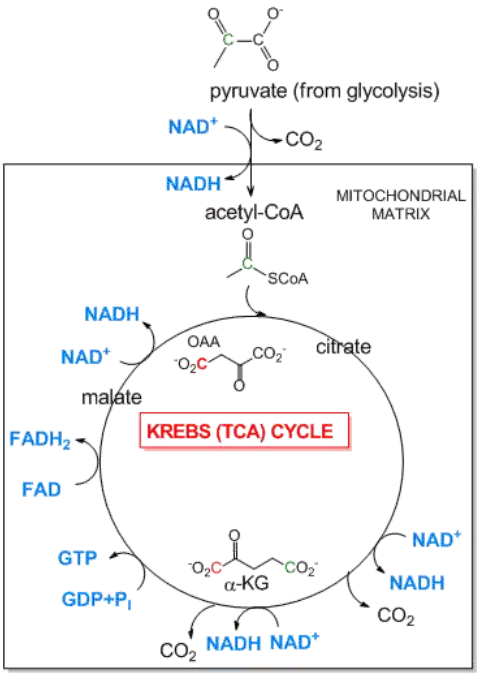


Figure 29.1.2: Pyruvate Dehydrogenase (mitochondrial) and the TCA Cycle. (Copyright; author via source)

- Mitochondrial Oxidative Phosphorylation/Electron Transport: The TCA cycle accomplishes what glycolysis didn't, that is the cleavage of all C-C bonds in glucose (in the form of pyruvate and acetyl-CoA, and the complete oxidation of all C atoms to CO2. Yet two problem remains. The pool of oxidizing molecules, NAD+ and FAD get converted to their reduced forms, NADH and FADH2. Unless NAD+ and FAD are regenerated, as was the case in anaerobic conditions when pyruvate gets converted to lacate, the pathway would again come to a grinding halt. In addition, not much ATP is made in the cycle (in the form of a related molecule GTP). Both these problems are resolved as the resulting NADH and FADH2 formed are reoxidized by mitochondrial membrane enzyme complexes which pass electrons from the oxidized NADH and FADH2 to increasingly potent oxidizing agents until they are accepted by the powerful oxidant O2,which is converted reduced to water. The net oxidation of NADH and FADH2 by dioxygen is greatly exergonic, and the energy released by the process drives the synthesis of ATP from ADP and Pi by an mitochondrial enzyme complex, the F0F1ATPase.

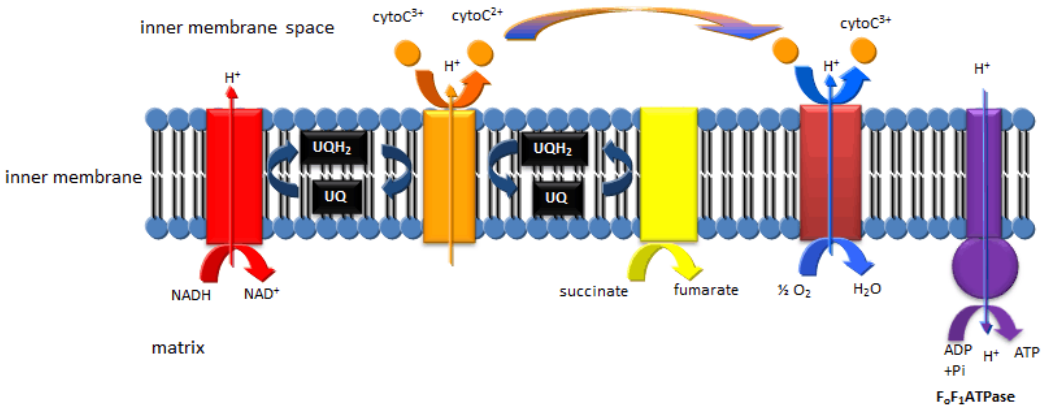
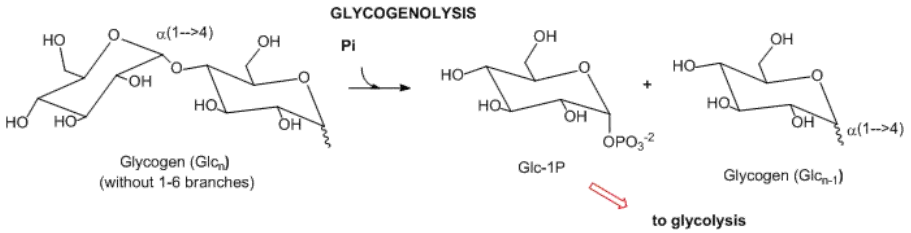


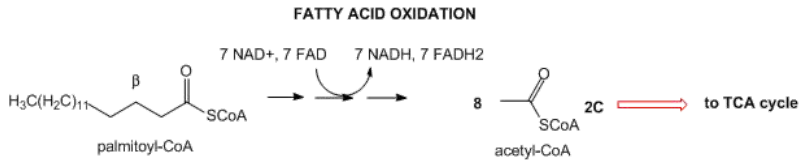
Figure 29.1.3: Mitochondrial Electron Transport/Oxidative Phosphorylation. (Copyright; author via source)

Feeder Pathways: Other catabolic pathways produce products that can enter glycolysis or the TCA cycle. Two examples are given below.

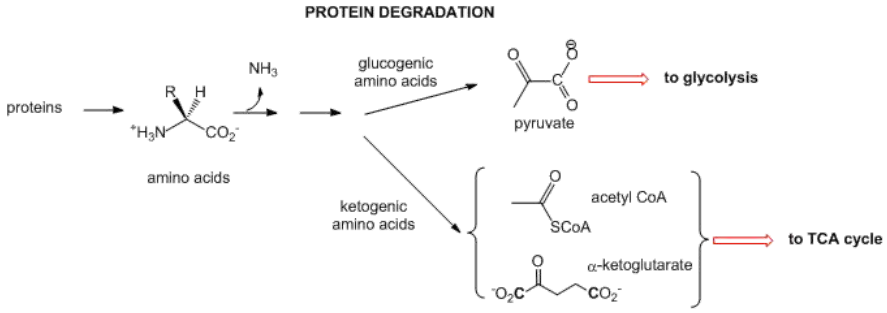
- Complex carbohydrates: In mammals, the major carbohydrate storage molecule is glycogen, a polymer of glucose linked a1-4 with a1-6 branches. The terminal acetal linkages in this highly branched polymer is cleaved sequentially at the ends not through hydrolysis but through phosphorolysis to produce lots of glucose-1-phosphate which can enter glycolysis.



- Lipids: Lipids are stored mostly as triacylglycerides in fat cells (adipocytes). When needed for energy, fatty acids are hydrolyzed from the glycerol backbone of the triacylglyceride, and send into cells where they broken down in an oxidative process to form acetyl-CoA with the concomitant production of lots of NADH and FADH2. These can then enter the mitochondrial oxidative phosphorylation/electrons transport system, which produces, under aerobic conditions, lots of ATP.



- Proteins: When intracellular proteins get degraded, they from individual amino acids. The amine N is lost as it enters the urea cycle. The rest of some amino acid structures can be ultimately converted to acetyl-CoA or keto acids (like alpha-ketoglutarate- a-KG) that are TCA intermediate. These amino acids are called ketogenic. Alternatively, some amino acids, after deamination, are coveted to pyruvate which can either enter the TCA cycle or in the liver be used to synthesize glucose in an anabolic process. These amino acids are called glucogenic. Chemical reactions such as these can be used to replenish intermediates in the TCA cycle which can become depleted as they are withdraw for other reactions.



29.1.1.1: ANABOLIC REACTIONS

Anabolic reactions are those that lead to the synthesis of biomolecules. In contrast to the catabolic reactions just discussed (glycolysis, TCA cycle and electron transport/oxidative phosphorylation) which lead to the oxidative degradation of carbohydrates and fatty acids and energy release, anabolic reactions lead to the synthesis of more complex biomolecules including biopolymers (glycogen, proteins, nucleic acids) and complex lipids. Many biosynthetic reactions, including those for fatty acid synthesis, are reductive and hence require reducing agents. Reductive biosynthesis and complex polymer formation require energy input, usually in the form of ATP whose exergonic cleavage is coupled to endergonic biosynthesis.

Cells have evolved interesting mechanism so as not to have oxidative degradation reactions (which release energy) proceed at the same time and in the same cell as reductive biosynthesis (which requires energy input). Consider this scenario. You dive into a liver cell and find palmitic acid, a 16C fatty acid. From where did it come? Was it just synthesized by the liver cell or did it just enter the cell from a distant location such as adipocytes (fat cells). Should it be oxidized, which should happen if there is a demand for energy production by the cell, or should the liver cell export it, perhaps to adipocytes, which might happen if there is an excess of energy storage molecules? Cells have devised many ways to distinguish these opposing needs. One is by using a slightly different pool of redox reagents for anabolic and catabolic reactions. Oxidative degradation reactions typically use the redox pair NAD^+/NADH (or FAD/FADH2) while reductive biosynthesis often uses phosphorylated variants of NAD^+, $NADP^+$/NADPH. In addition, cells often carry out competing reactions in different cellular compartments. Fatty acid oxidation of our example molecule (palmitic acid) occurs in the mitochondrial matrix, while reductive fatty acid synthesis occurs in the cytoplasm of the cell. Fatty acids entering the cell destined for oxidative degradation are transported into the mitochondria by the carnitine transport system. This transport system is inhibited under conditions when fatty acid synthesis is favored. We will discuss the regulation of metabolic pathways in a subsequent section. One of the main methods, as we will see,

is to activate or inhibit key enzymes in the pathways under a given set of cellular conditions. The key enzyme in fatty acid synthesis, acetyl-CoA carboxylase, is inhibited when cellular conditions require fatty acid oxidation.

The following examples give short descriptions of anabolic pathways. Compare them to the catabolic pathways from the previous section.

- Glucose synthesis, better known as Gluconeogenesis: In glycolysis, glucose ($C_6H_{12}O_6$), a 6C molecule, is converted to two, 3C molecules (pyruvate) in an oxidative process that requires NAD^+ and makes two net ATP molecules. In a few organs, most predominately in the liver, the reverse pathway can take place. The liver does this to provide glucose to the brain when the body is deficient in circulating glucose, for example, under fasting and starving conditions. (The liver under these conditions can get its energy from oxidation of fatty acids). The reactions in gluconeogenesis are the same reactions in glycolysis but run in reverse, with the exception of three glycolytic steps which are essentially irreversible. These three steps have bypass enzymes in the gluconeogenesis pathway. Although the synthesis of glucose is a reductive pathway, it uses NADH instead of NADPH as the redundant as the same enzyme used in glycolysis is simply run in reverse. Gluconeogenesis, which also occurs in the cortex of the kidney, is more than just a simple reversal of glycolysis, however. It can be thought of as the net synthesis of glucose from non-carbohydrate precursors. Pyruvate, as seen in the section on catabolism, can be formed from protein degradation to glucogenic amino acids which can be converted to pyruvate. It can also be formed from triacylglycerides from the 3C molecule glycerol formed and released from adipocytes after hydrolysis of three fatty acids from triacylglycerides. However, in humans, glucose can not be made in net fashion from fatty acids. Fatty acids can be converted to acetyl-CoA by fatty acid oxidation. The resulting acetyl-CoA can not form pyruvate since the enzyme that catalyzes the formation for acetyl-CoA from pyruvate, pyruvate dehydrogenase, is irreversible and there is no bypass reaction known. The acetyl-CoA can enter the TCA cycle but since the pathway is cyclic and proceeds in one direction, it can not form in net fashion oxaloacetate. Although oxaloacetate can be remove from the TCA cycle and be use to form phosphoenolpyuvate, a glycolytic intermediate, one acetyl-CoA condenses with one oxaloacetate to form citrate which leads back to one oxaloacetate. Hence fatty acids can not be converted to glucose and other sugars in a net fashion.

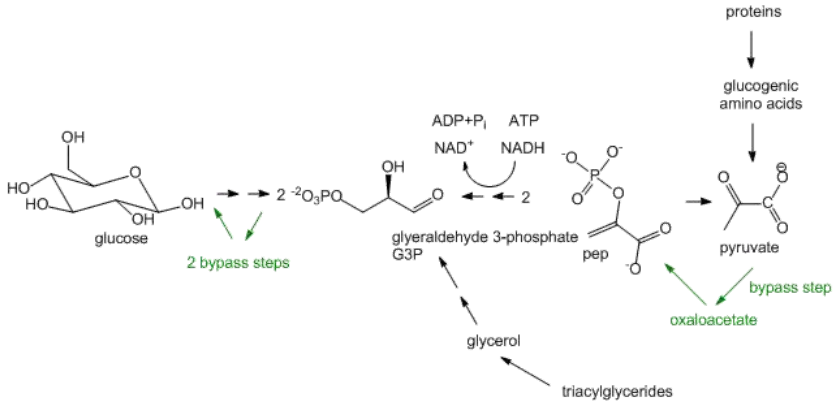
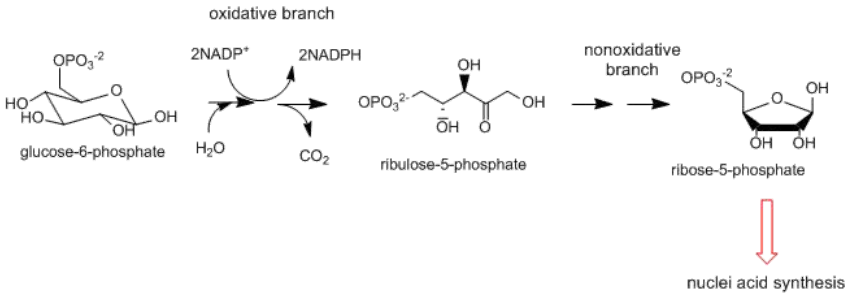


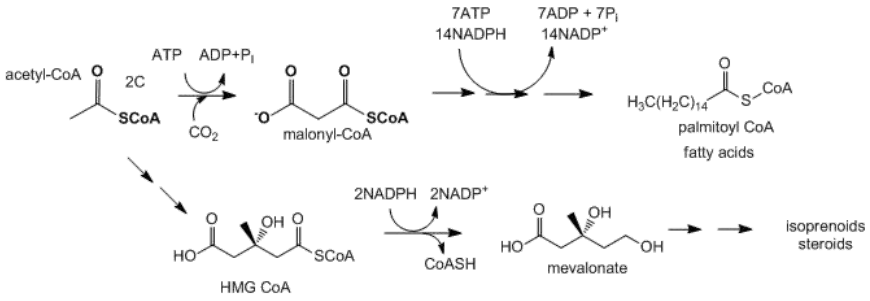
Figure 29.1.4: Gluconeogenesis. (Copyright; author via source)

- Pentose Phosphate Shunt: This two-part pathway doesn't appear to start as a reductive biosynthetic pathway as the first part is the oxidative conversion of a glycolytic intermediate, glucose-6-phosphate, to ribulose-5-phosphate. The next, nonoxidative branch leads to the formation of ribose-5-phosphate, a key biosynthetic intermediate in nucleic acid synthesis as well as erythrose-4-phosphate used for biosynthesis of aromatic amino acids . The oxidative branch is important in reductive biosynthesis as it is a major source of the reductant NADPH used in biosynthetic reactions.



- Fatty acid and isoprenoid/sterol biosynthesis: Acetyl-CoA is the source of carbon atoms for the synthesis of more complex lipids such as fatty acids, isoprenoids, and sterols. When energy needs in a cell are not high, citrate, the condensation product of oxaloacetate and acetyl-CoA in the TCA cycle, builds up in the mitochondrial matrix. It is then transported by the citrate transporter (an inner mitochondrial membrane protein) to the cytoplasm, where it is cleaved back to oxaloacetate and acetyl-CoA by the cytoplasmic enzyme citrate lyase. The oxaloacetate is returned to the mitochondria by conversion first to malate (reduction reaction using NADH), which can

move back into the mitochondria through the malate transporter, or further conversion to pyruate, using the cytosolic malic enzyme, which uses NADP+ to oxidize malate to pyruvate which then enters the mitochondria. The acetyl-CoA formed in the cytoplasm can then be used in reductive biosynthesis using NADPH as the reductant to form fatty acids, isoprenoids, and sterols. The NADPH for the reduction comes from the oxidative branch of the pentose phosphate pathway and from the reaction catalyzed by malic enzyme. The liver cells can still run the glycolytic pathway as the NADH/NAD+ ratio is low in the cytoplasm while NADPH/NADP+ ratio is high.



29.1.1.2: REGULATION AND INTEGRATION

None of these pathways exist in isolation. They are connected in a complicated web of interactions. Instead of cursory summaries of how these pathways are all interconnected, we have chosen to present a series of published articles that are freely available for reuse through Creative Common's licenses. They describe organ- and systems-specific metabolisms and emerging understanding of metabolism in disease states. These include:

- Regulation of glucose metabolism from a liver-centric perspective
- Emerging role of the brain in the homeostatic regulation of energy and glucose metabolism
- Skeletal Muscle Regulates Metabolism
- Intestinal Fructose and Glucose Metabolism in Health and Disease
- Metabolic consequences of obesity and type 2 diabetes: Balancing genes and environment for personalized care
- Lipid-Induced Mechanisms of Metabolic Syndrome
- Fundamentals of cancer metabolism
- Immunometabolism: Cellular Metabolism Turns Immune Regulator
- Ethanol metabolism: The good, the bad, and the ugly
- Gut microbiota-derived metabolites as central regulators in metabolic disorders

---

29.2: REGULATION OF GLUCOSE METABOLISM FROM A LIVER-CENTRIC PERSPECTIVE

Han, HS., Kang, G., Kim, J. *et al.* Regulation of glucose metabolism from a liver-centric perspective. *Exp Mol Med* **48,** e218 (2016). https://doi.org/10.1038/emm.2015.122

29.2.1: RIGHTS AND PERMISSIONS

- Hye-Sook Han,
- Geon Kang,
- Jun Seok Kim,
- Byeong Hoon Choi &
- Seung-Hoi Koo

Experimental & Molecular Medicine **volume 48**, pagee218 (2016)

29.2.2: ABSTRACT

Glucose homeostasis is tightly regulated to meet the energy requirements of the vital organs and maintain an individual's health. The liver has a major role in the control of glucose homeostasis by controlling various pathways of glucose metabolism, including glycogenesis, glycogenolysis, glycolysis and gluconeogenesis. Both the acute and chronic regulation of the enzymes involved in the pathways are required for the proper functioning of these complex interwoven systems. Allosteric control by various metabolic intermediates, as well as post-translational modifications of these metabolic enzymes constitute the acute control of these pathways, and the controlled expression of the genes encoding these enzymes is critical in mediating the longer-term regulation of these metabolic pathways. Notably, several key transcription factors are shown to be involved in the control of glucose metabolism including glycolysis and gluconeogenesis in the liver. In this review, we would like to illustrate the current understanding of glucose metabolism, with an emphasis on the transcription factors and their regulators that are involved in the chronic control of glucose homeostasis.

29.2.3: OVERVIEW OF GLUCOSE METABOLISM IN THE LIVER

Under feeding conditions, dietary carbohydrates are digested and processed by various glucosidases in the digestive tract, and the resultant monosaccharides, mainly hexose glucose, are transported into various tissues as a primary fuel for ATP generation.[1] In most mammalian tissues, the catabolism of glucose into pyruvate, termed glycolysis, is preserved as a major pathway in eliciting ATP. In tissues with abundant mitochondria, cytosolic pyruvate is transported into the mitochondrial matrix, converted to acetyl-CoA by pyruvate dehydrogenase complex, and incorporated into the tricarboxylic acid cycle in conjunction with oxaloacetate. The cycle generates energy equivalent to ATP (that is, GTP) as well as both NADH and $FADH_2$, which serve as important electron carriers for electron transport chain-oxidative phosphorylation, resulting in the generation of ATP.

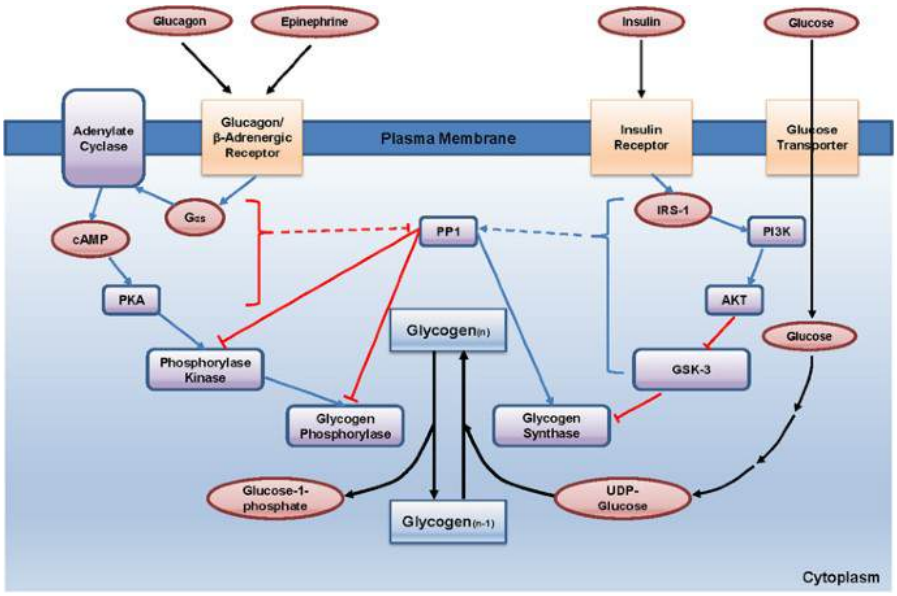
In some cases, such as red blood cells lacking mitochondria or cells under ischemic conditions, pyruvate is converted into lactate in the cytosol to regenerate the NAD^+ that is necessary for the continued generation of ATP by substrate-level phosphorylation via anaerobic glycolysis. Excessive carbohydrates in the liver are first converted into glycogen, a storage form of glucose in animals, by glycogenesis. In addition, in a carbohydrate-rich diet, the excessive carbohydrates are also converted into fatty acids via lipogenesis using the acetyl-CoA generated from glycolysis-driven pyruvate, which is incorporated into very low density lipoproteins for transport to white adipose tissue for the storage.[2] The regulation of glycogen metabolism is examined in detail in this section, and the transcriptional control of glycolysis and lipogenesis is delineated in the following section.

Under fasting conditions, the liver has a major role in generating glucose as a fuel for other tissues, such as the brain, red blood cells and muscles. Initially, an increase in the pancreatic hormone glucagon initiates the cascade of kinase action (stated below in detail) that releases glucose from the stored glycogen via glycogenolysis.[1] Normally, stored glycogen is critical for maintaining glucose homeostasis in mammals during an overnight fasting period. During a longer term fast or starvation, essentially all of the stored glycogen in the liver is depleted (after ~30 h of fasting), and *de novo* glucose synthesis or gluconeogenesis is responsible for the generation of glucose as a fuel for other tissues. Major non-carbohydrate precursors for gluconeogenesis are lactate, which is transported from peripheral tissues such as skeletal muscles or red blood cells, and glycerol, which is released from the adipose tissues via enhanced lipolysis during fasting. Most of the initial precursors for gluconeogenesis are generated in the mitochondria (except glycerol 3-phosphate via glycerol kinase activity), but the majority of the reaction occurs in the cytosolic part of the cell. The complex regulatory mechanism is delineated in detail in the following section, with an emphasis on the transcriptional control of key regulatory enzyme genes.

29.2.4: REGULATION OF GLYCOGEN METABOLISM IN THE LIVER

The accumulation of glycogen in the liver during feeding conditions provides a storage form of glucose that can be used in times of reduced food intake (Figure 1). Multiple layers of regulation are required for this process for both the activation of glycogen synthase, which is a key enzyme of glycogenesis (glycogen synthesis), and the inhibition of glycogen phosphorylase, which is a key enzyme of glycogenolysis (glycogen breakdown) in the liver. Glycogen synthase is a major enzyme that facilitates the elongation of glycogen chains by catalyzing the transfer of the glucose residue of UDP-glucose to the non-reducing end of a pre-existing glycogen branch to produce a new $\alpha 1 \rightarrow 4$ glycosidic linkage. The regulation of glycogen synthase has been mostly studied using a muscle-specific isoform. In the muscle, glycogen synthase is inactivated via phosphorylation on multiple serine residues by various serine/threonine kinases such as casein kinase-1, protein kinase A (PKA), and glycogen synthase kinase-3 (GSK-3). Most notably, the phosphorylation of glycogen synthase by GSK-3 at serine residues 640, 644 and 648 has been linked to the most important inhibitory post-translational modification for its catalytic activity.

Figure 1



Regulation of hepatic glycogen metabolism. Under fasting conditions, glucagon and epinephrine induce cAMP-dependent signaling cascades, leading to the activation of glycogen phosphorylase and glycogenolysis while inhibiting glycogenesis. Conversely, feeding enhances insulin-mediated signaling in the liver, leading to the activation of both PP1 and Akt, thus promoting glycogen synthesis in response to increased glucose uptake in the liver. See the main text for more specific regulatory pathways. cAMP, cyclic AMP.

Under fasting conditions, dephosphorylated and active GSK-3 phosphorylate and inactivate glycogen synthase, leading to the inhibition of hepatic glycogen synthesis. On feeding, increased insulin signaling activates Akt in the cell, which in turn phosphorylates and inactivates GSK-3, thus resulting in the activation of glycogen synthase. In addition, increased concentrations of glucose 6-phosphate allosterically activate this enzyme, thus potentiating its catalytic activity under feeding conditions.[3, 4] One recent publication argues against the role of GSK-3 in the regulation of the liver-specific isoform of glycogen synthase. In that study, Guinovart *et al.*[5] mutated the corresponding serine residues that are shown to be regulated by GSK-3 in the liver-isoform of glycogen synthase. They found that the mutation of those residues did not affect the overall enzyme activity but that the mutation of serine 7 to alanine, a site that is recognized and regulated by PKA, led to the increased activity of this enzyme. Further study is necessary to determine whether these results can be verified *in vivo* using animal models such as liver-specific knock-in mice for S7A liver glycogen synthase. The protein phosphatase 1 (PP1) may be responsible for the dephosphorylation and activation of glycogen synthase. Accordingly, both glucose and insulin have been shown to activate PP1 activity, whereas glucagon and epinephrine have been linked to the inhibition of its activity.

Glycogen phosphorylase is a major enzyme involved in glycogenolysis (Figure 1). This enzyme catalyzes the reaction of the removal of a glucose residue from the non-reducing end of a glycogen chain, leading to the generation of glucose 1-phosphate.[6] Glucose 1-phosphate can be converted into glucose 6-phosphate by phosphoglucomutase, and glucose 6-phosphate can be incorporated into glycolysis or further converted into glucose by glucose 6-phosphatase, depending on the energy status of the organism. Glycogen phosphorylase is active when it is phosphorylated at its serine 14 residue. The phosphorylation of glycogen phosphorylase requires a cascade mechanism of epinephrine and glucagon in the liver. On the activation of Gαs by the binding of hormones to cell surface G protein-coupled receptors (beta adrenergic receptor or glucagon receptor), the intracellular cyclic AMP (cAMP) levels increase via adenylate cyclase, leading to the activation of PKA. PKA is then responsible for the phosphorylation and activation of glycogen phosphorylase kinase, which in turn phosphorylates and activates glycogen phosphorylase to enhance glycogen breakdown. Under feeding conditions, this kinase cascade is inactive due to the lack of secretion of catabolic hormones. In addition, insulin promotes the activation of PP1, which dephosphorylates and inactivates glycogen

phosphorylase. In essence, the anabolic hormone insulin promotes glycogenesis and inhibits glycogenolysis via the activation of PP1, leading to the dephosphorylation of glycogen phosphorylase (inactivation) and glycogen synthase (activation), and via the activation of Akt, leading to the phosphorylation of GSK-3 (inactivation) that is unable to phosphorylate and inactivate glycogen synthase.

29.2.5: CONTROL OF HEPATIC GLYCOLYSIS

As stated above, glycolysis is critical to the catabolism of glucose in most cells to generate energy. The key rate-limiting enzymes for this pathway include glucokinase (GK, also termed hexokinase IV), which converts glucose into glucose 6-phosphate; phosphofructokinase-1 (PFK-1), which converts fructose 6-bisphosphate into fructose 1,6-bisphosphate; and liver-type pyruvate kinase (L-PK), which converts phosphoenolpyruvate (PEP) into pyruvate in the liver. These enzymes are tightly regulated by allosteric mediators that generally promote the catabolism of glucose in the cell.[2, 7, 8, 9]

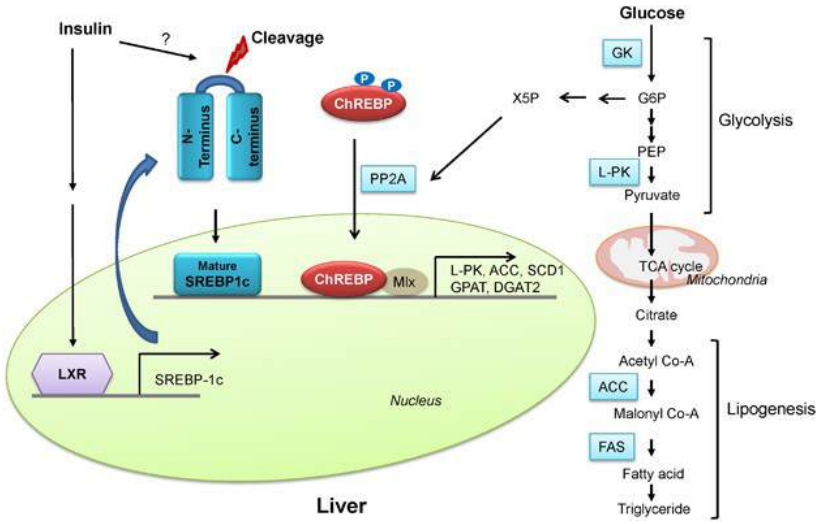
GK is a high Km hexokinase that is present in the liver and the pancreatic beta cells, thus functioning as a glucose sensor for each cell type. Unlike the other hexokinase isotypes, GK activity is not allosterically inhibited by its catalytic product, glucose 6-phosphate in the cell, thus enabling the liver to continuously utilize glucose for glycolysis during conditions of increased glucose availability, such as during feeding conditions. GK is regulated via its interaction with glucokinase regulatory protein (GKRP). In the low intracellular glucose concentration during fasting, the binding of GK and GKRP is enhanced by fructose 6-phosphate, leading to the nuclear localization of this protein complex. Higher concentrations of glucose during feeding compete with fructose 6-phosphate to bind this complex, which promotes the cytosolic localization of GK that is released from GKRP, thus causing the increased production of glucose 6-phosphate in this setting.[10]

PFK-1 catalyzes the metabolically irreversible step that essentially commits glucose to glycolysis. This enzyme activity is allosterically inhibited by ATP and citrate, which generally indicate a sign of energy abundance. Reciprocally, it is allosterically activated by ADP or AMP, making it more efficient to bring about glycolysis to produce more ATP in the cell. In addition, PFK-1 activity is allosterically activated by fructose 2,6-bisphosphate (F26BP), a non-glycolytic metabolite that is critical for the regulation of glucose metabolism in the liver. F26BP is generated from fructose 6-phosphate by the kinase portion of a bifunctional enzyme that contains both a kinase domain (phosphofructokinase-2, PFK-2) and a phosphatase domain (fructose 2,6-bisphosphatase, F-2,6-Pase). PFK-2 is activated by the insulin-dependent dephosphorylation of a bifunctional enzyme that activates PFK-2 activity and simultaneously inhibits F-2,6-Pase activity to promote the increased F26BP concentration. Glucagon-mediated activation of PKA is shown to be responsible for the phosphorylation and inactivation of the kinase portion of this enzyme.[7]

Unlike its muscle counterpart, L-PK is also a critical regulatory step in the control of glycolysis in the liver. As in the case of other glycolytic enzymes, L-PK activity is regulated by both allosteric mediators and post-translational modifications. L-PK activity is allosterically activated by fructose 1,6-bisphosphate, an indicator for the active glycolysis. By contrast, its activity is allosterically inhibited by ATP, acetyl-CoA, and long-chain fatty acids, all of which signal an abundant energy supply. Additionally, the amino acid alanine inhibits its activity, as it can be readily converted to pyruvate by a transamination reaction. L-PK is inhibited by PKA following a glucagon-mediated increase in intracellular cAMP during fasting and is activated by insulin-mediated dephosphorylation under feeding conditions.[7]

In addition to the acute regulation of key regulatory enzymes, glycolysis is regulated by a transcriptional mechanism that is activated during feeding conditions. Two major transcription factors, sterol regulatory element binding protein 1c (SREBP-1c) and carbohydrate response element binding protein (ChREBP), are responsible for the transcriptional activation of not only glycolytic enzyme genes but also the genes involved in fatty acid biosynthesis (such as fatty acid synthase (FAS), acetyl-CoA carboxylase (ACC), and stearoyl-CoA desaturase 1 (SCD1)) and triacylglycerol formation (such as glycerol 3-phosphate acyltransferase (GPAT) and diacylglycerol acyltransferase 2 (DGAT2)), a process that is normally activated by a carbohydrate-rich diet (Figure 2).[11] Because these processes are often coordinately regulated, that is activated during feeding and inhibited by fasting, they are sometimes collectively called lipogenesis.

Figure 2



Regulation of hepatic glycolysis. Under feeding conditions, increased glucose uptake in hepatocytes promotes glycolysis and lipogenesis to generate triglycerides as storage forms of fuel. This process is transcriptionally regulated by two major transcription factors in the liver, SREBP-1c and ChREBP-Mlx heterodimer, which mediate the insulin and glucose response, respectively. See the main text for more specific regulatory pathways.

Full size image

SREBPs are the major regulators of lipid metabolism in mammals. They are members of the basic helix-loop-helix leucine zipper (b/HLH/LZ) type transcription factor families comprising SREBP-1a, SREBP-1c, and SREBP-2. SREBP is translated as an endoplasmic reticulum (ER)-bound precursor form that contains the N-terminal transcription factor domain and the C-terminal regulatory domain linked with the central transmembrane domain.[12] Within this family of transcription factors, SREBP-2 is linked to the control of cholesterol uptake or biosynthesis in the liver by the transcriptional activation of the genes involved in the pathway including low density lipoprotein receptor (LDLR), 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMGCR), hydroxy-3-methylglutaryl-coenzyme A synthase 1 (HMGCS1), and farnesyl diphosphate synthase (FDPS). SREBP-1c, however, activates the genes encoding the enzymes for lipogenesis (FAS, ACC, SCD1, and DGAT2) as well as GK, which is a first enzyme in the commitment step of glucose utilization in the liver. Indeed, liver-specific SREBP-1c knockout mice showed an impaired activation of lipogenic genes in a high carbohydrate diet, thus confirming the importance of this transcription factor in the regulation of hepatic glycolysis and fatty acid biosynthesis.[13] SREBP-1a is not highly expressed in the liver but was shown to be involved in the formation of inflammasomes in response to lipopolysaccharide (LPS) treatment in macrophages by transcriptional activation of Nlrp1.[14] The regulation of SREBP-2 and SREBP-1c are quite distinct in the liver. The expression of SREBP-2 is not controlled by sterols, but its proteolytic processing is tightly regulated by intracellular concentrations of cholesterol. It is normally bound in the ER via the interaction of SREBP-cleavage-activating protein (SCAP) and insulin-induced gene protein (INSIG) in the presence of high intracellular cholesterol levels, and the reduction in the cholesterol concentration releases the interaction of SCAP and SREBP-2/INSIG complex, resulting in the translocation of the latter complex into the Golgi apparatus and the liberation of the active SREBP-2 factor by sequential proteolytic cleavages.[15] Unlike SREBP-2, SREBP-1c is mainly regulated at the transcription level by insulin. The exact transcription factor that mediates this insulin-dependent signal is not yet clear, although SREBP-1c itself might be involved in the process as part of an auto-regulatory loop. Interestingly, the oxysterol-sensing transcription factor liver X receptor (LXR) is shown to control the transcription of SREBP-1c, suggesting that SREBP-1c and SREBP-2 could be regulated differently in response to cellular cholesterol levels.[16] Recent studies have revealed the involvement of various kinases in the control of SREBP-1c activity. In HepG2 cells, PKA was shown to reduce the DNA binding ability of SREBP-1a by the phosphorylation of serine 338 (equivalent of serine 265 for SREBP-1c).[17] A report by Bengoechea and Ericsson suggested that GSK-3, a kinase known to reduce glycogen synthesis by targeting glycogen synthase, downregulates SREBP-1 activity via the phosphorylation of the C-terminal residue that promotes the ubiquitin ligase Fbw7-dependent degradation of SREBP-1 proteins.[18] In addition, both AMP activated protein kinase (AMPK) and its related kinase salt-inducible kinase (SIK) 1 are involved in the down-regulation of its activity through inhibitory phosphorylation (serine 372 for AMPK, which blocks proteolysis and nuclear localization of SREBP-1c, and serine 329 for SIK1, which directly reduces its transcriptional activity).[19, 20] These data suggest that the fine-tuning of SREBP-1c activity is critical to the maintenance of glucose and lipid homeostasis in the liver.

The other prominent transcription factor for controlling glycolysis and fatty acid biosynthesis in the liver is ChREBP. ChREBP was initially known as Williams-Beuren syndrome critical region 14 (WBSCR14) and was considered one of the potential genes that instigate Williams-Beuren syndrome. Later, by using a carbohydrate response element (ChoRE) from L-PK, ChREBP was isolated as a bona fide transcription factor for binding ChoRE of glycolytic promoters.[21] Indeed, ChREBP is highly expressed in tissues that are active in lipogenesis such as the liver, brown adipocytes, white adipocytes, small intestine, and kidney. As in the case for SREBP, ChREBP belongs to the b/HLH/LZ

transcription factor family and forms a heterodimer with another b/HLH/LZ transcription factor Max-like protein X (Mlx) on the glycolytic promoter.[22] As in the case for the SREBP-1c, the expression of ChREBP is increased in the liver as a result of a high carbohydrate diet, and the effect was recapitulated in primary hepatocytes with high glucose treatment.

A recent report indeed suggested a role for LXR in the transcriptional activation of ChREBP in response to glucose, although the study needs to be further verified because the transcriptional response is shown not only by the treatment of D-glucose, a natural form of glucose present in animals, but also by the treatment of unnatural L-glucose, a form of glucose that is not known to activate lipogenesis in the liver.[23] Moreover, studies performed in LXR knockout mice revealed no changes in ChREBP expression in the liver, arguing against the role of LXR in the control of ChREBP.[24] Glucose is also shown to regulate ChREBP activity by controlling its nuclear localization. There are three prominent serine/threonine residues that are targeted by serine/threonine kinases. PKA is shown to phosphorylate serine 196, which is critical for cellular localization, and threonine 666, which is critical for its DNA binding ability, whereas AMPK phosphorylate serine 568 dictates its DNA binding ability. All three sites are phosphorylated under fasting conditions by these kinases and are dephosphorylated under feeding by xylulose 5-phosphate (X5P)-mediated activity of protein phosphatase 2A (PP2A).[25, 26] However, the current model needs to be further verified due to the contrasting data that have been published regarding the role of these phosphorylations on ChREBP activity.

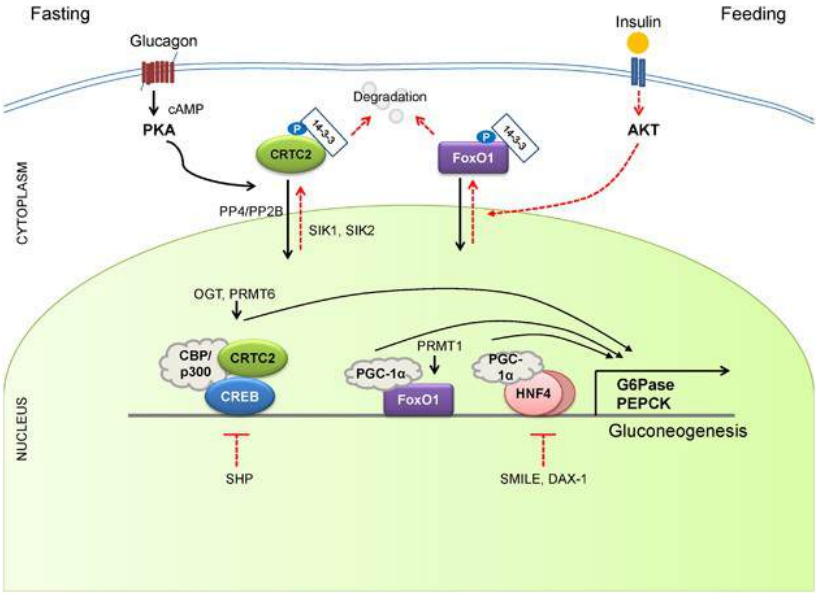
First, high glucose concentrations in primary hepatocytes do not result in decreased cAMP levels or PKA activity, suggesting that other signals might be necessary to mediate the high glucose-dependent nuclear translocation of ChREBP. In addition, a serine to alanine mutant of ChREBP still requires high glucose for its full activity, suggesting that additional actions are necessary to recapitulate the high glucose-mediated activation/nuclear localization of ChREBP in the liver.[27, 28] The physiological role of ChREBP in liver glucose metabolism was verified by *in vivo* studies. ChREBP knockout mice were born in a Mendelian ratio and showed no developmental problems. The knockout animals showed reduced liver triacylglycerol levels together with a reduction in lipogenic gene expression, thus confirming the role of ChREBP in the control of hepatic glycolysis and fatty acid synthesis.[29] Interestingly, the compensatory increase in glycogen was observed in the livers of these mice, suggesting that these mice adapted to store more glycogen as a storage form of fuel as opposed to triacylglycerol. In ob/ob mouse liver, increased accumulation of nuclear ChREBP was shown, suggesting that this phenomenon might be causal to the fatty liver phenotype in these mice. Indeed, knockdown of ChREBP in ob/ob mice reduced the rate of lipogenesis with decreased expression of most lipogenic genes.[30] Furthermore, the depletion of hepatic ChREBP in ob/ob mice improved hyperglycemia, hyperlipidemia, and hyperinsulinemia, suggesting that regulation of ChREBP might be critical in the control of metabolic disorders in the presence of obesity and insulin resistance.

29.2.6: CONTROL OF HEPATIC GLUCONEOGENESIS

Prolonged fasting or starvation induces de novo glucose synthesis from non-carbohydrate precursors, termed hepatic gluconeogenesis. This process initiates from the conversion of pyruvate to oxaloacetate by pyruvate carboxylase (PC) in the mitochondria and eventually concludes in the conversion into glucose via several enzymatic processes in the cytosol.[7, 8, 9] Among the substrates for gluconeogenesis are amino acids, which can be converted into either pyruvate or intermediates of the tricarboxylic acid cycle; lactate, which can be converted into pyruvate by lactate dehydrogenase; and glycerol (from increased lipolysis in the white adipocytes under fasting), which can be converted into dihydroxyacetone phosphate, a gluconeogenic intermediate (a two-step process that is catalyzed by glycerol kinase and glycerol 3-phosphate dehydrogenase). Key regulatory enzymes in that pathway, including glucose 6-phosphatase (G6Pase), fructose 1,6-bisphosphatase (Fbpase1), PC, and phosphoenolpyruvate carboxykinase (PEPCK), are activated under fasting conditions to enhance gluconeogenic flux in that setting.

Mitochondrial acetyl-CoA (derived from the increased fatty acid oxidation under fasting) functions as a key allosteric activator of PC, leading to the increased production of oxaloacetate for the gluconeogenesis. In addition, F26BP, which is a key allosteric regulator for glycolysis by activating PFK-1, was shown to inhibit gluconeogenesis via the allosteric inhibition of Fbpase1, which helps reciprocally control gluconeogenesis and glycolysis under different dietary statuses. Because Fbpase2 is activated but PFK-2 is inhibited under fasting, the lack of F26BP enables the activation of Fbpase1 and the increased production of fructose 6-phosphate in gluconeogenesis. The chronic activation of gluconeogenesis is ultimately achieved via transcriptional mechanisms. Major transcriptional factors that are shown to induce gluconeogenic genes include CREB, FoxO1, and several nuclear receptors (Figure 3).[31]

Figure 3



Regulation of hepatic gluconeogenesis. Under fasting conditions, hepatic gluconeogenesis is enhanced via a decreased concentration of insulin and an increased concentration of insulin counterregulatory hormones such as glucagon. CREB/CRTC2, FoxO1, and a family of nuclear receptors are critical in coordinating the fasting-mediated activation of gluconeogenesis in the liver. See the main text for more specific regulatory pathways. FoxO1, forkhead box O 1

Full size image

Under fasting conditions, glucagon and epinephrine can increase the cAMP concentration in the liver via the activation of adenylate cyclase, leading to the activation of PKA and the subsequent induction of CREB via its serine 133 phosphorylation. The phosphorylation event is crucial in the recruitment of histone acetyltransferases (HAT) CBP/p300, leading to the histone H3 and H4 acetylation and the transcriptional activation of target genes.[32, 33] CREB-dependent transcription is further enhanced by association with additional transcriptional coactivators CREB regulated transcription coactivators (CRTCs), which are a target for CBP/p300-mediated acetylation, which in turn promotes a tighter association of CREB, CBP/p300, and CRTC on the promoter.[34, 35, 36] The role of CREB in the control of hepatic gluconeogenesis has been confirmed by *in vivo* studies by utilizing albumin promoter-driven ACREB (CREB inhibitor) transgenic mice and siRNA-mediated CREB knockdown mouse models.[37, 38] In both mouse models, the inactivation of CREB reduced blood glucose levels and reduced the expression of gluconeogenic genes in mice, showing that CREB is a bona fide physiological transcriptional regulator of hepatic gluconeogenesis *in vivo*. In contrast, the role for CBP in gluconeogenesis is still controversial. Disruption of CREB-CBP interaction does not appear to affect glucose homeostasis because mice exhibiting a stable expression of mutant CBP that was unable to bind CREB showed normal glycemia.[36] Furthermore, mutant mice producing CH1 null products (ΔCH1-a domain that is critical for insulin-mediated depression of CBP activity) displayed normal fasting gluconeogenesis.[39] Thus, further studies are required to describe the potential role of HATs in the transcriptional control of CREB activity in this setting.

The CRTC family of transcriptional coactivators consists of CRTC1, CRTC2 and CRTC3, which were isolated by the expression library screening as activaters of CREB-dependent transcription.[34] CRTC activity is regulated by cellular localization, and the AMPK family of serine/threonine kinases, such as AMPK, SIK1 or SIK2, was shown to be involved in the inhibitory phosphorylation of this factor (serine 171 for CRTC2).[40] In addition, the phosphorylation status of CRTC is regulated by a pair of serine/threonine phosphatases (PP2B or PP4) in response to cAMP signaling or calcium concentration in the cell.[41, 42] CRTC activity is also further enhanced by *O*-GlcNAcylation on serine 171 and arginine methylation by protein arginine methyltransferase (PRMT) 6.[43, 44] Among the family members, CRTC2 is the prominent isoform in the liver. Recent studies have delineated the role of CRTC2 in the regulation of hepatic gluconeogenesis in vivo. Knockdown of CRTC2 in mice by RNAi reduced blood glucose levels and led to a concomitant repression of gluconeogenic gene expression.[36] In addition, CRTC2 knockout mice displayed lower plasma glucose levels and improved glucose tolerance, indeed showing that CRTC2 is crucial in controlling hepatic glucose metabolism *in vivo*.[45] A recent study indicated that CRTC2 could also coactivate other bZIP transcription factors that are implicated in the regulation of glucose homeostasis.[46, 47] Further study is required to delineate the potential contributions from other bZIP factors in the control of hepatic gluconeogenesis by using tissue-specific knockout mouse models.

The forkhead box O (FoxOs) belongs to a class of forkhead families of transcription, which recognize the AT-rich insulin response element on the promoter.[48, 49] Of the four major isoforms in mammals (FoxO1, FoxO3, FoxO4, and FoxO6), FoxO1 is the predominant isoform in the liver. The activity of this protein is also regulated by phosphorylation-dependent subcellular localization, and three major serine and threonine residues (threonine 24, serine 253 and serine 316 for murine FoxO1) are targeted by the insulin/Akt pathway. Following phosphorylation, FoxO1 moves to the cytosol via an association with 14-3-3, where it is degraded by the ubiquitin/proteasome-dependent

pathway.[50, 51, 52] In addition to phosphorylation, FoxO1 was shown to be regulated by the HAT-dependent acetylation of specific lysine residues (lysine 242, 245 and 262 for murine FoxO1), which also inhibit its transcriptional activity.[53] In the liver, FoxO1 regulates hepatic gluconeogenesis via the transcriptional regulation of key genes in the pathway such as PEPCK and G6Pase and is considered a major regulatory point for the insulin-mediated repression of hepatic gluconeogenesis.[54] Indeed, mice with liver-specific knockout of FoxO1 showed lower plasma glucose levels that those associated with reduced hepatic glucose output, thus underscoring the physiological significance of this factor in the control of glucose homeostasis *in vivo*.[54, 55] As in the case for CREB, FoxO1 requires transcriptional coactivators for optimal transcriptional activity.

Peroxisome proliferator-activated receptor gamma coactivator 1 alpha (PGC-1α), a known coactivator for nuclear receptors, functions as a key transcriptional coactivator for FoxO1 in hepatic gluconeogenesis.[56] PGC-1α was originally isolated in brown adipocytes and was shown to control adaptive thermogenesis in response to cold shock in that setting.[57] In the liver, the expression of PGC-1α is upregulated under fasting conditions via a CRTC2-CREB-dependent mechanism and is critical in maintaining prolonged gluconeogenesis under starvation by enhancing the transcriptional activity of FoxO1 as a coactivator.[38, 57, 58] Indeed, the depletion of hepatic PGC-1α in mice results in lower fasting glucose levels with a concomitant reduction in hepatic gluconeogenesis, thus showing the physiological relevance of this coactivator in the control of glucose homeostasis.[59, 60] As is the case for CRTC2, FoxO1 activity is enhanced by arginine methylation by PRMT. In this case, PRMT1 promotes the asymmetric dimethylation of arginine 248 and 250 in FoxO1, which blocks the binding of Akt and the subsequent Akt-mediated phosphorylation of the adjacent serine residue (serine 253), thus enhancing the nuclear localization of FoxO1.[61] Consequently, the chromatin occupancy of FoxO1 onto the gluconeogenic promoter and gluconeogenesis itself are increased due to the PRMT1-dependent arginine methylation.[62] Acute knockdown of hepatic PRMT1 in mice reduces FoxO1-mediated glucose production, confirming that PRMT1 is crucial in modulating FoxO1 activity and subsequent gluconeogenesis in the physiological context.

Nuclear receptors belong to the superfamily of transcription factors that possess two Cys2-His2 type zinc finger motifs as a DNA binding domain as well as both ligand-independent and ligand-dependent transactivation domains.[63] The latter activation domain also contains a ligand-binding domain. Nuclear receptors can be classified into one of three subgroups based on their dimer-forming potential. Homodimeric nuclear receptors are also called cytosolic receptors because they reside in the cytosol and associate with molecular chaperones such as heat-shock proteins. On binding to the ligand, they form homodimers and translocate to the nucleus to bind a specific response element termed the hormone response element to elicit the ligand-dependent transcriptional response. Most of the steroid hormone receptors, such as the glucocorticoid receptor (GR), estrogen receptor (ER), and progesterone receptor (PR), belong to this subfamily. By contrast, heterodimeric nuclear receptors reside in the nucleus and are bound to their cognate binding sites together with the universal binding partner retinoid X receptor (RXR). In the absence of the ligands, these factors repress the transcription of target genes in association with transcriptional corepressors such as histone deacetylase or nuclear receptor corepressor (NCoR)/silencing mediator of retinoid and thyroid hormone receptors (SMRT). Ligand binding initiates the conformational changes of these heterodimeric nuclear receptors, which promotes the dissociation of corepressors and the association of coactivators such as CBP/p300, p160 steroid receptor coactivator family, and PGC-1α.

Examples of this class of nuclear receptors include members of peroxisome proliferator-activated receptors, LXRs, vitamin D receptors and thyroid hormone receptors. The final subclasses of nuclear receptors are types that function as monomers. They usually lack specific endogenous ligands and are often called orphan nuclear receptors. Some of them also lack DNA binding domain and thus function as transcriptional repressors of various transcription factors, including members of nuclear receptors. They are called atypical orphan nuclear receptors. Among the homodimeric nuclear receptors, the role of GR has been linked to the control of hepatic gluconeogenesis. GR is activated by cortisol, which is released from the adrenal cortex in response to chronic stresses such as prolonged fasting.[64, 65] GR was shown to directly bind to the cognate binding sites found in the promoters of gluconeogenic genes such as PEPCK and G6Pase and to enhance transcription of these genes under fasting conditions. The same response elements were also shown to be recognized and regulated by hepatocyte nuclear factor 4 (HNF4), a member of heterodimeric nuclear receptors, which suggests that these nuclear receptors could coordinately function to control hepatic gluconeogenesis in response to fasting.[58]

In accordance with this idea, the activity of these nuclear receptors can be effectively integrated by the function of transcriptional co-activator PGC-1α. Recently, estrogen-related receptor gamma (ERRγ), a member of monomeric nuclear receptors, was shown to be involved in the regulation of hepatic gluconeogenesis.[66, 67] In the liver, ERRγ expression is increased under fasting or by insulin resistance in a CRTC2-CREB-dependent manner. This factor regulates hepatic gluconeogenesis by binding to unique response elements that are distinct from the known nuclear receptor-binding sites in the promoters of PEPCK and G6Pase. Inhibition of ERRγ activity by injecting either RNAi or the inverse agonist GSK5182 effectively reduced hyperglycemia in diabetic mice, suggesting that the control of this factor might potentially be beneficial in the treatment of patients with metabolic diseases. As is the case for other nuclear receptors that control hepatic gluconeogenesis, ERRγ activity is further enhanced by interaction with the transcriptional coactivator PGC-1α, showing that this coactivator functions as a master regulator for the hepatic glucose metabolism.

Three members of atypical orphan nuclear receptors, the small heterodimer partner (SHP, also known as NR0B2); the dosage-sensitive sex reversal, adrenal hypoplasia critical region, on chromosome X (DAX-1, also known as NR0B1); and the SHP-interacting leucine zipper protein (SMILE) are implicated in the transcriptional repression of hepatic gluconeogenesis.[68, 69, 70] SHP is ubiquitously expressed in mammalian tissues, with the highest expression occurring in the liver. Interestingly, metformin directly activates the transcription of SHP

via an AMPK-mediated pathway. SHP directly inhibits cAMP-dependent transcription by binding to CREB, resulting in the reduced association of CREB with CRTC2.[71, 72] The adenovirus-mediated overexpression of SHP could effectively reduce blood glucose levels in diabetic mice, thus showing the importance of this pathway in the control of hepatic glucose metabolism.

These results provide a dual mechanism for a metformin-AMPK dependent pathway to inhibit hepatic gluconeogenesis at the transcriptional level; an acute regulation of CRTC2 phosphorylation to inhibit the CRTC2-CREB-dependent transcriptional circuit; and a longer-term regulation of gluconeogenic transcription by enhanced SHP expression. Both DAX-1 and SMILE were shown to repress hepatic gluconeogenesis by inhibiting HNF4-dependent transcriptional events.[73, 74] SIK1, a member of the AMPK-related kinases, was shown to enhance DAX-1 expression in the liver, whereas Akt was shown to activate the transcription of SMILE to target the HNF4 pathway under feeding conditions. Interestingly, SMILE was shown to directly replace PGC-1α from HNF4 and the gluconeogenic promoters, suggesting that this factor could potentially function as a major transcriptional repressor of hepatic gluconeogenesis in response to insulin signaling. Further study is necessary to fully understand the relative contribution of these nuclear receptors in the control of glucose homeostasis in both physiological conditions and pathological settings.

29.2.7: CONCLUDING REMARKS

In this review, we attempted to describe the current understanding of the regulation of glucose metabolism in the mammalian liver. Under feeding conditions, glucose, a major hexose monomer of dietary carbohydrate, is taken up in the liver and oxidized via glycolysis. The excess glucose that is not utilized as an immediate fuel for energy is stored initially as glycogen and is later converted into triacylglycerols via lipogenesis. Glycogenesis is activated via the insulin-Akt-mediated inactivation of GSK-3, leading to the activation of glycogen synthase and the increased glycogen stores in the liver. Insulin is also critical in the activation of PP1, which functions to dephosphorylate and activate glycogen synthase. In addition, PP1 inhibits glycogenolysis via the dephosphorylation/inactivation of glycogen phosphorylase. Glycolysis is controlled by the regulation of three rate-limiting enzymes: GK, PFK-1 and L-PK. The activities of these enzymes are acutely regulated by allosteric regulators such as ATP, AMP, and F26BP but are also controlled at the transcription level. Two prominent transcription factors are SREBP-1c and ChREBP, which regulate not only the aforementioned glycolytic enzyme genes but also the genes encoding enzymes for fatty acid biosynthesis and triacylglycerol synthesis (collectively termed as lipogenesis).

The importance of these transcription factors in the control of glycolysis and fatty acid biosynthesis has been verified by knockout mouse studies, as described in the main text. The liver also has a critical role in controlling glucose homeostasis under fasting conditions. Initially, insulin counterregulatory hormones such as glucagon and epinephrine are critical in activating the PKA-driven kinase cascades that promote glycogen phosphorylase and glycogenolysis in the liver, thus enabling this tissue to provide enough fuel for peripheral tissues such as the brain, red blood cells and muscles. Subsequently, these hormones together with adrenal cortisol are crucial in initiating the transcriptional activation of gluconeogenesis such as PC, PEPCK and G6Pase. The major transcription factors involved in the pathway include CREB, FoxO1 and members of nuclear receptors, with aid from transcriptional coactivators such as CRTC, PGC-1α and PRMTs. These adaptive responses are critical for maintaining glucose homeostasis in times of starvation in mammals. Further study is necessary by using liver-specific knockout mice for each regulator of hepatic glucose metabolism to provide better insights into the intricate control mechanisms of glucose homeostasis in mammals.

29.2.8: REFERENCES

1. 1

   Nordlie RC, Foster JD, Lange AJ . Regulation of glucose production by the liver. *Annul Rev Nutr* 1999; **19**: 379–406.

   CAS Article Google Scholar

2. 2

   Towle HC, Kaytor EN, Shih HM . Regulation of the expression of lipogenic enzyme genes by carbohydrate. *Annu Rev Nutr* 1997; **17**: 405–433.

   CAS PubMed Article Google Scholar

3. 3

   Roach PJ . Glycogen and its metabolism. *Curr Mol Med* 2002; **2**: 101–120.

   CAS PubMed Article Google Scholar

4. 4

   van de Werve G, Jeanrenaud B . Liver glycogen metabolism: an overview. *Diabetes Metab Rev* 1987; **3**: 47–78.

   CAS PubMed Article Google Scholar

5. 5

   Ros S, Garcia-Rocha M, Dominguez J, Ferrer JC, Guinovart JJ . Control of liver glycogen synthase activity and intracellular distribution by phosphorylation. *The J Biol Chem* 2009; **284**: 6370–6378.

CAS PubMed Article Google Scholar

6. 6

Agius L . Role of glycogen phosphorylase in liver glycogen metabolism. *Mol Aspects Med* 2015; **46**: 34–45.

CAS PubMed Article Google Scholar

7. 7

Pilkis SJ, Claus TH . Hepatic gluconeogenesis/glycolysis: regulation and structure/function relationships of substrate cycle enzymes. *Annu Rev Nutr* 1991; **11**: 465–515.

CAS PubMed Article Google Scholar

8. 8

Pilkis SJ, Claus TH, el-Maghrabi MR . The role of cyclic AMP in rapid and long-term regulation of gluconeogenesis and glycolysis. *Adv Second Messenger Phosphoprotein Res* 1988; **22**: 175–191.

CAS PubMed Google Scholar

9. 9

Pilkis SJ, el-Maghrabi MR, Claus TH . Hormonal regulation of hepatic gluconeogenesis and glycolysis. *Annu Rev Biochem* 1988; **57**: 755–783.

CAS PubMed Article Google Scholar

10. 10

Brouwers MC, Jacobs C, Bast A, Stehouwer CD, Schaper NC . Modulation of glucokinase regulatory protein: a double-edged sword? *Trends Mol Med* 2015; **21**: 583–594.

CAS PubMed Article Google Scholar

11. 11

Dentin R, Girard J, Postic C . Carbohydrate responsive element binding protein (ChREBP) and sterol regulatory element binding protein-1c (SREBP-1c): two key regulators of glucose metabolism and lipid synthesis in liver. *Biochimie* 2005; **87**: 81–86.

CAS Article Google Scholar

12. 12

Horton JD, Goldstein JL, Brown MS . SREBPs: activators of the complete program of cholesterol and fatty acid synthesis in the liver. *J Clin Invest* 2002; **109**: 1125–1131.

CAS PubMed PubMed Central Article Google Scholar

13. 13

Shimano H, Yahagi N, Amemiya-Kudo M, Hasty AH, Osuga J, Tamura Y *et al*. Sterol regulatory element-binding protein-1 as a key transcription factor for nutritional induction of lipogenic enzyme genes. *J Biol Chem* 1999; **274**: 35832–35839.

CAS PubMed PubMed Central Article Google Scholar

14. 14

Im SS, Yousef L, Blaschitz C, Liu JZ, Edwards RA, Young SG *et al*. Linking lipid metabolism to the innate immune response in macrophages through sterol regulatory element binding protein-1a. *Cell Metab* 2011; **13**: 540–549.

CAS PubMed PubMed Central Article Google Scholar

15. 15

Jeon TI, Osborne TF . SREBPs: metabolic integrators in physiology and metabolism. *Trends Endocrinol Metab* 2012; **23**: 65–72.

CAS PubMed Article Google Scholar

16. 16

Repa JJ, Liang G, Ou J, Bashmakov Y, Lobaccaro JM, Shimomura I *et al*. Regulation of mouse sterol regulatory element-binding protein-1c gene (SREBP-1c) by oxysterol receptors, LXRalpha and LXRbeta. *Genes Dev* 2000; **14**: 2819–2830.

CAS PubMed PubMed Central Article Google Scholar

17. 17

Lu M, Shyy JY . Sterol regulatory element-binding protein 1 is negatively modulated by PKA phosphorylation. *Am J Physiol Cell Physiol* 2006; **290**: C1477–C1486.

CAS PubMed Article Google Scholar

18. 18

Bengoechea-Alonso MT, Ericsson J . A phosphorylation cascade controls the degradation of active SREBP1. *J Biol Chem* 2009; **284**: 5885–5895.

CAS Article Google Scholar

19. 19

Yoon YS, Seo WY, Lee MW, Kim ST, Koo SH . Salt-inducible kinase regulates hepatic lipogenesis by controlling SREBP-1c phosphorylation. *J Biol Chem* 2009; **284**: 10446–10452.

CAS PubMed PubMed Central Article Google Scholar

20. 20

Li Y, Xu S, Mihaylova MM, Zheng B, Hou X, Jiang B *et al*. AMPK phosphorylates and inhibits SREBP activity to attenuate hepatic steatosis and atherosclerosis in diet-induced insulin-resistant mice. *Cell Metab* 2011; **13**: 376–388.

CAS PubMed PubMed Central Article Google Scholar

21. 21

Yamashita H, Takenoshita M, Sakurai M, Bruick RK, Henzel WJ, Shillinglaw W *et al*. A glucose-responsive transcription factor that regulates carbohydrate metabolism in the liver. *Proc Natl Acad Sci USA* 2001; **98**: 9116–9121.

CAS PubMed Article Google Scholar

22. 22

Ma L, Tsatsos NG, Towle HC . Direct role of ChREBP.Mlx in regulating hepatic glucose-responsive genes. *J Biol Chem* 2005; **280**: 12019–12027.

CAS PubMed Article Google Scholar

23. 23

Mitro N, Mak PA, Vargas L, Godio C, Hampton E, Molteni V *et al*. The nuclear receptor LXR is a glucose sensor. *Nature* 2007; **445**: 219–223.

CAS PubMed Article Google Scholar

24. 24

Denechaud PD, Bossard P, Lobaccaro JM, Millatt L, Staels B, Girard J *et al*. ChREBP, but not LXRs, is required for the induction of glucose-regulated genes in mouse liver. *J Clin Invest* 2008; **118**: 956–964.

CAS PubMed PubMed Central Google Scholar

25. 25

Kawaguchi T, Osatomi K, Yamashita H, Kabashima T, Uyeda K . Mechanism for fatty acid "sparing" effect on glucose-induced transcription: regulation of carbohydrate-responsive element-binding protein by AMP-activated protein kinase. *J Biol Chem* 2002; **277**: 3829–3835.

CAS PubMed Article Google Scholar

26. 26

Kawaguchi T, Takenoshita M, Kabashima T, Uyeda K . Glucose and cAMP regulate the L-type pyruvate kinase gene by phosphorylation/dephosphorylation of the carbohydrate response element binding protein. *Proc Natl Acad Sci USA* 2001; **98**: 13710–13715.

CAS PubMed Article Google Scholar

27. 27

Tsatsos NG, Davies MN, O'Callaghan BL, Towle HC . Identification and function of phosphorylation in the glucose-regulated transcription factor ChREBP. *Biochem J* 2008; **411**: 261–270.

CAS PubMed Article Google Scholar

28. 28

Tsatsos NG, Towle HC . Glucose activation of ChREBP in hepatocytes occurs via a two-step mechanism. *Biochem Biophys Res Commun* 2006; **340**: 449–456.

CAS PubMed Article Google Scholar

29. 29

Iizuka K, Horikawa Y . ChREBP: a glucose-activated transcription factor involved in the development of metabolic syndrome. *Endocr J* 2008; **55**: 617–624.

CAS PubMed Article Google Scholar

30. 30

Dentin R, Benhamed F, Hainault I, Fauveau V, Foufelle F, Dyck JR *et al*. Liver-specific inhibition of ChREBP improves hepatic steatosis and insulin resistance in ob/ob mice. *Diabetes* 2006; **55**: 2159–2170.

CAS PubMed Article Google Scholar

31. 31

Oh KJ, Han HS, Kim MJ, Koo SH . CREB and FoxO1: two transcription factors for the regulation of hepatic gluconeogenesis. *BMB Rep* 2013; **46**: 567–574.

CAS PubMed PubMed Central Article Google Scholar

32. 32

Arias J, Alberts AS, Brindle P, Claret FX, Smeal T, Karin M *et al*. Activation of cAMP and mitogen responsive genes relies on a common nuclear factor. *Nature* 1994; **370**: 226–229.

CAS PubMed PubMed Central Article Google Scholar

33. 33

Chrivia JC, Kwok RP, Lamb N, Hagiwara M, Montminy MR, Goodman RH . Phosphorylated CREB binds specifically to the nuclear protein CBP. *Nature* 1993; **365**: 855–859.

CAS Article Google Scholar

34. 34

Altarejos JY, Montminy M . CREB and the CRTC co-activators: sensors for hormonal and metabolic signals. *Nat Rev Mol Cell Biol* 2011; **12**: 141–151.

CAS PubMed PubMed Central Article Google Scholar

35. 35

Conkright MD, Canettieri G, Screaton R, Guzman E, Miraglia L, Hogenesch JB *et al*. TORCs: transducers of regulated CREB activity. *Mol Cell* 2003; **12** (2): 413–423.

CAS PubMed Article Google Scholar

36. 36

Koo SH, Flechner L, Qi L, Zhang X, Screaton RA, Jeffries S *et al*. The CREB coactivator TORC2 is a key regulator of fasting glucose metabolism. *Nature* 2005; **437**: 1109–1111.

CAS PubMed PubMed Central Article Google Scholar

37. 37

Erion DM, Ignatova ID, Yonemitsu S, Nagai Y, Chatterjee P, Weismann D *et al*. Prevention of hepatic steatosis and hepatic insulin resistance by knockdown of cAMP response element-binding protein. *Cell Metab* 2009; **10**: 499–506.

CAS PubMed PubMed Central Article Google Scholar

38. 38

Herzig S, Long F, Jhala US, Hedrick S, Quinn R, Bauer A *et al*. CREB regulates hepatic gluconeogenesis through the coactivator PGC-1. *Nature* 2001; **413**: 179–183.

CAS PubMed PubMed Central Article Google Scholar

39. 39

Bedford DC, Kasper LH, Wang R, Chang Y, Green DR, Brindle PK . Disrupting the CH1 domain structure in the acetyltransferases CBP and p300 results in lean mice with increased metabolic control. *Cell Metab* 2011; **14**: 219–230.

CAS PubMed PubMed Central Article Google Scholar

40. 40

Screaton RA, Conkright MD, Katoh Y, Best JL, Canettieri G, Jeffries S *et al*. The CREB coactivator TORC2 functions as a calcium- and cAMP-sensitive coincidence detector. *Cell* 2004; **119**: 61–74.

CAS Article Google Scholar

41. 41

Wang Y, Li G, Goode J, Paz JC, Ouyang K, Screaton R *et al*. Inositol-1,4,5-trisphosphate receptor regulates hepatic gluconeogenesis in fasting and diabetes. *Nature* 2012; **485**: 128–132.

CAS PubMed PubMed Central Article Google Scholar

42. 42

Yoon YS, Lee MW, Ryu D, Kim JH, Ma H, Seo WY *et al*. Suppressor of MEK null (SMEK)/protein phosphatase 4 catalytic subunit (PP4C) is a key regulator of hepatic gluconeogenesis. *Proc Natl Acad Sci USA a* 2010; **107**: 17704–17709.

CAS Article Google Scholar

43. 43

Dentin R, Hedrick S, Xie J, Yates J 3rd, Montminy M . Hepatic glucose sensing via the CREB coactivator CRTC2. *Science* 2008; **319**: 1402–1405.

CAS PubMed PubMed Central Article Google Scholar

44. 44

Han HS, Jung CY, Yoon YS, Choi S, Choi D, Kang G *et al*. Arginine methylation of CRTC2 is critical in the transcriptional control of hepatic glucose metabolism. *Sci Signal* 2014; **7**: ra19.

PubMed Article Google Scholar

45. 45

Wang Y, Inoue H, Ravnskjaer K, Viste K, Miller N, Liu Y *et al*. Targeted disruption of the CREB coactivator Crtc2 increases insulin sensitivity. *Proc Natl Acad Sci USA* 2010; **107**: 3087–3092.

CAS PubMed Article Google Scholar

46. 46

Lee MW, Chanda D, Yang J, Oh H, Kim SS, Yoon YS *et al*. Regulation of hepatic gluconeogenesis by an ER-bound transcription factor, CREBH. *Cell Metab* 2010; **11**: 331–339.

CAS PubMed PubMed Central Article Google Scholar

47. 47

Wang Y, Vera L, Fischer WH, Montminy M . The CREB coactivator CRTC2 links hepatic ER stress and fasting gluconeogenesis. *Nature* 2009; **460**: 534–537.

CAS PubMed PubMed Central Article Google Scholar

48. 48

Accili D, Arden KC . FoxOs at the crossroads of cellular metabolism, differentiation, and transformation. *Cell* 2004; **117**: 421–426.

CAS PubMed Article Google Scholar

49. 49

Barthel A, Schmoll D, Unterman TG . FoxO proteins in insulin action and metabolism. *Trends Endocrinol Metab* 2005; **16**: 183–189.

CAS Article Google Scholar

50. 50

Biggs WH 3rd, Meisenhelder J, Hunter T, Cavenee WK, Arden KC . Protein kinase B/Akt-mediated phosphorylation promotes nuclear exclusion of the winged helix transcription factor FKHR1. *Proc Natl Acad Sci USA* 1999; **96**: 7421–7426.

CAS Article Google Scholar

51. 51

Brunet A, Bonni A, Zigmond MJ, Lin MZ, Juo P, Hu LS *et al*. Akt promotes cell survival by phosphorylating and inhibiting a Forkhead transcription factor. *Cell* 1999; **96**: 857–868.

CAS Article Google Scholar

52. 52

Nakae J, Park BC, Accili D . Insulin stimulates phosphorylation of the forkhead transcription factor FKHR on serine 253 through a Wortmannin-sensitive pathway. *J Biol Chem* 1999; **274**: 15982–15985.

CAS PubMed Article Google Scholar

53. 53

Daitoku H, Hatta M, Matsuzaki H, Aratani S, Ohshima T, Miyagishi M *et al*. Silent information regulator 2 potentiates Foxo1-mediated transcription through its deacetylase activity. *Proc Natl Acad Sci USA* 2004; **101**: 10042–10047.

CAS PubMed Article Google Scholar

54. 54

Haeusler RA, Kaestner KH, Accili D . FoxOs function synergistically to promote glucose production. *J Biol Chem* 2010; **285**: 35245–35248.

CAS PubMed PubMed Central Article Google Scholar

55. 55

Matsumoto M, Pocai A, Rossetti L, Depinho RA, Accili D . Impaired regulation of hepatic glucose production in mice lacking the forkhead transcription factor Foxo1 in liver. *Cell Metab* 2007; **6**: 208–216.

CAS PubMed Article Google Scholar

56. 56

Puigserver P, Rhee J, Donovan J, Walkey CJ, Yoon JC, Oriente F *et al*. Insulin-regulated hepatic gluconeogenesis through FOXO1-PGC-1alpha interaction. *Nature* 2003; **423**: 550–555.

CAS PubMed PubMed Central Article Google Scholar

57. 57

Puigserver P, Wu Z, Park CW, Graves R, Wright M, Spiegelman BM . A cold-inducible coactivator of nuclear receptors linked to adaptive thermogenesis. *Cell* 1998; **92**: 829–839.

CAS Article Google Scholar

58. 58

Yoon JC, Puigserver P, Chen G, Donovan J, Wu Z, Rhee J *et al*. Control of hepatic gluconeogenesis through the transcriptional coactivator PGC-1. *Nature* 2001; **413**: 131–138.

CAS PubMed PubMed Central Article Google Scholar

59. 59

Koo SH, Satoh H, Herzig S, Lee CH, Hedrick S, Kulkarni R *et al*. PGC-1 promotes insulin resistance in liver through PPAR-alpha-dependent induction of TRB-3. *Nat Med* 2004; **10**: 530–534.

CAS PubMed PubMed Central Article Google Scholar

60. 60

Lin J, Wu PH, Tarr PT, Lindenberg KS, St-Pierre J, Zhang CY *et al*. Defects in adaptive energy metabolism with CNS-linked hyperactivity in PGC-1alpha null mice. *Cell* 2004; **119**: 121–135.

CAS PubMed PubMed Central Article Google Scholar

61. 61

Yamagata K, Daitoku H, Takahashi Y, Namiki K, Hisatake K, Kako K *et al*. Arginine methylation of FOXO transcription factors inhibits their phosphorylation by Akt. *Mol Cell* 2008; **32**: 221–231.

CAS Article Google Scholar

62. 62

Choi D, Oh KJ, Han HS, Yoon YS, Jung CY, Kim ST *et al*. Protein arginine methyltransferase 1 regulates hepatic glucose production in a FoxO1-dependent manner. *Hepatology* 2012; **56**: 1546–1556.

CAS PubMed Article Google Scholar

63. 63

Evans RM, Mangelsdorf DJ . Nuclear Receptors, RXR, and the Big Bang. *Cell* 2014; **157**: 255–266.

CAS PubMed PubMed Central Article Google Scholar

64. 64

van Schaftingen E, Gerin I . The glucose-6-phosphatase system. *Biochem J* 2002; **362**: 513–532.

CAS PubMed PubMed Central Article Google Scholar

65. 65

Zinker B, Mika A, Nguyen P, Wilcox D, Ohman L, von Geldern TW *et al*. Liver-selective glucocorticoid receptor antagonism decreases glucose production and increases glucose disposal, ameliorating insulin resistance. *Metabolism* 2007; **56**: 380–387.

CAS PubMed Article Google Scholar

66. 66

Kim DK, Gang GT, Ryu D, Koh M, Kim YN, Kim SS *et al*. Inverse agonist of nuclear receptor ERRgamma mediates antidiabetic effect through inhibition of hepatic gluconeogenesis. *Diabetes* 2013; **62**: 3093–3102.

CAS PubMed PubMed Central Article Google Scholar

67. 67

Kim DK, Ryu D, Koh M, Lee MW, Lim D, Kim MJ *et al*. Orphan nuclear receptor estrogen-related receptor gamma (ERRgamma) is key regulator of hepatic gluconeogenesis. *J Biol Chem* 2012; **287**: 21628–21639.

CAS PubMed PubMed Central Article Google Scholar

68. 68

Iyer AK, McCabe ER . Molecular mechanisms of DAX1 action. *Mol Genet Metab* 2004; **83**: 60–73.

CAS PubMed Article Google Scholar

69. 69

Seol W, Choi HS, Moore DD . An orphan nuclear hormone receptor that lacks a DNA binding domain and heterodimerizes with other receptors. *Science* 1996; **272**: 1336–1339.

CAS Article Google Scholar

70. 70

Xie YB, Lee OH, Nedumaran B, Seong HA, Lee KM, Ha H *et al*. SMILE, a new orphan nuclear receptor SHP-interacting protein, regulates SHP-repressed estrogen receptor transactivation. *Biochem J* 2008; **416**: 463–473.

CAS PubMed Article Google Scholar

71. 71

Kim YD, Park KG, Lee YS, Park YY, Kim DK, Nedumaran B *et al*. Metformin inhibits hepatic gluconeogenesis through AMP-activated protein kinase-dependent regulation of the orphan nuclear receptor SHP. *Diabetes* 2008; **57**: 306–314.

CAS PubMed Article Google Scholar

72. 72

Lee JM, Seo WY, Song KH, Chanda D, Kim YD, Kim DK *et al*. AMPK-dependent repression of hepatic gluconeogenesis via disruption of CREB.CRTC2 complex by orphan nuclear receptor small heterodimer partner. *J Biol Chem* 2010; **285**: 32182–32191.

CAS PubMed PubMed Central Article Google Scholar

73. 73

Nedumaran B, Hong S, Xie YB, Kim YH, Seo WY, Lee MW *et al*. DAX-1 acts as a novel corepressor of orphan nuclear receptor HNF4alpha and negatively regulates gluconeogenic enzyme gene expression. *J Biol Chem* 2009; **284**: 27511–27523.

CAS PubMed PubMed Central Article Google Scholar

74. 74

Lee JM, Seo WY, Han HS, Oh KJ, Lee YS, Kim DK *et al*. Insulin-inducible SMILE inhibits hepatic gluconeogenesis. *Diabetes* 2016; **65**: 62–73.

CAS PubMed Google Scholar

Download references

29.2.9: ACKNOWLEDGEMENTS

This work is supported by the National Research Foundation of Korea (grant nos.: NRF-2012M3A9B6055345, NRF-2015R1A5A1009024 and NRF- 2015R1A2A1A01006687), funded by the Ministry of Science, ICT & Future Planning, Republic of Korea, a grant of the Korean Health technology R&D Project (grant no: HI13C1886), Ministry of Health & Welfare, Republic of Korea and a grant from Korea University, Seoul, Republic of Korea.

29.2.10: AUTHOR INFORMATION

29.2.10.1: AFFILIATIONS

1. **Division of Life Sciences, College of Life Sciences & Biotechnology, Korea University, Seoul, 136-713, Korea**

29.2.10.2: CORRESPONDING AUTHOR

Correspondence to Seung-Hoi Koo.

29.2.11: ETHICS DECLARATIONS

29.2.11.1: COMPETING INTERESTS

The authors declare no conflict of interest.

29.2.12: RIGHTS AND PERMISSIONS

29.3: EMERGING ROLE OF THE BRAIN IN THE HOMEOSTATIC REGULATION OF ENERGY AND GLUCOSE METABOLISM

Roh, E., Song, D. & Kim, MS. Emerging role of the brain in the homeostatic regulation of energy and glucose metabolism. *Exp Mol Med* **48,** e216 (2016). https://doi.org/10.1038/emm.2016.4

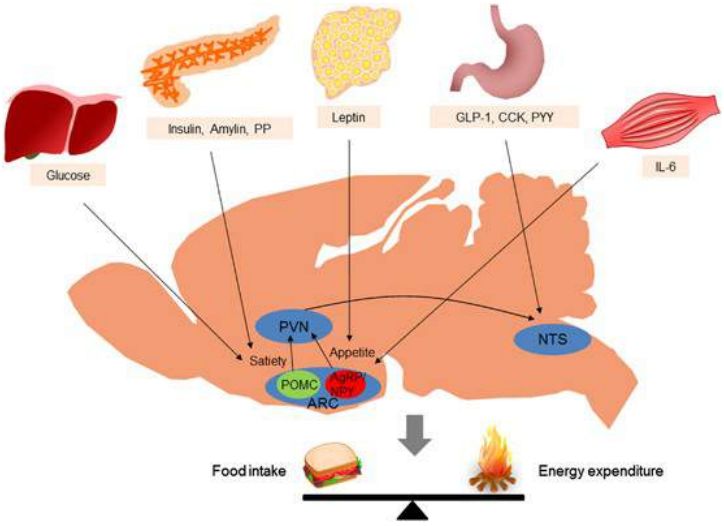
29.3.1: ABSTRACT

Accumulated evidence from genetic animal models suggests that the brain, particularly the hypothalamus, has a key role in the homeostatic regulation of energy and glucose metabolism. The brain integrates multiple metabolic inputs from the periphery through nutrients, gut-derived satiety signals and adiposity-related hormones. The brain modulates various aspects of metabolism, such as food intake, energy expenditure, insulin secretion, hepatic glucose production and glucose/fatty acid metabolism in adipose tissue and skeletal muscle. Highly coordinated interactions between the brain and peripheral metabolic organs are critical for the maintenance of energy and glucose homeostasis. Defective crosstalk between the brain and peripheral organs contributes to the development of obesity and type 2 diabetes. Here we comprehensively review the above topics, discussing the main findings related to the role of the brain in the homeostatic regulation of energy and glucose metabolism.

29.3.2: CENTRAL REGULATION OF ENERGY METABOLISM

In normal individuals, food intake and energy expenditure are tightly regulated by homeostatic mechanisms to maintain energy balance. Substantial evidence indicates that the brain, particularly the hypothalamus, is primarily responsible for the regulation of energy homeostasis.[1] The brain monitors changes in the body energy state by sensing alterations in the plasma levels of key metabolic hormones and nutrients. Specialized neuronal networks in the brain coordinate adaptive changes in food intake and energy expenditure in response to altered metabolic conditions (Figure 1).[2, 3]

Figure 1



Integration of peripheral metabolic signals andthe central nervous system maintains energy homeostasis. The brain integrates metabolic signals from peripheral tissues such as the liver, pancreas, adipose tissue, gut and muscle. Specialized neuronal networks in the brain coordinate adaptive changes in food intake and energy expenditure in response to altered metabolic conditions. Neuropeptide Y/agouti-related protein- and proopiomelanocortin-producing neurons in the hypothalamic arcuate nucleus primarily sense the body energy state. These neurons project to other hypothalamic nuclei and to the nucleus of the solitary tract in the brain stem to control multiple aspects of the homeostatic regulation of energy balance. ARC, arcuate nucleus; CCK, cholecystokinin; GLP-1, glucagon-like peptide-1; IL-6, interleukin-6; PP, pancreatic polypeptide; PVN, paraventricular nucleus; PYY, peptide YY.

Full size image

29.3.2.1: BRAIN REGULATION OF FOOD INTAKE

The hypothalamus is considered a key organ in the regulation of food intake. The hypothalamic arcuate nucleus (ARC) is adjacent to the median eminence, one of the circumventricular organs, and surrounds the third cerebroventricle. Thus, hormones and nutrients in the systemic circulation and the cerebrospinal fluid can easily access the ARC. Anatomically, the ARC is considered a hypothalamic area that primarily senses metabolic signals from the periphery via the systemic circulation.[4] In the ARC, there are two distinct neuronal populations: one group of neurons produces the orexigenic neuropeptides neuropeptide Y (NPY) and agouti-related peptide (AgRP), whereas the other subset of neurons expresses the anorexigenic neuropeptides proopiomelanocortin (POMC), and cocaine- and amphetamine-regulated transcript. These neurons are the first-order neurons on which peripheral metabolic hormones, including leptin, insulin, ghrelin and nutrients, primarily act.[5] The anorexigenic effect of monoamine serotonin is also mediated by the 5HT-2C receptor in POMC neurons.[6] POMC neurons project axonal processes to second-order neurons in hypothalamic areas such as the paraventricular nucleus (PVN), ventromedial hypothalamus (VMH) and lateral hypothalamus (LH), and to autonomic preganglionic neurons in the brain stem and spinal cord.[7]

The anorexigenic neuropeptide α-melanocyte-stimulating hormone (α-MSH) is produced by posttranscriptional processing of POMC and is released from the presynaptic terminals of POMC neurons. Upon binding to the melanocortin-3 and -4 receptors (MC3R and MC4R) on second-order neurons, α-MSH activates catabolic pathways, leading to reduced food intake and increased energy expenditure. Targeted deletion of the MC4R in mice induces hyperphagia, reduces energy expenditure and leads to obesity.[8] In humans, MC4R mutations account for ~6% of severe early-onset obesity cases,[9] suggesting an important role for the central melanocortin system in the maintenance of normal body weight.

The endogenous MC-3/4R antagonist AgRP is released from the terminals of NPY/AgRP-producing neurons to the synaptic space of second-order neurons where it competes with α-MSH for MC3Rs and MC4Rs and antagonizes its effects.[10] Selective ablation of NPY/AgRP neurons in young mice results in a significant decrease in food intake and body weight,[11] suggesting that these neurons are critical for promoting food intake and preventing weight loss. Administration of NPY stimulates food intake via Y1 or Y5 receptors.[12] NPY is required for the rapid stimulation of feeding, whereas AgRP stimulates feeding over a prolonged period.[13]

PVN neurons synthesize and secrete neuropeptides that have a net catabolic action, including corticotrophin-releasing hormone, thyrotropin-releasing hormone, somatostatin, vasopressin and oxytocin. On the other hand, PVN neurons control sympathetic outflow to peripheral metabolic organs, resulting in increased fatty acid oxidation and lipolysis.[14] Destruction of PVN and haploinsufficiency of Sim1, a critical transcriptional factor in the development of PVN, cause hyperphagia and obesity,[15] implying a inhibitory role of the PVN in food intake and weight gain.

The VMH mainly receives neuronal projections from the ARC and projects their axons to the ARC, dorsomedial nucleus (DMN), LH and brain stem regions. The VMH contains neurons that sense glucose and leptin.[16, 17] Moreover, the anorexigenic neuropeptide brain-derived neurotrophic factor is produced in the VMH.[18] Destruction of the VMH causes hyperphagia, obesity and hyperglycemia.[19] Thus, the VMH is regarded a pivotal area in generating satiety and maintaining glucose homeostasis. The DMN contains a high level of NPY terminals and α-MSH terminals originating from the ARC.[20] Destruction of the DMN also results in hyperphagia and obesity.[21]

In contrast to the PVN, VMH and DMN, destruction of the LH leads to hypophagia and weight loss. Therefore, LH is considered a feeding center. LH contains two neuronal populations producing the orexigenic neuropeptides melanin-concentrating hormone (MCH) and orexin, also called hypocretin. NPY/AgRP- and α-MSH-immunoreactive terminals from ARC neurons are in contact with MCH- and orexin-expressing neurons.[22] Orexin-producing neurons are also involved in glucose sensing and the regulation of sleep–awake cycles.[23] Alterations in the orexin receptor-2 and orexin genes produce narcolepsy in animal models and humans.[24] On the other hand, depletion of MCH or the MCH-1 receptor in mice attenuates weight gain, suggesting that MCH is an endogenous orexigenic molecule.[25]

The brain stem is another key brain area involved in the regulation of food intake. Satiety signals from the gastrointestinal tract are relayed to the nucleus tractus solitaries (NTS) through the sensory vagus nerve, a major neuronal connection between the gut and brain. Transection of sensory vagal fibers decreases meal size and meal duration, confirming that vagal afferents transfer meal-related signals to the brain.[26] Like the ARC, the NTS is anatomically close to the area postrema, another circumventricular organ.[27] Therefore, the NTS is perfectly located for receiving both humoral and neural signals. Meanwhile, the NTS receives extensive neuronal projections from the PVN and vice versa,[28] indicating that there are intimate communications between the hypothalamus and the brain stem. Like hypothalamic neurons, NTS neurons produce appetite-regulating glucagon-like peptide-1 (GLP-1), NPY and POMC, and sense peripheral metabolic signals.[29] For instance, NTS POMC neurons show the signal transducer and activator of transcription 3 (STAT3) activation in response to exogenous leptin.[30] Thus, circulating hormones and nutrients may relay metabolic signals to the brain by acting on both the hypothalamus and brain stem.

On the other hand, the brain reward system is involved in the control of hedonic feeding, that is, intake of palatable foods. Like other addiction behaviors, the mesolimbic and mesocortical dopaminergic pathways are involved in hedonic feeding. Intake of palatable foods elicits dopamine release in the ventral tegmental area (VTA), which in turn activates the neural pathways from the VTA to the nucleus accumbens via the medial forebrain bundles. Interestingly, hedonic feeding is modulated by metabolic signals. Leptin acts on the

dopaminergic neurons in the VTA to suppress feeding.[31] Conversely, hedonic feeding can override satiety signals. Mice lacking the D_2 receptor are more sensitive to leptin.[32]

29.3.2.1: BRAIN REGULATION OF ENERGY EXPENDITURE

The brain modulates various processes that consume energy, such as locomotor activity, fatty acid oxidation in the skeletal muscle and thermogenesis.[33] Tumor growth factor-α, produced in the suprachiasmatic nucleus in a circadian manner, strongly inhibits locomotor activity by acting on the epidermal growth factor receptors expressed in the hypothalamic subparaventricular zone.[34] Orexin-A produced by LH neurons promotes locomotor activity and wakefulness through orexin-1 and orexin-2 receptors.[35] A role for orexin in food-seeking behavior in food-deprived conditions has been suggested.[36] Leptin stimulates locomotor activity via a mechanism that depends on hypothalamic POMC neurons[37]. Leptin also enhances fatty acid oxidation in skeletal muscle via both central and peripheral mechanisms.[38]

Thermogenesis is theprocess that dissipates energy as heat to maintain body temperature. Thermogenesis mainly occurs in brown adipose tissue (BAT).[39] Brown fat-like adipocytes, so-called browning of white adipose tissue (WAT), are found in the subcutaneous inguinal WAT under certain circumstances. Cold exposure or intracerebroventricular (ICV) coinjection of insulin and leptin induces WAT browning.[40] Induction of WAT browning results in increased energy expenditure and attenuation of diet-induced obesity in mice. Conversely, inhibition of WAT browning by depletion of Prdm16 leads to obesity.[39]

The brain regulates BAT thermogenesis through modulation of the sympathetic nervous system. Norepinephrine released from sympathetic nervous terminals acts on the β3-adrenergic receptors in adipocytes in the BAT and inguinal fat pads. Activation of adrenergic receptors triggers cyclic-adenosine monophosphate signaling, which in turn increases the expression of uncoupling protein-1 in the mitochondria. BAT thermogenesis is important for maintaining body temperature in response to cold exposure and dissipating excess energy after high-calorie intake. Because metabolic fuel substrates such as glucose and fatty acid are consumed during BAT thermogenesis, BAT thermogenesis can affect body weight and body fat mass.[41] In the past, BAT was thought to be present only in human infants. However, ^{18}F-fluorodeoxyglucose positron emission tomography revealed the presence of BAT in the adult humans. Human BAT depots are distributed in the supraclavicular area and in perivascular and periviscus areas (for example, around the heart, airway, gut, liver and adrenal gland) of the chest and abdomen.[42] BAT activity, determined by ^{18}F-fluorodeoxyglucose positron emission tomography, is affected by outdoor temperature, age, sex, body mass index and the coexistence of diabetes. Because the amount of BAT is inversely correlated with body mass index, especially in older subjects, a potential role of BAT in adult human metabolism has been suggested.[43]

In thermogenic regulation, the hypothalamus integrates the sensation of body temperature with efferent sympathetic outflow. Hypothalamic areas such as the prooptic area, VMH, DMN and ARC modulate thermogenic activity by influencing the sympathetic nervous system.[44] The prooptic area is an important area in the control of body temperature.[45] VMH was the first hypothalamic nucleus to be studied regarding the regulation of BAT activity. The DMN also contains sympathoexcitatory neurons,[46] which regulate thermogenic activity.[47, 48] BAT thermogenesis is also related to the ARC melanocortin system because α-MSH stimulates BAT activity.[49]

Hormonal- and nutrient-mediated metabolic signals can influence sympathetic outflow to the BAT. Central administration of leptin, MC3/4R agonist, glucagon and GLP-1 stimulates BAT activity.[50, 51] Central administration of insulin either stimulates or inhibits BAT thermogenesis, depending on the insulin dose. Central administration of high doses of insulin increases sympathetic nerve activity in the BAT, whereas low doses of insulin decrease it.[52, 53] Food consumption or dietary composition also affects BAT thermogenesis. Although the mechanism of postprandial thermogenesis is unclear, norepinephrine turnover in the BAT is increased after a meal.[54] Glucose administration increases thermogenesis, whereas fasting or food restriction inhibits thermogenesis. Low-protein diet and high-fat diet increase BAT activity.[55]

29.3.2.1: PERIPHERAL SIGNALS MODULATING ENERGY METABOLISM

29.3.2.0.1: ADIPOSITY SIGNALS

Adiposity signals refer to the peripheral signals that circulate in proportion to the total amount of stored fat and inform the brain about the stored energy state. They modulate energy balance through the regulation of food intake and energy expenditure.[2, 56] Insulin is a hormone that was first identified as an adiposity signal.[5, 57] Insulin is secreted by β-cells in response to energy flux. Plasma insulin concentrations increase in proportion to the amount of stored fat.[58] When insulin is administered directly into the central nervous system, it induces a dose-dependent reduction in food intake and body weight.[59, 60] Thus, insulin is thought to signal adiposity to the brain. In hypothalamic neurons, insulin activates the insulin receptor substrate-2 (IRS2)–phosphatidylinositol 3-kinase (PI3K) signaling pathway. Neuronal deletion of insulin receptor and IRS2 results in increased food intake and susceptibility to diet-induced obesity.[61, 62]

The adipose tissue-derived hormone leptin was discovered by positional cloning of the obesity locus (*ob*) in 1994.[63] Leptin is now considered a representative adiposity signal.[64] The receptors activated by leptin are highly expressed in several regions of brain, including the hypothalamus.[65] Genetic deficiency in leptin or the long-form leptin receptor (LepRb) is associated with hyperphagia, hypoactivity and obesity.[66] Of several brain regions, the ARC is an important area that mediates leptin actions. Injection of leptin directly into the ARC reduces food intake and body weight.[67] Leptin also stimulates locomotion through signaling in POMC neurons.[37] Consistently, ICV administration of leptin in leptin-deficient (*ob*/*ob*) mice attenuates obesity.[66] In hypothalamic neurons, leptin provokes several signaling cascades such as the Janus kinase–STAT pathway, IRS–PI3K signaling, the mammalian target of rapamycin–S6 kinase signaling, AMP-

activated protein kinase (AMPK) signaling and ERK signaling.[68] Of those, STAT3 signaling represents hypothalamic leptin signaling and is frequently used as a marker of leptin signaling activity.

29.3.2.0.2: NUTRIENT SIGNALS

Nutrients such as glucose, fatty acids and amino acids provide information on nutrient availability to the brain. Glucose signals the presence of anenergy supply to the brain, whereas hypoglycemia signals an energy deficit.[69] Thus, central administration of glucose and long-chain fatty acids decreases food intake.[70] In contrast, ICV administration of the glucose anti-metabolite 2-deoxy-D-glucose increases food intake.[71] The malonyl-CoA content in hypothalamic neurons has been suggested to be a fuel gauge.[56, 72] Administration of the fatty acid synthase inhibitor C75 induces accumulation of malonyl-CoA in hypothalamic neurons, leading to decreased food intake and body weight.[73] Long-chain fatty acyl-CoA (LCFA-CoA) content in hypothalamic neurons also acts as a cellular nutrient sensor. An increased hypothalamic LCFA-CoA level due to ICV long-chain fatty acid (LCFA) administration leads to decreased food intake.[70] Hypothalamic inhibition of carnitine palmitoyltransferase-1 inhibits food intake by elevating LCFA-CoA content in hypothalamic neurons.[74]

29.3.2.0.3: GASTROINTESTINAL SIGNALS

Hormones secreted by the gut in response to a meal provide information on energy intake. Cholecystokinin, peptide YY and GLP-1 released from the gut induce satiety by acting on the vagus nerve or in the brain.[75] For example, GLP-1 is secreted from intestinal L-cells after a meal. GLP-1 receptors are prevalent in vagus nerve terminals,[76] as well as in the central nucleus of the amygdala, the PVN and ARC of the hypothalamus, and the caudal brain stem.[77] Both central and peripheral administration of GLP-1 promotes satiety.[78, 79] In contrast, ghrelin is secreted by the stomach during a fast and promotes food intake.[80]

29.3.2.0.4: SIGNALS FROM OTHER ORGANS

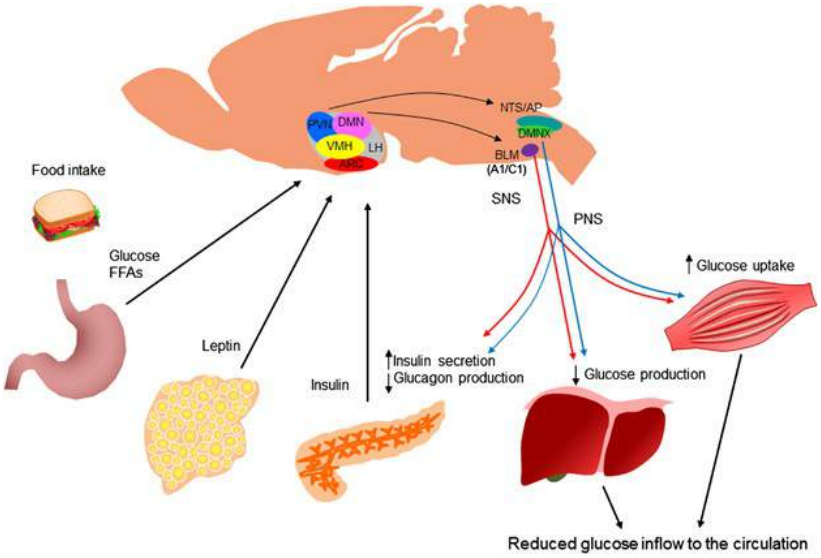
Interleukin-6 (IL-6) is synthesized and released from contracting skeletal muscle during exercise. The elevation in the plasma IL-6 concentration during exercise correlates with exercise intensity and duration and the muscle mass recruited.[81] IL-6 enters the brain across the blood–brain barrier. IL-6 may mobilize fat from storage sites to provide energy to the muscle. ICV administration of IL-6 stimulates energy expenditure, and mice lacking IL-6 develop mature-onset obesity.[82]

Hormones secreted from the endocrine pancreas are also involved in energy homeostasis. Insulin and amylin are co-secreted by β-cells. Like insulin, amylin acts as a satiety signal and reduces food intake via amylin receptors in the area postrema. Other brain sites mediating amylin action include the NTS and the lateral parabrachial nucleus.[83] Amylin also acts as an adiposity signal because amylin levels are well correlated with body fat content. Glucagon, a counter-regulatory hormone to insulin, is secreted from α-cells. Glucagon reduces meal size by acting on the vagus nerves and stimulates energy expenditure through central and peripheral mechanisms.[84] Pancreatic polypeptide is also secreted from the endocrine pancreas. Pancreatic polypeptide regulates gastric motility, pancreatic exocrine secretion and food intake. Systemic administration of pancreatic polypeptide reduces food intake and weight gain.[85] The anorectic effect of pancreatic polypeptide is mediated by Y4 receptors in the dorsal vagal complex.[86]

29.3.3: BRAIN REGULATION OF GLUCOSE METABOLISM

The earliest demonstration of the role of the brain in glucose homeostasis was provided by the physiologist Claude Bernard in 1854. Dr Bernard demonstrated that a puncture in the floor of the fourth ventricle of the rabbit brain resulted in glycosuria.[87] In the past few decades, the concept of central regulation of glucose metabolism has been further established by the subsequent discovery of glucose-sensing neurons in the hypothalamus[88, 89] and the demonstration of their roles in maintaining normal glucose levels.[90] A specialized neuronal population in the brain senses hormones (insulin and leptin) and nutrients (glucose and fatty acids) to regulate glucose homeostasis. The major sites of convergence of these metabolic signals are the hypothalamus and brain stem (Figure 2).

Figure 2



Brain regulation of glucose homeostasis. The brain senses peripheral metabolic signals through hormones (insulin, leptin and so on) and nutrients (glucose, free fatty acids and so on) to regulate glucose metabolism. The sites of the convergence of these metabolic signals are the hypothalamus and brain stem. The autonomic nervous system intervenes in the brain and peripheral metabolic organs to modulate pancreatic insulin/glucagon secretion, hepatic glucose production and skeletal muscle glucose uptake. AP, area postrema; ARC, arcuate nucleus; BLM, basolateral medulla; DMN, dorsomedial nucleus; DMNX, dorsal motor nucleus of the vagus; FFA, free fatty acids; LH, lateral hypothalamus; NTS, nucleus of the solitary tract; PNS, parasympathetic nervous system; PVN, paraventricular nucleus; SNS, sympathetic nervous system; VMH, ventromedial hypothalamus.

Full size image

29.3.3.1: NEURONAL POPULATIONS CONTROLLING GLUCOSE METABOLISM

Brain regions related to the control of glucose metabolism contain neurons whose excitability changes with alterations in glucose concentrations in the extracellular fluid. These glucose-sensing neurons are found in the hypothalamic nuclei and brain stem, which are also important areas in the control of energy balance. Glucose-sensing neurons are subgrouped into two types. Glucose-excited neurons are excited when extracellular glucose levels increase. In contrast, glucose-inhibited neurons are activated by a fall in extracellular glucose concentrations.[91] Glucose-excited neurons are mostly located in the VMH, the ARC and the PVN,[92] whereas glucose-inhibited neurons are distributed in the LH, ARC and PVN.[89, 91] Both types of neurons are also located in the dorsal vagal complex in the brain stem, which encompasses the NTS, area postrema and the dorsal motor nucleus of the vagus.[93, 94, 95]

29.3.3.1: PERIPHERAL SIGNALS AFFECTING BRAIN REGULATION OF GLUCOSE METABOLISM

29.3.3.0.1: INSULIN

During the past decade, the brain has been recognized to be a site of insulin action with regard to glucose homeostasis. Obici *et al.*[96] showed that insulin acts on the brain to modulate hepatic glucose metabolism. They showed, by injecting insulin receptor antisense oligonucleotides into the cerebroventricle, that inhibition of central insulin action impaired insulin-mediated suppression of hepatic glucose production (HGP) during hyperinsulinemic clamp studies in rats. They also demonstrated that infusion of insulin into the cerebroventricle suppressed HGP, irrespective of circulating insulin levels. Moreover, central administration of insulin antibodies or inhibitors of the downstream signaling of insulin diminished the ability of insulin to inhibit glucose production.[97] The hypothalamic insulin signaling pathway was investigated in subsequent studies. Overexpression of the insulin signaling molecules IRS2 and Akt in the hypothalamus enhances the glucose-lowering effect of insulin in streptozotocin-induced diabetic rats.[98] These data support a role for hypothalamic insulin actions in controlling glucose metabolism in peripheral organs.

The ATP-sensitive potassium (K_{ATP}) channel mediates insulin actions in hypothalamic neurons.[99] Activation of neuronal K_{ATP} channels by ICV injection of a K_{ATP} channel activator (diazoxide) lowers glucose production,[100] whereas infusion of a K_{ATP} blocker (sulfonylurea) negates the glucose production-lowering effect of centrally and peripherally administered insulin.[96, 100] Moreover, mice lacking the sulfonylurea receptor subunit SUR1 of the K_{ATP} channel show a diminished response to central insulin action.[100] Vagal efferent fibers constitute the brain–liver axis of insulin actions because hepatic vagotomy blocks central insulin actions.[100] Interestingly, ICV infusion of insulin increases hepatic IL-6 expression, which leads to the activation of hepatic STAT3 signaling.[101] Activated STAT3 inhibits FoxO1 activity and gluconeogenic gene expression in the liver. Collectively, central insulin actions are mediated via neuronal K_{ATP} channel–vagus nerve–hepatic IL6/STAT3 signaling, although the detailed mechanisms involved remain to be determined.

29.3.3.0.2: LEPTIN

Leptin has an important role in the control of glucose metabolism. A lack of leptin (*ob/ob* mice) or its functional receptor (*db/db* mice) leads not only to obesity, but also metabolic derangement, including insulin resistance and diabetes.[102] Leptin treatment of *ob/ob* mice improves glucose homeostasis.[103, 104] Notably, acute leptin treatment via both systemic and central routes in *ob/ob* mice restores glucose metabolism independently of changes in food intake and adiposity.[105, 106] Consistently, leptin-treated *ob/ob* mice display a marked reduction in serum glucose and insulin concentrations.[107] Leptin treatment in lipodystrophy mice improves insulin resistance and hyperglycemia independently of food intake.[108, 109] Thus, leptin regulates glucose homeostasis independently of its anorectic effects.

The hypothalamus is a key site of action of leptin-mediated control of glucose metabolism. ICV administration of leptin in the lipodystrophy mice model corrects insulin resistance and improves impaired insulin signaling in the liver. In contrast, peripheral injection of the same dose of leptin did not have a similar effect.[110] Acute ICV injection of leptin suppresses glycogenolysis and reduces hepatic insulin resistance induced by high-fat feeding.[111] Restoration of leptin signaling in the unilateral ARC by viral gene therapy in leptin receptor-null mice markedly improves hyperinsulinemia and normalizes blood glucose levels, with a mild decrease in body weight and food intake. These data demonstrate that leptin signaling in the ARC is critical for the maintenance of glucose homeostasis.[112]

Leptin-mediated regulation of glucose metabolism is mediated by hypothalamic STAT3 and PI3K signaling pathways. As in *db/db* mice, *s/s* mice with a mutated leptin receptor, which are unable to activate STAT3, exhibit severe hepatic insulin resistance.[113] Blockade of leptin-induced STAT3 activation in the hypothalamus abolishes the suppressive effect of leptin on HGP, confirming the importance of leptin-induced STAT3 signaling.[113] Conversely, hypothalamic deletion of suppressor of cytokine signaling 3, a negative regulator of STAT3 signaling, enhances leptin sensitivity and improves glucose metabolism.[114] On the other hand, reconstitution of leptin receptors in the ARC of leptin-receptor-deficient fa^k/fa^k rats improves insulin sensitivity, which is attenuated by ICV infusion of PI3K inhibitor. Consistently, ARC expression of constitutively active Aktin fa^k/fa^k rats mimics the effect of restored hypothalamic leptin signaling.[115] These findings indicate that PI3K–Akt signaling mediates leptin actions on glucose homeostasis.

29.3.3.0.3: GLUCOSE

Glucose sensing in the hypothalamus is important in glucose homeostasis. Injection of 2-deoxy-D-glucose into the VMH increases plasma glucose levels by elevating plasma glucagon and catecholamine levels.[116] Conversely, intra-VMH glucose infusion suppresses counter-regulatory hormonal responses to hypoglycemia.[90] The brain stem is also involved in glucoprivic feeding and counter-regulatory hormone secretion during hypoglycemia. Injection of another glucose anti-metabolite, 5-thio-D-glucose, into the NTS and the basolateral medulla, which contain A1/C1 catecholaminergic neurons projecting to the hypothalamic PVN and ARC, induces feeding and glucose responses, as seen in hypoglycemia.[117] Similarly, destruction of hindbrain catecholaminergic neurons by immunotoxins blocks 2-deoxy-D-glucose-induced feeding and blood glucose responses.[118]

The glucose-sensing mechanisms in hypothalamic neurons are similar to those in pancreatic β-cells.[119] Glucose signaling in glucose-excited neurons requires glucose uptake via the type 2 glucose transporter, which is followed by glucose phosphorylation by glucokinase, intramitochondrial glucose oxidation, and an increased cellular ATP/ADP ratio. This leads to the closure of ATP-sensitive K_{ATP} channels, depolarization of the membrane potential, and influx of Ca^{2+} through voltage-dependent calcium channels, which stimulate neuronal activity and neurotransmitter release.[120] The role of hypothalamic type 2 glucose transporter, glucokinase and K_{ATP} channels in sensing hypoglycemia and counter-regulatory hormone responses has been demonstrated.[121, 122, 123, 124, 125, 126] How glucose inhibits neuronal activity in glucose-inhibited neurons is unclear. One possibility is that glucose increases the ATP/ADP ratio, which stimulates the Na^+/K^+-ATPase pump and triggers hyperpolarizing currents.[127] Alternatively, glucose-induced activation of ATP-dependent Cl^- channels may induce hyperpolarization of the plasma membrane.[91, 128]

AMPK functions as a ‘fuel gauge’ that monitors cellular energy status and provokes adaptive responses to maintain cellular energy levels[129, 130]. ICV administration of glucose suppresses feeding via inhibition of hypothalamic AMPK activity.[131, 132] Hypothalamic AMPK activation is critical for feeding and counter-regulatory responses to hypoglycemia.[131] Intra-VMH administration of AICAR (5-aminoimidazole-4-carboxamide ribonucleotide), a chemical AMPK activator, increases HGP without changing the plasma levels of counter-regulatory hormones.[133] AMPK activation in the VMH restores reduced counter-regulatory responses induced by repeated hypoglycemia.[134] Consistent with these findings, inhibition of hypothalamic AMPK activity attenuates the counter-regulatory response during hypoglycemia.[135]

29.3.3.0.4: FATTY ACIDS

LCFA signals nutrient availability to the brain and modulates peripheral glucose metabolism.[70] ICV administration of oleic acid suppresses HGP during basal insulin clamping. ICV administration of K_{ATP} channel blocker attenuates the inhibitory effect of oleic acid on glucose production, indicating an involvement of brain K_{ATP} channels in this process.[70] Increased LCFA-CoA levels in hypothalamic neurons suppresses endogenous glucose production.[74] Pharmacological inhibition of hypothalamic esterification of fatty acids or surgical resection of the hepatic branch of the vagus nerve increases HGP.[136] Therefore, hypothalamic lipid sensing regulates glucose homeostasis via a mechanism involving the esterification of LCFAs to LCFA-CoAs, intact K_{ATP} channels and vagal outflow to the liver.

29.3.3.1: EFFECTOR PATHWAYS IN THE BRAIN CONTROL OF GLUCOSE METABOLISM

29.3.3.0.1: TO THE LIVER

In rodents, direct action of insulin on the liver is necessary, but is insufficient to inhibit HGP, unless the indirect brain pathway is not fully functional. Restoration of insulin receptor expression in either the liver or brain of insulin receptor-null mice does not completely restore the ability of insulin to inhibit HGP.[137] In contrast, restoration of insulin receptor expression in both the brain and liver normalizes insulin actions on HGP.[138] Whether neuronal control of HGP is unique to rodents remains uncertain. ICV insulin infusion in the dog augments hepatic glucose uptake and glycogen synthesis without altering HGP,[139] indicating that the regulation of gluconeogenesis by brain insulin signaling may differ among species. The basal HGP rate per weight is almost 5–10 times higher in rodents than in dogs and humans.[140] Dogs and humans maintain hepatic glycogen storage even after a 42-h fasting.[141, 142] In contrast, hepatic glycogen content is significantly depleted in rodents after a relatively short fast, which may be due to higher metabolic rates.[143] Therefore, the contribution of the gluconeogenic pathway to HGP may be greater in rodents than in animals with a larger body size.[139] Thus, changes in gluconeogenesis may be more easily detected in rodents.

29.3.3.0.2: TO THE SKELETAL MUSCLE

Electrical stimulation of VMH neurons and local injection of leptin into the VMH increases glucose uptake in the skeletal muscle of rats independently of circulating insulin levels.[144, 145] These effects appears to be mediated by the sympathetic nervous system as they are abolished by blockade of the sympathetic nervous system.[146, 147] Consistently, central infusion of leptin improves glucose tolerance and enhances insulin-stimulated Akt phosphorylation in skeletal muscle.[148, 149] Activated Akt leads to translocation of the glucose transporter GLUT4 from its sequestered cytoplasmic location to the cell membrane, facilitating glucose uptake[150].

In the skeletal muscle, AMPK activation is induced by muscle contraction and adrenergic agonist and mediates insulin-independent glucose uptake.[151] Leptin activates skeletal muscle AMPK through the hypothalamus and sympathetic nervous system.[152] Therefore, leptin may promote glucose uptake to the skeletal muscle through the sympathetic nervous system–muscle AMPK signaling pathway. On the other hand, orexin-producing neurons in the LH are activated by sweet foods. Orexin regulates skeletal muscle glucose uptake through VMH neurons expressing orexin receptors and the sympathetic nervous system.[153]

29.3.3.0.3: TO THE PANCREAS

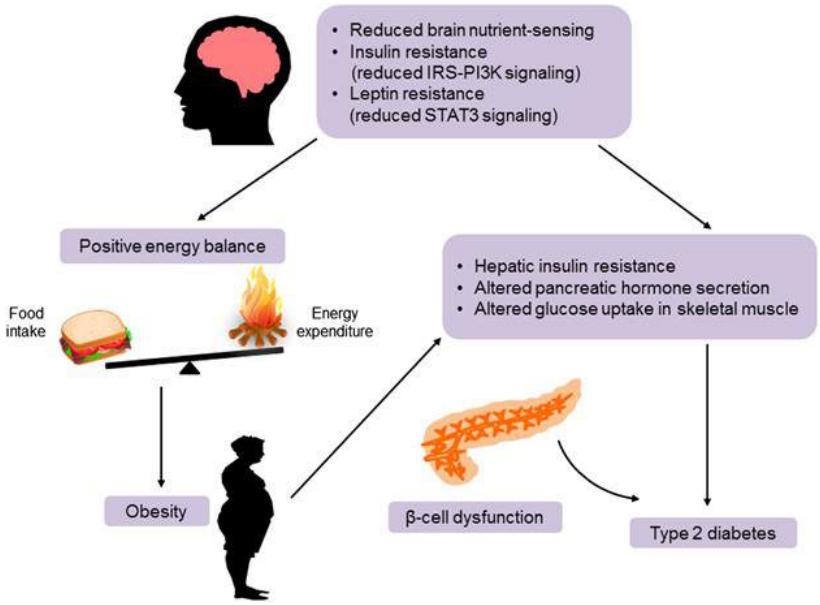
The autonomic nervous system controls the secretion of insulin and glucagon in the pancreas. Sympathetic and parasympathetic nerve endings are foundin pancreatic islets.[154] Moreover, α- and β-cells express neurotransmitter receptors.[155] Both sympathetic and parasympathetic nerve branches can stimulate glucagon secretion. In contrast, parasympathetic branches stimulate insulin secretion, whereas sympathetic branches inhibit it.[156] Neurons in the dorsal motor nucleus of the vagus project nerve terminals to the pancreatic ganglions via the vagus nerve, and thus vagus nerves connect the dorsal motor nucleus of the vagus and endocrine pancreas.[157]

Insulin regulates whole-body glucose metabolism by acting on the brain, and modulating insulin and glucagon secretion. ICV administration of insulin increases pancreatic insulin output, demonstrating that pancreatic β-cells are influenced by insulin-sensitive cells of the brain.[158] Moreover, insulin injection into the VMH inhibits glucagon secretion by pancreatic α-cells, indicating that insulin controls glucagon secretion via brain-mediated mechanisms.[159] Taken together, the brain, especially the hypothalamus and brain stem, modulates pancreatic insulin and glucagon secretion via the parasympathetic and sympathetic efferent nerves that innervate pancreatic α- and β-cells.[160]

29.3.4: DYSREGULATION OF ENERGY/GLUCOSE METABOLISM IN OBESITY AND DIABETES

In healthy conditions, energy intake matches energy expenditure to maintain normal body weight. Impaired ability of the brain to maintain energy homeostasis may underlie pathological weight gain and obesity (Figure 3). Several defects in the negative-feedback pathway in energy homeostatic mechanisms have been suggested.[3] Defects in the secretion of key metabolic hormones such as insulin and leptin may predispose weight gain. Because leptin primarily acts on hypothalamic neurons to regulate the energy balance, leptin transfer to the brain may be critical for its action.[5] Leptin concentrations in the plasma increase in proportion to body mass index, an indicator of fat mass. However, the increase in leptin concentrations in the cerebrospinal fluid of obese individuals is less than that of plasma leptin concentrations.[161] Therefore, reduced leptin transport to the brain may be due to reduced action of leptin in obesity.

Figure 3



Pathogenesis of obesity and type 2 diabetes due to defective central regulation of energy and glucose homeostasis. Reduced nutrient sensing and impaired insulin and leptin signaling in the hypothalamus may result in a positive energy balance and predispose weight gain, causing insulin resistance in peripheral metabolic organs. Obesity-associated insulin resistance may lead to type 2 diabetes when it is combined with β-cell dysfunction. IRS, insulin receptor substrate; PI3K, phosphatidylinositol 3 kinase; STAT3, signal transducer and activator of transcription 3.

Full size image

Defective hypothalamic sensing of these hormones favors a positive energy balance because loss of leptin receptors in the hypothalamus leads to obesity in mice.[13] Rats with diet-induced obesity have reduced expressions of leptin receptors in the hypothalamus.[162] Impaired postreceptor signaling in hypothalamic neurons may also result in pathological weight gain. Disruption of the hypothalamic IRS–PI3K signaling pathway causes resistance to peripheral metabolic signals and leads to obesity.[163] Likewise, mice with disrupted neuronal STAT3 signaling develop hyperphagia and obesity.[164]

In rodents, long-term high-fat feeding reduces the anorectic response and hypothalamic STAT3 activation induced by leptin, which is called leptin resistance.[165] Increased hypothalamic expression of suppressor of cytokine signaling 3 has been suggested to be a mechanism of hypothalamic leptin resistance. Ablation of suppressor of cytokine signaling 3 expression in neurons mitigates high-fat diet-induced weight gain and hyperleptinemia and improves glucose tolerance and insulin sensitivity.[166] Protein-tyrosine phosphatase 1B, a well-known negative regulator of insulin and leptin signaling,[167, 168] has also been suggested to cause leptin and insulin resistance in hypothalamic neurons. Neuronal Protein-tyrosine phosphatase 1B knockout mice are hypersensitive to exogenous leptin and insulin, and display improved glucose tolerance during chronic high-fat feeding.[169] Increased IKKβ-NFκB and endoplasmic stress have been found in the hypothalamus of obese rodents and shown to disrupt hypothalamic leptin and insulin signaling.[170, 171] However, a recent paper has shown, using a leptin receptor antagonist, that endogenous leptin signaling and actions in high-fat diet-fed obese mice treated are comparable to those of chow diet-fed mice, arguing against the concept of leptin resistance.[172] Thus, further studies are needed to clarify the issue of leptin resistance in obese humans and animals.

Diabetes mellitus is a metabolic disorder characterized by hyperglycemia that affects ~9% of adults worldwide.[173] It results from deficits in pancreatic insulin secretion and insulin signaling/actions in insulin target organs. Experimental evidence suggests that defective metabolic sensing in hypothalamic neurons may lead to dysregulation of glucose homeostasis and diabetes (Figure 3). Hypothalamic insulin–PI3K signaling is markedly impaired in rats with streptozotocin-induced diabetes.[98] Pharmacological inhibition of hypothalamic PI3K signaling attenuates the glucose-lowering effect of insulin. Conversely, enhanced hypothalamic PI3K signaling via adenoviral gene therapy potentiates insulin-induced glucose lowering.[98] Notably, central insulin actions are blunted by short-term high-fat diet feeding.[174] Thus, a fat-rich diet may contribute to the development of diabetes by disrupting insulin signaling in the hypothalamus.[174]

29.3.5: CONCLUDING REMARKS

This review highlights the role of the brain in the homeostatic regulation of energy and glucose metabolism. The brain detects energy intake by sensing gut hormones released in response to food intake and detecting nutrients in circulating blood. The brain also monitors body energy stores by sensing adiposity-related signals. Information on nutrient availability and stored fat is transferred to specialized neurons in the hypothalamus and brain stem. In the control of the energy balance, outflow pathways from the brain regulate food intake and energy expenditure (thermogenesis or locomotor activity).

The brain also has an important role in the maintenance of glucose homeostasis, which is achieved by the modulation of insulin/glucagon secretion in the endocrine pancreas, HGP, and skeletal muscle glucose uptake. The autonomic nervous system constitutes the outflow pathways from the brain to peripheral metabolic organs. Defective crosstalk between the brain and peripheral metabolic organs observed in the obese condition may lead to type 2 diabetes development and obesity progression. Therefore, better understanding of neural mechanisms involved in the regulation of glucose/energy homeostasis will provide us with the opportunity to develop new therapeutics combating obesity and diabetes.

29.3.6: REFERENCES

1. 1

   Morton GJ, Meek TH, Schwartz MW . Neurobiology of food intake in health and disease. *Nat Rev Neurosci* 2014; **15**: 367–378.

   CAS PubMed PubMed Central Article Google Scholar

2. 2

   Sandoval D, Cota D, Seeley RJ . The integrative role of CNS fuel-sensing mechanisms in energy balance and glucose regulation. *Annu Rev Physiol* 2008; **70**: 513–535.

   CAS PubMed Article Google Scholar

3. 3

   Schwartz MW, Porte D Jr . Diabetes, obesity, and the brain. *Science* 2005; **307**: 375–379.

   CAS PubMed Article Google Scholar

4. 4

   Broadwell RD, Brightman MW . Entry of peroxidase into neurons of the central and peripheral nervous systems from extracerebral and cerebral blood. *J Comp Neurol* 1976; **166**: 257–283.

   CAS PubMed Article Google Scholar

5. 5

   Schwartz MW, Woods SC, Porte D Jr, Seeley RJ, Baskin DG . Central nervous system control of food intake. *Nature* 2000; **404**: 661–671.

   CAS PubMed PubMed Central Article Google Scholar

6. 6

   Heisler LK, Cowley MA, Tecott LH, Fan W, Low MJ, Smart JL *et al*. Activation of central melanocortin pathways by fenfluramine. *Science* 2002; **297**: 609–611.

   CAS PubMed Article Google Scholar

7. 7

   Bouret SG, Draper SJ, Simerly RB . Formation of projection pathways from the arcuate nucleus of the hypothalamus to hypothalamic regions implicated in the neural control of feeding behavior in mice. *J Neurosci* 2004; **24**: 2797–2805.

   CAS PubMed Article Google Scholar

8. 8

   Huszar D, Lynch CA, Fairchild-Huntress V, Dunmore JH, Fang Q, Berkemeier LR *et al*. Targeted disruption of the melanocortin-4 receptor results in obesity in mice. *Cell* 1997; **88**: 131–141.

   CAS PubMed Article Google Scholar

9. 9

   Tao YX . Molecular mechanisms of the neural melanocortin receptor dysfunction in severe early onset obesity. *Mol Cell Endocrinol* 2005; **239**: 1–14.

   CAS PubMed Article Google Scholar

10. 10

    Ollmann MM, Wilson BD, Yang YK, Kerns JA, Chen Y, Gantz I *et al*. Antagonism of central melanocortin receptors *in vitro* and *in vivo* by agouti-related protein. *Science* 1997; **278**: 135–138.

    CAS PubMed Article Google Scholar

11. 11

    Bewick GA, Gardiner JV, Dhillo WS, Kent AS, White NE, Webster Z *et al*. Post-embryonic ablation of AgRP neurons in mice leads to a lean, hypophagic phenotype. *FASEB J* 2005; **19**: 1680–1682.

CAS PubMed Article Google Scholar

12. 12

Yulyaningsih E, Zhang L, Herzog H, Sainsbury A . NPY receptors as potential targets for anti-obesity drug development. *Br J Pharmacol* 2011; **163**: 1170–1202.

CAS PubMed PubMed Central Article Google Scholar

13. 13

Bingham NC, Anderson KK, Reuter AL, Stallings NR, Parker KL . Selective loss of leptin receptors in the ventromedial hypothalamic nucleus results in increased adiposity and a metabolic syndrome. *Endocrinology* 2008; **149**: 2138–2148.

CAS PubMed PubMed Central Article Google Scholar

14. 14

Foster MT, Song CK, Bartness TJ . Hypothalamic paraventricular nucleus lesion involvement in the sympathetic control of lipid mobilization. *Obesity (Silver Spring)* 2010; **18**: 682–689.

Article Google Scholar

15. 15

Leibowitz SF, Hammer NJ, Chang K . Hypothalamic paraventricular nucleus lesions produce overeating and obesity in the rat. *Physiol Behav* 1981; **27**: 1031–1040.

CAS PubMed Article Google Scholar

16. 16

Gonzalez JA, Reimann F, Burdakov D . Dissociation between sensing and metabolism of glucose in sugar sensing neurones. *J Physiol* 2009; **587**: 41–48.

CAS PubMed Article Google Scholar

17. 17

Fei H, Okano HJ, Li C, Lee GH, Zhao C, Darnell R *et al*. Anatomic localization of alternatively spliced leptin receptors (Ob-R) in mouse brain and other tissues. *Proc Natl Acad Sci USA* 1997; **94**: 7001–7005.

CAS PubMed Article Google Scholar

18. 18

Xu B, Goulding EH, Zang K, Cepoi D, Cone RD, Jones KR *et al*. Brain-derived neurotrophic factor regulates energy balance downstream of melanocortin-4 receptor. *Nat Neurosci* 2003; **6**: 736–742.

CAS PubMed PubMed Central Article Google Scholar

19. 19

Shimizu N, Oomura Y, Plata-Salaman CR, Morimoto M . Hyperphagia and obesity in rats with bilateral ibotenic acid-induced lesions of the ventromedial hypothalamic nucleus. *Brain Res* 1987; **416**: 153–156.

CAS PubMed Article Google Scholar

20. 20

Jacobowitz DM, O'Donohue TL . alpha-Melanocyte stimulating hormone: immunohistochemical identification and mapping in neurons of rat brain. *Proc Natl Acad Sci USA* 1978; **75**: 6300–6304.

CAS PubMed Article Google Scholar

21. 21

Bernardis LL, Bellinger LL . The dorsomedial hypothalamic nucleus revisited: 1986 update. *Brain Res* 1987; **434**: 321–381.

CAS PubMed Article Google Scholar

22. 22

Broberger C, De Lecea L, Sutcliffe JG, Hokfelt T . Hypocretin/orexin- and melanin-concentrating hormone-expressing cells form distinct populations in the rodent lateral hypothalamus: relationship to the neuropeptide Y and agouti gene-related protein systems. *J Comp Neurol* 1998; **402**: 460–474.

CAS PubMed Article Google Scholar

23. 23

Ohno K, Sakurai T . Orexin neuronal circuitry: role in the regulation of sleep and wakefulness. *Front Neuroendocrinol* 2008; **29**: 70–87.

CAS PubMed Article Google Scholar

24. 24

Hungs M, Mignot E . Hypocretin/orexin, sleep and narcolepsy. *Bioessays* 2001; **23**: 397–408.

CAS PubMed Article Google Scholar

25. 25

Marsh DJ, Weingarth DT, Novi DE, Chen HY, Trumbauer ME, Chen AS *et al*. Melanin-concentrating hormone 1 receptor-deficient mice are lean, hyperactive, and hyperphagic and have altered metabolism. *Proc Natl Acad Sci USA* 2002; **99**: 3240–3245.

CAS PubMed Article Google Scholar

26. 26

Schwartz GJ . The role of gastrointestinal vagal afferents in the control of food intake: current prospects. *Nutrition* 2000; **16**: 866–873.

CAS PubMed Article Google Scholar

27. 27

Stanley S, Wynne K, McGowan B, Bloom S . Hormonal regulation of food intake. *Physiol Rev* 2005; **85**: 1131–1158.

CAS PubMed Article Google Scholar

28. 28

Geerling JC, Shin JW, Chimenti PC, Loewy AD . Paraventricular hypothalamic nucleus: axonal projections to the brain stem. *J Comp Neurol* 2010; **518**: 1460–1499.

PubMed PubMed Central Article Google Scholar

29. 29

Ahima RS, Antwi DA . Brain regulation of appetite and satiety. *Endocrinol Metab Clin North Am* 2008; **37**: 811–823.

CAS PubMed PubMed Central Article Google Scholar

30. 30

Ellacott KL, Halatchev IG, Cone RD . Characterization of leptin-responsive neurons in the caudal brain stem. *Endocrinology* 2006; **147**: 3190–3195.

CAS PubMed Article Google Scholar

31. 31

Hommel JD, Trinko R, Sears RM, Georgescu D, Liu ZW, Gao XB *et al*. Leptin receptor signaling in midbrain dopamine neurons regulates feeding. *Neuron* 2006; **51**: 801–810.

CAS PubMed Article Google Scholar

32. 32

Kim KS, Yoon YR, Lee HJ, Yoon S, Kim SY, Shin SW *et al*. Enhanced hypothalamic leptin signaling in mice lacking dopamine D2 receptors. *J Biol Chem* 2010; **285**: 8905–8917.

CAS PubMed PubMed Central Article Google Scholar

33. 33

Spiegelman BM, Flier JS . Obesity and the regulation of energy balance. *Cell* 2001; **104**: 531–543.

CAS PubMed PubMed Central Article Google Scholar

34. 34

Kramer A, Yang FC, Snodgrass P, Li X, Scammell TE, Davis FC *et al*. Regulation of daily locomotor activity and sleep by hypothalamic EGF receptor signaling. *Science* 2001; **294**: 2511–2515.

CAS PubMed PubMed Central Article Google Scholar

35. 35

Samson WK, Bagley SL, Ferguson AV, White MM . Orexin receptor subtype activation and locomotor behaviour in the rat. *Acta Physiol (Oxf)* 2010; **198**: 313–324.

CAS Article Google Scholar

36. 36

Nakamachi T, Matsuda K, Maruyama K, Miura T, Uchiyama M, Funahashi H *et al*. Regulation by orexin of feeding behaviour and locomotor activity in the goldfish. *J Neuroendocrinol* 2006; **18**: 290–297.

CAS PubMed Article Google Scholar

37. 37

Huo L, Gamber K, Greeley S, Silva J, Huntoon N, Leng XH *et al*. Leptin-dependent control of glucose balance and locomotor activity by POMC neurons. *Cell Metab* 2009; **9**: 537–547.

CAS PubMed PubMed Central Article Google Scholar

38. 38

Minokoshi Y, Kim YB, Peroni OD, Fryer LG, Muller C, Carling D *et al*. Leptin stimulates fatty-acid oxidation by activating AMP-activated protein kinase. *Nature* 2002; **415**: 339–343.

CAS PubMed Article Google Scholar

39. 39

Seale P, Conroe HM, Estall J, Kajimura S, Frontini A, Ishibashi J *et al*. Prdm16 determines the thermogenic program of subcutaneous white adipose tissue in mice. *J Clin Invest* 2011; **121**: 96–105.

CAS PubMed Article Google Scholar

40. 40

Dodd GT, Decherf S, Loh K, Simonds SE, Wiede F, Balland E *et al*. Leptin and insulin act on POMC neurons to promote the browning of white fat. *Cell* 2015; **160**: 88–104.

CAS PubMed PubMed Central Article Google Scholar

41. 41

Morrison SF, Madden CJ, Tupone D . Central neural regulation of brown adipose tissue thermogenesis and energy expenditure. *Cell Metab* 2014; **19**: 741–756.

CAS PubMed PubMed Central Article Google Scholar

42. 42

Sacks H, Symonds ME . Anatomical locations of human brown adipose tissue: functional relevance and implications in obesity and type 2 diabetes. *Diabetes* 2013; **62**: 1783–1790.

CAS PubMed PubMed Central Article Google Scholar

43. 43

Ouellet V, Routhier-Labadie A, Bellemare W, Lakhal-Chaieb L, Turcotte E, Carpentier AC *et al*. Outdoor temperature, age, sex, body mass index, and diabetic status determine the prevalence, mass, and glucose-uptake activity of 18F-FDG-detected BAT in humans. *J Clin Endocrinol Metab* 2011; **96**: 192–199.

CAS PubMed Article Google Scholar

44. 44

Seoane-Collazo P, Ferno J, Gonzalez F, Dieguez C, Leis R, Nogueiras R *et al*. Hypothalamic-autonomic control of energy homeostasis. *Endocrine* 2015; **50**: 276–291.

CAS PubMed Article Google Scholar

45. 45

Imai-Matsumura K, Matsumura K, Nakayama T . Involvement of ventromedial hypothalamus in brown adipose tissue thermogenesis induced by preoptic cooling in rats. *Jpn J Physiol* 1984; **34**: 939–943.

CAS PubMed Article Google Scholar

46. 46

Yoshida K, Li X, Cano G, Lazarus M, Saper CB . Parallel preoptic pathways for thermoregulation. *J Neurosci* 2009; **29**: 11954–11964.

CAS PubMed PubMed Central Article Google Scholar

47. 47

Zhang Y, Kerman IA, Laque A, Nguyen P, Faouzi M, Louis GW *et al*. Leptin-receptor-expressing neurons in the dorsomedial hypothalamus and median preoptic area regulate sympathetic brown adipose tissue circuits. *J Neurosci* 2011; **31**: 1873–1884.

CAS PubMed PubMed Central Article Google Scholar

48. 48

Chao PT, Yang L, Aja S, Moran TH, Bi S . Knockdown of NPY expression in the dorsomedial hypothalamus promotes development of brown adipocytes and prevents diet-induced obesity. *Cell Metab* 2011; **13**: 573–583.

CAS PubMed PubMed Central Article Google Scholar

49. 49

Brito MN, Brito NA, Baro DJ, Song CK, Bartness TJ . Differential activation of the sympathetic innervation of adipose tissues by melanocortin receptor stimulation. *Endocrinology* 2007; **148**: 5339–5347.

CAS PubMed Article Google Scholar

50. 50

Haynes WG, Morgan DA, Djalali A, Sivitz WI, Mark AL . Interactions between the melanocortin system and leptin in control of sympathetic nerve traffic. *Hypertension* 1999; **33**: 542–547.

CAS PubMed Article Google Scholar

51. 51

Lockie SH, Heppner KM, Chaudhary N, Chabenne JR, Morgan DA, Veyrat-Durebex C *et al*. Direct control of brown adipose tissue thermogenesis by central nervous system glucagon-like peptide-1 receptor signaling. *Diabetes* 2012; **61**: 2753–2762.

CAS PubMed PubMed Central Article Google Scholar

52. 52

Rahmouni K, Morgan DA, Morgan GM, Liu X, Sigmund CD, Mark AL *et al*. Hypothalamic PI3K and MAPK differentially mediate regional sympathetic activation to insulin. *J Clin Invest* 2004; **114**: 652–658.

CAS PubMed PubMed Central Article Google Scholar

53. 53

Rothwell NJ, Stock MJ . A role for insulin in the diet-induced thermogenesis of cafeteria-fed rats. *Metabolism* 1981; **30**: 673–678.

CAS PubMed Article Google Scholar

54. 54

Schwartz RS, Jaeger LF, Veith RC . Effect of clonidine on the thermic effect of feeding in humans. *Am J Physiol* 1988; **254**: R90–R94.

CAS PubMed Google Scholar

55. 55

Cannon B, Nedergaard J . Brown adipose tissue: function and physiological significance. *Physiol Rev* 2004; **84**: 277–359.

CAS PubMed Article Google Scholar

56. 56

Seeley RJ, Woods SC . Monitoring of stored and available fuel by the CNS: implications for obesity. *Nat Rev Neurosci* 2003; **4**: 901–909.

CAS PubMed Article Google Scholar

57. 57

Woods SC, Lotter EC, McKay LD, Porte D Jr . Chronic intracerebroventricular infusion of insulin reduces food intake and body weight of baboons. *Nature* 1979; **282**: 503–505.

CAS PubMed Article Google Scholar

58. 58

Bagdade JD, Bierman EL, Porte D Jr . The significance of basal insulin levels in the evaluation of the insulin response to glucose in diabetic and nondiabetic subjects. *J Clin Invest* 1967; **46**: 1549–1557.

CAS PubMed PubMed Central Article Google Scholar

59. 59

Air EL, Benoit SC, Blake Smith KA, Clegg DJ, Woods SC . Acute third ventricular administration of insulin decreases food intake in two paradigms. *Pharmacol Biochem Behav* 2002; **72**: 423–429.

CAS PubMed Article Google Scholar

60. 60

Chavez M, Kaiyala K, Madden LJ, Schwartz MW, Woods SC . Intraventricular insulin and the level of maintained body weight in rats. *Behav Neurosci* 1995; **109**: 528–531.

CAS PubMed Article Google Scholar

61. 61

Bruning JC, Gautam D, Burks DJ, Gillette J, Schubert M, Orban PC *et al*. Role of brain insulin receptor in control of body weight and reproduction. *Science* 2000; **289**: 2122–2125.

CAS Article PubMed Google Scholar

62. 62

White MF . Insulin signaling in health and disease. *Science* 2003; **302**: 1710–1711.

CAS PubMed Article Google Scholar

63. 63

Maffei M, Stoffel M, Barone M, Moon B, Dammerman M, Ravussin E *et al*. Absence of mutations in the human OB gene in obese/diabetic subjects. *Diabetes* 1996; **45**: 679–682.

CAS PubMed Article Google Scholar

64. 64

Considine RV, Sinha MK, Heiman ML, Kriauciunas A, Stephens TW, Nyce MR *et al*. Serum immunoreactive-leptin concentrations in normal-weight and obese humans. *N Engl J Med* 1996; **334**: 292–295.

CAS Article PubMed Google Scholar

65. 65

Tartaglia LA, Dembski M, Weng X, Deng N, Culpepper J, Devos R *et al*. Identification and expression cloning of a leptin receptor, OB-R. *Cell* 1995; **83**: 1263–1271.

CAS Article PubMed Google Scholar

66. 66

Coppari R, Ichinose M, Lee CE, Pullen AE, Kenny CD, McGovern RA *et al*. The hypothalamic arcuate nucleus: a key site for mediating leptin's effects on glucose homeostasis and locomotor activity. *Cell Metab* 2005; **1**: 63–72.

CAS PubMed Article Google Scholar

67. 67

Satoh N, Ogawa Y, Katsuura G, Hayase M, Tsuji T, Imagawa K *et al*. The arcuate nucleus as a primary site of satiety effect of leptin in rats. *Neurosci Lett* 1997; **224**: 149–152.

CAS PubMed Article Google Scholar

68. 68

Oswal A, Yeo G . Leptin and the control of body weight: a review of its diverse central targets, signaling mechanisms, and role in the pathogenesis of obesity. *Obesity (Silver Spring)* 2010; **18**: 221–229.

Article Google Scholar

69. 69

Myers MG Jr, Olson DP . Central nervous system control of metabolism. *Nature* 2012; **491**: 357–363.

CAS PubMed Article Google Scholar

70. 70

Obici S, Feng Z, Morgan K, Stein D, Karkanias G, Rossetti L . Central administration of oleic acid inhibits glucose production and food intake. *Diabetes* 2002; **51**: 271–275.

CAS PubMed Article Google Scholar

71. 71

Miselis RR, Epstein AN . Feeding induced by intracerebroventricular 2-deoxy-D-glucose in the rat. *Am J Physiol* 1975; **229**: 1438–1447.

CAS PubMed Article Google Scholar

72. 72

Foster DW . Malonyl-CoA: the regulator of fatty acid synthesis and oxidation. *J Clin Invest* 2012; **122**: 1958–1959.

CAS PubMed PubMed Central Article Google Scholar

73. 73

Clegg DJ, Wortman MD, Benoit SC, McOsker CC, Seeley RJ . Comparison of central and peripheral administration of C75 on food intake, body weight, and conditioned taste aversion. *Diabetes* 2002; **51**: 3196–3201.

CAS PubMed Article Google Scholar

74. 74

Obici S, Feng Z, Arduini A, Conti R, Rossetti L . Inhibition of hypothalamic carnitine palmitoyltransferase-1 decreases food intake and glucose production. *Nat Med* 2003; **9**: 756–761.

CAS PubMed Article Google Scholar

75. 75

Kennedy GC . The role of depot fat in the hypothalamic control of food intake in the rat. *Proc R Soc Lond B Biol Sci* 1953; **140**: 578–596.

CAS PubMed Article Google Scholar

76. 76

Baggio LL, Drucker DJ . Glucagon-like peptide-1 receptors in the brain: controlling food intake and body weight. *J Clin Invest* 2014; **124**: 4223–4226.

CAS PubMed PubMed Central Article Google Scholar

77. 77

Merchenthaler I, Lane M, Shughrue P . Distribution of pre-pro-glucagon and glucagon-like peptide-1 receptor messenger RNAs in the rat central nervous system. *J Comp Neurol* 1999; **403**: 261–280.

CAS PubMed Article Google Scholar

78. 78

Chelikani PK, Haver AC, Reidelberger RD . Intravenous infusion of glucagon-like peptide-1 potently inhibits food intake, sham feeding, and gastric emptying in rats. *Am J Physiol Regul Integr Comp Physiol* 2005; **288**: R1695–R1706.

CAS PubMed Article Google Scholar

79. 79

Tang-Christensen M, Larsen PJ, Goke R, Fink-Jensen A, Jessop DS, Moller M *et al*. Central administration of GLP-1-(7-36) amide inhibits food and water intake in rats. *Am J Physiol* 1996; **271**: R848–R856.

CAS PubMed Google Scholar

80. 80

Cummings DE, Clement K, Purnell JQ, Vaisse C, Foster KE, Frayo RS *et al*. Elevated plasma ghrelin levels in Prader Willi syndrome. *Nat Med* 2002; **8**: 643–644.

CAS PubMed Article Google Scholar

81. 81

Febbraio MA, Pedersen BK . Muscle-derived interleukin-6: mechanisms for activation and possible biological roles. *FASEB J* 2002; **16**: 1335–1347.

CAS PubMed Article Google Scholar

82. 82

Wallenius V, Wallenius K, Ahren B, Rudling M, Carlsten H, Dickson SL *et al*. Interleukin-6-deficient mice develop mature-onset obesity. *Nat Med* 2002; **8**: 75–79.

CAS PubMed Article Google Scholar

83. 83

Lutz TA . Control of energy homeostasis by amylin. *Cell Mol Life Sci* 2012; **69**: 1947–1965.

CAS PubMed Article Google Scholar

84. 84

Campbell JE, Drucker DJ . Islet α cells and glucagon − critical regulators of energy homeostasis. *Nat Rev Endocrinol* 2015; **11**: 329–338.

CAS PubMed Article Google Scholar

85. 85

Asakawa A, Inui A, Yuzuriha H, Ueno N, Katsuura G, Fujimiya M *et al*. Characterization of the effects of pancreatic polypeptide in the regulation of energy balance. *Gastroenterology* 2003; **124**: 1325–1336.

CAS PubMed Article Google Scholar

86. 86

Woods SC, Lutz TA, Geary N, Langhans W . Pancreatic signals controlling food intake; insulin, glucagon and amylin. *Philos Trans R Soc Lond B Biol Sci* 2006; **361**: 1219–1235.

CAS PubMed PubMed Central Article Google Scholar

87. 87

Bernard C . *Leçons de physiologie expérimentale appliquée à la médecine*. Ballière et Fils: Paris, France, 1855.

Google Scholar

88. 88

Anand B, Chhina G, Sharma K, Dua S, Singh B . Activity of single neurons in the hypothalamic feeding centers: effect of glucose. *Am J Physiol* 1964; **207**: 1146–1154.

CAS PubMed Article Google Scholar

89. 89

Oomura Y, Ono T, Ooyama H, Wayner M . Glucose and osmosensitive neurones of the rat hypothalamus. *Nature* 1969; **222**: 282–284.

CAS PubMed Article Google Scholar

90. 90

Borg MA, Sherwin RS, Borg WP, Tamborlane WV, Shulman GI . Local ventromedial hypothalamus glucose perfusion blocks counterregulation during systemic hypoglycemia in awake rats. *J Clin Invest* 1997; **99**: 361–365.

CAS PubMed PubMed Central Article Google Scholar

91. 91

Routh VH . Glucose-sensing neurons: are they physiologically relevant? *Physiol Behav* 2002; **76**: 403–413.

CAS PubMed Article Google Scholar

92. 92

Dunn-Meynell AA, Rawson NE, Levin BE . Distribution and phenotype of neurons containing the ATP-sensitive K+ channel in rat brain. *Brain Res* 1998; **814**: 41–54.

CAS PubMed Article Google Scholar

93. 93

Mizuno Y, Oomura Y . Glucose responding neurons in the nucleus tractus solitarius of the rat: *in vitro* study. *Brain Res* 1984; **307**: 109–116.

CAS PubMed Article Google Scholar

94. 94

Funahashi M, Adachi A . Glucose-responsive neurons exist within the area postrema of the rat: *in vitro* study on the isolated slice preparation. *Brain Res Bull* 1993; **32**: 531–535.

CAS PubMed Article Google Scholar

95. 95

Yettefti K, Orsini J-C, Perrin J . Characteristics of glycemia-sensitive neurons in the nucleus tractus solitarii: possible involvement in nutritional regulation. *Physiol Behav* 1997; **61**: 93–100.

CAS PubMed Article Google Scholar

96. 96

Obici S, Feng Z, Karkanias G, Baskin DG, Rossetti L . Decreasing hypothalamic insulin receptors causes hyperphagia and insulin resistance in rats. *Nat Neurosci* 2002; **5**: 566–572.

CAS PubMed Article Google Scholar

97. 97

Obici S, Zhang BB, Karkanias G, Rossetti L . Hypothalamic insulin signaling is required for inhibition of glucose production. *Nat Med* 2002; **8**: 1376–1382.

CAS PubMed Article Google Scholar

98. 98

Gelling RW, Morton GJ, Morrison CD, Niswender KD, Myers MG, Rhodes CJ *et al*. Insulin action in the brain contributes to glucose lowering during insulin treatment of diabetes. *Cell Metab* 2006; **3**: 67–73.

CAS PubMed Article Google Scholar

99. 99

Spanswick D, Smith M, Mirshamsi S, Routh V, Ashford M . Insulin activates ATP-sensitive K^+ channels in hypothalamic neurons of lean, but not obese rats. *Nat Neurosci* 2000; **3**: 757–758.

CAS PubMed Article Google Scholar

100. 100

Pocai A, Lam TK, Gutierrez-Juarez R, Obici S, Schwartz GJ, Bryan J *et al*. Hypothalamic KATP channels control hepatic glucose production. *Nature* 2005; **434**: 1026–1031.

CAS PubMed Article Google Scholar

101. 101

Inoue H, Ogawa W, Asakawa A, Okamoto Y, Nishizawa A, Matsumoto M *et al*. Role of hepatic STAT3 in brain-insulin action on hepatic glucose production. *Cell Metab* 2006; **3**: 267–275.

CAS PubMed Article Google Scholar

102. 102

Coleman D . Obese and diabetes: two mutant genes causing diabetes-obesity syndromes in mice. *Diabetologia* 1978; **14**: 141–148.

CAS PubMed Article Google Scholar

103. 103

Halaas JL, Gajiwala KS, Maffei M, Cohen SL, Chait BT, Rabinowitz D *et al*. Weight-reducing effects of the plasma protein encoded by the obese gene. *Science* 1995; **269**: 543–546.

CAS PubMed Article Google Scholar

104. 104

Pelleymounter MA, Cullen MJ, Baker MB, Hecht R, Winters D, Boone T *et al*. Effects of the obese gene product on body weight regulation in ob/ob mice. *Science* 1995; **269**: 540–543.

CAS PubMed Article Google Scholar

105. 105

Rossetti L, Massillon D, Barzilai N, Vuguin P, Chen W, Hawkins M *et al*. Short term effects of leptin on hepatic gluconeogenesis and *in vivo* insulin action. *J Biol Chem* 1997; **272**: 27758–27763.

CAS PubMed Article Google Scholar

106. 106

Liu L, Karkanias GB, Jose'C M, Hawkins M, Barzilai N, Wang J *et al*. Intracerebroventricular leptin regulates hepatic but not peripheral glucose fluxes. *J Biol Chem* 1998; **273**: 31160–31167.

CAS PubMed Article Google Scholar

107. 107

Schwartz MW, Baskin DG, Bukowski TR, Kuijper JL, Foster D, Lasser G *et al*. Specificity of leptin action on elevated blood glucose levels and hypothalamic neuropeptide Y gene expression in ob/ob mice. *Diabetes* 1996; **45**: 531–535.

CAS PubMed Article Google Scholar

108. 108

Shimomura I, Hammer RE, Ikemoto S, Brown MS, Goldstein JL . Leptin reverses insulin resistance and diabetes mellitus in mice with congenital lipodystrophy. *Nature* 1999; **401**: 73–76.

CAS Article PubMed Google Scholar

109. 109

Ebihara K, Ogawa Y, Masuzaki H, Shintani M, Miyanaga F, Aizawa-Abe M *et al*. Transgenic overexpression of leptin rescues insulin resistance and diabetes in a mouse model of lipoatrophic diabetes. *Diabetes* 2001; **50**: 1440–1448.

CAS PubMed Article Google Scholar

110. 110

Asilmaz E, Cohen P, Miyazaki M, Dobrzyn P, Ueki K, Fayzikhodjaeva G *et al*. Site and mechanism of leptin action in a rodent form of congenital lipodystrophy. *J Clin Invest* 2004; **113**: 414–424.

CAS PubMed PubMed Central Article Google Scholar

111. 111

Pocai A, Morgan K, Buettner C, Gutierrez-Juarez R, Obici S, Rossetti L . Central leptin acutely reverses diet-induced hepatic insulin resistance. *Diabetes* 2005; **54**: 3182–3189.

CAS PubMed Article Google Scholar

112. 112

Coppari R, Ichinose M, Lee CE, Pullen AE, Kenny CD, McGovern RA *et al*. The hypothalamic arcuate nucleus: a key site for mediating leptin's effects on glucose homeostasis and locomotor activity. *Cell Metab* 2005; **1**: 63–72.

CAS PubMed Article Google Scholar

113. 113

Buettner C, Pocai A, Muse ED, Etgen AM, Myers MG, Rossetti L . Critical role of STAT3 in leptin's metabolic actions. *Cell Metab* 2006; **4**: 49–60.

CAS PubMed PubMed Central Article Google Scholar

114. 114

Kievit P, Howard JK, Badman MK, Balthasar N, Coppari R, Mori H *et al*. Enhanced leptin sensitivity and improved glucose homeostasis in mice lacking suppressor of cytokine signaling-3 in POMC-expressing cells. *Cell Metab* 2006; **4**: 123–132.

CAS PubMed Article Google Scholar

115. 115

Morton GJ, Gelling RW, Niswender KD, Morrison CD, Rhodes CJ, Schwartz MW . Leptin regulates insulin sensitivity via phosphatidylinositol-3-OH kinase signaling in mediobasal hypothalamic neurons. *Cell Metab* 2005; **2**: 411–420.

CAS PubMed Article Google Scholar

116. 116

Borg WP, Sherwin RS, During MJ, Borg MA, Shulman GI . Local ventromedial hypothalamus glucopenia triggers counterregulatory hormone release. *Diabetes* 1995; **44**: 180–184.

CAS PubMed Article Google Scholar

117. 117

Ritter S, Dinh TT, Zhang Y . Localization of hindbrain glucoreceptive sites controlling food intake and blood glucose. *Brain Res* 2000; **856**: 37–47.

CAS PubMed Article Google Scholar

118. 118

Ritter S, Bugarith K, Dinh TT . Immunotoxic destruction of distinct catecholamine subgroups produces selective impairment of glucoregulatory responses and neuronal activation. *J Comp Neurol* 2001; **432**: 197–216.

CAS PubMed Article Google Scholar

119. 119

Schuit FC, Huypens P, Heimberg H, Pipeleers DG . Glucose sensing in pancreatic β-cells a model for the study of other glucose-regulated cells in gut, pancreas, and hypothalamus. *Diabetes* 2001; **50**: 1–11.

CAS PubMed Article Google Scholar

120. 120

Marty N, Dallaporta M, Thorens B . Brain glucose sensing, counterregulation, and energy homeostasis. *Physiology* 2007; **22**: 241–251.

CAS PubMed Article Google Scholar

121. 121

Burcelin R, Thorens B . Evidence that extrapancreatic GLUT2-dependent glucose sensors control glucagon secretion. *Diabetes* 2001; **50**: 1282–1289.

CAS PubMed Article Google Scholar

122. 122

Marty N, Dallaporta M, Foretz M, Emery M, Tarussio D, Bady I *et al*. Regulation of glucagon secretion by glucose transporter type 2 (glut2) and astrocyte-dependent glucose sensors. *J Clin Invest* 2005; **115**: 3545–3553.

CAS PubMed PubMed Central Article Google Scholar

123. 123

Sanders NM, Dunn-Meynell AA, Levin BE . Third ventricular alloxan reversibly impairs glucose counterregulatory responses. *Diabetes* 2004; **53**: 1230–1236.

CAS PubMed Article Google Scholar

124. 124

Miki T, Liss B, Minami K, Shiuchi T, Saraya A, Kashima Y *et al*. ATP-sensitive K+ channels in the hypothalamus are essential for the maintenance of glucose homeostasis. *Nat Neurosci* 2001; **4**: 507–512.

CAS PubMed Article Google Scholar

125. 125

Evans ML, McCrimmon RJ, Flanagan DE, Keshavarz T, Fan X, McNay EC *et al*. Hypothalamic ATP-sensitive K+ channels play a key role in sensing hypoglycemia and triggering counterregulatory epinephrine and glucagon responses. *Diabetes* 2004; **53**: 2542–2551.

CAS PubMed Article Google Scholar

126. 126

McCrimmon RJ, Evans ML, Fan X, McNay EC, Chan O, Ding Y *et al*. Activation of ATP-sensitive K+ channels in the ventromedial hypothalamus amplifies counterregulatory hormone responses to hypoglycemia in normal and recurrently hypoglycemic rats. *Diabetes* 2005; **54**: 3169–3174.

CAS PubMed Article Google Scholar

127. 127

Oomura Y, Ooyama H, Sugimori M, Nakamura T, Yamada Y . Glucose inhibition of the glucose-sensitive neurone in the rat lateral hypothalamus. *Nature* 1974; **247**: 284–286.

CAS PubMed Article Google Scholar

128. 128

Song Z, Levin BE, McArdle JJ, Bakhos N, Routh VH . Convergence of pre-and postsynaptic influences on glucosensing neurons in the ventromedial hypothalamic nucleus. *Diabetes* 2001; **50**: 2673–2681.

CAS PubMed Article Google Scholar

129. 129

Hardie DG, Carling D, Carlson M . The AMP-activated/SNF1 protein kinase subfamily: metabolic sensors of the eukaryotic cell? *Annu Rev Biochem* 1998; **67**: 821–855.

CAS PubMed Article Google Scholar

130. 130

Rutter G, daSilva Xavier G, Leclerc I . Roles of 5′-AMP-activated protein kinase (AMPK) in mammalian glucose homoeostasis. *Biochem J* 2003; **375**: 1–16.

CAS PubMed PubMed Central Article Google Scholar

131. 131

Kim M-S, Park J-Y, Namkoong C, Jang P-G, Ryu J-W, Song H-S *et al*. Anti-obesity effects of α-lipoic acid mediated by suppression of hypothalamic AMP-activated protein kinase. *Nat Med* 2004; **10**: 727–733.

CAS Article PubMed Google Scholar

132. 132

Minokoshi Y, Alquier T, Furukawa N, Kim Y-B, Lee A, Xue B *et al*. AMP-kinase regulates food intake by responding to hormonal and nutrient signals in the hypothalamus. *Nature* 2004; **428**: 569–574.

CAS Article PubMed Google Scholar

133. 133

McCrimmon RJ, Fan X, Ding Y, Zhu W, Jacob RJ, Sherwin RS . Potential role for AMP-activated protein kinase in hypoglycemia sensing in the ventromedial hypothalamus. *Diabetes* 2004; **53**: 1953–1958.

CAS PubMed Article Google Scholar

134. 134

McCrimmon RJ, Fan X, Cheng H, McNay E, Chan O, Shaw M *et al*. Activation of AMP-activated protein kinase within the ventromedial hypothalamus amplifies counterregulatory hormone responses in rats with defective counterregulation. *Diabetes* 2006; **55**: 1755–1760.

CAS PubMed Article Google Scholar

135. 135

Han S-M, Namkoong C, Jang P, Park I, Hong S, Katakami H *et al*. Hypothalamic AMP-activated protein kinase mediates counter-regulatory responses to hypoglycaemia in rats. *Diabetologia* 2005; **48**: 2170–2178.

CAS PubMed Article Google Scholar

136. 136

Lam TK, Pocai A, Gutierrez-Juarez R, Obici S, Bryan J, Aguilar-Bryan L *et al*. Hypothalamic sensing of circulating fatty acids is required for glucose homeostasis. *Nat Med* 2005; **11**: 320–327.

CAS PubMed Article Google Scholar

137. 137

Okamoto H, Obici S, Accili D, Rossetti L . Restoration of liver insulin signaling in Insr knockout mice fails to normalize hepatic insulin action. *J Clin Invest* 2005; **115**: 1314–1322.

CAS PubMed PubMed Central Article Google Scholar

138. 138

Lin HV, Plum L, Ono H, Gutiérrez-Juárez R, Shanabrough M, Borok E *et al*. Divergent regulation of energy expenditure and hepatic glucose production by insulin receptor in agouti-related protein and POMC neurons. *Diabetes* 2010; **59**: 337–346.

CAS PubMed Article Google Scholar

139. 139

Ramnanan CJ, Saraswathi V, Smith MS, Donahue EP, Farmer B, Farmer TD *et al*. Brain insulin action augments hepatic glycogen synthesis without suppressing glucose production or gluconeogenesis in dogs. *J Clin Invest* 2011; **121**: 3713–3723.

CAS PubMed PubMed Central Article Google Scholar

140. 140

Cherrington A, Moore M, Sindelar D, Edgerton D . Insulin action on the liver *in vivo*. *Biochem Soc Trans* 2007; **35**: 1171–1174.

CAS PubMed Article Google Scholar

141. 141

Hendrick GK, Frizzell RT, Williams PE, Cherrington AD . Effect of hyperglucagonemia on hepatic glycogenolysis and gluconeogenesis after a prolonged fast. *Am J Physiol* 1990; **258**: E841–E849.

CAS PubMed Article Google Scholar

142. 142

Nuttall FQ, Ngo A, Gannon MC . Regulation of hepatic glucose production and the role of gluconeogenesis in humans: is the rate of gluconeogenesis constant? *Diabetes Metab Res Rev* 2008; **24**: 438–458.

CAS PubMed Article Google Scholar

143. 143

Kokubun E, Hirabara SM, Fiamoncini J, Curi R, Haebisch H . Changes of glycogen content in liver, skeletal muscle, and heart from fasted rats. *Cell Biochem Funct* 2009; **27**: 488–495.

CAS PubMed Article Google Scholar

144. 144

Shimazu T, Sudo M, Minokoshi Y, Takahashi A . Role of the hypothalamus in insulin-independent glucose uptake in peripheral tissues. *Brain Res Bull* 1991; **27**: 501–504.

CAS PubMed Article Google Scholar

145. 145

Sudo M, Minokoshi Y, Shimazu T . Ventromedial hypothalamic stimulation enhances peripheral glucose uptake in anesthetized rats. *Am J Physiol* 1991; **261**: E298–E303.

CAS PubMed Google Scholar

146. 146

Minokoshi Y, Okano Y, Shimazu T . Regulatory mechanism of the ventromedial hypothalamus in enhancing glucose uptake in skeletal muscles. *Brain Res* 1994; **649**: 343–347.

CAS PubMed Article Google Scholar

147. 147

Minokoshi Y, Haque MS, Shimazu T . Microinjection of leptin into the ventromedial hypothalamus increases glucose uptake in peripheral tissues in rats. *Diabetes* 1999; **48**: 287–291.

CAS PubMed Article Google Scholar

148. 148

Roman EA, Reis D, Romanatto T, Maimoni D, Ferreira EA, Santos GA *et al*. Central leptin action improves skeletal muscle AKT, AMPK, and PGC1α activation by hypothalamic PI3K-dependent mechanism. *Mol Cell Endocrinol* 2010; **314**: 62–69.

CAS PubMed Article Google Scholar

149. 149

Koch C, Augustine RA, Steger J, Ganjam GK, Benzler J, Pracht C *et al*. Leptin rapidly improves glucose homeostasis in obese mice by increasing hypothalamic insulin sensitivity. *J Neurosci* 2010; **30**: 16180–16187.

CAS PubMed PubMed Central Article Google Scholar

150. 150

Funai K, Cartee GD . Inhibition of contraction-stimulated AMP-activated protein kinase inhibits contraction-stimulated increases in PAS-TBC1D1 and glucose transport without altering PAS-AS160 in rat skeletal muscle. *Diabetes* 2009; **58**: 1096–1104.

CAS PubMed PubMed Central Article Google Scholar

151. 151

Hutchinson DS, Bengtsson T . AMP-activated protein kinase activation by adrenoceptors in L6 skeletal muscle cells mediation by α1-adrenoceptors causing glucose uptake. *Diabetes* 2006; **55**: 682–690.

CAS PubMed Article Google Scholar

152. 152

Minokoshi Y, Kim Y-B, Peroni OD, Fryer LG, Müller C, Carling D *et al*. Leptin stimulates fatty-acid oxidation by activating AMP-activated protein kinase. *Nature* 2002; **415**: 339–343.

CAS PubMed Article Google Scholar

153. 153

Shiuchi T, Haque MS, Okamoto S, Inoue T, Kageyama H, Lee S *et al*. Hypothalamic orexin stimulates feeding-associated glucose utilization in skeletal muscle via sympathetic nervous system. *Cell Metab* 2009; **10**: 466–480.

CAS PubMed Article Google Scholar

154. 154

Ahrén B . Autonomic regulation of islet hormone secretion–implications for health and disease. *Diabetologia* 2000; **43**: 393–410.

PubMed Article Google Scholar

155. 155

Satin LS, Kinard TA . Neurotransmitters and their receptors in the islets of Langerhans of the pancreas. *Endocrine* 1998; **8**: 213–223.

CAS PubMed Article Google Scholar

156. 156

Thorens B . Central control of glucose homeostasis: the brain–endocrine pancreas axis. *Diabetes Metab* 2010; **36**: S45–S49.

CAS PubMed Article Google Scholar

157. 157

Ionescu E, Rohner-Jeanrenaud F, Berthoud H-R, Jeanrenaud B . Increases in plasma insulin levels in response to electrical stimulation of the dorsal motor nucleus of the vagus nerve. *Endocrinology* 1983; **112**: 904–910.

CAS PubMed Article Google Scholar

158. 158

Chen M, Woods SC, Porte D . Effect of cerebral intraventricular insulin on pancreatic insulin secretion in the dog. *Diabetes* 1975; **24**: 910–914.

CAS PubMed Article Google Scholar

159. 159

Paranjape SA, Chan O, Zhu W, Horblitt AM, McNay EC, Cresswell JA *et al*. Influence of insulin in the ventromedial hypothalamus on pancreatic glucagon secretion *in vivo*. *Diabetes* 2010; **59**: 1521–1527.

CAS PubMed PubMed Central Article Google Scholar

160. 160

Thorens B . Brain glucose sensing and neural regulation of insulin and glucagon secretion. *Diabetes Obes Metab* 2011; **13**: 82–88.

CAS PubMed Article Google Scholar

161. 161

Caro JF, Kolaczynski JW, Nyce MR, Ohannesian JP, Opentanova I, Goldman WH *et al*. Decreased cerebrospinal-fluid/serum leptin ratio in obesity: a possible mechanism for leptin resistance. *Lancet* 1996; **348**: 159–161.

CAS PubMed Article Google Scholar

162. 162

Maes HH, Neale MC, Eaves LJ . Genetic and environmental factors in relative body weight and human adiposity. *Behav Genet* 1997; **27**: 325–351.

CAS PubMed Article Google Scholar

163. 163

Kubota N, Terauchi Y, Tobe K, Yano W, Suzuki R, Ueki K *et al*. Insulin receptor substrate 2 plays a crucial role in beta cells and the hypothalamus. *J Clin Invest* 2004; **114**: 917–927.

CAS PubMed PubMed Central Article Google Scholar

164. 164

Gao Q, Wolfgang MJ, Neschen S, Morino K, Horvath TL, Shulman GI *et al*. Disruption of neural signal transducer and activator of transcription 3 causes obesity, diabetes, infertility, and thermal dysregulation. *Proc Natl Acad Sci USA* 2004; **101**: 4661–4666.

CAS PubMed Article Google Scholar

165. 165

El-Haschimi K, Pierroz DD, Hileman SM, Bjørbæk C, Flier JS . Two defects contribute to hypothalamic leptin resistance in mice with diet-induced obesity. *J Clin Invest* 2000; **105**: 1827–1832.

CAS PubMed PubMed Central Article Google Scholar

166. 166

Mori H, Hanada R, Hanada T, Aki D, Mashima R, Nishinakamura H *et al*. Socs3 deficiency in the brain elevates leptin sensitivity and confers resistance to diet-induced obesity. *Nat Med* 2004; **10**: 739–743.

CAS PubMed Article Google Scholar

167. 167

Egawa K, Maegawa H, Shimizu S, Morino K, Nishio Y, Bryer-Ash M *et al*. Protein-tyrosine phosphatase-1B negatively regulates insulin signaling in l6 myocytes and Fao hepatoma cells. *J Biol Chem* 2001; **276**: 10207–10211.

CAS PubMed Article Google Scholar

168. 168

Kaszubska W, Falls HD, Schaefer VG, Haasch D, Frost L, Hessler P *et al*. Protein tyrosine phosphatase 1B negatively regulates leptin signaling in a hypothalamic cell line. *Mol Cell Endocrinol* 2002; **195**: 109–118.

CAS PubMed Article Google Scholar

169. 169

Bence KK, Delibegovic M, Xue B, Gorgun CZ, Hotamisligil GS, Neel BG *et al*. Neuronal PTP1B regulates body weight, adiposity and leptin action. *Nat Med* 2006; **12**: 917–924.

CAS PubMed Article Google Scholar

170. 170

Zhang X, Zhang G, Zhang H, Karin M, Bai H, Cai D . Hypothalamic IKKβ/NF-κB and ER stress link overnutrition to energy imbalance and obesity. *Cell* 2008; **135**: 61–73.

CAS PubMed PubMed Central Article Google Scholar

171. 171

Ozcan L, Ergin AS, Lu A, Chung J, Sarkar S, Nie D *et al*. Endoplasmic reticulum stress plays a central role in development of leptin resistance. *Cell Metab* 2009; **9**: 35–51.

CAS PubMed Article Google Scholar

172. 172

Ottaway N, Mahbod P, Rivero B, Norman LA, Gertler A, D'Alessio DA *et al*. Diet-induced obese mice retain endogenous leptin action. *Cell Metab* 2015; **21**: 877–882.

CAS PubMed PubMed Central Article Google Scholar

173. 173

Alwan A . *Global Status Report on Noncommunicable Diseases 2010*. World Health Organization: Geneva, Switzerland, 2011.

Google Scholar

174. 174

Ono H, Pocai A, Wang Y, Sakoda H, Asano T, Backer JM *et al*. Activation of hypothalamic S6 kinase mediates diet-induced hepatic insulin resistance in rats. *J Clin Invest* 2008; **118**: 2959–2968.

CAS PubMed PubMed Central Google Scholar

Download references

29.3.7: ACKNOWLEDGEMENTS

This work was supported by grants from the National Research Foundation (NRF-2014R1A6A3A01057664, NRF-2013M3C7A1056024 for M-SK) and the Asan Institute for Life Sciences (2013-326).

29.3.8: AUTHOR INFORMATION

Author notes

1. Eun Roh and Do Kyeong Song: These authors contributed equally to this work.

29.3.8.1: AFFILIATIONS

1. **Appeptite Regulation Laboratory, Asan Institute for Life Sciences, University of Ulsan College of Medicine, Seoul, Korea**
2. **Department of Medicine, University of Ulsan College of Medicine, Seoul, Korea**
3. **Division of Endocrinology and Metabolism, Asan Medical Center, Seoul, Korea**

29.3.8.2: CORRESPONDING AUTHOR

Correspondence to Min-Seon Kim.

29.3.9: ETHICS DECLARATIONS

29.3.9.1: COMPETING INTERESTS

The authors declare no conflict of interest.

29.3.10: RIGHTS AND PERMISSIONS

29.4: SKELETAL MUSCLE REGULATES METABOLISM

Josep M. Argilés, Nefertiti Campos, José M. Lopez-Pedrosa, Ricardo Rueda, Leocadio Rodriguez-Mañas. Journal of the American Medical Directors Association,
Volume 17, Issue 9, 2016, Pages 789-796,ISSN 1525-8610,https://doi.org/10.1016/j.jamda.2016.04.019.

29.4.0.1: ABSTRACT

Skeletal muscle is recognized as vital to physical movement, posture, and breathing. In a less known but critically important role, muscle influences energy and protein metabolism throughout the body. Muscle is a primary site for glucose uptake and storage, and it is also a reservoir of amino acids stored as protein. Amino acids are released when supplies are needed elsewhere in the body. These conditions occur with acute and chronic diseases, which decrease dietary intake while increasing metabolic needs. Such metabolic shifts lead to the muscle loss associated with sarcopenia and cachexia, resulting in a variety of adverse health and economic consequences. With loss of skeletal muscle, protein and energy availability is lowered throughout the body. Muscle loss is associated with delayed recovery from illness, slowed wound healing, reduced resting metabolic rate, physical disability, poorer quality of life, and higher health care costs. These adverse effects can be combatted with exercise and nutrition. Studies suggest dietary protein and leucine or its metabolite β-hydroxy β-methylbutyrate (HMB) can improve muscle function, in turn improving functional performance. Considerable evidence shows that use of high-protein oral nutritional supplements (ONS) can help maintain and rebuild muscle mass and strength. We review muscle structure, function, and role in energy and protein balance. We discuss how disease- and age-related malnutrition hamper muscle accretion, ultimately causing whole-body deterioration. Finally, we describe how specialized nutrition and exercise can restore muscle mass, strength, and function, and ultimately reverse the negative health and economic outcomes associated with muscle loss.

29.4.0.1: KEYWORDS

Muscle

glucose

amino acid

sarcopenia

HMB

ONS

Skeletal muscle is integral to physical movement, posture, and vital actions, such as chewing, swallowing, and breathing.[1, 2] Skeletal muscle also serves as a regulator of interorgan crosstalk for energy and protein metabolism throughout the body, a less recognized but critically important role. As such, skeletal muscle is a key site for glucose uptake and storage.[3] Skeletal muscle is likewise a reservoir of amino acids that can support protein synthesis or energy production elsewhere in the body when other sources are depleted.[4]

This review of muscle metabolism describes how amino acids stored as protein in muscle can be broken down through proteolysis for ultimate use in energy production. Such breakdown occurs when energy demands are high (as with stress-induced hypermetabolism), or when supplies are low (as in severe starvation or longer-term protein energy malnutrition). Both of these states can be hallmarks of many diseases, either directly as a result of disease-related dysregulation of metabolism (such as in the extreme case of cancer-cachexia) or, more subtly, as a result of the general illness-associated loss of appetite. Muscle is therefore crucially important during illness, both for its role in balancing the metabolic needs of other organs and for its reserves of protein for use in energy production. Yet, during illness, the maintenance of muscle mass through exercise and nutrition are often overlooked or difficult to address, and muscle atrophy develops. Even more subtle is aging-related muscle loss, which can dramatically increase morbidity and mortality of otherwise survivable illnesses in the aged. This review also illustrates the consequences of muscle atrophy in aging and illness and proposes steps to combat these challenges.

29.4.0.1: MUSCLE BASICS

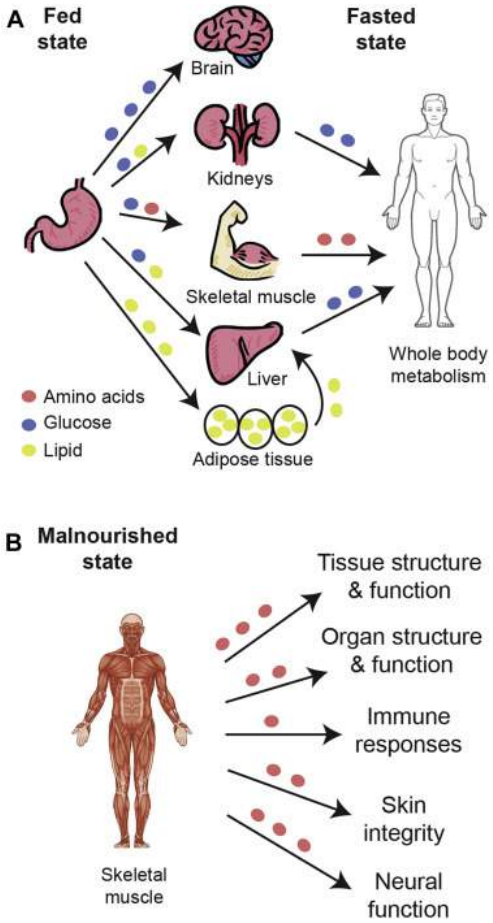
29.4.0.0.1: MUSCLE STRUCTURE AND CLASSIFICATION

Skeletal muscle comprises the fibrillar proteins myosin (a thick filament) and actin (a thin filament) that interact to cause muscle contraction, a process requiring energy in the form of adenosine triphosphate (ATP). Different muscle types have been classified according to histochemical features, structural protein composition, and major metabolic properties.[5, 6] Most commonly, skeletal muscles are referred to as either “slow” or “fast” to reflect speeds of contraction, or the shortening of myosin heavy chain (MHC) protein.[6] The velocity of this shortening is dependent on the MHC isoform present; “fast” fiber isoforms MHCIIa and IIb demonstrate a higher shortening velocity than their “slow” fiber MHCI counterparts.[6, 7] Classic histochemical staining methods also classify muscle as type I (slow) and type II (fast) based on the myosin ATPase enzyme forms revealed. Recently, these types have been further distinguished based on histology (types I, IC, IIC, IIAC, IIA, IIAB, and IIB).[6]

29.4.0.0.2: MUSCLE METABOLISM AND INTERORGAN CROSSTALK

Glucose regulation is central to energy balance both within muscle fibers and throughout the body. In the cytoplasm of most cells, glucose undergoes glycolysis to produce the substrate for ATP generation. Muscle fibers are also characterized on the basis of the speed and manner in which they metabolize glucose. The terms “fast” and “slow” can indicate the type of glucose metabolism occurring within the fiber. Slow muscles, which use aerobic metabolism, contain a high density of capillaries and oxidative enzymes that allow a greater resistance to fatigue.[7] Fast muscles, which depend on anaerobic metabolism, or glycolysis, can quickly generate ATP and therefore contract more readily. Fast muscles also fatigue sooner than slow fibers, as the conversion of glucose to pyruvate generates less ATP than can be generated by using the rest of central metabolism, ultimately generating CO_2.

Muscle has the ability to store glucose in the form of glycogen, which facilitates the rapid initiation of energy production for contraction even when glucose is not readily available from the diet. This unique capacity, shared also by the liver and kidneys, makes skeletal muscle an important metabolic organ that helps all organs have access to essential energy substrates during fasting. Furthermore, the amino acids stored in muscle as protein can be broken down as a last resort during times of starvation or extreme energy shortfalls.[4] Patterns of glucose utilization throughout the body as a whole reflect feeding status (Figure 1; Table 1). Based on a classic study of the fed state (measurement within 3 hours of eating), researchers estimated that 25% to 35% of an ingested carbohydrate load was quickly extracted from circulation and stored by the liver.[3] Of the remaining glucose, approximately 40% was disposed in the muscle and 10% in the kidney.[3] The brain used 15% to 20% of post-meal glucose.[3]



1. Download : Download high-res image (512KB)
2. Download : Download full-size image

Fig. 1. Glucose metabolism in fed, fasted and malnourished states. A, Glucose, lipids, and amino acids from the diet circulate during the fed state for use or storage in body organs. In the fasted state, glucose is released from the muscles, kidneys, and liver for whole-body metabolism, along with lipids from adipose tissue and amino acids from the muscle. B, When glucose stores have been depleted, amino acids are provided by the muscles to support crucial bodily functions.

Table 1. Glucose Metabolism in Fasted and Fed States

Fed State	Fasted State
Diet-sourced glucose (exogenous glucose) is absorbed from the intestine to circulate in blood; glucose serves as an energy source in cells throughout the body. Cytoplasmic glucose undergoes glycolysis, in turn producing ATP.	Little or no blood glucose from dietary sources; alternative energy sources are needed for function of tissues body-wide.
Glucose is primarily taken up by muscle and liver, where it can be used for energy or stored as glycogen (Glycogen synthesis).	Glycogen stored in liver, kidney, and muscle is broken down to provide glucose as energy source (Glycogenolysis). Muscle uses glycogen-sourced glucose internally; liver and kidney can supply glucose to circulation.
Gluconeogenic substrates are stored in various organs (eg, pyruvate in liver, glycerol in fat, and amino acids in muscle).	Endogenous generation of glucose from noncarbohydrate carbon substrates, such as pyruvate, lactate, glycerol, and glucogenic amino acids (Gluconeogenesis); occurs primarily in the liver and muscle, and to a lesser extent in the kidney.

In the fasted state (after 14 to 16 h without eating), the liver provides approximately 80% of glucose that is released into circulation. About half of this glucose comes from the breakdown of stored glycogen, and the rest from the metabolism of sources other than carbohydrate or glycogen, including certain amino acids, through a process known as gluconeogenesis.[8] Interactions between muscle and liver are largely responsible for regulating carbohydrate metabolism and for achieving energy balance in normal fed and fasted states; the kidneys play a role similar to that of the liver, but to a lesser extent.3, 8 In addition, muscle tissue stores amino acids as protein, and adipose tissue serves as a depot of glycerol and fatty acids. As needed, amino acids and fatty acids can be metabolized to form acetyl coenzyme A for the tricarboxylic acid (TCA) cycle.

As glycogen stores become depleted, increasingly more glucose is produced by gluconeogenesis. Gluconeogenesis provides 70% of glucose released into the body 24 hours after eating, and 90% by 48 hours.[8] As fasting is prolonged, the kidneys contribute increasingly higher amounts of glucose from gluconeogenesis.

Ultimately, amino acids stored in skeletal muscle are metabolized when the need for gluconeogenesis substrate is greatest. Skeletal muscle houses nearly 75% of all protein in the body and constitutes an important contributor to gluconeogenesis in states of drastic depletion. Maintenance of muscle protein content depends on the balance between protein synthesis and degradation.[5] Under normal conditions, muscle protein mass gains during the fed state balance losses during the fasted state.[4] However, under severe metabolic stress generated by serious illness or injury, muscle protein can become depleted by catabolism, and this can lead to harmful functional limitations.

Skeletal muscle proteolysis can provide amino acid substrates for glucose and glycogen formation, notably glutamine and alanine. Alanine is released into circulation and reaches the liver, where it serves as an excellent substrate for gluconeogenesis. Glutamine also has a beneficial role in this process: the carbon skeleton of glutamine is a gluconeogenic precursor that can regulate gluconeogenesis independently of the insulin/glucagon ratio. Therefore, glutamine supplementation may also enhance glycogen synthesis and increase muscle glycogen stores even when insulin levels are low or when insulin resistance is present.[9]

In summary, dietary glucose is supplied by meals, and glucose is stored as glycogen in liver, kidney, and muscle for metabolic energy functions, as needed (Table 1). At times when glucose supplies are not sufficient to meet energy needs, breakdown of glycogen (glycogenolysis) occurs. When stored glucose products are no longer available, energy is released by breakdown of substrates other than glucose. In this review of muscle metabolism, we emphasize that amino acids stored as protein in muscle can be broken down by way of gluconeogenesis, ultimately entering the TCA cycle for energy production. Such breakdown occurs when energy demands are high, as with stress-induced hypermetabolism of disease, or when supplies are low, as in severe starvation or disease-associated loss of appetite. Such use can become problematic in that it reduces skeletal muscle mass and produces waste nitrogen, which requires further energy to sequester and secrete. Prolonged reliance on these processes can accelerate existing health problems and must be addressed by the health care provider.

29.4.0.1: MUSCLE PLASTICITY: CHANGES IN MUSCLE MASS, STRENGTH, AND FUNCTION

Skeletal muscle is remarkably plastic. It changes continuously in response to calorie and nutrient intake, illness, and physical stress. Changes in adult skeletal muscle also may occur as fiber-type switching, which is influenced by changes in physical activity, loading, nerve stimulation, or hormone and cytokine levels.7, 10, 11, 12

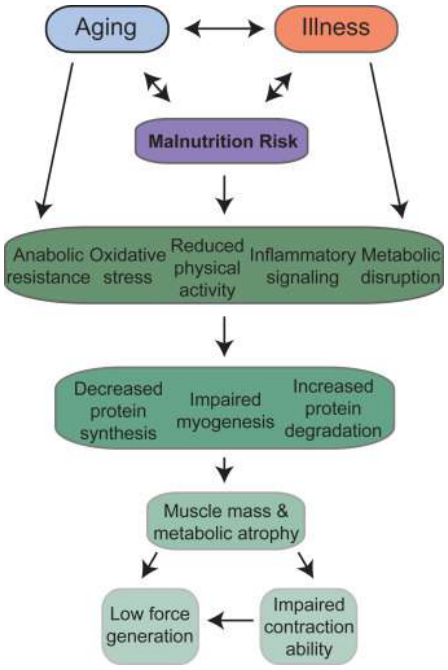
29.4.0.0.1: MECHANISMS OF MUSCLE GROWTH AND STRENGTH INCREASE

Muscle adapts positively to demands placed on it, such as the increased contractile activity associated with endurance training or the increased loading attributable to strength training. This tremendous plasticity is evident as muscle tissue accretion, specific changes within muscle, and muscle tissue breakdown.[1] Muscle growth, or hypertrophy, occurs when protein synthesis within the muscles outpaces protein degradation. This process can be positively regulated by mammalian target of rapamycin (mTOR) signaling induced by insulin after calorie ingestion, by hormones such as testosterone, and by exercise.[10]

29.4.0.0.2: MECHANISMS OF MUSCLE LOSS IN AGING, INACTIVITY, SICKNESS, AND FRAILTY

This ability of skeletal muscle to change dynamically in response to body conditions is also manifest as changes resulting from injury, illness, or aging. When the metabolic demands placed on muscle outweigh the protein synthesis that occurs from dietary intake and after exercise, muscle mass is lost, metabolic storage products are depleted, and muscle fiber balance changes.

Aging may lead to a loss of muscle mass resulting from both the shrinking of muscle fibers (atrophy) and the elimination of fibers altogether (Figure 2).[6] This condition is known as primary sarcopenia, the age-related loss of muscle mass and function. Although both fiber types I and II lose mass, aging causes preferential atrophy of type II fibers; the net change is thus from type II to type I fibers, or from fast to slow muscle fibers.[6, 13] Because fast muscle fibers mobilize ATP and create tension more readily than slow fibers, this shift can leave older adults without the energy to perform daily tasks.[14] This shift to type I slow fibers leads to a corresponding increase in their characteristic oxidative metabolism relative to the glycolytic metabolism that occurs in type II fast fibers. Exacerbating the problems caused by muscle degradation in aging, it is possible that type I oxidative fibers normally experience higher protein turnover (i. protein synthesis and degradation), are less able to grow in size, and have different responses to insufficient nutrient intake, although these fiber-type differences remain poorly understood.[7, 15, 16]



1. Download : Download high-res image (362KB)
2. Download : Download full-size image

Fig. 2. Effects of aging and illness on muscle mass.

Beyond aging, muscle wasting is associated with many pathological states and chronic diseases, such as malnutrition, cancer, chronic kidney disease, chronic obstructive pulmonary disease, burns, muscular dystrophies, acquired immunodeficiency syndrome, sepsis, and immune disorders, and forced immobilization and bed rest are devastating to patients who are already challenged by these factors (Figure 2).[14, 17] Most of these pathological conditions are associated with variable degrees of local and/or systemic chronic inflammation, which plays a crucial role in the onset of muscle atrophy. Loss of muscle mass is frequently associated with increased production of proinflammatory cytokines. Systemic inflammation is associated with reduced rates of protein synthesis paralleled by enhanced protein breakdown, both accounting for the loss of muscle mass. The effects exerted by proinflammatory cytokines on muscle mass are partially mediated by activating the transcription factor nuclear factor κB (NF-κB).[18] The transcriptional activity is regulated by the phosphorylation and consequent degradation of the inhibitor Iκ-Bα, allowing the positive regulation of muscle RING-finger protein-1 (MuRF1) and other atrophy-related genes. Proinflammatory cytokines act on muscle protein metabolism not only by activating catabolic pathways, but also by downregulating the anabolic pathways.[19] Elevated tumor necrosis factor-alpha (TNF-α) and interleukin-1 (IL-1) lead to inhibition of the Akt/mTOR signal transduction pathway and a subsequent reduction in protein synthesis. The inflammatory process that takes place during trauma or fractures is controlled and finely regulated. In the short term, it can facilitate complete and efficient reconstruction of muscle fibers through the stimulation of myogenesis. However, chronic inflammation can be deleterious, driving uncontrolled muscle atrophy and affecting contraction ability. Balance between pro- and anti-inflammatory cytokines is well known to be important in regulating physiological muscle protein turnover and myogenesis, and evidence suggesting that inflammation can impair force generation in muscles is also growing.[20, 21]

As inflammation accelerates muscle catabolism, resting energy expenditure increases and amino acids are released from muscles to serve as substrates for gluconeogenesis in liver and elsewhere in the body (Table 2).[22] The efficiency of energy production is low when amino acids are used to generate energy, so muscle is at further risk for breakdown to meet needs.[23] In addition, the liver changes metabolic priorities, using amino acids to produce acute phase reactant proteins instead of normal proteins, such as serum albumin, and to support

gluconeogenesis. This process continues until the cause of stress has subsided. Thus, when the dietary proteins supplied are inadequate to meet needs, muscle protein is broken down to supply amino acids throughout the body. This reaction releases waste nitrogen, which requires further energy to convert to urea, thereby exacerbating the problem of the energy shortfall.[24]

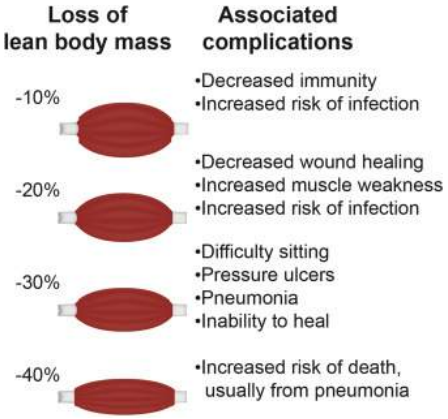
Table 2. Major Molecular Pathways Influencing Muscle Accretion

Effector	Mediator	Major Pathway(s)	Consequence
Mammalian target of rapamycin (mTOR)	+Induced by BCAAs, HMB	Interacts with protein translation machinery to facilitate initiation and elongation	mTOR stimulation by a number of pathways increases protein synthesis
Insulinlike growth factor (IGF1)	+Stimulated by meal-induced insulin	IGF1R → PI3K → AKT → mTOR	Reduced IGF1 from decreased eating and/or exercise leads to reduced protein synthesis and to muscle wasting
	+Stimulated by exercise		
Myostatin/Activin	+Produced by skeletal muscle	Activin receptors (ACTRIIA/B) → Smad2/3 –I mTOR	Myostatins negatively regulate protein synthesis
	–Inhibited by Follistatin	ACTRIIA/B → FoxO → UPS	
Inflammatory cytokines (TNFα, IL-1)	+Upregulated by illness, injury	Cytokine receptors → NFKB, p38, JAK, Caspases, E3 ligases	Inflammation leads to apoptosis or autophagy-mediated muscle cell loss
	–Inhibited by exercise	FoxO transcription factors → MAFBX; MURF1 → UPS (ubiquitin-proteasome system)	
Vitamin D	+Levels are increased by diet and sunlight	Vitamin D receptor → gene expression or repression in myogenic cells	Vitamin D positively influences muscle growth

AKT, protein kinase B; BCAA, branched-chain amino acid; FoxO, forkhead box protein O; HMB, β-Hydroxy β-Methylbutyrate; IGF1R, insulinlike growth factor 1 receptor; JAK, janus kinase; MURF1, muscle RING-finger protein-1 p38, mitogen-activated protein kinase; PI3K, phosphatidylinositide 3-kinase; UPS, ubiquitin-proteasome system.

29.4.0.0.3: COMPLICATIONS ASSOCIATED WITH LOSS OF MUSCLE

As aging and illness lead to muscle breakdown and atrophy, reduced muscle mass leaves patients without a crucial reservoir of amino acids and effector molecules, such as myokines, cytokines released by muscle, to help the body combat illness, infection, and wasting (Figure 3).[23, 25, 26, 27] Therefore, muscle atrophy is associated with a wide range of harmful health effects that can be life changing, especially for older people.[28, 29, 30, 31, 32, 33, 34, 35, 36] The most relevant condition associated with the presence of sarcopenia in this population is a clinical syndrome called frailty. The most accepted physiological framework explaining frailty and its consequences was proposed by Walston and Fried,[37] who described a relationship between sarcopenia and energy imbalance called the "frailty cycle." This cycle affects multiple systems, especially those susceptible to changes in hormones (mainly sexual hormones, IGF-1, and insulin) and the progressive development of a proinflammatory state.[38, 39, 40] Additional biomarkers have recently been identified for roles in frailty, such as those related to endothelial dysfunction or micro RNAs central to the aging process.[41, 42] Frailty can be defined as an age-associated biological syndrome characterized by a decreased biological reserve resulting from a decline in multiple physiological systems that leaves the individual at risk for developing poor outcomes (disability, death, and hospitalization) in the presence of stressors.[43, 44] The prevalence of frailty in people older than 65 is approximately 10%, increases with age, and is greater in women.[45]



1. Download : Download high-res image (273KB)
2. Download : Download full-size image

Fig. 3. Complications of lean body mass (muscle) loss.

Frailty is now a recognized clinical medical syndrome that provides a biological framework for understanding vulnerabilities resulting from aging or chronic conditions.[44, 46] It is clinically important to detect frailty in those at risk of developing disability. As aging progresses, frailty increases as the prognostic factor for death and incident disability.[47, 48] Frailty and its underlying sarcopenia have been shown to

predict risk of death, disability, and other adverse outcomes, including muscle mass atrophy, metabolic deterioration, slowed wound or postsurgical healing, and delayed recovery from illness.[32, 34, 35] Frailty and the weakness that follows muscle loss lead to higher risk of falls, fractures,[30] physical disability,[29] need for institutional care,[29] reduced quality of life,[36] and heightened mortality.[29, 33] Early identification of frailty risk provides the opportunity to provide interventions and avoid or delay disability as well as enhance recovery.

Loss of muscle associated with disease, injury, disuse, or aging significantly increases the cost of health care.[34, 49, 50] Results of a recent study showed that older adults (mean age=70 years) who were very frail spent (Euro) 1917 more on total health costs in an interval of 3 months than did those who were not frail.[51] In the United States, the direct cost of cachexia/sarcopenia to health care was reported to be 1.5% of annual total health care expenses.[50] Such costs arise from the increased rate of hospitalization, incidence of complications, lengths of stay, and likelihood of readmission.[52, 53] In the face of an aging population, the importance of identifying, preventing, and treating muscle loss cannot be overstated.

29.4.0.1: DETECTION AND TREATMENT OF MUSCLE LOSS

29.4.0.0.1: WHO IS AT RISK OF MUSCLE ATROPHY, AND HOW DO WE IDENTIFY IT?

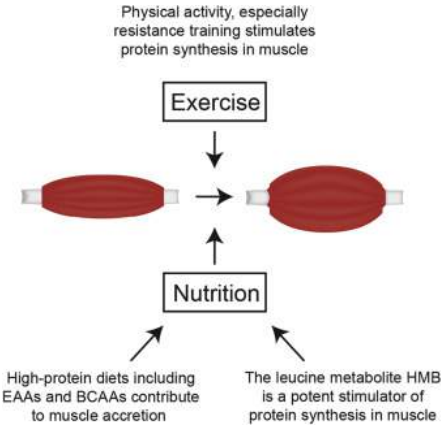
Screening is crucial for predicting risk, and proper, timely intervention can reduce or eliminate the ensuing muscle mass and metabolic atrophy, substantially affecting morbidity, mortality, and cost. Special attention should be paid to the main risk categories: people who are malnourished or at risk of malnutrition for any reason[33, 54]; frail adults, especially the very old; people who become deconditioned and lose muscle due to age- and disability-related physical inactivity[35]; those with diseases or conditions with inflammatory components, such as chronic heart failure,[55] chronic or acute kidney disease,[56] cancer,[57, 58, 59] severe infection and sepsis,[60] insulin resistance/diabetes,[61] intensive care unit–acquired weakness,[25] and wound/surgical recovery.[34]

Reaching an accurate diagnosis of age- or disease-related muscle atrophy is difficult, and a number of criteria have been proposed but have not yet assessed in the clinical setting.[14] Nonetheless, specific criteria and measures can be used to diagnose sarcopenia or cachexia.[13, 27, 62, 63] Sarcopenia can be diagnosed when a patient has muscle mass that is at least 2 SDs below the relevant population mean and also presents with a low gait (walking) speed. In addition, low muscle strength and general physical performance may be taken into consideration.[14] Cachexia can be diagnosed when at least 5% of body weight is lost within 12 months in the presence of underlying illness, and 3 of the following criteria are also met: decreased muscle strength, increased fatigue, anorexia, low fat-free mass index, abnormal biochemistry, increased inflammatory markers C-reactive protein (>5.0 mg/L) or IL-6 (>4.0 pg/mL), anemia (<12 g/dL), or low serum albumin (<3.2 g/dL).

Recent research into the molecular adaptations associated with the development of or that result from muscle atrophy and metabolic depletion may lead to the identification of biomarkers and, therefore, improvements in early detection (Table 2). A variety of signaling pathways known to positively influence muscle growth (bone morphogenetic proteins, brain-derived neurotrophic factors, follistatin, and irisin), as well as those known to negatively regulate muscle growth (transforming growth factor β, myostatin, activins, and growth and differentiation factor-15) and factors associated with muscle function and dysfunction (C-terminal agrin fragment and skeletal muscle specific troponin T) may emerge as biomarkers for muscle atrophy in aging and disease.[64] To date, there is no universally recognized biomarker for muscle atrophy, but recent research in the field suggests that the combination of several biomarkers may facilitate the adequate diagnosis of muscle atrophy. Identification of such biomarkers and their incorporation into validated testing instruments should allow early identification of muscle atrophy (improving prognosis, and likely reducing cost to health care systems), but may also provide exciting targets for the development of new medications.

29.4.0.0.2: NUTRITIONAL STRATEGIES FOR MAINTAINING AND REBUILDING MUSCLE

Treatment of patients at risk can prevent or delay onset of muscle atrophy, or even target rebuilding of muscle when muscle atrophy is already evident (Figure 4).[65] As a first step, treatment must provide adequate energy so that muscle proteins and their constituent amino acids are spared as an energy source. In addition, high protein intake is vital to treatment of muscle atrophy or for delaying its onset.[7, 66, 67, 68, 69] It should be noted that the range of protein needs can vary widely from patient to patient. Because muscle mass may decrease or remain the same (based largely on how much protein synthesis outpaces protein degradation), the most direct way to prevent muscle loss is to ensure that sufficient protein is ingested. Use of high-protein oral nutritional supplements (ONS; ≥20% of total calories as protein) may be beneficial to such patients.[70]



1. Download : Download high-res image (190KB)
2. Download : Download full-size image

Fig. 4. Treatments for sarcopenia. It is currently recommended that patients at risk of or suffering from sarcopenia consume a diet high in protein, engage in resistance exercise, and take supplements of the leucine metabolite HMB.

By definition, the essential amino acids (EAAs) play a central role in protein nutritional status. Some amino acids play roles that are distinct from the traditional one of protein building blocks; many of these have little or nothing to do with protein synthesis, and are thus not included here. However, of central importance to the current discussion are the branched-chain amino acids (BCAAs), especially leucine.65, 71, 72 BCAAs promote protein synthesis in the muscles through a number of pathways.[66] In particular, they are now known to have a key role in altering tissue response to a meal, the post-prandial response, especially in muscle, where they signal a reduction in protein breakdown and an increase in protein synthesis, resulting in net accretion of protein in muscle and helping to regulate blood amino acid levels. However, aspects of this postprandial regulation are not as robust in aged muscle, and muscle in hypercatabolic conditions, such as cancer, is challenged and its normal system is overwhelmed. In these cases, a substantial body of research suggests that significantly more of these amino acids are required in the diet to overcome resistance to protein anabolism; very high doses, such as 10 to 15 g of BCAAs, or 3 g or more of leucine per meal, have been studied to combat muscle loss in the elderly,[44] although this may be a result of improved protein synthesis that does not lead to muscle mass accretion.73, 74, 75

This resistance to the normal of BCAAs in muscle protein homeostasis has prompted studies into leucine's mechanism of action. These have identified the leucine metabolite β-hydroxy β-methylbutyrate (HMB) as a potent stimulator of protein synthesis as well as an inhibitor of protein breakdown in the extreme case of cachexia.65, 72, 76, 77, 78, 79, 80, 81, 82, 83, 84 A growing body of evidence suggests HMB may help slow, or even reverse, the muscle loss experienced in sarcopenia and improve measures of muscle strength.44, 65, 72, 76, 77, 78, 79, 80, 81, 82, 83, 84 However, dietary leucine does not provide a large amount of HMB: only a small portion, as little as 5%, of catabolized leucine is metabolized into HMB.[85] Thus, although dietary leucine itself can lead to a modest stimulation of protein synthesis by producing a small amount of HMB, direct ingestion of HMB more potently affects such signaling, resulting in demonstrable muscle mass accretion.71, 80 Indeed, a vast number of studies have found that supplementation of HMB to the diet may reverse some of the muscle loss seen in sarcopenia and in hypercatabolic disease.65, 72, 83, 86, 87 The overall treatment of muscle atrophy should include dietary supplementation with HMB, although the optimal dosage for each condition is still under investigation.[68]

In addition to dietary protein, EAAs/BCAAs including leucine, the leucine metabolite HMB, a number of other dietary or supplemental components have been explored for their ability to positively influence muscle mass during sarcopenia. These include creatine monohydrate, a variety of antioxidants, ornithine α-ketoglutarate, omega-3 fatty acids, ursolic acid, and nitrates.68, 88, 89, 90 Given the length of this list, its growing nature, and the difficulty many aged individuals experience ingesting proper calories and nutrients, additional studies will be needed to determine which components are most beneficial to maintaining muscle mass, as well as their optimal doses and administration routes both in isolation and in combination.

29.4.0.0.3: PHYSICAL ACTIVITY IS ALSO KEY

Nutrition is important and can counteract metabolic alterations induced during periods of significant stress and inflammation; however, sufficient exogenous provisions of protein and energy substrates alone cannot completely eliminate or reverse the deteriorations associated with aging or the deleterious impact inadequate control and regulation of inflammation have on muscle.23, 66, 69 Protein synthesis occurs in muscle fibers following their contraction,[91] and physical activity has been shown to induce a number of anabolic signaling pathways.[92] Physical activity can likewise reduce degradation of muscle protein.93, 94 Even more, a lack of physical activity increases the resistance of muscle to anabolism, particularly the synthesis of proteins from amino acids.[95] An exercise component to muscle atrophy treatment is therefore highly recommended, and exercise also may prevent the onset of sarcopenia later in life, possibly by increasing the presence of type I fibers that are less susceptible to degradation during sarcopenia.7, 25, 66, 89 Although aerobic and other types of exercise are all

preferable to a lack of physical activity, resistance exercise in particular has been shown repeatedly to improve rates of protein synthesis and reverse muscle loss.[88, 89, 94] This may be attributable to differential effects on muscle fiber types.[7] It is therefore recommended that patients with muscle atrophy or at risk of developing muscle atrophy engage a regular exercise program containing both aerobic and anaerobic components, and the importance of appropriate resistance training cannot be overstated. Although this must be tailored to the individual's current physical status, it should also periodically be reviewed and increased to maximize its impact.

Although resistance exercise and general physical activity are important to the stimulation of protein synthesis from amino acids in the diet, some aged and ill individuals experiencing extensive muscle atrophy are likely to have difficulty engaging in physical activities because of low energy and other medical complications. Nutrition and some supplements can be used to bolster results of exercise, both preventively and during sarcopenia. For example, bioactive substances known as nutraceuticals that mimic the molecular effects of exercise can induce signaling pathways that are thought to support or even underlie exercise's effects on health and muscle mass accretion. Found in a variety of foods, including some common fruits, green tea, and even red wine, these compounds can be isolated and added to nutritional supplements used in the treatment of muscle atrophy.[44]

29.4.0.1: SUMMARY AND CONCLUSIONS

The classic physical functions of skeletal muscle are well known, but skeletal muscle is increasingly recognized as a one of the key regulators of energy and protein metabolism by way of metabolic crosstalk between body organs. Skeletal muscle is the primary site for glucose uptake and storage, and it is likewise a reservoir of amino acids that sustain protein synthesis in all other body sites. When dietary glucose intake decreases or metabolic needs increase, stored glucose is mobilized from liver, while energy is released from fat depots. When these energy supplies are depleted, the muscle reservoir of amino acids stores is tapped, and muscle proteins are broken down to provide amino acids for gluconeogenesis, thereby supplying energy to other parts of the body.

Undernutrition and resultant muscle loss (muscle atrophy), as associated with aging and disease, can lead to adverse health and economic consequences. Conditions and diseases that lower dietary intake and increase nutrient needs are associated with catabolism of skeletal muscle, which in turn limits availability of protein and energy throughout the body. Loss of muscle mass, strength, and function has adverse consequences: slowed wound healing and recovery from illness, physical disability (due both to overall reduction of muscle status), as well as selective losses in type I fibers, which are essential for balance recovery (and thus fall prevention), poorer quality of life, and higher health care costs.

Nutrition and exercise are key to growth and maintenance of muscle promoting overall health, well-being, and recovery from disease. A wealth of research underscores the importance of a few key dietary components: protein (EAAs/BCAAs in particular), and the leucine metabolite HMB. Others will very likely be added to this list as our knowledge base grows. In addition, physical activity, especially resistance strength training, is essential to the treatment of muscle atrophy. Others will very likely be added to this list as our knowledge base grows. Considerable evidence shows that ONS and enteral feeding formulations can help maintain and rebuild muscle mass and strength. Further studies are needed to show support for functional outcomes, such as ability to perform activities of daily living and maintain or restore independence.

29.4.0.1: ACKNOWLEDGMENTS

The authors thank Jeffrey H. Baxter and Abby Sauer from ANR&D for their critical review of this article, as well as Cecilia Hofmann, PhD, and Hilary North Scheler, PhD (C Hofmann & Associates, Western Springs, IL), for valuable assistance with efficient compilation of the medical literature and with editing this English-language review article.

29.4.0.1: REFERENCES

1

J.A. Chromiak, J. Antonio

Skeletal muscle plasticity

J. Antonio, D. Kalman, J. Stout (Eds.), Essentials of Sports Nutrition and Supplements, Humana Press, Totowa, NJ (2008)

Google Scholar

2

H. Shiozu, M. Higashijima, T. Koga

Association of sarcopenia with swallowing problems, related to nutrition and activities of daily living of elderly individuals

J Phys Ther Sci, 27 (2015), pp. 393-396

CrossRefView Record in ScopusGoogle Scholar

3

C. Meyer, J.M. Dostou, S.L. Welle, J.E. Gerich

Role of human liver, kidney, and skeletal muscle in postprandial glucose homeostasis

Am J Physiol Endocrinol Metab, 282 (2002), pp. E419-E427
CrossRefView Record in ScopusGoogle Scholar

4

R.R. Wolfe
The underappreciated role of muscle in health and disease
Am J Clin Nutr, 84 (2006), pp. 475-482
CrossRefView Record in ScopusGoogle Scholar

5

W.R. Frontera, J. Ochala
Skeletal muscle: a brief review of structure and function
Calcif Tissue Int, 96 (2015), pp. 183-195
CrossRefView Record in ScopusGoogle Scholar

6

W. Scott, J. Stevens, S.A. Binder-Macleod
Human skeletal muscle fiber type classifications
Phys Ther, 81 (2001), pp. 1810-1816
CrossRefView Record in ScopusGoogle Scholar

7

A. Matsakas, K. Patel
Skeletal muscle fibre plasticity in response to selected environmental and physiological stimuli
Histol Histopathol, 24 (2009), pp. 611-629
View Record in ScopusGoogle Scholar

8

M. Shrayyef, J. Gerich
Normal glucose homeostasis
L. Poretsky (Ed.), Principles of Diabetes Mellitus (2nd ed), Springer, New York, NY, US (2010)
Google Scholar

9

G. Perriello, N. Nurjhan, M. Stumvoll, *et al.*
Regulation of gluconeogenesis by glutamine in normal postabsorptive humans
Am J Physiol, 272 (1997), pp. E437-E445
View Record in ScopusGoogle Scholar

10

S. Schiaffino, C. Reggiani
Fiber types in mammalian skeletal muscles
Physiol Rev, 91 (2011), pp. 1447-1531
CrossRefView Record in ScopusGoogle Scholar

11

T. Eken, K. Gundersen
Electrical stimulation resembling normal motor-unit activity: effects on denervated fast and slow rat muscles
J Physiol, 402 (1988), pp. 651-669
CrossRefView Record in ScopusGoogle Scholar

12

D. Pette
The adaptive potential of skeletal muscle fibers
Can J Appl Physiol, 27 (2002), pp. 423-448
CrossRefView Record in ScopusGoogle Scholar

13
S. von Haehling, J.E. Morley, S.D. Anker
From muscle wasting to sarcopenia and myopenia: update 2012
J Cachexia Sarcopenia Muscle, 3 (2012), pp. 213-217
CrossRefView Record in ScopusGoogle Scholar
14
M.J. Tisdale
Catabolic mediators of cancer cachexia
Curr Opin Support Palliat Care, 2 (2008), pp. 256-261
View Record in ScopusGoogle Scholar
15
B. Mittendorfer, J.L. Andersen, P. Plomgaard, *et al.*
Protein synthesis rates in human muscles: Neither anatomical location nor fibre-type composition are major determinants
J Physiol, 563 (2005), pp. 203-211
CrossRefView Record in ScopusGoogle Scholar
16
C.A. Goodman, J.A. Kotecki, B.L. Jacobs, T.A. Hornberger
Muscle fiber type-dependent differences in the regulation of protein synthesis
PLoS One, 7 (2012), p. e37890
CrossRefView Record in ScopusGoogle Scholar
17
S. Ventadour, D. Attaix
Mechanisms of skeletal muscle atrophy
Curr Opin Rheumatol, 18 (2006), pp. 631-635
CrossRefView Record in ScopusGoogle Scholar
18
H. Li, S. Malhotra, A. Kumar
Nuclear factor-kappa B signaling in skeletal muscle atrophy
J Mol Med (Berl), 86 (2008), pp. 1113-1126
CrossRefView Record in ScopusGoogle Scholar
19
H.L. Eley, M.J. Tisdale
Skeletal muscle atrophy, a link between depression of protein synthesis and increase in degradation
J Biol Chem, 282 (2007), pp. 7087-7097
ArticleDownload PDFView Record in ScopusGoogle Scholar
20
D. Costamagna, P. Costelli, M. Sampaolesi, F. Penna
Role of inflammation in muscle homeostasis and myogenesis
Mediators Inflamm, 2015 (2015), p. 805172
View Record in ScopusGoogle Scholar
21
R.C. Langen, A.M. Schols, M.C. Kelders, *et al.*
Inflammatory cytokines inhibit myogenic differentiation through activation of nuclear factor-kappaB
FASEB J, 15 (2001), pp. 1169-1180
View Record in ScopusGoogle Scholar
22
M. Winkler, A. Malone

Medical nutrition therapy for metabolic stress: Sepsis, trauma, burns, and surgery

K. Mahon, J. Raymond, S. Escott-Stump (Eds.), Krause's Food and the Nutrition Care Process, Elsevier, St Louis, MO (2012)

Google Scholar

23

M. Fischer, A. JeVenn, P. Hipskind

Evaluation of muscle and fat loss as diagnostic criteria for malnutrition

Nutr Clin Pract, 30 (2015), pp. 239-248

CrossRefView Record in ScopusGoogle Scholar

24

DL Nelson, MM Cox

Lehninger Principles of Biochemistry

(6th ed), W. H. Freeman & Co, New York, NY (2013), pp. 884-900

Google Scholar

25

D.K. Heyland, R.D. Stapleton, M. Mourtzakis, *et al.*

Combining nutrition and exercise to optimize survival and recovery from critical illness: Conceptual and methodological issues

Clin Nutr (2015), 10.1016/j.clnu.2015.07.003

Google Scholar

26

E. Curtis, A. Litwic, C. Cooper, E. Dennison

Determinants of muscle and bone aging

J Cell Physiol, 230 (2015), pp. 2618-2625

CrossRefView Record in ScopusGoogle Scholar

27

A.J. Cruz-Jentoft, J.P. Baeyens, J.M. Bauer, *et al.*

Sarcopenia: European consensus on definition and diagnosis: Report of the European working group on sarcopenia in older people

Age Ageing, 39 (2010), pp. 412-423

CrossRefView Record in ScopusGoogle Scholar

28

B.C. Clark, T.M. Manini

Functional consequences of sarcopenia and dynapenia in the elderly

Curr Opin Clin Nutr Metab Care, 13 (2010), pp. 271-276

View Record in ScopusGoogle Scholar

29

V. Hirani, F. Blyth, V. Naganathan, *et al.*

Sarcopenia is associated with incident disability, institutionalization, and mortality in community-dwelling older men: The Concord Health and Ageing in Men Project

J Am Med Dir Assoc, 16 (2015), pp. 607-613

ArticleDownload PDFView Record in ScopusGoogle Scholar

30

D. Scott, R.M. Daly, K.M. Sanders, P.R. Ebeling

Fall and fracture risk in sarcopenia and dynapenia with and without obesity: The role of lifestyle interventions

Curr Osteoporos Rep, 13 (2015), pp. 235-244

View Record in ScopusGoogle Scholar

31

S. Yu, K. Umapathysivam, R. Visvanathan

Sarcopenia in older people

Int J Evid Based Healthc, 12 (2014), pp. 227-243

View Record in ScopusGoogle Scholar

32

R.H. Demling

Nutrition, anabolism, and the wound healing process: An overview

Eplasty, 9 (2009), p. e9

Google Scholar

33

A.P. Cerri, G. Bellelli, A. Mazzone, *et al.*

Sarcopenia and malnutrition in acutely ill hospitalized elderly: Prevalence and outcomes

Clin Nutr, 34 (2015), pp. 745-751

ArticleDownload PDFView Record in ScopusGoogle Scholar

34

J. Friedman, A. Lussiez, J. Sullivan, *et al.*

Implications of sarcopenia in major surgery

Nutr Clin Pract, 30 (2015), pp. 175-179

CrossRefView Record in ScopusGoogle Scholar

35

B.T. Wall, M.L. Dirks, L.J. van Loon

Skeletal muscle atrophy during short-term disuse: Implications for age-related sarcopenia

Ageing Res Rev, 12 (2013), pp. 898-906

ArticleDownload PDFView Record in ScopusGoogle Scholar

36

R. Rizzoli, J.Y. Reginster, J.F. Arnal, *et al.*

Quality of life in sarcopenia and frailty

Calcif Tissue Int, 93 (2013), pp. 101-120

CrossRefView Record in ScopusGoogle Scholar

37

J. Walston, L.P. Fried

Frailty and the older man

Med Clin North Am, 83 (1999), pp. 1173-1194

ArticleDownload PDFView Record in ScopusGoogle Scholar

38

L. Carcaillon, F.J. Garcia-Garcia, J.A. Tresguerres, *et al.*

Higher levels of endogenous estradiol are associated with frailty in postmenopausal women from the Toledo study for healthy aging

J Clin Endocrinol Metab, 97 (2012), pp. 2898-2906

CrossRefView Record in ScopusGoogle Scholar

39

B.W. Penninx, S.B. Kritchevsky, A.B. Newman, *et al.*

Inflammatory markers and incident mobility limitation in the elderly

J Am Geriatr Soc, 52 (2004), pp. 1105-1113

View Record in ScopusGoogle Scholar

40

J. Walston, M.A. McBurnie, A. Newman, *et al.*

Frailty and activation of the inflammation and coagulation systems with and without clinical comorbidities: Results from the Cardiovascular Health Study

Arch Intern Med, 162 (2002), pp. 2333-2341

View Record in ScopusGoogle Scholar

41

C. Alonso-Bouzon, L. Carcaillon, F.J. Garcia-Garcia, *et al.*

Association between endothelial dysfunction and frailty: The Toledo Study for Healthy Aging

Age (Dordr), 36 (2014), pp. 495-505

CrossRefView Record in ScopusGoogle Scholar

42

L. Rodriguez Manas

Determinants of frailty and longevity: Are they the same ones?

Nestle Nutr Inst Workshop Ser, 83 (2015), pp. 29-39

View Record in ScopusGoogle Scholar

43

L. Rodriguez-Manas, C. Feart, G. Mann, *et al.*

Searching for an operational definition of frailty: A Delphi method based consensus statement: The frailty operative definition-consensus conference project

J Gerontol A Biol Sci Med Sci, 68 (2013), pp. 62-67

CrossRefView Record in ScopusGoogle Scholar

44

A. Clegg, J. Young, S. Iliffe, *et al.*

Frailty in elderly people

Lancet, 381 (2013), pp. 752-762

ArticleDownload PDFView Record in ScopusGoogle Scholar

45

R.M. Collard, H. Boter, R.A. Schoevers, R.C. Oude Voshaar

Prevalence of frailty in community-dwelling older persons: A systematic review

J Am Geriatr Soc, 60 (2012), pp. 1487-1492

CrossRefView Record in ScopusGoogle Scholar

46

L. Rodriguez-Manas, L.P. Fried

Frailty in the clinical scenario

Lancet, 385 (2015), pp. e7-e9

ArticleDownload PDFView Record in ScopusGoogle Scholar

47

P. Abizanda, L. Romero, P.M. Sanchez-Jurado, *et al.*

Age, frailty, disability, institutionalization, multimorbidity or comorbidity. Which are the main targets in older adults?

J Nutr Health Aging, 18 (2014), pp. 622-627

CrossRefView Record in ScopusGoogle Scholar

48

N. Sourial, H. Bergman, S. Karunananthan, *et al.*

Implementing frailty into clinical practice: A cautionary tale

J Gerontol A Biol Sci Med Sci, 68 (2013), pp. 1505-1511

CrossRefView Record in ScopusGoogle Scholar

49

C. Beaudart, R. Rizzoli, O. Bruyere, *et al.*

Sarcopenia: Burden and challenges for public health

Arch Public Health, 72 (2014), p. 45

View Record in ScopusGoogle Scholar

50

I. Janssen, D.S. Shepard, P.T. Katzmarzyk, R. Roubenoff

The healthcare costs of sarcopenia in the United States

J Am Geriatr Soc, 52 (2004), pp. 80-85

View Record in ScopusGoogle Scholar

51

J.O. Bock, H.H. Konig, H. Brenner, *et al.*

Associations of frailty with health care costs - results of the ESTHER cohort study

BMC Health Serv Res, 16 (2016), p. 128

View Record in ScopusGoogle Scholar

52

S. Gariballa, A. Alessa

Sarcopenia: Prevalence and prognostic significance in hospitalized patients

Clin Nutr, 32 (2013), pp. 772-776

ArticleDownload PDFView Record in ScopusGoogle Scholar

53

S. Joglekar, P.N. Nau, J.J. Mezhir

The impact of sarcopenia on survival and complications in surgical oncology: A review of the current literature

J Surg Oncol, 112 (2015), pp. 503-509

CrossRefView Record in ScopusGoogle Scholar

54

P. Abizanda, A. Sinclair, N. Barcons, *et al.*

Costs of malnutrition in institutionalized and community-dwelling older adults: A systematic review

J Am Med Dir Assoc, 17 (2016), pp. 17-23

ArticleDownload PDFView Record in ScopusGoogle Scholar

55

S. von Haehling

The wasting continuum in heart failure: From sarcopenia to cachexia

Proc Nutr Soc, 74 (2015), pp. 1-11

View Record in ScopusGoogle Scholar

56

P. Stenvinkel, J.J. Carrero, F. von Walden, *et al.*

Muscle wasting in end-stage renal disease promulgates premature death: established, emerging and potential novel treatment strategies

Nephrol Dial Transplant (2015)

Google Scholar

57

J.M. Argiles

Cancer-associated malnutrition

Eur J Oncol Nurs, 9 (2005), pp. S39-S50

ArticleDownload PDFView Record in ScopusGoogle Scholar

58

J.M. Argiles, S. Busquets, B. Stemmler, F.J. Lopez-Soriano

Cancer cachexia: Understanding the molecular basis

Nat Rev Cancer, 14 (2014), pp. 754-762
CrossRefView Record in ScopusGoogle Scholar

59

J.M. Argiles, C.C. Fontes-Oliveira, M. Toledo, *et al.*

Cachexia: A problem of energetic inefficiency

J Cachexia Sarcopenia Muscle, 5 (2014), pp. 279-286
CrossRefView Record in ScopusGoogle Scholar

60

C.E. Baldwin, A.D. Bersten

Myopathic characteristics in septic mechanically ventilated patients

Curr Opin Clin Nutr Metab Care, 18 (2015), pp. 240-247
View Record in ScopusGoogle Scholar

61

C. Guillet, Y. Boirie

Insulin resistance: A contributing factor to age-related muscle mass loss?

Diabetes Metab, 31 (2005)
5S20–5S26
Google Scholar

62

B.P. Martinez, A.K. Batista, I.B. Gomes, *et al.*

Frequency of sarcopenia and associated factors among hospitalized elderly patients

BMC Musculoskelet Disord, 16 (2015), p. 108
View Record in ScopusGoogle Scholar

63

W.J. Evans, J.E. Morley, J. Argiles, *et al.*

Cachexia: A new definition

Clin Nutr, 27 (2008), pp. 793-799
ArticleDownload PDFView Record in ScopusGoogle Scholar

64

A. Kalinkovich, G. Livshits

Sarcopenia—The search for emerging biomarkers

Ageing Res Rev, 22 (2015), pp. 58-71
ArticleDownload PDFView Record in ScopusGoogle Scholar

65

A.J. Cruz-Jentoft, F. Landi, S.M. Schneider, *et al.*

Prevalence of and interventions for sarcopenia in ageing adults: A systematic review. Report of the International Sarcopenia Initiative (EWGSOP and IWGS)

Age Ageing, 43 (2014), pp. 748-759
CrossRefGoogle Scholar

66

N.E. Deutz, J.M. Bauer, R. Barazzoni, *et al.*

Protein intake and exercise for optimal muscle function with aging: Recommendations from the ESPEN Expert Group

Clin Nutr, 33 (2014), pp. 929-936
ArticleDownload PDFView Record in ScopusGoogle Scholar

67

E. Arentson-Lantz, S. Clairmont, D. Paddon-Jones, *et al.*

Protein: A nutrient in focus

Appl Physiol Nutr Metab, 40 (2015), pp. 755-761
CrossRefView Record in ScopusGoogle Scholar

68

R. Calvani, A. Miccheli, F. Landi, *et al.*

Current nutritional recommendations and novel dietary strategies to manage sarcopenia

J Frailty Aging, 2 (2013), pp. 38-53
View Record in ScopusGoogle Scholar

69

J. Bauer, G. Biolo, T. Cederholm, *et al.*

Evidence-based recommendations for optimal dietary protein intake in older people: A position paper from the PROT-AGE Study Group

J Am Med Dir Assoc, 14 (2013), pp. 542-559
ArticleDownload PDFView Record in ScopusGoogle Scholar

70

A.C. Milne, J. Potter, A. Vivanti, A. Avenell

Protein and energy supplementation in elderly people at risk from malnutrition

Cochrane Database Syst Rev (2) (2009), p. CD003288
View Record in ScopusGoogle Scholar

71

M. Girón, J. Vílchez, R. Salto, *et al.*

Conversion of leucine to β-hydroxy-β-methylbutyrate by α-keto isocaproate dioxygenase is required for a potent stimulation of protein synthesis in L6 rat myotubes

J Cachexia, Sarcopenia, and Muscle, 7 (2016), pp. 68-78
CrossRefView Record in ScopusGoogle Scholar

72

D.J. Wilkinson, T. Hossain, D.S. Hill, *et al.*

Effects of leucine and its metabolite beta-hydroxy-beta-methylbutyrate on human skeletal muscle protein metabolism

J Physiol, 591 (2013), pp. 2911-2923
CrossRefView Record in ScopusGoogle Scholar

73

S.L. Casperson, M. Sheffield-Moore, S.J. Hewlings, D. Paddon-Jones

Leucine supplementation chronically improves muscle protein synthesis in older adults consuming the RDA for protein

Clin Nutr, 31 (2012), pp. 512-519
ArticleDownload PDFView Record in ScopusGoogle Scholar

74

M. Leenders, L.B. Verdijk, L. van der Hoeven, *et al.*

Prolonged leucine supplementation does not augment muscle mass or affect glycemic control in elderly type 2 diabetic men

J Nutr, 141 (2011), pp. 1070-1076
CrossRefView Record in ScopusGoogle Scholar

75

S. Verhoeven, K. Vanschoonbeek, L.B. Verdijk, *et al.*

Long-term leucine supplementation does not increase muscle mass or strength in healthy elderly men

Am J Clin Nutr, 89 (2009), pp. 1468-1475
CrossRefView Record in ScopusGoogle Scholar

76

T. Alon, D. Bagchi, H.G. Preuss

Supplementing with beta-hydroxy-beta-methylbutyrate (HMB) to build and maintain muscle mass: A review

Res Commun Mol Pathol Pharmacol, 111 (2002), pp. 139-151
View Record in ScopusGoogle Scholar

77

N.E. Deutz, S.L. Pereira, N.P. Hays, *et al.*

Effect of beta-hydroxy-beta-methylbutyrate (HMB) on lean body mass during 10 days of bed rest in older adults

Clin Nutr, 32 (2013), pp. 704-712
ArticleDownload PDFView Record in ScopusGoogle Scholar

78

P.J. Fitschen, G.J. Wilson, J.M. Wilson, K.R. Wilund

Efficacy of beta-hydroxy-beta-methylbutyrate supplementation in elderly and clinical populations

Nutrition, 29 (2013), pp. 29-36
ArticleDownload PDFView Record in ScopusGoogle Scholar

79

S.R. Kimball, L.S. Jefferson

Signaling pathways and molecular mechanisms through which branched-chain amino acids mediate translational control of protein synthesis

J Nutr, 136 (2006), pp. 227S-231S
Google Scholar

80

S. Nissen, R. Sharp, M. Ray, *et al.*

Effect of leucine metabolite beta-hydroxy-beta-methylbutyrate on muscle metabolism during resistance-exercise training

J Appl Physiol (1985), 81 (1996), pp. 2095-2104
CrossRefView Record in ScopusGoogle Scholar

81

S.M. Pasiakos, L.M. Margolis, J.S. Orr

Optimized dietary strategies to protect skeletal muscle mass during periods of unavoidable energy deficit

FASEB J, 29 (2015), pp. 1136-1142
CrossRefView Record in ScopusGoogle Scholar

82

S. Portal, A. Eliakim, D. Nemet, *et al.*

Effect of HMB supplementation on body composition, fitness, hormonal profile and muscle damage indices

J Pediatr Endocrinol Metab, 23 (2010), pp. 641-650
View Record in ScopusGoogle Scholar

83

K.A. Szczesniak, P. Ostaszewski, J.C. Fuller Jr., *et al.*

Dietary supplementation of beta-hydroxy-beta-methylbutyrate in animals—a review

J Anim Physiol Anim Nutr (Berl), 99 (2015), pp. 405-417
CrossRefView Record in ScopusGoogle Scholar

84

G.J. Wilson, J.M. Wilson, A.H. Manninen

Effects of beta-hydroxy-beta-methylbutyrate (HMB) on exercise performance and body composition across varying levels of age, sex, and training experience: A review

Nutr Metab (Lond), 5 (2008), p. 1
CrossRefView Record in ScopusGoogle Scholar

85

M. Van Koevering, S. Nissen

Oxidation of leucine and alpha-ketoisocaproate to beta-hydroxy-beta-methylbutyrate in vivo

Am J Physiol, 262 (1992), pp. E27-E31
CrossRefView Record in ScopusGoogle Scholar

86

S. Baier, D. Johannsen, N. Abumrad, *et al.*

Year-long changes in protein metabolism in elderly men and women supplemented with a nutrition cocktail of beta-hydroxy-beta-methylbutyrate (HMB), L-arginine, and L-lysine

J Parenter Enteral Nutr, 33 (2009), pp. 71-82
CrossRefView Record in ScopusGoogle Scholar

87

M.D. Giron, J.D. Vilchez, S. Shreeram, *et al.*

beta-Hydroxy-beta-methylbutyrate (HMB) normalizes dexamethasone-induced autophagy-lysosomal pathway in skeletal muscle

PLoS One, 10 (2015), p. e0117520
CrossRefView Record in ScopusGoogle Scholar

88

L.A. Burton, D. Sumukadas

Optimal management of sarcopenia

Clin Interv Aging, 5 (2010), pp. 217-228
View Record in ScopusGoogle Scholar

89

H.J. Denison, C. Cooper, A.A. Sayer, S.M. Robinson

Prevention and optimal management of sarcopenia: A review of combined exercise and nutrition interventions to improve muscle outcomes in older people

Clin Interv Aging, 10 (2015), pp. 859-869
View Record in ScopusGoogle Scholar

90

J.E. Morley

Pharmacologic options for the treatment of sarcopenia

Calcif Tissue Int, 94 (2016), pp. 319-333
CrossRefView Record in ScopusGoogle Scholar

91

B.F. Miller, J.L. Olesen, M. Hansen, *et al.*

Coordinated collagen and muscle protein synthesis in human patella tendon and quadriceps muscle after exercise

J Physiol, 567 (2005), pp. 1021-1033
CrossRefView Record in ScopusGoogle Scholar

92

E. Pasini, S. Le Douairon Lahaye, V. Flati, *et al.*

Effects of treadmill exercise and training frequency on anabolic signaling pathways in the skeletal muscle of aged rats

Exp Gerontol, 47 (2012), pp. 23-28
ArticleDownload PDFView Record in ScopusGoogle Scholar

93

T. Wenz, S.G. Rossi, R.L. Rotundo, *et al.*

Increased muscle PGC-1alpha expression protects from sarcopenia and metabolic disease during aging

Proc Natl Acad Sci U S A, 106 (2009), pp. 20405-20410
CrossRefView Record in ScopusGoogle Scholar

94

T.S. Bowen, G. Schuler, V. Adams

Skeletal muscle wasting in cachexia and sarcopenia: Molecular pathophysiology and impact of exercise training

J Cachexia Sarcopenia Muscle, 6 (2015), pp. 197-207
CrossRefView Record in ScopusGoogle Scholar

95

K. Dideriksen, S. Reitelseder, L. Holm
Influence of amino acids, dietary protein, and physical activity on muscle mass development in humans
Nutrients, 5 (2013), pp. 852-876
CrossRefView Record in ScopusGoogle Scholar

29.5: INTESTINAL FRUCTOSE AND GLUCOSE METABOLISM IN HEALTH AND DISEASE

Merino, B.; Fernández-Díaz, C.M.; Cózar-Castellano, I.; Perdomo, G. Intestinal Fructose and Glucose Metabolism in Health and Disease. *Nutrients* **2020**, *12*, 94. https://doi.org/10.3390/nu12010094

29.5.1: ABSTRACT

The worldwide epidemics of obesity and diabetes have been linked to increased sugar consumption in humans. Here, we review fructose and glucose metabolism, as well as potential molecular mechanisms by which excessive sugar consumption is associated to metabolic diseases and insulin resistance in humans. To this end, we focus on understanding molecular and cellular mechanisms of fructose and glucose transport and sensing in the intestine, the intracellular signaling effects of dietary sugar metabolism, and its impact on glucose homeostasis in health and disease. Finally, the peripheral and central effects of dietary sugars on the gut–brain axis will be reviewed.

Keywords: **fructose**; **glucose**; **small intestine**; **liver**; **brain**; **gut-brain axis**; **non-alcoholic fatty liver**; **insulin resistance**; **metabolic syndrome**; **type 2 diabetes mellitus**

29.5.2: 1. INTRODUCTION

According to the World Health Organization, obesity is the epidemic of the 21th century. About 13% of the world's adult population is obese [1]. Worldwide, between 1975–2016, the global obesity rate was nearly triplicated, increasing from 1% up to 6%–8% among children and adolescents [1]. As a major public health issue, clinical interventions based on low-fat diets attracted significant interest. However, over decades, the consumption of sugars has risen significatively worldwide, and has been partially associated to the rapid increase in the prevalence of obesity [2].

From the Industrial Revolution, the consumption of sweeteners has increased dramatically, causing a dietary switch in the world population [3]. Most of this increase in sugar consumption is derived from refined or processed fructose, which is obtained from the conversion of glucose from sugar cane and corn through a chemical process developed in 1957 [4,5]. Fructose constitutes a significant portion of the caloric intake in many countries [3]. The average daily consumption of added sugars (13%–17% of daily energy intake), of which about half is fructose, is above the recommended limit of 10% in many countries [6]. Of note, 16% of total energy in children's diets comes from added sugars [7]. The increase in total fructose intake parallels a decrease in the proportion of dietary fructose coming from fruits, but augmented proportion from fructose-based sweeteners [3]. In the past, fructose was considered sweeter, more soluble, and less gluconeogenic than glucose and sucrose, and was proposed as a substitute for these sugars [8]. Over time, this idea has been reconsidered in view of the impact of high-fructose consumption on whole-body metabolism, and because it is a risk factor for developing obesity and diabetes [9,10,11].

Although fructose and glucose share the same molecular formula ($C_6H_{12}O_6$) and caloric value (4 kcal/g), fructose tastes sweeter than glucose (relative to sucrose, which by consensus agreement is equal to one; the sweetness of glucose is 0.75, and fructose is 1.7), and has a lower glycemic index than glucose (23 versus 100, respectively) [12]. In addition, fructose is less satiating than glucose, increasing food intake [13]. As reviewed in detail below, intestinal fructose and glucose absorption are also quite different, because glucose transport is an energy-requiring process mediated by the sodium-glucose co-transporter 1 (SLGT1), whereas fructose moves through a facilitated passive transport mediated by GLUT5 [14]. Furthermore, fructose metabolism has a negligible impact on circulating insulin levels compared to glucose metabolism, which is related to insufficient leptin (the satiety hormone) secretion, and suppression of ghrelin (the hunger-promoting hormone) [15].

29.5.3: 2. INTESTINAL FRUCTOSE TRANSPORT AND METABOLISM: IMPLICATIONS FOR HEALTH AND DISEASE

Whole body fructose homeostasis results from two main processes: Intestinal absorption and clearance, the latter is commonly assumed to be mainly mediated by the liver (~55%–71%) and, to a lesser extent, by kidneys (<20%) [16]. Dietary fructose moves from the intestinal lumen to the circulation through a facilitated passive transport [3] across enterocyte membranes by members of the facilitative glucose transport (GLUT; Slc2a) family [14]. Upon its intestinal absorption, fructose reaches the liver through the hepatic portal vein and undergoes metabolization in hepatocytes [16]. Fructose transport and metabolism has been extensively reviewed (see refs. [3,14,17]). Exhaustive description of fructose hepatic metabolism is out of the scope of this review. Here, we briefly describe the regulation of intestinal fructose transport and transporters and its intracellular metabolism. We focus on revisiting the role of the liver and small intestine in fructose clearance, the relevance of endogenous fructose production in human diseases, and plant extract inhibitors of fructose transporters.

29.5.3.0.1: 2.1. INTESTINAL FRUCTOSE TRANSPORT

Fructose uptake into enterocytes is an insulin-independent process [18]. Among the members of the GLUT family able to transport fructose (GLUT5, GLUT8 and GLUT11), GLUT5 (*Slc2a5*) is primarily responsible for fructose uptake into the enterocyte at the apical side of the membrane, whereas GLUT2 (*Slc2a2*) moves most of fructose from the cytosol into blood vessels at the basolateral side of the enterocyte [14,19,20,21] (Figure 1). Although GLUT5 belongs to the GLUT family, it only transports fructose without the ability to transport glucose or galactose. Conversely, GLUT2 can transport glucose and galactose in addition to fructose, with an affinity (K_m) for fructose more than five-fold higher than that of GLUT5 [22,23].

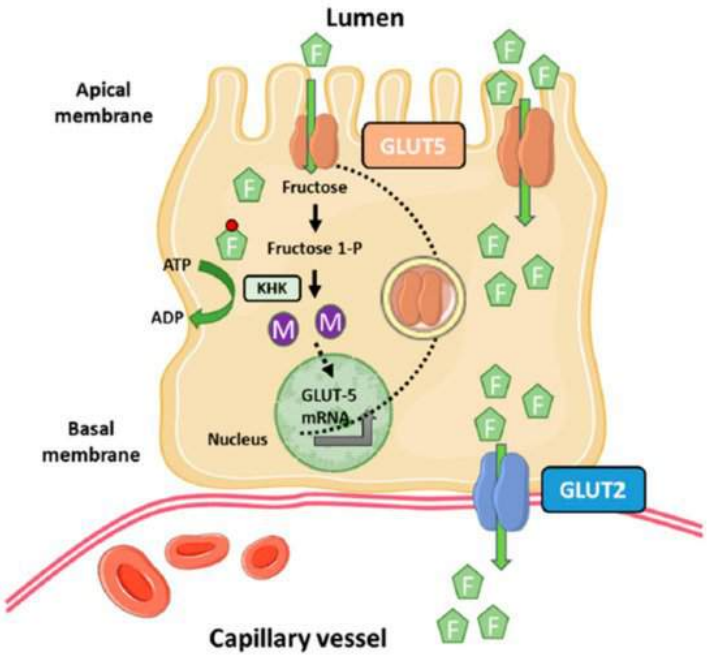


Figure 1. Fructose transport in the small intestine. Fructose (**F**) is transported by GLUT5 through the brush border membrane and enters cytosol of enterocytes, where it is rapidly phosphorylated by ketohexokinase (KHK), leading to a rapid depletion of intracellular ATP levels. A pool of phosphorylated fructose is partially or totally metabolized yielding metabolites (**M**) that induce *Slc2a5* gene expression. The remaining fructose is released across the basolateral membrane into portal circulation down the concentration gradient by GLUT2. In this conceptual model, the small intestine passively transport fructose to portal circulation and the liver is a major organ for fructose metabolism. This figure was created using Servier Medical Art (available at https://smart.servier.com/).

The main site of GLUT5 expression is the apical membrane of intestinal epithelial cells, although to a much lower extent is also expressed in kidneys, brain, fat, testes, and muscle [24]. However, the physiological relevance of GLUT5 expression in these extraintestinal human tissues is uncertain. On the other hand, GLUT2, in addition to the basolateral membranes of intestinal epithelial cells, is highly expressed in hepatocytes, pancreatic β-cells, and the basolateral membranes of kidney epithelial cells [22].

The K_m of GLUT5 for fructose varies depending on the study model and the species used for its assessment. Thus, Burant et al. reported a K_m of ~6 mM in *Xenopus* oocytes expressing the mammalian GLUT5 [19]. In contrast, Kane et al. reported a K_m of 11–15 mM using the same expression system for human GLUT5 [25]. Similar values (K_m of 11–13 mM) were found for mouse and rabbit GLUT5 transporter expressed in oocytes [26,27]. Finally, Mate et al. reported a K_m of ~8–11 mM in ileal brush border membrane vesicles of normotensive Wistar-Kyoto rats and their spontaneously hypertensive rats [28]. Assuming a K_m value ranging from 11–15 mM for GLUT5, this K_m is similar to that reported for intestinal luminal fructose concentrations (26 mM) in rats fed dietary fructose [29]. On the other hand, the K_m of GLUT2 for fructose is ~11–17 mM [22,23].

29.5.3.0.1: 2.2. DIETARY FRUCTOSE METABOLISM

High concentrations of dietary fructose in foods and drinks lead to elevated intestinal luminal fructose concentrations that are needed for driving the facilitated fructose transport across the enterocyte membrane, and fluctuate around the K_m of GLUT5 for fructose [3]. Unlike the high fructose concentration in luminal small intestine, fructose concentrations in systemic circulation are relatively low as a result of intestinal absorption and liver clearance rates. In humans, estimates of fasting systemic blood fructose concentrations are low (<0.05 mM), even in those healthy humans consuming high-fructose or sucrose diets (~0.2–0.5 mM) [30,31,32,33], which is still very low compared to fasting blood glucose levels (5.5 mM). Finally, type 1 and type 2 diabetic patients exhibited 0.016 mM and 0.009–0.013 mM fasting fructose concentrations, respectively [34,35]. The low fructose concentrations in peripheral blood support the notion that the liver and kidneys are much more sensitive to small changes in circulating fructose levels than the small intestine. Nonetheless, it is unclear how hepatocytes or nephrons reabsorb fructose from the sinusoidal capillaries or glomerular filtrates, respectively, containing very low fructose levels.

Metabolization of dietary fructose in the small intestine is a process regulated at various steps. In the first step of the classical Hers pathway for fructose metabolism, fructose is mobilized from intestinal lumen into the cytosol of enterocytes by GLUT5, where it is rapidly phosphorylated by the ketohexokinase (KHK, *Khk*), also known as fructokinase, to fructose-1-phosphate using ATP as a phosphate donor [36]. The *Khk* gene encodes two isoforms of the enzyme as a result of alternative splicing of the adjacent exons 3A and 3C of the gene leading to the KHK-A and KHK-C isoforms, respectively [37,38]. Studies of expression analysis in several human and rat tissues indicated that only one mRNA variant is expressed in each tissue [38], but the pancreas is an exception to this pattern because, although KHK-C expression predominates, some KHK-A is also expressed. The KHK-C mRNA variant is expressed at high levels in the liver, kidneys, and duodenum, and is considered more physiological than the KHK-A variant because its K_m for fructose is <1 mM [39,40]. On the other hand, KHK-A is expressed at a low level in a wide range of tissues including skeletal muscle and adipose tissue [39,40]. KHK-A K_m for fructose is 8 mM, suggesting that it poorly phosphorylates fructose at physiological concentrations and that it may have a more important role when fructose intake is excessive [41].

In the second step of the fructose pathway, fructose-1-phosphate is split into glyceraldehyde and dihydroxyacetone phosphate by aldolase B (ALDOB; *Aldob*) [16]. In the third and final step of the pathway, the triokinase (TKFC; ATP:D-glyceraldehyde 3-phosphotransferase) catalyzes the phosphorylation of glyceraldehyde by ATP to form glyceraldehyde-3-phosphate [16,36]. Both ALDOB and TKFC are highly expressed in the liver, kidneys, and small intestine, relative to other organs [42].

Unlike glycolysis, the catabolism of fructose (fructolysis) bypasses major regulatory steps of glycolysis and gluconeogenesis (i.e., phosphofructokinase and fructose-1,6-bisphosphatase), and it is not regulated by feedback inhibition [16,43]. In addition, fructolysis bypasses the glucose-6-phosphate and fructose-6-phosphate production from the pentose phosphate pathway for de novo synthesis of nucleotides and nucleic acids [44]. Thus, it is plausible that in conditions of excessive fructose consumption, KHK-mediated fructolysis leads to increased glyceraldehyde, dihydroxy-acetone-phosphate, and glyceraldehyde-3-phosphate production, which are the source of gluconeogenic and lipogenic substrates (e.g., pyruvate, lactate, acetyl-CoA, and glycerol-3-phosphate), leading to elevated rates of gluconeogenesis, glycogenesis, and/or lipogenesis. Another consequence of the fructolysis is a rapid ATP and Pi intracellular depletion [45,46].

29.5.3.0.1: 2.3. REGULATION OF GLUT5

In addition to dietary fructose catabolism, metabolism of fructose comprises its biosynthesis from glucose through the polyol pathway [16]. This two-step pathway becomes active when intracellular glucose concentrations are elevated. In the first step, glucose undergoes reduction by NADPH to sorbitol (polyol) by the rate-limiting enzyme in the pathway, the aldose reductase (AR), followed by metabolization of sorbitol into fructose by sorbitol dehydrogenase (SDH) in the presence of NAD+ as a cofactor [16].

Intestinal fructose metabolism is not only important for the metabolic fate of fructose but for the up-regulation of GLUT5, KHK, ALDOB, TKFC, fructose-1,6-bisphosphatease, and glucose-6-phosphate (**Figure 1**) [47,48]. Thus, it has been extensively shown that chronic or acute fructose exposure increases GLUT5 levels and activity in rodents and human proximal intestine regions [3,17]. The response of GLUT5 to its substrate requires partial or total metabolization of fructose because the nonmetabolizable fructose analog 3-*O*-methylfructose has a modest effect on GLUT5 expression [49], and blocking intracellular fructose metabolism in the $HKH^{-/-}$ mouse model prevents fructose up-regulation of GLUT5 [47]. Furthermore, these effects of fructose on GLUT5 expression are very specific because fructose, glucose, and nonmetabolizable glucose analogs have similar changes on GLUT2 expression in intestinal cells [49]. The molecular mechanisms underlying fructose-mediated regulation of GLUT5 in enterocytes remain incompletely understood. In rats, fructose-induced cAMP stimulates fructose uptake without affecting transcriptional regulation of *Slc2a5* [50], whereas in human Caco-2 cells, fructose increases *Slc2a5* mRNA stability mediated by the cAMP pathway [51]. On the other hand, the use of inhibitors or activators of the phosphatidylinositol 3-kinase (PI3K) and/or protein kinase B (PKB) have demonstrated that this signaling pathway mediates the fructose-induced increase in fructose transport without affecting transcriptional regulation of GLUT5 [52]. How does the PI3K/AKT signaling pathway mediate the effects of fructose on GLUT5 upregulation? It is known that Class II PI3Ks control the endocytic trafficking of transporters through the production of phosphatidylinositol 3-phosphate (PtdIns3*P*). This second messenger is required for Rab11 activation, a small GTPase of the Rab family that coordinates endosome recycling to the plasma membrane [53]. Enterocyte-specific $Rab11a^{\Delta IEC}$ ablation (Rab11a-KO mouse model) blunted fructose-induced upregulation of GLUT5 in the small intestine, most likely by impairing endosomal trafficking of the fructose transporter towards the apical membrane of the enterocyte [47].

The expression of GLUT5 in the intestine can also be regulated by the carbohydrate response element-binding protein (ChREBP), a liver glucose-responsive basic helix-loop-helix-leucine zipper transcriptional factor [54]. High fructose diet feeding increases intestinal ChREBP protein levels, accompanied by increased fructose transport (GLUT5), fructolytic (fructokinase, ALDOB, TKFC, and lactate dehydrogenase) and gluconeogenic (glucose-6-phosphatae and fructose-1,6-bisphosphatase) gene expression in mice [55]. Conversely, genetic ablation of ChREBP (ChREBP-KO mice) leads to fructose intolerance due to insufficient induction of these genes involved in fructose transport and metabolism [55,56,57,58,59]. The molecular mechanism by which fructose mediates ChREBP-induction of *Slc2a5* gene expression involves direct interaction of ChREBP with the promoter of *Slc2a5* [55] in mice, whereas ectopic co-expression of ChREBP and its heterodimer partner Max-like protein X (MLX) binds to carbohydrate response elements (ChoREs) and activates *Slc2a5* promoter in Caco-2BBE human cells [55]. Further work is required to confirm whether, similarly to glucose, fructose might regulate

ChREBP activity by posttranslational modifications such as *O*-glycosylation, phosphorylation and conformational changes in intestinal cells [57].

Another identified regulatory protein of intestinal fructose transport is the thioredoxin-interacting protein (TXNIP, *Txnip*), an arrestin-like protein that can bind to thioredoxin protein that regulates cellular metabolism and redox state [60,61]. In response to glucose, the transcriptional complex ChREBP/MLX and MondoA/MLX binds to the ChoRE on the *Txnip* promoter to induce mRNA expression [62,63]. Glucose-induced TXNIP inhibits glucose transport through interaction with GLUT1 and inducing its internalization through clathrin-coated pits, as well as reducing the expression of GLUT1, whereas energy stress leads to TXNIP degradation through phosphorylation by AMP-dependent protein kinase (AMPK), resulting in increased GLUT1 function and mRNA expression [61,64]. Dotimas et al. demonstrated that TXNIP regulates fructose absorption in the small intestine [65]. Although the precise mechanisms remains elusive, TXNIP is upregulated in response to fructose consumption and co-immunoprecipitates with GLUT2 and GLUT5. It may be possible that the link between fructose transport and TXNIP may be mediated by phosphorylation of the protein mediated by AMPK, similar to what we described above for GLUT1 [65].

The expression of GLUT5 and its activity is also regulated by early development in the intestine of mammalians (i.e., rat, rabbit, and humans). In rats, under normal conditions (suckling and weaning), intestinal fructose transport and GLUT5 mRNA levels are very low due to the fact that maternal milk is fructose-free, unless there is a precocious exposure to luminal intestine fructose signal, which in turn stimulates GLUT5 expression and activity [17]. The mechanism by which fructose increases GLUT5 expression and activity during weaning is complex and involves systemic levels of glucocorticoids, but not thyroxine [17,66,67,68]. In addition, the diurnal rhythm regulates GLUT5 mRNA and protein expression in adult rats, but this regulation is not present in neonates [69]. Independently of fructose uptake, 3–4 h before the onset of peak feeding, GLUT5 levels increase by four-fold. This diurnal rhythm is also accompanied by upregulation of GLUT2 [8].

29.5.3.0.1: 2.4. FRUCTOSE METABOLISM IN HUMAN DISEASES

Major pathways of fructose metabolism are conversion to glucose and lipids [16]. Therefore, excessive fructose intake would result in increased portal fructose concentrations that stimulates endogenous glucose production and lipid synthesis in the liver, which is associated with metabolic syndrome (MetS) [70,71,72], non-alcoholic fatty liver disease (NAFLD) [73,74,75,76,77,78,79,80,81], obesity, and type 2 diabetes mellitus (T2DM) [9,10,11,82,83,84,85,86,87,88]. Although there is mounting epidemiological and experimental evidence linking fructose consumption to metabolic diseases, the relative contribution of fructose to these human diseases remains controversial [87,89,90,91].

29.5.3.0.1: 2.5. REVISITING THE ROLE OF LIVER AND SMALL INTESTINE IN FRUCTOSE CLEARANCE

Traditionally, the liver has been considered as the main organ that metabolizes fructose before entering systemic circulation [16]. This assumption is based on the following evidences: (1) Intestinal absorption of fructose is primarily driven to the liver through portal circulation; (2) peripheral tissues, such as skeletal muscle, have low capacity for fructose metabolism; and (3) the ketohexokinase isoform KHK-C is expressed at the highest level in the liver relative to extrahepatic tissues, leading to a high capacity for fructose phosphorylation and extraction from the blood. In this way, the liver would prevent high fructose doses to spill over peripheral tissues [16].

The current notion that the liver is the main site of dietary fructose metabolism and clearance has been recently challenged by Jang et al. [92]. They used sophisticated and elegant isotopic tracing techniques and arterio-venous blood sampling to demonstrate that most ingested fructose is metabolized by the small intestine in mice. At low-doses of fructose (<0.5 g kg^{-1}), ~90% of fructose phosphorylation occurs in the jejunum, duodenum, or ileum. Most of this fructose is metabolized in the small intestine, appearing in the portal circulation as glucose and lactate (~60%), and the remaining as fructose (<20%). In contrast, high-doses of fructose (≥1 g kg^{-1}) saturate the absorption and catabolism of fructose in the small intestine, leading to fructose spill-over into the liver (>30%) and the colonic microbiota in mice [92] (**Figure 2**). This work challenges our current knowledge about the role of the small intestine in dietary fructose metabolism and spurs the notion that the small intestine shields the liver from toxic fructose exposure. However, several questions arise from this work and remain to be fully addressed: (1) A limitation of the study is regarding the dose-response to fructose, which may vary between mice and humans. Humans may saturate the capacity for fructose metabolism in the small intestine at relatively lower doses than mice. It is necessary to understand the associated dose-response pattern in humans. (2) The role of the small intestine in fructose metabolism in mice and humans may have diverged across evolution. In fact, humans have a relative shorter gut and smaller intestinal area than rodents [93]. (3) The long standing view is that the liver and kidneys are the only gluconeogenic organs in humans, but not the small intestine because it does not express glucose-6-phosphatase (G-6-Pase) [16]. This critical issue is important to translate experimental evidences from mice to humans. In this line, one study have shown the expression of G-6-Pase in the small intestine of humans [94], and another one showed some evidence of the existence of a conversion of fructose to glucose in human jejunum [95].

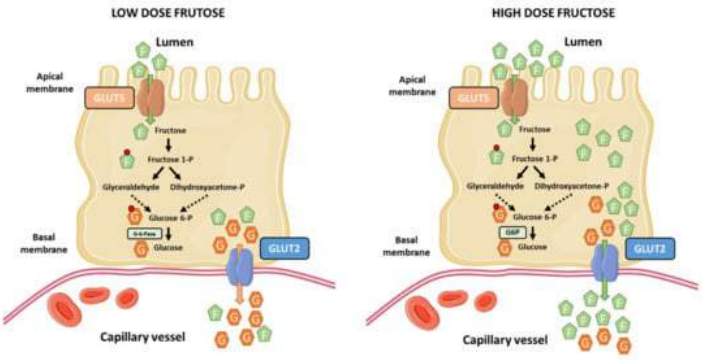


Figure 2. Current view of intestinal fructose metabolism in mice. Under conditions of low-dose dietary fructose (**F**) consumption, most of the fructose is metabolized by the Hers pathway appearing into portal circulation mainly as glucose (**G**), and the remaining as unmetabolized fructose. Conversely, under high-dose dietary fructose consumption, intestinal capacity is overwhelmed, leading to most of the fructose to spill over to the liver. Further work is needed to validate or refute this model in humans. G-6-Pase; glucose-6-phosphatase. This figure was created using Servier Medical Art.

29.5.3.0.1: 2.6. RELEVANCE OF ENDOGENOUS FRUCTOSE PRODUCTION IN HUMAN DISEASES

In addition to exogenous fructose, fructose can be synthesized from glucose through the polyol pathway [16,96], which has drawn attention on the potential role of endogenous fructose production in metabolic diseases.

The biosynthetic fructose pathway is constituted by two enzymes; the aldose reductase that converts glucose into sorbitol, and the sorbitol dehydrogenase that converts sorbitol into fructose [16]. Under physiological conditions, this pathway is mostly inactive in the majority of body tissues and organs, which has been associated to lower fasting and postprandial circulating fructose levels [97]. However, this pathway can be activated after ingestion of a drink containing glucose (~30 g) and fructose (~30 g) in healthy individuals. Tracer dilution analysis estimated endogenous fructose production ~ 55 mug kg^{-1}·min^{-1}. This work evidenced, for the first time, the capacity for endogenous fructose production in humans [97]. Further research demonstrated the presence of an active polyol pathway in tissues other than those involved in metabolizing dietary fructose, such as the human brain [98,99,100]. Numerous studies using animal models have linked the polyol pathway to metabolic alterations such as obesity, insulin resistance, diabetes, diabetic nephropathy, chronic kidney disease, acute kidney injury, blood pressure, and MetS [101,102,103,104]. Nonetheless, although the presence of an active polyol pathway has been described in humans, and mounting evidences obtained in animal models of the importance of this pathway in diseases, its significance in human metabolic diseases awaits further confirmation.

29.5.3.0.1: 2.7. PLANT EXTRACTS INHIBITORS OF FRUCTOSE TRANSPORTERS

As described above, multiple studies in humans and animal models have linked fructose consumption with diseases, which has spurred the notion of the potential use of GLUT5 inhibitors for preventing fructose-induced diseases. So far, no potent and specific inhibitors of GLUT5 have been discovered, although phloretin and cytochalasin B are used to inhibit GLUT2 for assessing fructose transport in vitro, whereas GLUT5 is insensitive to both inhibitors [22,25].

In the last decade, plant extracts have been used to screen compounds with inhibitory effects on intestinal GLUT5 transporters. Thus, green tea catechins inhibited D-fructose transport in *Xenopus laevis* oocytes expressing the mammalian GLUT5. Inhibition of D-fructose transport via GLUT5 was more efficient by catechins containing a gallate group [apparent *Ki* values between ~113 and ~117 μM for (−)-epigallocatechin-gallate and (−)-epicatechin-gallate, respectively] than by catechins lacking this group [apparent *Ki* values >500 μM for (−)-epicatechin and (−)-epigallocatechin] [105]. In this line of evidence, it has been shown that chamomile tea and green tea [containing (−)-epigallocatechin gallate (240 mg/g extract), (−)-epigallocatechin (70 mg/g extract), (−)-epicatechin (40 mg/g extract), and (+)-catechin (17 mg/g extract)] effectively inhibited fructose transport through GLUT2 in differentiated Caco-2 cells [106]. In addition, chamomile also inhibits D-fructose transport via GLUT5 in Caco-2 cells and in *Xenopus* oocytes expressing the mammalian GLUT5 [106]. Likewise, Satsu et al. demonstrated that epicatechin gallate inhibited fructose uptake in Caco-2 cells. Interestingly, this reduction in fructose uptake was not related to changes in the affinity (K_m) of GLUT5 for fructose, but with a decrease in the maximal velocity (V_{max}) [107]. Furthermore, authors demonstrated that epicatechin gallate suppressed fructose permeation in Caco-2 cells, suggesting that this compound suppressed the transepithelial transport of fructose across epithelial cell monolayers, in addition to its effect on fructose uptake. Lastly, authors reported that similar effects on fructose uptake and permeation were observed with nobiletin, another phytochemical tested in this study [107].

An additional compound extracted from the Chinese blackberry tea (rubusoside) inhibited GLUT5-mediated fructose transport in liposomes reconstituted with human GLUT5 purified from insect cells transduced with baculoviruses [18]. Likewise, astragalin-6-glucoside (a glycosylated derivative of astragalin) inhibited GLUT5-mediated fructose transport in these proteoliposomes [18]. The same group performed a virtual screening (in silico) for potential GLUT5 inhibitors using a 3D inward-facing GLUT5 model against a library of >600,000 chemicals [108]. The ability of the top ranked compounds for inhibiting GLUT5-mediated fructose transport were tested in GLUT5 proteoliposomes, identifying the N-[4-(methylsulfonyl)-2-nitrophenyl]-1,3-benzodioxol-5-amine (MSNBA) as an specific inhibitor, which did not affect the fructose transport of human GLUT2 or the glucose transport of human GLUT1-4 [108]. Additionally, whole-cell

systems for high-throughput screening of potential GLUT5 inhibitors and activators have been developed using a yeast strain deficient in fructose uptake [109].

The ability of culinary plant extracts containing phytochemicals to inhibit fructose transport has also been assessed in Caco-2 cells. Lee et al. found that demethoxycurcumin and curcumin from turmeric extracts inhibited fructose transport by GLUT2- and GLUT5-mediated fructose uptake, respectively [110]. Similarly, catechin from guava leaf (*Psidium guajava*) inhibited GLUT5-mediated fructose uptake, whereas quercetin inhibited both GLUT5- and GLUT2-mediated fructose transport [110]. In addition, the ability of guava leaf and guava fruit extracts to inhibit glucose transport have also been demonstrated by Müller et al. in Caco-2 cells and mice (C57BL/6N) [111]. The effect of these extracts on glucose uptake in Caco-2 cells were related to inhibition of GLUT2, although the effects on fructose uptake were not assessed [111]. More recently, König et al. demonstrated that fruit extracts prepared from guava inhibited intestinal glucose resorption in a clinical trial [112].

The effects of hesperidin, a flavonoid present in orange juice, on fructose uptake in Caco-2 cell monolayers was studied by Kerimi et al. [113]. They showed that hesperidin inhibited fructose uptake in these cells using fructose (130 mM) as the only source of sugars. Of note, the inhibitory effect of hesperidin on fructose uptake was abolished in the presence of other sugars, such as glucose and sucrose, at high concentrations (120 mM and 130 mM, respectively). Using *Xenopus laevis* oocytes expressing human GLUT2 or GLUT5, they gained insights into the molecular mechanisms by which hesperidin inhibited fructose transport. Thus, hesperidin inhibited the uptake of fructose by GLUT5 expressed in *Xenopus* oocytes. In addition to its effects on fructose uptake, hesperidin lowered glucose uptake in Caco-2 cells and inhibited GLUT2 and GLUT5 transporters when expressed in *Xenopus* oocytes. Lastly, in an attempt to reproduce in vivo these previously observed effects of hesperidin, authors conducted three separated human intervention studies on healthy volunteers using orange juice with different amounts of added hesperidin and a control drink containing equivalent amounts of glucose, fructose, and sucrose, and measured the postprandial glycemic response as biomarker for the effect of hesperidin. They observed that the biggest difference in postprandial blood glucose between orange juice and the control drink was when the juice was diluted [113]. The inhibitory effects of other flavonoids, such as apigenin, on fructose uptake have also been investigated by Gauer et al. in *Xenopus* oocytes. Apigenin, as well as (−)-epigallocatechin gallate, inhibited fructose uptake in oocytes expressing GLUT5 [114].

Finally, acarbose, an α-glucosidase inhibitor that improves insulin sensitivity and decreases postprandial hyperglycemia [115], does not inhibit fructose transport in human Caco-2 cells or in *Xenopus* oocytes expressing the mammalian GLUT2 and GLUT5 [106]. These results suggest that the effects of acarbose on fructose absorption would be mediated by its well-known effects on attenuating sucrose digestion [116], rather than direct effects on fructose transport across the intestinal epithelium.

29.5.4: 3. INTESTINAL GLUCOSE TRANSPORT AND METABOLISM: IMPLICATIONS FOR HEALTH AND DISEASE

Glucose is the main catabolic and anabolic substrate for the great majority of complex organs that controls energy homeostasis in the body. Glucose homeostasis is the result of three physiological events: Intestinal glucose absorption in the post-prandial state, hepatic glucose production (which accounts for ~90% of endogenous glucose production and is the net balance between gluconeogenesis, glycogenolysis, glycogen synthesis, glycolysis, and other pathways), and extrahepatic glucose usage, mainly by the brain, the skeletal muscle, and the adipose tissue. Glucose controls hormonal secretion in endocrine pancreas (i.e., insulin, glucagon, and somatostatin) [117,118,119] and neuronal signaling involved in glucose homeostasis, feeding regulation, and energy expenditure [120].

29.5.4.0.1: 3.1. INTESTINAL GLUCOSE TRANSPORT.

Gastric emptying and intestinal glucose absorption determine the glucose appearance rate in the bloodstream after a meal. Intestinal enterocytes are polarized cells responsible for glucose uptake from the intestinal lumen to capillary blood vessels, which is the main mechanism of glucose entrance into the body. Enterocytes express two glucose transporters named sodium-glucose co-transporter 1 (SGLT1; expressed in the brush border membrane) and GLUT2 (localized in the basolateral membrane). SGLT1 couples the transport of one glucose molecule and two sodium ions, which provides the energy to drive glucose accumulation in the enterocyte against its concentration gradient due to the energy stored in the sodium electrochemical potential gradient across the brush border membrane generated by the sodium transport. Sodium is then transported out into the blood vessels by the Na^+/K^+-ATPase in the basolateral membrane, maintaining the driving force to transport glucose. As a result, glucose accumulates within the enterocyte and diffuse out of the cell through GLUT2 into the blood stream. This process is ATP-dependent [121,122] (**Figure 3**).

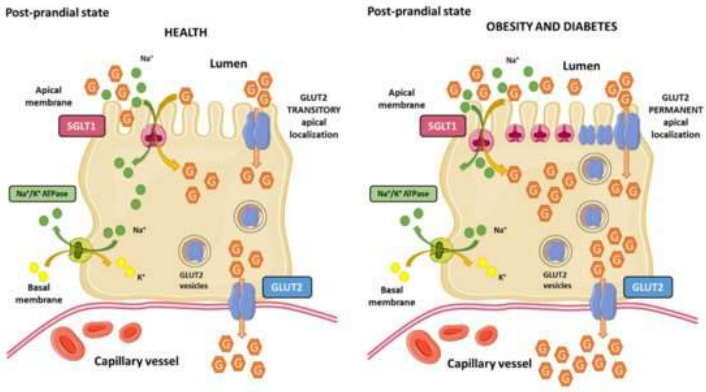


Figure 3. Glucose transport in enterocytes in health and disease. (**Left panel**): After a meal, luminal glucose (**G**) is transported across the apical membrane by SGLT1, and the Na^+ is then transported out the enterocyte through the basolateral membrane by the Na^+/K^+-ATPase. Glucose is phosphorylated and accumulates within the cell. Dephosphorylated glucose passively is transported out of the cell through the basolateral membrane by GLUT2. Alternatively, in response to high glucose luminal concentrations, a pool of endosomal GLUT2 is rapidly and transiently translocated to the apical membrane leading to increased glucose uptake. (**Right panel**): In the setting of obesity and/or diabetes, insulin resistance provokes the loss of GLUT2 trafficking control leading to a permanent localization of GLUT2 in the apical and/or endosomal enterocyte membranes and to increased transepithelial glucose transport from lumen to blood circulation. This figure was created using Servier Medical Art.

Intestinal SGLT1 is a high-affinity (K_m ~0.4 mM), low-capacity transporter able to transport glucose or galactose. It is a monomeric integral membrane protein embedded in the lipid bilayer composed by 664 amino acids with 14 transmembrane-spanning regions and it has one glucose binding-site and two sodium binding-sites in the center of the protein. In humans, SGLT1 is encoded by the *Slc5a1* gene, and it is highly expressed in the duodenum and skeletal muscle [123,124]. SGLT1 activity varies diurnally to meet fluctuating availability of glucose. The maximal transport capacity occurs when food is anticipated, and it could be regulated by clock genes [125,126].

After energy-dependent glucose uptake via SGLT1, glucose exits the enterocyte passively through GLUT2 located in basolateral membrane. Intestinal GLUT2 is a facilitative glucose uniporter with low glucose affinity (K_m ~17 mM), but high transport capacity, located in basolateral membrane of the enterocytes. GLUT2 can also transport galactose, mannose, and fructose (with low affinity), and glucosamine with high affinity (K_m ~ 0.8 mM) [127].

29.5.4.0.1: 3.2. REGULATION OF SGLT1 AND GLUT2

SGLT1 expression in the intestinal lumen is regulated by dietary carbohydrate content. Thus, luminal glucose, but not intravenous administration of glucose, increases intestinal SGLT1 expression. High-diet glucose feeding increases SGLT1 expression and activity (rat, mouse, and sheep), which is accompanied by increased glucose transport. Similarly, obese mice exhibit increased intestinal glucose transport mediated by augmented SGLT1 transporters, without increased activity [128,129,130,131,132].

In addition to glucose-mediated regulation of SGLT1, phosphorylation by protein kinase A (PKA) and protein kinase C (PKC) regulates its activity. In humans, SGLT1 contains one consensus site for regulation by PKA and five consensus sites for PKC. The number of consensus sites and conserved sequences varies between species (rat, rabbit, and humans) [133,134]. In Chinese hamster ovary (CHO) cells overexpressing human SGLT1, activation of PKA increased the amount of SGLT1 in the membrane [135]. In contrast, stimulation of human embryonic kidney cells expressing human SGLT1 with 8-Br-cAMP (a brominated derivative of cyclic adenosine monophosphate that activates cAMP-dependent protein kinase) significantly reduced glucose transport [136]. On the other hand, PKC activation in the absence of RS1 increases transport capacity of human SGLT1, while in the presence of RS1, glucose transport is decreased [137].

The adipocyte-derived hormone leptin also regulates SGLT1. Although leptin is not required for intestinal SGLT1 expression, hyperleptinemia or leptin administration drastically reduce intestinal SGLT1 expression. The intracellular signaling pathways by which leptin regulates intestinal SGLT1 remain incompletely understood, but may include PKA, PKC, and the leptin receptor isoform b [138,139]. Finally, as in the case of GLUT5, green tea catechins markedly inhibit SGLT1-mediated glucose transport in the small intestine, being more pronounced by catechins containing a gallate group [(−)-epigallocatechin-gallate and (−)-epicatechin-gallate] than by catechins lacking this group [140].

The classical view of intestinal glucose absorption is underlined by the evidence that SGLT1 is in the apical membrane of enterocytes, while GLUT2 is located exclusively in the basolateral membrane, leading to the transepithelial glucose transport from the lumen into the portal circulation. This classical theory explains glucose absorption at low luminal glucose concentrations (≤10 mM) but it fails to explain the marked increase at glucose concentrations that surpass SGLT1 (≥25 mM) transport capacity. GLUT2 levels are also regulated by glucose concentrations in enterocytes. As part of an adaptive physiological mechanism in response to increased luminal glucose concentrations, caloric demand, and glucagon-like peptide 2 (GLP-2); GLUT2 is rapidly and transiently recruited to the brush border membrane of the enterocyte, leading to a three-fold enhancement of glucose transport [141,142] (Figure 3). This adaptive mechanism that is known as the

"GLUT2 translocation" theory, which in addition to other theories, such as the "solvent drag" theory, have been proposed to explain the marked increase in glucose absorption in response to high luminal glucose concentrations [143].

Conversely, it has also been demonstrated that in addition to high luminal glucose concentrations, insulin decreases GLUT2 membrane levels as a result of the internalization of GLUT2 from plasma membranes back into intracellular pools, leading to the inhibition of glucose transport [144]. The regulation of intestinal glucose absorption by insulin is probably another physiological mechanism at the enterocyte level by which the hormone limits sugar excursions in the blood circulation during a sugar-rich meal. This evidence raised the idea that insulin resistance may provoke a loss of insulin-mediated control of GLUT2 membrane trafficking, leading to unleash intestinal glucose absorption upon high-sugar diets consumption. Tobin et al. demonstrated that insulin resistance in mice provoked a loss of GLUT2 trafficking control, where GLUT2 levels remain permanently elevated in the brush border membrane and low in the basolateral membrane of the enterocyte [144]. Ait-Omar et al. investigated the relevance of these previously described mechanisms in the small intestine of morbidly obese insulin resistant humans and lean control subjects. They found that GLUT2 was accumulated in apical and/or endosomal membranes of enterocytes in obese subjects. Interpretation of these findings is complex, but authors proposed that permanent apical GLUT2 localization in obese subjects would mediate blood-to-lumen glucose flux during fasting hyperglycemia, leading to glucose secretion into the intestinal lumen. In contrast, after consumption of a sugar-rich meal, permanent apical GLUT2 localization would provide a large glucose uptake from the intestinal lumen to the portal circulation [145].

29.5.4.0.1: 3.3. INTESTINAL GLUCOSE METABOLISM IN HUMAN DISEASES

29.5.4.0.1: 3.3.1. RELEVANCE OF GLYCEMIC INDEX AND GLYCEMIC LOAD FOR T2DM

The glycemic response (GR) is the appearance of glucose in blood after a meal. It depends on the amount of glucose absorbed, the rate of glucose entry into circulation, the rate of disappearance due to tissue uptake from circulation, and the regulation of hepatic glucose production [146]. Blood glucose concentrations will rise and fall rapidly or slowly depending on the carbohydrate content of food. The glycemic index (GI) is a tool developed to compare the postprandial responses to constants amounts of different carbohydrate-containing food. It is a useful tool for people with diabetes, providing information on the GR that might be expected when a person consumes the quantity of a food containing a fixed amount of carbohydrates [147]. The glycemic load (GL) concept was introduced as a mean of predicting the GR, considering the GI and the amount of available carbohydrate in a portion of the food eaten [148]. Thus, foods have been classified by GI into low (GI ≤ 55), medium (GI 56–69), and high (GI ≥ 70) categories, and classified by GL as being low (GL ≤ 10), medium (GL 11–19), and high (GL ≥ 20). Since these concepts were introduced, numerous studies have been performed to ascertain how GI and GL relate to health and disease. Of note, the American Diabetes Association (ADA) indicated that current knowledge is insufficient to relate low–GL diet with a reduction on diabetes risk, and that it has not been demonstrated that one method of assessing the relationship between carbohydrate intake and blood glucose response is better than other methods [149].

To shed light into this issue, Livesey et al. performed a review meta-analysis of prospective cohort studies for a comprehensive examination of evidence on the dose-response that links GL to T2DM. The analysis concluded that a GL over a dose range of 100 g/2000 kcal, increases the risk of T2DM by 45%, supporting the notion that GL is an important and underestimated dietary characteristic that contributes to the incidence of T2DM [150]. Greenwood et al., in a systematic review and dose-response meta-analysis of prospective studies, showed that there is a protective effect of low dietary GI and GL and risk of T2DM [151]. In addition, two previous systematic reviews concluded that there is evidence of a positive association between both dietary GI and GL and risk of T2DM [152,153]. In summary, despite the fact that epidemiological studies of GI and GL in relation to diabetes risk have yielded inconsistent results, there is important research in support of significantly positive associations between dietary GI and GL and the risk of T2DM, thus reducing the intake of high-GI foods may bring benefits in diabetes prevention.

29.5.4.0.1: 3.3.2. REGULATION OF SGLT1 IN DIABETES MELLITUS

Several studies in rodent models of T2DM and type 1 diabetes mellitus (T1DM) have shown a link between intestinal SGLT1 expression and diabetes. Streptozotozin (STZ)-induced diabetes in mice and rats (STZ; a toxic drug that produces a destruction of pancreatic β-cells causing insulin deficiency and hyperglycemia [154]) produces increased SGLT1 intestinal expression [155]. Likewise, a rat model of T2DM (Otsuka Log-Evans Tokushima Fatty rats) exhibited increased intestinal mRNA expression of SGLT1 associated with impaired glucose tolerance and occurred before the onset of insulin resistance and hyperinsulinemia [156]. Similar results were confirmed in patients with noninsulin-dependent diabetes mellitus where mRNA and protein levels were increased three- to four-fold in brush border membranes of enterocytes in the small intestine [157]. Finally, in morbid obese non-diabetic patients, increased SGLT1 expression in the intestine was found and it correlated with accelerated intestinal absorption [158].

Taken together, these findings are consistent with the concept that SGLT1-mediated glucose absorption in the intestine underlies the rapid post-prandial rise in blood glucose levels observed in obesity and T2DM. This knowledge has prompted the concept that pharmacological inhibition of SGLT1 in the small intestine can lower hyperglycemia by inhibiting glucose absorption and increasing GLP-1. The pharmacological tools that have been used to determine the potential of SGLT1 inhibition include phlorizin (or phloridzin), canagliflozin, LX4211 (or sotagliflozin), LP-925219, KGA-2727, and GSK-1614235 [159]. The use of these inhibitors in rodent models of T2DM and in humans has lent support to this pharmacological approach in the treatment of T2DM, but more studies are needed on long-term safety of SGLT1 inhibition.

29.5.5: 4. PERIPHERAL AND CENTRAL EFFECTS OF DIETARY SUGARS IN THE GUT–BRAIN AXIS IN HEALTH AND DISEASE

29.5.5.0.1: 4.1. THE GUT–BRAIN AXIS

In the early 20th century, Ivan Pavlov discovered the existence of a close interaction between the gut and the brain. Pavlov observed in dogs how a stimulus associated with feeding induced vagal-dependent gastric acid secretion [160]. Since then, this interaction has been widely described and is enclosed in the term of “gut–brain axis”, a complex bidirectional communication system that maintains constant crosstalk between the gastrointestinal system and the enteric and central nervous system. This intimate connection involves numerous endocrine, immune, and neuronal pathways [161]. Through this complex system, the gut can send modulating signals to the brain via visceral messages that influence emotional and cognitive brain centers producing different psychobehavioural responses [161]. In the other direction, the brain is able to send orders for proper maintenance of gastrointestinal homeostasis (such as by modulating intestinal motility and mucin production) and can also modulate the immune system (such as by modulating cytokine production by mucosal cells) [161].

The gut–brain axis uses mostly four major information carriers to communicate with each other: Neural messages via vagal and spinal afferent neurons, immune mediators carried by cytokines, endocrine signals carried by gut hormones, and microbiota-related factors that reach the brain directly from the blood stream [162,163]. The integration of all these signals allows the maintenance of a large number of vital functions such as the control of food intake and satiety, the repulsion of harmful foods, and the adaptation of our gastrointestinal system to the environment, giving rise in pathological conditions to the sensation of nausea, pain, or even may result in gastrointestinal dysfunction [164,165].

29.5.5.0.1: 4.2. REGULATION OF THE GUT–BRAIN AXIS BY ENTEROENDOCRINE CELLS AND SENSING OF INTESTINAL SUGARS

Enteroendocrine cells (EECs) form the largest endocrine organ in the body and play a key role in regulating nutrients intake and postprandial metabolism. Following a meal, EECs in the small intestine sense luminal and circulating levels of nutrients, and simultaneously are stimulated by prevailing nutrients through multiple nutrient transporters and G protein-coupled receptors (GPCRs), leading to activation of intracellular signaling pathways that produce secretion of peptides and hormones. These hormones enter circulation and act on multiple distant tissues such as the brain, gallbladder, and pancreas, as well as, on neighboring enteric neurons, endothelial cells, and the gastrointestinal epithelium. Thus, the physiological role of the enteroendocrine system in response to ingested glucose and fructose is to detect nutrients in the intestinal lumen, to monitor energy status of the body, and to elaborate an appropriate response, through the production of more than 30 different hormones and neurotransmitters to control postprandial whole-body metabolic homeostasis [166].

EECs are endoderm-derived epithelial cells widely distributed in the villi and crypts, where they are interspersed between non endocrine cells [166]. The intestinal epithelium is in a constant turnover that is replenished from pluripotent stem cells at the base of intestinal crypts and their progenies migrating up the crypt–villus axis [167]. The spatial distribution and differentiation of EECs is regulated by an interplay of the surface protein Notch and three basic helix-loop-helix transcriptional factors (Math1, Neurogenin 3, and NeuroD), among other factors [167,168]. EECs are classified depending on their morphology and position in the gastrointestinal mucosa into “open-type” with a bottle neck shape and an apical prolongation with microvilli facing towards the intestinal lumen or “closed type” that are located close to the basal membrane, do not reach the lumen of the gut, and lack microvilli [169,170,171]. The open-type cells are activated by luminal content through the microvilli, whereas the close type cells are activated by luminal content indirectly through neuronal or humoral pathways. In both cases, hormones and peptides accumulate into cytoplasmatic secretory granules that are released by exocytosis at the basolateral membrane upon chemical, mechanical, or neural stimulation [170,172].

29.5.5.0.1: 4.2.1. FRUCTOSE-INDUCED HORMONAL SECRETION IN INTESTINAL CELLS

Using specialized organoid cultures enriched in a single intestinal cell type, primarily enterocytes, Paneth or goblet, but not intestinal stem cells, Kishida et al. demonstrated that fructose can be sensed by absorptive enterocytes and secretory goblet and Paneth cells, but not stem cells [173]. In response to fructose there was an increased expression of fructolytic genes without affecting non-fructolytic gene expression. Sensing was independent of Notch, Wnt, and glucose concentrations in the culture medium, but required fructose uptake and metabolism. Stronger responses were found in more mature enterocyte- and goblet-enriched organoids. Of note, the response to fructose in enterocyte organoids was retained upon forced dedifferentiation to reacquire stem cells characteristics [173].

Fructose increases secretion of human peptide tyrosine tyrosine (PYY), cholecystokinin (CCK), neurotensin, and serotonin (5-HT) in EECs subtypes L, I, N, and enterochromaffin cells (EC), respectively [174,175]. Likewise, fructose stimulates secretion of glucagon-like peptide 1 (GLP-1) from L-subtype EECs in humans, rats and mice, but not glucose-dependent insulinotropic polypeptide (GIP; glucose-dependent insulinotropic polypeptide or gastric inhibitory peptide) [174,176]. On the other hand, fructose induced the secretion of GIP from K-subtype EECs in mice [176] but is unaffected or reduced in rats and humans [174,177].

29.5.5.0.1: 4.2.2. GLUCOSE-INDUCED HORMONAL SECRETION IN INTESTINAL CELLS

Oral glucose, but not intravenous glucose, leads to a greater stimulation of insulin secretion and modulation of glucagon secretion in the pancreas. This physiological response to glucose is called the incretin effect, which is due to the release of incretin hormones (GIP and GLP-1) from specialized EECs [178,179]. Of the three signals originating in the gut (glucose, incretins, and neutral signals transmitted by

the autonomic nervous system) that regulate pancreatic insulin secretion, the incretin effect makes a substantial contribution to maintenance of glucose homeostasis [178,179].

GIP is secreted in response to glucose by K-cells located in the duodenum and upper jejunum. GIP is synthesized as a precursor pro-peptide (pro-GIP), which is cleaved to GIP by posttranslational processing. GLP-1 and GLP-2 are secreted in response to glucose by L-cells in the small and large intestine, with a gradient from low density in the duodenum to high density in the ileum, but also in the colon and rectum [180,181]. The proglucagon gene is cleaved to GLP-1 and GLP-2 by posttranslational processing. The biological active forms of GLP-1 are GLP-1 [7-36 amide] (amidated GLP-1), and GLP-1 [7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37] (glycine-extended GLP-1), which are "truncated" forms in comparison to the originally proposed sequences GLP-1 [1-36 amide] by the N-terminal six amino acids [182,183,184]. In pancreatic α-cells, the same proglucagon gene is processed in a different manner, yielding glucagon and a "major proglucagon fragment that is not further processed to GLP-1 and GLP-2 [185]. Finally, in addition to its role in the regulation of pancreatic insulin secretion, GLP-1 and GLP-2 promotes nutrient absorption [180].

5-HT is secreted in response to glucose by EC cells located throughout the gastrointestinal tract, and regulates intestinal motility and brain control of appetite [186,187]. 5-HT and GLP-1 activate 5-HT_3 and GLP-1 receptors in the intestinal vagus nerve, respectively, leading to vagal reflexes, which in turn slow the subsequent emptying of carbohydrates from the stomach and induce satiation [188,189].

29.5.5.0.1: 4.2.3. INTESTINAL SWEET SENSING AND GLYCEMIC CONTROL

The gastrointestinal tract is a major determinant of metabolic homeostasis. Sensing of nutrients, and particularly glucose, in the EECs provides feedback signals from the intestine to slow the rate of gastric emptying, limit postprandial glycemic excursions, and induce satiation.

Intestinal sweet sensing is regulated by the sweet taste receptor (STR), which has been described on K-cells, L-cells, and EECs in humans. Additionally, STRs have been described in metabolic tissues that sense and respond to carbohydrates, such as hypothalamic neurons, hepatocytes, adipocytes, and β-cells (for review see refs. [190,191,192,193]). STR senses hexose sugars, D-amino acids, sweet proteins, and low-calorie sweeteners. The receptor is comprised of a heterodimer of class C, G-protein coupled receptors T1R2 and T1R3. The mechanisms of sweet taste transduction have been mostly studied in lingual sweet taste cells (this topic is out of the scope of this manuscript, for review see refs. [194,195,196,197]). Briefly, the interaction of sweet tastings with STR initiates the dissociation of the gustducin (the G-protein) into Gα and Gβγ subunits and activation of phospholipase C. Then, intracellular Ca^{2+} is released from inositol 1,4,5-triphosphate (IP3)-sensitive stores, leading to opening of the melastatine type-5 transient receptor potential cation (TRPM5) channel allowing sodium influx. Increases in intracellular Na^{+} and Ca^{2+} levels lead to depolarization of the basolateral membrane, which via 5-HT and ATP-dependent pathways activate intermediary taste cells and nerves involved in lingual sweet taste that convey information centrally to the cortex.

Numerous studies in rodents and human cells support the notion that intestinal STR is a glucose sensor on the gut luminal membrane responsible for the regulation of SGLT1 expression and GLP-1 secretion. First, it was demonstrated that T1R2, T1R3, and the α-subunit of gustducin were co-expressed in K- and L-endocrine cells in rodents and humans [198], and to a lesser extend in EC cells containing serotonin in pig intestine [199]. Second, Parker et al. proposed that secretion of GLP-1 by L-cells and GIP by K-cells was through uptake of glucose by SGLT1, suggesting that SGLT1 was likely the mediator of the direct responsiveness of K- and L-cells to luminal glucose [200]. Third, genetic deletion of T1R3 or gustducin in mice abolished the ability of mouse intestine to upregulate SGLT1 expression in response to increased dietary carbohydrate, providing convincing evidence for the involvement of the STR in intestinal sweet transduction [198]. Fourth, genetic deletion of T1R3 and gustducin exhibited deficiencies in secretion of GLP-1 [201]. Fifth, luminal glucose above a threshold results in secretion of GPL-1, GLP-2, and GIP through a signaling pathway involving STR in enteroendocrine cells [198]. These evidences beg an important question: How does glucose activation of STR in EECs cause increased expression of SGLT1 in enterocytes? The communication between EECs and neighboring enterocytes likely resides in the involvement of intermediaries such as GLP-1 and/or GLP-2 and enteric neurons. Thus, GLP-2 receptors are present on enteric neurons [202,203], while enterocytes respond to GLP-2 in an enteric neuron-dependent manner [203]. In addition, GLP-2 upregulates SGLT1 expression [204,205,206], and STR-dependent release of GLP-1 and GLP-2 is detected at higher concentrations in the portal and lymphatic circulation in rodents [204,207].

All these evidences have led to a model of intestinal dietary glucose sensing. Luminal glucose is sensed by STR expressed on the luminal membrane of enteroendocrine cells. Above a threshold level of luminal glucose, the hexose binds to and activates STR, initiated by dissociation of gustducin into Gα and Gβγ subunits, which leads to activation of phospholipase C β_2. Then, IP3-sensitive stores release intracellular Ca^{2+} that opens the TRPM5 channel increasing sodium influx. Intracellular elevation of Ca^{2+} and Na^{+} depolarizes the basolateral membrane resulting in the release of GLP-2. GLP-2 binds to its receptor on enteric neurons evoking an action potential that triggers the release of an unknown neuropeptide to the vicinity of neighboring enterocytes. The neuropeptide binds to its receptor located on basolateral membranes of enterocytes leading to a rise in intracellular levels of cAMP, which increases stabilization of the 3′end of *Slc5a1* mRNA and ultimately augmented SGLT1 translation and insertion into the apical brush border membrane of the enterocyte (**Figure 4**).

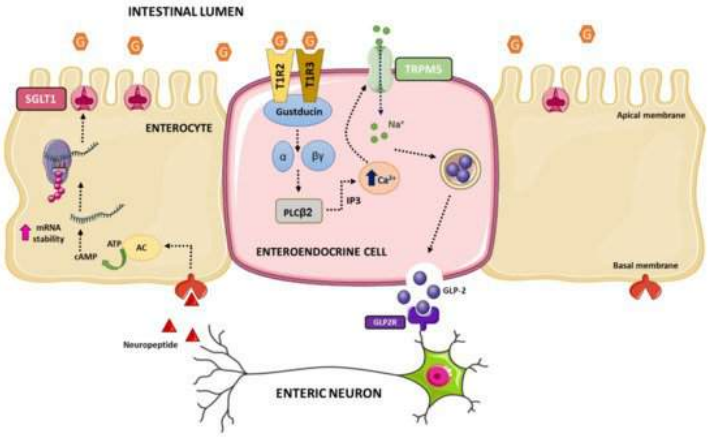


Figure 4. Model of intestinal glucose sensing and signaling pathways. Glucose (G) binds and activates the taste receptor type 1 member (STR) comprised of a heterodimer T1R2+T1R3 and G-protein gustducin, leading to its dissociation into Gα and Gβγ subunits and activation of phospholipase C β_2 (PLCβ_2) in enteroendocrine cells. Inositol 1,4,5-triphosphate (IP3) triggers intracellular calcium release resulting in increased sodium flux through type-5 transient receptor potential cation channel (TRPM5). Depolarization of the basal membrane results in GLP-2 release, which triggers release of an unidentified peptide from enteric neurons at neighbor enterocytes. Signal transduction in enterocyte leads to cAMP-mediated adenylate cyclase (AC) production that increases stabilization of *Slc5a1* mRNA and ultimately augmented translation and insertion of SGLT1 into apical membrane. This figure was created using Servier Medical Art.

29.5.5.0.1: 4.3. CENTRAL EFFECTS OF GLUCOSE AND FRUCTOSE CONSUMPTION

Sugar overconsumption has been associated with detrimental metabolic effects, such as obesity, dyslipidemia, MetS, and impaired insulin sensitivity [71,208,209]. Therefore, it is necessary to understand the specific molecular mechanisms by which dietary sugars cause an addictive eating behavior and how sugar intake affects the gut–brain axis. Herein, we will review the effects of the two main dietary monosaccharides: Glucose and fructose, the latter of which is usually consumed in the form of sucrose disaccharide (50% glucose, 50% fructose) or in the form of high-fructose corn syrup (HFCS) (range 47%–65% fructose, and 53%–35% glucose) [210], the major component of sweetened soft drinks.

Appetite control is a complex crosstalk between the periphery and the central nervous system that involves a large number of peptides and hormones [211]. Disturbances in food intake control will ultimately be responsible for large changes in energy balance and different metabolic effects. The appetite regulatory hormones are secreted from peripheral tissues such as the pancreas (e.g., insulin), adipose tissue (e.g., leptin), or the gastrointestinal tract [e.g., ghrelin, CCK, PYY, GLP-1 and GIP], and bind to receptors located in the arcuate nucleus of the hypothalamus, where they inhibit or stimulate appetite or satiety [212].

Many studies have demonstrated that circulating levels of satiety hormones are regulated by the type of sugar consumed. In response to glucose stimuli, a cascade of hormonal secretion is triggered. Thus, glucose produces a repression of the hunger hormone ghrelin (secreted by the stomach), whereas there is a stimulation of the secretion of satiety hormones such as leptin, insulin, GIP, GLP-1, and PYY. However, fructose produces lower repression of ghrelin and a decreased stimulation of satiety hormones (leptin, insulin, GIP, GLP-1, and PYY) than glucose [13,83,212,213,214,215,216,217]. These effects may be related to different explanations such as the lower ratio of intestinal fructose uptake, the lower intestinal levels of GLUT5 compared to the high levels of GLUT2, and also due to the low expression of GLUT5 in pancreatic β-cells leading to decreased insulin release [218,219,220].

Some of these hormones regulated differentially by fructose or glucose convey signals to brain structures. Specifically, there are two neuronal types in the arcuate nucleus that integrate signals from the periphery, acting as metabolic sensors: Neurons co-expressing agouti-related peptide (AgRP) and neuropeptide Y (NPY), whose activation triggers orexigenic effects; and neurons expressing pro-opiomelanocortin (POMC), whose activation triggers anorexigenic effects [221,222,223]. These different types of neurons are sensitive to changes in hormone levels promoting or suppressing food intake. Therefore, the differential effect of dietary sugars on hormonal levels affects neuronal stimulation causing both short-term and long-term central effects in the regulation of food intake and energy homeostasis [224,225]. The low stimulatory capacity of fructose on satiety hormones such as leptin and insulin will lead to low stimulation of POMC neurons and the maintenance of the signal on NPY/AgRP neurons, thus promoting less satiety than glucose, and therefore increased food intake. In the same way, the hypothalamic AMPK functions as a ‘fuel gauge’ to monitor cellular energy status, and its inhibition promotes anorexigenic effect [226]. AMPK activity is inhibited by leptin and insulin. Intracerebroventricular glucose administration in rodents inhibits hypothalamic AMPK activity and suppresses food intake, whereas fructose activates it, thus promoting an orexigenic effect [227,228,229,230] (Figure 5).

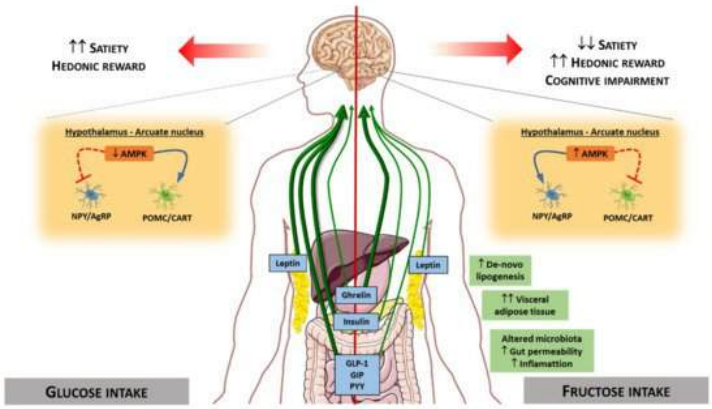


Figure 5. Peripheral and central effects of glucose and fructose on food intake. Upon nutrient ingestion, glucose triggers the secretion of anorexigenic peptides such as leptin, insulin, GLP-1, GIP, PYY, and blocks the secretion of the orexigenic hormone ghrelin. These peptides will induce AMPK inhibition, which will lead to the stimulation of POMC/CART neurons, contributing to satiety response. However, upon fructose consumption, the secretion of anorexigenic peptides is decreased, as well as the repression of ghrelin secretion. These facts will cause the increase in AMPK activity, which will lead to the repression of POMC/CART neurons and the activation of NPY/AgRP, leading to reduced satiety suppression, as well as higher hedonic reward to food and increasing food intake, in addition to playing a possible role in cognitive functions. This figure was created using Servier Medical Art.

With the use of new technological advances, it is possible to evaluate the brain activity produced by the intake of different nutrients. In humans, differences in cerebral blood flow have been reported between subjects undergoing glucose and fructose infusions [231], and compared to glucose, fructose causes poor satiety stimulation in specific appetite-regulating regions (e.g., hypothalamus) [13]. It has also been observed that fructose ingestion compared to glucose resulted in a significantly greater incentive value of food cues [232]. These findings suggest that fructose promotes effects on brain activity that affect appetite, probably promoting less satiety than other sugars in humans.

In addition to the above-mentioned findings, it has been described that high-fructose intake may affect central appetite regulation by altering specific components of the endocannabinoid system in rats. Fructose consumption has been reported to significantly increase the mRNA expression of the cannabinoid 1 receptor (CB1) [233], and induces an increase in fatty acid amide hydrolase (FAAH) and diacylglycerol lipase (DAG) 1β, but a decrease in DAG1α mRNA [234]. These changes in the endocannabinoid system suggest that fructose consumption may lead to increased hedonic reward for food, thus leading to disturbances in the eating behavior pattern.

The consumption of dietary sugars has not only been related to central effects that control appetite and satiety, but also to disturbances in cognitive functions. In rodents, studies have shown that fructose consumption reduced phosphorylation levels of the insulin receptor, leading to impaired brain insulin signaling [235,236], a harmful feature associated with cognitive impairment [237]. Moreover, diminished phosphorylation of cAMP-response element binding and synapsin I, and reduced synaptophysin levels have been observed after fructose intake [236]. Together, these findings indicate that excessive fructose consumption could lead not only to detrimental effects in eating behavior, but also can trigger impaired cognitive function. Further work is required to investigate these evidences.

29.5.5.0.1: THE FRUCTOSE HYPOTHESIS

In view of the association between fructose consumption in Western diets and MetS, fructose has been suggested as one of etiological factor of MetS. The “fructose hypothesis” proposed that a high amount of fructose consumption is a leading risk factor for the development and progression of MetS, covering obesity, insulin resistance, dyslipidemia, fatty liver, and cardiovascular disease [238,239,240].

Fructose may cause insulin resistance by accumulation of triglycerides in the liver. There are two metabolic pathways to increased hepatic lipid content, i.e., lipogenesis and/or reduced mitochondrial fatty acid oxidation. Hepatic fructolysis leads to increased gluconeogenic sources resulting in elevated rates of lipogenesis [16,45,46]. Hepatic accumulation of toxic intermediary lipid metabolites, such as diacylglycerol (DAG) results in PKCε activation that impairs hepatic insulin signaling through phosphorylation of serine residues on the insulin receptor substrate 1 and 2 (IRS1/2). When hepatic insulin signaling is impaired, gluconeogenesis and glycogenolysis are unleashed, contributing to hyperglycemia and hyperinsulinemia. Under these circumstances, hepatic lipid synthesis is enhanced due to hyperinsulinemia [241,242]. Likewise, reduced fatty acid oxidation leads to hepatic triglycerides accumulation. Of note, Ohashi et al. demonstrated that excessive amounts of fructose consumption lead to epigenetic modifications, such as DNA hypermethylation of promoter regions of peroxisome proliferator-activated receptor alpha (PPARα) and carnitine palmitoyl transferase 1A (CPT1A) that results in lower amounts of mRNA levels [243]. Hepatic triglyceride accumulation results in augmented secretion of very low-density lipoprotein (VLDL) leading to increased lipid uptake in skeletal muscle and peripheral tissues. Similarly to what happens in the liver, intramyocellular lipid accumulation (particularly DAG) activates the PKCθ isoform that phosphorylates and inactivates IRS1 resulting in impaired insulin-stimulated glucose uptake, contributing to hyperglycemia, increased delivery of glucose to the liver, and hyperinsulinemia [241,242].

Fructose-induced hyperuricemia has also been proposed as a causal agent in the etiology of insulin resistance [244,245,246]. This notion arises from the observation that lowering uric acid levels prevents the development of MetS induced by fructose [244,245,246], defective endothelial NO production in mice leads to development of MetS [247], and that uric acid inhibits endothelial NO in in vitro and in vivo [248]. Two mechanisms have been proposed: The first mechanism proposed is that uric acid inhibits endothelial nitric oxide (NO) release, and NO increases blood flow ensuing enhanced insulin delivery and glucose disposal in skeletal muscle and peripheral tissues [249]. The second mechanism states that uric acid promotes inflammation and oxidative stress within the adipocyte [250,251,252]. In addition, uric acid-mediated insulin resistance in the adipose tissue, via the classical mechanisms (i.e., low-grade chronic inflammation mediated by proinflammatory cytokines secreted by the adipocytes, increased lipolysis, and reduced lipogenesis), may result in MetS [241,242].

Additionally, persistent high fructose consumption leads to higher levels of leptin and leptin resistance, which in turn increases food and energy intake [253]. Potential molecular mechanisms underlying leptin resistance may be related to impaired leptin transport across the blood-brain barrier and/or reduced basal levels of phosphorylated signal transducer and activator of transcription 3 (STAT3; a downstream component of the leptin receptor signaling cascade), despite equivalent expression of leptin receptors, in the hypothalamus [253].

However, the fructose hypothesis is not universally accepted. It has been argued that fructose is rarely consumed in its pure form and many published studies have used fructose levels that far exceed dietary composition [254]. Likewise, many animal studies have used extremely high-fructose doses or unusual glucose to fructose ratio that are not representative of actual human diets, which makes it difficult to extrapolate this phenomenon to humans. Therefore, caution in interpreting studies of the effects of fructose on health should be taken into consideration [254]. Another proposed argument to refute the fructose hypothesis is that the causative role of fructose in increasing the risk for the development and progression of MetS is not fully demonstrated. Carefully designed studies should be performed to tease apart the contribution of each risk factor associated to MetS (e.g., obesity, diabetes, or insulin resistance) from fructose, per se [254].

29.5.5.0.1: 4.4. PERIPHERAL EFFECTS OF GLUCOSE AND FRUCTOSE CONSUMPTION

29.5.5.0.1: GUT MICROBIOTA, LIPID METABOLISM, AND LIVER DISEASE

The gut microbiota is a complex and dynamic population of microorganisms that, in addition to acting as an immune barrier and protecting against pathogens, plays a crucial role as a metabolic organ itself modulating intestinal permeability, and therefore the nutrient availability [255]. It is generally known that diet exerts a large effect on the gut microbiota, which may affect intestinal permeability and ultimately cause a great metabolic impact [255,256,257,258].

High-fructose or high-glucose diets have been described as an intestinal microbiota modulator that increases inflammation, gut permeability, and metabolic endotoxemia, causing metabolic disturbances such as hepatic lipid accumulation, liver damage, and insulin resistance [258,259]. Likewise, sugar overconsumption also affects lipid metabolism. In obese and overweight subjects, the consumption of glucose-sweetened beverages leads to a lower increase in plasma triglycerides, de novo lipogenesis, and visceral adipose tissue compared to those that consumed fructose-sweetened beverage [71]. However, in rodents, both high-glucose and -fructose diets stimulated similar hepatic lipogenic gene expression [260].

Liver is the principal metabolic organ within the human body and has a major role in regulating carbohydrate metabolism [261]. Many studies point out to the direct implication of high-sugar diets in the development of serious liver diseases, such as NAFLD, hepatic steatosis, liver fibrosis, and dysfunction [262,263,264]. Multiple studies showed that fructose more potently stimulates hepatic de novo lipogenesis than glucose [78,265,266], and the effect is much higher when both monosaccharides were consumed simultaneously [265]. These differences in de novo lipogenesis between both sugars can be explained by differences in their hepatic metabolism. Fructose is directly phosphorylated by fructokinase, bypassing the enzyme phosphofructokinase, a major rate-limiting step in glucose metabolism, providing a larger available substrate for de novo lipogenesis than glucose [261,267].

Regarding the effect of isocaloric diets with different sugar composition, various studies have observed no differences in liver fat content between high-fructose or high-glucose diets [268,269], nor between isocaloric diets with high-fructose corn syrup or sucrose [270]. However, when comparing different doses of fructose in the diet, liver fat content was increased in high-fructose diet, probably associated with increased de novo lipogenesis and reduced whole-body fatty acid oxidation [266,270]. In the same direction as the previous findings, when comparing hypercaloric diets enriched in either fructose or glucose, no significant changes are observed between both diets, suggesting that high-glucose and high-fructose diets provide the same risk for the development of NAFLD [269,271,272].

Taken together, these data would point to the detrimental effect of fructose compared to glucose in terms of hepatic and lipid metabolism. However, there is controversy between different studies, probably due to differences in the doses of sugars administered and their form of administration (oral, intraperitoneal injection, etc.). Many of the above-mentioned studies were performed using supra-doses of fructose in rodents. Since humans typically do not consume fructose as a single sugar, and it is frequently consumed in the form of HFCS, the direct relationship with the real effect of fructose human consumption is not entirely clear. Therefore, more detailed studies on the pattern of sugar consumption in humans should be carried out.

29.5.5.0.1: 4.5. IMPACT OF EXCESSIVE DIETARY SUGARS CONSUMPTION ON INCRETIN SECRETION

There are many associations reported between high-sugar consumption and the development of pathologies such as diabetes, obesity, and MetS [10,273,274,275]. These associations are mainly due to the current consumption of sugar-sweetened beverages, whose main sweetener

is the HFCS. HFCS represents >40% of caloric sweeteners and its consumption has been increased by >1000% between 1970 and 1990. This sugar overconsumption can lead to important changes in the secretion of gut hormones, and therefore, lead to central effects that affect appetite and satiety control.

Many authors have focused their studies on the effect that different GIs and GLs have on incretin secretion. Runchey et al. observed that 28-days consumption of a high-GL diet in weight-maintained healthy individuals led to statistically significant increased post-prandial GIP and lower GLP-1 concentrations compared with low-GL diets [276]. However, other authors did not corroborate these findings clearly. One study performed in healthy sedentary women reported that GLP-1 concentrations did not differ significantly following high- or low-GI meals [277]. In the same way, another study in overweight subjects observed no differences in GLP-1 concentrations when comparing consumption of low- and high-GI beverages [278]. Other authors suggest that the rate of small intestinal glucose exposure (i.e., GL) is a major determinant of the magnitude of the incretin effect, since they observed that the incretin effect was stronger when they administered larger intraduodenal glucose load [279].

However, it is necessary to be cautious with this upregulation of incretins in response to high-sugar diets, because it has been described that in diabetic patients, who have increased levels of incretins, a reduced incretin effect is observed, which suggest the development of an "incretin resistance" process [280,281,282,283].

29.5.6: 5. FUTURE DIRECTIONS

During human evolution, ancestral human diets contained low carbohydrate levels and most of the sugars were derived from fruits and honey. In the last century, changes in lifestyle, nutritional habits in the world population, and the abusive use of sweeteners by the food industry have dramatically increased dietary sugar consumption, particularly constituent monomers, such as glucose and fructose, and fructose-based sweeteners. International and national health organizations have called attention into this issue and recommend reductions in sugars consumption due to concerns in their potential role as risk factors for developing human diseases such as obesity and T2DM.

In the last decades, the scientific community has made great efforts to understand intestinal sugar absorption, identifying molecular and physiological mechanisms of fructose and glucose sensing and transport. In the case of fructose metabolism, the current notion that fructose is mainly metabolized by the liver has been challenged, and the new paradigm proposes that the small intestine shields the liver from toxic fructose exposure. This provocative view of intestinal fructose metabolism is awaiting confirmation in humans. Similarly, the finding that humans can synthesize fructose by the polyol pathway leaves open the question about the significance of this pathway in human metabolic diseases. In addition, more meta-analysis studies should be performed to clearly demonstrate the causal role of dietary fructose and glucose in developing human metabolic diseases.

On the other hand, the role of intestinal glucose metabolism on the etiology of hyperglycemia remains incompletely clarified. Despite studies relating chronic hyperglycemia with impaired glucose transport and metabolism in the small intestine, more studies in humans are required to reveal if chronic hyperglycemia is a cause or consequence of impaired glucose homeostasis in the small intestine. The identification of molecular mechanisms by which glucose and insulin regulate SGLT1 have set this transporter, and its potential role in the physiopathology of hyperglycemia and intestinal insulin resistance, in the spotlight. In this line of thinking, further research is required to demonstrate the efficacy of SGLT1 inhibitors in the treatment of T2DM and obesity. Likewise, it remains to be clarified whether the apical localization of GLUT2 in response to high glucose levels in obese and/or diabetic patients is an adaptive mechanism to protect the body from excessive glucose concentrations, or if it is a consequence of hyperglycemia and insulin resistance.

Finally, the differentiated effects of glucose and fructose on eating behavior and impaired cognitive function observed in rodent models are difficult to extrapolate to humans due to the use of extremely high-fructose diets or unusual glucose to fructose ratio. To clarify the causality of fructose in human eating disorders leading to metabolic diseases, it is necessary to develop new research tools and experimental approaches in humans.

29.5.7: AUTHOR CONTRIBUTIONS

I.C.-C. and G.P. conceptualized the manuscript. B.M., C.M.F.-D., I.C.-C., and G.P. drafted the manuscript. B.M. and C.M.F.-D. prepared the figures. I.C.-C. and G.P. revised and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

29.5.8: FUNDING

This research was funded by the Spanish MINISTERIO DE ECONOMÍA, INDUSTRIA Y COMPETITIVIDAD, grant numbers SAF2016-77871-C2-1-R and SAF2016-77871-C2-2-R to I.C-C. and G.P. respectively; the EFSD European Research Programme on New Targets for Type 2 Diabetes supported by an educational research grant from MSD to I.C-C. and G.P.; the FUNDACIÓN LA-CAIXA Y FUNDACIÓN CAJA DE BURGOS, grant number CAIXA-UBU001 to G.P.

29.5.9: CONFLICTS OF INTEREST

The authors declare no conflict of interest.

29.5.10: REFERENCES

1. *Obesity and Overweight, Fact Sheet 311*; World Health Organization: Geneva, Switzerland, 2018.
2. Mooradian, A.D.; Smith, M.; Tokuda, M. The role of artificial and natural sweeteners in reducing the consumption of table sugar: A narrative review. *Clin. Nutr. ESPen* **2017**, *18*, 1–8. [Google Scholar] [CrossRef] [PubMed]
3. Douard, V.; Ferraris, R.P. The role of fructose transporters in diseases linked to excessive fructose intake. *J. Physiol.* **2013**, *591*, 401–414. [Google Scholar] [CrossRef] [PubMed]
4. Marshall, R.O.; Kooi, E.R. Enzymatic conversion of D-glucose to D-fructose. *Science* **1957**, *125*, 648–649. [Google Scholar] [CrossRef] [PubMed]
5. Gross, L.S.; Li, L.; Ford, E.S.; Liu, S. Increased consumption of refined carbohydrates and the epidemic of type 2 diabetes in the United States: An ecologic assessment. *Am. J. Clin. Nutr.* **2004**, *79*, 774–779. [Google Scholar] [CrossRef] [PubMed]
6. Powell, E.S.; Smith-Taillie, L.P.; Popkin, B.M. Added Sugars Intake Across the Distribution of US Children and Adult Consumers: 1977–2012. *J. Acad. Nutr. Diet.* **2016**, *116*, 1543–1550. [Google Scholar] [CrossRef]
7. Ervin, R.B.; Ogden, C.L. Consumption of added sugars among U.S. adults, 2005–2010. *NCHS Data Brief* **2013**, *122*, 1–8. [Google Scholar]
8. Corpe, C.P.; Basaleh, M.M.; Affleck, J.; Gould, G.; Jess, T.J.; Kellett, G.L. The regulation of GLUT5 and GLUT2 activity in the adaptation of intestinal brush-border fructose transport in diabetes. *Pflug. Arch.* **1996**, *432*, 192–201. [Google Scholar] [CrossRef]
9. DiNicolantonio, J.J.; O'Keefe, J.H.; Lucan, S.C. Added fructose: A principal driver of type 2 diabetes mellitus and its consequences. *Mayo Clin. Proc.* **2015**, *90*, 372–381. [Google Scholar] [CrossRef]
10. Imamura, F.; O'Connor, L.; Ye, Z.; Mursu, J.; Hayashino, Y.; Bhupathiraju, S.N.; Forouhi, N.G. Consumption of sugar sweetened beverages, artificially sweetened beverages, and fruit juice and incidence of type 2 diabetes: Systematic review, meta-analysis, and estimation of population attributable fraction. *BMJ* **2015**, *351*, h3576. [Google Scholar] [CrossRef]
11. Montonen, J.; Jarvinen, R.; Knekt, P.; Heliovaara, M.; Reunanen, A. Consumption of sweetened beverages and intakes of fructose and glucose predict type 2 diabetes occurrence. *J. Nutr.* **2007**, *137*, 1447–1454. [Google Scholar] [CrossRef]
12. Grembecka, M. Natural sweeteners in a human diet. *Rocz. Panstw. Zakl. Hig.* **2015**, *66*, 195–202. [Google Scholar] [PubMed]
13. Page, K.A.; Chan, O.; Arora, J.; Belfort-Deaguiar, R.; Dzuira, J.; Roehmholdt, B.; Cline, G.W.; Naik, S.; Sinha, R.; Constable, R.T.; et al. Effects of fructose vs glucose on regional cerebral blood flow in brain regions involved with appetite and reward pathways. *JAMA* **2013**, *309*, 63–70. [Google Scholar] [CrossRef] [PubMed]
14. Ferraris, R.P.; Choe, J.Y.; Patel, C.R. Intestinal Absorption of Fructose. *Annu. Rev. Nutr.* **2018**, *38*, 41–67. [Google Scholar] [CrossRef] [PubMed]
15. Melanson, K.J.; Angelopoulos, T.J.; Nguyen, V.; Zukley, L.; Lowndes, J.; Rippe, J.M. High-fructose corn syrup, energy intake, and appetite regulation. *Am. J. Clin. Nutr.* **2008**, *88*, 1738S–1744S. [Google Scholar] [CrossRef] [PubMed]
16. Mayes, P.A. Intermediary metabolism of fructose. *Am. J. Clin. Nutr.* **1993**, *58*, 754S–765S. [Google Scholar] [CrossRef] [PubMed]
17. Douard, V.; Ferraris, R.P. Regulation of the fructose transporter GLUT5 in health and disease. *Am. J. Physiol. Endocrinol. Metab.* **2008**, *295*, E227–E237. [Google Scholar] [CrossRef]
18. George Thompson, A.M.; Iancu, C.V.; Nguyen, T.T.; Kim, D.; Choe, J.Y. Inhibition of human GLUT1 and GLUT5 by plant carbohydrate products; insights into transport specificity. *Sci. Rep.* **2015**, *5*, 12804. [Google Scholar] [CrossRef]
19. Burant, C.F.; Takeda, J.; Brot-Laroche, E.; Bell, G.I.; Davidson, N.O. Fructose transporter in human spermatozoa and small intestine is GLUT5. *J. Biol. Chem.* **1992**, *267*, 14523–14526. [Google Scholar]
20. Shepherd, P.R.; Gibbs, E.M.; Wesslau, C.; Gould, G.W.; Kahn, B.B. Human small intestine facilitative fructose/glucose transporter (GLUT5) is also present in insulin-responsive tissues and brain. Investigation of biochemical characteristics and translocation. *Diabetes* **1992**, *41*, 1360–1365. [Google Scholar] [CrossRef]
21. Ebert, K.; Ewers, M.; Bisha, I.; Sander, S.; Rasputniac, T.; Daniel, H.; Antes, I.; Witt, H. Identification of essential amino acids for glucose transporter 5 (GLUT5)-mediated fructose transport. *J. Biol. Chem.* **2018**, *293*, 2115–2124. [Google Scholar] [CrossRef]
22. Manolescu, A.R.; Witkowska, K.; Kinnaird, A.; Cessford, T.; Cheeseman, C. Facilitated hexose transporters: New perspectives on form and function. *Physiology* **2007**, *22*, 234–240. [Google Scholar] [CrossRef] [PubMed]
23. Thorens, B. Molecular and cellular physiology of GLUT-2, a high-Km facilitated diffusion glucose transporter. *Int. Rev. Cytol.* **1992**, *137*, 209–238. [Google Scholar] [CrossRef] [PubMed]
24. Mueckler, M.; Thorens, B. The SLC2 (GLUT) family of membrane transporters. *Mol. Asp. Med.* **2013**, *34*, 121–138. [Google Scholar] [CrossRef] [PubMed]
25. Kane, S.; Seatter, M.J.; Gould, G.W. Functional studies of human GLUT5: Effect of pH on substrate selection and an analysis of substrate interactions. *Biochem. Biophys. Res. Commun.* **1997**, *238*, 503–505. [Google Scholar] [CrossRef] [PubMed]
26. Miyamoto, K.; Tatsumi, S.; Morimoto, A.; Minami, H.; Yamamoto, H.; Sone, K.; Taketani, Y.; Nakabou, Y.; Oka, T.; Takeda, E. Characterization of the rabbit intestinal fructose transporter (GLUT5). *Biochem. J.* **1994**, *303 Pt 3*, 877–883. [Google Scholar] [CrossRef]

27. Corpe, C.P.; Bovelander, F.J.; Munoz, C.M.; Hoekstra, J.H.; Simpson, I.A.; Kwon, O.; Levine, M.; Burant, C.F. Cloning and functional characterization of the mouse fructose transporter, GLUT5. *Biochim. Biophys. Acta* **2002**, *1576*, 191–197. [Google Scholar] [CrossRef]
28. Mate, A.; Barfull, A.; Hermosa, A.M.; Planas, J.M.; Vazquez, C.M. Regulation of D-fructose transporter GLUT5 in the ileum of spontaneously hypertensive rats. *J. Membr. Biol.* **2004**, *199*, 173–179. [Google Scholar] [CrossRef]
29. Douard, V.; Choi, H.I.; Elshenawy, S.; Lagunoff, D.; Ferraris, R.P. Developmental reprogramming of rat GLUT5 requires glucocorticoid receptor translocation to the nucleus. *J. Physiol.* **2008**, *586*, 3657–3673. [Google Scholar] [CrossRef]
30. Prieto, P.G.; Cancelas, J.; Villanueva-Penacarrillo, M.L.; Valverde, I.; Malaisse, W.J. Plasma D-glucose, D-fructose and insulin responses after oral administration of D-glucose, D-fructose and sucrose to normal rats. *J. Am. Coll. Nutr.* **2004**, *23*, 414–419. [Google Scholar] [CrossRef]
31. Preston, G.M.; Calle, R.A. Elevated Serum Sorbitol and not Fructose in Type 2 Diabetic Patients. *Biomark. Insights* **2010**, *5*, 33–38. [Google Scholar] [CrossRef]
32. Wahjudi, P.N.; Patterson, M.E.; Lim, S.; Yee, J.K.; Mao, C.S.; Lee, W.N. Measurement of glucose and fructose in clinical samples using gas chromatography/mass spectrometry. *Clin. Biochem.* **2010**, *43*, 198–207. [Google Scholar] [CrossRef] [PubMed]
33. Macdonald, I.; Keyser, A.; Pacy, D. Some effects, in man, of varying the load of glucose, sucrose, fructose, or sorbitol on various metabolites in blood. *Am. J. Clin. Nutr.* **1978**, *31*, 1305–1311. [Google Scholar] [CrossRef] [PubMed]
34. Kawasaki, T.; Igarashi, K.; Ogata, N.; Oka, Y.; Ichiyanagi, K.; Yamanouchi, T. Markedly increased serum and urinary fructose concentrations in diabetic patients with ketoacidosis or ketosis. *Acta Diabetol.* **2012**, *49*, 119–123. [Google Scholar] [CrossRef] [PubMed]
35. Kawasaki, T.; Ogata, N.; Akanuma, H.; Sakai, T.; Watanabe, H.; Ichiyanagi, K.; Yamanouchi, T. Postprandial plasma fructose level is associated with retinopathy in patients with type 2 diabetes. *Metabolism* **2004**, *53*, 583–588. [Google Scholar] [CrossRef] [PubMed]
36. Hers, H.G. [Liver fructokinase]. *Biochim. Biophys. Acta* **1952**, *8*, 416–423. [Google Scholar] [CrossRef]
37. Bonthron, D.T.; Brady, N.; Donaldson, I.A.; Steinmann, B. Molecular basis of essential fructosuria: Molecular cloning and mutational analysis of human ketohexokinase (fructokinase). *Hum. Mol. Genet.* **1994**, *3*, 1627–1631. [Google Scholar] [CrossRef]
38. Hayward, B.E.; Bonthron, D.T. Structure and alternative splicing of the ketohexokinase gene. *Eur. J. Biochem.* **1998**, *257*, 85–91. [Google Scholar] [CrossRef]
39. Diggle, C.P.; Shires, M.; Leitch, D.; Brooke, D.; Carr, I.M.; Markham, A.F.; Hayward, B.E.; Asipu, A.; Bonthron, D.T. Ketohexokinase: Expression and localization of the principal fructose-metabolizing enzyme. *J. Histochem. Cytochem.* **2009**, *57*, 763–774. [Google Scholar] [CrossRef]
40. Asipu, A.; Hayward, B.E.; O'Reilly, J.; Bonthron, D.T. Properties of normal and mutant recombinant human ketohexokinases and implications for the pathogenesis of essential fructosuria. *Diabetes* **2003**, *52*, 2426–2432. [Google Scholar] [CrossRef]
41. Ishimoto, T.; Lanaspa, M.A.; Le, M.T.; Garcia, G.E.; Diggle, C.P.; Maclean, P.S.; Jackman, M.R.; Asipu, A.; Roncal-Jimenez, C.A.; Kosugi, T.; et al. Opposing effects of fructokinase C and A isoforms on fructose-induced metabolic syndrome in mice. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 4320–4325. [Google Scholar] [CrossRef]
42. Giroix, M.H.; Jijakli, H.; Courtois, P.; Zhang, Y.; Sener, A.; Malaisse, W.J. Fructokinase activity in rat liver, ileum, parotid gland, pancreas, pancreatic islet, B and non-B islet cell homogenates. *Int. J. Mol. Med.* **2006**, *17*, 517–522. [Google Scholar] [CrossRef] [PubMed]
43. Underwood, A.H.; Newsholme, E.A. Properties of Phosphofructokinase from Rat Liver and Their Relation to the Control of Glycolysis and Gluconeogenesis. *Biochem. J.* **1965**, *95*, 868–875. [Google Scholar] [CrossRef] [PubMed]
44. Li, X.; Qian, X.; Peng, L.X.; Jiang, Y.; Hawke, D.H.; Zheng, Y.; Xia, Y.; Lee, J.H.; Cote, G.; Wang, H.; et al. A splicing switch from ketohexokinase-C to ketohexokinase-A drives hepatocellular carcinoma formation. *Nat. Cell Biol.* **2016**, *18*, 561–571. [Google Scholar] [CrossRef] [PubMed]
45. Karczmar, G.S.; Tavares, N.J.; Weiner, M.W. A 31P NMR study of the GI tract: Effect of fructose loading and measurement of transverse relaxation times. *Magn. Reson. Med.* **1989**, *9*, 8–15. [Google Scholar] [CrossRef] [PubMed]
46. Tharabenjasin, P.; Douard, V.; Patel, C.; Krishnamra, N.; Johnson, R.J.; Zuo, J.; Ferraris, R.P. Acute interactions between intestinal sugar and calcium transport in vitro. *Am. J. Physiol. Gastrointest. Liver Physiol.* **2014**, *306*, G1–G12. [Google Scholar] [CrossRef] [PubMed]
47. Patel, C.; Douard, V.; Yu, S.; Tharabenjasin, P.; Gao, N.; Ferraris, R.P. Fructose-induced increases in expression of intestinal fructolytic and gluconeogenic genes are regulated by GLUT5 and KHK. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **2015**, *309*, R499–R509. [Google Scholar] [CrossRef] [PubMed]
48. Cui, X.L.; Soteropoulos, P.; Tolias, P.; Ferraris, R.P. Fructose-responsive genes in the small intestine of neonatal rats. *Physiol. Genom.* **2004**, *18*, 206–217. [Google Scholar] [CrossRef]
49. Jiang, L.; Ferraris, R.P. Developmental reprogramming of rat GLUT-5 requires de novo mRNA and protein synthesis. *Am. J. Physiol. Gastrointest. Liver Physiol.* **2001**, *280*, G113–G120. [Google Scholar] [CrossRef]
50. Cui, X.L.; Ananian, C.; Perez, E.; Strenger, A.; Beuve, A.V.; Ferraris, R.P. Cyclic AMP stimulates fructose transport in neonatal rat small intestine. *J. Nutr.* **2004**, *134*, 1697–1703. [Google Scholar] [CrossRef]

51. Gouyon, F.; Onesto, C.; Dalet, V.; Pages, G.; Leturque, A.; Brot-Laroche, E. Fructose modulates GLUT5 mRNA stability in differentiated Caco-2 cells: Role of cAMP-signalling pathway and PABP (polyadenylated-binding protein)-interacting protein (Paip) 2. *Biochem. J.* **2003**, *375*, 167–174. [Google Scholar] [CrossRef]
52. Cui, X.L.; Schlesier, A.M.; Fisher, E.L.; Cerqueira, C.; Ferraris, R.P. Fructose-induced increases in neonatal rat intestinal fructose transport involve the PI3-kinase/Akt signaling pathway. *Am. J. Physiol. Gastrointest. Liver Physiol.* **2005**, *288*, G1310–G1320. [Google Scholar] [CrossRef]
53. Franco, I.; Gulluni, F.; Campa, C.C.; Costa, C.; Margaria, J.P.; Ciraolo, E.; Martini, M.; Monteyne, D.; De Luca, E.; Germena, G.; et al. PI3K class II alpha controls spatially restricted endosomal PtdIns3P and Rab11 activation to promote primary cilium function. *Dev. Cell.* **2014**, *28*, 647–658. [Google Scholar] [CrossRef]
54. Ortega-Prieto, P.; Postic, C. Carbohydrate Sensing Through the Transcription Factor ChREBP. *Front. Genet.* **2019**, *10*, 472. [Google Scholar] [CrossRef]
55. Oh, A.R.; Sohn, S.; Lee, J.; Park, J.M.; Nam, K.T.; Hahm, K.B.; Kim, Y.B.; Lee, H.J.; Cha, J.Y. ChREBP deficiency leads to diarrhea-predominant irritable bowel syndrome. *Metabolism* **2018**, *85*, 286–297. [Google Scholar] [CrossRef]
56. Lee, H.J.; Cha, J.Y. Recent insights into the role of ChREBP in intestinal fructose absorption and metabolism. *BMB Rep.* **2018**, *51*, 429–436. [Google Scholar] [CrossRef]
57. Iizuka, K. The Role of Carbohydrate Response Element Binding Protein in Intestinal and Hepatic Fructose Metabolism. *Nutrients* **2017**, 9. [Google Scholar] [CrossRef]
58. Kim, M.; Astapova, I.I.; Flier, S.N.; Hannou, S.A.; Doridot, L.; Sargsyan, A.; Kou, H.H.; Fowler, A.J.; Liang, G.; Herman, M.A. Intestinal, but not hepatic, ChREBP is required for fructose tolerance. *JCI Insight* **2017**, 2. [Google Scholar] [CrossRef]
59. Kato, T.; Iizuka, K.; Takao, K.; Horikawa, Y.; Kitamura, T.; Takeda, J. ChREBP-Knockout Mice Show Sucrose Intolerance and Fructose Malabsorption. *Nutrients* **2018**, *10*, 340. [Google Scholar] [CrossRef]
60. Shalev, A. Minireview: Thioredoxin-interacting protein: Regulation and function in the pancreatic beta-cell. *Mol. Endocrinol.* **2014**, *28*, 1211–1220. [Google Scholar] [CrossRef]
61. Patwari, P.; Lee, R.T. An expanded family of arrestins regulate metabolism. *Trends Endocrinol. Metab.* **2012**, *23*, 216–222. [Google Scholar] [CrossRef]
62. Cha-Molstad, H.; Saxena, G.; Chen, J.; Shalev, A. Glucose-stimulated expression of Txnip is mediated by carbohydrate response element-binding protein, p300, and histone H4 acetylation in pancreatic beta cells. *J. Biol. Chem.* **2009**, *284*, 16898–16905. [Google Scholar] [CrossRef] [PubMed]
63. Stoltzman, C.A.; Peterson, C.W.; Breen, K.T.; Muoio, D.M.; Billin, A.N.; Ayer, D.E. Glucose sensing by MondoA:Mlx complexes: A role for hexokinases and direct regulation of thioredoxin-interacting protein expression. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 6912–6917. [Google Scholar] [CrossRef] [PubMed]
64. Wu, N.; Zheng, B.; Shaywitz, A.; Dagon, Y.; Tower, C.; Bellinger, G.; Shen, C.H.; Wen, J.; Asara, J.; McGraw, T.E.; et al. AMPK-dependent degradation of TXNIP upon energy stress leads to enhanced glucose uptake via GLUT1. *Mol. Cell* **2013**, *49*, 1167–1175. [Google Scholar] [CrossRef] [PubMed]
65. Dotimas, J.R.; Lee, A.W.; Schmider, A.B.; Carroll, S.H.; Shah, A.; Bilen, J.; Elliott, K.R.; Myers, R.B.; Soberman, R.J.; Yoshioka, J.; et al. Diabetes regulates fructose absorption through thioredoxin-interacting protein. *Elife* **2016**, 5. [Google Scholar] [CrossRef] [PubMed]
66. Douard, V.; Cui, X.L.; Soteropoulos, P.; Ferraris, R.P. Dexamethasone sensitizes the neonatal intestine to fructose induction of intestinal fructose transporter (Slc2A5) function. *Endocrinology* **2008**, *149*, 409–423. [Google Scholar] [CrossRef] [PubMed]
67. Henning, S.J. Postnatal development: Coordination of feeding, digestion, and metabolism. *Am. J. Physiol.* **1981**, *241*, G199–G214. [Google Scholar] [CrossRef] [PubMed]
68. Monteiro, I.M.; Jiang, L.; Ferraris, R.P. Dietary modulation of intestinal fructose transport and GLUT5 mRNA expression in hypothyroid rat pups. *J. Pediatr. Gastroenterol. Nutr.* **1999**, *29*, 563–570. [Google Scholar] [CrossRef]
69. Shu, R.; David, E.S.; Ferraris, R.P. Luminal fructose modulates fructose transport and GLUT-5 expression in small intestine of weaning rats. *Am. J. Physiol.* **1998**, *274*, G232–G239. [Google Scholar] [CrossRef]
70. Dhingra, R.; Sullivan, L.; Jacques, P.F.; Wang, T.J.; Fox, C.S.; Meigs, J.B.; D'Agostino, R.B.; Gaziano, J.M.; Vasan, R.S. Soft drink consumption and risk of developing cardiometabolic risk factors and the metabolic syndrome in middle-aged adults in the community. *Circulation* **2007**, *116*, 480–488. [Google Scholar] [CrossRef]
71. Stanhope, K.L.; Schwarz, J.M.; Keim, N.L.; Griffen, S.C.; Bremer, A.A.; Graham, J.L.; Hatcher, B.; Cox, C.L.; Dyachenko, A.; Zhang, W.; et al. Consuming fructose-sweetened, not glucose-sweetened, beverages increases visceral adiposity and lipids and decreases insulin sensitivity in overweight/obese humans. *J. Clin. Investig.* **2009**, *119*, 1322–1334. [Google Scholar] [CrossRef]
72. Teff, K.L.; Grudziak, J.; Townsend, R.R.; Dunn, T.N.; Grant, R.W.; Adams, S.H.; Keim, N.L.; Cummings, B.P.; Stanhope, K.L.; Havel, P.J. Endocrine and metabolic effects of consuming fructose- and glucose-sweetened beverages with meals in obese men and women: Influence of insulin resistance on plasma triglyceride responses. *J. Clin. Endocrinol. Metab.* **2009**, *94*, 1562–1569. [Google Scholar] [CrossRef] [PubMed]

73. Chong, M.F.; Fielding, B.A.; Frayn, K.N. Mechanisms for the acute effect of fructose on postprandial lipemia. *Am. J. Clin. Nutr.* **2007**, *85*, 1511–1520. [Google Scholar] [CrossRef] [PubMed]
74. Donnelly, K.L.; Smith, C.I.; Schwarzenberg, S.J.; Jessurun, J.; Boldt, M.D.; Parks, E.J. Sources of fatty acids stored in liver and secreted via lipoproteins in patients with nonalcoholic fatty liver disease. *J. Clin. Investig.* **2005**, *115*, 1343–1351. [Google Scholar] [CrossRef] [PubMed]
75. Dushay, J.R.; Toschi, E.; Mitten, E.K.; Fisher, F.M.; Herman, M.A.; Maratos-Flier, E. Fructose ingestion acutely stimulates circulating FGF21 levels in humans. *Mol. Metab.* **2015**, *4*, 51–57. [Google Scholar] [CrossRef]
76. Lambert, J.E.; Ramos-Roman, M.A.; Browning, J.D.; Parks, E.J. Increased de novo lipogenesis is a distinct characteristic of individuals with nonalcoholic fatty liver disease. *Gastroenterology* **2014**, *146*, 726–735. [Google Scholar] [CrossRef]
77. Parks, E.J.; Skokan, L.E.; Timlin, M.T.; Dingfelder, C.S. Dietary sugars stimulate fatty acid synthesis in adults. *J. Nutr.* **2008**, *138*, 1039–1046. [Google Scholar] [CrossRef]
78. Softic, S.; Cohen, D.E.; Kahn, C.R. Role of Dietary Fructose and Hepatic De Novo Lipogenesis in Fatty Liver Disease. *Dig. Dis. Sci.* **2016**, *61*, 1282–1293. [Google Scholar] [CrossRef]
79. Sun, S.Z.; Empie, M.W. Fructose metabolism in humans—What isotopic tracer studies tell us. *Nutr. Metab.* **2012**, *9*, 89. [Google Scholar] [CrossRef]
80. Ter Horst, K.W.; Serlie, M.J. Fructose Consumption, Lipogenesis, and Non-Alcoholic Fatty Liver Disease. *Nutrients* **2017**, 9. [Google Scholar] [CrossRef]
81. Theytaz, F.; de Giorgi, S.; Hodson, L.; Stefanoni, N.; Rey, V.; Schneiter, P.; Giusti, V.; Tappy, L. Metabolic fate of fructose ingested with and without glucose in a mixed meal. *Nutrients* **2014**, *6*, 2632–2649. [Google Scholar] [CrossRef]
82. Bizeau, M.E.; Pagliassotti, M.J. Hepatic adaptations to sucrose and fructose. *Metabolism* **2005**, *54*, 1189–1201. [Google Scholar] [CrossRef] [PubMed]
83. Havel, P.J. Dietary fructose: Implications for dysregulation of energy homeostasis and lipid/carbohydrate metabolism. *Nutr. Rev.* **2005**, *63*, 133–157. [Google Scholar] [CrossRef] [PubMed]
84. Taskinen, M.R.; Packard, C.J.; Boren, J. Dietary Fructose and the Metabolic Syndrome. *Nutrients* **2019**, 11. [Google Scholar] [CrossRef] [PubMed]
85. Johnson, R.J.; Sanchez-Lozada, L.G.; Andrews, P.; Lanaspa, M.A. Perspective: A Historical and Scientific Perspective of Sugar and Its Relation with Obesity and Diabetes. *Adv. Nutr.* **2017**, *8*, 412–422. [Google Scholar] [CrossRef]
86. Malik, V.S.; Popkin, B.M.; Bray, G.A.; Despres, J.P.; Willett, W.C.; Hu, F.B. Sugar-sweetened beverages and risk of metabolic syndrome and type 2 diabetes: A meta-analysis. *Diabetes Care* **2010**, *33*, 2477–2483. [Google Scholar] [CrossRef]
87. Tsilas, C.S.; de Souza, R.J.; Mejia, S.B.; Mirrahimi, A.; Cozma, A.I.; Jayalath, V.H.; Ha, V.; Tawfik, R.; Di Buono, M.; Jenkins, A.L.; et al. Relation of total sugars, fructose and sucrose with incident type 2 diabetes: A systematic review and meta-analysis of prospective cohort studies. *CMAJ* **2017**, *189*, E711–E720. [Google Scholar] [CrossRef]
88. Xi, B.; Li, S.; Liu, Z.; Tian, H.; Yin, X.; Huai, P.; Tang, W.; Zhou, D.; Steffen, L.M. Intake of fruit juice and incidence of type 2 diabetes: A systematic review and meta-analysis. *PLoS ONE* **2014**, *9*, e93471. [Google Scholar] [CrossRef]
89. van Buul, V.J.; Tappy, L.; Brouns, F.J. Misconceptions about fructose-containing sugars and their role in the obesity epidemic. *Nutr. Res. Rev.* **2014**, *27*, 119–130. [Google Scholar] [CrossRef]
90. Caliceti, C.; Calabria, D.; Roda, A.; Cicero, A.F.G. Fructose Intake, Serum Uric Acid, and Cardiometabolic Disorders: A Critical Review. *Nutrients* **2017**, 9. [Google Scholar] [CrossRef]
91. Jegatheesan, P.; De Bandt, J.P. Fructose and NAFLD: The Multifaceted Aspects of Fructose Metabolism. *Nutrients* **2017**, 9. [Google Scholar] [CrossRef]
92. Jang, C.; Hui, S.; Lu, W.; Cowan, A.J.; Morscher, R.J.; Lee, G.; Liu, W.; Tesz, G.J.; Birnbaum, M.J.; Rabinowitz, J.D. The Small Intestine Converts Dietary Fructose into Glucose and Organic Acids. *Cell Metab.* **2018**, *27*, 351–361. [Google Scholar] [CrossRef] [PubMed]
93. Casteleyn, C.; Rekecki, A.; Van der Aa, A.; Simoens, P.; Van den Broeck, W. Surface area assessment of the murine intestinal tract as a prerequisite for oral dose translation from mouse to man. *Lab. Anim.* **2010**, *44*, 176–183. [Google Scholar] [CrossRef] [PubMed]
94. Rajas, F.; Bruni, N.; Montano, S.; Zitoun, C.; Mithieux, G. The glucose-6 phosphatase gene is expressed in human and rat small intestine: Regulation of expression in fasted and diabetic rats. *Gastroenterology* **1999**, *117*, 132–139. [Google Scholar] [CrossRef]
95. Ockerman, P.A.; Lundborg, H. Conversion of fructose to glucose by human jejunum absence of galactose-to-glucose conversion. *Biochim. Biophys. Acta* **1965**, *105*, 34–42. [Google Scholar] [CrossRef]
96. Hwang, J.J.; Johnson, A.; Cline, G.; Belfort-DeAguiar, R.; Snegovskikh, D.; Khokhar, B.; Han, C.S.; Sherwin, R.S. Fructose levels are markedly elevated in cerebrospinal fluid compared to plasma in pregnant women. *PLoS ONE* **2015**, *10*, e0128582. [Google Scholar] [CrossRef]
97. Francey, C.; Cros, J.; Rosset, R.; Creze, C.; Rey, V.; Stefanoni, N.; Schneiter, P.; Tappy, L.; Seyssel, K. The extra-splanchnic fructose escape after ingestion of a fructose-glucose drink: An exploratory study in healthy humans using a dual fructose isotope method. *Clin. Nutr. Espen.* **2019**, *29*, 125–132. [Google Scholar] [CrossRef]

98. Hwang, J.J.; Jiang, L.; Hamza, M.; Dai, F.; Belfort-DeAguiar, R.; Cline, G.; Rothman, D.L.; Mason, G.; Sherwin, R.S. The human brain produces fructose from glucose. *JCI Insight* **2017**, *2*, e90508. [Google Scholar] [CrossRef]
99. Oppelt, S.A.; Zhang, W.; Tolan, D.R. Specific regions of the brain are capable of fructose metabolism. *Brain Res.* **2017**, *1657*, 312–322. [Google Scholar] [CrossRef]
100. Song, Z.; Roncal-Jimenez, C.A.; Lanaspa-Garcia, M.A.; Oppelt, S.A.; Kuwabara, M.; Jensen, T.; Milagres, T.; Andres-Hernando, A.; Ishimoto, T.; Garcia, G.E.; et al. Role of fructose and fructokinase in acute dehydration-induced vasopressin gene expression and secretion in mice. *J. Neurophysiol.* **2017**, *117*, 646–654. [Google Scholar] [CrossRef]
101. Yan, L.J. Redox imbalance stress in diabetes mellitus: Role of the polyol pathway. *Anim. Model. Exp. Med.* **2018**, *1*, 7–13. [Google Scholar] [CrossRef]
102. Lanaspa, M.A.; Kuwabara, M.; Andres-Hernando, A.; Li, N.; Cicerchi, C.; Jensen, T.; Orlicky, D.J.; Roncal-Jimenez, C.A.; Ishimoto, T.; Nakagawa, T.; et al. High salt intake causes leptin resistance and obesity in mice by stimulating endogenous fructose production and metabolism. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 3138–3143. [Google Scholar] [CrossRef] [PubMed]
103. Lanaspa, M.A.; Ishimoto, T.; Cicerchi, C.; Tamura, Y.; Roncal-Jimenez, C.A.; Chen, W.; Tanabe, K.; Andres-Hernando, A.; Orlicky, D.J.; Finol, E.; et al. Endogenous fructose production and fructokinase activation mediate renal injury in diabetic nephropathy. *J. Am. Soc. Nephrol.* **2014**, *25*, 2526–2538. [Google Scholar] [CrossRef] [PubMed]
104. Johnson, R.J.; Bakris, G.L.; Borghi, C.; Chonchol, M.B.; Feldman, D.; Lanaspa, M.A.; Merriman, T.R.; Moe, O.W.; Mount, D.B.; Sanchez Lozada, L.G.; et al. Hyperuricemia, Acute and Chronic Kidney Disease, Hypertension, and Cardiovascular Disease: Report of a Scientific Workshop Organized by the National Kidney Foundation. *Am. J. Kidney Dis.* **2018**, *71*, 851–865. [Google Scholar] [CrossRef] [PubMed]
105. Slavic, K.; Derbyshire, E.T.; Naftalin, R.J.; Krishna, S.; Staines, H.M. Comparison of effects of green tea catechins on apicomplexan hexose transporters and mammalian orthologues. *Mol. Biochem. Parasitol.* **2009**, *168*, 113–116. [Google Scholar] [CrossRef] [PubMed]
106. Villa-Rodriguez, J.A.; Aydin, E.; Gauer, J.S.; Pyner, A.; Williamson, G.; Kerimi, A. Green and Chamomile Teas, but not Acarbose, Attenuate Glucose and Fructose Transport via Inhibition of GLUT2 and GLUT5. *Mol. Nutr. Food Res.* **2017**, 61. [Google Scholar] [CrossRef]
107. Satsu, H.; Awara, S.; Unno, T.; Shimizu, M. Suppressive effect of nobiletin and epicatechin gallate on fructose uptake in human intestinal epithelial Caco-2 cells. *Biosci. Biotechnol. Biochem.* **2018**, *82*, 636–646. [Google Scholar] [CrossRef]
108. George Thompson, A.M.; Ursu, O.; Babkin, P.; Iancu, C.V.; Whang, A.; Oprea, T.I.; Choe, J.Y. Discovery of a specific inhibitor of human GLUT5 by virtual screening and in vitro transport evaluation. *Sci. Rep.* **2016**, *6*, 24240. [Google Scholar] [CrossRef]
109. Tripp, J.; Essl, C.; Iancu, C.V.; Boles, E.; Choe, J.Y.; Oreb, M. Establishing a yeast-based screening system for discovery of human GLUT5 inhibitors and activators. *Sci. Rep.* **2017**, *7*, 6197. [Google Scholar] [CrossRef]
110. Lee, Y.; Lim, Y.; Kwon, O. Selected Phytochemicals and Culinary Plant Extracts Inhibit Fructose Uptake in Caco-2 Cells. *Molecules* **2015**, *20*, 17393–17404. [Google Scholar] [CrossRef]
111. Muller, U.; Stubl, F.; Schwarzinger, B.; Sandner, G.; Iken, M.; Himmelsbach, M.; Schwarzinger, C.; Ollinger, N.; Stadlbauer, V.; Hoglinger, O.; et al. In Vitro and In Vivo Inhibition of Intestinal Glucose Transport by Guava (Psidium Guajava) Extracts. *Mol. Nutr. Food Res.* **2018**, *62*, e1701012. [Google Scholar] [CrossRef]
112. Konig, A.; Schwarzinger, B.; Stadlbauer, V.; Lanzerstorfer, P.; Iken, M.; Schwarzinger, C.; Kolb, P.; Schwarzinger, S.; Morwald, K.; Brunner, S.; et al. Guava (*Psidium guajava*) Fruit Extract Prepared by Supercritical CO2 Extraction Inhibits Intestinal Glucose Resorption in a Double-Blind, Randomized Clinical Study. *Nutrients* **2019**, 11. [Google Scholar] [CrossRef] [PubMed]
113. Kerimi, A.; Gauer, J.S.; Crabbe, S.; Cheah, J.W.; Lau, J.; Walsh, R.; Cancalon, P.F.; Williamson, G. Effect of the flavonoid hesperidin on glucose and fructose transport, sucrase activity and glycaemic response to orange juice in a crossover trial on healthy volunteers. *Br. J. Nutr.* **2019**, *121*, 782–792. [Google Scholar] [CrossRef] [PubMed]
114. Gauer, J.S.; Tumova, S.; Lippiat, J.D.; Kerimi, A.; Williamson, G. Differential patterns of inhibition of the sugar transporters GLUT2, GLUT5 and GLUT7 by flavonoids. *Biochem. Pharm.* **2018**, *152*, 11–20. [Google Scholar] [CrossRef] [PubMed]
115. Chiasson, J.L.; Josse, R.G.; Gomis, R.; Hanefeld, M.; Karasik, A.; Laakso, M.; Group, S.-N.T.R. Acarbose for prevention of type 2 diabetes mellitus: The STOP-NIDDM randomised trial. *Lancet* **2002**, *359*, 2072–2077. [Google Scholar] [CrossRef]
116. Madariaga, H.; Lee, P.C.; Heitlinger, L.A.; Lebenthal, E. Effects of graded alpha-glucosidase inhibition on sugar absorption in vivo. *Dig. Dis Sci.* **1988**, *33*, 1020–1024. [Google Scholar] [CrossRef]
117. Rorsman, P.; Braun, M. Regulation of insulin secretion in human pancreatic islets. *Annu. Rev. Physiol.* **2013**, *75*, 155–179. [Google Scholar] [CrossRef]
118. Quesada, I.; Tuduri, E.; Ripoll, C.; Nadal, A. Physiology of the pancreatic alpha-cell and glucagon secretion: Role in glucose homeostasis and diabetes. *J. Endocrinol.* **2008**, *199*, 5–19. [Google Scholar] [CrossRef]
119. Rorsman, P.; Huising, M.O. The somatostatin-secreting pancreatic delta-cell in health and disease. *Nat. Rev. Endocrinol.* **2018**, *14*, 404–414. [Google Scholar] [CrossRef]
120. Marty, N.; Dallaporta, M.; Thorens, B. Brain glucose sensing, counterregulation, and energy homeostasis. *Physiology* **2007**, *22*, 241–251. [Google Scholar] [CrossRef]

121. Kellett, G.L.; Brot-Laroche, E.; Mace, O.J.; Leturque, A. Sugar absorption in the intestine: The role of GLUT2. *Annu. Rev. Nutr.* **2008**, *28*, 35–54. [Google Scholar] [CrossRef]
122. Wright, E.M.; Hirayama, B.A.; Loo, D.F. Active sugar transport in health and disease. *J. Intern. Med.* **2007**, *261*, 32–43. [Google Scholar] [CrossRef] [PubMed]
123. Hediger, M.A.; Coady, M.J.; Ikeda, T.S.; Wright, E.M. Expression cloning and cDNA sequencing of the Na+/glucose co-transporter. *Nature* **1987**, *330*, 379–381. [Google Scholar] [CrossRef] [PubMed]
124. Chen, J.; Williams, S.; Ho, S.; Loraine, H.; Hagan, D.; Whaley, J.M.; Feder, J.N. Quantitative PCR tissue expression profiling of the human SGLT2 gene and related family members. *Diabetes* **2010**, *1*, 57–92. [Google Scholar] [CrossRef] [PubMed]
125. Balakrishnan, A.; Stearns, A.T.; Ashley, S.W.; Rhoads, D.B.; Tavakkolizadeh, A. PER1 modulates SGLT1 transcription in vitro independent of E-box status. *Dig. Dis. Sci.* **2012**, *57*, 1525–1536. [Google Scholar] [CrossRef] [PubMed]
126. Balakrishnan, A.; Stearns, A.T.; Ashley, S.W.; Tavakkolizadeh, A.; Rhoads, D.B. Restricted feeding phase shifts clock gene and sodium glucose cotransporter 1 (SGLT1) expression in rats. *J. Nutr.* **2010**, *140*, 908–914. [Google Scholar] [CrossRef] [PubMed]
127. Uldry, M.; Ibberson, M.; Hosokawa, M.; Thorens, B. GLUT2 is a high affinity glucosamine transporter. *FEBS Lett.* **2002**, *524*, 199–203. [Google Scholar] [CrossRef]
128. Solberg, D.H.; Diamond, J.M. Comparison of different dietary sugars as inducers of intestinal sugar transporters. *Am. J. Physiol.* **1987**, *252*, G574–G584. [Google Scholar] [CrossRef]
129. Shirazi-Beechey, S.P.; Hirayama, B.A.; Wang, Y.; Scott, D.; Smith, M.W.; Wright, E.M. Ontogenic development of lamb intestinal sodium-glucose co-transporter is regulated by diet. *J. Physiol.* **1991**, *437*, 699–708. [Google Scholar] [CrossRef]
130. Dyer, J.; Hosie, K.B.; Shirazi-Beechey, S.P. Nutrient regulation of human intestinal sugar transporter (SGLT1) expression. *Gut* **1997**, *41*, 56–59. [Google Scholar] [CrossRef]
131. Shirazi-Beechey, S.P.; Smith, M.W.; Wang, Y.; James, P.S. Postnatal development of lamb intestinal digestive enzymes is not regulated by diet. *J. Physiol.* **1991**, *437*, 691–698. [Google Scholar] [CrossRef]
132. Moran, A.W.; Al-Rammahi, M.A.; Arora, D.K.; Batchelor, D.J.; Coulter, E.A.; Ionescu, C.; Bravo, D.; Shirazi-Beechey, S.P. Expression of Na+/glucose co-transporter 1 (SGLT1) in the intestine of piglets weaned to different concentrations of dietary carbohydrate. *Br. J. Nutr.* **2010**, *104*, 647–655. [Google Scholar] [CrossRef] [PubMed]
133. Hirsch, J.R.; Loo, D.D.; Wright, E.M. Regulation of Na+/glucose cotransporter expression by protein kinases in Xenopus laevis oocytes. *J. Biol. Chem.* **1996**, *271*, 14740–14746. [Google Scholar] [CrossRef] [PubMed]
134. Wright, E.M.; Hirsch, J.R.; Loo, D.D.; Zampighi, G.A. Regulation of Na+/glucose cotransporters. *J. Exp. Biol.* **1997**, *200*, 287–293. [Google Scholar] [PubMed]
135. Subramanian, S.; Glitz, P.; Kipp, H.; Kinne, R.K.; Castaneda, F. Protein kinase-A affects sorting and conformation of the sodium-dependent glucose co-transporter SGLT1. *J. Cell Biochem.* **2009**, *106*, 444–452. [Google Scholar] [CrossRef] [PubMed]
136. Ghezzi, C.; Wright, E.M. Regulation of the human Na^+-dependent glucose cotransporter hSGLT2. *Am. J. Physiol. Cell Physiol.* **2012**, *303*, C348–C354. [Google Scholar] [CrossRef] [PubMed]
137. Veyhl, M.; Wagner, C.A.; Gorboulev, V.; Schmitt, B.M.; Lang, F.; Koepsell, H. Downregulation of the Na^+-D-glucose cotransporter SGLT1 by protein RS1 (RSC1A1) is dependent on dynamin and protein kinase C. *J. Membr. Biol.* **2003**, *196*, 71–81. [Google Scholar] [CrossRef] [PubMed]
138. Poulsen, S.B.; Fenton, R.A.; Rieg, T. Sodium-glucose cotransport. *Curr. Opin. Nephrol. Hypertens.* **2015**, *24*, 463–469. [Google Scholar] [CrossRef]
139. Ducroc, R.; Guilmeau, S.; Akasbi, K.; Devaud, H.; Buyse, M.; Bado, A. Luminal leptin induces rapid inhibition of active intestinal absorption of glucose mediated by sodium-glucose cotransporter 1. *Diabetes* **2005**, *54*, 348–354. [Google Scholar] [CrossRef]
140. Kobayashi, Y.; Suzuki, M.; Satsu, H.; Arai, S.; Hara, Y.; Suzuki, K.; Miyamoto, Y.; Shimizu, M. Green tea polyphenols inhibit the sodium-dependent glucose transporter of intestinal epithelial cells by a competitive mechanism. *J. Agric. Food Chem.* **2000**, *48*, 5618–5623. [Google Scholar] [CrossRef]
141. Gouyon, F.; Caillaud, L.; Carriere, V.; Klein, C.; Dalet, V.; Citadelle, D.; Kellett, G.L.; Thorens, B.; Leturque, A.; Brot-Laroche, E. Simple-sugar meals target GLUT2 at enterocyte apical membranes to improve sugar absorption: A study in GLUT2-null mice. *J. Physiol.* **2003**, *552*, 823–832. [Google Scholar] [CrossRef]
142. Chaudhry, R.M.; Scow, J.S.; Madhavan, S.; Duenes, J.A.; Sarr, M.G. Acute enterocyte adaptation to luminal glucose: A posttranslational mechanism for rapid apical recruitment of the transporter GLUT2. *J. Gastrointest. Surg.* **2012**, *16*, 312–319; discussion 319. [Google Scholar] [CrossRef] [PubMed]
143. Zheng, Y.; Scow, J.S.; Duenes, J.A.; Sarr, M.G. Mechanisms of glucose uptake in intestinal cell lines: Role of GLUT2. *Surgery* **2012**, *151*, 13–25. [Google Scholar] [CrossRef] [PubMed]
144. Tobin, V.; Le Gall, M.; Fioramonti, X.; Stolarczyk, E.; Blazquez, A.G.; Klein, C.; Prigent, M.; Serradas, P.; Cuif, M.H.; Magnan, C.; et al. Insulin internalizes GLUT2 in the enterocytes of healthy but not insulin-resistant mice. *Diabetes* **2008**, *57*, 555–562. [Google Scholar] [CrossRef] [PubMed]
145. Ait-Omar, A.; Monteiro-Sepulveda, M.; Poitou, C.; Le Gall, M.; Cotillard, A.; Gilet, J.; Garbin, K.; Houllier, A.; Chateau, D.; Lacombe, A.; et al. GLUT2 accumulation in enterocyte apical and intracellular membranes: A study in morbidly obese human subjects and ob/ob

and high fat-fed mice. *Diabetes* **2011**, *60*, 2598–2607. [Google Scholar] [CrossRef]
146. Triplitt, C.L. Examining the mechanisms of glucose regulation. *Am. J. Manag. Care* **2012**, *18*, S4–S10. [Google Scholar]
147. Jenkins, D.J.; Wolever, T.M.; Taylor, R.H.; Barker, H.; Fielden, H.; Baldwin, J.M.; Bowling, A.C.; Newman, H.C.; Jenkins, A.L.; Goff, D.V. Glycemic index of foods: A physiological basis for carbohydrate exchange. *Am. J. Clin. Nutr.* **1981**, *34*, 362–366. [Google Scholar] [CrossRef]
148. Salmeron, J.; Manson, J.E.; Stampfer, M.J.; Colditz, G.A.; Wing, A.L.; Willett, W.C. Dietary fiber, glycemic load, and risk of non-insulin-dependent diabetes mellitus in women. *JAMA* **1997**, *277*, 472–477. [Google Scholar] [CrossRef]
149. Bantle, J.P.; Wylie-Rosett, J.; Albright, A.L.; Apovian, C.M.; Clark, N.G.; Franz, M.J.; Hoogwerf, B.J.; Lichtenstein, A.H.; Mayer-Davis, E.; Mooradian, A.D.; et al. Nutrition recommendations and interventions for diabetes: A position statement of the American Diabetes Association. *Diabetes Care* **2008**, *31* (Suppl. S1), S61–S78. [Google Scholar] [CrossRef]
150. Livesey, G.; Taylor, R.; Livesey, H.; Liu, S. Is there a dose-response relation of dietary glycemic load to risk of type 2 diabetes? Meta-analysis of prospective cohort studies. *Am. J. Clin. Nutr.* **2015**, *97*, 584–596. [Google Scholar] [CrossRef]
151. Greenwood, D.C.; Threapleton, D.E.; Evans, C.E.; Cleghorn, C.L.; Nykjaer, C.; Woodhead, C.; Burley, V.J. Glycemic index, glycemic load, carbohydrates, and type 2 diabetes: Systematic review and dose-response meta-analysis of prospective studies. *Diabetes Care* **2013**, *36*, 4166–4171. [Google Scholar] [CrossRef]
152. Dong, J.Y.; Zhang, L.; Zhang, Y.H.; Qin, L.Q. Dietary glycaemic index and glycaemic load in relation to the risk of type 2 diabetes: A meta-analysis of prospective cohort studies. *Br. J. Nutr.* **2011**, *106*, 1649–1654. [Google Scholar] [CrossRef] [PubMed]
153. Barclay, A.W.; Petocz, P.; McMillan-Price, J.; Flood, V.M.; Prvan, T.; Mitchell, P.; Brand-Miller, J.C. Glycemic index, glycemic load, and chronic disease risk—A meta-analysis of observational studies. *Am. J. Clin. Nutr.* **2008**, *87*, 627–637. [Google Scholar] [CrossRef] [PubMed]
154. Furman, B.L. Streptozotocin-Induced Diabetic Models in Mice and Rats. *Curr. Protoc. Pharm.* **2015**, *70*, 5–47. [Google Scholar] [CrossRef] [PubMed]
155. Miyamoto, K.; Hase, K.; Taketani, Y.; Minami, H.; Oka, T.; Nakabou, Y.; Hagihira, H. Diabetes and glucose transporter gene expression in rat small intestine. *Biochem. Biophys. Res. Commun.* **1991**, *181*, 1110–1117. [Google Scholar] [CrossRef]
156. Fujita, Y.; Kojima, H.; Hidaka, H.; Fujimiya, M.; Kashiwagi, A.; Kikkawa, R. Increased intestinal glucose absorption and postprandial hyperglycaemia at the early step of glucose intolerance in Otsuka Long-Evans Tokushima Fatty rats. *Diabetologia* **1998**, *41*, 1459–1466. [Google Scholar] [CrossRef]
157. Dyer, J.; Wood, I.S.; Palejwala, A.; Ellis, A.; Shirazi-Beechey, S.P. Expression of monosaccharide transporters in intestine of diabetic humans. *Am. J. Physiol. Gastrointest. Liver Physiol.* **2002**, *282*, G241–G248. [Google Scholar] [CrossRef]
158. Nguyen, N.Q.; Debreceni, T.L.; Bambrick, J.E.; Chia, B.; Wishart, J.; Deane, A.M.; Rayner, C.K.; Horowitz, M.; Young, R.L. Accelerated intestinal glucose absorption in morbidly obese humans: Relationship to glucose transporters, incretin hormones, and glycemia. *J. Clin. Endocrinol. Metab.* **2015**, *100*, 968–976. [Google Scholar] [CrossRef]
159. Song, P.; Onishi, A.; Koepsell, H.; Vallon, V. Sodium glucose cotransporter SGLT1 as a therapeutic target in diabetes mellitus. *Expert Opin. Targets* **2016**, *20*, 1109–1125. [Google Scholar] [CrossRef]
160. Smith, G.P. Pavlov and integrative physiology. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **2000**, *279*, R743–R755. [Google Scholar] [CrossRef]
161. Carabotti, M.; Scirocco, A.; Maselli, M.A.; Severi, C. The gut-brain axis: Interactions between enteric microbiota, central and enteric nervous systems. *Ann. Gastroenterol.* **2015**, *28*, 203–209. [Google Scholar]
162. urness, J.B.; Rivera, L.R.; Cho, H.J.; Bravo, D.M.; Callaghan, B. The gut as a sensory organ. *Nat. Rev. Gastroenterol. Hepatol.* **2013**, *10*, 729–740. [Google Scholar] [CrossRef] [PubMed]
163. Mayer, E.A. Gut feelings: The emerging biology of gut-brain communication. *Nat. Rev. Neurosci.* **2011**, *12*, 453–466. [Google Scholar] [CrossRef] [PubMed]
164. Chaudhri, O.B.; Salem, V.; Murphy, K.G.; Bloom, S.R. Gastrointestinal satiety signals. *Annu. Rev. Physiol.* **2008**, *70*, 239–255. [Google Scholar] [CrossRef] [PubMed]
165. Holzer, P.; Reichmann, F.; Farzi, A. Neuropeptide Y, peptide YY and pancreatic polypeptide in the gut-brain axis. *Neuropeptides* **2012**, *46*, 261–274. [Google Scholar] [CrossRef]
166. Mace, O.J.; Tehan, B.; Marshall, F. Pharmacology and physiology of gastrointestinal enteroendocrine cells. *Pharm. Res. Perspect.* **2015**, *3*, e00155. [Google Scholar] [CrossRef]
167. Li, H.J.; Ray, S.K.; Singh, N.K.; Johnston, B.; Leiter, A.B. Basic helix-loop-helix transcription factors and enteroendocrine cell differentiation. *Diabetes Obes Metab* **2011**, *13* (Suppl. S1), 5–12. [Google Scholar] [CrossRef]
168. Artavanis-Tsakonas, S.; Rand, M.D.; Lake, R.J. Notch signaling: Cell fate control and signal integration in development. *Science* **1999**, *284*, 770–776. [Google Scholar] [CrossRef]
169. Gribble, F.M.; Reimann, F. Enteroendocrine Cells: Chemosensors in the Intestinal Epithelium. *Annu. Rev. Physiol.* **2016**, *78*, 277–299. [Google Scholar] [CrossRef]
170. Hofer, D.; Asan, E.; Drenckhahn, D. Chemosensory Perception in the Gut. *News Physiol. Sci.* **1999**, *14*, 18–23. [Google Scholar] [CrossRef]

171. Sternini, C.; Anselmi, L.; Rozengurt, E. Enteroendocrine cells: A site of taste in gastrointestinal chemosensing. *Curr. Opin. Endocrinol. Diabetes Obes.* **2008**, *15*, 73–78. [Google Scholar] [CrossRef]
172. Psichas, A.; Reimann, F.; Gribble, F.M. Gut chemosensing mechanisms. *J. Clin. Investig.* **2015**, *125*, 908–917. [Google Scholar] [CrossRef] [PubMed]
173. Kishida, K.; Pearce, S.C.; Yu, S.; Gao, N.; Ferraris, R.P. Nutrient sensing by absorptive and secretory progenies of small intestinal stem cells. *Am. J. Physiol. Gastrointest. Liver Physiol.* **2017**, *312*, G592–G605. [Google Scholar] [CrossRef] [PubMed]
174. Kuhre, R.E.; Gribble, F.M.; Hartmann, B.; Reimann, F.; Windelov, J.A.; Rehfeld, J.F.; Holst, J.J. Fructose stimulates GLP-1 but not GIP secretion in mice, rats, and humans. *Am. J. Physiol. Gastrointest. Liver Physiol.* **2014**, *306*, G622–G630. [Google Scholar] [CrossRef] [PubMed]
175. Martin, A.M.; Lumsden, A.L.; Young, R.L.; Jessup, C.F.; Spencer, N.J.; Keating, D.J. Regional differences in nutrient-induced secretion of gut serotonin. *Physiol. Rep.* **2017**, 5. [Google Scholar] [CrossRef] [PubMed]
176. Seino, Y.; Ogata, H.; Maekawa, R.; Izumoto, T.; Iida, A.; Harada, N.; Miki, T.; Seino, S.; Inagaki, N.; Tsunekawa, S.; et al. Fructose induces glucose-dependent insulinotropic polypeptide, glucagon-like peptide-1 and insulin secretion: Role of adenosine triphosphate-sensitive K^+ channels. *J. Diabetes Investig.* **2015**, *6*, 522–526. [Google Scholar] [CrossRef] [PubMed]
177. Yau, A.M.; McLaughlin, J.; Gilmore, W.; Maughan, R.J.; Evans, G.H. The Acute Effects of Simple Sugar Ingestion on Appetite, Gut-Derived Hormone Response, and Metabolic Markers in Men. *Nutrients* **2017**, 9. [Google Scholar] [CrossRef] [PubMed]
178. Nauck, M.A.; Meier, J.J. The incretin effect in healthy individuals and those with type 2 diabetes: Physiology, pathophysiology, and response to therapeutic interventions. *Lancet Diabetes Endocrinol.* **2016**, *4*, 525–536. [Google Scholar] [CrossRef]
179. Creutzfeldt, W. The incretin concept today. *Diabetologia* **1979**, *16*, 75–85. [Google Scholar] [CrossRef]
180. Drucker, D.J. Glucagon-like peptide 2. *J. Clin. Endocrinol. Metab.* **2001**, *86*, 1759–1764. [Google Scholar] [CrossRef]
181. Roder, P.V.; Geillinger, K.E.; Zietek, T.S.; Thorens, B.; Koepsell, H.; Daniel, H. The role of SGLT1 and GLUT2 in intestinal glucose transport and sensing. *PLoS ONE* **2014**, *9*, e89977. [Google Scholar] [CrossRef]
182. Bell, G.I.; Sanchez-Pescador, R.; Laybourn, P.J.; Najarian, R.C. Exon duplication and divergence in the human preproglucagon gene. *Nature* **1983**, *304*, 368–371. [Google Scholar] [CrossRef] [PubMed]
183. Holst, J.J.; Orskov, C.; Nielsen, O.V.; Schwartz, T.W. Truncated glucagon-like peptide I, an insulin-releasing hormone from the distal gut. *FEBS Lett.* **1987**, *211*, 169–174. [Google Scholar] [CrossRef]
184. Mojsov, S.; Weir, G.C.; Habener, J.F. Insulinotropin: Glucagon-like peptide I (7–37) co-encoded in the glucagon gene is a potent stimulator of insulin release in the perfused rat pancreas. *J. Clin. Investig.* **1987**, *79*, 616–619. [Google Scholar] [CrossRef] [PubMed]
185. Drucker, D.J. The biology of incretin hormones. *Cell Metab.* **2006**, *3*, 153–165. [Google Scholar] [CrossRef] [PubMed]
186. Beattie, D.T.; Smith, J.A. Serotonin pharmacology in the gastrointestinal tract: A review. *Naunyn Schmiedebergs Arch. Pharm.* **2008**, *377*, 181–203. [Google Scholar] [CrossRef] [PubMed]
187. Sjolund, K.; Sanden, G.; Hakanson, R.; Sundler, F. Endocrine cells in human intestine: An immunocytochemical study. *Gastroenterology* **1983**, *85*, 1120–1130. [Google Scholar] [PubMed]
188. Imeryuz, N.; Yegen, B.C.; Bozkurt, A.; Coskun, T.; Villanueva-Penacarrillo, M.L.; Ulusoy, N.B. Glucagon-like peptide-1 inhibits gastric emptying via vagal afferent-mediated central mechanisms. *Am. J. Physiol.* **1997**, *273*, G920–G927. [Google Scholar] [CrossRef]
189. Raybould, H.E.; Glatzle, J.; Robin, C.; Meyer, J.H.; Phan, T.; Wong, H.; Sternini, C. Expression of 5-HT3 receptors by extrinsic duodenal afferents contribute to intestinal inhibition of gastric emptying. *Am. J. Physiol. Gastrointest. Liver Physiol.* **2003**, *284*, G367–G372. [Google Scholar] [CrossRef]
190. Young, R.L.; Chia, B.; Isaacs, N.J.; Ma, J.; Khoo, J.; Wu, T.; Horowitz, M.; Rayner, C.K. Disordered control of intestinal sweet taste receptor expression and glucose absorption in type 2 diabetes. *Diabetes* **2013**, *62*, 3532–3541. [Google Scholar] [CrossRef]
191. Laffitte, A.; Neiers, F.; Briand, L. Functional roles of the sweet taste receptor in oral and extraoral tissues. *Curr. Opin. Clin. Nutr. Metab. Care* **2014**, *17*, 379–385. [Google Scholar] [CrossRef]
192. Rother, K.I.; Conway, E.M.; Sylvetsky, A.C. How Non-nutritive Sweeteners Influence Hormones and Health. *Trends Endocrinol. Metab.* **2018**, *29*, 455–467. [Google Scholar] [CrossRef] [PubMed]
193. Dyer, J.; Salmon, K.S.; Zibrik, L.; Shirazi-Beechey, S.P. Expression of sweet taste receptors of the T1R family in the intestinal tract and enteroendocrine cells. *Biochem. Soc. Trans* **2005**, *33*, 302–305. [Google Scholar] [CrossRef] [PubMed]
194. Nelson, G.; Hoon, M.A.; Chandrashekar, J.; Zhang, Y.; Ryba, N.J.; Zuker, C.S. Mammalian sweet taste receptors. *Cell* **2001**, *106*, 381–390. [Google Scholar] [CrossRef]
195. Thompson, M.D.; Cole, D.E.; Jose, P.A.; Chidiac, P. G protein-coupled receptor accessory proteins and signaling: Pharmacogenomic insights. *Methods Mol. Biol.* **2014**, *1175*, 121–152. [Google Scholar] [CrossRef] [PubMed]
196. Yarmolinsky, D.A.; Zuker, C.S.; Ryba, N.J. Common sense about taste: From mammals to insects. *Cell* **2009**, *139*, 234–244. [Google Scholar] [CrossRef] [PubMed]
197. Liu, D.; Liman, E.R. Intracellular Ca2+ and the phospholipid PIP2 regulate the taste transduction ion channel TRPM5. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 15160–15165. [Google Scholar] [CrossRef] [PubMed]
198. Margolskee, R.F.; Dyer, J.; Kokrashvili, Z.; Salmon, K.S.; Ilegems, E.; Daly, K.; Maillet, E.L.; Ninomiya, Y.; Mosinger, B.; Shirazi-Beechey, S.P. T1R3 and gustducin in gut sense sugars to regulate expression of Na^+-glucose cotransporter 1. *Proc. Natl. Acad. Sci. USA*

2007, *104*, 15075–15080. [Google Scholar] [CrossRef]
199. Moran, A.W.; Al-Rammahi, M.A.; Arora, D.K.; Batchelor, D.J.; Coulter, E.A.; Daly, K.; Ionescu, C.; Bravo, D.; Shirazi-Beechey, S.P. Expression of Na+/glucose co-transporter 1 (SGLT1) is enhanced by supplementation of the diet of weaning piglets with artificial sweeteners. *Br. J. Nutr.* **2010**, *104*, 637–646. [Google Scholar] [CrossRef]
200. Parker, H.E.; Reimann, F.; Gribble, F.M. Molecular mechanisms underlying nutrient-stimulated incretin secretion. *Expert Rev. Mol. Med.* **2010**, *12*, e1. [Google Scholar] [CrossRef]
201. Kokrashvili, Z.; Mosinger, B.; Margolskee, R.F. T1r3 and alpha-gustducin in gut regulate secretion of glucagon-like peptide-1. *Ann. N. Y. Acad. Sci.* **2009**, *1170*, 91–94. [Google Scholar] [CrossRef]
202. Baldassano, S.; Liu, S.; Qu, M.H.; Mule, F.; Wood, J.D. Glucagon-like peptide-2 modulates neurally evoked mucosal chloride secretion in guinea pig small intestine in vitro. *Am. J. Physiol. Gastrointest. Liver Physiol.* **2009**, *297*, G800–G805. [Google Scholar] [CrossRef] [PubMed]
203. Bjerknes, M.; Cheng, H. Modulation of specific intestinal epithelial progenitors by enteric neurons. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 12497–12502. [Google Scholar] [CrossRef] [PubMed]
204. Sato, S.; Hokari, R.; Kurihara, C.; Sato, H.; Narimatsu, K.; Hozumi, H.; Ueda, T.; Higashiyama, M.; Okada, Y.; Watanabe, C.; et al. Dietary lipids and sweeteners regulate glucagon-like peptide-2 secretion. *Am. J. Physiol. Gastrointest. Liver Physiol.* **2013**, *304*, G708–G714. [Google Scholar] [CrossRef] [PubMed]
205. Cheeseman, C.I. Upregulation of SGLT-1 transport activity in rat jejunum induced by GLP-2 infusion in vivo. *Am. J. Physiol.* **1997**, *273*, R1965–R1971. [Google Scholar] [CrossRef]
206. Ramsanahie, A.; Duxbury, M.S.; Grikscheit, T.C.; Perez, A.; Rhoads, D.B.; Gardner-Thorpe, J.; Ogilvie, J.; Ashley, S.W.; Vacanti, J.P.; Whang, E.E. Effect of GLP-2 on mucosal morphology and SGLT1 expression in tissue-engineered neointestine. *Am. J. Physiol. Gastrointest. Liver Physiol.* **2003**, *285*, G1345–G1352. [Google Scholar] [CrossRef]
207. Pal, A.; Rhoads, D.B.; Tavakkoli, A. Foregut exclusion disrupts intestinal glucose sensing and alters portal nutrient and hormonal milieu. *Diabetes* **2015**, *64*, 1941–1950. [Google Scholar] [CrossRef]
208. Johnson, R.J.; Nakagawa, T.; Sanchez-Lozada, L.G.; Shafiu, M.; Sundaram, S.; Le, M.; Ishimoto, T.; Sautin, Y.Y.; Lanaspa, M.A. Sugar, uric acid, and the etiology of diabetes and obesity. *Diabetes* **2013**, *62*, 3307–3315. [Google Scholar] [CrossRef]
209. McDevitt, R.M.; Bott, S.J.; Harding, M.; Coward, W.A.; Bluck, L.J.; Prentice, A.M. De novo lipogenesis during controlled overfeeding with sucrose or glucose in lean and obese women. *Am. J. Clin. Nutr.* **2001**, *74*, 737–746. [Google Scholar] [CrossRef]
210. Ventura, E.E.; Davis, J.N.; Goran, M.I. Sugar content of popular sweetened beverages based on objective laboratory analysis: Focus on fructose content. *Obesity* **2011**, *19*, 868–874. [Google Scholar] [CrossRef]
211. Anderson, G.H.; Aziz, A.; Abou Samra, R. Physiology of food intake regulation: Interaction with dietary components. In Nestle Nutrition workshop series. *Pediatr. Program.* **2006**, *58*, 133–143. [Google Scholar] [CrossRef]
212. Heijboer, A.C.; Pijl, H.; Van den Hoek, A.M.; Havekes, L.M.; Romijn, J.A.; Corssmit, E.P. Gut-brain axis: Regulation of glucose metabolism. *J. Neuroendocr.* **2006**, *18*, 883–894. [Google Scholar] [CrossRef]
213. Shima, K.; Suda, T.; Nishimoto, K.; Yoshimoto, S. Relationship between molecular structures of sugars and their ability to stimulate the release of glucagon-like peptide-1 from canine ileal loops. *Acta Endocrinol.* **1990**, *123*, 464–470. [Google Scholar] [CrossRef]
214. Kong, M.F.; Chapman, I.; Goble, E.; Wishart, J.; Wittert, G.; Morris, H.; Horowitz, M. Effects of oral fructose and glucose on plasma GLP-1 and appetite in normal subjects. *Peptides* **1999**, *20*, 545–551. [Google Scholar] [CrossRef]
215. Steinert, R.E.; Frey, F.; Topfer, A.; Drewe, J.; Beglinger, C. Effects of carbohydrate sugars and artificial sweeteners on appetite and the secretion of gastrointestinal satiety peptides. *Br. J. Nutr.* **2011**, *105*, 1320–1328. [Google Scholar] [CrossRef]
216. Teff, K.L.; Elliott, S.S.; Tschop, M.; Kieffer, T.J.; Rader, D.; Heiman, M.; Townsend, R.R.; Keim, N.L.; D'Alessio, D.; Havel, P.J. Dietary fructose reduces circulating insulin and leptin, attenuates postprandial suppression of ghrelin, and increases triglycerides in women. *J. Clin. Endocrinol. Metab.* **2004**, *89*, 2963–2972. [Google Scholar] [CrossRef] [PubMed]
217. Havel, P.J. Glucose but not fructose infusion increases circulating leptin in proportion to adipose stores in Rhesus monkeys. *Exp. Clin. Endocrinol. Diabetes* **1997**, *105*, 37–38. [Google Scholar] [CrossRef]
218. Sato, Y.; Ito, T.; Udaka, N.; Kanisawa, M.; Noguchi, Y.; Cushman, S.W.; Satoh, S. Immunohistochemical localization of facilitated-diffusion glucose transporters in rat pancreatic islets. *Tissue Cell* **1996**, *28*, 637–643. [Google Scholar] [CrossRef]
219. Jones, H.F.; Butler, R.N.; Brooks, D.A. Intestinal fructose transport and malabsorption in humans. *Am. J. Physiol. Gastrointest. Liver Physiol.* **2011**, *300*, G202–G206. [Google Scholar] [CrossRef]
220. Bjorkman, O.; Crump, M.; Phillips, R.W. Intestinal metabolism of orally administered glucose and fructose in Yucatan miniature swine. *J. Nutr.* **1984**, *114*, 1413–1420. [Google Scholar] [CrossRef]
221. Hillebrand, J.J.; de Wied, D.; Adan, R.A. Neuropeptides, food intake and body weight regulation: A hypothalamic focus. *Peptides* **2002**, *23*, 2283–2306. [Google Scholar] [CrossRef]
222. Lee, H.M.; Wang, G.; Englander, E.W.; Kojima, M.; Greeley, G.H., Jr. Ghrelin, a new gastrointestinal endocrine peptide that stimulates insulin secretion: Enteric distribution, ontogeny, influence of endocrine, and dietary manipulations. *Endocrinology* **2002**, *143*, 185–190. [Google Scholar] [CrossRef] [PubMed]

223. Broglio, F.; Arvat, E.; Benso, A.; Gottero, C.; Muccioli, G.; Papotti, M.; van der Lely, A.J.; Deghenghi, R.; Ghigo, E. Ghrelin, a natural GH secretagogue produced by the stomach, induces hyperglycemia and reduces insulin secretion in humans. *J. Clin. Endocrinol. Metab.* **2001**, *86*, 5083–5086. [Google Scholar] [CrossRef] [PubMed]
224. Havel, P.J. Peripheral signals conveying metabolic information to the brain: Short-term and long-term regulation of food intake and energy homeostasis. *Exp. Biol. Med.* **2001**, *226*, 963–977. [Google Scholar] [CrossRef] [PubMed]
225. Figlewicz, D.P.; Bennett, J.; Evans, S.B.; Kaiyala, K.; Sipols, A.J.; Benoit, S.C. Intraventricular insulin and leptin reverse place preference conditioned with high-fat diet in rats. *Behav. Neurosci.* **2004**, *118*, 479–487. [Google Scholar] [CrossRef] [PubMed]
226. Minokoshi, Y.; Alquier, T.; Furukawa, N.; Kim, Y.B.; Lee, A.; Xue, B.; Mu, J.; Foufelle, F.; Ferre, P.; Birnbaum, M.J.; et al. AMP-kinase regulates food intake by responding to hormonal and nutrient signals in the hypothalamus. *Nature* **2004**, *428*, 569–574. [Google Scholar] [CrossRef] [PubMed]
227. Cha, S.H.; Wolfgang, M.; Tokutake, Y.; Chohnan, S.; Lane, M.D. Differential effects of central fructose and glucose on hypothalamic malonyl-CoA and food intake. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 16871–16875. [Google Scholar] [CrossRef]
228. Burmeister, M.A.; Ayala, J.; Drucker, D.J.; Ayala, J.E. Central glucagon-like peptide 1 receptor-induced anorexia requires glucose metabolism-mediated suppression of AMPK and is impaired by central fructose. *Am. J. Physiol. Endocrinol. Metab.* **2013**, *304*, E677–E685. [Google Scholar] [CrossRef]
229. Cawley, N.X. Sugar making sugar: Gluconeogenesis triggered by fructose via a hypothalamic-adrenal-corticosterone circuit. *Endocrinology* **2012**, *153*, 3561–3563. [Google Scholar] [CrossRef]
230. Wang, R.; Liu, X.; Hentges, S.T.; Dunn-Meynell, A.A.; Levin, B.E.; Wang, W.; Routh, V.H. The regulation of glucose-excited neurons in the hypothalamic arcuate nucleus by glucose and feeding-relevant peptides. *Diabetes* **2004**, *53*, 1959–1965. [Google Scholar] [CrossRef]
231. Purnell, J.Q.; Klopfenstein, B.A.; Stevens, A.A.; Havel, P.J.; Adams, S.H.; Dunn, T.N.; Krisky, C.; Rooney, W.D. Brain functional magnetic resonance imaging response to glucose and fructose infusions in humans. *Diabetes Obes. Metab.* **2011**, *13*, 229–234. [Google Scholar] [CrossRef]
232. Luo, S.; Monterosso, J.R.; Sarpelleh, K.; Page, K.A. Differential effects of fructose versus glucose on brain and appetitive responses to food cues and decisions for food rewards. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 6509–6514. [Google Scholar] [CrossRef] [PubMed]
233. Lindqvist, A.; Baelemans, A.; Erlanson-Albertsson, C. Effects of sucrose, glucose and fructose on peripheral and central appetite signals. *Regul. Pept.* **2008**, *150*, 26–32. [Google Scholar] [CrossRef] [PubMed]
234. Erlanson-Albertsson, C.; Lindqvist, A. Fructose affects enzymes involved in the synthesis and degradation of hypothalamic endocannabinoids. *Regul. Pept.* **2010**, *161*, 87–91. [Google Scholar] [CrossRef] [PubMed]
235. Mielke, J.G.; Taghibiglou, C.; Liu, L.; Zhang, Y.; Jia, Z.; Adeli, K.; Wang, Y.T. A biochemical and functional characterization of diet-induced brain insulin resistance. *J. Neurochem.* **2005**, *93*, 1568–1578. [Google Scholar] [CrossRef]
236. Agrawal, R.; Gomez-Pinilla, F. Metabolic syndrome in the brain: Deficiency in omega-3 fatty acid exacerbates dysfunctions in insulin receptor signalling and cognition. *J. Physiol.* **2012**, *590*, 2485–2499. [Google Scholar] [CrossRef]
237. Lowette, K.; Roosen, L.; Tack, J.; Vanden Berghe, P. Effects of high-fructose diets on central appetite signaling and cognitive function. *Front. Nutr.* **2015**, *2*, 5. [Google Scholar] [CrossRef]
238. Rizkalla, S.W. Health implications of fructose consumption: A review of recent data. *Nutr. Metab.* **2010**, *7*, 82. [Google Scholar] [CrossRef]
239. Stanhope, K.L.; Havel, P.J. Fructose consumption: Recent results and their potential implications. *Ann. N. Y. Acad. Sci.* **2010**, *1190*, 15–24. [Google Scholar] [CrossRef]
240. Bray, G.A.; Nielsen, S.J.; Popkin, B.M. Consumption of high-fructose corn syrup in beverages may play a role in the epidemic of obesity. *Am. J. Clin. Nutr.* **2004**, *79*, 537–543. [Google Scholar] [CrossRef]
241. Nagarajan, P.; Samuel, V.T.; Shulman, G.I. The pathogenesis of insulin resistance: Integrating signaling pathways and substrate flux. *J. Clin. Investig.* **2016**, *126*, 12–22. [Google Scholar] [CrossRef]
242. Roden, M.; Shulman, G.I. The integrative biology of type 2 diabetes. *Nature* **2019**, *576*, 51–60. [Google Scholar] [CrossRef] [PubMed]
243. Ohashi, K.; Munetsuna, E.; Yamada, H.; Ando, Y.; Yamazaki, M.; Taromaru, N.; Nagura, A.; Ishikawa, H.; Suzuki, K.; Teradaira, R.; et al. High fructose consumption induces DNA methylation at PPARalpha and CPT1A promoter regions in the rat liver. *Biochem. Biophys. Res. Commun.* **2015**, *468*, 185–189. [Google Scholar] [CrossRef] [PubMed]
244. Nakagawa, T.; Hu, H.; Zharikov, S.; Tuttle, K.R.; Short, R.A.; Glushakova, O.; Ouyang, X.; Feig, D.I.; Block, E.R.; Herrera-Acosta, J.; et al. A causal role for uric acid in fructose-induced metabolic syndrome. *Am. J. Physiol. Ren. Physiol.* **2006**, *290*, F625–F631. [Google Scholar] [CrossRef] [PubMed]
245. Reungjui, S.; Roncal, C.A.; Mu, W.; Srinivas, T.R.; Sirivongs, D.; Johnson, R.J.; Nakagawa, T. Thiazide diuretics exacerbate fructose-induced metabolic syndrome. *J. Am. Soc. Nephrol.* **2007**, *18*, 2724–2731. [Google Scholar] [CrossRef]
246. Sanchez-Lozada, L.G.; Tapia, E.; Bautista-Garcia, P.; Soto, V.; Avila-Casado, C.; Vega-Campos, I.P.; Nakagawa, T.; Zhao, L.; Franco, M.; Johnson, R.J. Effects of febuxostat on metabolic and renal alterations in rats with fructose-induced metabolic syndrome. *Am. J. Physiol. Ren. Physiol.* **2008**, *294*, F710–F718. [Google Scholar] [CrossRef]

247. Duplain, H.; Burcelin, R.; Sartori, C.; Cook, S.; Egli, M.; Lepori, M.; Vollenweider, P.; Pedrazzini, T.; Nicod, P.; Thorens, B.; et al. Insulin resistance, hyperlipidemia, and hypertension in mice lacking endothelial nitric oxide synthase. *Circulation* **2001**, *104*, 342–345. [Google Scholar] [CrossRef]
248. Khosla, U.M.; Zharikov, S.; Finch, J.L.; Nakagawa, T.; Roncal, C.; Mu, W.; Krotova, K.; Block, E.R.; Prabhakar, S.; Johnson, R.J. Hyperuricemia induces endothelial dysfunction. *Kidney Int.* **2005**, *67*, 1739–1742. [Google Scholar] [CrossRef]
249. Roy, D.; Perreault, M.; Marette, A. Insulin stimulation of glucose uptake in skeletal muscles and adipose tissues in vivo is NO dependent. *Am. J. Physiol.* **1998**, *274*, E692–E699. [Google Scholar] [CrossRef]
250. Sautin, Y.Y.; Nakagawa, T.; Zharikov, S.; Johnson, R.J. Adverse effects of the classic antioxidant uric acid in adipocytes: NADPH oxidase-mediated oxidative/nitrosative stress. *Am. J. Physiol. Cell Physiol.* **2007**, *293*, C584–C596. [Google Scholar] [CrossRef]
251. Furukawa, S.; Fujita, T.; Shimabukuro, M.; Iwaki, M.; Yamada, Y.; Nakajima, Y.; Nakayama, O.; Makishima, M.; Matsuda, M.; Shimomura, I. Increased oxidative stress in obesity and its impact on metabolic syndrome. *J. Clin. Investig.* **2004**, *114*, 1752–1761. [Google Scholar] [CrossRef]
252. Cheung, K.J.; Tzameli, I.; Pissios, P.; Rovira, I.; Gavrilova, O.; Ohtsubo, T.; Chen, Z.; Finkel, T.; Flier, J.S.; Friedman, J.M. Xanthine oxidoreductase is a regulator of adipogenesis and PPARgamma activity. *Cell Metab.* **2007**, *5*, 115–128. [Google Scholar] [CrossRef]
253. Shapiro, A.; Mu, W.; Roncal, C.; Cheng, K.Y.; Johnson, R.J.; Scarpace, P.J. Fructose-induced leptin resistance exacerbates weight gain in response to subsequent high-fat feeding. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **2008**, *295*, R1370–R1375. [Google Scholar] [CrossRef]
254. White, J.S. Challenging the fructose hypothesis: New perspectives on fructose consumption and metabolism. *Adv. Nutr.* **2013**, *4*, 246–256. [Google Scholar] [CrossRef]
255. Thursby, E.; Juge, N. Introduction to the human gut microbiota. *Biochem. J.* **2017**, *474*, 1823–1836. [Google Scholar] [CrossRef]
256. Donaldson, G.P.; Lee, S.M.; Mazmanian, S.K. Gut biogeography of the bacterial microbiota. *Nat. Rev. Microbiol.* **2016**, *14*, 20–32. [Google Scholar] [CrossRef]
257. David, L.A.; Maurice, C.F.; Carmody, R.N.; Gootenberg, D.B.; Button, J.E.; Wolfe, B.E.; Ling, A.V.; Devlin, A.S.; Varma, Y.; Fischbach, M.A.; et al. Diet rapidly and reproducibly alters the human gut microbiome. *Nature* **2014**, *505*, 559–563. [Google Scholar] [CrossRef]
258. Frazier, T.H.; DiBaise, J.K.; McClain, C.J. Gut microbiota, intestinal permeability, obesity-induced inflammation, and liver injury. *J. Parenter. Enter. Nutr.* **2011**, *35*, 14S–20S. [Google Scholar] [CrossRef]
259. Do, M.H.; Lee, E.; Oh, M.-J.; Kim, Y.; Park, H.-Y. High-Glucose or -Fructose Diet Cause Changes of the Gut Microbiota and Metabolic Disorders in Mice without Body Weight Change. *Nutrients* **2018**, *10*, 761. [Google Scholar] [CrossRef]
260. Roglans, N.; Vila, L.; Farre, M.; Alegret, M.; Sanchez, R.M.; Vazquez-Carrera, M.; Laguna, J.C. Impairment of hepatic Stat-3 activation and reduction of PPARalpha activity in fructose-fed rats. *Hepatology* **2007**, *45*, 778–788. [Google Scholar] [CrossRef]
261. Parry, S.A.; Hodson, L. Influence of dietary macronutrients on liver fat accumulation and metabolism. *J. Investig. Med.* **2017**, *65*, 1102–1115. [Google Scholar] [CrossRef]
262. Jia, Q.; Xia, Y.; Zhang, Q.; Wu, H.; Du, H.; Liu, L.; Wang, C.; Shi, H.; Guo, X.; Liu, X.; et al. Dietary patterns are associated with prevalence of fatty liver disease in adults. *Eur. J. Clin. Nutr.* **2015**, *69*, 914–921. [Google Scholar] [CrossRef] [PubMed]
263. Volynets, V.; Kuper, M.A.; Strahl, S.; Maier, I.B.; Spruss, A.; Wagnerberger, S.; Konigsrainer, A.; Bischoff, S.C.; Bergheim, I. Nutrition, intestinal permeability, and blood ethanol levels are altered in patients with nonalcoholic fatty liver disease (NAFLD). *Dig. Dis. Sci.* **2012**, *57*, 1932–1941. [Google Scholar] [CrossRef] [PubMed]
264. Assy, N.; Nasser, G.; Kamayse, I.; Nseir, W.; Beniashvili, Z.; Djibre, A.; Grosovski, M. Soft drink consumption linked with fatty liver in the absence of traditional risk factors. *Can. J. Gastroenterol.* **2008**, *22*, 811–816. [Google Scholar] [CrossRef] [PubMed]
265. Hudgins, L.C.; Parker, T.S.; Levine, D.M.; Hellerstein, M.K. A dual sugar challenge test for lipogenic sensitivity to dietary fructose. *J. Clin. Endocrinol. Metab.* **2011**, *96*, 861–868. [Google Scholar] [CrossRef] [PubMed]
266. Schwarz, J.M.; Noworolski, S.M.; Wen, M.J.; Dyachenko, A.; Prior, J.L.; Weinberg, M.E.; Herraiz, L.A.; Tai, V.W.; Bergeron, N.; Bersot, T.P.; et al. Effect of a High-Fructose Weight-Maintaining Diet on Lipogenesis and Liver Fat. *J. Clin. Endocrinol. Metab.* **2015**, *100*, 2434–2442. [Google Scholar] [CrossRef]
267. Frayn, K.N.; Kingman, S.M. Dietary sugars and lipid metabolism in humans. *Am. J. Clin. Nutr.* **1995**, *62*, 250S–261S. [Google Scholar] [CrossRef] [PubMed]
268. Jin, R.; Welsh, J.A.; Le, N.-A.; Holzberg, J.; Sharma, P.; Martin, D.R.; Vos, M.B. Dietary fructose reduction improves markers of cardiovascular disease risk in Hispanic-American adolescents with NAFLD. *Nutrients* **2014**, *6*, 3187–3201. [Google Scholar] [CrossRef]
269. Johnston, R.D.; Stephenson, M.C.; Crossland, H.; Cordon, S.M.; Palcidi, E.; Cox, E.F.; Taylor, M.A.; Aithal, G.P.; Macdonald, I.A. No difference between high-fructose and high-glucose diets on liver triacylglycerol or biochemistry in healthy overweight men. *Gastroenterology* **2013**, *145*, 1016–1025. [Google Scholar] [CrossRef]
270. Bravo, S.; Lowndes, J.; Sinnett, S.; Yu, Z.; Rippe, J. Consumption of sucrose and high-fructose corn syrup does not increase liver fat or ectopic fat deposition in muscles. *Appl. Physiol. Nutr. Metab.* **2013**, *38*, 681–688. [Google Scholar] [CrossRef]

271. Ngo Sock, E.T.; Le, K.A.; Ith, M.; Kreis, R.; Boesch, C.; Tappy, L. Effects of a short-term overfeeding with fructose or glucose in healthy young males. *Br. J. Nutr.* **2010**, *103*, 939–943. [Google Scholar] [CrossRef]
272. Silbernagel, G.; Machann, J.; Unmuth, S.; Schick, F.; Stefan, N.; Haring, H.U.; Fritsche, A. Effects of 4-week very-high-fructose/glucose diets on insulin sensitivity, visceral fat and intrahepatic lipids: An exploratory trial. *Br. J. Nutr.* **2011**, *106*, 79–86. [Google Scholar] [CrossRef] [PubMed]
273. Malik, V.S.; Pan, A.; Willett, W.C.; Hu, F.B. Sugar-sweetened beverages and weight gain in children and adults: A systematic review and meta-analysis. *Am. J. Clin. Nutr.* **2013**, *98*, 1084–1102. [Google Scholar] [CrossRef] [PubMed]
274. Hu, F.B. Resolved: There is sufficient scientific evidence that decreasing sugar-sweetened beverage consumption will reduce the prevalence of obesity and obesity-related diseases. *Obes. Rev.* **2013**, *14*, 606–619. [Google Scholar] [CrossRef] [PubMed]
275. Wang, H.; Steffen, L.M.; Zhou, X.; Harnack, L.; Luepker, R.V. Consistency between increasing trends in added-sugar intake and body mass index among adults: The Minnesota Heart Survey, 1980–1982 to 2007–2009. *Am. J. Public Health* **2013**, *103*, 501–507. [Google Scholar] [CrossRef]
276. Runchey, S.S.; Valsta, L.M.; Schwarz, Y.; Wang, C.; Song, X.; Lampe, J.W.; Neuhouser, M.L. Effect of low- and high-glycemic load on circulating incretins in a randomized clinical trial. *Metab. Clin. Exp.* **2013**, *62*, 188–195. [Google Scholar] [CrossRef]
277. Stevenson, E.J.; Astbury, N.M.; Simpson, E.J.; Taylor, M.A.; Macdonald, I.A. Fat oxidation during exercise and satiety during recovery are increased following a low-glycemic index breakfast in sedentary women. *J. Nutr.* **2009**, *139*, 890–897. [Google Scholar] [CrossRef]
278. Milton, J.E.; Sananthanan, C.S.; Patterson, M.; Ghatei, M.A.; Bloom, S.R.; Frost, G.S. Glucagon-like peptide-1 (7–36) amide response to low versus high glycaemic index preloads in overweight subjects with and without type II diabetes mellitus. *Eur. J. Clin. Nutr.* **2007**, *61*, 1364–1372. [Google Scholar] [CrossRef]
279. Marathe, C.S.; Rayner, C.K.; Bound, M.; Checklin, H.; Standfield, S.; Wishart, J.; Lange, K.; Jones, K.L.; Horowitz, M. Small intestinal glucose exposure determines the magnitude of the incretin effect in health and type 2 diabetes. *Diabetes* **2014**, *63*, 2668–2675. [Google Scholar] [CrossRef]
280. Young, R.L.; Lumsden, A.L.; Martin, A.M.; Schober, G.; Pezos, N.; Thazhath, S.S.; Isaacs, N.J.; Cvijanovic, N.; Sun, E.W.L.; Wu, T.; et al. Augmented capacity for peripheral serotonin release in human obesity. *Int. J. Obes.* **2018**, *42*, 1880–1889. [Google Scholar] [CrossRef]
281. Alssema, M.; Rijkelijkhuizen, J.M.; Holst, J.J.; Teerlink, T.; Scheffer, P.G.; Eekhoff, E.M.; Gastaldelli, A.; Mari, A.; Hart, L.M.; Nijpels, G.; et al. Preserved GLP-1 and exaggerated GIP secretion in type 2 diabetes and relationships with triglycerides and ALT. *Eur. J. Endocrinol.* **2013**, *169*, 421–430. [Google Scholar] [CrossRef]
282. Grasset, E.; Puel, A.; Charpentier, J.; Collet, X.; Christensen, J.E.; Terce, F.; Burcelin, R. A Specific Gut Microbiota Dysbiosis of Type 2 Diabetic Mice Induces GLP-1 Resistance through an Enteric NO-Dependent and Gut-Brain Axis Mechanism. *Cell Metab.* **2017**, *25*, 1075–1090. [Google Scholar] [CrossRef] [PubMed]
283. Yeow, T.P.; Pacini, G.; Tura, A.; Hor, C.P.; Lim, S.L.; Tan, F.H.; Tong, C.V.; Hong, J.Y.; Md Zain, F.; Holst, J.J.; et al. Preserved glucagon-like peptide-1 responses to oral glucose, but reduced incretin effect, insulin secretion and sensitivity in young Asians with type 2 diabetes mellitus. *BMJ Open Diabetes Res. Care* **2017**, *5*, e000352. [Google Scholar] [CrossRef] [PubMed]

29.6: METABOLIC CONSEQUENCES OF OBESITY AND TYPE 2 DIABETES-BALANCING GENES AND ENVIRONMENT FOR PERSONALIZED CARE

29.6.1: METABOLIC CONSEQUENCES OF OBESITY AND TYPE 2 DIABETES: BALANCING GENES AND ENVIRONMENT FOR PERSONALIZED CARE

Nicolas J. Pillon, Ruth J.F. Loos, Sally M. Marshall, Juleen R. Zierath,. Metabolic consequences of obesity and type 2 diabetes: Balancing genes and environment for personalized care. Cell, Volume 184, Issue 6, 2021, Pages 1530-1544, ISSN 0092-8674, https://doi.org/10.1016/j.cell.2021.02.012.

SUMMARY

The prevalence of type 2 diabetes and obesity has risen dramatically for decades and is expected to rise further, secondary to the growing aging, sedentary population. The strain on global health care is projected to be colossal. This review explores the latest work and emerging ideas related to genetic and environmental factors influencing metabolism. Translational research and clinical applications, including the impact of the COVID-19 pandemic, are highlighted. Looking forward, strategies to personalize all aspects of prevention, management and care are necessary to improve health outcomes and reduce the impact of these metabolic diseases.

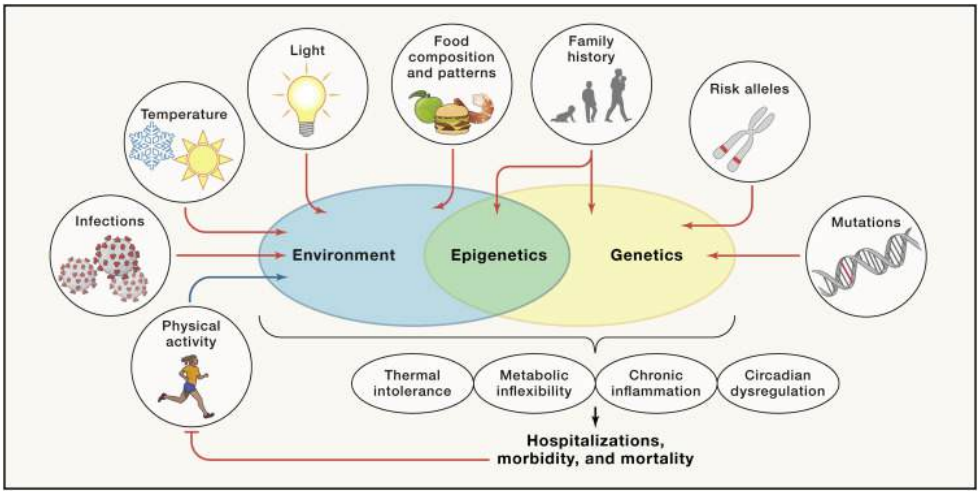
29.6.1.1: INTRODUCTION

The COVID-19 pandemic has brought the deleterious health consequences of obesity and type 2 diabetes into sharp focus. Individuals with type 2 diabetes and/or obesity are more likely to have severe disease and to die than are individuals without diabetes (Barron et al., 2020). Fasting glucose level at the time of hospital admission predicts 28-day mortality even in those without a previous diagnosis of diabetes (Wang et al., 2020a). Glycemic control and body mass index along with older age, male sex, socio-economic deprivation, non-white ethnicity, and pre-existing renal and cardiovascular disease all independently increase mortality (Holman et al., 2020). COVID-19 is also a timely reminder that diabetes is not merely a state of glucose dysregulation but a multi-faceted syndrome driven by many medical and social risk factors and associated with pathophysiological changes throughout the body.

The World Health Organization estimates that worldwide, 422 million people have diabetes, the majority living in low- and middle-income countries, and most having type 2 diabetes (who.int/health-topics/diabetes). The prevalence has risen dramatically for decades, as the population ages and becomes less active and more overweight (GBD 2019 Risk Factors Collaborators, 2020). Early detection is vital, particularly as long-term complications, such as referable diabetic retinopathy, may be present at diagnosis of type 2 diabetes (Kohner et al., 1998). Many developed countries have systematic screening programs of individuals deemed to be at high risk (American Diabetes Association, 2020). However, there is disagreement as to how to define “high risk” and how to screen (oral glucose tolerance test, fasting glucose or glycated hemoglobin, HbA1c). Glucose-based tests and HbA1c each identify slightly different populations. We do not know if these differences in diagnoses lead to important clinically different outcomes or if they signal slightly different pathological metabolic forms of glucose dysregulation (American Diabetes Association, 2020).

The WHO defines overweight and obesity as body mass indexes (BMI) ^{3} 25 and 30 kg/m^2, respectively and estimated that 1.9 billion adults were overweight and 650 million obese in 2016 (https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight). Obesity is now regarded as a chronic, progressive disease with remissions and relapses (Bray et al., 2017) and an important driver of the development of diabetes and many of its associated features (GBD 2019 Risk Factors Collaborators, 2020). The deleterious effects of obesity and type 2 diabetes are seen in most, if not all, tissues in the body, with consequences resulting in significantly increased premature morbidity and mortality (GBD 2019 Risk Factors Collaborators, 2020). Social and cultural factors are also extremely important in the development, management, and clinical outcomes of obesity and type 2 diabetes.

Despite advances in diabetes care over the recent decades, there remain vast challenges: developing an improved understanding of the heterogeneity of obesity and diabetes, how best to assess risk, to screen, to select individualized treatments and vitally how to engage the relevant populations in these programs. This review explores the genetic and metabolic aspects of diabetes and obesity (Figure 1) and discusses some of the latest work and emerging ideas related to basic biological mechanisms, translational research, and clinical applications.



1. Download : Download high-res image (569KB)
2. Download : Download full-size image

Figure 1. Gene-environment interactions regulating disease risk of obesity and type 2 diabetes

Individual genetic predispositions and environmental factors interact to promote or impair molecular processes, such as circadian regulation, thermal tolerance, and/or chronic inflammation. Accumulation of genetic and environment risk factors eventually leads to the development of complications, reducing both healthspan and lifespan.

29.6.1.1.1: GENETICS AND METABOLISM

The current obesogenic environment, favoring high-calorie foods and physical inactivity, is a major driver of the growing obesity and diabetes epidemic. However, not everyone exposed to this environment gains weight or develops type 2 diabetes. The way people respond to environmental factors is, at least in part, determined by their genetic predisposition to obesity and type 2 diabetes. Traditionally, the genetic contribution has been quantified by the heritability, which is a population-level estimate of how much of the variation in disease susceptibility is attributable to genetic variation. For obesity and type 2 diabetes, the heritability has been estimated to be moderate-to-high, ranging between 30% and 70% (Elks et al., 2012; Willemsen et al., 2015). The search for contributing genes started in the 1990's with early success largely confined to monogenic forms of obesity and diabetes. Mutations that segregate in families or occur *de novo* were found to cause major disruptions in the function of genes in which they are located, providing the first insights in the pathophysiology of body-weight regulation and glucose metabolism (Hattersley and Patel, 2017; van der Klaauw and Farooqi, 2015). The search for genetic variants that contribute to common forms of obesity and diabetes began slowly with candidate gene and genome-wide linkage studies. However, the advent of genome-wide association studies (GWASs) in the mid-2000's accelerated the pace of gene discovery.

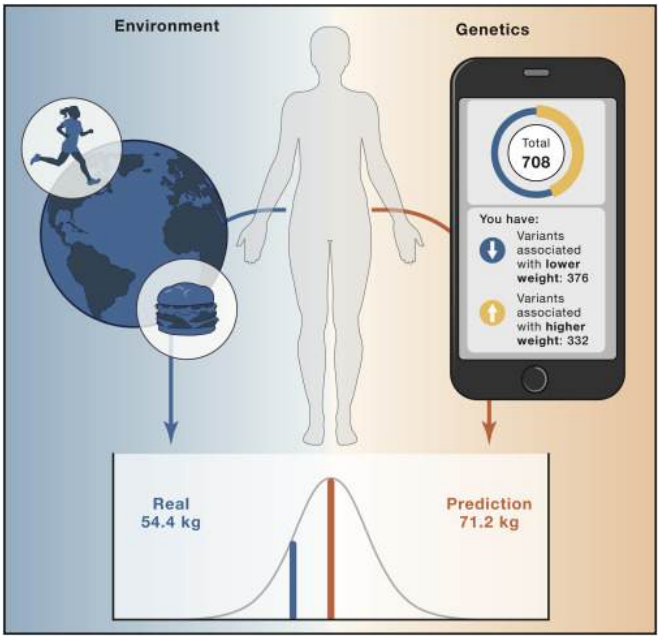
GWASs have identified thousands of genetic loci that are robustly associated with complex diseases and traits, including 700 for obesity (Yengo et al., 2018) and at least 400 for type 2 diabetes (Mahajan et al., 2018). From the earliest GWAS, tissue enrichment and pathway analyses for BMI-associated loci have suggested that the central nervous system plays a key role in body weight regulation (Locke et al., 2015). Loci associated with type 2 diabetes act predominantly through the perturbation of insulin secretion, pointing to the importance of beta cell function or mass, whereas few loci affect insulin resistance through an effect on body weight or fat distribution (Barroso and McCarthy, 2019).

Despite the success of GWASs, pinpointing the causal gene(s) and variant(s) within each locus remains an ongoing challenge. So far, about 20% of loci associated with type 2 diabetes and a handful of loci associated with obesity have been mapped to the most likely causal variant, whereas the underlying biology of hundreds of additional loci remain to be elucidated (Larder et al., 2017; Mahajan et al., 2018; Rathjen et al., 2017). However, with increasing availability of high-throughput genome-scale technologies for mapping regulatory elements, comprehensive multi-omics databases, advanced computational tools, and the latest genetic engineering and molecular phenotyping approaches, we are poised to accelerate the translation of GWAS loci into meaningful biology in the years ahead.

With GWASs, genetic susceptibility to disease can be assessed using polygenic scores. A polygenic score represents an individual's overall genetic susceptibility to disease and is calculated by summing the number of disease-increasing alleles that were inherited from either parent, weighted by each variant's effect size observed in a GWAS. Even though each locus has a small effect on disease risk and explains only a fraction of the variation in disease susceptibility, when aggregated in a polygenic score, their contribution can be substantial. Polygenic scores are normally distributed, with most individuals having an average score, and thus an average genetic susceptibility, whereas individuals at the extremes of the distribution have a (very) high or low genetic risk of disease. For example, in the UK Biobank, the average BMI of individuals with a high polygenic score (top decile) is 2.9 kg/m^2 (equivalent to 8 kg in body weight) higher and their odds of severe obesity (BMI ^{3} 40 kg/m^2) is 4.2-fold higher, compared to those with a lower polygenic score (bottom 9 deciles) (Khera et al.,

2019). Similarly, individuals with a very high polygenic score (top 5%) for type 2 diabetes have a 2.75-fold increased risk of disease compared to the remainder of the population (Udler et al., 2019).

These observations have fueled expectations that genotype information, including polygenic scores, can soon be used in clinical care for early diagnosis of high-risk individuals, to tailor prevention and treatments strategies, and to improve disease prognostics. In fact, many online direct-to-consumer genomic companies are already informing customers about their risks and predispositions for a range of common diseases and traits based solely on genetic profiling, including for obesity and type 2 diabetes (Figure 2). However, even though the genetic associations observed in GWASs are robust, their ability to predict who will be at a high risk of obesity or type 2 diabetes is still low-to-moderate, and not ready for use in clinical settings (Udler et al., 2019). For example, a recent polygenic score applied to individuals of European ancestry of the UK Biobank explains only 8%–9% of the variation in BMI and is a weak predictor of obesity, with an area under the receiver operating characteristic curve (AUC_{ROC}) of 0.64 (Khera et al., 2019). Findings are similar for polygenic scores for type 2 diabetes, with AUC_{ROC} of 0.64–0.66 (Udler et al., 2019). The predictive ability of polygenic scores are expected to improve as GWASs increase in sample size and the per-variant effect estimates become more precise, and as algorithms to aggregate millions of genetic variants across the genome improve. Nevertheless, given the importance of socio-demographic, lifestyle and clinical risk factors in the etiology of obesity and type 2 diabetes, it is unlikely that a polygenic score on its own will ever be able to accurately predict obesity or type 2 diabetes. More comprehensive approaches that include a broad spectrum of genetic, demographic, environmental, clinical, and possibly also molecular markers are needed to accurately predict who is at risk of gaining weight and/or developing type 2 diabetes.



1. Download : Download high-res image (446KB)
2. Download : Download full-size image

Figure 2. Genetic prediction of body weight—context matters

A 49-year old woman, who eats a balanced diet, runs 6 miles/day, and commutes by bike to work, does a direct-to-consumer genetic test with one of the many online personalized genomics companies. The company claims to provide genetic insight into her health to make it easier for her to take action. She provides a saliva sample and completes numerous questionnaires on her physical and mental health, family medical history, and more. Her reported results show that, based on the genetic variants tested, she carries 376 weight-lowering variants and 332 weight-increasing variants, predisposing her to weigh about “average” or 157 lbs (71.2 kg, based on the company’s customers’ weight of the same age, height, and sex). However, the woman’s real weight was 120 lbs (54.4 kg). A likely reason for why this genetic test overpredicted the woman’s weight by 30% is because her lifestyle—even though information was shared in detail—was not appropriately incorporated in the prediction models.

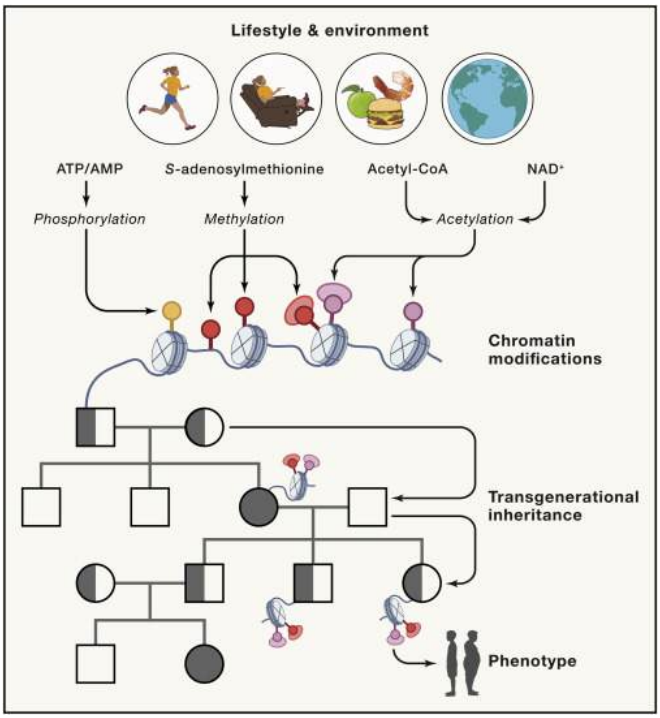
The vast amount of new genetic information generated by GWASs is being used in sub-typing disease at a population level. Obesity and type 2 diabetes are highly heterogeneous diseases, and the diagnosis of these metabolic diseases is unrefined, based on a single marker (BMI ^{3} 30 kg/m^2 and hyperglycemia, respectively). Consequently, individuals with the same diagnosis may differ considerably in disease pathogenesis, clinical presentation, disease course and response to treatments. Subtypes of obesity and type 2 diabetes have been typically based on phenotypic differences and similarities. As the number of GWAS-identified loci continues to increase, subtyping of obesity and type 2 diabetes based on genetic information has become possible. In a recent study, 141 variants previously identified for diabetes and diabetes-related traits were clustered in five groups, based on their association with more than 75 traits (Udler et al., 2018). Variants with a

similar association profile cluster in the same group, and the group-specific association profile can inform about the mechanisms underlying a given subtype of type 2 diabetes. For example, two of the five groups identified for diabetes-related traits represent reduced beta-cell function, of which one cluster is characterized by high and the other by low proinsulin levels. The three other groups of variants show features of insulin resistance, of which one group represents obesity-mediated insulin resistance, a second group represents abnormal body fat distribution ("lipodystrophy-like"), and a third group represents disrupted liver lipid metabolism. Genetic risk scores based on variants in each cluster are associated with distinct clinical outcomes (Udler et al., 2018). Further for obesity, genotype information has been used to identify individuals who are predisposed to increased adiposity and, concomitantly, are protected from cardiometabolic outcomes (representing the so-called metabolically healthy obesity phenotype) (Ji et al., 2019). Subtyping of heterogenous diseases, like obesity and type 2 diabetes, is key to precision medicine. Indeed, these more homogeneous subgroups are characterized by distinct underlying biological mechanisms, such that diagnosis and prognosis will be more precise and optimization of treatment more efficient (Chung et al., 2020). As GWASs continue to identify more loci, additional and possibly better-defined clusters may be identified to more accurately represent the heterogenous group of individuals with obesity and type 2 diabetes.

As more GWAS loci are being discovered, Mendelian Randomization (MR) becomes an increasingly powerful approach to determine causality between an exposure (e.g., health-related behaviors, biomarkers [e.g., lipid levels, metabolites]) and an outcome (e.g., obesity, type 2 diabetes). Genetic variants that are robustly associated with the exposure are used to randomize a population in individuals with high exposure (i.e., carriers of the risk alleles) and those with low exposure (i.e., carriers of non-risk alleles). If the same genetic variants also associate with the disease outcome, through their association with the exposure, then causality between exposure and disease is inferred. For example, a large-scale MR study examined the causal role of a wide range of possible risk factors for type 2 diabetes, mostly confirming established risk factors, but also revealing new ones (e.g., insomnia) (Yuan and Larsson, 2020). As more GWAS data becomes available for a range of multi-omics biomarkers, MR analyses may reveal novel disease-causing biomarkers, broadening insights in the pathogenesis of obesity and type 2 diabetes.

29.6.1.1.2: EPIGENETIC IMPACT ON METABOLISM

Beyond genetic risk, the genes we inherit and the environmental factors we are exposed to can interact synergistically to modify our physiology and risk for obesity and type 2 diabetes through epigenetic modifications (Figure 3). Epigenetic modifications are biochemical processes that influence gene activity and expression, and ultimately modify cellular and whole-body physiology, without altering the DNA sequence of an organism's genome (Barrès and Zierath, 2016). Mechanistically, epigenetic modifications can arise from chemical alterations of nucleosides in the DNA molecule itself by methylation or hydroxymethylation, alterations in chromatin structure or post-translational modifications of histones (i.e., methylation, phosphorylation, acetylation, ubiquitylation, and sumoylation) or RNA-associated gene silencing (Bošković and Rando, 2018). Although epigenetic modifications are generally thought to be fixed during development and maintained over an organism's lifetime, there is some degree of plasticity in the epigenome, which engenders organismal adaptation to rapid environmental changes.



1. Download : Download high-res image (458KB)
2. Download : Download full-size image

Figure 3. Epigenetic modifications in response to environmental factors lead to transgenerational effects on the phenotypes of offspring

Diet and exercise influence the cellular availability of nutrients impacting methylation, acetylation and phosphorylation of chromatin. Paternal or maternal environmental exposure can therefore influence metabolism and manifest obesity- or type 2 diabetes-related traits in the offspring through transgenerational epigenetic inheritance.

Alterations in nutritional status, food supply, physical activity/exercise, thermal stress, toxins, or other environmental insults can trigger epigenetic modifications and lead to genomic changes in somatic cells within an individual that directly disrupt metabolic homeostasis (Barrès and Zierath, 2016; Bošković and Rando, 2018). These same factors may also modify the physiology of an organism by transgenerational epigenetic inheritance, whereby paternal or maternal environmental exposure can influence metabolism and manifest obesity- or type 2 diabetes-related traits in the offspring. Prenatal undernutrition affects glucose tolerance and risk of diabetes in the offspring, as demonstrated by epidemiological studies of several famines over the past century (Li et al., 2010; Ravelli et al., 1998). In rodents, paternal and maternal diet and exercise influence metabolic and cardiovascular outcomes in offspring over several generations (de Castro Barbosa et al., 2015; Murashov et al., 2016; Stanford et al., 2015). Thus, nutritional status *in utero* during fetal development may affect the epigenome for several generations, but the molecular transducers remain to be clarified. Additionally, food restriction during childhood, at different growth phases around puberty, also leads to epigenetic changes that influence the risk of cardiovascular and metabolic disease of offspring over several generations (Kaati et al., 2002). Accordingly, epigenetic factors passed on by the gametes may contribute to the global increase in obesity and type 2 diabetes. Thus, an area of emerging interest is the influence of the environment on epigenetic mechanisms, and how this modifies metabolic disease risk.

A variety of dietary agents, as well as micronutrients and metabolites synthesized *de novo*, can serve as substrates or co-factors to influence the epigenome and potentially affect metabolic disease risk in humans, in part by affecting genomic plasticity (Tiffon, 2018). One-carbon metabolism encompasses folate and methionine cycles, which transfer one-carbon moieties and methyl groups for nucleotide synthesis, methylation reactions and reductive metabolism (Newman and Maddocks, 2017). Metabolites including acetyl-coA, AMP, NAD+, and *S*-adenosylmethionine are required for histone modifications (acetylation, phosphorylation) and methylation of DNA and histones. The extent to which nutritional factors, metabolites, and other co-factors directly modify the epigenome within a generation remains to be fully substantiated in humans.

While it is important to stress that type 2 diabetes and obesity are complex multi-factorial, progressive metabolic diseases with diverse etiology, and not simply “lifestyle disorders,” diet and exercise regimes can prevent or delay disease progression. Changes in the concentration of cellular metabolites, nucleotides, or calcium levels in skeletal muscle in response to acute exercise alter DNA methylation or histone modifications and influence gene expression through epigenetic mechanisms (Barrès and Zierath, 2016). In humans, acute exercise alters DNA methylation of the promoters of genes involved in metabolic regulation in skeletal muscle (Barrès et al., 2012; Nitert et

al., 2012). Epigenetic modifications have also been observed in skeletal muscle and adipose tissue in humans with obesity and weight loss (Barres et al., 2013; Multhaup et al., 2015). Thus, the impact of environmental exposures and epigenetic influences on the risk for metabolic diseases throughout the lifespan is an important aspect of biology to unravel.

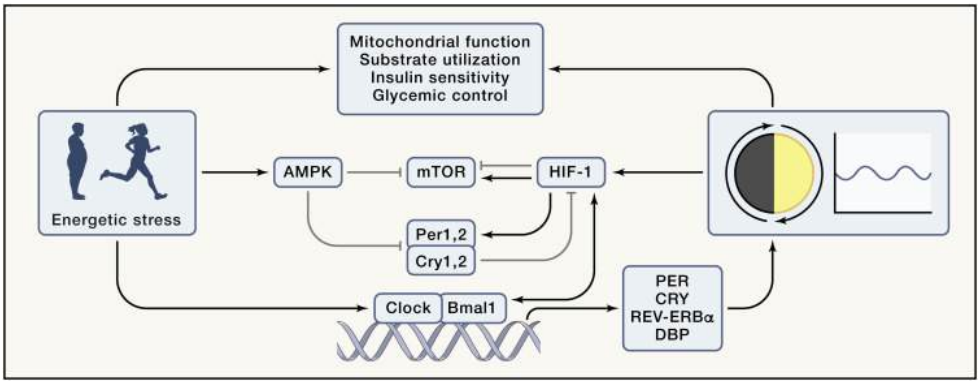
29.6.1.1.3: CIRCADIAN CONTROL OF METABOLISM

An evolutionarily conserved mechanism by which environmental factors can impact whole-body physiology is through internal biological clocks and the control of circadian rhythms (Young, 2018). Circadian rhythms are driven by cell-autonomous intrinsic clocks that anticipate day/night cycles in order to optimize the physiology and behavior of organisms. Circadian programs are regulated at both the central and peripheral level with the master clock, located in the suprachiasmatic nucleus region of the hypothalamus, acting as conductor to synchronize and direct peripheral oscillators (Young, 2018). Synchronization of these intrinsic circadian clocks can be achieved in response to photic and non-photic *zeitgebers* (time-givers). The most powerful *zeitgeber* is light, which synchronizes the central clock. In addition to receiving cues from the central clock, peripheral clocks are synchronized by external *zeitgebers*, including food intake, temperature, energetic stressors, and drive the expression of a broad network of genes, many of which are involved in metabolic homeostasis (Gabriel and Zierath, 2019). The precise mechanism by which circadian clocks coordinate whole-body homeostatic processes is an area of emerging interest given the importance of external *zeitgebers* and the regulation of gene programs controlling metabolism and development.

One mechanism by which the circadian machinery influences metabolism is through the diurnal patterns of hormone secretion (Gamble et al., 2014). Endocrine organs release a variety of hormones in response to diverse environmental factors including diurnal cycles of light/dark, fasting/feeding, and temperature changes. For example, there are diurnal or circadian patterns of secretion of cortisol, growth hormone, prolactin, thyroid hormone, gonadal steroids, and melatonin related to sleep/wake cycles, whereas metabolic hormones including insulin, leptin, ghrelin, and glucagon vary in response to nutritional cues related to fasting/feeding cycles (Gamble et al., 2014). Many of these hormones including insulin, insulin-like growth factor 1, and glucocorticoids can act as *zeitgebers* to reset or fine tune the clock (Balsalobre et al., 2000; Crosby et al., 2019). Thus, an intimate relationship between circadian clocks and endocrine systems exists. This relationship is clinically relevant since disruption of the circadian clock is linked to metabolic disease.

In humans, long duration of shift work is associated with an increased risk of type 2 diabetes, which is only partly explained by lifestyle factors and BMI (Vimalananda et al., 2015). Epidemiological studies show that disruption of the sleep/wake cycle through extended periods of rotating night shift work is associated with obesity and increased risk of type 2 diabetes (Lin et al., 2009; Pan et al., 2011). Chronic jet lag in mouse models disrupts exergy homeostasis and leptin signaling and leads to circadian dysfunction-induced obesity (Kettner et al., 2015). Similarly, a population-based cohort study indicates that social jet lag, defined as the discrepancy between circadian and social clocks, is associated with increased risk of metabolic syndrome and diabetes/prediabetes (Koopman et al., 2017). Thus, chronobiology has implications for obesity and type 2 diabetes pathogenesis.

A basic paradigm of circadian regulation of metabolism is that oscillations of gene expression generate daily rhythms in cellular metabolism (Kim and Lazar, 2020). At the molecular level, circadian rhythms are generated by a cell autonomous and self-sustained transcriptional auto-regulatory feedback loop that is composed of transcriptional activators and their target genes, which rhythmically accumulate and form a repressor complex to inhibit transcriptional activity (Figure 4). Energy, nutrient, and oxygen sensors interact with the circadian clock machinery to control metabolic outputs including mitochondrial function, substrate utilization, insulin sensitivity, and glycemic control (Lamia et al., 2009; Peek et al., 2017; Sato et al., 2019). These sensors monitor oxygen availability and energy stress via hypoxia-inducible factor-1 alpha (HIF1α) and AMP-activated protein kinase (AMPK), respectively. Cells also integrate signals from nutrients and growth factors via mammalian target of rapamycin (mTOR). These energetic sensors not only exhibit circadian rhythmicity, but also regulate components of the core clock machinery through epigenetic modifications, mainly involving histone modifications (Kim and Lazar, 2020). Thus, cross-talk exists between the circadian clock and epigenetic factors that influence the genomic plasticity of organs controlling metabolic homeostasis. In rodents, dysregulation of the intrinsic molecular clock in a variety of tissues leads to obesity, insulin resistance, and altered glucose homeostasis (Rudic et al., 2004; Turek et al., 2005). Nevertheless, the mechanisms underlying disrupted circadian rhythmicity in people with type 2 diabetes are unknown. There is potential to coordinate behavioral changes with the body's daily rhythm to improve metabolic homeostasis. Timing of exercise training bouts or meals and distribution of calories throughout the day may lead to improved outcomes for people with obesity or type 2 diabetes (Lundell et al., 2020; Savikj et al., 2019).



1. Download : Download high-res image (357KB)
2. Download : Download full-size image

Figure 4. Circadian control and influence of energy sensing pathways

Mitochondrial function, substrate utilization, insulin sensitivity, and glycemic control exhibit diurnal rhythms that are influenced by a variety of factors including energetic stressors such as diet, exercise, and metabolic disease, as well as intrinsic clocks. The molecular circadian clock is composed of transcriptional activators, circadian locomotor output cycles kaput (CLOCK), and brain and muscle ARNTL-like protein 1 (BMAL1), and their target genes period (*PER*), cryptochrome (*CRY*), *NR1D1* (which encodes REV-ERBα) and *DBP*, which rhythmically accumulate and form a repressor complex that interacts with CLOCK and BMAL1 to inhibit transcriptional activity. Energetic stressors influence the circadian program and metabolism. AMPK-mediated phosphorylation of CRY and PER promotes their destabilization and degradation, while mTOR activation induces CRY1 expression. PER2 inhibits mTOR complex activity via the tuberous sclerosis complex 1. HIF1α regulates PER2 transcription and interacts with BMAL1 at the chromatin level. CRY1 reduces HIF1α half-life by interacting with its basic-helix-loop-helix domain.

29.6.1.1.4: IMPACT OF ENERGETIC STRESSORS ON THE CONTROL OF METABOLISM

Obesity, diabetes, exercise, and food restriction are energetic stressors that represent major challenges to organismal homeostasis, triggering wide-ranging responses in numerous cells and tissues controlling glucose and energy metabolism. An essential component of an organism's survival is the ability to sense energy availability and to adapt accordingly. Metabolic flexibility, the ability to shift between fat and glucose oxidation with fasting and feeding, is reduced in individuals with metabolic diseases and contributes to the overall insulin resistance phenotype (Kelley et al., 1992). Skeletal muscle exhibits metabolic flexibility in fuel preference, likely due to its crucial role in hunting and surviving predation, situations requiring movement even if nutrient availability is not optimal (Freese et al., 2017). A body of literature supports the idea that metabolic flexibility can be directly influenced by physical activity, independent of changes in energy balance (Rynders et al., 2018). Physical exercise enhances skeletal muscle insulin sensitivity and improves whole-body glucose metabolism in people with type 2 diabetes (Savikj and Zierath, 2020). However, recent findings, based on stable-isotope tracer and liquid chromatography tandem mass spectrometry, demonstrate that skeletal muscle mitochondrial substrate preference is not altered in insulin resistant rodents and humans, calling into question the central role of metabolic flexibility in the pathogenesis of metabolic diseases (Song et al., 2020). Nevertheless, there is growing appreciation that insulin resistance, obesity, and type 2 diabetes can be avoided or at least delayed by lifestyle intervention strategies, including diet and exercise, which initiate diverse homeostatic responses across multiple organs (Savikj and Zierath, 2020).

The concept of "time-restricted feeding" has gained traction as a dietary means to restore metabolic homeostasis, enhance insulin sensitivity, and curb obesity. Time-restricted feeding refers to restricting daily food intake to a few hours, without caloric restriction (Chaix et al., 2014). In rodents, time-restricted feeding synchronizes the feeding/fasting cycle with the central clock, thereby promoting robust circadian and metabolic cycles, which mitigates obesity and metabolic dysfunction (Hatori et al., 2012). Thus, timing of food intake with the molecular circadian clock may fine-tune metabolism. In humans, time-restricted feeding paradigms improve cardiometabolic health in people with obesity or metabolic disease (Cienfuegos et al., 2020; Wilkinson et al., 2020). Short-term time-restricted feeding schedules in men with obesity modulate the diurnal rhythm of lipid and amino acid metabolism, without affecting core clock gene expression in skeletal muscle (Lundell et al., 2020). Furthermore, the timing and type of nutritional intake throughout a day influences carbohydrate metabolism and protein synthesis (Areta et al., 2013). Whether this is dependent upon the release of hormones, metabolites, or thermogenesis warrants further investigation. Moreover, the weight and cardiometabolic benefits achieved with time-restricted feeding schedules may be related to reductions in calorie intake, rather than meal timing. Concordantly, a prospective randomized clinical trial including 116 men and women with overweight or obesity found that modest reductions in weight loss and energy intake from time-restricted eating did not differ from the control group (Lowe et al., 2020), hinting at the possibility that benefits of time-restricted feeding programs are mainly due to reductions in calorie intake.

Diet and exercise have a synergistic effect on insulin sensitivity, which may be influenced by altering the timing of the meal or an exercise bout throughout the day. In rodents, there is a time-of-day-dependent effect of acute exercise on the diurnal oscillations of skeletal muscle metabolites and transcripts, with a greater reliance on glycolytic metabolism when exercise is performed during the early active phase of the day (Sato et al., 2019). Moreover, in a preliminary clinical investigation comparing the time-of-day impact of high intensity exercise in men with type 2 diabetes, greater blood glucose control was achieved with afternoon versus morning exercise (Savikj et al., 2019).

The oxygen-sensitive transcription factor HIF1α links time-of-day-specific effects of exercise on gene expression and carbohydrate metabolism in mice models (Peek et al., 2017; Sato et al., 2019). This finding has clinical relevance, since intense exercise acutely increases skeletal muscle protein abundance and DNA binding activity of HIF1α in humans (Ameln et al., 2005). Moreover, energetic stressors, such as exercise and hypoxia, increase skeletal muscle glucose uptake in healthy and insulin resistant humans and rodents (Ranheim et al., 1997; Ryder et al., 2000). Thus, perturbing energy, nutrient, and/or oxygen sensors may have a varied response on cellular metabolism depending on the time of day. Collectively, these studies provide evidence that the timing of exercise bouts throughout the day is clinically relevant for the diurnal control of glycemia or systemic metabolism. Adjusting the timing of external cues (i.e., meal/exercise timing) may sustain or amplify circadian clock signals to prevent or mitigate metabolic disease.

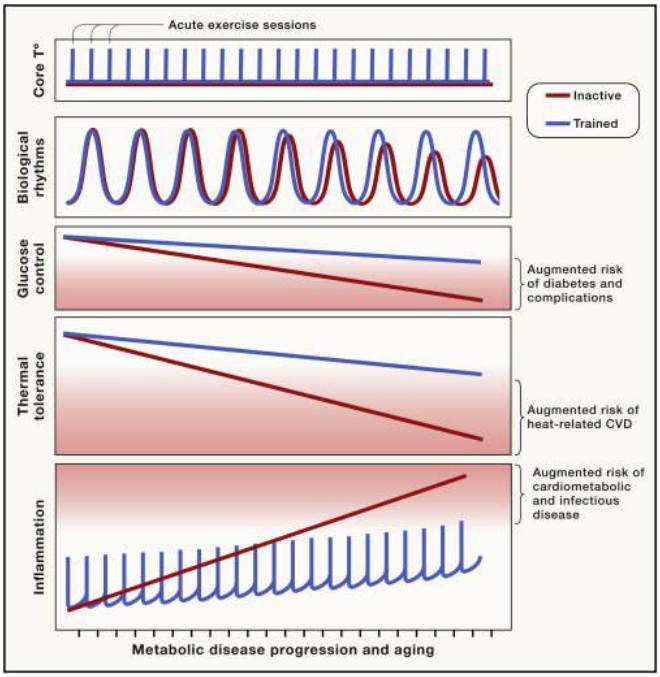
29.6.1.1.5: THERMAL TOLERANCE

Excess energy can be dissipated in the form of heat, a process that occurs in brown adipose tissue and is stimulated by food intake and cold exposure (Chouchani et al., 2019). Feedback loops involving temperature sensors, thermogenesis, sweating, and the control of blood circulation are tightly regulated to maintain body temperature in humans at ∼37°C. Alterations in ambient temperature trigger acute and chronic changes in whole-body physiology, making climate a major environmental stressor that affects all individuals on the planet. Acute exposure to cold triggers shivering in skeletal muscle, where ATP is used to generate movement and its associated production of heat. Chronic adaptation to cold involves different mechanisms, the main one being activation of brown adipose tissue thermogenesis (Chouchani et al., 2019). Uncoupling protein 1 dissipates the proton gradient in the mitochondria to generate heat instead of ATP. Consequently, oxidative phosphorylation increases to maintain mitochondrial membrane potential. Therefore, exposure to cold temperatures increases the metabolic rate during sleep cycles, as well as diet-induced thermogenesis, thereby increasing total energy expenditure (Chouchani et al., 2019). A rise in ambient temperature above thermoneutrality also increases metabolism by promoting heat dissipation (Chouchani et al., 2019).

The processes involved in heat acclimation have been extensively studied in humans and involve an increase in total body water, increased sweat volume and decreased sweat concentration, as well as adaptations of heart rate and skin blood flow (Périard et al., 2015). Mechanisms involved in heat acclimation and associated cardiovascular events are related to increased central heat production and dehydration and the ensuing deleterious consequences on blood pressure and cardiovascular function (Meade et al., 2020).

Acute exposure to extreme ambient temperatures, often referred to as “cold stress” or “heat stress,” is associated with an increased risk of cardio-pulmonary mortality (Achebak et al., 2019). In this context, age, weight, obesity, and type 2 diabetes are major risk factors (Hajat et al., 2017; Huang et al., 2012). The mechanisms for increased risk of cardiovascular events secondary to extreme temperatures in people with metabolic diseases are poorly understood. Reduced heat tolerance in obesity might be due to impairments in blood flow and sweat production (Vroman et al., 1983). The reduced sweating ability is possibly linked to a decreased body surface area-to-body mass ratio in a person with obesity as compared to a leaner person. Individuals with type 2 diabetes also exhibit reduced skin blood flow in response to local and whole-body heating, likely due to impaired endothelial function (Meade et al., 2020). However, chronic exposure to mild electrical stimulation with heat shock improves visceral adiposity, glucose homeostasis, and insulin sensitivity in people with type 2 diabetes (Kondo et al., 2014). This paradox suggests that increasing heat tolerance by repeated acute exposures to heat might mitigate heat-induced cardiovascular events in individuals with metabolic diseases.

Heat stress from both exercise and environmental factors can increase thermal strain in unacclimated individuals (Figure 5). Acute exercise increases core body temperature and high-intensity exercise can lead to heat illness consisting of symptoms ranging from minor cramps and syncope to major heat stroke, even in highly trained athletes (Charlot et al., 2017). The capacity to dissipate an exercise-induced elevation in body temperature is reduced in people with type 2 diabetes, but this can be overcome by regular exercise training, which is associated with improved heat tolerance (Kenny et al., 2016). Regular exercise training also reduces cardiovascular mortality and improves glucose control in people with type 2 diabetes (Savikj and Zierath, 2020). At a molecular level, exercise training increases heat shock protein abundance, a process that could contribute to the beneficial effects of exercise to enhance insulin sensitivity (Archer et al., 2018). Individuals with obesity or type 2 diabetes exhibit decreased levels of heat shock proteins in skeletal muscle (Chung et al., 2008). This decrease is reversible, and induction of heat shock proteins by mild electrical stimulation with heat shock improves visceral adiposity as well as plasma glucose and insulin levels (Kondo et al., 2014). Regular exposure to thermal stressors, such as exercise or environmental temperature, may improve heat tolerance through overlapping adaptive mechanistic responses (sweat volume and composition, body water, heart rate), thereby improving metabolism and decreasing risk of cardio-pulmonary events in individuals exposed to extreme ambient temperatures.



1. Download : Download high-res image (630KB)
2. Download : Download full-size image

Figure 5. Beneficial effects of exercise training on metabolic risk

Decreased thermal tolerance, increased chronic inflammation, deregulated circadian rhythms, and poor glucose control worsen with age and disease development, increasing the susceptibility to life-threatening infections and extreme temperatures and eventually the risk of cardiovascular events (CVD). Exercise triggers acute and transient changes in inflammation and body temperature. These acute events are required for the beneficial effects of exercise training. Exercise training slows the progression of chronic inflammation, limits the decrease in heat-tolerance, and helps synchronize circadian clocks, thereby improving metabolism.

Presently, most of the human population lives under conditions of thermoneutrality, which is made possible due to appropriate clothing and heating systems in homes and workplaces. Reduced energy expenditure, due to the comforts of our modern society and the decline in our prolonged exposure to cold environments, may contribute to the worldwide rise in obesity, although this is difficult to firmly establish. However, there is a clear link between thermal regulation, metabolic diseases, and associated complications. Most of the temperature-related cardio-pulmonary events occur on moderately hot and moderately cold days (Gasparrini et al., 2015), suggesting that steady increases in the average global temperature has the potential to impact the numbers of these events worldwide. The combination of an epidemic of obesity and type 2 diabetes, juxtaposed with an aging population and climate change, may potentially lead to a dramatic increase in cardio-vascular morbimortality. Understanding the molecular basis of heat intolerance in people with obesity or type 2 diabetes could open novel preventative and therapeutic perspectives.

29.6.1.1.6: INFLAMMATORY RESPONSES

The stress that temperature, obesity, diabetes, exercise, and food exert on organismal homeostasis triggers activation of the immune system and different states of metabolic inflammation. The immune system is composed of specialized cells present in every organ that protect against a wide variety of insults, including infections, mechanical injuries, and a variety of diseases. The immune response comes in waves, starting with a pro-inflammatory activation and finishing with a resolving anti-inflammatory phase (Feehan and Gilroy, 2019). When the immune system fails to recover after an insult, a chronic inflammatory state occurs, leading to long-term deleterious consequences. This typically happens in obesity, where immune cells infiltrate tissues and lead to chronic low-grade inflammation, associated with increased risk of cardiovascular complications. The association of inflammation with type 2 diabetes and obesity has been extensively studied, as evidenced by the rapid development of the field of "*immunometabolism*," which includes the analysis of the complex interactions between metabolic and inflammatory pathways in immune and metabolic tissues (Lee et al., 2018).

Obesity and type 2 diabetes are associated with an accumulation of immune cells in key tissues involved in metabolic homeostasis. A link between metabolic diseases and immunology emerged with the detection of macrophage infiltration in adipose tissue, followed by the discovery that lymphocytes, neutrophils, and other specific subtypes of immune cells accumulate not only in adipose tissue but also in skeletal muscle and liver (Hotamisligil, 2017). Even neuroinflammation is part of the systemic inflammatory syndrome in metabolic

diseases (Cai, 2013). The accumulation of triglycerides in adipocytes increases adipocyte size (hypertrophy) and number (hyperplasia), resulting in the rapid expansion of adipose tissue, which triggers hypoxia and the production of soluble mediators likely responsible for the attraction of immune cells. The first immune cells reaching adipose tissue are likely attracted to support tissue remodeling in a beneficial manner, but the chronic increase in adipose tissue volume and the establishment of a new obese steady-state leads to increasing lipolysis and circulating free fatty acids, which activate immune cells toward a pro-inflammatory phenotype and promote the establishment of chronic inflammation (Lee et al., 2018). Immune cells also respond to metabolic changes and are susceptible to the deleterious effects of an excessive lipid or glucose accumulation (i.e., "*lipotoxicity*" or "*glucotoxicity*"), as well as other metabolism-related danger signals that are released by tissues during metabolic stress (Wang et al., 2020b). The composition and phenotype of circulating immune cells is altered in blood of individuals with obesity, with an increase in CD16+ monocytes, and immune cell activation in response to high concentrations of glucose or fatty acids (Pillon et al., 2016). These findings suggest that the immune system is profoundly affected by whole-body glucose and lipid homeostasis.

The mechanisms by which non-adipose tissues establish a state of inflammation is unclear. However, lipotoxicity, including the excessive accumulation of toxic lipid mediators such as ceramides, diacylglycerol, or acylcarnitine, and increased levels of circulating free fatty acids likely play a role. In addition, activated immune cells primed to respond to metabolism-related danger signals can impair whole-body metabolism. There is ample evidence to suggest that inflammation is associated with the development of metabolic diseases and the ensuing complications, but pharmacological targeting of pathways controlling immunometabolism has shown limited benefits for the treatment of metabolic diseases (Pålsson-McDermott and O'Neill, 2020). Perhaps the key to successful clinical intervention will be to identify relevant patient groups early, before the manifestation of a chronic low-grade inflammatory state.

Acute exercise, especially intense and/or eccentric exercise, triggers an acute inflammatory response, which is necessary for skeletal muscle repair and adaptations to exercise training. Exercise training has beneficial anti-inflammatory effects (Gleeson et al., 2011). Thus, repeated peaks of inflammation triggered by acute bouts of moderate intensity exercise may be beneficial to reduce long-term basal concentrations of pro-inflammatory mediators. In severely obese individuals, combining exercise and dietary interventions can reduce macrophage infiltration and pro-inflammatory polarization in adipose tissue (Bruun et al., 2006). The anti-inflammatory effects of exercise could be secondary to an increased capacity for fatty acid utilization, as exercise training in people with obesity or type 2 diabetes reduces the level of deleterious lipid species such as DAG, acetylcarnitines, and ceramides in skeletal muscle (Lancaster and Febbraio, 2014). However, exercise training in healthy individuals also improves insulin sensitivity without changes in these lipid species, making the role of intramyocellular lipids on insulin sensitivity ambiguous and perhaps more relevant in an obesity context (Reidy et al., 2020).

29.6.1.1.7: UNSUSPECTED CAUSES

Currently known genetic, lifestyle, and environmental risk factors only partly explain the development of obesity and diabetes. Other yet unknown factors must be at play. A recent example of potential novel causes of diabetes is the high prevalence of extreme hyperglycemia/ketoacidosis in patients not known to have diabetes admitted to hospital with COVID-19 (Rubino et al., 2020). This seems both more common and more severe than has been seen with other infections/serious illnesses, so it may not represent "stress hyperglycemia" or unmasking of pre-existing, undiagnosed diabetes. Instead, these observations may suggest a specific pathological entity. The SARS-CoV-2 spike protein penetrates cell membranes by binding to the angiotensin converting enzyme (ACE) 2 receptor. This receptor is present on pancreatic beta cells (Hamming et al., 2004). Infection may result in acute loss of insulin secretory capacity and/or beta cell destruction (Apicella et al., 2020). ACE2 receptor is also present on adipocytes so that SARS-CoV-2 may also exacerbate chronic inflammation in adipose tissue (Kassir, 2020).

The mechanisms whereby widely accepted risk factors such as obesity result in disease may have novel aspects. It is generally assumed that individuals with type 2 diabetes who are not obese have a different pathophysiological cause unrelated to weight. However, this belief has been challenged recently. The concept of a "personal fat threshold" arose from observations that the median BMI in the UK Prospective Diabetes Study was only 28 kg/m^2 (Taylor and Holman, 2015) and that reversal of type 2 diabetes by weight loss could be achieved equally successfully in individuals with higher and lower BMI (Lim et al., 2011). The underlying mechanism appears to be lipotoxicity, an individual's propensity to accumulate liver and pancreas fat, and their susceptibility to the adverse effects of fat accumulation. At any given body weight or BMI, at-risk individuals will accumulate more liver fat and be more susceptible to developing hepatic insulin resistance at any given liver fat content. The subsequent increase in VLDL-TG export from the liver drives fat accumulation in the pancreas and declining insulin secretion, both also dependent on the individual's susceptibility. Remission of type 2 diabetes by weight loss is accompanied by reduction in liver and pancreatic fat, decreased hepatic VLDL export, and increased insulin secretion (Al-Mrabeh et al., 2020). Conversely, weight re-gain leading to re-emergence of diabetes is associated with increased liver fat export and pancreatic fat, with recurrent pancreatic dysfunction. The importance of these observations underscores the usefulness of weight loss in the management of diabetic individuals even of normal weight. However, diabetes does not remit in everyone following substantial weight loss, so weight loss is not a universal panacea. Further work is needed to establish if this relates to longer duration diabetes, perhaps with irreversible beta cell damage, or to different pathological mechanisms of disease.

29.6.1.1.8: BENDING THE CURVE

Weight loss is clearly the key to reducing rates of obesity and type 2 diabetes, with considerable individual and societal benefits. There is a continuum of action required in prevention of obesity and diabetes, and management and care if they develop (Chan et al., 2020). Many intervention programs have demonstrated successful short-term weight loss and reversal of diabetes, but perhaps the bigger challenge is in preventing weight re-gain (Forouhi et al., 2018). There may be a weight "set-point," at which compensatory hormonal, metabolic and neurochemical mechanisms prevent further weight loss and drive weight regain (Blüher, 2019). However, a significant proportion of individuals who lose a substantial amount of weight, whether by diet or bariatric surgery, do not regain weight over years and maintain the metabolic benefits of the initial weight loss. Thus, weight loss programs must have two parts: an initial phase of weight loss, followed by a weight maintenance program. Obviously, reduction in energy intake by some means is essential for weight loss. Exactly how this is achieved is probably less important than an individual's ability to adhere to the program long term (Johnston et al., 2014). The benefits of one regimen over another have been debated (Forouhi et al., 2018), but no one size fits all, and many different approaches are needed.

Understanding the influence of social and cultural aspects in the development and management of obesity and diabetes is also crucial (Blüher, 2019). Individuals from socially deprived backgrounds are more likely to be at high risk to develop obesity and type 2 diabetes, to have poorer glycemic control, to develop more complications, and to have a greater reduction in life expectancy (Chan et al., 2020). Identifying and overcoming barriers to participation in screening and prevention programs and in diabetes and obesity care are vital. Most programs do not reach individuals from ethnic minorities or low socioeconomic class, who are most at need (Timpel et al., 2019). Involving overweight and obese individuals from a wide diversity of backgrounds in the identification of barriers to adherence and then in the design of weight loss and diabetes prevention/reversal programs is essential to improve engagement.

29.6.1.1.9: PERSONALIZED MEDICINE

Although we talk about type 2 diabetes as one disease, this "blanket" diagnosis covers important heterogeneity (Ahlqvist et al., 2018). Only rarely is the heterogeneity obvious and explainable: slim rather than obese, or young age at presentation with a striking family history in monogenic diabetes. On many occasions, individuals with apparently similar phenotypes have very different clinical courses and respond quite differently to glucose lowering agents. Dissecting out particular forms of "type 2 diabetes," whether by genetic analyses and risk scores or by improving our understanding of the underlying pathophysiological bases to dysglycemia, is currently possible at the population level but remains extremely difficult at an individual level. As a result, the selection of glucose-lowering agents for individuals is a "best guess" approach, far removed from personalized medicine.

Personalized medicine is defined simply as the right treatment for the right person at the right time. The recent American Diabetes Association/European Association for the Study of Diabetes consensus report describes the ambition to personalize all aspects of an individual's diabetes, including precision diagnosis, lifestyle and pharmacological management, and prognosis (Chung et al., 2020). Currently, for a very small number of individuals (for example, those with congenital leptin deficiency [Montague et al., 1997] and GCK-MODY [MODY 2] [Froguel et al., 1992; Hattersley et al., 1992]), precision diagnosis is possible. However, there are major challenges in precision diagnosis for individuals with the polygenic common forms of obesity and type 2 diabetes. Likewise, there are only a small number of examples of precision therapeutics (for example, leptin for management of severe obesity in congenital leptin deficiency [Farooqi et al., 2002] and sulphonylureas rather than insulin for individuals with neonatal diabetes due to mutations in the genes encoding the potassium channel [KCNJ11 and ABCC8] [Pearson et al., 2006]). For many individuals with obesity and type 2 diabetes, we have extremely blunt "precision" tools. For example, analysis of data from participants in the RECORD and ADOPT trials demonstrated that individuals with insulin resistance have a greater sustained fall in HbA1c on thiazolidenediones compared to sulphonylureas (Dennis et al., 2018). Additionally, there are benefits of SGLT2 inhibitors in individuals with high cardiovascular risk and/or renal disease (Lo et al., 2020). Work is beginning to examine possibilities for personalization of lifestyle measures.

29.6.1.1.10: GRAND CHALLENGES

Prevention of obesity is probably the most important factor in reducing the prevalence of obesity and related metabolic diseases. This will require action at an individual and societal level (Chan et al., 2020). Societal action is necessary in many areas, including changes to road, rail, and cycling transport plans to encourage increased physical activity, as well as negotiations with the food industry (Chan et al., 2020). Governments also need effective communication plans that reach all sections of society (Timpel et al., 2019). Moreover, different strategies are required for different life stages. The lifestyles of almost everyone must change radically, and this must be facilitated by appropriate action by governments and many branches of industry. The challenges to overcome the status quo and vested interests are considerable.

There is abundant evidence that many individuals with obesity at high risk of metabolic disease can lose substantial amounts of weight, reversing pre-diabetes and diabetes. A significant proportion then maintain the weight loss and improved metabolic status for years. Weight loss programs are projected to be more effective per quality-adjusted life year and cost-saving over a lifetime compared to standard care in individuals with type 2 diabetes (Xin et al., 2020). The challenge then is to expand and adapt these successful programs so that all individuals can access them and be supported through them. We must work with individuals who find current programs inappropriate for them, identifying barriers to participation and working together to develop practical solutions. Most current weight loss programs center on improving basic diet, with or without advice on exercise (Forouhi et al., 2018). Further incremental benefit may well be obtained by incorporating additional "personalized" measures, perhaps based on genes, occupation, and inflammatory status, such as advice on specific

micronutrients, timing of food intake and exercise, and light exposure. However, the challenge will be to ensure that the message does not become so complex that adherence falls.

There is a particular challenge for young people (Chan et al., 2020). The WHO estimated in 2016 that world-wide, 340 million children and adolescents aged 5–19 years were overweight or obese and, in 2019, that 38 million children aged <5 years were overweight/obese (https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight). Associated with this, type 2 diabetes is increasingly diagnosed in children, adolescents, and young adults (IDF Diabetes Atlas 9th edition 2019, www.diabetesatlas.org). A recent meta-analysis has demonstrated the greater impact of type 2 diabetes presenting at younger age: each one-year increase in age at diabetes diagnosis was associated with a 4%, 3%, and 5% decreased risk of all-cause mortality, macrovascular, and microvascular disease respectively (Nanayakkara et al., 2020). These changes underscore the need to prevent obesity and/or manage it appropriately in young people.

Paralleling the rise in obesity in younger people is the rise in the number of women with hyperglycemia during pregnancy. The IDF estimated that, in 2017, 16% of women with live births had some form of hyperglycemia during pregnancy, and that 86% of them had gestational diabetes (IDF Diabetes Atlas 9th edition 2019, www.diabetesatlas.org). In addition to the immediate maternal and fetal adverse effects of hyperglycemia during pregnancy, many of these women will develop type 2 diabetes in the subsequent 5-10 years. There are also longer-term consequences to the offspring of increased risk of obesity, type 2 diabetes, hypertension, and cardiovascular disease (Catalano and Shankar, 2017). Some of these adverse consequences are now being reported over several generations of offspring, implicating an epigenetic influence (Catalano and Shankar, 2017). Thus, in addition to the immediate management of the index pregnancy, it is extremely important that further studies of the index in women, their children, and potentially subsequent generations are conducted urgently.

Low levels of fitness are a risk factor for hospitalizations and all-cause mortality, and predict morbidity after surgical interventions (West et al., 2016). During the COVID-19 pandemic, public health recommendations regarding confinement and closure of recreation areas decreased daily activity in the general population (Sánchez-Sánchez et al., 2020), aggravating already high levels of inactivity in most countries (https://www.who.int/news-room/fact-sheets/detail/physical-activity). In young adults who only developed mild symptoms, COVID-19 decreased the predicted maximal aerobic capacity (Crameri et al., 2020), while persons more severely infected with SARS-CoV-2 exhibited anorexia and skeletal muscle loss, aggravated by long hospital stays, raising the question of whether COVID-19 could be a major cause of cachexia and sarcopenia (Morley et al., 2020). As obesity and type 2 diabetes are risk factors for COVID-19 complications, the underlying inflammatory conditions in combination with impaired skeletal muscle function may contribute to worse outcomes after infection (Guisado-Vasco et al., 2020). Whether exercise can protect against viral infection or influence disease severity is unclear, but the benefits of physical activity to prevent skeletal muscle wasting are important factors for prevention and rehabilitation of people in risks groups. Understanding the molecular mechanisms underlying the beneficial effects of physical exercise as an inflammatory modulator could thus potentially prevent or mitigate complications due to unexpected infections (da Silveira et al., 2021; Krause et al., 2020).

29.6.1.1.11: ON THE HORIZON

How do we move forward? Progress will only come if we tackle the problems at both a population and an individual level. Putting into practice what we already know will benefit many individuals (Chan et al., 2020). Incorporating the newer evidence described above—for example, around the timing of eating and exercise—and light exposure, in ways that do not overwhelm people, will bring added benefits. Better personalization of all aspects of prevention, management, and care should help adherence. In-depth large-scale analysis of genetic and environmental factors may help clarify why people respond differently to the whole gamut of care, allow stratification into refined sub-groups with specific risk factors and genetic predispositions, and potentially thus optimize the efficacy of both lifestyle and pharmacological interventions. Ongoing initiatives like the Innovative Medicines Initiative (www.imi.europa.eu) have demonstrated that combining large databases from multiple public and private organizations is possible, generating power to tackle relevant genetic and biomarker questions. Such initiatives, bringing together diverse stakeholders with people with obesity or diabetes, are essential in our efforts to provide personalized, timely, affordable, and equitable access to high-quality health interventions, with the aim of improving health outcomes for all.

29.6.1.2: ACKNOWLEDGMENTS

N.J.P. was supported by an Individual Fellowship from the Marie Skłodowska-Curie Actions (European Commission, 704978). J.R.Z. was supported from the Swedish Research Council (Vetenskapsrådet) (2015-00165), a Novo Nordisk Foundation Challenge Grant (NNF14OC0011493), and the Novo Nordisk Foundation Center for Basic Metabolic Research at the University of Copenhagen (NNF18CC0034900). R.J.F.L. was supported by the National Institutes of Health (R01DK110113; R01DK107786; R01HL142302; R01 DK124097) and the Novo Nordisk Foundation Center for Basic Metabolic Research at the University of Copenhagen (Alliance 190503).

29.6.1.2.1: AUTHOR CONTRIBUTIONS

N.J.P., R.L., S.M., and J.R.Z. wrote the manuscript. All authors read and approved the final version of the manuscript.

29.6.1.2.2: DECLARATION OF INTERESTS

The authors declare no competing interests. J.R.Z. is member of Cell Advisory Board.

29.6.1.3: REFERENCES

Achebak et al., 2019

H. Achebak, D. Devolder, J. Ballester

Trends in temperature-related age-specific and sex-specific mortality from cardiovascular diseases in Spain: a national time-series analysis

Lancet Planet. Health, 3 (2019), pp. e297-e306

ArticleDownload PDFView Record in ScopusGoogle Scholar

Ahlqvist et al., 2018

E. Ahlqvist, P. Storm, A. Käräjämäki, M. Martinell, M. Dorkhan, A. Carlsson, P. Vikman, R.B. Prasad, D.M. Aly, P. Almgren, *et al.*

Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables

Lancet Diabetes Endocrinol., 6 (2018), pp. 361-369

ArticleDownload PDFView Record in ScopusGoogle Scholar

Al-Mrabeh et al., 2020

A. Al-Mrabeh, S.V. Zhyzhneuskaya, C. Peters, A.C. Barnes, S. Melhem, A. Jesuthasan, B. Aribisala, K.G. Hollingsworth, G. Lietz, J.C. Mathers, *et al.*

Hepatic lipoprotein export and remission of human type 2 diabetes after weight loss

Cell Metab., 31 (2020), pp. 233-249 e234

View Record in ScopusGoogle Scholar

Ameln et al., 2005

H. Ameln, T. Gustafsson, C.J. Sundberg, K. Okamoto, E. Jansson, L. Poellinger, Y. Makino

Physiological activation of hypoxia inducible factor-1 in human skeletal muscle

FASEB J., 19 (2005), pp. 1009-1011

CrossRefView Record in ScopusGoogle Scholar

American Diabetes Association, 2020

American Diabetes Association

2. Classification and diagnosis of diabetes: Standards of medical care in Diabetes-2020

Diabetes Care, 43 (Suppl 1) (2020), pp. S14-S31

Google Scholar

Apicella et al., 2020

M. Apicella, M.C. Campopiano, M. Mantuano, L. Mazoni, A. Coppelli, S. Del Prato

COVID-19 in people with diabetes: understanding the reasons for worse outcomes

Lancet Diabetes Endocrinol., 8 (2020), pp. 782-792

ArticleDownload PDFView Record in ScopusGoogle Scholar

Archer et al., 2018

A.E. Archer, A.T. Von Schulze, P.C. Geiger

Exercise, heat shock proteins and insulin resistance

Philos. Trans. R. Soc. Lond. B Biol. Sci., 373 (2018), 10.1098/rstb.2016.0529

Google Scholar

Areta et al., 2013

J.L. Areta, L.M. Burke, M.L. Ross, D.M. Camera, D.W. West, E.M. Broad, N.A. Jeacocke, D.R. Moore, T. Stellingwerff, S.M. Phillips, *et al.*

Timing and distribution of protein ingestion during prolonged recovery from resistance exercise alters myofibrillar protein synthesis

J. Physiol., 591 (2013), pp. 2319-2331

CrossRefView Record in ScopusGoogle Scholar

Balsalobre et al., 2000

A. Balsalobre, S.A. Brown, L. Marcacci, F. Tronche, C. Kellendonk, H.M. Reichardt, G. Schütz, U. Schibler

Resetting of circadian time in peripheral tissues by glucocorticoid signaling

Science, 289 (2000), pp. 2344-2347

CrossRefView Record in ScopusGoogle Scholar

Barrès and Zierath, 2016

R. Barrès, J.R. Zierath

The role of diet and exercise in the transgenerational epigenetic landscape of T2DM

Nat. Rev. Endocrinol., 12 (2016), pp. 441-451

CrossRefView Record in ScopusGoogle Scholar

Barrès et al., 2012

R. Barrès, J. Yan, B. Egan, J.T. Treebak, M. Rasmussen, T. Fritz, K. Caidahl, A. Krook, D.J. O'Gorman, J.R. Zierath

Acute exercise remodels promoter methylation in human skeletal muscle

Cell Metab., 15 (2012), pp. 405-411

ArticleDownload PDFView Record in ScopusGoogle Scholar

Barres et al., 2013

R. Barres, H. Kirchner, M. Rasmussen, J. Yan, F.R. Kantor, A. Krook, E. Näslund, J.R. Zierath

Weight loss after gastric bypass surgery in human obesity remodels promoter methylation

Cell Rep., 3 (2013), pp. 1020-1027

ArticleDownload PDFView Record in ScopusGoogle Scholar

Barron et al., 2020

E. Barron, C. Bakhai, P. Kar, A. Weaver, D. Bradley, H. Ismail, P. Knighton, N. Holman, K. Khunti, N. Sattar, *et al.*

Associations of type 1 and type 2 diabetes with COVID-19-related mortality in England: a whole-population study

Lancet Diabetes Endocrinol., 8 (2020), pp. 813-822

ArticleDownload PDFView Record in ScopusGoogle Scholar

Barroso and McCarthy, 2019

I. Barroso, M.I. McCarthy

The genetic basis of metabolic disease

Cell, 177 (2019), pp. 146-161

ArticleDownload PDFView Record in ScopusGoogle Scholar

Blüher, 2019

M. Blüher

Obesity: global epidemiology and pathogenesis

Nat. Rev. Endocrinol., 15 (2019), pp. 288-298

CrossRefView Record in ScopusGoogle Scholar

Bošković and Rando, 2018

A. Bošković, O.J. Rando

Transgenerational epigenetic inheritance

Annu. Rev. Genet., 52 (2018), pp. 21-41

CrossRefView Record in ScopusGoogle Scholar

Bray et al., 2017

G.A. Bray, K.K. Kim, J.P.H. Wilding, World Obesity Federation

Obesity: a chronic relapsing progressive disease process. A position statement of the World Obesity Federation

Obes. Rev., 18 (2017), pp. 715-723

CrossRefView Record in ScopusGoogle Scholar

Bruun et al., 2006

J.M. Bruun, J.W. Helge, B. Richelsen, B. Stallknecht

Diet and exercise reduce low-grade inflammation and macrophage infiltration in adipose tissue but not in skeletal muscle in severely obese subjects

Am. J. Physiol. Endocrinol. Metab., 290 (2006), pp. E961-E967
CrossRefView Record in ScopusGoogle Scholar

Cai, 2013
D. Cai
Neuroinflammation and neurodegeneration in overnutrition-induced diseases
Trends Endocrinol. Metab., 24 (2013), pp. 40-47
ArticleDownload PDFView Record in ScopusGoogle Scholar

Catalano and Shankar, 2017
P.M. Catalano, K. Shankar
Obesity and pregnancy: mechanisms of short term and long term adverse consequences for mother and child
BMJ, 356 (2017), p. j1
CrossRefView Record in ScopusGoogle Scholar

Chaix et al., 2014
A. Chaix, A. Zarrinpar, P. Miu, S. Panda
Time-restricted feeding is a preventative and therapeutic intervention against diverse nutritional challenges
Cell Metab., 20 (2014), pp. 991-1005
ArticleDownload PDFView Record in ScopusGoogle Scholar

Chan et al., 2020
J.C.N. Chan, L.L. Lim, N.J. Wareham, J.E. Shaw, T.J. Orchard, P. Zhang, E.S.H. Lau, B. Eliasson, A.P.S. Kong, M. Ezzati, *et al.*
The Lancet Commission on diabetes: using data to transform diabetes care and patient lives
Lancet, 396 (10267) (2020), pp. 2019-2082, 10.1016/S0140-6736(20)32374-6
ArticleDownload PDFView Record in ScopusGoogle Scholar

Charlot et al., 2017
K. Charlot, C. Faure, S. Antoine-Jonville
Influence of hot and cold environments on the regulation of energy balance following a single exercise session: A mini-review
Nutrients, 9 (2017), p. 9
Google Scholar

Chouchani et al., 2019
E.T. Chouchani, L. Kazak, B.M. Spiegelman
New advances in adaptive thermogenesis: UCP1 and beyond
Cell Metab., 29 (2019), pp. 27-37
ArticleDownload PDFView Record in ScopusGoogle Scholar

Chung et al., 2008
J. Chung, A.K. Nguyen, D.C. Henstridge, A.G. Holmes, M.H. Chan, J.L. Mesa, G.I. Lancaster, R.J. Southgate, C.R. Bruce, S.J. Duffy, *et al.*
HSP72 protects against obesity-induced insulin resistance
Proc. Natl. Acad. Sci. USA, 105 (2008), pp. 1739-1744
CrossRefView Record in ScopusGoogle Scholar

Chung et al., 2020
W.K. Chung, K. Erion, J.C. Florez, A.T. Hattersley, M.F. Hivert, C.G. Lee, M.I. McCarthy, J.J. Nolan, J.M. Norris, E.R. Pearson, *et al.*
Precision medicine in diabetes: a Consensus Report from the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD)
Diabetologia, 63 (2020), pp. 1671-1693
CrossRefView Record in ScopusGoogle Scholar

Cienfuegos et al., 2020
S. Cienfuegos, K. Gabel, F. Kalam, M. Ezpeleta, E. Wiseman, V. Pavlou, S. Lin, M.L. Oliveira, K.A. Varady

Effects of 4- and 6-h time-restricted feeding on weight and cardiometabolic health: A Randomized Controlled Trial in Adults with Obesity.

Cell Metab., 32 (2020), pp. 366-378.e363

View Record in ScopusGoogle Scholar

Crameri et al., 2020

G.A.G. Crameri, M. Bielecki, R. Züst, T.W. Buehrer, Z. Stanga, J.W. Deuel

Reduced maximal aerobic capacity after COVID-19 in young adult recruits, Switzerland, May 2020

Euro Surveill., 25 (2020), p. 2001542

Google Scholar

Crosby et al., 2019

P. Crosby, R. Hamnett, M. Putker, N.P. Hoyle, M. Reed, C.J. Karam, E.S. Maywood, A. Stangherlin, J.E. Chesham, E.A. Hayter, *et al.*

Insulin/IGF-1 drives PERIOD synthesis to entrain circadian rhythms with feeding time

Cell, 177 (2019), pp. 896-909 e820

View Record in ScopusGoogle Scholar

da Silveira et al., 2021

M.P. da Silveira, K.K. da Silva Fagundes, M.R. Bizuti, E. Starck, R.C. Rossi, E.S.D.T. de Resende

Physical exercise as a tool to help the immune system against COVID-19: an integrative review of the current literature

Clin. Exp. Med., 21 (2021), pp. 15-28, 10.1007/s10238-020-00650-3

CrossRefView Record in ScopusGoogle Scholar

de Castro Barbosa et al., 2015

T. de Castro Barbosa, L.R. Ingerslev, P.S. Alm, S. Versteyhe, J. Massart, M. Rasmussen, I. Donkin, R. Sjögren, J.M. Mudry, L. Vetterli, *et al.*

High-fat diet reprograms the epigenome of rat spermatozoa and transgenerationally affects metabolism of the offspring

Mol. Metab., 5 (2015), pp. 184-197

Google Scholar

Dennis et al., 2018

J.M. Dennis, B.M. Shields, A.G. Jones, E.R. Pearson, A.T. Hattersley, W.E. Henley, MASTERMIND consortium

Evaluating associations between the benefits and risks of drug therapy in type 2 diabetes: a joint modeling approach

Clin. Epidemiol., 10 (2018), pp. 1869-1877

CrossRefView Record in ScopusGoogle Scholar

Elks et al., 2012

C.E. Elks, M. den Hoed, J.H. Zhao, S.J. Sharp, N.J. Wareham, R.J. Loos, K.K. Ong

Variability in the heritability of body mass index: a systematic review and meta-regression

Front. Endocrinol. (Lausanne), 3 (2012), p. 29

Google Scholar

Farooqi et al., 2002

I.S. Farooqi, G. Matarese, G.M. Lord, J.M. Keogh, E. Lawrence, C. Agwu, V. Sanna, S.A. Jebb, F. Perna, S. Fontana, *et al.*

Beneficial effects of leptin on obesity, T cell hyporesponsiveness, and neuroendocrine/metabolic dysfunction of human congenital leptin deficiency

J. Clin. Invest., 110 (2002), pp. 1093-1103

View Record in ScopusGoogle Scholar

Feehan and Gilroy, 2019

K.T. Feehan, D.W. Gilroy

Is resolution the end of inflammation?

Trends Mol. Med., 25 (2019), pp. 198-214

ArticleDownload PDFView Record in ScopusGoogle Scholar

Forouhi et al., 2018

N.G. Forouhi, A. Misra, V. Mohan, R. Taylor, W. Yancy

Dietary and nutritional approaches for prevention and management of type 2 diabetes

BMJ, 361 (2018), p. k2234

CrossRefView Record in ScopusGoogle Scholar

Freese et al., 2017

J. Freese, R.J. Klement, B. Ruiz-Núñez, S. Schwarz, H. Lötzerich

The sedentary (r)evolution: Have we lost our metabolic flexibility?

F1000Res., 6 (2017), p. 1787

CrossRefView Record in ScopusGoogle Scholar

Froguel et al., 1992

P. Froguel, M. Vaxillaire, F. Sun, G. Velho, H. Zouali, M.O. Butel, S. Lesage, N. Vionnet, K. Clément, F. Fougerousse, *et al.*

Close linkage of glucokinase locus on chromosome 7p to early-onset non-insulin-dependent diabetes mellitus

Nature, 356 (1992), pp. 162-164

View Record in ScopusGoogle Scholar

Gabriel and Zierath, 2019

B.M. Gabriel, J.R. Zierath

Circadian rhythms and exercise - re-setting the clock in metabolic disease

Nat. Rev. Endocrinol., 15 (2019), pp. 197-206

CrossRefView Record in ScopusGoogle Scholar

Gamble et al., 2014

K.L. Gamble, R. Berry, S.J. Frank, M.E. Young

Circadian clock control of endocrine factors

Nat. Rev. Endocrinol., 10 (2014), pp. 466-475

CrossRefView Record in ScopusGoogle Scholar

Gasparrini et al., 2015

A. Gasparrini, Y. Guo, M. Hashizume, E. Lavigne, A. Zanobetti, J. Schwartz, A. Tobias, S. Tong, J. Rocklöv, B. Forsberg, *et al.*

Mortality risk attributable to high and low ambient temperature: a multicountry observational study

Lancet, 386 (2015), pp. 369-375

ArticleDownload PDFView Record in ScopusGoogle Scholar

Gleeson et al., 2011

M. Gleeson, N.C. Bishop, D.J. Stensel, M.R. Lindley, S.S. Mastana, M.A. Nimmo

The anti-inflammatory effects of exercise: mechanisms and implications for the prevention and treatment of disease

Nat. Rev. Immunol., 11 (2011), pp. 607-615

CrossRefGoogle Scholar

Guisado-Vasco et al., 2020

P. Guisado-Vasco, M. Cano-Megías, M. Rodríguez-López, I.M. de-Luna-Boquera, D. Carnevali-Ruiz, Immunosuppressants Against COVID-19 Working Team

COVID-19 and metabolic syndrome: NF-κB activation. Crossroads

Trends Endocrinol. Metab., 31 (2020), pp. 802-803

ArticleDownload PDFView Record in ScopusGoogle Scholar

Hajat et al., 2017

S. Hajat, A. Haines, C. Sarran, A. Sharma, C. Bates, L.E. Fleming

The effect of ambient temperature on type-2-diabetes: case-crossover analysis of 4+ million GP consultations across England

Environ. Health, 16 (2017), p. 73

View Record in ScopusGoogle Scholar

Hamming et al., 2004

I. Hamming, W. Timens, M.L. Bulthuis, A.T. Lely, G. Navis, H. van Goor

Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis

J. Pathol., 203 (2004), pp. 631-637

View Record in ScopusGoogle Scholar

Hatori et al., 2012

M. Hatori, C. Vollmers, A. Zarrinpar, L. DiTacchio, E.A. Bushong, S. Gill, M. Leblanc, A. Chaix, M. Joens, J.A. Fitzpatrick, *et al.*

Time-restricted feeding without reducing caloric intake prevents metabolic diseases in mice fed a high-fat diet

Cell Metab., 15 (2012), pp. 848-860

ArticleDownload PDFView Record in ScopusGoogle Scholar

Hattersley and Patel, 2017

A.T. Hattersley, K.A. Patel

Precision diabetes: learning from monogenic diabetes

Diabetologia, 60 (2017), pp. 769-777

CrossRefView Record in ScopusGoogle Scholar

Hattersley et al., 1992

A.T. Hattersley, R.C. Turner, M.A. Permutt, P. Patel, Y. Tanizawa, K.C. Chiu, S. O'Rahilly, P.J. Watkins, J.S. Wainscoat

Linkage of type 2 diabetes to the glucokinase gene

Lancet, 339 (1992), pp. 1307-1310

ArticleDownload PDFView Record in ScopusGoogle Scholar

Holman et al., 2020

N. Holman, P. Knighton, P. Kar, J. O'Keefe, M. Curley, A. Weaver, E. Barron, C. Bakhai, K. Khunti, N.J. Wareham, *et al.*

Risk factors for COVID-19-related mortality in people with type 1 and type 2 diabetes in England: a population-based cohort study

Lancet Diabetes Endocrinol., 8 (2020), pp. 823-833

ArticleDownload PDFView Record in ScopusGoogle Scholar

Hotamisligil, 2017

G.S. Hotamisligil

Inflammation, metaflammation and immunometabolic disorders

Nature, 542 (2017), pp. 177-185

CrossRefView Record in ScopusGoogle Scholar

Huang et al., 2012

C. Huang, A.G. Barnett, X. Wang, S. Tong

Effects of extreme temperatures on years of life lost for cardiovascular deaths: a time series study in Brisbane, Australia

Circ. Cardiovasc. Qual. Outcomes, 5 (2012), pp. 609-614

CrossRefView Record in ScopusGoogle Scholar

Ji et al., 2019

Y. Ji, A.M. Yiorkas, F. Frau, D. Mook-Kanamori, H. Staiger, E.L. Thomas, N. Atabaki-Pasdar, A. Campbell, J. Tyrrell, S.E. Jones, *et al.*

Genome-wide and abdominal MRI data provide evidence that a genetically determined favorable adiposity phenotype is characterized by lower ectopic liver fat and lower risk of type 2 diabetes, heart disease, and hypertension

Diabetes, 68 (2019), pp. 207-219

CrossRefView Record in ScopusGoogle Scholar

Johnston et al., 2014

B.C. Johnston, S. Kanters, K. Bandayrel, P. Wu, F. Naji, R.A. Siemieniuk, G.D. Ball, J.W. Busse, K. Thorlund, G. Guyatt, *et al.*

Comparison of weight loss among named diet programs in overweight and obese adults: a meta-analysis

JAMA, 312 (2014), pp. 923-933

CrossRefView Record in ScopusGoogle Scholar

Kaati et al., 2002

G. Kaati, L.O. Bygren, S. Edvinsson

Cardiovascular and diabetes mortality determined by nutrition during parents' and grandparents' slow growth period

Eur. J. Hum. Genet., 10 (2002), pp. 682-688

CrossRefView Record in ScopusGoogle Scholar

Kassir, 2020

R. Kassir

Risk of COVID-19 for patients with obesity

Obes. Rev., 21 (2020), p. e13034

View Record in ScopusGoogle Scholar

Kelley et al., 1992

D.E. Kelley, M. Mokan, L.J. Mandarino

Intracellular defects in glucose metabolism in obese patients with NIDDM

Diabetes, 41 (1992), pp. 698-706

CrossRefView Record in ScopusGoogle Scholar

Kenny et al., 2016

G.P. Kenny, R.J. Sigal, R. McGinn

Body temperature regulation in diabetes

Temperature (Austin), 3 (2016), pp. 119-145

CrossRefView Record in ScopusGoogle Scholar

Kettner et al., 2015

N.M. Kettner, S.A. Mayo, J. Hua, C. Lee, D.D. Moore, L. Fu

Circadian dysfunction induces leptin resistance in mice

Cell Metab., 22 (2015), pp. 448-459

ArticleDownload PDFView Record in ScopusGoogle Scholar

Khera et al., 2019

A.V. Khera, M. Chaffin, K.H. Wade, S. Zahid, J. Brancale, R. Xia, M. Distefano, O. Senol-Cosar, M.E. Haas, A. Bick, *et al.*

Polygenic prediction of weight and obesity trajectories from birth to adulthood

Cell, 177 (2019), pp. 587-596.e589

View Record in ScopusGoogle Scholar

Kim and Lazar, 2020

Y.H. Kim, M.A. Lazar

Transcriptional control of circadian rhythms and metabolism: A matter of time and space

Endocr. Rev., 41 (2020), pp. 707-732

CrossRefGoogle Scholar

Kohner et al., 1998

E.M. Kohner, S.J. Aldington, I.M. Stratton, S.E. Manley, R.R. Holman, D.R. Matthews, R.C. Turner

United Kingdom Prospective Diabetes Study, 30: diabetic retinopathy at diagnosis of non-insulin-dependent diabetes mellitus and associated risk factors

Arch. Ophthalmol., 116 (1998), pp. 297-303

CrossRefView Record in ScopusGoogle Scholar

Kondo et al., 2014

T. Kondo, K. Ono, S. Kitano, R. Matsuyama, R. Goto, M.A. Suico, S. Kawasaki, M. Igata, J. Kawashima, H. Motoshima, *et al.*

Mild electrical stimulation with heat shock reduces visceral adiposity and improves metabolic abnormalities in subjects with metabolic syndrome or type 2 diabetes: Randomized Crossover Trials

EBioMedicine, 1 (2014), pp. 80-89

ArticleDownload PDFView Record in ScopusGoogle Scholar

Koopman et al., 2017

A.D.M. Koopman, S.P. Rauh, E. van 't Riet, L. Groeneveld, A.A. van der Heijden, P.J. Elders, J.M. Dekker, G. Nijpels, J.W. Beulens, F. Rutters

The association between social jetlag, the metabolic syndrome, and type 2 diabetes mellitus in the general population: The New Hoorn study

J. Biol. Rhythms, 32 (2017), pp. 359-368

CrossRefView Record in ScopusGoogle Scholar

Krause et al., 2020

M. Krause, F. Gerchman, R. Friedman

Coronavirus infection (SARS-CoV-2) in obesity and diabetes comorbidities: is heat shock response determinant for the disease complications?

Diabetol. Metab. Syndr., 12 (2020), p. 63

View Record in ScopusGoogle Scholar

Lamia et al., 2009

K.A. Lamia, U.M. Sachdeva, L. DiTacchio, E.C. Williams, J.G. Alvarez, D.F. Egan, D.S. Vasquez, H. Juguilon, S. Panda, R.J. Shaw, *et al.*

AMPK regulates the circadian clock by cryptochrome phosphorylation and degradation

Science, 326 (2009), pp. 437-440

CrossRefView Record in ScopusGoogle Scholar

Lancaster and Febbraio, 2014

G.I. Lancaster, M.A. Febbraio

The immunomodulating role of exercise in metabolic disease

Trends Immunol., 35 (2014), pp. 262-269

ArticleDownload PDFView Record in ScopusGoogle Scholar

Larder et al., 2017

R. Larder, M.F.M. Sim, P. Gulati, R. Antrobus, Y.C.L. Tung, D. Rimmington, E. Ayuso, J. Polex-Wolf, B.Y.H. Lam, C. Dias, *et al.*

Obesity-associated gene *TMEM18* has a role in the central control of appetite and body weight regulation

Proc. Natl. Acad. Sci. USA, 114 (2017), pp. 9421-9426

CrossRefView Record in ScopusGoogle Scholar

Lee et al., 2018

Y.S. Lee, J. Wollam, J.M. Olefsky

An integrated view of immunometabolism

Cell, 172 (2018), pp. 22-40

ArticleDownload PDFView Record in ScopusGoogle Scholar

Li et al., 2010

Y. Li, Y. He, L. Qi, V.W. Jaddoe, E.J. Feskens, X. Yang, G. Ma, F.B. Hu

Exposure to the Chinese famine in early life and the risk of hyperglycemia and type 2 diabetes in adulthood

Diabetes, 59 (2010), pp. 2400-2406

CrossRefView Record in ScopusGoogle Scholar

Lim et al., 2011

E.L. Lim, K.G. Hollingsworth, B.S. Aribisala, M.J. Chen, J.C. Mathers, R. Taylor

Reversal of type 2 diabetes: normalisation of beta cell function in association with decreased pancreas and liver triacylglycerol

Diabetologia, 54 (2011), pp. 2506-2514

CrossRefView Record in ScopusGoogle Scholar

Lin et al., 2009

Y.C. Lin, T.J. Hsiao, P.C. Chen

Persistent rotating shift-work exposure accelerates development of metabolic syndrome among middle-aged female employees: a five-year follow-up

Chronobiol. Int., 26 (2009), pp. 740-755

CrossRefView Record in ScopusGoogle Scholar

Lo et al., 2020

K.B. Lo, F. Gul, P. Ram, A.Y. Kluger, K.M. Tecson, P.A. McCullough, J. Rangaswami

The effects of SGLT2 inhibitors on cardiovascular and renal outcomes in diabetic patients: A Systematic Review and Meta-Analysis

Cardiorenal Med., 10 (2020), pp. 1-10

CrossRefView Record in ScopusGoogle Scholar

Locke et al., 2015

A.E. Locke, B. Kahali, S.I. Berndt, A.E. Justice, T.H. Pers, F.R. Day, C. Powell, S. Vedantam, M.L. Buchkovich, J. Yang, *et al.*, LifeLines Cohort Study, ADIPOGen Consortium, AGEN-BMI Working Group, CARDIOGRAMplusC4D Consortium, CKDGen Consortium, GLGC, ICBP, MAGIC Investigators, MuTHER Consortium, MIGen Consortium, PAGE Consortium, ReproGen Consortium, GENIE Consortium, International Endogene Consortium

Genetic studies of body mass index yield new insights for obesity biology

Nature, 518 (2015), pp. 197-206

CrossRefView Record in ScopusGoogle Scholar

Lowe et al., 2020

D.A. Lowe, N. Wu, L. Rohdin-Bibby, A.H. Moore, N. Kelly, Y.E. Liu, E. Philip, E. Vittinghoff, S.B. Heymsfield, J.E. Olgin, *et al.*

Effects of time-restricted eating on weight loss and other metabolic parameters in women and men with overweight and obesity: The TREAT Randomized Clinical Trial

JAMA Intern. Med., 180 (2020), pp. 1491-1499

CrossRefGoogle Scholar

Lundell et al., 2020

L.S. Lundell, E.B. Parr, B.L. Devlin, L.R. Ingerslev, A. Altıntaş, S. Sato, P. Sassone-Corsi, R. Barrès, J.R. Zierath, J.A. Hawley

Time-restricted feeding alters lipid and amino acid metabolite rhythmicity without perturbing clock gene expression

Nat. Commun., 11 (2020), p. 4643

View Record in ScopusGoogle Scholar

Mahajan et al., 2018

A. Mahajan, D. Taliun, M. Thurner, N.R. Robertson, J.M. Torres, N.W. Rayner, A.J. Payne, V. Steinthorsdottir, R.A. Scott, N. Grarup, *et al.*

Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps

Nat. Genet., 50 (2018), pp. 1505-1513

CrossRefView Record in ScopusGoogle Scholar

Meade et al., 2020

R.D. Meade, A.P. Akerman, S.R. Notley, R. McGinn, P. Poirier, P. Gosselin, G.P. Kenny

Physiological factors characterizing heat-vulnerable older adults: A narrative review

Environ. Int., 144 (2020), p. 105909

ArticleDownload PDFView Record in ScopusGoogle Scholar

Montague et al., 1997

C.T. Montague, I.S. Farooqi, J.P. Whitehead, M.A. Soos, H. Rau, N.J. Wareham, C.P. Sewter, J.E. Digby, S.N. Mohammed, J.A. Hurst, *et al.*

Congenital leptin deficiency is associated with severe early-onset obesity in humans

Nature, 387 (1997), pp. 903-908

View Record in ScopusGoogle Scholar

Morley et al., 2020

J.E. Morley, K. Kalantar-Zadeh, S.D. Anker

COVID-19: a major cause of cachexia and sarcopenia?

J. Cachexia Sarcopenia Muscle, 11 (2020), pp. 863-865

CrossRefView Record in ScopusGoogle Scholar

Multhaup et al., 2015

M.L. Multhaup, M.M. Seldin, A.E. Jaffe, X. Lei, H. Kirchner, P. Mondal, Y. Li, V. Rodriguez, A. Drong, M. Hussain, *et al.*

Mouse-human experimental epigenetic analysis unmasks dietary targets and genetic liability for diabetic phenotypes

Cell Metab., 21 (2015), pp. 138-149

ArticleDownload PDFView Record in ScopusGoogle Scholar

Murashov et al., 2016

A.K. Murashov, E.S. Pak, M. Koury, A. Ajmera, M. Jeyakumar, M. Parker, O. Williams, J. Ding, D. Walters, P.D. Neufer

Paternal long-term exercise programs offspring for low energy expenditure and increased risk for obesity in mice

FASEB J., 30 (2016), pp. 775-784

CrossRefView Record in ScopusGoogle Scholar

Nanayakkara et al., 2020

N. Nanayakkara, A.J. Curtis, S. Heritier, A. Gadowski, M.E. Pavkov, T. Kenealy, D.R. Owens, R.L. Thomas, S. Song, J. Wong, *et al.*

Impact of age at type 2 diabetes mellitus diagnosis on mortality and vascular complications: systematic review and meta-analyses

Diabetologia, 64 (2020), pp. 275-287, 10.1007/s00125-020-05319-w

Published online December 14, 2020

Google Scholar

Newman and Maddocks, 2017

A.C. Newman, O.D.K. Maddocks

One-carbon metabolism in cancer

Br. J. Cancer, 116 (2017), pp. 1499-1504

CrossRefView Record in ScopusGoogle Scholar

Nitert et al., 2012

M.D. Nitert, T. Dayeh, P. Volkov, T. Elgzyri, E. Hall, E. Nilsson, B.T. Yang, S. Lang, H. Parikh, Y. Wessman, *et al.*

Impact of an exercise intervention on DNA methylation in skeletal muscle from first-degree relatives of patients with type 2 diabetes

Diabetes, 61 (2012), pp. 3322-3332

CrossRefView Record in ScopusGoogle Scholar

Pålsson-McDermott and O'Neill, 2020

E.M. Pålsson-McDermott, L.A.J. O'Neill

Targeting immunometabolism as an anti-inflammatory strategy

Cell Res., 30 (2020), pp. 300-314

CrossRefView Record in ScopusGoogle Scholar

Pan et al., 2011

A. Pan, E.S. Schernhammer, Q. Sun, F.B. Hu

Rotating night shift work and risk of type 2 diabetes: two prospective cohort studies in women

PLoS Med., 8 (2011), p. e1001141

CrossRefView Record in ScopusGoogle Scholar

Pearson et al., 2006

E.R. Pearson, I. Flechtner, P.R. Njølstad, M.T. Malecki, S.E. Flanagan, B. Larkin, F.M. Ashcroft, I. Klimes, E. Codner, V. Iotova, *et al.*, Neonatal Diabetes International Collaborative Group

Switching from insulin to oral sulfonylureas in patients with diabetes due to Kir6.2 mutations

N. Engl. J. Med., 355 (2006), pp. 467-477

CrossRefView Record in ScopusGoogle Scholar

Peek et al., 2017

C.B. Peek, D.C. Levine, J. Cedernaes, A. Taguchi, Y. Kobayashi, S.J. Tsai, N.A. Bonar, M.R. McNulty, K.M. Ramsey, J. Bass

Circadian clock interaction with HIF1alpha mediates oxygenic metabolism and anaerobic glycolysis in skeletal muscle

Cell Metab., 25 (2017), pp. 86-92

ArticleDownload PDFView Record in ScopusGoogle Scholar

Périard et al., 2015

J.D. Périard, S. Racinais, M.N. Sawka

Adaptations and mechanisms of human heat acclimation: Applications for competitive athletes and sports

Scand. J. Med. Sci. Sports, 25 (Suppl 1) (2015), pp. 20-38

CrossRefView Record in ScopusGoogle Scholar

Pillon et al., 2016

N.J. Pillon, K.L. Chan, S. Zhang, M. Mejdani, M.R. Jacobson, A. Ducos, P.J. Bilan, W. Niu, A. Klip

Saturated fatty acids activate caspase-4/5 in human monocytes, triggering IL-1β and IL-18 release

Am. J. Physiol. Endocrinol. Metab., 311 (2016), pp. E825-E835

CrossRefView Record in ScopusGoogle Scholar

Ranheim et al., 1997

T. Ranheim, C. Dumke, K.L. Schueler, G.D. Cartee, A.D. Attie

Interaction between BTBR and C57BL/6J genomes produces an insulin resistance syndrome in (BTBR x C57BL/6J) F1 mice

Arterioscler. Thromb. Vasc. Biol., 17 (1997), pp. 3286-3293

View Record in ScopusGoogle Scholar

Rathjen et al., 2017

T. Rathjen, X. Yan, N.L. Kononenko, M.C. Ku, K. Song, L. Ferrarese, V. Tarallo, D. Puchkov, G. Kochlamazashvili, S. Brachs, *et al.*

Regulation of body weight and energy homeostasis by neuronal cell adhesion molecule 1

Nat. Neurosci., 20 (2017), pp. 1096-1103

CrossRefView Record in ScopusGoogle Scholar

Ravelli et al., 1998

A.C. Ravelli, J.H. van der Meulen, R.P. Michels, C. Osmond, D.J. Barker, C.N. Hales, O.P. Bleker

Glucose tolerance in adults after prenatal exposure to famine

Lancet, 351 (1998), pp. 173-177

ArticleDownload PDFView Record in ScopusGoogle Scholar

Reidy et al., 2020

P.T. Reidy, Z.S. Mahmassani, A.I. McKenzie, J.J. Petrocelli, S.A. Summers, M.J. Drummond

Influence of exercise training on skeletal muscle insulin resistance in aging: Spotlight on muscle ceramides

Int. J. Mol. Sci., 21 (2020), p. 1514

CrossRefView Record in ScopusGoogle Scholar

GBD 2019 Risk Factors Collaborators, 2020

GBD 2019 Risk Factors Collaborators

Global burden of 87 risk factors in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019

Lancet, 396 (2020), pp. 1223-1249

Google Scholar

Rubino et al., 2020

F. Rubino, S.A. Amiel, P. Zimmet, G. Alberti, S. Bornstein, R.H. Eckel, G. Mingrone, B. Boehm, M.E. Cooper, Z. Chai, *et al.*

New-onset diabetes in Covid-19

N. Engl. J. Med., 383 (2020), pp. 789-790

CrossRefGoogle Scholar

Rudic et al., 2004

R.D. Rudic, P. McNamara, A.M. Curtis, R.C. Boston, S. Panda, J.B. Hogenesch, G.A. Fitzgerald

BMAL1 and CLOCK, two essential components of the circadian clock, are involved in glucose homeostasis

PLoS Biol., 2 (2004), p. e377

View Record in ScopusGoogle Scholar

Ryder et al., 2000

J.W. Ryder, J. Yang, D. Galuska, J. Rincón, M. Björnholm, A. Krook, S. Lund, O. Pedersen, H. Wallberg-Henriksson, J.R. Zierath, G.D. Holman

Use of a novel impermeable biotinylated photolabeling reagent to assess insulin- and hypoxia-stimulated cell surface GLUT4 content in skeletal muscle from type 2 diabetic patients

Diabetes, 49 (2000), pp. 647-654

CrossRefView Record in ScopusGoogle Scholar

Rynders et al., 2018

C.A. Rynders, S. Blanc, N. DeJong, D.H. Bessesen, A. Bergouignan

Sedentary behaviour is a key determinant of metabolic inflexibility

J. Physiol., 596 (2018), pp. 1319-1330

View Record in ScopusGoogle Scholar

Sánchez-Sánchez et al., 2020

E. Sánchez-Sánchez, G. Ramírez-Vargas, Y. Avellaneda-López, J.I. Orellana-Pecino, E. García-Marín, J. Díaz-Jimenez

Eating habits and physical activity of the Spanish population during the COVID-19 pandemic period

Nutrients, 12 (2020), p. 2826

CrossRefGoogle Scholar

Sato et al., 2019

S. Sato, A.L. Basse, M. Schonke, S. Chen, M. Samad, A. Altintas, R.C. Laker, E. Dalbram, R. Barres, P. Baldi, *et al.*

Time of exercise specifies the impact on muscle metabolic pathways and systemic energy homeostasis.

Cell Metab., 30 (2019), pp. 92-110.e114

View Record in ScopusGoogle Scholar

Savikj and Zierath, 2020

M. Savikj, J.R. Zierath

Train like an athlete: applying exercise interventions to manage type 2 diabetes

Diabetologia, 63 (2020), pp. 1491-1499

CrossRefView Record in ScopusGoogle Scholar

Savikj et al., 2019

M. Savikj, B.M. Gabriel, P.S. Alm, J. Smith, K. Caidahl, M. Björnholm, T. Fritz, A. Krook, J.R. Zierath, H. Wallberg-Henriksson

Afternoon exercise is more efficacious than morning exercise at improving blood glucose levels in individuals with type 2 diabetes: a randomised crossover trial

Diabetologia, 62 (2019), pp. 233-237

CrossRefView Record in ScopusGoogle Scholar

Song et al., 2020

J.D. Song, T.C. Alves, D.E. Befroy, R.J. Perry, G.F. Mason, X.M. Zhang, A. Munk, Y. Zhang, D. Zhang, G.W. Cline, *et al.*

Dissociation of muscle insulin resistance from alterations in mitochondrial substrate preference

Cell Metab., 32 (2020), pp. 726-735.e725

View Record in ScopusGoogle Scholar

Stanford et al., 2015

K.I. Stanford, M.Y. Lee, K.M. Getchell, K. So, M.F. Hirshman, L.J. Goodyear

Exercise before and during pregnancy prevents the deleterious effects of maternal high-fat feeding on metabolic health of male offspring

Diabetes, 64 (2015), pp. 427-433

CrossRefView Record in ScopusGoogle Scholar

Taylor and Holman, 2015

R. Taylor, R.R. Holman

Normal weight individuals who develop type 2 diabetes: the personal fat threshold

Clin. Sci. (Lond.), 128 (2015), pp. 405-410

View Record in ScopusGoogle Scholar

Tiffon, 2018

C. Tiffon

The Impact of nutrition and environmental epigenetics on human health and disease

Int. J. Mol. Sci., 19 (2018), p. 3425

CrossRefView Record in ScopusGoogle Scholar

Timpel et al., 2019

P. Timpel, L. Harst, D. Reifegerste, S. Weihrauch-Blüher, P.E.H. Schwarz

What should governments be doing to prevent diabetes throughout the life course?

Diabetologia, 62 (2019), pp. 1842-1853

CrossRefView Record in ScopusGoogle Scholar

Turek et al., 2005

F.W. Turek, C. Joshu, A. Kohsaka, E. Lin, G. Ivanova, E. McDearmon, A. Laposky, S. Losee-Olson, A. Easton, D.R. Jensen, *et al.*

Obesity and metabolic syndrome in circadian Clock mutant mice

Science, 308 (2005), pp. 1043-1045

CrossRefView Record in ScopusGoogle Scholar

Udler et al., 2018

M.S. Udler, J. Kim, M. von Grotthuss, S. Bonàs-Guarch, J.B. Cole, J. Chiou, M. Boehnke, M. Laakso, G. Atzmon, B. Glaser, *et al.*, Christopher D. Anderson on behalf of METASTROKE and the ISGC

Type 2 diabetes genetic loci informed by multi-trait associations point to disease mechanisms and subtypes: A soft clustering analysis

PLoS Med., 15 (2018), p. e1002654

CrossRefView Record in ScopusGoogle Scholar

Udler et al., 2019

M.S. Udler, M.I. McCarthy, J.C. Florez, A. Mahajan

Genetic risk scores for diabetes diagnosis and precision medicine

Endocr. Rev., 40 (2019), pp. 1500-1520

CrossRefView Record in ScopusGoogle Scholar

van der Klaauw and Farooqi, 2015

A.A. van der Klaauw, I.S. Farooqi

The hunger genes: pathways to obesity

Cell, 161 (2015), pp. 119-132

ArticleDownload PDFView Record in ScopusGoogle Scholar

Vimalananda et al., 2015

V.G. Vimalananda, J.R. Palmer, H. Gerlovin, L.A. Wise, J.L. Rosenzweig, L. Rosenberg, E.A. Ruiz Narváez

Night-shift work and incident diabetes among African-American women

Diabetologia, 58 (2015), pp. 699-706

CrossRefView Record in ScopusGoogle Scholar

Vroman et al., 1983

N.B. Vroman, E.R. Buskirk, J.L. Hodgson

Cardiac output and skin blood flow in lean and obese individuals during exercise in the heat

J. Appl. Physiol., 55 (1983), pp. 69-74

CrossRefView Record in ScopusGoogle Scholar

Wang et al., 2020a

S. Wang, P. Ma, S. Zhang, S. Song, Z. Wang, Y. Ma, J. Xu, F. Wu, L. Duan, Z. Yin, *et al.*

Fasting blood glucose at admission is an independent predictor for 28-day mortality in patients with COVID-19 without previous diagnosis of diabetes: a multi-centre retrospective study

Diabetologia, 63 (2020), pp. 2102-2111

CrossRefView Record in ScopusGoogle Scholar

Wang et al., 2020b

X. Wang, Y. Wang, V. Antony, H. Sun, G. Liang

Metabolism-associated molecular patterns (MAMPs)

Trends Endocrinol. Metab., 31 (2020), pp. 712-724

ArticleDownload PDFCrossRefView Record in ScopusGoogle Scholar

West et al., 2016

M.A. West, R. Asher, M. Browning, G. Minto, M. Swart, K. Richardson, L. McGarrity, S. Jack, M.P. Grocott, Perioperative Exercise Testing and Training Society

Validation of preoperative cardiopulmonary exercise testing-derived variables to predict in-hospital morbidity after major colorectal surgery

Br. J. Surg., 103 (2016), pp. 744-752

CrossRefView Record in ScopusGoogle Scholar

Wilkinson et al., 2020

M.J. Wilkinson, E.N.C. Manoogian, A. Zadourian, H. Lo, S. Fakhouri, A. Shoghi, X. Wang, J.G. Fleischer, S. Navlakha, S. Panda, *et al.*

Ten-hour time-restricted eating reduces weight, blood pressure, and atherogenic lipids in patients with metabolic syndrome

Cell Metab., 31 (2020), pp. 92-104.e105

View Record in ScopusGoogle Scholar

Willemsen et al., 2015

G. Willemsen, K.J. Ward, C.G. Bell, K. Christensen, J. Bowden, C. Dalgård, J.R. Harris, J. Kaprio, R. Lyle, P.K. Magnusson, *et al.*

The concordance and heritability of type 2 diabetes in 34,166 twin pairs from international twin registers: The Discordant Twin (Dros. Inf. Serv.COTWIN) Consortium

Twin Res. Hum. Genet., 18 (2015), pp. 762-771

CrossRefView Record in ScopusGoogle Scholar

Xin et al., 2020

Y. Xin, A. Davies, A. Briggs, L. McCombie, C.M. Messow, E. Grieve, W.S. Leslie, R. Taylor, M.E.J. Lean

Type 2 diabetes remission: 2 year within-trial and lifetime-horizon cost-effectiveness of the Diabetes Remission Clinical Trial (DiRECT)/Counterweight-Plus weight management programme

Diabetologia, 63 (2020), pp. 2112-2122

CrossRefView Record in ScopusGoogle Scholar

Yengo et al., 2018

L. Yengo, J. Sidorenko, K.E. Kemper, Z. Zheng, A.R. Wood, M.N. Weedon, T.M. Frayling, J. Hirschhorn, J. Yang, P.M. Visscher, GIANT Consortium

Meta-analysis of genome-wide association studies for height and body mass index in :700000 individuals of European ancestry

Hum. Mol. Genet., 27 (2018), pp. 3641-3649

CrossRefView Record in ScopusGoogle Scholar

Young, 2018

M.W. Young

Time travels: A 40-Year journey from drosophila's clock mutants to human circadian disorders (Nobel Lecture)

Angew. Chem. Int. Ed. Engl., 57 (2018), pp. 11532-11539

CrossRefView Record in ScopusGoogle Scholar

Yuan and Larsson, 2020

S. Yuan, S.C. Larsson

An atlas on risk factors for type 2 diabetes: a wide-angled Mendelian randomisation study

Diabetologia, 63 (2020), pp. 2359-2371
CrossRefView Record in ScopusGoogle Scholar

29.7: LIPID-INDUCED MECHANISMS OF METABOLIC SYNDROME

Yulia K. Denisenko, Oxana Yu Kytikova,Tatyana P. Novgorodtseva, Marina V. Antonyuk, Tatyana A. Gvozdenko,and Tatyana A. Kantur.
VJournal of Obesity, Volume 2020 |Article ID 5762395 | https://doi.org/10.1155/2020/5762395

29.7.1: COPYRIGHT

29.7.2: ABSTRACT

Metabolic syndrome (MetS) has a worldwide tendency to increase and depends on many components, which explains the complexity of diagnosis, approaches to the prevention, and treatment of this pathology. Insulin resistance (IR) is the crucial cause of the MetS pathogenesis, which develops against the background of abdominal obesity. In light of recent evidence, it has been shown that lipids, especially fatty acids (FAs), are important signaling molecules that regulate the signaling pathways of insulin and inflammatory mediators. On the one hand, the lack of n-3 polyunsaturated fatty acids (PUFAs) in the body leads to impaired molecular mechanisms of glucose transport, the formation of unresolved inflammation. On the other hand, excessive formation of free fatty acids (FFAs) underlies the development of oxidative stress and mitochondrial dysfunction in MetS. Understanding the molecular mechanisms of the participation of FAs and their metabolites in the pathogenesis of MetS will contribute to the development of new diagnostic methods and targeted therapy for this disease. The purpose of this review is to highlight recent advances in the study of the effect of fatty acids as modulators of insulin response and inflammatory process in the pathogenesis and treatment for MetS.

29.7.3: 1. INTRODUCTION

Metabolic syndrome (MetS) is a complex of several disorders (abdominal obesity, hyperglycemia, hypertriglyceridemia, and hypertension), which together dramatically raise the risk of developing atherosclerotic cardiovascular disease, insulin resistance, and diabetes mellitus [1, 2]. Because the prevalence of obesity has doubly increased worldwide over the past 30 years, the prevalence of MetS has markedly boosted in tandem [2–5]. Currently, clinicians and researchers have not identified an optimal treatment for MetS, and consequently, it is critical to identify new ways of approaching this syndrome in order to identify efficacious methods of diagnosing, screening, and treating MetS. Most researchers believe that hyperinsulinemia and/or insulin resistance (IR) is the first link in the chain of clinical-metabolic disturbances of MetS [5–7]. The development of IR is the result of a long chain of pathological events. Lipids play the crucial role in the pathogenesis of IR and the subsequent development of MetS [8–18]. All lipids are no longer considered the same. It is well known that excessive consumption of saturated fats contributes to the development of obesity and related diseases [19]. It has now been shown that high plasma levels of free fatty acids (FFAs), particularly saturated fatty acids (SFAs), may be associated with insulin resistance in obese patients with type 2 diabetes mellitus [17]. The lack of polyunsaturated fatty acids (PUFAs), especially n-3 PUFAs, some phospholipids, and plasmalogens in the cell membrane, is the cause of changes in glucose-insulin homeostasis and the development of inflammation [10, 13, 20–23]. Conversely, multiple investigations have established a connection between inflammation and changes in lipids and their derivatives in the setting of MetS [24–29]. Alteration in the metabolism of fatty acids affects the synthesis of eicosanoids and pro-resolving lipid mediators responsible for immune-metabolic homeostasis [30–33]. Recent studies have further elucidated the role of these metabolites in the contribution to the chronic, low-grade inflammatory state in MetS [34–36]. A comprehensive understanding of the importance of lipids in the pathogenesis of MetS contributes to the development of preventive and targeted lipid-correcting therapy. The aim of the review is to analyse the modern views on the role of lipids, particularly PUFAs and FFAs, in the pathogenesis of MetS. In this review, we summarized the molecular mechanisms of the relationship between fatty acids and glucose transport, inflammatory response, mitochondrial dysfunction, and endoplasmic reticulum stress in the development of MetS.

29.7.4: 2. METABOLIC SYNDROME: DEFINITIONS AND CRITERIA

Metabolic syndrome (MetS) has become a widely debated scientific, medical, and social problem worldwide. Indeed, the definition of metabolic syndrome is important for clinical practice and deserves serious scientific and medical research. MetS is characterized by the following clinical criteria: abdominal obesity, decreased peripheral tissue sensitivity to insulin, and hyperinsulinemia, which cause metabolic disorders of carbohydrates, lipids, and purines [1, 2]. This combination of metabolic disorders is often found in one person and, thus, significantly increases the risk of cardiovascular disease (CVD), type 2 diabetes mellitus (T2DM), arthritis, chronic kidney disease, schizophrenia, nonalcoholic fatty liver disease (NAFLD), and several types of cancer [37–42].

MetS is characterized by a steadily increasing prevalence [3, 4]. However, its prevalence rates vary depending on the criteria used to determine MetS, genetic component, gender, age, population and area of residence, education, level of physical activity, nutrition, and lifestyle [39]. Approximately one-fourth of world's adult population have MetS [43, 44]. Urbanization and its associated sedentary lifestyle and surplus nutrition are the root cause of this global epidemic.

The determination of MetS uses the criteria of the following medical communities: WHO (World Health Organization), NCEPATP III (National Cholesterol Education Program-Adult Treatment Panel III), AACE (American Association of Clinical Endocrinologists), IDF (International Diabetes Federation), EGIR (European Group for the Study of Insulin Resistance), The International Diabetes Federation (IDF), American Heart Association/National Heart, Lung and Blood Institute (AHA/NLHBI), World Heart Federation (WHF), International Atherosclerosis Society (IAS), and The International Association for the Study of Obesity (IASO) [3]. A guideline was made in 2009 to unify the criteria for the diagnosis of MetS. According to this guide, three of the five criteria are necessary for diagnosing MetS: (1) waist circumference ≥102 cm for males and ≥88 cm for females (for Asians ≥90 cm for males and ≥80 cm for females); (2) systolic blood pressure ≥130 mmHg or diastolic blood pressure ≥85 mmHg or antihypertensive medication; (3) fasting plasma glucose ≥5.6 mmol/L or on medication for high blood glucose; (4) HDL cholesterol <1.03 mmol/L for males and <1.30 mmol/L for females or on medications for reduced HDL cholesterol; and (5) triglycerides ≥1.7 mmol/L or on medications for elevated triglycerides [3].

Although the exact etiology of the MetS is not clearly understood, insulin resistance (IR) is considered as the principal factor for the pathogenesis of this syndrome [6, 11, 18, 45]. As found by the insulin-modified, frequently sampled intravenous glucose tolerance assay, insulin sensitivity is significantly lower in patients with two or more components of the MetS compared to those with none of these components [2]. Dysregulation of lipid metabolism is considered as an important link in the overall development chain of IR. It is well known that lipids play a critical role in the regulation of energy metabolism, glucose transport, and immune process in many organs and tissues such as the liver, adipose tissue, muscle, heart, and gastrointestinal tract [46]. However, the molecular mechanisms of this regulation remain largely unexplored. The study of lipid metabolism disorders is a promising direction for the development of methods for the effective treatment of this pathology.

29.7.5: 3. POLYUNSATURATED FATTY ACIDS AND METABOLIC SYNDROME RISK

Fatty acids (FAs) play multiple roles in humans and other organisms. Most importantly, FAs are a substantial part of lipids. Fatty acids are either saturated or unsaturated carboxylic acids with carbon chains varying between 2 and 36 carbon atoms. Polyunsaturated fatty acids with an acid end containing the functional carboxylic acid group and a methyl end are also known as omega end. In omega-3 (ω-3 or n-3) and omega-6 (ω-6 or n-6) fatty acids, the first site of desaturation is located after the third and the sixth carbon from the omega end, respectively. Our body cannot synthesize some PUFAs, such as alpha-linolenic acid (18:3n3) and linoleic acid (18:2n6). These essential PUFAs enter our bodies only through diet. The dietary sources of n-3 PUFAs include fish oils rich in eicosapentaenoic acid (20:5n3) and docosahexaenoic acid (22:6n3), whereas the n-6 PUFA linoleic acid is mostly found in plants and vegetable oils [47].

Nowadays, there is growing evidence showing that dietary n-3 PUFAs have a variety of healthy properties such as the reduction of plasma atherogenic lipids and inflammation [16, 21, 25, 41, 47–49]. The associations between n-3 PUFAs and metabolic syndrome risk demonstrate inconsistent results [50]. Several cross-sectional and case-control studies have indicated that plasma/serum n-3 PUFAs were significantly higher in healthy subjects compared with those in patients with MetS, while some studies have suggested opposite and null associations [51]. Meanwhile, Guo et al. showed that higher circulating n-3 PUFAs were significantly associated with decreased MetS risk [52]. A study by Kim et al. demonstrated that in healthy individuals the level of long-chain n-3 PUFAs is positively correlated with insulin sensitivity [53]. A decrease in the level of PUFAs has been established in patients with T2DM and diabetic retinopathy. It was found that the development of insulin resistance is preceded by a reduction of essential n-3 PUFAs in the cell membranes [6]. There are several reasons why n-3 PUFAs are important in the pathogenesis and prevention of MetS. PUFAs perform a structural function, being important components of the cell membrane and determining its physical and chemical properties [10, 16]. The efficiency of glucose transport and expression of many receptors depend on the composition and ratio of PUFAs in the cell membrane [17]. Also, PUFAs are the precursors for inflammatory and pro-resolving lipids mediators' synthesis [29, 32]. The imbalance between the synthesis of inflammatory and pro-resolving lipids mediators determines the development of chronic inflammation in MetS. Furthermore, we will summarize the main molecular mechanisms underlying the ability of n-3 PUFAs to prevent and/or ameliorate insulin resistance and inflammation in MetS.

29.7.5.1: 3.1. POLYUNSATURATED FATTY ACIDS AND GLUCOSE TRANSPORT

The identification of a causal relationship between the composition of fatty acids of cell membranes and MetS pathogenesis significantly contributes to an understanding of the main pathophysiological mechanisms of the disease.

Polyunsaturated fatty acids affect the fundamental properties of the cell membrane, including its fluidity, elasticity, receptor expression activity, the functionality of embedded proteins, and signal transmission through lipid rafts, which leads to changes in cell signaling and modification of gene expression [54, 55]. The length and degree of the FA chain unsaturation have a profound effect on the physical and chemical properties of cell membranes [48]. The ratio of polyunsaturated to saturated fatty acids determines the membrane flexibility, which affects the efficiency of glucose transport using insulin-independent glucose transporters (GLUTs) and insulin-dependent GLUT4 [54].

GLUT1 is a monomeric protein with 12 transmembrane helical segments [56]. One molecule GLUT1 covers an area of about 17 molecules of a phosphatidylcholine bilayer with saturated fatty acids (SFAs), which requires a high membrane flexibility for pore formation. GLUT4 is inserted into the membrane of intracellular vesicles, which demands the flexibility of the vesicular membrane. The GLUT4 containing vesicles take part in a fusion process with the cell membrane. The increased flexibility of the membrane provides a smooth bending of the cell membrane bilayer and the fused pores formation [54]. Thus, decreased membrane flexibility causes a reduction in all Class 1 glucose

transporters which, in turn, reduces the glucose flux and increases the plasma glucose concentration. Therefore, high membrane flexibility is a crucial factor in glucose transport. Changes in the fatty acid composition of membranes will result in disturbance in the physicochemical properties of the bilayer, such as flexibility and fluidity. Tighter membrane packaging due to increased saturated fatty acids in it leads to a reduction in the capacity for GLUT4 glucose transport [54].

A number of other studies have also revealed that an increase in the level of saturated fatty acids in the cell membrane is associated with a growth in blood glucose level and the development of insulin resistance [13, 14]. The important role of PUFAs in maintaining glucose-insulin homeostasis is confirmed by many studies [13–15, 21, 25, 47, 49, 51, 53]. Comprehensive evidence shows that diet n-3 PUFAs can improve insulin signal transduction in adipocytes, affecting in turn the insulin-stimulated glucose uptake through the regulation of the expression or the translocation of the GLUT4 [11]. In vitro studies have found that adipocytes from n-3 PUFAs-depleted rats had lower basal and insulin-stimulated glucose incorporation, while cultured adipocytes supplemented with fish oil increased levels of GLUT4 and GLUT1 [41]. González-Périz et al. [11] reported that feeding with a marine n-3 PUFAs-enriched diet improved insulin resistance in association with an increased expression of Irs-1 and Glut4mRNA in the adipose tissue of genetically obese ob/ob mice. The above indicates the huge importance of n-3 PUFAs in the development and regulation of components in the MetS, such as insulin resistance and glucose tolerance.

29.7.5.2: 3.2. POLYUNSATURATED FATTY ACIDS AND INFLAMMATION

Numerous studies have suggested that MetS, like its downstream sequelae of atherosclerotic cardiovascular disease and T2DM, is largely an inflammatory disease [24]. A chronic, low-grade inflammatory state caused by obesity leads to metabolic alterations responsible for multiple organ damage [57, 58]. This metabolic dysfunction could determine clinical conditions such as hypertension, hypercholesterinemia, and insulin resistance [40]. The contribution of inflammation to insulin resistance has been widely studied, and immunological changes occurring in various tissues are thought to be etiological factors affecting the development of insulin resistance [58]. A characteristic of obese people is a chronic, low-grade inflammation state promoted by the release of many inflammatory mediators by the adipose tissue and, more importantly, by infiltrating macrophages. PUFAs and their oxidized metabolites are important participants of the inflammatory processes of MetS [13, 15, 27, 28]. Understanding the molecular mechanisms of the participation of PUFAs and their metabolites in the pathogenesis of MetS will contribute to the development of new diagnostic methods and targeted therapy for this disease.

29.7.5.2.1: 3.2.1. SPECIALIZED PRO-RESOLVING MEDIATORS

Inflammation is a complex, multifactorial adaptive process with different periods of development. Inflammation is a natural reaction to harmful irritants, such as bacterial infections, virus infections, and tissue damage. This is a host's defensive reaction in which immune and endothelial cells and proinflammatory mediators are attracted to eliminate inflammatory agents, clear damaged cells and tissues, and initiate tissue repair. This response, when properly functioning, is self-limiting and leads to the cessation of the inflammatory response and a return to homeostasis, a process called the resolution of inflammation [16, 48]. Resolution of inflammation is now known to be an active process involving the activation of negative feedback mechanisms, such as anti-inflammatory cytokine secretion, reduction in receptor expression, activation of regulatory cells, and production of pro-resolving lipid mediators [57]. However, when acute inflammation is intense or prolonged, the resolution process is not successful, which leads to excessive tissue damage and ultimately resulting in chronic inflammation [16]. Many studies have confirmed that unresolved inflammation is the main mechanism for the pathogenesis of MetS [15, 16, 24, 48, 57, 58].

PUFAs are a source of synthesis of inflammatory and pro-resolving lipid mediators. The major substrate for the synthesis of inflammatory lipid mediators is arachidonic acid (20:4n6) (see Figure 1). The high content of 20:4n6 provides a direct link with inflammation since 20:4n6 released from cell membrane phospholipids acts as a substrate for cyclooxygenase (COX), lipoxygenase (LOX), and cytochrome P450 enzymes [29]. Eicosanoids are important regulators and mediators of acute inflammatory processes and include prostaglandins (PGs), thromboxanes (TBs), and leukotrienes (LTs). Many anti-inflammatory therapies, such as nonsteroidal anti-inflammatory drugs and COX inhibitors, target arachidonic acid metabolism [29, 35, 59, 60].

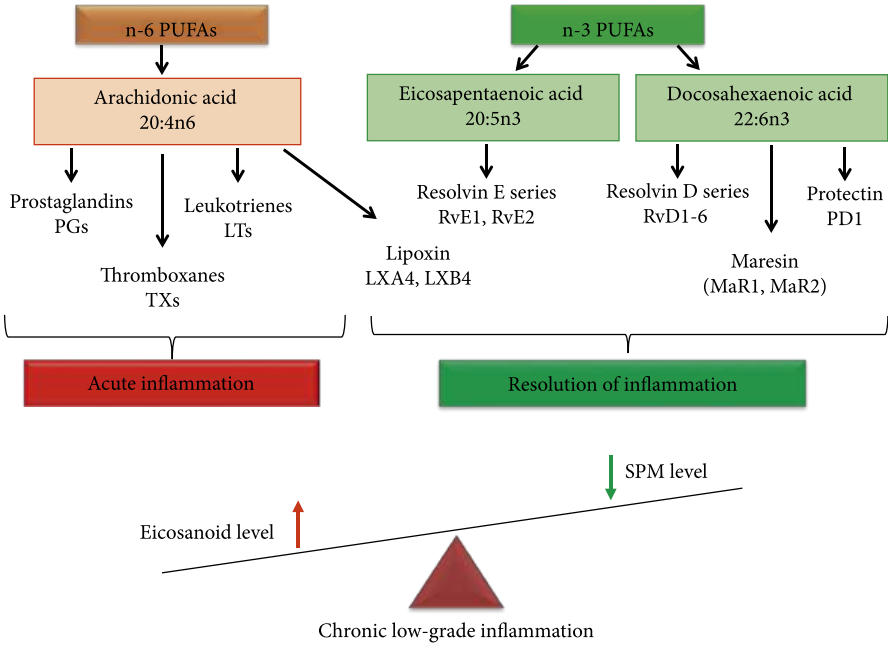


Figure 1

PUFA pathway and role of lipid mediators in the development and resolution of inflammation.

Eicosapentaenoic acid (20:5n3) and docosahexaenoic acid (22:6n3) from the n-3 PUFAs family are a source of synthesizing specialized pro-resolving mediators (SPMs): maresins, resolvins, and protectins (see Figure 1) [32–35]. SPMs are a class of cell compounds generated at a later stage of the inflammation and initiate the resolution of the inflammatory process [33, 34, 61]. SPMs actively facilitate the resolution stage of acute inflammation unlike eicosanoids, which mainly act during the first stage of inflammation. A balanced n-6 : n-3 PUFAs ratio (where 1 : 1 to 2 : 1 is optimal) is important for homeostasis and normal development throughout the lifespan. High n-6 PUFA intake in the Western diet increases the n-6 : n-3 ratio to a range from 10 : 1 to 20 : 1 and may play a role in the pathogenesis of MetS and related diseases [36]. The balance between n-3 PUFAs and n-6 PUFAs determines the path of inflammatory response. The prevalence of n-6 PUFAs and the shortage of n-3 PUFAs may contribute to impaired inflammation resolution [31].

Docosahexaenoic acid- and eicosapentaenoic acid-derived SPMs are identified in the adipose tissue. At the same time, the levels of certain SPMs are markedly reduced with obesity, suggesting adipose SPM deficiency, potentially resulting in unresolved inflammation [36].

Resolvins are synthesized spontaneously from eicosapentaenoic and docosahexaenoic acids during inflammation and thus are designated as E-series (RvE) and D-series (RvD), respectively [32]. The anti-inflammatory effect of RvE1 is due to interaction with peroxisome proliferator-activated receptors (PPARs), which are classified as nuclear transcription factors with anti-inflammatory activity. Leukotriene B4 receptor 1 (BLT1) and G protein-coupled receptor, Chemerin Receptor 23 (ChemR23), are receptors for RvE1. RvE2 has a similar biologic effect; it regulates neutrophil chemotaxis and activates phagocytosis and proinflammatory cytokines synthesis [31, 62].

Protectins (PD) are another class of pro-resolving molecules produced from 22:6n3 during the resolution of inflammation. Protectins are synthesized by a number of cells including brain cells, monocytes, and CD4+ lymphocytes [33]. PD1, the key representative of the protectin family, demonstrates a strong anti-inflammatory and neuroprotective effect. This mediator functioning is based on PPARs interacting and NF-κB blocking [62].

An alternative process for docosahexaenoic acid oxygenation is found in human macrophages and platelets, leading to the synthesis of maresin 1 (MaR1). In addition, 13S, 14S-epoxy-maresin, which has important biological activities of its own, is the precursor for maresin 2 (MaR2) [36].

Lipoxins (LXs) are powerful anti-inflammatory bioregulators suppressing inflammation and activating resolution and recovery processes, in particular in MetS [34]. The substrate for LXs synthesis is arachidonic acids. Two members of the LXs family, LXA4 and LXB4, have been well studied [31]. In general, LXs are a branch of the leukotriene family. For example, their production by platelets is catalyzed by 12-LOX through converting LTA4 [32]. Unlike proinflammatory LTs, LXs act as powerful anti-inflammatory bioregulators, suppressing the inflammation and activating the processes of resolution and recovery. The result of their action is the inhibition of chemotaxis and migration of macrophages and neutrophils to the inflammatory focus, blocking of the lipid peroxidation, the activation of NF-κB, and the suppression of the synthesis of proinflammatory cytokines. In addition, LXs are actively involved in functioning of macrophages that are associated with homeostasis restoration processes [32].

There is a considerable amount of evidence regarding the contribution of n-3 PUFAs to diseases with inflammatory conditions, such as metabolic syndrome [25, 34–36, 47, 49, 51, 53]. It was reported that SPM levels reduced in metabolic syndrome as well as sensitivity to

SPM of the adipose tissue [36]. Obesity reduces the levels of PD1, intermediates in the synthesis of D-series resolvins and protectins (17-HDHA), and intermediates in the maresin biosynthesis (14-HDHA) for the adipose tissues from diet- and genetically-induced obese mice [36]. One of the mechanisms resulting in a decrease in the SPM level in obesity is a change in the enzyme activities involved in biosynthesis or conversion of certain SPMs. N-3 PUFAs supplementation increased the level of SPM in the blood of individuals with obesity and MetS. The effects of n-3 PUFAs are mediated by their ability to interfere with arachidonic acid metabolism and promote the synthesis of SPMs. The supply of n-3 PUFAs increases the levels of resolvins, enhances resolution, and improves insulin sensitivity in an experiment with fat-1 mice. In addition, n-3 PUFAs prevent macrophage increase, adipokine secretion, and insulin resistance induced by a high-fat diet. Synthetic pro-resolving lipid mediators (17-hydroxy-DHA) or n-3 PUFAs added to the treatment contributed to higher levels of pro-resolving lipid mediators in the adipose tissue, reduced inflammation, and increased insulin sensitivity [31]. N-3 PUFAs increased RvE-series levels in patients with MetS but did not affect RvD-series, which requires further studies into the mechanism of n-3 PUFAs influence in MetS. For instance, intraperitoneal administration of 17-HDHA or RvD1 significantly reduced adipose inflammation and improved the glucose tolerance in diet-induced obese mice and in db/db mice [63]. Treatment with either RvD1 or RvD2 also reduced the secretion of proinflammatory cytokines including TNF-α, IL-1β, and IL-12 in the adipose tissue [64]. The MaR1 treatment improved insulin sensitivity, determined with an insulin tolerance test. MaR1 also increased adiponectin gene expression and Akt phosphorylation in the adipose tissues and attenuated adipose tissue inflammation in both ob/ob and diet-induced obese mice [36]. PD1 treatment acutely increased the adiponectin transcripts in adipose tissue explants isolated from ob/ob mice. A potent ability to induce adiponectin expression/secretion has been demonstrated with synthetical RvD1, RvD2, and PD1 and their biosynthetic intermediate, 17-HDHA [63].

Therefore, one of the pathogenetic mechanisms of the development of MetS is a reduction of the processes of resolving inflammation and the development of chronic, low-grade inflammatory. A decrease in the synthesis of specialized pro-resolving lipid mediators is the basis of the above disorders [65]. Thus, the anti-inflammatory effect of n-3 PUFAs in MetS can be mediated through the regulation of the SPM synthesis.

29.7.5.2.2: 3.2.2. TOLL-LIKE RECEPTOR 4

The inflammatory process observed in individuals with metabolic syndrome differs from the classical inflammatory response and this type of inflammation characterized by a chronic, low-intensity reaction [58]. The toll-like receptor 4 (TLR4) signaling pathway is acknowledged as one of the main triggers of the obesity-induced inflammatory response [57]. TLR4 plays a significant role in the pathogenesis of inflammation mediated by insulin resistance in MetS [57]. Toll-like receptors, including TLR4, are type 1 transmembrane proteins with three domains: (1) extracellular domain with leucine-rich repeats (LRRs) responsible for ligand recognition; (2) transmembrane domain; and (3) intracellular toll/interleukin-1 receptor (TIR) domain. These provide signal transmission from the cell surface to adapter proteins. TLR4 was the first TLR reported in humans; it is expressed in innate immune cells, including monocytes, macrophages, and dendritic cells, as well as in other cell types, such as adipocytes, enterocytes, and muscle cells. TLR4 is a membrane-associated receptor involved in lipid recognition [66]. TLRs are activated both by the influence of endogenous ligands and by the participation of lipids—cholesterol, SFAs, and oxidized forms of phospholipids [67].

Humans with type I diabetes exhibit greater expression of TLR4 in the cellular membrane in monocytes. Individuals with T2DM show increased cellular membrane levels of TLR4 in blood monocytes, as well as a higher concentration of IL-1, IL-6, IL-8, and TNF in serum. Similarly, TLR4 is more highly expressed in blood mononuclear cells and in the abdominal subcutaneous white adipose tissue of obese and diabetic individuals [57, 68].

Lipids from foods change the expression of TLRs by cells [69]. On the one hand, SFAs activate the TLR4 signaling pathway (see Figure 2). Among the SFAs, lauric acid (12 : 0) and palmitic acid (16 : 0) had the strongest activation capacity through TLR4 [69]. On the other hand, TLRs can be inhibited by PUFAs [70]. Consumption of n-3 PUFAs, particularly 22:6n3, is associated with anti-inflammatory and cardioprotective effects. It is believed that the use of n-3 PUFAs is associated with anti-inflammatory activity due to inhibition of arachidonic acid metabolism [71]. The molecular effect of n-3 PUFAs, especially 20:5n3 and 22:6n3, on inflammatory-response modulation are based on the ability of these PUFAs to inhibit the expression of inflammatory genes, such as COX-2, iNOS, and IL-1 in macrophages [72]. PUFAs of the n-3 family reduce the activation of the NF-κB transcription factor pathway that is induced by various agonists [70].

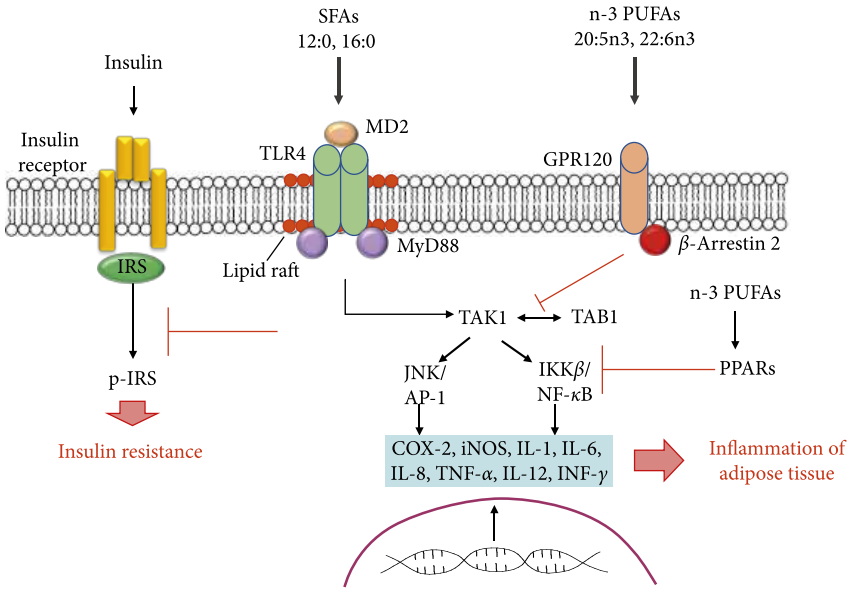


Figure 2

The role of TLRs and FAs in the signaling mechanisms of inflammation in the adipose tissue and insulin resistance. The saturated fatty acids (SFAs) act as nonmicrobial TLR4 agonists or indirectly promote the TLR4 activation, triggering its inflammatory response and inflammation of the adipose tissue. Inflammatory signaling caused by saturated fatty acids via TLR4/MD-2 inhibits the phosphorylation of the insulin receptor, leading to the development of insulin resistance. GPR120 activation induced by n-3 PUFA leads to a decrease in the activity of IKK-β/NF-κB and JNK/AP-1 signaling pathways, which reduces the expression of proinflammatory genes. The anti-inflammatory properties of PPARs are achieved by inhibiting nuclear factor-kappa B (NF-κB). N-3 PUFAs directly interact with PPARs and modulate the expression of proinflammatory genes.

Other mechanisms modulate the inflammatory response by fatty acids based on binding G protein-coupled receptor 120 (GPR120) [66]. GPR120 is a free fatty acid 4 receptor (FFAR4), and GPR120 activation induced by n-3 PUFA leads to β-arrestin 2 recruitment to the plasma membrane where this protein binds to GPR120 (see Figure 2) The GPR120/β-arrestin 2 complex is internalized into the cytoplasmic compartment where this complex binds to the TAK1-binding protein (TAB1). This process impairs the association between TAB1 and the kinase activated by the growth factor beta (TAK1) and, consequently, results in reduced TAK1 activation and decreases the activity of the IKK-β/NF-κB and JNK/AP-1 signaling pathways. The mitigation of TAK-1 activation by n-3 PUFAs leads to the reduced expression of TNF-α and IL-6 genes with proinflammatory actions [17, 57].

One more important molecular mechanism that is associated with the n-3 PUFA effects concerns their capacities to bind to PPARs [62]. Three isoforms of PPARs are known: PPARα (NR1C1), PPARβ/δ (NR1C2), and PPARγ (NR1C3). PPARs are involved in the regulation of inflammatory reactions and lipid metabolism. The anti-inflammatory properties of PPARs are mainly achieved by inhibiting nuclear factor-kappa B (NF-κB) which, in turn, is the proinflammatory nuclear transcription factor [73]. The interactions between PPARs, NF-κB, and toll-like receptors (TLRs) are of great interest. Along with the anti-inflammatory mechanism of action of PPARs, the proinflammatory activity of some isoforms of PPARs is also being studied. For example, PPARγ is considered a mediator of interactions between dendritic and T cells in the development of type 2 (or T2) inflammation [73]. N-3 PUFAs directly interact with PPARs and, therefore, modulate the expression of genes that are involved in lipid metabolism and the inflammatory response [57]. Anti-inflammatory effects of 20:5n3 and 22:6n3 on this signaling pathway can occur due to diminished nicotinamide adenine dinucleotide phosphate (NADPH) oxidase activity, which leads to lower TLR4 recruitment for lipid rafts and TLR4 dimerization [16]. Also, another possible mechanism of action of the n-3 PUFA concerns the capacity of incorporating 22:6n3 into the plasma membrane, which can lead to reduced TLR4 translocation for lipid rafts formation [74, 75]. The variety of molecular mechanisms in lipids and TLR4 signaling pathway interaction indicates the complexity of the pathogenesis of MetS and associated diseases.

29.7.5.3: 3.3. POLYUNSATURATED FATTY ACIDS AND PLASMALOGENS

Permanent exogenous use of PUFA is a necessary condition for maintaining immune-metabolic homeostasis. The profile of fatty acids that are present in the Western diet consists of a high level of saturated fatty acids and trans fatty acids. While the total consumption of marine and plant n-3 polyunsaturated fatty acids in contemporary society is significantly reduced [7, 76].

Another reason for PUFAs reduction is deterioration in the plasmalogen synthesis [77]. Plasmalogens are a subclass of phospholipids characterized by having a vinyl ether bond linking the fatty aldehyde to the glycerol molecule in the 1-position and a fatty acyl bond in the 2-position. The sn-1 position consists of palmitic acid (16 : 0), stearic acid (18 : 0), or oleic acid (18 : 1) carbon chains, and the head group is

usually either ethanolamine or choline. Thus, there are two main types of plasmalogens: ethanolamine plasmalogens and choline plasmalogens. The sn-2 position is generally occupied by PUFAs, specifically arachidonic acid or docosahexaenoic acid [78, 79].

The highest concentrations of plasmalogens are found in the brain, red blood cells, skeletal muscle, and spermatozoa and can represent as much as 18–20% of the total phospholipids in cell membranes [78, 79]. Plasmalogens are either derived from dietary sources and/or are synthesized mainly in the liver and gastrointestinal epithelium. Plasmalogens are not only important structural phospholipids in the cell membranes but they are also reservoirs of secondary messages and mediators of membrane dynamics and involved in membrane fusion, ion transport, cholesterol efflux, membrane-bound enzyme activity, and diffusion of signal-transduction molecules [80].

Secondary deficiency of plasmalogens triggered by their synthesis reduction or their degradation growth is associated with metabolic and inflammatory disorders such as cardiac diseases and diabetes mellitus [77]. The specificity of choline plasmalogens as a sensitive biomarker of an atherogenic state was confirmed. On the one hand, positive correlations of the choline plasmalogen content with serum adiponectin concentration and high-density lipoproteins (HDL), and on the other hand, inverse relationships with waist circumference, including triacylglycerides and low-density lipoproteins (LDL) content, have been identified. Reduced levels of ethanolamine plasmalogens in plasma have been shown to be also closely associated with cardiovascular, metabolic, and cancer diseases [81]. The content of plasmalogens is relatively stable in all lipoprotein fractions. However, the correlation between the levels of choline plasmalogens and HDL is stronger than that between the levels of ethanolamine plasmalogens and HDL. In the study by Pietiläinen et al., a decrease in the level of plasmalogens in adipocyte membranes in obese twins was established compared with metabolically healthy twins. Conversely, plasmalogen levels increase in trained people and dietary patients [82]. At the same time, it was found that the level of plasmalogens increases in the liver of rats receiving a high-fat diet [83].

The adaptation of the phospholipid composition of cells to exogenous lipid changes has been verified [83]. The compensatory response to a decrease in the plasmalogen level is the regulation of the level of phosphatidylethanolamine [84]. However, with plasmalogen deficiency, the total amount of PUFAs in phosphatidylethanolamine remains constant in human fibroblasts and in the brains of mice. Plasmalogens have been noted to play an important role as neuroprotectors and modulators of the signaling mechanisms of cell membranes [85]. Plasmalogens also act as endogenous antioxidants, protecting lipids and lipoproteins from oxidative stress [86]. This can be attributed to the fact that the hydrogen atoms adjacent to the vinyl ether bond are more susceptible to oxidation, protecting PUFAs from it that are found in the sn-2 position of the glycerol residue. Plasmalogen oxidation products are not capable of further initiation of lipid peroxidation processes. Another important function of plasmalogens is their participation in cell metabolism and transmembrane transport of FAs. The presence of PUFAs in the side chains of plasmalogens preconfigures their significant depositing function [77]. Cholesterol esterification depends on the level of plasmalogens. So, for example, the cells characterized by plasmalogen deficiency demonstrated a lower level of esterified cholesterol and a higher level of free and total cholesterol [84].

Therefore, the important role of plasmalogens as modulators of signaling mechanisms in protecting cells from lipid peroxidation and participation in PUFA metabolism has been made clear. However, the exact biological functions of plasmalogens and the underlying molecular mechanisms still remain to be discovered [52, 55].

29.7.6: 4. FREE FATTY ACIDS AND METABOLIC SYNDROME RISK

Free fatty acids (FFAs), or nonesterified fatty acids (NEFAs), in circulating plasma are derived from the ingestion of dietary fat or from the triglycerides stored in adipose tissue that are distributed to cells to serve as fuel for muscle contraction and systemic metabolism [87]. As FAs are insoluble in water, they are transported by binding to plasma albumin. FFAs can be taken up from circulating plasma by all mitochondria-containing cells, and they are metabolized by β-oxidation [17]. FFAs carry out many important biological functions in the body, and they are a source of energy, signal molecules, and structural components of cell membranes [17]. FFAs are involved in the pathogenesis of insulin resistance and subsequent development of metabolic syndrome [12, 17, 87]. Chronic energy imbalance can trigger adipocyte hypertrophy, endoplasmic reticulum stress, and mitochondrial dysfunction, which lead to the systemic release of FFAs [17, 88–90]. When plasma FFA levels rise, as occurs in obesity, a lipotoxicity state is induced, which induce activation of different cell responses: oxidative stress, apoptosis, and inflammation [17]. Consequently, FFAs play a highly important role in the association between obesity and insulin resistance.

29.7.6.1: 4.1. FFAS AND MITOCHONDRIA

There is an interesting hypothesis that IR is associated with the development of mitochondrial dysfunction [87, 90]. Lipid degradation occurs in mitochondria. On top of that, the normal functioning of mitochondria provides glucose-stimulated insulin secretion from β-cells of the pancreas. Initially, the theories have suggested that impaired mitochondrial function leads to impaired β-oxidation of lipids, which is accompanied by the accumulation of FFAs in peripheral tissues (lipotoxicity theory) [91]. The accumulation of lipid metabolites brings about the activation of kinases involved in the disruption of insulin signaling at the level of insulin receptor substrate 1 (IRS-1). In skeletal muscles, insulin signaling pathway disorder is accompanied by a decrease in the production of GLUT4 and glucose uptake by cells. In this case, an improvement in insulin sensitivity can be achieved by enhancing the β-oxidation of lipids. This theory was supported by studies that proved an increase in the rate of β-oxidation of lipids to be followed by protection against the development of IR [17].

Nevertheless, the early stages of obesity and IR development are characterized by an increase in β-oxidation of lipids. Besides, an impairment of fat oxidation results in higher insulin production. Therefore, mitochondrial dysfunction in skeletal muscles cannot be the only reason for the development of IR [92].

An alternative explanation of the relationship between mitochondria and insulin resistance is focused on the production of a reactive oxygen species (ROS) by mitochondria as a result of excess accumulation of FFAs in them [92]. Oxidative stress is known to be a pathogenetic component of chronic inflammation development and IR [88]. An oxidized redox environment can induce insulin resistance by directly affecting the protein involved in glucose uptake [89].

On the other hand, changes in redox cell homeostasis have been argued to step up the activity of the serine-/threonine-sensitive stress kinases that inhibit the transmission of insulin signals, inducing the development of IR [93, 94].

Oxidative stress also can stimulate the activation of transcriptional factors, such as nuclear factor-kappa B (NF-κB), activator protein 1 (AP-1), and hypoxia-inducible factor 1 (HIF-1), which promote the synthesis of inflammatory cytokines (IL-1β, IL-6, and TNF-α) (see Figure 3). These inflammatory cytokines contribute to obesity-associated local inflammation and directly induce insulin resistance. Also, chronic prolonged FFAs excess is the cause of pancreatic β-cells dysfunction. In addition, FFAs inhibit insulin gene expression and induce apoptosis in these cells [17].

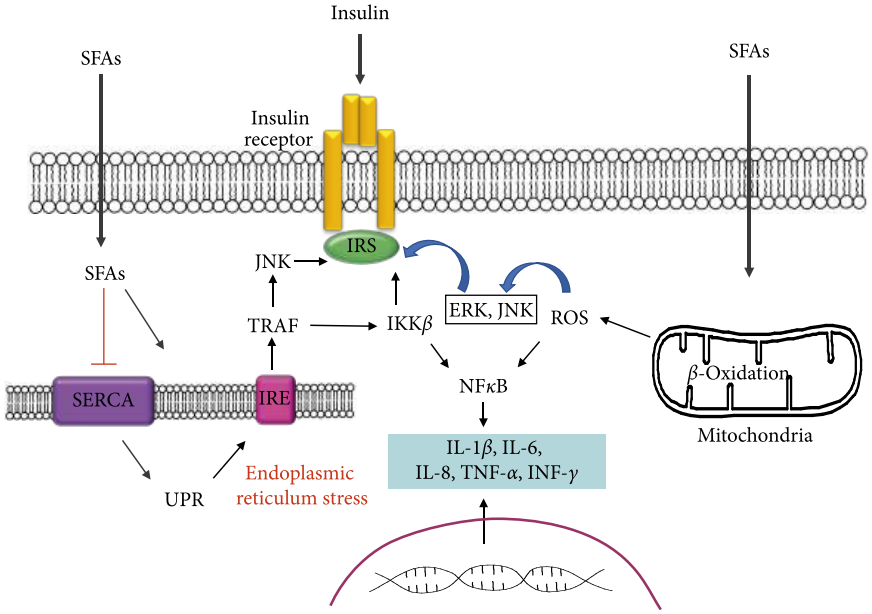


Figure 3

FFA-induced insulin resistance through endoplasmic reticulum stress and oxidative stress. A high level of FFA induces an increase in the production of ROS by mitochondria and the formation of oxidative stress. ROS stimulates NF-κB, which promotes the synthesis of L-1β, IL-6, and TNF-α. These inflammatory cytokines contribute to obesity-associated local inflammation and directly induce insulin resistance. In response to the enhanced level of FFAs and other nutrients in fats, adipose cells can develop signs of ER stress. A decrease in SERCA expression promotes the development of ER stress. UPR triggers the activation of IRE. Activation of IRE induces interaction with TRAF protein, which stimulations activation of IKKβ and JNK kinases. Its reaction can phosphorylate the IRS, thus blocking insulin signaling. JNK and IKKβ also lead to NF-κB activation and the development of inflammation.

Although discussing the role of mitochondrial skeletal muscle dysfunction in the pathogenesis of IR and type 2 diabetes is still underway [93], it is generally accepted that a mitochondrial defect does occur in these diseases. The connection between IR and mitochondrial dysfunction of liver cells, visceral, and subcutaneous adipose tissue has been proved [95]. Moreover, in the mitochondria of individuals suffering from obesity and type 2 diabetes, ATP synthesis is reduced, which correlates with the accumulation of FFAs and inhibition of insulin-stimulated glucose utilization.

29.7.6.2: 4.2. FFAS AND ENDOPLASMIC RETICULUM STRESS

Results of numerous studies establish that dysregulation of the endoplasmic reticulum (ER) function contributes to the development of MetS [96, 97]. Mitochondria are known to be both functionally and structurally associated with ER [97]. Obviously, the changes of the structure and function of these organelles can serve as a trigger for the development of metabolic homeostasis disorders [96]. ER is involved in maintaining Ca2+ homeostasis and participates in maturation and expression of membrane and secretion proteins. Cell stress conditions that increase ER demand and entail an overload of its functional capacity cause a series of alterations known as “endoplasmic reticulum stress.” Under these conditions, the ER activates a compensatory mechanism called the “unfolding protein response” (UPR), which attempts to restore the homeostasis of ER functions. With the stressful effects lasting for a long time, ER stress results in cell death (apoptosis) [17, 98].

UPR triggers activation of inositol-requiring endoribonuclease enzyme (IRE) (see Figure 3). The activation of IRE induces interaction with TRAF protein, which stimulations activation of IKKβ and JNK kinases. Its reaction can phosphorylate IRS, thus blocking insulin signaling. In response to the enhanced level of FFAs and other nutrients in fats, adipose cells can develop signs of ER stress [17]. ER stress produces insulin resistance mainly through JNK activation. JNK activity has been detected to be elevated in animal models of obesity, and JNK isoforms 1 and 2 deletion protects mice from insulin resistance induced by a fat-rich diet. Experimental evidence indicates that, on the one hand, JNK phosphorylates serine IRS-1, and on the other, it phosphorylates IKKβ, which leads to NF-κB activation and to inflammation development [96]. Remarkably, that change in expression of sarco/endoplasmic reticulum Ca2+ ATPase (SERCA), which has calcium elimination from the cytosol and returns it to the ER as their function, is associated with ER stress and subsequently with insulin resistance. The treatment of people with diabetes mellitus by rosiglitazone, an antidiabetic drug, increased SERCA expression, thus restoring the pump expression reduction observed in diabetic patients with altered hyperglycemia [17]. This way, the decrease in SERCA expression promotes the development of ER stress, with JNK ensuing activation, which desensitizes the insulin signal, thus generating a state of insulin resistance and contributing to chronic metabolic deterioration.

29.7.6.3: 4.3. FFAS AS LIGANDS FOR FFAR

FFAs serve not only as energy sources but also as natural ligands for a group of orphan G protein-coupled receptors (GPCRs) termed free fatty acid receptors (FFARs) [99]. The GPCR superfamily is the largest one in the human genome and encompasses some subfamilies (Gq, Gi, Gs, and G12/13) [100]. These receptors respond to various ligands and, therefore, are involved in the pathogenesis of many diseases, e.g., MetS, and are the target for more than half of pharmaceutical products [101–106]. There are four main members of FFAR family: FFAR1 (GPR40), FFAR2 (GPR43), FFAR3 (GPR41), and FFAR4 (GPR120 and GPR84) (see Table 1) [75].

Table 1

Family of FFARs and their ligands.

FFAR1 expression was revealed in neurons and in pancreas β-cells [99]. FFAR2 and FFAR3 are common in leukocytes and adipose tissues. Besides, FFAR3 is also expressed by pancreas cells, in the sympathetic nervous system and vessel plain muscles [100]. FFAR4 is expressed in adipocytes, the intestinal tract, macrophages, and in the central nervous system [105]. There are other specific receptors for FFA: GPR119 and GRP84. GPR119 is expressed in intestinal endocrine cells and pancreatic β-cells and activates the synthesis of GLP-1 and insulin. GPR84 is expressed in the spleen, thymus, leukocytes, and macrophages [99]. Long- and medium-chain length fatty acids are endogenous ligands for FFAR1, FFAR4, and GPR84. FFAR2 and FFAR3 are activated by short-chain FAs. FFAR2 is capable of binding with Gq and Gi proteins, whereas FFAR3 binds only with Gi. FFAR4 is activated by n-3 or n-6 PUFAs [99]. Thus, each FFAR can act as an FFA sensor with selectivity for a particular FFA carbon chain length derived from food or food-derived metabolites. FFARs have been reported to have physiological functions such as facilitation of insulin and incretin hormone secretion, adipocyte differentiation, anti-inflammatory effects, neuronal responses, and taste preferences [106]. Dysfunction of FFARs underlies the pathogenesis of many metabolic diseases, such as MetS and diabetes mellitus.

It has been found that FFAR4 acts as an anti-inflammatory receptor in proinflammatory macrophages and mature adipocytes. Signaling of FFAR4 activated by n-3 PUFAs inhibits TLR signaling and TNF-α-induced inflammatory responses. FFAR4 dysfunction leads to obesity and glucose intolerance in humans and mice [107]. Many results strongly support that FFAR4-mediated anti-inflammatory effects reduce the infiltration of proinflammatory macrophages into the adipose tissue and improve insulin sensitivity [102].

The activation of FFAR1 signaling enhances glucose-stimulated insulin secretion (GSIS) directly via stimulation of insulin secretion from pancreatic β cells and indirectly via the production of incretin hormones. Also, the activation of FFAR1 signaling reduces the expression of inflammatory cytokines such as TNF-α and IL-8. It has been shown that α-linolenic (18:3n3) and oleic (18:1n9) acids improve insulin resistance in obesity and type 2 diabetes [108].

There is some scientific evidence that short-chain fatty acids (SCFAs) are a substantial modulator of MetS inflammation [109]. SCFAs are the end products of metabolic fermentation of dietary fibers by gut microbiota. FFAR2 is a receptor for SCFAs and is expressed in enteroendocrine cells, adipose tissues, and pancreatic β-cells [99]. Dietary fiber intake reduces the risk of obesity, diabetes, inflammatory bowel disease, colon cancer, and cardiovascular disease. SCFAs supplementation with a high-fat diet improved insulin sensitivity and increased energy expenditure in a mouse model of diet-induced obesity [110, 111]. SCFAs are involved in intestinal immune homeostasis due to their role in regulating T cell polarization and differentiation. In human monocytes, SCFAs decrease the production of TNF-α and increase the production of PGE2 [109]. Activation of FFAR2 by SCFAs regulates metabolic disorders, increases energy expenditure, and preferentially enables fat consumption by inhibition of insulin signaling in adipose tissues. The expression of FFAR2 in neutrophils and mononuclear cells regulates intestinal homeostasis and inflammation. In light of this evidence, regulation FFAR2 expression and/or high fiber consumption may be a potential target for therapeutic intervention of MetS.

FFAR3 is also a receptor for SCFAs. FFAR3 is widely expressed in enteroendocrine cells, adipose tissues, the peripheral nervous system, peripheral blood mononuclear cells, monocytes, and macrophages [99]. FFAR3 expression in intestinal epithelial cells enhances the synthesis of proinflammatory mediators through extracellular signal-regulated kinase 1/2 and p38 MAPK [100]. Since these pathways help

to protect against bacterial infection, FFAR3 can stimulate acute inflammatory reactions in the intestine that have beneficial effects on host homeostasis [112]. Thus, FFAR 3 can exhibit proinflammatory properties.

29.7.7: 5. CONCLUSION

The wide phenotypic heterogeneity of MetS and its complex pathogenesis make it difficult to identify a therapeutic target. This syndrome is considered as a cluster of pathogenetically related conditions caused by metabolic disorders and the development of chronic, low-grade inflammation. In this review, we examined the molecular mechanisms of the development of MetS driven by impaired lipid metabolism. PUFAs and FFAs have been shown to play an important role in both the pathogenesis and treatment of MetS. Fatty acids perform structural, energy, signaling, and immunoregulatory functions. These FAs properties underlie the pathogenetic mechanisms of glucose transport disturbance, the development of IR and chronic inflammation, the formation of oxidative stress, and mitochondrial dysfunction in MetS. Correction in lifestyle and nutrition is considered as the main way to minimize complications caused by an imbalance in the body between saturated and polyunsaturated fatty acids. At the same time, there are controversial data about the therapeutic efficacy of dietary n-3 PUFAs in MetS [50]. SPMs have shown potent pro-resolving actions in different disease models, including MetS [61]. SPM-based therapeutics could be one of the most optimistic treatments for MetS. Further studies are needed to detail the mechanisms of FA participation and their oxidized metabolites in the development of inflammation and pathogenesis of MetS.

29.7.8: CONFLICTS OF INTEREST

The authors declare that there are no conflicts of interest regarding the publication of this paper.

29.7.9: ACKNOWLEDGMENTS

The study was funded by the Ministry of Education and Science of the Russian Federation.

29.7.10: REFERENCES

1. P. M. Nilsson, J. Tuomilehto, and L. Rydén, “The metabolic syndrome—what is it and how should it be managed?,,” *European Journal of Preventive Cardiology*, vol. 26, no. 2_suppl, pp. 33–46, 2019.View at: Publisher Site | Google Scholar
2. H. H. Wang, D. K. Lee, M. Liu, P. Portincasa, and D. Q.-H. Wang, “Novel insights into the pathogenesis and management of the metabolic syndrome,” *Pediatric Gastroenterology, Hepatology & Nutrition*, vol. 23, no. 3, pp. 189–230, 2020.View at: Publisher Site | Google Scholar
3. P. Ranasinghe, Y. Mathangasinghe, R. Jayawardena, A. P. Hills, and A. Misra, “Prevalence and trends of metabolic syndrome among adults in the asia-pacific region: a systematic review,” *BMC Public Health*, vol. 17, no. 1, p. 101, 2017.View at: Publisher Site | Google Scholar
4. F. S. Sigit, D. L. Tahapary, S. Trompet et al., “The prevalence of metabolic syndrome and its association with body fat distribution in middle-aged individuals from Indonesia and the Netherlands: a cross-sectional analysis of two population-based studies,” *Diabetology & Metabolic Syndrome*, vol. 12, no. 1, p. 2, 2020.View at: Publisher Site | Google Scholar
5. M.-K. Lee, K. Han, M. K. Kim et al., “Changes in metabolic syndrome and its components and the risk of type 2 diabetes: a nationwide cohort study,” *Scientific Reports*, vol. 10, no. 1, p. 2313, 2020.View at: Publisher Site | Google Scholar
6. Y. Denisenko, T. Novgorodtseva, N. Zhukova et al., “Metabolic syndrome: modification of the fatty acid composition and glucose-insulin homeostasis,” *British Journal of Medicine and Medical Research*, vol. 8, no. 11, pp. 975–987, 2015.View at: Publisher Site | Google Scholar
7. P. Silva Figueiredo, A. Carla Inada, G. Marcelino et al., “Fatty acids consumption: the role metabolic aspects involved in obesity and its associated disorders,” *Nutrients*, vol. 9, no. 10, p. 1158, 2017.View at: Publisher Site | Google Scholar
8. I. Grgurevic, K. PodrugStéphanie Monnerie, B. Comte et al., “Metabolomic and lipidomic signatures of metabolic syndrome and its physiological components in adults: a systematic review,” *Scientific Reports*, vol. 10, p. 669, 2020.View at: Publisher Site | Google Scholar
9. I. Surowiec, R. Noordam, K. Bennett et al., “Metabolomic and lipidomic assessment of the metabolic syndrome in Dutch middle-aged individuals reveals novel biological signatures separating health and disease,” *Metabolomics*, vol. 15, no. 2, p. 23, 2019.View at: Publisher Site | Google Scholar
10. T. P. Novgorodtseva, Y. K. Karaman, N. V. Zhukova, E. G. Lobanova, M. V. Antonyuk, and T. A. Kantur, “Composition of fatty acids in plasma and erythrocytes and eicosanoids level in patients with metabolic syndrome,” *Lipids in Health and Disease*, vol. 10, no. 1, p. 82, 2011.View at: Publisher Site | Google Scholar
11. A. González-Périz, R. Horrillo, N. Ferré et al., “Obesity-induced insulin resistance and hepatic steatosis are alleviated by ω3 fatty acids: a role for resolvins and protectins,” *The FASEB Journal*, vol. 23, no. 6, pp. 1946–1957, 2009.View at: Publisher Site | Google Scholar
12. R. N. Weijers, “Lipid composition of cell membranes and its relevance in type 2 diabetes mellitus,” *Current Diabetes Reviews*, vol. 8, no. 5, pp. 390–400, 2012.View at: Google Scholar
13. L. Martínez-Fernández, L. M. Laiglesia, A. E. Huerta, J. A. Martínez, and M. J. Moreno-Aliaga, “Omega-3 fatty acids and adipose tissue function in obesity and metabolic syndrome,” *Prostaglandins & Other Lipid Mediators*, vol. 121, pp. 24–41, 2015.View at:

Publisher Site | Google Scholar
14. A. J. Tremblay, J.-P. Després, M.-È. Piché et al., “Associations between the fatty acid content of triglyceride, visceral adipose tissue accumulation, and components of the insulin resistance syndrome,” *Metabolism*, vol. 53, no. 3, pp. 310–317, 2004.View at: Publisher Site | Google Scholar
15. S. D. Phinney, “Fatty acids, inflammation, and the metabolic syndrome,” *The American Journal of Clinical Nutrition*, vol. 82, no. 6, pp. 1151-1152, 2005.View at: Publisher Site | Google Scholar
16. P. C. Calder, “Omega-3 fatty acids and inflammatory processes: from molecules to man,” *Biochemical Society Transactions*, vol. 45, no. 5, pp. 1105–1115, 2017.View at: Publisher Site | Google Scholar
17. J. G. Vázquez-Jiménez, A. Roura-Guiberna, L. R. Jiménez-Mena, and J. A. Olivares-Reyes, “Role of free fatty acids on insulin resistance,” *Gaceta Médica de México*, vol. 153, pp. 773–783, 2017.View at: Publisher Site | Google Scholar
18. D. S. Lark, K. H. Fisher-Wellman, and P. D. Neufer, “High-fat load: mechanism(s) of insulin resistance in skeletal muscle,” *International Journal of Obesity Supplements*, vol. 2, no. S2, pp. S31–S36, 2012.View at: Publisher Site | Google Scholar
19. H. Cena and P. C. Calder, “Defining a healthy diet: evidence for the role of contemporary dietary patterns in health and disease,” *Nutrients*, vol. 12, no. 2, p. 334, 2020.View at: Publisher Site | Google Scholar
20. C. K. Glass and J. M. Olefsky, “Inflammation and lipid signaling in the etiology of insulin resistance,” *Cell Metabolism*, vol. 15, no. 5, pp. 635–645, 2012.View at: Publisher Site | Google Scholar
21. M. Lepretti, S. Martucciello, M. Burgos Aceves, R. Putti, and L. Lionetti, “Omega-3 fatty acids and insulin resistance: focus on the regulation of mitochondria and endoplasmic reticulum stress,” *Nutrients*, vol. 10, no. 3, p. 350, 2018.View at: Publisher Site | Google Scholar
22. P. Koehrer, N. Acar, S. Saab et al., “Polyunsaturated fatty acids and plasmalogens in diabetics,” *Investigative Ophthalmology & Visual Science*, vol. 54, p. 1158, 2013.View at: Google Scholar
23. O. Y. Kytikova, T. P. Novgorodtseva, M. V. Antonyuk, and T. A. Gvozdenko, “Plasmalogenes in the pathophysiology and therapy of age-specific diseases,” *Advances in Gerontology*, vol. 32, no. 6, pp. 948–958, 2019.View at: Google Scholar
24. P. Reddy, D. Lent-Schochet, N. Ramakrishnan, M. Mclaughlin, and I. Jialal, “Metabolic syndrome is an inflammatory disorder: a conspiracy between adipose tissue and phagocytes,” *Clinica Chimica Acta*, vol. 496, pp. 35–44, 2019.View at: Publisher Site | Google Scholar
25. C. Lambert, J. Cubedo, T. Padró et al., “Phytosterols and omega 3 supplementation exert novel regulatory effects on metabolic and inflammatory pathways: a proteomic study,” *Nutrients*, vol. 9, no. 6, p. 599, 2017.View at: Publisher Site | Google Scholar
26. O.-K. Kim, W. Jun, and J. Lee, “Mechanism of ER stress and inflammation for hepatic insulin resistance in obesity,” *Annals of Nutrition and Metabolism*, vol. 67, no. 4, pp. 218–227, 2015.View at: Publisher Site | Google Scholar
27. G. Grandl and C. Wolfrum, “Hemostasis, endothelial stress, inflammation, and the metabolic syndrome,” *Seminars in Immunopathology*, vol. 40, no. 2, pp. 215–224, 2018.View at: Publisher Site | Google Scholar
28. M. S. Han, A. White, R. J. Perry et al., “Regulation of adipose tissue inflammation by interleukin 6,” *Proceedings of the National Academy of Sciences*, vol. 117, no. 6, pp. 2751–2760, 2020.View at: Publisher Site | Google Scholar
29. D. W. Gilroy and D. Bishop-Bailey, “Lipid mediators in immune regulation and resolution,” *British Journal of Pharmacology*, vol. 176, no. 8, pp. 1009–1023, 2019.View at: Publisher Site | Google Scholar
30. S. E. Headland and L. V. Norling, “The resolution of inflammation: principles and challenges,” *Seminars in Immunology*, vol. 27, no. 3, pp. 149–160, 2015.View at: Publisher Site | Google Scholar
31. O. Kytikova, T. Novgorodtseva, Y. Denisenko, M. Antonyuk, and T. Gvozdenko, “Pro-resolving lipid mediators in the pathophysiology of asthma,” *Medicina*, vol. 55, no. 6, p. 284, 2019.View at: Publisher Site | Google Scholar
32. C. N. Serhan, N. Chiang, and J. Dalli, “New pro-resolving n-3 mediators bridge resolution of infectious inflammation to tissue regeneration,” *Molecular Aspects of Medicine*, vol. 64, pp. 1–17, 2018.View at: Publisher Site | Google Scholar
33. M. Aursnes, J. E. Tungen, A. Vik et al., “Total synthesis of the lipid mediator PD1n-3 DPA: configurational assignments and anti-inflammatory and pro-resolving actions,” *Journal of Natural Products*, vol. 77, no. 4, pp. 910–916, 2014.View at: Publisher Site | Google Scholar
34. E. S. K. Doğan, B. Doğan, Ö Fentoğlu, and F. Y. Kırzıoğlu, “The role of serum lipoxin A4 levels in the association between periodontal disease and kwonlic syndrome,” *Journal of Periodontal & Implant Science*, vol. 49, no. 2, pp. 105–113, 2019.View at: Publisher Site | Google Scholar
35. A. E. Barden, E. Mas, K. D. Croft, M. Phillips, and T. A. Mori, “Specialized proresolving lipid mediators in humans with the metabolic syndrome after n-3 fatty acids and aspirin,” *The American Journal of Clinical Nutrition*, vol. 102, no. 6, pp. 1357–1364, 2015.View at: Publisher Site | Google Scholar
36. Y. Kwon, “Immuno-resolving ability of resolvins, protectins, and maresins derived from omega-3 fatty acids in metabolic syndrome,” *Molecular Nutrition & Food Research*, vol. 64, no. 4, Article ID e1900824, 2020.View at: Publisher Site | Google Scholar
37. Y. Xiang, W. Zhou, X. Duan et al., “Metabolic syndrome, and particularly the hypertriglyceridemic-waist phenotype, increases breast cancer risk, and adiponectin is a potential mechanism: a case-control study in Chinese women,” *Front Endocrinol (Lausanne)*, vol. 10, no. 905, 2019.View at: Google Scholar

38. J. H. Lee, K. S. Lee, H. Kim et al., “The relationship between metabolic syndrome and the incidence of colorectal cancer,” *Environmental Health and Preventive Medicine*, vol. 25, no. 6, 2020.View at: Publisher Site | Google Scholar
39. Y. Wang, R. Tu, H. Yuan et al., “Associations of unhealthy lifestyles with metabolic syndrome in Chinese rural aged females,” *Scientific Reports*, vol. 10, no. 1, p. 2718, 2020.View at: Publisher Site | Google Scholar
40. K. O. Ju, M. V. Antonjuk, T. A. Gvozdenko, and T. P. Novgorodceva, “Metabolic aspects of the relationship of obesity and bronchial asthma,” *Obesity and Metabolism*, vol. 15, no. 4, pp. 9–14, 2018.View at: Publisher Site | Google Scholar
41. M. Amatruda, G. Ippolito, S. Vizzuso, G. Vizzari, G. Banderali, and E. Verduci, “Epigenetic effects of n-3 LCPUFAs: a role in pediatric metabolic syndrome,” *International Journal of Molecular Sciences*, vol. 20, no. 9, p. 2118, 2019.View at: Publisher Site | Google Scholar
42. J. B. Halter, N. Musi, F. McFarland Horne et al., “Diabetes and cardiovascular disease in older adults: current status and future directions,” *Diabetes*, vol. 63, no. 8, pp. 2578–2589, 2014.View at: Publisher Site | Google Scholar
43. D. Yach, D. Stuckler, and K. D. Brownell, “Epidemiologic and economic consequences of the global epidemics of obesity and diabetes,” *Nature Medicine*, vol. 12, no. 1, pp. 62–66, 2006.View at: Publisher Site | Google Scholar
44. WHO—World Health Organization, “World Health Organization obesity and overweight fact sheet,” 2020, https://www.who.int/en/news-room/fact-sheets/detail/obesity-and-overweight.View at: Google Scholar
45. J. Jiang, X. Cai, Y. Pan et al., “Relationship of obesity to adipose tissue insulin resistance,” *BMJ Open Diabetes Research & Care*, vol. 8, no. 1, Article ID e000741, 2020.View at: Publisher Site | Google Scholar
46. T. Harayama and H. Riezman, “Understanding the diversity of membrane lipid composition,” *Nature Reviews Molecular Cell Biology*, vol. 19, no. 5, pp. 281–296, 2018.View at: Publisher Site | Google Scholar
47. E. Tortosa-Caparrós, D. Navas-Carrillo, F. Marín, and E. Orenes-Piñero, “Anti-inflammatory effects of omega 3 and omega 6 polyunsaturated fatty acids in cardiovascular disease and metabolic syndrome,” *Critical Reviews in Food Science and Nutrition*, vol. 57, no. 16, pp. 3421–3429, 2017.View at: Publisher Site | Google Scholar
48. P. C. Calder, “n-3 PUFA and inflammation: from membrane to nucleus and from bench to bedside,” *Proceedings of the Nutrition Society*, pp. 1–37, 2020.View at: Publisher Site | Google Scholar
49. Y. Iwase, N. Kamei, and M. Takeda-Morishita, “Antidiabetic effects of omega-3 polyunsaturated fatty acids: from mechanism to therapeutic possibilities,” *Pharmacology & Pharmacy*, vol. 6, no. 3, pp. 190–200, 2015.View at: Publisher Site | Google Scholar
50. J. E. Radcliffe, J. Thomas, A. L. Bramley et al., “Controversies in omega-3 efficacy and novel concepts for application,” *Journal of Nutrition & Intermediary Metabolism*, vol. 5, pp. 11–22, 2016.View at: Publisher Site | Google Scholar
51. K. Albracht-Schulte, N. S. Kalupahana, L. Ramalingam et al., “Omega-3 fatty acids in obesity and metabolic syndrome: a mechanistic update,” *The Journal of Nutritional Biochemistry*, vol. 58, pp. 1–16, 2018.View at: Publisher Site | Google Scholar
52. X.-f. Guo, X. Li, M. Shi, and D. Li, “n-3 polyunsaturated fatty acids and metabolic syndrome risk: a meta-analysis,” *Nutrients*, vol. 9, p. 703, 2017.View at: Publisher Site | Google Scholar
53. Y.-S. Kim, P. Xun, and K. He, “Fish consumption, long-chain omega-3 polyunsaturated fatty acid intake and risk of metabolic syndrome: a meta-analysis,” *Nutrients*, vol. 7, no. 4, pp. 2085–2100, 2015.View at: Publisher Site | Google Scholar
54. R. N. M. Weijers, “Membrane flexibility, free fatty acids, and the onset of vascular and neurological lesions in type 2 diabetes,” *Journal of Diabetes & Metabolic Disorders*, vol. 15, no. 1, p. 13, 2016.View at: Publisher Site | Google Scholar
55. S. W. Wong, M.-J. Kwon, A. M. K. Choi, H.-P. Kim, K. Nakahira, and D. H. Hwang, “Fatty acids modulate toll-like receptor 4 activation through regulation of receptor dimerization and recruitment into lipid rafts in a reactive oxygen species-dependent manner,” *Journal of Biological Chemistry*, vol. 284, no. 40, pp. 27384–27392, 2009.View at: Publisher Site | Google Scholar
56. A. Salas-Burgos, P. Iserovich, F. Zuniga, J. C. Vera, and J. Fischbarg, “Predicting the three-dimensional structure of the human facilitative glucose transporter Glut1 by a novel evolutionary homology strategy: insights on the molecular mechanism of substrate migration, and binding sites for glucose and inhibitory molecules,” *Biophysical Journal*, vol. 87, no. 5, pp. 2990–2999, 2004.View at: Publisher Site | Google Scholar
57. M. Rogero and P. C. Calder, “Obesity, inflammation, toll-like receptor 4 and fatty acids,” *Nutrients*, vol. 10, p. 432, 2018.View at: Publisher Site | Google Scholar
58. A. M. Minihane, S. Vinoy, W. R. Russell et al., “Low-grade inflammation, diet composition and health: current research evidence and its translation,” *British Journal Of Nutrition*, vol. 114, no. 7, pp. 999–1012, 2015.View at: Publisher Site | Google Scholar
59. R. S. Peebles, “Prostaglandins in asthma and allergic diseases,” *Pharmacology & Therapeutics*, vol. 193, pp. 1–19, 2019.View at: Publisher Site | Google Scholar
60. S. Marcone, P. Evans, and D. J. Fitzgerald, “15-deoxy-Δ12,14-prostaglandin J2 modifies components of the proteasome and inhibits inflammatory responses in human endothelial cells,” *Frontiers in Immunology*, vol. 7, p. 459, 2016.View at: Publisher Site | Google Scholar
61. C. N. Serhan and B. D. Levy, “Resolvins in inflammation: emergence of the pro-resolving superfamily of mediators,” *Journal of Clinical Investigation*, vol. 128, no. 7, pp. 2657–2669, 2018.View at: Publisher Site | Google Scholar
62. O. Y. Kytikova, J. M. Perelman, T. P. Novgorodtseva et al., “Peroxisome proliferator-activated receptors as a therapeutic target in asthma,” *PPAR Research*, vol. 2020, Article ID 8906968, 18 pages, 2020.View at: Publisher Site | Google Scholar

63. A. Neuhofer, M. Zeyda, D. Mascher et al., “Impaired local production of proresolving lipid mediators in obesity and 17-HDHA as a potential treatment for obesity-associated inflammation,” *Diabetes*, vol. 62, no. 6, pp. 1945–1956, 2013.View at: Publisher Site | Google Scholar
64. J. Clària, J. Dalli, S. Yacoubian, F. Gao, and C. N. Serhan, “Resolvin D1 and resolvin D2 govern local inflammatory tone in obese fat,” *The Journal of Immunology*, vol. 189, no. 5, pp. 2597–2605, 2012.View at: Publisher Site | Google Scholar
65. N. Morshedzadeh, A. Saedisomeolia, M. Djalali, M. R. Eshraghian, S. Hantoushzadeh, and M. Mahmoudi, “Resolvin D1 impacts on insulin resistance in women with polycystic ovary syndrome and healthy women,” *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, vol. 13, no. 1, pp. 660–664, 2019.View at: Publisher Site | Google Scholar
66. K. Sidletskaya, T. Vitkina, and Y. Denisenko, “The role of toll-like receptors 2 and 4 in the pathogenesis of chronic obstructive pulmonary disease,” *International Journal of Chronic Obstructive Pulmonary Disease*, vol. 15, pp. 1481–1493, 2020.View at: Publisher Site | Google Scholar
67. J. M. Ruysschaert and C. Lonez, “Role of lipid microdomains in TLR-mediated signalling,” *Biochimica et Biophysica Acta*, vol. 1848, pp. 1860–1867, 2015.View at: Publisher Site | Google Scholar
68. B. Prajapati, P. K. Jena, P. Rajput et al., “Understanding and modulating the toll like receptors (TLRs) and NOD like receptors (NLRs) cross talk in type 2 diabetes,” *Current Diabetes Reviews*, vol. 10, no. 3, pp. 190–200, 2014.View at: Publisher Site | Google Scholar
69. S. Huang, J. M. Rutkowsky, R. G. Snodgrass et al., “Saturated fatty acids activate TLR-mediated proinflammatory signaling pathways,” *Journal of Lipid Research*, vol. 53, pp. 2002–2013, 2012.View at: Publisher Site | Google Scholar
70. D. H. Hwang, J. A. Kim, and J. Y. Lee, “Mechanisms for the activation of Toll-like receptor 2/4 by saturated fatty acids and inhibition by docosahexaenoic acid,” *European Journal of Pharmacology*, vol. 785, pp. 24–35, 2016.View at: Publisher Site | Google Scholar
71. A. Z. Lalia and I. R. Lanza, “Insulin-sensitizing effects of omega-3 fatty acids: lost in translation?,,” *Nutrients*, vol. 8, p. 329, 2016.View at: Publisher Site | Google Scholar
72. Y. Li, S. L. Deng, Z. X. Lian et al., “Roles of toll-like receptors in nitroxidative stress in mammals,” *Cells*, vol. 8, no. 6, p. E576, 2019.View at: Publisher Site | Google Scholar
73. D. Huang, Q. Zhao, H. Liu, Y. Guo, and H. Xu, “PPAR-α agonist WY-14643 inhibits LPS-induced inflammation in synovial fibroblasts via NF-kB pathway,” *Journal of Molecular Neuroscience*, vol. 59, no. 4, pp. 544–553, 2016.View at: Publisher Site | Google Scholar
74. S. R. Shaikh, J. J. Kinnun, and S. R. Wassall, “How polyunsaturated fatty acids modify molecular organization in membranes: insight from NMR studies of model systems,” *Biochimica et Biophysica Acta*, vol. 1848, pp. 211–219, 2015.View at: Publisher Site | Google Scholar
75. S. R. Wassall, X. Leng, S. W. Canner, E. R. Pennington, J. J. Kinnun, and A. T. Cavazos, “Docosahexaenoic acid regulates the formation of lipid rafts: a unified view from experiment and simulation,” *Biochimica et Biophysica Acta (BBA)—Biomembranes*, vol. 1860, pp. 1985–1993, 2018.View at: Publisher Site | Google Scholar
76. M. G. Saklayen, “The global epidemic of the metabolic syndrome,” *Current Hypertension Reports*, vol. 20, p. 12, 2018.View at: Publisher Site | Google Scholar
77. O. Y. Kytikova, T. P. Novgorodtseva, M. V. Antonyuk et al., “Molecular targets of fatty acid ethanolamides in asthma,” *Medicina*, vol. 55, no. 87, p. 4, 2019.View at: Publisher Site | Google Scholar
78. M. Honsho, Y. Abe, and Y. Fujiki, “Plasmalogen biosynthesis is spatiotemporally regulated by sensing plasmalogens in the inner leaflet of plasma membranes,” *Scientific Reports*, vol. 7, p. 43936, 2017.View at: Publisher Site | Google Scholar
79. M. S. F. Messias, G. C. Mecatti, D. G. Priolli, and O. Carvalho, “Plasmalogen lipids: functional mechanism and their involvement in gastrointestinal cancer,” *Lipids in Health and Disease*, vol. 17, p. 41, 2018.View at: Publisher Site | Google Scholar
80. X. Su, J. Wang, and A. J. Sinclair, “Plasmalogens and Alzheimer’s disease: a review,” *Lipids in Health and Disease*, vol. 18, p. 100, 2019.View at: Publisher Site | Google Scholar
81. M. Barchuk, A. Dutour, P. Ancel et al., “Untargeted lipidomics reveals a specific enrichment in plasmalogens in epicardial adipose tissue and a specific signature in coronary artery disease,” *Arteriosclerosis, Thrombosis, and Vascular Biology*, vol. 40, no. 4, pp. 986–1000, 2020.View at: Publisher Site | Google Scholar
82. K. H. Pietiläinen, M. Sysi-Aho, A. Rissanen et al., “Acquired obesity is associated with changes in the serum lipidomic profile independent of genetic effects—a monozygotic twin study,” *PLoS One*, vol. 2, no. 2, Article ID e218, 2007.View at: Publisher Site | Google Scholar
83. J. M. Dean and I. J. Lodhi, “Structural and functional roles of ether lipids,” *Protein & Cell*, vol. 9, pp. 196–206, 2018.View at: Publisher Site | Google Scholar
84. S. Wallner, E. Orsó, M. Grandl, T. Konovalova, G. Liebisch, and G. Schmitz, “Phosphatidylcholine and phosphatidylethanolamine plasmalogens in lipid loaded human macrophages,” *PLoS One*, vol. 13, Article ID e0205706, 2018.View at: Publisher Site | Google Scholar
85. A. West, V. Zoni, W. E. Teague Jr. et al., “How do ethanolamine plasmalogens contribute to order and structure of neurological membranes?,,” *The Journal of Physical Chemistry B*, vol. 124, no. 5, pp. 828–839, 2020.View at: Publisher Site | Google Scholar
86. A. Broniec, A. Żądło, and A. Pawlak, “Interaction of plasmenylcholine with free radicals in selected model systems,” *Free Radical Biology and Medicine*, vol. 106, pp. 368–378, 2017.View at: Publisher Site | Google Scholar

87. C. Suiter, S. K. Singha, R. Khalili, and Z. Shariat-Madar, “Free fatty acids: circulating contributors of metabolic syndrome,” *Cardiovascular & Hematological Agents in Medicinal Chemistry*, vol. 16, no. 1, pp. 20–34, 2018.View at: Publisher Site | Google Scholar
88. A. Ghosh, L. Gao, A. Thakur, P. M. Siu, and C. W. K. Lai, “Role of free fatty acids in endothelial dysfunction,” *Journal of Biomedical Science*, vol. 24, no. 1, p. 50, 2017.View at: Publisher Site | Google Scholar
89. M. Alicka and K. Marycz, “The effect of chronic inflammation and oxidative and endoplasmic reticulum stress in the course of metabolic syndrome and its therapy,” *Stem Cells International*, vol. 2018, Article ID 4274361, 13 pages, 2018.View at: Publisher Site | Google Scholar
90. M. K. Montgomery and N. Turner, “Mitochondrial dysfunction and insulin resistance: an update,” *Endocrine Connections*, vol. 4, pp. R1–R15, 2015.View at: Publisher Site | Google Scholar
91. P. Ježek, M. Jabůrek, B. Holendová, and L. Plecitá-Hlavatá, “Fatty acid-stimulated insulin secretion vs. lipotoxicity,” *Molecules*, vol. 23, p. 1483, 2018.View at: Publisher Site | Google Scholar
92. J. O. Holloszy, ““Deficiency” of mitochondria in muscle does not cause insulin resistance,” *Diabetes*, vol. 62, pp. 1036–1040, 2013.View at: Publisher Site | Google Scholar
93. S. D. Martin and S. L. McGee, “The role of mitochondria in the aetiology of insulin resistance and type 2 diabetes,” *Biochimica et Biophysica Acta*, vol. 1840, pp. 1303–1312, 2014.View at: Publisher Site | Google Scholar
94. O. Zhenyukh, M. González-Amor, R. R. Rodrigues-Diez et al., “Branched-chain amino acids promote endothelial dysfunction through increased reactive oxygen species generation and inflammation,” *Journal of Cellular and Molecular Medicine*, vol. 22, no. 10, pp. 4948–4962, 2018.View at: Publisher Site | Google Scholar
95. S. Bouderba, M. N. Sanz, C. Sánchez-Martín et al., “Hepatic mitochondrial alterations and increased oxidative stress in nutritional diabetes-prone Psammomys obesus model,” *Experimental Diabetes Research*, vol. 2012, Article ID 430176, 8 pages, 2012.View at: Publisher Site | Google Scholar
96. J. Rieusset, “Role of endoplasmic reticulum-mitochondria communication in type 2 diabetes,” *Advances in Experimental Medicine and Biology*, vol. 997, pp. 171–186, 2017.View at: Publisher Site | Google Scholar
97. J. L. Maiers and H. Malhi, “Endoplasmic reticulum stress in metabolic liver diseases and hepatic fibrosis,” *Seminars in Liver Disease*, vol. 39, pp. 235–248, 2019.View at: Publisher Site | Google Scholar
98. L. D. Ly, S. Xu, S.-K. Choi et al., “Oxidative stress and calcium dysregulation by palmitate in type 2 diabetes,” *Experimental & Molecular Medicine*, vol. 49, p. e291, 2017.View at: Publisher Site | Google Scholar
99. I. Kimura, A. Ichimura, R. Ohue-Kitano, and M. Igarashi, “Free fatty acid receptors in health and disease,” *Physiological Reviews*, vol. 100, no. 1, pp. 171–210, 2020.View at: Publisher Site | Google Scholar
100. J. Miyamoto, S. Hasegawa, M. Kasubuchi, A. Ichimura, A. Nakajima, and I. Kimura, “Nutritional signaling via free fatty acid receptors,” *International Journal of Molecular Sciences*, vol. 17, p. 450, 2016.View at: Publisher Site | Google Scholar
101. T. Yonezawa, R. Kurata, K. Yoshida, M. A. Murayama, X. Cui, and A. Hasegawa, “Free fatty acids-sensing G protein-coupled receptors in drug targeting and therapeutics,” *Current Medicinal Chemistry*, vol. 20, no. 31, pp. 3855–3871, 2013.View at: Publisher Site | Google Scholar
102. M. Congreve, C. de Graaf, N. A. Swain, and C. G. Tate, “Impact of GPCR structures on drug discovery,” *Cell*, vol. 181, no. 1, pp. 81–91, 2020.View at: Publisher Site | Google Scholar
103. J. B. Pujol, N. Christinat, Y. Ratinaud, C. Savoia, S. E. Mitchell, and E. H. M. Dioum, “Coordination of GPR40 and ketogenesis signaling by medium chain fatty acids regulates beta cell function,” *Nutrients*, vol. 10, no. 4, p. E473, 2018.View at: Publisher Site | Google Scholar
104. Q. Zhang, H. Yang, J. Li, and X. Xie, “Discovery and characterization of a novel small-molecule agonist for medium-chain free fatty acid receptor g protein-coupled receptor 84,” *Journal of Pharmacology and Experimental Therapeutics*, vol. 357, pp. 337–344, 2016.View at: Publisher Site | Google Scholar
105. G. Milligan, B. Shimpukade, T. Ulven, and B. D. Hudson, “Complex pharmacology of free fatty acid receptors,” *Chemical Reviews*, vol. 117, pp. 67–110, 2017.View at: Publisher Site | Google Scholar
106. A. Bartoszek, E. V. Moo, A. Binienda et al., “Free fatty acid receptors as new potential therapeutic target in inflammatory bowel diseases,” *Pharmacological Research*, vol. 152, Article ID 104604, 2020.View at: Publisher Site | Google Scholar
107. J. Miyamoto, M. Kasubuchi, A. Nakajima, and I. Kimura, “Anti-inflammatory and insulin-sensitizing effects of free fatty acid receptors,” in *Free Fatty Acid Receptors. Handbook of Experimental Pharmacology*, G. Milligan and I. Kimura, Eds., vol. 236, Springer, Cham, Switzerland, 2016.View at: Google Scholar
108. V. Oliveira, “Diets containing α-linolenic (ω3) or oleic (ω9) fatty acids rescues obese mice from insulin resistance,” *Endocrinology*, vol. 156, no. 11, pp. 4033–4046, 2015.View at: Publisher Site | Google Scholar
109. J. Hu, S. Lin, B. Zheng, and P. C. K. Cheung, “Short-chain fatty acids in control of energy metabolism,” *Critical Reviews in Food Science and Nutrition*, vol. 58, no. 8, pp. 1243–1249, 2018.View at: Publisher Site | Google Scholar
110. M. Sun, W. Wu, Z. Liu, and Y. Cong, “Microbiota metabolite short chain fatty acids, GPCR, and inflammatory bowel diseases,” *Journal of Gastroenterology*, vol. 52, no. 1, pp. 1–8, 2017.View at: Publisher Site | Google Scholar

111. S. Sivaprakasam, P. D. Prasad, and N. Singh, “Benefits of short-chain fatty acids and their receptors in inflammation and carcinogenesis,” *Pharmacology & Therapeutics*, vol. 164, pp. 144–151, 2016.View at: Publisher Site | Google Scholar
112. M. H. Kim, “Short-chain fatty acids activate GPR41 and GPR43 on intestinal epithelial cells to promote inflammatory responses in mice,” *Gastroenterology*, vol. 145, no. 2, pp. 396–406, 2013.View at: Publisher Site | Google Scholar

29.7.11: COPYRIGHT

29.8: FUNDAMENTALS OF CANCER METABOLISM

29.8.1: FUNDAMENTALS OF CANCER METABOLISM

1. **Ralph J. DeBerardinis**[1],* **and**
2. **Navdeep S. Chandel**[2],*

Science Advances 27 May 2016: Vol. 2, no. 5, e1600200. DOI: 10.1126/sciadv.1600200

29.8.1.1: ABSTRACT

Tumors reprogram pathways of nutrient acquisition and metabolism to meet the bioenergetic, biosynthetic, and redox demands of malignant cells. These reprogrammed activities are now recognized as hallmarks of cancer, and recent work has uncovered remarkable flexibility in the specific pathways activated by tumor cells to support these key functions. In this perspective, we provide a conceptual framework to understand how and why metabolic reprogramming occurs in tumor cells, and the mechanisms linking altered metabolism to tumorigenesis and metastasis. Understanding these concepts will progressively support the development of new strategies to treat human cancer.

29.8.1.1: INTRODUCTION AND OVERARCHING PRINCIPLES

Cancer metabolism is one of the oldest areas of research in cancer biology, predating the discovery of oncogenes and tumor suppressors by some 50 years. The field is based on the principle that metabolic activities are altered in cancer cells relative to normal cells, and that these alterations support the acquisition and maintenance of malignant properties. Because some altered metabolic features are observed quite generally across many types of cancer cells, reprogrammed metabolism is considered a hallmark of cancer (*1*, *2*). Precisely how metabolism becomes reprogrammed in cancer cells, whose functions or malignant properties are enabled by these activities, and how to exploit metabolic changes for therapeutic benefit are among the key questions driving research in the field.

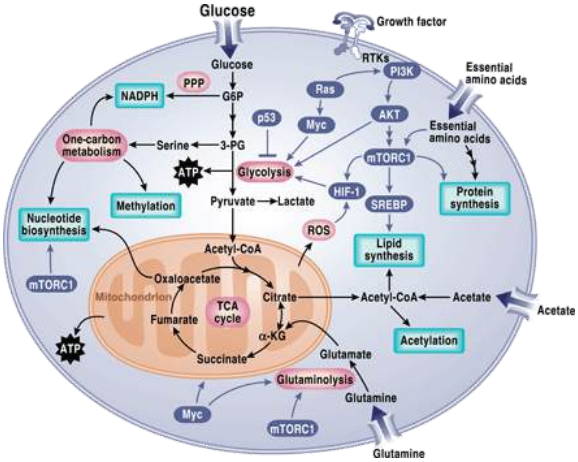
This review covers several fundamental principles in cancer metabolism, with the goal of introducing non-experts to the concepts motivating ongoing research. With the explosion of research in cancer metabolism over the past decade, no single review can possibly cover it all. The sections below highlight some of the essential, recent papers supporting these core principles. An overarching theme in cancer metabolism is that reprogrammed activities improve cellular fitness to provide a selective advantage during tumorigenesis. Most of the classical examples of reprogrammed activities either support cell survival under stressful conditions or allow cells to grow and proliferate at pathologically elevated levels. Three of these—altered bioenergetics, enhanced biosynthesis, and redox balance—are discussed at length below. It logically flows that if these activities provide benefit to the malignant cell, then some of them might be suitable therapeutic targets. This rendering of cancer metabolism is supported by many examples in which inhibition of an enhanced metabolic activity results in impaired growth of experimental tumors (*3*, *4*). In some cases, the particular metabolic liabilities of cancer cells have been translated into effective therapies in human cancer. Asparaginase, an enzyme that converts the amino acid asparagine to aspartic acid and ammonia, is an essential component of treatment for acute lymphoblastic leukemia (ALL) (*5*). Because of their high rates of protein synthesis and poor ability to synthesize asparagine de novo, ALL cells require a constant supply of asparagine from the plasma. This supply is essentially eliminated by systemic administration of asparaginase. Ultimately, effective metabolic therapy will require defining the stage of tumor progression in which each pathway provides its benefit to the cancer cell. Some activities first become essential very early in tumorigenesis as the nascent tumor begins to experience nutrient limitations (*6*). In other cases, altered pathways may be dispensable in primary tumors but essential for metastasis (*7*, *8*). Because new therapeutic targets are nominated from simple experimental models like cultured cells, it will be essential to define their context-specific roles in biologically accurate models of tumor initiation and progression.

29.8.1.1: METABOLIC REPROGRAMMING AND ONCOMETABOLITES IN CANCER

Altered metabolic activity supports anabolic growth during nutrient-replete conditions, catabolism to support cell survival during nutrient limitation, and fortification of redox homeostatic systems to counteract the metabolic effects of oncogene activation, tumor suppressor loss, and other stresses (*9*). Discovery and characterization of reprogrammed activities may provide opportunities to image tumor tissue noninvasively, predict tumor behavior, and prevent tumor progression by blocking essential pathways. It is important to differentiate "metabolic reprogramming" from "oncometabolites," two terms widely used in the recent cancer metabolism literature (*10*). We propose that the term metabolic reprogramming be used to describe conventional metabolic pathways whose activities are enhanced or suppressed in tumor cells relative to benign tissues as a consequence of tumorigenic mutations and/or other factors. Oncometabolite is a relatively new term that refers to metabolites whose abundance increases markedly in tumors. We suggest that this term be reserved for metabolites for which (i) there is a clear mechanism connecting a specific mutation in the tumor to accumulation of the metabolite, and (ii) there is compelling evidence for involvement of the metabolite in the development of malignancy.

The classical example of a reprogrammed metabolic pathway in cancer is the Warburg effect or aerobic glycolysis (*11*). Glycolysis is a physiological response to hypoxia in normal tissues, but Otto Warburg in the 1920s observed that tumor slices and ascites cancer cells constitutively take up glucose and produce lactate regardless of oxygen availability, an observation that has been seen in many types of cancer cells and tumors (*12*). The increase in glycolytic flux allows glycolytic intermediates to supply subsidiary pathways to fulfill the metabolic demands of proliferating cells (*11*). Like glycolytic intermediates, tricarboxylic acid (TCA) cycle intermediates are also used as precursors for macromolecule synthesis (*13*). Their utilization in biosynthetic pathways requires that carbon be resupplied to the cycle so that intermediate pools are maintained; pathways that “refill” the cycle are termed anaplerotic pathways, and they generate TCA cycle intermediates that can enter the cycle at sites other than acetyl-CoA (coenzyme A) (*14*). Two activities that provide anaplerotic fluxes in cancer cells are glutaminolysis, which produces α-ketoglutarate from glutamine, and pyruvate carboxylation, which produces oxaloacetate from glucose/pyruvate. Oxidation of the branched-chain amino acids (BCAAs) isoleucine and valine also provides an anaplerotic flux in some tissues.

Despite the incredible genetic and histological heterogeneity of tumors, malignancy seems to involve the common induction of a finite set of pathways to support core functions like anabolism, catabolism, and redox balance (*15*). The general induction of these pathways may reflect their regulation by signaling pathways that are commonly perturbed in cancer cells (Fig. 1). Normal cells, upon stimulation by growth factors, activate phosphatidylinositol 3-kinase (PI3K) and its downstream pathways AKT and mammalian target of rapamycin (mTOR), thereby promoting a robust anabolic program involving increased glycolytic flux and fatty acid synthesis through activation of hypoxia-inducible factor–1 (HIF-1) and sterol regulatory element–binding protein (SREBP), respectively (*16*). Tumor cells very frequently contain mutations that allow the PI3K-AKT-mTOR network to achieve high levels of signaling with minimal dependence on extrinsic stimulation by growth factors (*17*). Many of the best-characterized oncogenes and tumor suppressors reside in the PI3K-AKT-mTOR network, and aberrant activation of this pathway is among the most frequent alterations seen in a diverse set of cancers.



- **Download high-res image**
- **Open in new tab**
- **Download Powerpoint**

Fig. 1 Signaling pathways that regulate cancer metabolism.

Tumor cells have aberrant activation of mTORC1 that induces an anabolic growth program resulting in nucleotide, protein, and lipid synthesis. Loss of tumor suppressors like p53 or activation of oncogenes like MYC further promotes anabolism through transcriptional regulation of metabolic genes. Metabolism controls signaling through regulating reactive oxygen species (ROS), acetylation, and methylation. PPP, pentose phosphate pathway; G6P, glucose-6-phosphate; 3-PG, 3-phosphoglycerate; ATP, adenosine 5′-triphosphate; mTORC1, mTOR complex 1; α-KG, α-ketoglutarate; RTK, receptor tyrosine kinase.

Another commonly deregulated pathway in cancer is gain of function of MYC by chromosomal translocations, gene amplification, and single-nucleotide polymorphisms. MYC increases the expression of many genes that support anabolic growth, including transporters and enzymes involved in glycolysis, fatty acid synthesis, glutaminolysis, serine metabolism, and mitochondrial metabolism (*18*). Oncogenes like Kras, which is frequently mutated in lung, colon, and pancreatic cancers, co-opt the physiological functions of PI3K and MYC pathways to promote tumorigenicity. Aside from oncogenes, tumor suppressors such as the p53 transcription factor can also regulate metabolism (*19*). The p53 protein–encoding gene *TP53* (tumor protein p53) is mutated or deleted in 50% of all human cancers. The tumor-suppressive functions of p53 have been ascribed to execution of DNA repair, cell cycle arrest, senescence, and apoptosis. However, recent studies indicate that p53 tumor-suppressive actions might be independent of these canonical p53 activities but rather dependent on the regulation of metabolism and oxidative stress (*20*, *21*). Loss of p53 increases glycolytic flux to promote anabolism and redox balance, two key processes that promote tumorigenesis (*19*).

A salient feature of many tumors is that they reside in a low-oxygen environment (hypoxia) ranging from 0 to 2% O_2 because the tumor cell proliferation rate often exceeds the rate of new blood vessel formation (angiogenesis) (*22*). The metabolic adaptation to hypoxia is coordinated by HIF-1, which induces metabolic genes involved in increasing glycolytic flux (*23*). Some tumors display constitutive activation of HIF-1 under normoxic conditions through a variety of mechanisms, including (i) hyperactivation of mTORC1, (ii) loss of von Hippel–Lindau, (iii) accumulation of ROS, and (iv) accumulation of the TCA cycle metabolites succinate or fumarate due to cancer-specific mutations in succinate dehydrogenase (SDH) or fumarate hydratase (FH), respectively (*24*).

The robust coordinated induction of metabolic pathways that support tumorigenesis by combination of deregulation of PI3K-AKT-mTOR signaling pathways, loss of tumor suppressors, and activation of oncogenes alleviates the necessity of having mutations or amplifications in metabolic enzymes per se. Thus, examples of metabolic enzyme deregulation through genetic mutation are rare. One example is the elevated expression of phosphoglycerate dehydrogenase (PHGDH) due to amplification in a fraction of breast cancer and melanoma (*25*, *26*). PHGDH catalyzes the conversion of the glycolytic intermediate 3-phosphoglycerate to 3-phosphohydroxypyruvate in the first step of the serine biosynthesis pathway. Serine metabolism supplies methyl groups to the one-carbon and folate pools contributing to nucleotide synthesis, methylation reactions, and NADPH (reduced nicotinamide adenine dinucleotide phosphate) production (*27*). Inhibiting serine biosynthesis by silencing PHGDH in cells with high levels of this enzyme results in growth suppression, and the enzyme displays oncogenic properties in gain of function assays (*25*, *26*).

The other examples of mutational deregulation of metabolic enzymes are those that generate oncometabolites. The current list of true oncometabolites is short (*28*). The term is most commonly and appropriately applied to D-2-hydroxyglutarate (D2HG), a reduced form of the TCA cycle intermediate α-ketoglutarate. D2HG is scarce in normal tissues but rises to millimolar concentrations in tumors with mutations in isocitrate dehydrogenase 1 or 2 (IDH1 or IDH2). These mutations occur commonly in gliomas, acute myelogenous leukemias (AMLs), and other types of cancer (*29*–*31*). D2HG and its relationship to mutant IDH1 and IDH2 have been reviewed extensively elsewhere (*32*). The most relevant point here is that D2HG production requires a neomorphic enzyme activity imparted to IDH1/IDH2 by specific active-site mutations (*33*, *34*). High levels of D2HG interfere with the function of dioxygenases requiring α-ketoglutarate as a cosubstrate. These include prolyl hydroxylases, cytosine hydroxylases, and histone demethylases, whose inhibition influences gene expression in part via an altered epigenetic state characterized by a failure to express differentiation programs (*35*–*41*). Thus, although D2HG arises from an alteration of the metabolic network, its role in cancer seems to depend on nonmetabolic effects. Currently, D2HG is being used as a biomarker for disease monitoring, and inhibitors specific to mutants IDH1/IDH2 are in clinical trials for AML and solid tumors.

The metabolite 2HG also exists as the enantiomer L-2HG (L2HG), which is not produced by mutant forms of IDH1/IDH2. This metabolite arises from the noncanonical activity of various dehydrogenases, including malate dehydrogenase and lactate dehydrogenase, whose promiscuous behavior reduces α-ketoglutarate to L2HG (*42*–*44*). L2HG may be oxidized back to α-ketoglutarate by a FAD-linked enzyme, L2HG dehydrogenase (L2HGDH). L2HGDH deficiency, also called L2HG aciduria, is a rare neurometabolic disease of childhood caused by the inheritance of biallelic mutations in the gene encoding L2HGDH (*45*). Affected children have seizures, mental retardation, white matter abnormalities of the brain, and systemically elevated levels of L2HG. Remarkably, a number of these children have developed malignant brain tumors (*46*), providing an early clue to the significance of D2HG in IDH1/IDH2-mutant gliomas and raising the question of whether L2HG is also an oncometabolite. L2HG and D2HG exhibit different effects on dioxygenase function (*38*), suggesting that the sensitivity of a particular tissue to the presence of either metabolite may depend on which dioxygenases are expressed. Recent work has demonstrated modest accumulation of L2HG in cells experiencing hypoxia or electron transport chain (ETC) dysfunction (*42*, *47*) and in human renal cell carcinomas, which frequently display epigenetic silencing of L2HGDH (*48*). It is unknown whether reducing L2HG levels in these settings will promote cellular differentiation or suppress tumor progression.

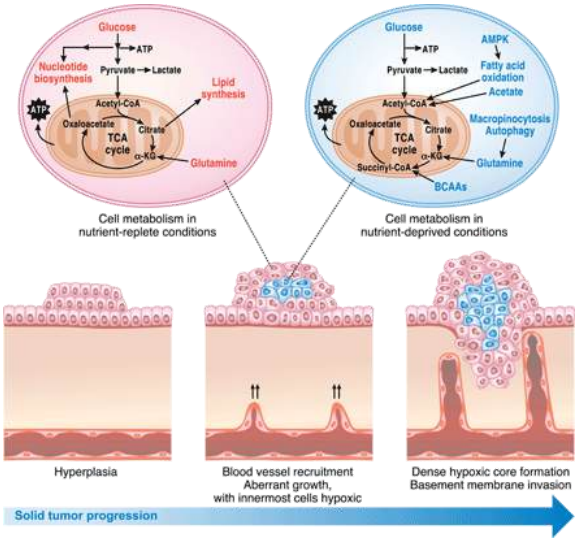
The principle that oncometabolites exert their effects outside of the conventional metabolic network also applies to the other two molecules that can reasonably be called oncometabolites: fumarate and succinate (*28*). Both are TCA cycle intermediates found throughout the body, but some tumors accumulate massive levels of fumarate and/or succinate as a consequence of loss-of-function mutations in FH or the SDH complex, respectively (*49*–*51*). Although these mutations markedly reprogram metabolism by impairing TCA cycle flux, the extent to which fumarate and succinate participate in cancer development likely involves their nonmetabolic functions (*28*). Like D2HG, evidence indicates that succinate and fumarate interfere with dioxygenase activity, supporting the notion that a general property of oncometabolites is the ability to regulate epigenetics (*52*, *53*). PHGDH overexpression and mutations in IDH1/IDH2, SDH, and FH all alter metabolite levels that control epigenetics (*54*). Several other metabolites, including acetyl-CoA, α-ketoglutarate, and *S*-adenosylmethionine also participate in epigenetic reprogramming, and time will tell whether genetically specific alterations of these metabolites in tumors promote tumorigenesis (*54*). Some metabolites, notably fumarate, covalently bind to sulfhydryl groups in glutathione, proteins, and peptides, altering their function and perhaps accounting for another mechanism by which oncometabolites promote or perpetuate malignant phenotypes (*55*–*58*).

29.8.1.1: BIOENERGETICS

Otto Warburg's hypothesis that cancer cells take up glucose and generate a substantial amount of lactate in the presence of ambient oxygen due to impaired mitochondrial function led to the widely held misconception that cancer cells rely on glycolysis as their major source of ATP (*59*, *60*). Today, it is clear that cancer cells exhibit aerobic glycolysis due to activation of oncogenes, loss of tumor suppressors, and up-

regulation of the PI3K pathway, and that one advantage of high glycolytic rates is the availability of precursors for anabolic pathways (2, 61). Warburg's observation that tumors display a high rate of glucose consumption has been validated in many human cancers with fluorodeoxyglucose positron emission tomography, which uses a radioactive glucose analog to image glucose uptake in tumors and adjacent normal tissue. Nevertheless, many studies have demonstrated that the great majority of tumor cells have the capacity to produce energy through glucose oxidation (that is, the process by which glucose-derived carbons enter the TCA cycle and are oxidized to CO_2, producing ATP through oxidative phosphorylation). Furthermore, limiting glycolytic ATP production by inhibiting the activity of pyruvate kinase fails to prevent tumorigenesis, suggesting that the major role of glycolysis is not to supply ATP (62). Moreover, mitochondrial metabolism is necessary for cancer cell proliferation and tumorigenesis (63–65). Thus, despite their high glycolytic rates, most cancer cells generate the majority of ATP through mitochondrial function, with the likely exception of tumors bearing mutations in enzymes involved in mitochondrial respiration (for example, SDH and FH) (66). Nevertheless, cells harboring mutations in FH or SDH still rely on mitochondrial metabolism by metabolically rewiring themselves to provide the necessary TCA cycle intermediates and ROS for cell proliferation (55, 67–70).

In addition to pyruvate derived from glycolysis, fatty acids and amino acids can supply substrates to the TCA cycle to sustain mitochondrial ATP production in cancer cells (Fig. 2). The breakdown of fatty acids (β-oxidation) in the mitochondria generates acetyl-CoA and the reducing equivalents NADH and $FADH_2$, which are used by the ETC to produce mitochondrial ATP. The amino acid glutamine can generate glutamate and subsequently α-ketoglutarate to fuel the TCA cycle through a series of biochemical reactions termed glutaminolysis (71). Furthermore, the BCAAs isoleucine, valine, and leucine, which are elevated in plasma of patients with pancreatic cancers, can be converted into acetyl-CoA and other organic molecules that also enter the TCA cycle (72). The metabolic flexibility afforded by multiple inputs into the TCA cycle allows cancer cells to adequately respond to the fuels available in the changing microenvironment during the evolution of the tumor (9). A combination of the local tumor microenvironment and oncogenic lesions is likely to dictate the fuel used by mitochondria to sustain tumor growth.



- Download high-res image
- Open in new tab
- Download Powerpoint

Fig. 2 Metabolic pathways under nutrient-replete and nutrient-deprived conditions.

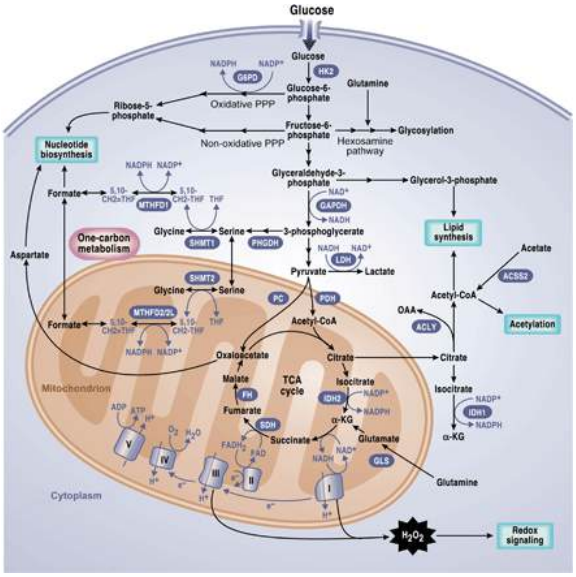
Accessibility to nutrients within solid tumors is regulated by proximity to the vasculature. Cells located adjacent to the vasculature use nutrients and oxygen to fuel anabolic pathways that support proliferation. However, cells distant from the vasculature have diminished accessibility to nutrients and oxygen and may engage in alternative forms of metabolism including oxidation of fatty acids and BCAAs as well as macromolecular degradation through autophagy and macropinocytosis to support cell viability.

Solid tumors contain significant heterogeneity of perfusion, such that many tumor cells reside in nutrient- and oxygen-poor environments. Cancer cells have therefore adapted multiple mechanisms to sustain mitochondrial function for survival. For example, the mitochondrial ETC can function optimally at oxygen levels as low as 0.5% (73). Moreover, hypoxic tumor cells (<2% O_2) can continue to cycle and use glutamine as a fuel for oxidative ATP production (74–76). Kras-driven pancreatic cancer cells in nutrient-depleted conditions use proteins scavenged from the extracellular space to produce glutamine and other amino acids to fuel the TCA cycle for cell survival and growth (Fig. 2) (77). Furthermore, if pyruvate oxidation to acetyl-CoA is compromised by hypoxia or ETC impairment, glutamine can provide acetyl-CoA as a biosynthetic precursor to sustain tumor growth (69, 78, 79).

When tumor cells become nutrient-deprived, they adapt to the microenvironment by decreasing their demand for ATP. The resultant increase in ATP availability maintains an adequate ATP/ADP (adenosine 5´-diphosphate) ratio to drive unfavorable biological reactions. The anabolic kinase mTOR, discussed in greater detail below, drives the energetically demanding growth of tumor cells. This kinase is inhibited when amino acids and oxygen levels are diminished (*80*). Furthermore, decreased mTOR activity increases autophagic flux. In oncogenic Kras- or Braf-driven non–small-cell lung cancer (NSCLC) cells, autophagy provides an intracellular glutamine supply to sustain mitochondrial function (*81*, *82*). To survive the hypoxic tumor microenvironment, cancer cells also diminish their demand for ATP by decreasing highly demanding ATP-dependent processes, such as running the Na/K-dependent adenosine triphosphatase. If diminishing ATP demand is not sufficient to maintain ATP/ADP ratio, the rise in ADP activates adenylate kinase, a phosphotransferase enzyme that buffers the fall in ATP levels by converting two ADP molecules into adenosine 5´-monophosphate (AMP) and ATP (*83*). The rise in AMP during nutrient deprivation triggers the activation of AMP kinase (AMPK), which activates catabolic pathways like fatty acid oxidation to stimulate ATP production (*84*). In conditions of metabolic stress, certain Ras-driven cancer cells scavenge lipids to support ATP production (*85*). Ovarian cancer cells use fatty acids from neighboring adipocytes to fuel mitochondrial ATP production (*86*). Thus, there are multiple mechanisms by which cancer cells maintain their ATP/ADP ratio to sustain viability in nutrient- and oxygen-poor environments.

29.8.1.1: BIOSYNTHESIS OF MACROMOLECULES

Biosynthetic or anabolic pathways are an essential aspect of cancer metabolism because they enable cells to produce the macromolecules required for replicative cell division and tumor growth. As a general theme, these pathways involve the acquisition of simple nutrients (sugars, essential amino acids, etc.) from the extracellular space, followed by their conversion into biosynthetic intermediates through core metabolic pathways like glycolysis, the PPP, the TCA cycle, and nonessential amino acid synthesis, and finally the assembly of larger and more complex molecules through ATP-dependent processes (Fig. 3). The three macromolecular classes most commonly studied in cancer metabolism are proteins, lipids, and nucleic acids, which comprise approximately 60, 15, and 5% of the dry mass of mammalian cells, respectively. Evidence indicates that biosynthesis of all three classes is under the control of the same signaling pathways that govern cell growth and are activated in cancer via tumorigenic mutations, particularly PI3K-mTOR signaling.



- **Download high-res image**
- **Open in new tab**
- **Download Powerpoint**

Fig. 3 Anabolic pathways that promote growth.

Glucose metabolism generates glycolytic intermediates that can supply subsidiary pathways including the hexosamine pathway, PPP, and one-carbon metabolism, all of which support cell growth. Mitochondrial TCA cycle intermediates such as oxaloacetate (OAA) and citrate are used to generate cytosolic aspartate and acetyl-CoA for nucleotide and lipid synthesis, respectively. Mitochondria also generate H_2O_2 and acetyl-CoA for redox signaling and acetylation, respectively. NADPH is used to drive anabolic reactions and to maintain antioxidant capacity. Cytosolic sources of NADPH include the oxidative PPP, IDH1, and enzymes from one-carbon metabolism including MTHFD1. Mitochondrial sources of NADPH include MTHFD2, MTHF2L, and IDH2. HK2, hexokinase 2; G6PDH, glucose-6-phosphate dehydrogenase; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; LDH, lactate dehydrogenase; ACLY, ATP citrate lyase; GLS, glutaminase; SHMT, serine hydroxymethyltransferase; MTHFD2, methylenetetrahydrofolate dehydrogenase 2; MTHFD2L, MTHFD2-like; ACSS2, acyl-CoA synthetase short-chain family member 2; THF, tetrahydrofolate.

Protein biosynthesis is highly regulated and requires access to a full complement of essential and nonessential amino acids. Cancer cells and other cells under the influence of growth factor signaling express surface transporters that allow them to acquire amino acids from the extracellular space (*87*). This not only provides cells with the raw materials needed for protein synthesis but also allows them to maintain activity of the mTOR signaling system, specifically mTORC1. mTORC1 is stimulated by the presence of amino acids and activates protein synthesis via its effects on translation and ribosome biogenesis (*80*). Most nonessential amino acids are produced through transamination reactions, which transfer the amino group from glutamate to a ketoacid. Proliferating cancer cells take up glutamine and convert it to glutamate through a variety of deamidation and transamidation reactions, most notably the mitochondrial amidohydrolase glutaminase (*71*). Together, these enzymes generate a large intracellular glutamate pool available for nonessential amino acid synthesis. Both glutamine uptake and glutaminase activity are stimulated by mTORC1, providing glutamate for transamination reactions and/or maintenance of the TCA cycle, which also contributes to amino acid synthesis. Furthermore, when the intracellular glutamine supply exceeds the cell's demands, glutamine can be exported in exchange for essential amino acids to stimulate mTORC1 and protein synthesis (*88*). Thus, growth conditions in which glutamine and essential amino acids are abundant enable mTORC1-mediated activation of protein synthesis.

When nutrients are scarce, cells have access to a number of catabolic pathways to degrade macromolecules and resupply key pools of intracellular metabolic intermediates. Protein degradation pathways have been characterized extensively as mechanisms to supply amino acids in cancer cells. Intracellular proteins and other macromolecules can be recycled through autophagy, a highly regulated process through which proteins and organelles are delivered to the lysosome and degraded (*89*). Autophagy is an essential survival pathway during nutrient or growth factor deprivation, and genetic studies demonstrate that it contributes to some forms of cancer in mice (*90*, *91*). However, because autophagy supplies amino acids through protein degradation, it does not serve as a source of net protein synthesis. It is also potently suppressed by mTORC1. Macropinocytosis allows cells to internalize proteins and other components of the extracellular milieu and deliver them for degradation via the endocytic pathway. Under conditions of nutrient depletion, macropinocytosis supplies both nitrogen and carbon to central metabolic pathways (*92*). Evidence indicates that extracellular protein degradation, like autophagy, is suppressed by mTORC1 (*93*). Suppressing pathways of protein degradation may help maximize rates of net protein synthesis when extracellular amino acids are available and mTORC1 is active.

Tumor cells rapidly produce fatty acids for membrane biosynthesis, lipidation reactions, and cellular signaling. Fatty acid synthesis requires sources of acetyl-CoA and reducing power in the form of cytosolic NADPH; effective fatty acid synthesis therefore requires integration with other pathways of carbon and redox metabolism. In most cultured cells, glucose is the most prominent acetyl-CoA source for fatty acid synthesis (*94*, *95*). Glutamine and acetate have been demonstrated to provide alternative carbon sources when access to glucose-derived acetyl-CoA is impaired by hypoxia or mitochondrial dysfunction (*69*, *78*, *79*, *96*). Leucine degradation can also supply acetyl-CoA in some cell lines (*97*). The relative importance of these nutrients for fatty acid synthesis in vivo is unknown, although early studies suggested that most fatty acyl carbon in experimental tumors is derived from glucose (*98*, *99*). Isotopic tracing experiments designed to assess the cytosolic NADPH pool have recently suggested that most NADPH used for fatty acid synthesis arises from the PPP (*100*, *101*).

Transcription of genes involved in fatty acid synthesis is regulated by the SREBP-1 transcription factor (*102*). SREBP-1 regulates not only the enzymes needed to convert acetyl-CoA into fatty acids but also the enzymes of the PPP and pathways required to convert acetate and glutamine into acetyl-CoA (*103*). This transcription factor therefore regulates genes encoding proteins that catalyze or facilitate fatty acid synthesis. In lipid-replete conditions, SREBP-1's transcriptional activity is suppressed by its sequestration in the endoplasmic reticulum. Under conditions of sterol depletion, proteolytic cleavage releases the transcriptionally active domain, which travels to the nucleus and binds to sterol response elements in the promoters of lipogenic genes (*104*).

In cancer cells with constitutively high rates of fatty acid synthesis, additional mechanisms help keep SREBP-1 in a transcriptionally active state. mTORC1 signaling, via its effector S6 kinase (S6K), activates a transcriptional program that includes both SREBP-1 and the related protein SREBP-2, which regulates transcription of genes in sterol biosynthesis (*103*). Both SREBP-1 and SREBP-2 are required for mTORC1-mediated cell proliferation. The mechanism of SREBP activation by mTORC1 is incompletely understood but involves nuclear entry of the phosphatidic acid phosphatase Lipin-1, which enhances nuclear SREBP abundance and activity on the promoters of lipogenic genes (*105*).

Both fatty acids and lipids can also be acquired from the extracellular space to supply membrane biosynthesis. PI3K signaling activates fatty acid uptake and suppresses fatty acid oxidation, thereby maximizing lipogenesis in proliferating cells under the control of growth factors (*106*). Lipid uptake may acquire further importance during conditions of metabolic stress, when the ability to meet oncogene-driven demands for biosynthesis is compromised. The ability to scavenge lysophospholipids (lipid intermediates containing a glycerophosphate backbone with one acyl chain) is required for maximal cancer cell growth during hypoxia, which suppresses de novo fatty acid synthesis from glucose (*85*). Furthermore, in cancer cells with constitutive mTORC1 signaling, hypoxia induces a state of dependence on access to extracellular desaturated fatty acids to maintain endoplasmic reticulum integrity in support of protein biosynthesis (*107*). Notably, SREBP-1 was first identified as the transcription factor responsible for expression of the low-density lipoprotein receptor (LDLR) (*108*), implying that enhanced de novo lipogenesis occurs concomitantly with enhanced import of lipids from the extracellular space. These parallel pathways appear to be essential in glioma, where oncogenic activation of epidermal growth factor receptor (EGFR) signaling stimulates SREBP-1 and expression of LDLR (*109*). These cancer cells are highly sensitive to inhibitors of fatty acid and cholesterol biosynthesis. Inhibition of the EGFR-PI3K signaling axis but not of mTORC1 suppresses nuclear translocation of SREBP-1 in glioma cells with oncogenic EGFR,

suggesting an alternate, mTORC1-independent mode of SREBP-1 activation in glioma cells (*109*). This transcriptional program includes LDLR expression and induces reliance on cholesterol uptake to maintain the intracellular pool (*110*). Impairing intracellular cholesterol availability by activating liver X receptor induced cell death both in culture and in vivo, suggesting a pharmacological approach to silence lipogenic programs in glioma (*110*).

Purine and pyrimidine nucleotides are required for synthesis of RNA and DNA. De novo biosynthesis of nucleotides is complex, requiring input from several pathways in a coordinated fashion. The phosphoribosylamine backbone of these molecules is produced from ribose-5-phosphate, an intermediate of the PPP, and an amide donation reaction using glutamine as a substrate (*111*). The purine and pyrimidine bases are constructed from various nonessential amino acids and methyl groups donated from the one-carbon/folate pool. The TCA cycle contributes oxaloacetate, which is transaminated to aspartate, an intermediate required to synthesize both purine and pyrimidine bases. Conversion of ribonucleotides to deoxynucleotides by ribonucleotide reductase requires a source of NADPH. Well-characterized mechanisms of feedback inhibition exist to prevent excessive accumulation of nucleotides, and mutations interrupting these mechanisms can produce pathological accumulation of nucleotide intermediates (for example, precipitation of uric acid crystals in gout).

Clearly, nucleotide biosynthesis is a targetable vulnerability in cancer cells because nucleoside analogs and antifolates have been a mainstay of chemotherapeutic regimens for decades (*112*). However, relatively little is known about how oncogenic signaling interfaces with nucleotide biosynthesis. It is likely that the effects of numerous signaling pathways on glucose and amino acid metabolism influence the availability of precursors for purines and pyrimidines. In the case of mTORC1, evidence points to a distinct mechanism by which activation of the signaling pathway enables nucleotide biosynthesis. The mTORC1 effector ribosomal S6K1 phosphorylates the trifunctional enzyme CAD (carbamoyl-phosphate synthetase 2, aspartate transcarbamoylase, dihydroorotase), which catalyzes the first three steps of pyrimidine synthesis (*113*). Phosphorylation on CAD S1859 is required for mTORC1-dependent stimulation of pyrimidine biosynthesis (*113*). Additional work is needed to determine how other aspects of de novo nucleotide synthesis, purine and pyrimidine salvage pathways, and accessory activities like folate metabolism are regulated in cancer cells in support of cell proliferation.

29.8.1.1: REDOX BALANCE

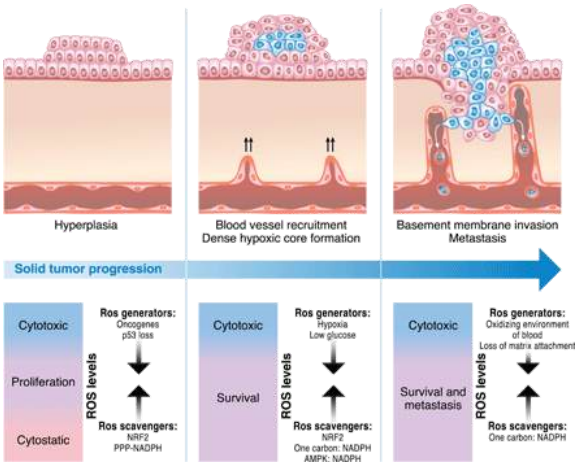
ROS are intracellular chemical species that contain oxygen and include the superoxide anion (O_2^-), hydrogen peroxide (H_2O_2), and the hydroxyl radical (OH·) (*114*). The mitochondria and cytosolic NADPH oxidases (NOXs) produce O_2^- from the one-electron reduction of oxygen (*115*). O_2^- is converted into H_2O_2 by the enzymatic activity of superoxide dismutase 1 or 2, which are localized to the cytosol or mitochondrial matrix, respectively. H_2O_2 is subsequently detoxified to water by the enzymatic activity of mitochondrial and cytosolic peroxiredoxins (PRXs), which, as a consequence, undergo H_2O_2-mediated oxidation of their active-site cysteines (*116*). Thioredoxin (TXN), thioredoxin reductase (TrxR), and the reducing equivalent NADPH reduce oxidized PRXs to complete the catalytic cycle (*117*). Glutathione peroxidases (GPXs) can also convert H_2O_2 to water in the mitochondrial matrix and cytosol through H_2O_2-mediated oxidation of reduced glutathione (GSH) (*118*). Glutathione reductase (GR) and NADPH reduce oxidized glutathione (GSSG) back to GSH. Additionally, catalase, an abundant antioxidant in peroxisomes, can detoxify H_2O_2 to water without any cofactors. However, in the presence of ferrous or cuprous ions, H_2O_2 can become OH· and quickly cause the oxidation of lipids, proteins, and DNA, resulting in cellular damage. NADPH is required to maintain multiple antioxidant defense systems. The cytosol has multiple sources of NADPH generation, including the oxidative PPP, malic enzyme 1, IDH1, and one-carbon metabolism. NADPH generation in the mitochondria, in part, is controlled by one-carbon metabolism and IDH2.

Historically, ROS have been thought of as lethal metabolic by-products of cellular respiration and protein folding. However, studies over the past two decades have unveiled a previously underappreciated role of ROS in cellular signaling. Low levels of ROS, particularly H_2O_2, can reversibly oxidize the cysteine residues of proteins to positively regulate cell proliferation and cellular adaptation to metabolic stress (*119*). As H_2O_2 levels increase, however, cell death signaling pathways are initiated, and H_2O_2 is converted to OH·, which can directly damage DNA, proteins, and lipids. Cancer cells have an increased rate of spatially localized mitochondria- and NOX-dependent ROS production compared to normal cells. This allows for the proximal activation of signaling pathways [PI3K and mitogen-activated protein kinase/extracellular signal–regulated kinase (MAPK/ERK)] and transcription factors [HIF and nuclear factor κB (NF-κB)] necessary for tumorigenesis. The cancer cell–specific increased rate of spatially localized ROS production is due to a combination of oncogenic lesions and the tumor microenvironment. For example, the activation of oncogenes, PI3K signaling pathway induction, and hypoxia (low-oxygen levels) stimulate the increased rate of ROS production from the mitochondria and NOXs in cancer cells (*120*–*122*). Thus, mitochondria-targeted antioxidants and NOX inhibitors can prevent cancer cell proliferation, hypoxic activation of HIF, tumorigenesis, and metastasis (*64*, *123*–*125*).

The increased localized ROS in cancer cells, which activates signaling pathways and transcription factors to promote tumorigenesis, needs to be buffered from reaching levels of ROS that incur cellular damage by the increased expression of antioxidant proteins (*126*). Thus, cancer cells have higher levels of ROS scavenging enzymes than normal cells, preventing ROS-mediated activation of death-inducing pathways like c-Jun N-terminal kinase (JNK) and p38 MAPK and oxidation of lipids, proteins, and DNA, resulting in irreversible damage and cell death. One mechanism by which cancer cells increase their antioxidant capacity is by activating the transcription factor nuclear factor (erythroid-derived 2)–related factor-2 (NRF2) (*127*). Specifically, NRF2 is activated following disruption of the interaction of NRF2 with its binding partner Kelch-like ECH-associated protein 1 (KEAP1). Critical cysteine residues within KEAP1 can undergo oxidation,

succination, and glutathionylation, thereby inhibiting the KEAP1-NRF2 interaction, leading to the proteasomal degradation of NRF2. Additionally, NRF2 activation can occur independently of KEAP1 (*128*). Once activated, NRF2 induces the transcription of many antioxidant proteins including GPXs and TXNs as well as enzymes involved in GSH synthesis and cysteine import through the cystine/glutamate antiporter. Furthermore, to maintain the antioxidant capacity of GPXs and TXNs, NADPH is required. NRF2 plays an important role in activating enzymes that increase cytosolic NADPH levels. NRF2 also regulates the serine biosynthesis pathway, generating NADPH in the mitochondria, which is critical for redox balance under hypoxic conditions (*129*, *130*). Therefore, inactivating NRF2 or disabling antioxidant proteins in cancer cells would allow for the accumulation of excessive amounts of ROS to levels that initiate toxicity and reduce tumorigenesis (*128*, *131*, *132*).

During tumorigenesis and metastasis, redox homeostasis is required (Fig. 4). An emerging model of redox balance is that as a tumor initiates, the metabolic activity of cancer cells is increased, resulting in an increase in ROS production and subsequent activation of signaling pathways that support cancer cell proliferation, survival, and metabolic adaptation (*126*). Accordingly, to prevent toxic levels of ROS, tumor cells increase their antioxidant capacity to allow for cancer progression (*133*). The harsh tumor microenvironment increases ROS levels due to hypoxia, and the low glucose levels limit flux through the cytosolic oxidative PPP, thus decreasing cytosolic NADPH levels. Cells in these nutrient-deprived conditions activate AMPK to increase NADPH levels by stimulating PPP-dependent NADPH and diminishing anabolic pathways, such as lipid synthesis, that require high levels of NADPH (*134*, *135*). ROS-dependent signaling and increased mitochondrial respiration are also necessary for tumor metastasis (*124*, *136*). However, when tumor cells detach from a matrix, they encounter high levels of ROS that incur cellular damage and require activation of adaptive ROS-mitigating pathways to survive and grow (*137*, *138*). The ability to up-regulate antioxidant proteins and increase flux through NADPH-producing metabolic pathways enables distant metastasis to occur (*8*). These findings suggest that perhaps disabling antioxidant capacity in cancer cells to raise ROS levels might be beneficial in preventing metastasis.



- Download high-res image
- Open in new tab
- Download Powerpoint

Fig. 4 Cancer cells maintain redox balance.

Cancer cells have increased rates of ROS production due to activation of oncogenes and loss of tumor suppressors that promote signaling pathways supporting proliferation and survival. However, cancer cells prevent the buildup of ROS to levels that incur damage by increasing antioxidant capacity through induction of NRF2-dependent genes and, in glucose replete conditions, the use of PPP to generate NADPH. As cells encounter hypoxia and low glucose due to limited vasculature accessibility, the levels of ROS further increase, requiring AMPK and one-carbon metabolism to enhance NADPH production to raise antioxidant capacity. Loss of matrix attachment and escape of cancer cells into the blood for dissemination to distant sites incur further increases in ROS levels, which require additional enhancements of antioxidant defenses to avoid cell death. It is important to note that too little ROS or too high steady-state ROS levels within cancer cells result in failure for solid tumor progression and metastasis.

29.8.1.1: TARGETING METABOLISM FOR CANCER THERAPY

There are a few things to consider when determining what makes a good metabolic target for cancer therapy. First, inhibition of some metabolic enzymes is likely to be systemically toxic because of their physiological functions in normal tissues (*139*). The feasibility of targeting these pathways therapeutically depends on whether systemic blockade of the pathway can be tolerated. Normal proliferating cells, such as immune cells and stem cells, also reprogram their metabolism in a manner similar to cancer cells (*140*, *141*). Metabolic inhibitors should likely not interfere with the adaptive immune system. Nevertheless, there are excellent examples of pathways whose reprogramming does provide an adequate therapeutic window in cancer. Enhanced nucleotide and DNA synthesis in tumor cells is targeted by antifolates

(methotrexate, pemetrexed, and others) (112). Although these drugs do produce toxicity in normal proliferative tissues like the intestinal epithelium and bone marrow, they are essential components of highly successful chemotherapeutic regimens. Thus, it is critical to elucidate in normal cells any toxic effects of metabolic enzyme inhibition. To circumvent this challenge, one approach is to target a metabolic enzyme in a deregulated pathway specific to cancer cells. To date, many of the genetic and pharmacologic interventions on metabolic enzymes have been conducted using human cancer cells subcutaneously injected into athymic mice. Therefore, it will be important for future experiments to not only use patient-derived xenograft (PDX) models but also make use of genetically engineered mouse cancer models and syngeneic mouse models that have intact immune systems, especially given the promising results from immunotherapy. An emerging theme is that cancer cells display metabolic plasticity and can alter their metabolic profile during the course of tumorigenesis and metastasis. Thus, it is conceivable that cancer cells could develop resistance to inhibition of a particular metabolic pathway by expressing alternate protein isoforms or up-regulating compensatory pathways. Therefore, a rational cancer therapeutic strategy should involve targeting multiple metabolic pathways simultaneously or targeting a particular metabolic pathway in combination with therapies against oncogenic or signaling pathways. Here, we highlight a few promising metabolic targets in glycolytic, one-carbon, mitochondrial, and redox metabolism.

Glycolysis was an early attractive target for cancer therapy given the clinical observation that many tumors exhibit a significant increase in glucose uptake compared with adjacent normal tissue (112). LDH-A, a metabolic enzyme that converts pyruvate (the final product of glycolysis) to lactate, was identified as the first metabolic target of the oncogene MYC (142). Genetic or pharmacologic inhibition of LDH-A has been shown to diminish MYC-driven tumors in xenograft models (143, 144). Furthermore, recent studies indicate that inhibition of LDH-A leads to the regression of established tumors in genetically engineered mouse models of NSCLC without systemic toxicity (145). Genetic ablation of LDH-A also delays the progression of myeloid leukemia (146). Thus, the increased expression of LDH-A, specifically in MYC-mutant cancer cells, may prove to be an attractive target. Another potential therapeutic target is the glycolytic protein HK2. Many tumor cells overexpress HK2, and preclinical mouse models of genetically engineered NSCLC and breast cancer demonstrate that HK2 inhibition delays tumor progression (3). Furthermore, systemic HK2 deletion in mice does not cause adverse physiological consequences. However, the effect of LDH-A and HK2 on the adaptive immune system is currently unknown. Lactate has been shown to inhibit cytotoxic T cells; thus, LDH-A inhibition may cooperate with immune checkpoint inhibitors to unleash host inflammatory T cells that will specifically attack tumor cells (147). Lactate can also reprogram macrophages to promote tumorigenesis (148). Thus, it may be efficacious to target LDH-A or HK2 in highly glycolytic tumors that overexpress these proteins.

Another potential glucose-dependent target is PHGDH, an enzyme in the de novo serine synthesis pathway. High levels of PHGDH have been found in a subset of human melanoma and breast cancers, and these cancer cells require PHGDH for their growth in vitro (25, 26). Serine starvation in mice diminishes tumorigenicity of p53-null cancers (149). De novo synthesis or exogenous uptake of serine can enter the mitochondria where SHMT2 converts it into glycine to generate folate intermediates (101, 150). In many cancer types, SHMT2 expression is elevated and correlates with a poor prognosis. Furthermore, the transcription factors MYC and HIF induce SHMT2 under hypoxia to promote survival (130, 151). Currently, it is not known whether targeting PHGDH, SHMT2, or other enzymes in the one-carbon metabolism pathway would be effective in delaying or regressing tumor progression in genetically engineered, PDX, or syngeneic mouse models of cancer without incurring systemic toxicity. However, given the importance of one-carbon metabolism in supporting the anabolic needs of cancer cells and its up-regulation in cancer cells, it is likely that this pathway is needed for in vivo tumor progression (152).

Mitochondrial metabolism has also emerged as a key target for cancer therapy, in part, due to the revelation that the antidiabetic drug metformin is an anticancer agent (153). Numerous epidemiological studies first suggested that diabetic patients taking metformin, to control their blood glucose levels, were less likely to develop cancer and had an improved survival rate if cancer was already present (154). Laboratory-based studies have also provided evidence that metformin may serve as an anticancer agent (155–157). Biochemists recognized that metformin reversibly inhibits mitochondrial complex I (158–160). Recent studies indicate that metformin acts as an anticancer agent by inhibiting mitochondrial ETC complex I (161). Specifically, metformin inhibits mitochondrial ATP production, inducing cancer cell death when glycolytic ATP levels diminish as a result of limited glucose availability. Metformin also inhibits the biosynthetic capacity of the mitochondria to generate macromolecules (lipids, amino acids, and nucleotides) within cancer cells (162). The remarkable safety profile of metformin is due to its uptake by organic cation transporters (OCTs), which are only present in a few tissues, such as the liver and kidney (163). Certain tumor cells also express OCTs to allow the uptake of metformin (164). However, in the absence of OCTs, tumors would not accumulate metformin to inhibit mitochondrial complex I. Ongoing clinical trials using metformin as an anticancer agent should assess the expression levels of OCTs to identify the tumors with highest expression, which are likely to be susceptible to metformin. Moreover, it is not clear whether the current antidiabetic dosing of metformin used in clinical trials allows for metformin accumulation to levels necessary to inhibit mitochondrial complex I in tumors. Thus, it is possible that metformin at doses higher than those currently used for diabetes might be more efficacious without causing toxicity. Like metformin, the biguanide phenformin exhibits anticancer properties by inhibiting mitochondrial complex I (165). In contrast to metformin, phenformin is readily transported into tumor cells and has been withdrawn from human use in most parts of the world due to its clinical increase in the incidence of lactic acidosis. However, it is worth considering phenformin as a possible cancer therapy because lactic acidosis can be monitored. Biguanide sensitivity can be improved in mice starved for serine or in tumors that have lost p53 or LKB1 (155, 166, 167). Thus, biguanides, and possibly other mitochondrial complex I inhibitors, may be effective anticancer agents.

Another potential therapeutic strategy to inhibit mitochondrial metabolism in certain tumors would be to use autophagy or glutaminase inhibitors. Autophagy provides amino acids, such as glutamine, that fuel the TCA cycle in NSCLC and pancreatic cancers, and short-term autophagy inhibition has been shown to decrease tumor progression without incurring systemic toxicity in mouse models of NSCLC (*168*, *169*). Some tumors are addicted to using glutamine to support TCA cycle metabolism even in the absence of autophagy; thus, glutaminase inhibitors can reduce tumor burden in these models (*4*, *75*, *170*). An alternative approach is to target acetate metabolism. Although a major function of the mitochondria is to provide acetyl-CoA to the cell, cancer cells can also use acetate to support cell growth and survival during metabolic stress (hypoxia or nutrient deprivation) (*96*, *171*). The cytosolic enzyme acetyl-CoA synthase 2 (ACCS2), which converts acetate to acetyl-CoA, is dispensable for normal development; thus, ACCS2 is a promising target of acetate metabolism. ACCS2 knockout mice do not display overt pathologies, but genetic loss of ACCS2 reduces tumor burden in models of hepatocellular carcinoma (*171*). Human glioblastomas can oxidize acetate and may be sensitive to inhibitors of this process (*172*). Thus, targeting metabolism with inhibitors of autophagy, acetate metabolism, and other pathways that supply key metabolic intermediates may be efficacious in some contexts.

Because mitochondrial inhibitors are unlikely to be effective cancer therapies as single agents, combination therapy is likely the best approach. For example, the use of metformin with the current clinical PI3K inhibitors, which reduce glucose uptake and glycolysis (*173*), is one approach that would impair both sources of ATP within cells. Targeted therapies against oncogenes such as KRAS, BRAF, and NOTCH1 kill a large majority of cancer cells but ultimately yield resistant cells that exhibit an increased sensitivity to inhibitors that impair mitochondrial metabolism (*174*–*176*). Cancer-initiating cells also have increased sensitivity to mitochondrial inhibitors, adding further evidence that inhibiting mitochondrial metabolism may suppress tumor recurrence (*177*, *178*).

Furthermore, to counterbalance the increased production of ROS encountered during tumorigenesis and metastasis, cancer cells increase their antioxidant capacity (*126*). Thus, an additional therapeutic approach is to target redox metabolism, that is, selectively disable the antioxidant capacity of cancer cells causing ROS levels to rise and induce cancer cell death (*133*). The reducing equivalent NADPH is required to maintain multiple antioxidant defense systems. The cytosol has multiple sources of NADPH generation, including the oxidative PPP, malic enzyme 1, IDH1, and one-carbon metabolism. By contrast, NADPH generation in the mitochondria is controlled in part by one-carbon metabolism and IDH2. Many of these NADPH-generating systems are critical for normal cell survival and function. Nevertheless, there are two NADPH-generating systems that may serve as potential therapeutic targets. It is estimated that 400 million people worldwide are deficient in G6PDH, an enzyme in the oxidative PPP that converts $NADP^+$ to NADPH. However, certain tumors rely on this pathway as a major source of cytosolic NADPH; therefore, it may be therapeutic to disable this pathway and induce oxidative stress and diminish tumor growth. Moreover, RNA profiling of metabolic enzymes identified the mitochondrial one-carbon metabolism protein MTHFD2, which can generate NADPH, as being highly expressed in 19 different cancer types but not in normal adult proliferating cells (*152*). Loss of MTHFD2 in cancer cells increases ROS levels and sensitizes the cells to oxidant-induced cell death in vitro. An interesting approach to depleting NADPH levels and increasing ROS is to administer high doses of vitamin C (ascorbate). Vitamin C is imported into cells through sodium-dependent vitamin C transporters, whereas the oxidized form of vitamin C, dehydroascorbate (DHA), is imported into cells through glucose transporters such as GLUT1 (*179*). When the cell takes up DHA, it is reduced back to vitamin C by glutathione (GSH), which consequently becomes GSSG. Subsequently, GSSG is converted back to GSH by NADPH-dependent GR. Because the blood is an oxidizing environment, vitamin C becomes oxidized to DHA before being taken up by the cell. Thus, high doses of vitamin C diminish the tumorigenesis of colorectal tumors that harbor oncogenic KRAS mutations and express high levels of GLUT1 by depleting the NADPH and GSH pools and consequently increasing ROS levels to induce cancer cell death (*179*, *180*). Vitamin C administered at high doses intravenously is safe in humans and, in conjunction with conventional paclitaxel-carboplatin therapy, demonstrated a benefit in a small number of patients (*181*). Additional strategies to diminish GSH include the administration of buthionine sulfoximine, an irreversible inhibitor of γ-glutamylcysteine synthetase, which can be safely administered to humans and is efficacious in preclinical tumor models (*182*). Moreover, glutathione is a tripeptide consisting of cysteine, glutamate, and glycine. Thus, decreasing glutamate levels using glutaminase inhibitors or diminishing cysteine levels by preventing extracellular cystine (two linked cysteine molecules) uptake can also raise ROS levels in cancer cells to induce cell death.

An important consideration is that normal stem cells are sensitive to ROS levels; thus, it is important to stratify patients on the basis of their expression levels of a particular antioxidant protein or antioxidant pathway. It is critical to determine which antioxidant pathways are likely up-regulated as a result of the high rate of ROS production within cancer cells. Many cancer types use the NRF2 pathway to maintain redox balance; therefore, targeting this pathway may provide a viable therapeutic opportunity (*128*). Additionally, superoxide dismutase 1 (SOD1) is overexpressed in NSCLC, and its inhibition kills human NSCLC cells and decreases the tumor burden in mouse models of NSCLC (*183*). Because NRF2 and SOD1 knockout mice develop normally, short-term inhibition of these pathways might be an effective way to kill cancer cells.

29.8.1.1: TECHNOLOGIES ENABLING DISCOVERY IN CANCER METABOLISM

Many recent advances in our understanding of cancer metabolism have been propelled by advanced technologies to detect metabolites and metabolic activities (*184*). A key concept is that quantifying metabolites (that is, metabolomics) is a more distinct form of metabolic analysis than measuring the activities of metabolic pathways [that is, metabolic flux analysis (*185*)]. Although these two approaches can provide complementary types of information, they are not interchangeable. One cannot infer metabolic activity from changes in metabolite

levels, and altered metabolic fluxes may or may not cause changes in metabolite levels (186). Both of these approaches have provided important recent insights into cancer metabolism, and using the two techniques together provides the most complete assessment of metabolic phenotypes.

Metabolomics experiments seek to characterize and quantify the metabolites in a biological sample, usually by nuclear magnetic resonance (NMR) or, more commonly, mass spectrometry. Depending on the methods of extraction, separation, and detection, metabolomics experiments may focus on particular classes of metabolites or provide a comprehensive analysis of as many metabolites as possible. Targeted approaches typically detect a few dozen to a few hundred molecules, whereas untargeted analyses may detect more than 1000. Detecting alterations of metabolite levels in cancer can be extremely valuable. The massive accumulation of D2HG in IDH1-mutant gliomas was initially discovered through a metabolomics approach (33). Because altered metabolite levels can be detected noninvasively using ^{1}H magnetic resonance spectroscopy (MRS), perturbed metabolite levels discovered through metabolomics can sometimes be translated into clinical diagnostic techniques. Elevated levels of lactate, choline, glycine, and other metabolites are detected by MRS in glioma. More recently, MRS techniques have been developed to monitor specific metabolic states programmed by tumor-specific mutations in metabolic enzymes. Applications include elevated 2HG in IDH1/IDH2-mutated gliomas (187) and elevated succinate in SDH-deficient paragangliomas (188).

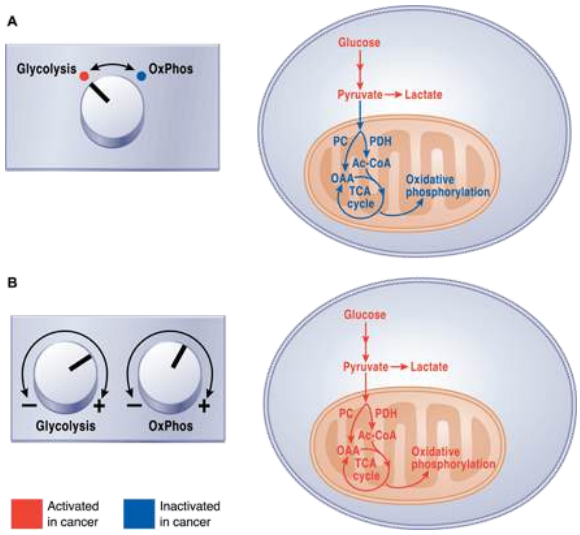
Metabolic flux studies use isotope tracers like ^{13}C, ^{15}N, and ^{2}H to track flow through metabolic pathways. Typically, a nutrient of interest is labeled by an isotope (for example, ^{13}C-glucose) and supplied to cancer cells in the culture medium. Metabolites extracted from the culture are analyzed for isotope enrichment using mass spectrometry or NMR. The extent and distribution of labeling within informative metabolites encode information about which pathways are active in the cells. Incorporating additional data (for example, definitive rates of nutrient consumption, waste secretion, and biomass production) allows quantitative fluxes to be determined across a metabolic network.

Isotope tracing studies provide information about metabolic alterations in cancer cells that cannot be detected by metabolite levels alone. For example, hypoxia and mutations in the ETC induce a restructuring of the TCA cycle in which many of the intermediates are produced in the reverse order from the conventional form of the cycle. The key reaction in this pathway involves the reductive carboxylation of α-ketoglutarate to isocitrate in a NADPH-dependent carboxylation reaction catalyzed by IDH1 and/or IDH2. Although metabolomics experiments can detect altered levels of TCA cycle metabolites in cells using the reductive carboxylation pathway or in cells with deficiencies in pyruvate import into mitochondria, the marked restructuring of the cycle is apparent only through isotope tracing experiments, particularly experiments using ^{13}C-glutamine as the tracer (69, 78, 79, 189–191). An example of the use of isotope tracers to identify metabolic liabilities involves the surprising discovery that a significant fraction of cellular NADPH, particularly in the mitochondria, is produced through folate metabolism (100, 101). These studies involved a sophisticated combination of ^{13}C and ^{2}H tracers, coupled with quantitative measurements of metabolic flux.

Several recent studies have begun to use stable isotopes to investigate metabolism in intact tumors. Because these isotopes do not undergo radioactive decay, they are safe for administration to animals and human subjects. Systemic administration of ^{13}C-labeled nutrients through either boluses or continuous infusions has been shown to generate substantial labeling of glycolytic and TCA cycle intermediates in tumors. In mice bearing orthotopic transplants of high-grade human gliomas, continuous infusion of ^{13}C-glucose was demonstrated to produce steady-state labeling of metabolites from the TCA cycle within the tumor, enabling the assessment of several metabolic pathways (192). Here, tumors with diverse oncogenotypes oxidized glucose-derived pyruvate in the mitochondria and synthesized glutamine from glucose carbon. In contrast to most cultured glioma cell lines, these tumors did not demonstrate significant levels of ^{13}C-glutamine oxidation in vivo, and primary cell lines derived from the tumors did not require glutamine for survival or proliferation. In another study, metabolism of ^{13}C-glucose and ^{13}C-glutamine in autochthonous models of *MYC*- or *MET*-driven tumorigenesis revealed that metabolic phenotypes depend not only on the tumor's genetic driver but also on the tissue or origin. *MYC* but not *MET* stimulated glutamine catabolism in liver tumors, whereas *MYC*-driven lung tumors expressed glutamine synthetase and accumulated glutamine (193). Thus, in vivo isotope tracing can detect metabolic activities of intact tumors and characterize some of the factors that specify the metabolic phenotype.

Administration of ^{13}C-labeled nutrients has also proven to be valuable in human cancer (172, 194–197). Fan *et al.* (196) used ^{13}C-glucose to demonstrate that human non–small cell lung tumors metabolize glucose through glycolysis and the TCA cycle concurrently, with metabolites from both pathways demonstrating higher levels of labeling in tumors relative to adjacent lung tissue. In a subsequent study, these investigators demonstrated that the anaplerotic enzyme pyruvate carboxylase (PC) was highly expressed in lung tumors and contributed to ^{13}C labeling in TCA cycle intermediates (195). Enhanced glucose oxidation involving both PC and pyruvate dehydrogenase (PDH) was demonstrated in a separate cohort of non–small cell lung tumors, in which formal analysis of metabolic fluxes was used to complement measurements of ^{13}C labeling (197). An important conclusion from these studies, and from a similar study in mice bearing *KRAS*-driven tumors (198), is that non–small cell lung tumors demonstrate higher levels of both glycolysis and glucose oxidation relative to adjacent, benign lung. This finding sharply contrasts with the frequently invoked "switch" from oxidative metabolism to glycolysis in malignant tissue, commonly used to explain the Warburg effect (Fig. 5A). Rather, the data support a model in which the amplitude of both pathways is increased simultaneously, perhaps through increased substrate delivery and enzyme expression in tumor cells (Fig. 5B). It is also significant that human tumors exhibit substantial heterogeneity of metabolic phenotypes, both between tumors and even within distinct regions of the same tumor (197). The extent of glucose-dependent labeling of TCA cycle intermediates is predicted by noninvasive

assessment of tumor perfusion by magnetic resonance imaging, providing an approach to identify areas of regional metabolic heterogeneity in human cancer (*197*).



- Download high-res image
- Open in new tab
- Download Powerpoint

Fig. 5 Relationship between glycolysis and oxidative phosphorylation in cancer cells.
(**A**) A common view of cancer cell metabolism invokes a switch from glucose oxidation in normal tissues toward glycolysis and suppressed oxidative phosphorylation (OxPhos) in cancer. (**B**) Analysis of metabolic activity in intact tumors from humans and mice argues against a switch. Rather, tumors appear to enhance both glycolysis and glucose oxidation simultaneously relative to surrounding tissue.

Metabolomics and metabolic flux analysis can be integrated with functional genomics to identify and understand metabolic vulnerabilities in cancer cells. This approach has produced several good examples of screens that identified potential therapeutic targets while stimulating entirely new lines of investigation in cancer cell biology. For example, the serine biosynthetic enzyme PHGDH was first identified as a metabolic vulnerability in breast cancer cells through a large-scale in vivo short hairpin RNA screen targeting thousands of metabolic enzymes (*25*). *PHGDH* is frequently amplified at the genomic level in breast tumors and melanomas and exhibits oncogene-like features in cell culture (*25*, *26*). Subsequent work on serine biosynthesis, much of it involving metabolomics and metabolic flux analysis, has uncovered novel functions and liabilities of this pathway in cancer cell growth and stress resistance (*129*, *150*, *151*). Combining functional screens with metabolic analysis can also identify context-specific vulnerabilities that may be therapeutically actionable. A CRISPR (clustered regularly interspaced short palindromic repeats)–based loss-of-function screen identified GOT1, the cytosolic aspartate aminotransferase, as conditionally essential for survival during treatment with the ETC inhibitor phenformin (*199*). Isotope labeling then demonstrated that ETC blockade caused the direction of this enzyme to reverse from aspartate consumption in untreated cells to aspartate synthesis during ETC blockade (*200*). In addition to the discovery of synthetic lethality between ETC and GOT1 inhibition, these studies led to the novel biological concept that a major function of the ETC in proliferating cells is to support the synthesis of aspartate for nucleotide and protein synthesis (*199*, *200*).

29.8.1.1: CONCLUSIONS AND CURRENT CHALLENGES

Substantial progress has been made in the past decade toward understanding the mechanisms, biological consequences, and liabilities associated with metabolic reprogramming in cancer. Several common themes have emerged from this research (Box 1). First, metabolic reprogramming is essential for the biology of malignant cells, particularly their ability to survive and grow by using conventional metabolic pathways to produce energy, synthesize biosynthetic precursors, and maintain redox balance. Second, metabolic reprogramming is the result of mutations in oncogenes and tumor suppressors, leading to activation of PI3K and mTORC1 signaling pathways and transcriptional networks involving HIFs, MYC, and SREBP-1. Third, alterations in metabolite levels can affect cellular signaling, epigenetics, and gene expression through posttranslational modifications such as acetylation, methylation, and thiol oxidation. Fourth, taken together, studies on cultured cells have demonstrated a remarkable diversity of anabolic and catabolic pathways in cancer, with induction of autophagy and

utilization of extracellular lipids and proteins complementing the classical pathways like glycolysis and glutaminolysis. We have exited the period when cancer metabolism could be considered synonymous with the Warburg effect.

Box 1

Key Principles and Lessons Learned

Reprogrammed metabolic pathways are essential for cancer cell survival and growth.

Frequently reprogrammed activities include those that allow tumor cells to take up abundant nutrients and use them to produce ATP, generate biosynthetic precursors and macromolecules, and tolerate stresses associated with malignancy (for example, redox stress and hypoxia).

An emerging class of reprogrammed pathways includes those allowing cancer cells to tolerate nutrient depletion by catabolizing macromolecules from inside or outside the cell (for example, autophagy, macropinocytosis, and lipid scavenging).

Reprogramming may be regulated intrinsically by tumorigenic mutations in cancer cells or extrinsically by influences of the microenvironment.

Oncometabolites (for example, 2HG) accumulate as a consequence of genetic changes within a tumor and contribute to the molecular process of malignant transformation.

Many metabolites exert their biological effects outside of the classical metabolic network, affecting signal transduction, epigenetics, and other functions.

New approaches to assess metabolism in living tumors in humans and mice may improve our ability to understand how metabolic reprogramming is regulated and which altered pathways hold opportunities to improve care of cancer patients.

Several challenges will likely shape research over the next decade. First, the studies cited above were performed primarily in cancer cell lines rather than intact tumors. These straightforward experimental models have been highly informative about the molecular mechanisms of metabolic reprogramming, particularly those linking aberrant signaling to altered metabolic fluxes. But it is challenging (perhaps impossible) to model an accurate tumor microenvironment in culture. Direct analysis of metabolic fluxes in intact tumors should begin to play a more prominent role in the field and may prove essential in determining precisely how to deploy metabolic inhibitors in clinical trials. Along these lines, it is remarkable that some tumor cell metabolic vulnerabilities observed in vivo are absent from cultured cell models (*198*) and that metabolic phenotypes are inconsistent even across single solid tumors in patients (*197*). Developing rational therapeutic strategies will be aided by learning how to derive metabolic information efficiently and comprehensively from both preclinical and clinical models of intact tumor growth. A further challenge for these in vivo studies will be to develop analytical or computational approaches to deconvolute the distinct metabolic phenotypes of discrete cell types (cancer cells, cancer-associated fibroblasts, lymphocytes, and endothelial cells) within solid tumors. This may allow us to understand the metabolic cooperativity among populations of cells within a tumor and whether metabolic reprogramming of stromal cells provides therapeutic opportunities. Second, by far the best recent candidate for a targetable, tumor-specific metabolic activity is the neomorphic function of mutant IDH1/IDH2. This has stimulated intense interest in finding other metabolic alterations for which the therapeutic window may be wide enough for real clinical opportunities. Third, although we have learned a great deal about the metabolic pathways that support cancer cell proliferation, we know much less about the metabolism that supports survival of nonproliferating tumor cells, which constitute the bulk of the malignant cells in most solid tumors. Along these lines, the metabolism of tumor-initiating cells/cancer stem cells is just now beginning to be investigated, and it will be of major interest to devise strategies to target metabolism in these cells. Finally, we still know relatively little about metabolic interactions between tumor and host. This area has the potential for enormous impact on public health. It is clear that obesity and diabetes, both of which are reaching epidemic proportions in the developed world, increase cancer risk, but we lack insight into how to break these links.

29.8.1.1: REFERENCES AND NOTES

1. ↵
    1. D. Hanahan,
    2. R. A. Weinberg

    , *Hallmarks of cancer: The next generation. Cell 144, 646–674 (2011).*

    CrossRefPubMedWeb of ScienceGoogle Scholar
2. ↵
    1. N. N. Pavlova,
    2. C. B. Thompson

    , *The emerging hallmarks of cancer metabolism. Cell Metab. 23, 27–47 (2016).*

CrossRefPubMedGoogle Scholar

3. ↵
   1. K. C. Patra,
   2. Q. Wang,
   3. P. T. Bhaskar,
   4. L. Miller,
   5. Z. Wang,
   6. W. Wheaton,
   7. N. Chandel,
   8. M. Laakso,
   9. W. J. Muller,
   10. E. L. Allen,
   11. A. K. Jha,
   12. G. A. Smolen,
   13. M. F. Clasquin,
   14. R. B. Robey,
   15. N. Hay

   , *Hexokinase 2 is required for tumor initiation and maintenance and its systemic deletion is therapeutic in mouse models of cancer. Cancer Cell 24, 213–228 (2013).*

   CrossRefPubMedWeb of ScienceGoogle Scholar

4. ↵
   1. E. H. Shroff,
   2. L. S. Eberlin,
   3. V. M. Dang,
   4. A. M. Gouw,
   5. M. Gabay,
   6. S. J. Adam,
   7. D. I. Bellovin,
   8. P. T. Tran,
   9. W. M. Philbrick,
   10. A. Garcia-Ocana,
   11. S. C. Casey,
   12. Y. Li,
   13. C. V. Dang,
   14. R. N. Zare,
   15. D. W. Felshera

   , *MYC oncogene overexpression drives renal cell carcinoma in a mouse model through glutamine metabolism. Proc. Natl. Acad. Sci. U.S.A. 112, 6539–6544 (2015).*

   Abstract/FREE Full TextGoogle Scholar

5. ↵
   1. L. A. Clavell,
   2. R. D. Gelber,
   3. H. J. Cohen,
   4. S. Hitchcock-Bryan,
   5. J. R. Cassady,
   6. N. J. Tarbell,
   7. S. R. Blattner,
   8. R. Tantravahi,
   9. P. Leavitt,
   10. S. E. Sallan

   , *Four-agent induction and intensive asparaginase therapy for treatment of childhood acute lymphoblastic leukemia. N. Engl. J. Med. 315, 657–663 (1986).*

   CrossRefPubMedWeb of ScienceGoogle Scholar

6. ↵
   1. J. Yun,
   2. C. Rago,
   3. I. Cheong,
   4. R. Pagliarini,
   5. P. Angenendt,
   6. H. Rajagopalan,
   7. K. Schmidt,
   8. J. K. Willson,
   9. S. Markowitz,
   10. S. Zhou,
   11. L. A. Diaz Jr.,
   12. V. E. Velculescu,
   13. C. Lengauer,
   14. K. W. Kinzler,
   15. B. Vogelstein,
   16. N. Papadopoulos

   , *Glucose deprivation contributes to the development of KRAS pathway mutations in tumor cells. Science 325, 1555–1559 (2009).*

   Abstract/FREE Full TextGoogle Scholar
7. ↵
   1. J. M. Loo,
   2. A. Scherl,
   3. A. Nguyen,
   4. F. Y. Man,
   5. E. Weinberg,
   6. Z. Zeng,
   7. L. Saltz,
   8. P. B. Paty,
   9. S. F. Tavazoie

   , *Extracellular metabolic energetics can promote cancer progression. Cell 160, 393–406 (2015).*

   CrossRefPubMedGoogle Scholar
8. ↵
   1. E. Piskounova,
   2. M. Agathocleous,
   3. M. M. Murphy,
   4. Z. Hu,
   5. S. E. Huddlestun,
   6. Z. Zhao,
   7. A. M. Leitch,
   8. T. M. Johnson,
   9. R. J. DeBerardinis,
   10. S. J. Morrison

   , *Oxidative stress inhibits distant metastasis by human melanoma cells. Nature 527, 186–191 (2015).*

   CrossRefPubMedGoogle Scholar
9. ↵
   1. L. K. Boroughs,
   2. R. J. DeBerardinis

   , *Metabolic pathways promoting cancer cell survival and growth. Nat. Cell Biol. 17, 351–359 (2015).*

   CrossRefPubMedGoogle Scholar
10. ↵
   1. P. S. Ward,
   2. C. B. Thompson

, *Metabolic reprogramming: A cancer hallmark even Warburg did not anticipate. Cancer Cell 21, 297–308 (2012).*
CrossRefPubMedWeb of ScienceGoogle Scholar

11. ↵
    1. S. Y. Lunt,
    2. M. G. Vander Heiden

    , *Aerobic glycolysis: Meeting the metabolic requirements of cell proliferation. Annu. Rev. Cell Dev. Biol. 27, 441–464 (2011).*
    CrossRefPubMedGoogle Scholar
12. ↵
    1. W. H. Koppenol,
    2. P. L. Bounds,
    3. C. V. Dang

    , *Otto Warburg's contributions to current concepts of cancer metabolism. Nat. Rev. Cancer 11, 325–337 (2011).*
    CrossRefPubMedWeb of ScienceGoogle Scholar
13. ↵
    1. C. S. Ahn,
    2. C. M. Metallo

    , *Mitochondria as biosynthetic factories for cancer proliferation. Cancer Metab. 3, 1 (2015).*
    CrossRefPubMedGoogle Scholar
14. ↵
    1. O. E. Owen,
    2. S. C. Kalhan,
    3. R. W. Hanson

    , *The key role of anaplerosis and cataplerosis for citric acid cycle function. J. Biol. Chem. 277, 30409–30412 (2002).*
    FREE Full TextGoogle Scholar
15. ↵
    1. J. R. Cantor,
    2. D. M. Sabatini

    , *Cancer cell metabolism: One hallmark, many faces. Cancer Discov. 2, 881–898 (2012).*
    Abstract/FREE Full TextGoogle Scholar
16. ↵
    1. C. C. Dibble,
    2. B. D. Manning

    , *Signal integration by mTORC1 coordinates nutrient input with biosynthetic output. Nat. Cell Biol. 15, 555–564 (2013).*
    CrossRefPubMedWeb of ScienceGoogle Scholar
17. ↵
    1. T. L. Yuan,
    2. L. C. Cantley

    , *PI3K pathway alterations in cancer: Variations on a theme. Oncogene 27, 5497–5510 (2008).*
    CrossRefPubMedWeb of ScienceGoogle Scholar
18. ↵
    1. Z. E. Stine,
    2. Z. E. Walton,
    3. B. J. Altman,
    4. A. L. Hsieh,
    5. C. V. Dang

    , *MYC, metabolism, and cancer. Cancer Discov. 5, 1024–1039 (2015).*
    Abstract/FREE Full TextGoogle Scholar
19. ↵

1. F. Kruiswijk,
2. C. F. Labuschagne,
3. K. H. Vousden

, *p53 in survival, death and metabolic health: A lifeguard with a licence to kill. Nat. Rev. Mol. Cell Biol. 16, 393–405 (2015).*

CrossRefPubMedGoogle Scholar

20. ↵

1. L. Jiang,
2. N. Kon,
3. T. Li,
4. S. J. Wang,
5. T. Su,
6. H. Hibshoosh,
7. R. Baer,
8. W. Gu

, *Ferroptosis as a p53-mediated activity during tumour suppression. Nature 520, 57–62 (2015).*

CrossRefPubMedGoogle Scholar

21. ↵

1. T. Li,
2. N. Kon,
3. L. Jiang,
4. M. Tan,
5. T. Ludwig,
6. Y. Zhao,
7. R. Baer,
8. W. Gu

, *Tumor suppression in the absence of p53-mediated cell-cycle arrest, apoptosis, and senescence. Cell 149, 1269–1283 (2012).*

CrossRefPubMedWeb of ScienceGoogle Scholar

22. ↵

1. R. K. Jain,
2. L. L. Munn,
3. D. Fukumura

, *Dissecting tumour pathophysiology using intravital microscopy. Nat. Rev. Cancer 2, 266–276 (2002).*

CrossRefPubMedWeb of ScienceGoogle Scholar

23. ↵

1. G.L. Semenza

, *Hypoxia-inducible factors in physiology and medicine. Cell 148, 399–408 (2012).*

CrossRefPubMedWeb of ScienceGoogle Scholar

24. ↵

1. W. G. Kaelin Jr.,
2. P. J. Ratcliffe

, *Oxygen sensing by metazoans: The central role of the HIF hydroxylase pathway. Mol. Cell 30, 393–402 (2008).*

CrossRefPubMedWeb of ScienceGoogle Scholar

25. ↵

1. R. Possemato,
2. K. M. Marks,
3. Y. D. Shaul,
4. M. E. Pacold,
5. D. Kim,
6. K. Birsoy,
7. S. Sethumadhavan,
8. H.-K. Woo,

9. H. G. Jang,
10. A. K. Jha,
11. W. W. Chen,
12. F. G. Barrett,
13. N. Stransky,
14. Z.-Y. Tsun,
15. G. S. Cowley,
16. J. Barretina,
17. N. Y. Kalaany,
18. P. P. Hsu,
19. K. Ottina,
20. A. M. Chan,
21. B. Yuan,
22. L. A. Garraway,
23. D. E. Root,
24. M. Mino-Kenudson,
25. E. F. Brachtel,
26. E. M. Driggers,
27. D. M. Sabatini

, *Functional genomics reveal that the serine synthesis pathway is essential in breast cancer. Nature 476, 346–350 (2011).*

CrossRefPubMedWeb of ScienceGoogle Scholar

26. ↵

1. J. W. Locasale,
2. A. R. Grassian,
3. T. Melman,
4. C. A. Lyssiotis,
5. K. R. Mattaini,
6. A. J. Bass,
7. G. Heffron,
8. C. M. Metallo,
9. T. Muranen,
10. H. Sharfi,
11. A. T. Sasaki,
12. D. Anastasiou,
13. E. Mullarky,
14. N. I. Vokes,
15. M. Sasaki,
16. R. Beroukhim,
17. G. Stephanopoulos,
18. A. H. Ligon,
19. M. Meyerson,
20. A. L. Richardson,
21. L. Chin,
22. G. Wagner,
23. J. M. Asara,
24. J. S. Brugge,
25. L. C. Cantley,
26. M. G. Vander Heiden

, *Phosphoglycerate dehydrogenase diverts glycolytic flux and contributes to oncogenesis. Nat. Genet. 43, 869–874 (2011).*

CrossRefPubMedGoogle Scholar

27. ↵

1. J. W. Locasale

, *Serine, glycine and one-carbon units: Cancer metabolism in full circle. Nat. Rev. Cancer 13, 572–583 (2013).*

CrossRefPubMedWeb of ScienceGoogle Scholar

28. ↵
1. M. Yang,
2. T. Soga,
3. P. J. Pollard

, *Oncometabolites: Linking altered metabolism with cancer. J. Clin. Invest. 123, 3652–3658 (2013).*

CrossRefPubMedWeb of ScienceGoogle Scholar

29. ↵
1. H. Yan,
2. D. W. Parsons,
3. G. Jin,
4. R. McLendon,
5. B. A. Rasheed,
6. W. Yuan,
7. I. Kos,
8. I. Batinic-Haberle,
9. S. Jones,
10. G. J. Riggins,
11. H. Friedman,
12. A. Friedman,
13. D. Reardon,
14. J. Herndon,
15. K. W. Kinzler,
16. V. E. Velculescu,
17. B. Vogelstein,
18. D. D. Bigner

, *IDH1 and IDH2 mutations in gliomas. N. Engl. J. Med. 360, 765–773 (2009).*

CrossRefPubMedWeb of ScienceGoogle Scholar

30. 1. E. R. Mardis,
2. L. Ding,
3. D. J. Dooling,
4. D. E. Larson,
5. M. D. McLellan,
6. K. Chen,
7. D. C. Koboldt,
8. R. S. Fulton,
9. K. D. Delehaunty,
10. S. D. McGrath,
11. L. A. Fulton,
12. D. P. Locke,
13. V. J. Magrini,
14. R. M. Abbott,
15. T. L. Vickery,
16. J. S. Reed,
17. J. S. Robinson,
18. T. Wylie,
19. S. M. Smith,
20. L. Carmichael,
21. J. M. Eldred,
22. C. C. Harris,
23. J. Walker,
24. J. B. Peck,
25. F. Du,
26. A. F. Dukes,
27. G. E. Sanderson,

28. A. M. Brummett,
29. E. Clark,
30. J. F. McMichael,
31. R. J. Meyer,
32. J. K. Schindler,
33. C. S. Pohl,
34. J. W. Wallis,
35. X. Shi,
36. L. Lin,
37. H. Schmidt,
38. Y. Tang,
39. C. Haipek,
40. M. E. Wiechert,
41. J. V. Ivy,
42. J. Kalicki,
43. G. Elliott,
44. R. E. Ries,
45. J. E. Payton,
46. P. Westervelt,
47. M. H. Tomasson,
48. M. A. Watson,
49. J. Baty,
50. S. Heath,
51. W. D. Shannon,
52. R. Nagarajan,
53. D. C. Link,
54. M. J. Walter,
55. T. A. Graubert,
56. J. F. DiPersio,
57. R. K. Wilson,
58. T. J. Ley

, *Recurring mutations found by sequencing an acute myeloid leukemia genome. N. Engl. J. Med. 361, 1058–1066 (2009).*

CrossRefPubMedWeb of ScienceGoogle Scholar

31. ↵

1. M. R. Kang,
2. M. S. Kim,
3. J. E. Oh,
4. Y. R. Kim,
5. S. Y. Song,
6. S. I. Seo,
7. J. Y. Lee,
8. N. J. Yoo,
9. S. H. Lee

, *Mutational analysis of IDH1 codon 132 in glioblastomas and other common cancers. Int. J. Cancer 125, 353–355 (2009).*

CrossRefPubMedWeb of ScienceGoogle Scholar

32. ↵

1. J.-A. Losman,
2. W. G. Kaelin Jr.

, *What a difference a hydroxyl makes: Mutant IDH, (R)-2-hydroxyglutarate, and cancer. Genes Dev. 27, 836–852 (2013).*

Abstract/FREE Full TextGoogle Scholar

33. ↵

1. L. Dang,
2. D. W. White,

3. S. Gross,
4. B. D. Bennett,
5. M. A. Bittinger,
6. E. M. Driggers,
7. V. R. Fantin,
8. H. G. Jang,
9. S. Jin,
10. M. C. Keenan,
11. K. M. Marks,
12. R. M. Prins,
13. P. S. Ward,
14. K. E. Yen,
15. L. M. Liau,
16. J. D. Rabinowitz,
17. L. C. Cantley,
18. C. B. Thompson,
19. M. G. Vander Heiden,
20. S. M. Su

, *Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. Nature 465, 966 (2010).*

CrossRefPubMedWeb of ScienceGoogle Scholar

34. ↵

1. P. S. Ward,
2. J. Patel,
3. D. R. Wise,
4. O. Abdel-Wahab,
5. B. D. Bennett,
6. H. A. Coller,
7. J. R. Cross,
8. V. R. Fantin,
9. C. V. Hedvat,
10. A. E. Perl,
11. J. D. Rabinowitz,
12. M. Carroll,
13. S. M. Su,
14. K. A. Sharp,
15. R. L. Levine,
16. C. B. Thompson

, *The common feature of leukemia-associated IDH1 and IDH2 mutations is a neomorphic enzyme activity converting α-ketoglutarate to 2-hydroxyglutarate. Cancer Cell 17, 225–234 (2010).*

CrossRefPubMedWeb of ScienceGoogle Scholar

35. ↵

1. M. E. Figueroa,
2. O. Abdel-Wahab,
3. C. Lu,
4. P. S. Ward,
5. J. Patel,
6. A. Shih,
7. Y. Li,
8. N. Bhagwat,
9. A. Vasanthakumar,
10. H. F. Fernandez,
11. M. S. Tallman,
12. Z. Sun,
13. K. Wolniak,

14. J. K. Peeters,
15. W. Liu,
16. S. E. Choe,
17. V. R. Fantin,
18. E. Paietta,
19. B. Löwenberg,
20. J. D. Licht,
21. L. A. Godley,
22. R. Delwel,
23. P. J. Valk,
24. C. B. Thompson,
25. R. L. Levine,
26. A. Melnick

, *Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. Cancer Cell 18, 553–567 (2010).*

CrossRefPubMedWeb of ScienceGoogle Scholar

36. 1. C. Lu,
2. P. S. Ward,
3. G. S. Kapoor,
4. D. Rohle,
5. S. Turcan,
6. O. Abdel-Wahab,
7. C. R. Edwards,
8. R. Khanin,
9. M. E. Figueroa,
10. A. Melnick,
11. K. E. Wellen,
12. D. M. O'Rourke,
13. S. L. Berger,
14. T. A. Chan,
15. R. L. Levine,
16. I. K. Mellinghoff,
17. C. B. Thompson

, *IDH mutation impairs histone demethylation and results in a block to cell differentiation. Nature 483, 474–478 (2012).*

CrossRefPubMedWeb of ScienceGoogle Scholar

37. 1. J.-A. Losman,
2. R. E. Looper,
3. P. Koivunen,
4. S. Lee,
5. R. K. Schneider,
6. C. McMahon,
7. G. S. Cowley,
8. D. E. Root,
9. B. L. Ebert,
10. W. G. Kaelin Jr.

, *(R)-2-hydroxyglutarate is sufficient to promote leukemogenesis and its effects are reversible. Science 339, 1621–1625 (2013).*

Abstract/FREE Full TextGoogle Scholar

38. ↵
1. P. Koivunen,
2. S. Lee,
3. C. G. Duncan,
4. G. Lopez,
5. G. Lu,
6. S. Ramkissoon,

7. J. A. Losman,
8. P. Joensuu,
9. U. Bergmann,
10. S. Gross,
11. J. Travins,
12. S. Weiss,
13. R. Looper,
14. K. L. Ligon,
15. R. G. W. Verhaak,
16. H. Yan,
17. W. G. Kaelin Jr.

, *Transformation by the (R)-enantiomer of 2-hydroxyglutarate linked to EGLN activation. Nature 483, 484–488 (2012).*

CrossRefPubMedGoogle Scholar

39. 1. S. Turcan,
2. D. Rohle,
3. A. Goenka,
4. L. A. Walsh,
5. F. Fang,
6. E. Yilmaz,
7. C. Campos,
8. A. W. M. Fabius,
9. C. Lu,
10. P. S. Ward,
11. C. B. Thompson,
12. A. Kaufman,
13. O. Guryanova,
14. R. Levine,
15. A. Heguy,
16. A. Viale,
17. L. G. T. Morris,
18. J. T. Huse,
19. I. K. Mellinghoff,
20. T. A. Chan

, *IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. Nature 483, 479–483 (2012).*

CrossRefPubMedWeb of ScienceGoogle Scholar

40. 1. R. Chowdhury,
2. K. K. Yeoh,
3. Y.-M. Tian,
4. L. Hillringhaus,
5. E. A. Bagg,
6. N. R. Rose,
7. I. K. H. Leung,
8. X. S. Li,
9. E. C. Y. Woon,
10. M. Yang,
11. M. A. McDonough,
12. O. N. King,
13. I. J. Clifton,
14. R. J. Klose,
15. T. D. W. Claridge,
16. P. J. Ratcliffe,
17. C. J. Schofield,
18. A. Kawamura

, *The oncometabolite 2-hydroxyglutarate inhibits histone lysine demethylases. EMBO Rep. 12, 463–469 (2011).*

Abstract/FREE Full TextGoogle Scholar

41. ↵

1. W. Xu,
2. H. Yang,
3. Y. Liu,
4. Y. Yang,
5. P. Wang,
6. S.-H. Kim,
7. S. Ito,
8. C. Yang,
9. P. Wang,
10. M.-T. Xiao,
11. L.-X. Liu,
12. W.-. Jiang,
13. J. Liu,
14. J.-. Zhang,
15. B. Wang,
16. S. Frye,
17. Y. Zhang,
18. Y.-. Xu,
19. Q.-. Lei,
20. K.-L. Guan,
21. S.-. Zhao,
22. Y. Xiong

, *Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α-ketoglutarate-dependent dioxygenases. Cancer Cell 19, 17–30 (2011).*

CrossRefPubMedWeb of ScienceGoogle Scholar

42. ↵

1. A. M. Intlekofer,
2. R. G. Dematteo,
3. S. Venneti,
4. L. W. S. Finley,
5. C. Lu,
6. A. R. Judkins,
7. A. S. Rustenburg,
8. P. B. Grinaway,
9. J. D. Chodera,
10. J. R. Cross,
11. C. B. Thompson

, *Hypoxia induces production of L-2-hydroxyglutarate. Cell Metab. 22, 304–311 (2015).*

CrossRefPubMedGoogle Scholar

43. 1. R. Rzem,
2. M.-F. Vincent,
3. E. Van Schaftingen,
4. M. Veiga-da-Cunha

, *L-2-hydroxyglutaric aciduria, a defect of metabolite repair. J. Inherit. Metab. Dis. 30, 681–689 (2007).*

CrossRefPubMedGoogle Scholar

44. ↵

1. W. M. Oldham,
2. C. B. Clish,
3. Y. Yang,
4. J. Loscalzo

, *Hypoxia-mediated increases in L-2-hydroxyglutarate coordinate the metabolic response to reductive stress. Cell Metab. 22, 291–303 (2015).*

CrossRefPubMedGoogle Scholar

45. ↵

1. M. Topçu,
2. F. Jobard,
3. S. Halliez,
4. T. Coskun,
5. C. Yalçinkayal,
6. F. O. Gerceker,
7. R. J. A. Wanders,
8. J.-F. Prud'homme,
9. M. Lathrop,
10. M. Özguc,
11. J. Fischer

, *L-2-Hydroxyglutaric aciduria: Identification of a mutant gene C14orf160, localized on chromosome 14q22.1. Hum. Mol. Genet. 13, 2803–2811 (2004).*

Abstract/FREE Full TextGoogle Scholar

46. ↵

1. M. Aghili,
2. F. Zahedi,
3. E. Rafiee

, *Hydroxyglutaric aciduria and malignant brain tumor: A case report and literature review. J. Neurooncol 91, 233–236 (2009).*

CrossRefPubMedGoogle Scholar

47. ↵

1. A. R. Mullen,
2. Z. Hu,
3. X. Shi,
4. L. Jiang,
5. L. K. Boroughs,
6. Z. Kovacs,
7. R. Boriack,
8. D. Rakheja,
9. L. B. Sullivan,
10. W. M. Linehan,
11. N. S. Chandel,
12. R. J. DeBerardinis

, *Oxidation of alpha-ketoglutarate is required for reductive carboxylation in cancer cells with mitochondrial defects. Cell Rep. 7, 1679–1690 (2014).*

CrossRefPubMedGoogle Scholar

48. ↵

1. E.-H. Shim,
2. C. B. Livi,
3. D. Rakheja,
4. J. Tan,
5. D. Benson,
6. V. Parekh,
7. E.-Y. Kho,
8. A. P. Ghosh,
9. R. Kirkman,
10. S. Velu,
11. S. Dutta,
12. B. Chenna,

13. S. L. Rea,
14. R. J. Mishur,
15. Q. Li,
16. T. L. Johnson-Pais,
17. L. Guo,
18. S. Bae,
19. S. Wei,
20. K. Block,
21. S. Sudarshan

, *L-2-Hydroxyglutarate: An epigenetic modifier and putative oncometabolite in renal cancer. Cancer Discov. 4, 1290–1298 (2014).*

Abstract/FREE Full TextGoogle Scholar

49. ↵

1. I. P. M. Tomlinson,
2. N. A. Alam,
3. A. J. Rowan,
4. E. Barclay,
5. E. E. M. Jaeger,
6. D. Kelsell,
7. I. Leigh,
8. P. Gorman,
9. H. Lamlum,
10. S. Rahman,
11. R. R. Roylance,
12. S. Olpin,
13. S. Bevan,
14. K. Barker,
15. N. Hearle,
16. R. S. Houlston,
17. M. Kiuru,
18. R. Lehtonen,
19. A. Karhu,
20. S. Vilkki,
21. P. Laiho,
22. C. Eklund,
23. O. Vierimaa,
24. K. Aittomäki,
25. M. Hietala,
26. P. Sistonen,
27. A. Paetau,
28. R. Salovaara,
29. R. Herva,
30. V. Launonen,
31. L. A. Aaltonen, Multiple Leiomyoma Consortium

, *Germline mutations in FH predispose to dominantly inherited uterine fibroids, skin leiomyomata and papillary renal cell cancer. Nat. Genet. 30, 406–410 (2002).*

CrossRefPubMedWeb of ScienceGoogle Scholar

50. 1. E. Gottlieb,
2. I. P. M. Tomlinson

, *Mitochondrial tumour suppressors: A genetic and biochemical update. Nat. Rev. Cancer 5, 857–866 (2005).*

CrossRefPubMedWeb of ScienceGoogle Scholar

51. ↵

1. B. E. Baysal,
2. R. E. Ferrell,

3. J. E. Willett-Brozick,
4. E. C. Lawrence,
5. D. Myssiorek,
6. A. Bosch,
7. A. van der Mey,
8. P. E. M. Taschner,
9. W. S. Rubinstein,
10. E. N. Myers,
11. C. W. Richard III.,
12. C. J. Cornelisse,
13. P. Devilee,
14. B. Devlin

, *Mutations in SDHD, a mitochondrial complex II gene, in hereditary paraganglioma. Science 287, 848–851 (2000).*

Abstract/FREE Full TextGoogle Scholar

52. ↵

1. T. Laukka,
2. C. J. Mariani,
3. T. Ihantola,
4. J. Z. Cao,
5. J. Hokkanen,
6. W. G. Kaelin Jr.,
7. L. A. Godley,
8. P. Koivunen

, *Fumarate and succinate regulate expression of hypoxia-inducible genes via TET enzymes. J. Biol. Chem. 291, 4256–4265 (2016).*

Abstract/FREE Full TextGoogle Scholar

53. ↵

1. M. Xiao,
2. H. Yang,
3. W. Xu,
4. S. Ma,
5. H. Lin,
6. H. Zhu,
7. L. Liu,
8. Y. Liu,
9. C. Yang,
10. Y. Xu,
11. S. Zhao,
12. D. Ye,
13. Y. Xiong,
14. K.-L. Guan

, *Inhibition of α-KG-dependent histone and DNA demethylases by fumarate and succinate that are accumulated in mutations of FH and SDH tumor suppressors. Genes Dev. 26, 1326–1338 (2012).*

Abstract/FREE Full TextGoogle Scholar

54. ↵

1. W. G. Kaelin Jr.,
2. S. L. McKnight

, *Influence of metabolism on epigenetics and disease. Cell 153, 56–69 (2013).*

CrossRefPubMedWeb of ScienceGoogle Scholar

55. ↵

1. L. B. Sullivan,
2. E. Martinez-Garcia,
3. H. Nguyen,

4. A. R. Mullen,
5. E. Dufour,
6. S. Sudarshan,
7. J. D. Licht,
8. R. J. Deberardinis,
9. N. S. Chandel

, *The proto-oncometabolite fumarate binds glutathione to amplify ROS-dependent signaling. Mol. Cell 51, 236–248 (2013).*

CrossRefPubMedWeb of ScienceGoogle Scholar

56. 1. J. Adam,
2. E. Hatipoglu,
3. L. O'Flaherty,
4. N. Ternette,
5. N. Sahgal,
6. H. Lockstone,
7. D. Baban,
8. E. Nye,
9. G. W. Stamp,
10. K. Wolhuter,
11. M. Stevens,
12. R. Fischer,
13. P. Carmeliet,
14. P. H. Maxwell,
15. C. W. Pugh,
16. N. Frizzell,
17. T. Soga,
18. B. M. Kessler,
19. M. El-Bahrawy,
20. P. J. Ratcliffe,
21. P. J. Pollard

, *Renal cyst formation in Fh1-deficient mice is independent of the Hif/Phd pathway: Roles for fumarate in KEAP1 succination and Nrf2 signaling. Cancer Cell 20, 524–537 (2011).*

CrossRefPubMedWeb of ScienceGoogle Scholar

57. 1. C. Bardella,
2. M. El-Bahrawy,
3. N. Frizzell,
4. J. Adam,
5. N. Ternette,
6. E. Hatipoglu,
7. K. Howarth,
8. L. O'Flaherty,
9. I. Roberts,
10. G. Turner,
11. J. Taylor,
12. K. Giaslakiotis,
13. V. M. Macaulay,
14. A. L. Harris,
15. A. Chandra,
16. H. J. Lehtonen,
17. V. Launonen,
18. L. A. Aaltonen,
19. C. W. Pugh,
20. R. Mihai,
21. D. Trudgian,
22. B. Kessler,

23. J. W. Baynes,
24. P. J. Ratcliffe,
25. I. P. Tomlinson,
26. P. J. Pollard

, *Aberrant succination of proteins in fumarate hydratase-deficient mice and HLRCC patients is a robust biomarker of mutation status. J. Pathol. 225, 4–11 (2011).*

CrossRefPubMedWeb of ScienceGoogle Scholar

58. ↵

1. A. Ooi,
2. J.-C. Wong,
3. D. Petillo,
4. D. Roossien,
5. V. Perrier-Trudova,
6. D. Whitten,
7. B. W. H. Min,
8. M.-H. Tan,
9. Z. Zhang,
10. X. J. Yang,
11. M. Zhou,
12. B. Gardie,
13. V. Molinié,
14. S. Richard,
15. P. H. Tan,
16. B. T. Teh,
17. K. A. Furge

, *An antioxidant response phenotype shared between hereditary and sporadic type 2 papillary renal cell carcinoma. Cancer Cell 20, 511–523 (2011).*

CrossRefPubMedGoogle Scholar

59. ↵

1. O. Warburg

, *On respiratory impairment in cancer cells. Science 124, 269–270 (1956).*

PubMedWeb of ScienceGoogle Scholar

60. ↵

1. O. Warburg

, *On the origin of cancer cells. Science 123, 309–314 (1956).*

FREE Full TextGoogle Scholar

61. ↵

1. C. V. Dang

, *Links between metabolism and cancer. Genes Dev. 26, 877–890 (2012).*

Abstract/FREE Full TextGoogle Scholar

62. ↵

1. W. J. Israelsen,
2. T. L. Dayton,
3. S. M. Davidson,
4. B. P. Fiske,
5. A. M. Hosios,
6. G. Bellinger,
7. J. Li,
8. Y. Yu,
9. M. Sasaki,
10. J. W. Horner,
11. L. N. Burga,

12. J. Xie,
13. M. J. Jurczak,
14. R. A. DePinho,
15. C. B. Clish,
16. T. Jacks,
17. R. G. Kibbey,
18. G. M. Wulf,
19. D. Di Vizio,
20. G. B. Mills,
21. L. C. Cantley,
22. M. G. Vander Heiden

, *PKM2 isoform-specific deletion reveals a differential requirement for pyruvate kinase in tumor cells. Cell 155, 397–409 (2013).*

CrossRefPubMedWeb of ScienceGoogle Scholar

63. ↵

1. S. Joshi,
2. D. Tolkunov,
3. H. Aviv,
4. A. A. Hakimi,
5. M. Yao,
6. J. J. Hsieh,
7. S. Ganesan,
8. C. S. Chan,
9. E. White

, *The genomic landscape of renal oncocytoma identifies a metabolic barrier to tumorigenesis. Cell Rep. 13, 1895–1908 (2015).*

CrossRefPubMedGoogle Scholar

64. ↵

1. F. Weinberg,
2. R. Hamanaka,
3. W. W. Wheaton,
4. S. Weinberg,
5. J. Joseph,
6. M. Lopez,
7. B. Kalyanaraman,
8. G. M. Mutlu,
9. G. R. S. Budinger,
10. N. S. Chandel

, *Mitochondrial metabolism and ROS generation are essential for Kras-mediated tumorigenicity. Proc. Natl. Acad. Sci. U.S.A. 107, 8788–8793 (2010).*

Abstract/FREE Full TextGoogle Scholar

65. ↵

1. I. Martinez-Reyes,
2. L. P. Diebold,
3. H. Kong,
4. M. Schieber,
5. H. Huang,
6. C. T. Hensley,
7. M. M. Mehta,
8. T. Wang,
9. J. H. Santos,
10. R. Woychik,
11. E. Dufour,
12. J. N. Spelbrink,
13. S. E. Weinberg,

14. Y. Zhao,
15. R. J. DeBerardinis,
16. N. S. Chandel

, *TCA cycle and mitochondrial membrane potential are necessary for diverse biological functions. Mol. Cell 61, 199–209 (2016).*

CrossRefPubMedGoogle Scholar

66. ↵

1. X. L. Zu,
2. M. Guppy

, *Cancer metabolism: Facts, fantasy, and fiction. Biochem. Biophys. Res. Commun. 313, 459–465 (2004).*

CrossRefPubMedWeb of ScienceGoogle Scholar

67. ↵

1. C. Lussey-Lepoutre,
2. K. E. R. Hollinshead,
3. C. Ludwig,
4. M. Menara,
5. A. Morin,
6. L.-J. Castro-Vega,
7. S. J. Parker,
8. M. Janin,
9. C. Martinelli,
10. C. Ottolenghi,
11. C. Metallo,
12. A.-P. Gimenez-Roqueplo,
13. J. Favier,
14. D. A. Tennant

, *Loss of succinate dehydrogenase activity results in dependency on pyruvate carboxylation for cellular anabolism. Nat. Commun. 6, 8784 (2015).*

CrossRefPubMedGoogle Scholar

68. 1. S. Cardaci,
2. L. Zheng,
3. G. MacKay,
4. N. J. F. van den Broek,
5. E. D. MacKenzie,
6. C. Nixon,
7. D. Stevenson,
8. S. Tumanov,
9. V. Bulusu,
10. J. J. Kamphorst,
11. A. Vazquez,
12. S. Fleming,
13. F. Schiavi,
14. G. Kalna,
15. K. Blyth,
16. D. Strathdee,
17. E. Gottlieb

, *Pyruvate carboxylation enables growth of SDH-deficient cells by supporting aspartate biosynthesis. Nat. Cell Biol. 17, 1317–1326 (2015).*

CrossRefPubMedGoogle Scholar

69. ↵

1. A. R. Mullen,
2. W. W. Wheaton,
3. E. S. Jin,
4. P.-H. Chen,

5. L. B. Sullivan,
6. T. Cheng,
7. Y. Yang,
8. W. M. Linehan,
9. N. S. Chandel,
10. R. J. DeBerardinis

, *Reductive carboxylation supports growth in tumour cells with defective mitochondria. Nature 481, 385–388 (2012).*

CrossRefPubMedWeb of ScienceGoogle Scholar

70. ↵

1. R. Guzy,
2. B. Sharma,
3. E. Bell,
4. N. Chandel,
5. P. Schumacker

, *Loss of the SdhB, but Not the SdhA, subunit of complex II triggers reactive oxygen species-dependent hypoxia-inducible factor activation and tumorigenesis. Mol. Cell. Biol. 28, 718–731 (2008).*

Abstract/FREE Full TextGoogle Scholar

71. ↵

1. C. T. Hensley,
2. A. T. Wasti,
3. R. J. DeBerardinis

, *Glutamine and cancer: Cell biology, physiology, and clinical opportunities. J. Clin. Invest. 123, 3678–3684 (2013).*

CrossRefPubMedWeb of ScienceGoogle Scholar

72. ↵

1. J. R. Mayers,
2. C. Wu,
3. C. B. Clish,
4. P. Kraft,
5. M. E. Torrence,
6. B. P. Fiske,
7. C. Yuan,
8. Y. Bao,
9. M. K. Townsend,
10. S. S. Tworoger,
11. S. M. Davidson,
12. T. Papagiannakopoulos,
13. A. Yang,
14. T. L. Dayton,
15. S. Ogino,
16. M. J. Stampfer,
17. E. L. Giovannucci,
18. Z. R. Qian,
19. D. A. Rubinson,
20. J. Ma,
21. H. D. Sesso,
22. J. M. Gaziano,
23. B. B. Cochrane,
24. S. Liu,
25. J. Wactawski-Wende,
26. J. E. Manson,
27. M. N. Pollak,
28. A. C. Kimmelman,
29. A. Souza,

30. K. Pierce,
31. T. J. Wang,
32. R. E. Gerszten,
33. C. S. Fuchs,
34. M. G. Vander Heiden,
35. B. M. Wolpin

, *Elevation of circulating branched-chain amino acids is an early event in human pancreatic adenocarcinoma development. Nat. Med. 20, 1193–1198 (2014).*

CrossRefPubMedGoogle Scholar

73. ↵

1. N. Chandel,
2. G. R. S. Budinger,
3. S. H. Choe,
4. P. T. Schumacker

, *Cellular respiration during hypoxia. Role of cytochrome oxidase as the oxygen sensor in hepatocytes. J. Biol. Chem. 272, 18808–18816 (1997).*

Abstract/FREE Full TextGoogle Scholar

74. ↵

1. J. Fan,
2. J. J. Kamphorst,
3. R. Mathew,
4. M. K. Chung,
5. E. White,
6. T. Shlomi,
7. J. D. Rabinowitz

, *Glutamine-driven oxidative phosphorylation is a major ATP source in transformed mammalian cells in both normoxia and hypoxia. Mol. Syst. Biol. 9, 712 (2013).*

CrossRefPubMedGoogle Scholar

75. ↵

1. A. Le,
2. A. N. Lane,
3. M. Hamaker,
4. S. Bose,
5. A. Gouw,
6. J. Barbi,
7. T. Tsukamoto,
8. C. J. Rojas,
9. B. S. Slusher,
10. H. Zhang,
11. L. J. Zimmerman,
12. D. C. Liebler,
13. R. J. C. Slebos,
14. P. K. Lorkiewicz,
15. R. M. Higashi,
16. T. W. M. Fan,
17. C. V. Dang

, *Glucose-independent glutamine metabolism via TCA cycling for proliferation and survival in B cells. Cell Metab. 15, 110–121 (2012).*

CrossRefPubMedWeb of ScienceGoogle Scholar

76. ↵

1. A. Le,
2. Z. E. Stine,
3. C. Nguyen,

4. J. Afzal,
5. P. Sun,
6. M. Hamaker,
7. N. M. Siegel,
8. A. M. Gouw,
9. B.-h. Kang,
10. S.-H. Yu,
11. R. L. Cochran,
12. K. A. Sailor,
13. H. Song,
14. C. V. Dang

, *Tumorigenicity of hypoxic respiring cancer cells revealed by a hypoxia–cell cycle dual reporter. Proc. Natl. Acad. Sci. U.S.A. 111, 12486–12491 (2014).*

Abstract/FREE Full TextGoogle Scholar

77. ↵

1. J. J. Kamphorst,
2. M. Nofal,
3. C. Commisso,
4. S. R. Hackett,
5. W. Lu,
6. E. Grabocka,
7. M. G. Vander Heiden,
8. G. Miller,
9. J. A. Drebin,
10. D. Bar-Sagi,
11. C. B. Thompson,
12. J. D. Rabinowitz

, *Human pancreatic cancer tumors are nutrient poor and tumor cells actively scavenge extracellular protein. Cancer Res. 75, 544–553 (2015).*

Abstract/FREE Full TextGoogle Scholar

78. ↵

1. C. M. Metallo,
2. P. A. Gameiro,
3. E. L. Bell,
4. K. R. Mattaini,
5. J. Yang,
6. K. Hiller,
7. C. M. Jewell,
8. Z. R. Johnson,
9. D. J. Irvine,
10. L. Guarente,
11. J. K. Kelleher,
12. M. G. Vander Heiden,
13. O. Iliopoulos,
14. G. Stephanopoulos

, *Reductive glutamine metabolism by IDH1 mediates lipogenesis under hypoxia. Nature 481, 380–384 (2012).*

PubMedWeb of ScienceGoogle Scholar

79. ↵

1. D. R. Wise,
2. P. S. Ward,
3. J. E. S. Shay,
4. J. R. Cross,
5. J. J. Gruber,

6. U. M. Sachdeva,
7. J. M. Platt,
8. R. G. DeMatteo,
9. M. C. Simon,
10. C. B. Thompson

, *Hypoxia promotes isocitrate dehydrogenase-dependent carboxylation of α-ketoglutarate to citrate to support cell growth and viability. Proc. Natl. Acad. Sci. U.S.A. 108, 19611–19616 (2011).*

Abstract/FREE Full TextGoogle Scholar

80. ↵

1. M. Laplante,
2. D. M. Sabatini

, *mTOR signaling in growth control and disease. Cell 149, 274–293 (2012).*

CrossRefPubMedWeb of ScienceGoogle Scholar

81. ↵

1. J. Y. Guo,
2. H.-Y. Chen,
3. R. Mathew,
4. J. Fan,
5. A. M. Strohecker,
6. G. Karsli-Uzunbas,
7. J. J. Kamphorst,
8. G. Chen,
9. J. M. S. Lemons,
10. V. Karantza,
11. H. A. Coller,
12. R. S. DiPaola,
13. C. Gelinas,
14. J. D. Rabinowitz,
15. E. White

, *Activated Ras requires autophagy to maintain oxidative metabolism and tumorigenesis. Genes Dev. 25, 460–470 (2011).*

Abstract/FREE Full TextGoogle Scholar

82. ↵

1. A. M. Strohecker,
2. E. White

, *Autophagy promotes Braf^{V600E}-driven lung tumorigenesis by preserving mitochondrial metabolism. Autophagy 10, 384–385 (2014).*

CrossRefPubMedWeb of ScienceGoogle Scholar

83. ↵

1. N. J. Lanning,
2. B. D. Looyenga,
3. A. L. Kauffman,
4. N. M. Niemi,
5. J. Sudderth,
6. R. J. DeBerardinis,
7. J. P. MacKeigan

, *A mitochondrial RNAi screen defines cellular bioenergetic determinants and identifies an adenylate kinase as a key regulator of ATP levels. Cell Rep. 7, 907–917 (2014).*

CrossRefPubMedGoogle Scholar

84. ↵

1. D. G. Hardie,
2. B. E. Schaffer,
3. A. Brunet

, *AMPK: An energy-sensing pathway with multiple inputs and outputs. Trends Cell Biol. 26, 190–201 (2016).*

CrossRefPubMedGoogle Scholar

85. ↵

1. J. J. Kamphorst,
2. J. R. Cross,
3. J. Fan,
4. E. de Stanchina,
5. R. Mathew,
6. E. P. White,
7. C. B. Thompson,
8. J. D. Rabinowitz

, *Hypoxic and Ras-transformed cells support growth by scavenging unsaturated fatty acids from lysophospholipids. Proc. Natl. Acad. Sci. U.S.A. 110, 8882–8887 (2013).*

Abstract/FREE Full TextGoogle Scholar

86. ↵

1. K. M. Nieman,
2. H. A. Kenny,
3. C. V. Penicka,
4. A. Ladanyi,
5. R. Buell-Gutbrod,
6. M. R. Zillhardt,
7. I. L. Romero,
8. M. S. Carey,
9. G. B. Mills,
10. G. S. Hotamisligil,
11. S. D. Yamada,
12. M. E. Peter,
13. K. Gwin,
14. E. Lengyel

, *Adipocytes promote ovarian cancer metastasis and provide energy for rapid tumor growth. Nat. Med. 17, 1498–1503 (2011).*

CrossRefPubMedGoogle Scholar

87. ↵

1. A. N. McCracken,
2. A. L. Edinger

, *Nutrient transporters: The Achilles' heel of anabolism. Trends Endocrinol. Metab. 24, 200–208 (2013).*

CrossRefPubMedGoogle Scholar

88. ↵

1. P. Nicklin,
2. P. Bergman,
3. B. Zhang,
4. E. Triantafellow,
5. H. Wang,
6. B. Nyfeler,
7. H. Yang,
8. M. Hild,
9. C. Kung,
10. C. Wilson,
11. V. E. Myer,
12. J. P. MacKeigan,
13. J. A. Porter,
14. Y. K. Wang,
15. L. C. Cantley,
16. P. M. Finan,

17. L. O. Murphy
, *Bidirectional transport of amino acids regulates mTOR and autophagy. Cell 136, 521–534 (2009).*
CrossRefPubMedWeb of ScienceGoogle Scholar

89. ↵
1. L. Galluzzi,
2. F. Pietrocola,
3. B. Levine,
4. G. Kroemer
, *Metabolic control of autophagy. Cell 159, 1263–1276 (2014).*
CrossRefPubMedGoogle Scholar

90. ↵
1. E. White
, *The role for autophagy in cancer. J. Clin. Invest. 125, 42–46 (2015).*
CrossRefPubMedGoogle Scholar

91. ↵
1. L. Galluzzi,
2. F. Pietrocola,
3. J. M. Bravo-San Pedro,
4. R. K. Amaravadi,
5. E. H. Baehrecke,
6. F. Cecconi,
7. P. Codogno,
8. J. Debnath,
9. D. A. Gewirtz,
10. V. Karantza,
11. A. Kimmelman,
12. S. Kumar,
13. B. Levine,
14. M. C. Maiuri,
15. S. J. Martin,
16. J. Penninger,
17. M. Piacentini,
18. D. C. Rubinsztein,
19. H.-U. Simon,
20. A. Simonsen,
21. A. M. Thorburn,
22. G. Velasco,
23. K. M. Ryan,
24. G. Kroeme
r, Autophagy in malignant transformation and cancer progression. EMBO J. 34, 856–880 (2015).
Abstract/FREE Full TextGoogle Scholar

92. ↵
1. C. Commisso,
2. S. M. Davidson,
3. R. G. Soydaner-Azeloglu,
4. S. J. Parker,
5. J. J. Kamphorst,
6. S. Hackett,
7. E. Grabocka,
8. M. Nofal,
9. J. A. Drebin,
10. C. B. Thompson,

11. J. D. Rabinowitz,
12. C. M. Metallo,
13. M. G. Vander Heiden,
14. D. Bar-Sagi

, *Macropinocytosis of protein is an amino acid supply route in Ras-transformed cells. Nature 497, 633–637 (2013).*

CrossRefPubMedWeb of ScienceGoogle Scholar

93. ↵

1. W. Palm,
2. Y. Park,
3. K. Wright,
4. N. N. Pavlova,
5. D. A. Tuveson,
6. C. B. Thompson

, *The utilization of extracellular proteins as nutrients is suppressed by mTORC1. Cell 162, 259–270 (2015).*

CrossRefPubMedGoogle Scholar

94. ↵

1. H. Yoo,
2. G. Stephanopoulos,
3. J. K. Kelleher

, *Quantifying carbon sources for de novo lipogenesis in wild-type and IRS-1 knockout brown adipocytes. J. Lipid Res. 45, 1324–1332 (2004).*

Abstract/FREE Full TextGoogle Scholar

95. ↵

1. R. J. DeBerardinis,
2. A. Mancuso,
3. E. Daikhin,
4. I. Nissim,
5. M. Yudkoff,
6. S. Wehrli,
7. C. B. Thompson

, *Beyond aerobic glycolysis: Transformed cells can engage in glutamine metabolism that exceeds the requirement for protein and nucleotide synthesis. Proc. Natl. Acad. Sci. U.S.A. 104, 19345–19350 (2007).*

Abstract/FREE Full TextGoogle Scholar

96. ↵

1. Z. T. Schug,
2. B. Peck,
3. D. T. Jones,
4. Q. Zhang,
5. S. Grosskurth,
6. I. S. Alam,
7. L. M. Goodwin,
8. E. Smethurst,
9. S. Mason,
10. K. Blyth,
11. L. McGarry,
12. D. James,
13. E. Shanks,
14. G. Kalna,
15. R. E. Saunders,
16. M. Jiang,
17. M. Howell,
18. F. Lassailly,

19. M. Z. Thin,
20. B. Spencer-Dene,
21. G. Stamp,
22. N. J. F. van den Broek,
23. G. Mackay,
24. V. Bulusu,
25. J. J. Kamphorst,
26. S. Tardito,
27. D. Strachan,
28. A. L. Harris,
29. E. O. Aboagye,
30. S. E. Critchlow,
31. M. J. O. Wakelam,
32. A. Schulze,
33. E. Gottlieb

, *Acetyl-CoA synthetase 2 promotes acetate utilization and maintains cancer cell growth under metabolic stress. Cancer Cell 27, 57–71 (2015).*

CrossRefPubMedGoogle Scholar

97. ↵

1. C. R. Green,
2. M. Wallace,
3. A. S. Divakaruni,
4. S. A. Phillips,
5. A. N. Murphy,
6. T. P. Ciaraldi,
7. C. M. Metallo

, *Branched-chain amino acid catabolism fuels adipocyte differentiation and lipogenesis. Nat. Chem. Biol. 12, 15–21 (2016).*

CrossRefPubMedGoogle Scholar

98. ↵

1. R. Kannan,
2. I. Lyon,
3. N. Baker

, *Dietary control of lipogenesis in vivo in host tissues and tumors of mice bearing Ehrlich ascites carcinoma. Cancer Res. 40, 4606–4611 (1980).*

Abstract/FREE Full TextGoogle Scholar

99. ↵

1. M. Ookhtens,
2. R. Kannan,
3. I. Lyon,
4. N. Baker

, *Liver and adipose tissue contributions to newly formed fatty acids in an ascites tumor. Am. J. Physiol. 247, R146–R153 (1984).*

Web of ScienceGoogle Scholar

100. ↵

1. J. Fan,
2. J. Ye,
3. J. J. Kamphorst,
4. T. Shlomi,
5. C. B. Thompson,
6. J. D. Rabinowitz

, *Quantitative flux analysis reveals folate-dependent NADPH production. Nature 510, 298–302 (2014).*

CrossRefPubMedWeb of ScienceGoogle Scholar

101. ↵

1. C. A. Lewis,
2. S. J. Parker,
3. B. P. Fiske,
4. D. McCloskey,
5. D. Y. Gui,
6. C. R. Green,
7. N. I. Vokes,
8. A. M. Feist,
9. M. G. Vander Heiden,
10. C. M. Metallo

, *Tracing compartmentalized NADPH metabolism in the cytosol and mitochondria of mammalian cells. Mol. Cell 55, 253–263 (2014).*

CrossRefPubMedWeb of ScienceGoogle Scholar

102. ↵

1. J. D. Horton,
2. J. L. Goldstein,
3. M. S. Brown

, *SREBPs: Activators of the complete program of cholesterol and fatty acid synthesis in the liver. J. Clin. Invest. 109, 1125–1131 (2002).*

CrossRefPubMedWeb of ScienceGoogle Scholar

103. ↵

1. K. Düvel,
2. J. L. Yecies,
3. S. Menon,
4. P. Raman,
5. A. I. Lipovsky,
6. A. L. Souza,
7. E. Triantafellow,
8. Q. Ma,
9. R. Gorski,
10. S. Cleaver,
11. M. G. Vander Heiden,
12. J. P. MacKeigan,
13. P. M. Finan,
14. C. B. Clish,
15. L. O. Murphy,
16. B. D. Manning

, *Activation of a metabolic gene regulatory network downstream of mTOR complex 1. Mol. Cell 39, 171–183 (2010).*

CrossRefPubMedWeb of ScienceGoogle Scholar

104. ↵

1. J. L. Goldstein,
2. R. A. DeBose-Boyd,
3. M. S. Brown

, *Protein sensors for membrane sterols. Cell 124, 35–46 (2006).*

CrossRefPubMedWeb of ScienceGoogle Scholar

105. ↵

1. T. R. Peterson,
2. S. S. Sengupta,
3. T. E. Harris,
4. A. E. Carmack,
5. S. A. Kang,
6. E. Balderas,
7. D. A. Guertin,
8. K. L. Madden,

9. A. E. Carpenter,
10. B. N. Finck,
11. D. M. Sabatini

, *mTOR complex 1 regulates lipin 1 localization to control the SREBP pathway. Cell 146, 408–420 (2011).*

CrossRefPubMedWeb of ScienceGoogle Scholar

106. ↵

1. R. J. Deberardinis,
2. J. J. Lum,
3. C. B. Thompson

, *Phosphatidylinositol 3-kinase-dependent modulation of carnitine palmitoyltransferase 1A expression regulates lipid metabolism during hematopoietic cell growth. J. Biol. Chem. 281, 37372–37380 (2006).*

Abstract/FREE Full TextGoogle Scholar

107. ↵

1. R. M. Young,
2. D. Ackerman,
3. Z. L. Quinn,
4. A. Mancuso,
5. M. Gruber,
6. L. Liu,
7. D. N. Giannoukos,
8. E. Bobrovnikova-Marjon,
9. J. A. Diehl,
10. B. Keith,
11. M. C. Simon

, *Dysregulated mTORC1 renders cells critically dependent on desaturated lipids for survival under tumor-like stress. Genes Dev. 27, 1115–1131 (2013).*

Abstract/FREE Full TextGoogle Scholar

108. ↵

1. C. Yokoyama,
2. X. Wang,
3. M. R. Briggs,
4. A. Admon,
5. J. Wu,
6. X. Hua,
7. J. L. Goldstein,
8. M. S. Brown

, *SREBP-1, a basic-helix-loop-helix-leucine zipper protein that controls transcription of the low density lipoprotein receptor gene. Cell 75, 187–197 (1993).*

CrossRefPubMedWeb of ScienceGoogle Scholar

109. ↵

1. D. Guo,
2. R. M. Prins,
3. J. Dang,
4. D. Kuga,
5. A. Iwanami,
6. H. Soto,
7. K. Y. Lin,
8. T. T. Huang,
9. D. Akhavan,
10. M. B. Hock,
11. S. Zhu,
12. A. A. Kofman,

13. S. J. Bensinger,
14. W. H. Yong,
15. H. V. Vinters,
16. S. Horvath,
17. A. D. Watson,
18. J. G. Kuhn,
19. H. I. Robins,
20. M. P. Mehta,
21. P. Y. Wen,
22. L. M. DeAngelis,
23. M. D. Prados,
24. I. K. Mellinghoff,
25. T. F. Cloughesy,
26. P. S. Mischel

, *EGFR signaling through an Akt-SREBP-1–dependent, rapamycin-resistant pathway sensitizes glioblastomas to antilipogenic therapy. Sci. Signal. 2, ra82 (2009).*

Abstract/FREE Full TextGoogle Scholar

110. ↵

1. D. Guo,
2. F. Reinitz,
3. M. Youssef,
4. C. Hong,
5. D. Nathanson,
6. D. Akhavan,
7. D. Kuga,
8. A. N. Amzajerdi,
9. H. Soto,
10. S. Zhu,
11. I. Babic,
12. K. Tanaka,
13. J. Dang,
14. A. Iwanami,
15. B. Gini,
16. J. DeJesus,
17. D. D. Lisiero,
18. T. T. Huang,
19. R. M. Prins,
20. P. Y. Wen,
21. H. I. Robins,
22. M. D. Prados,
23. L. M. DeAngelis,
24. I. K. Mellinghoff,
25. M. P. Mehta,
26. C. D. James,
27. A. Chakravarti,
28. T. F. Cloughesy,
29. P. Tontonoz,
30. P. S. Mischel

, *An LXR agonist promotes glioblastoma cell death through inhibition of an EGFR/AKT/SREBP-1/LDLR–dependent pathway. Cancer Discov. 1, 442–456 (2011).*

Abstract/FREE Full TextGoogle Scholar

111. ↵

1. A. Stincone,
2. A. Prigione,

3. T. Cramer,
4. M. M. C. Wamelink,
5. K. Campbell,
6. E. Cheung,
7. V. Olin-Sandoval,
8. N.-M. Grüning,
9. A. Krüger,
10. M. Tauqeer Alam,
11. M. A. Keller,
12. M. Breitenbach,
13. K. M. Brindle,
14. J. D. Rabinowitz,
15. M. Ralser

, *The return of metabolism: Biochemistry and physiology of the pentose phosphate pathway. Biol. Rev. Camb. Philos. Soc. 90, 927–963 (2014).*

Google Scholar

112. ↵

1. M. G. Vander Heiden

, *Targeting cancer metabolism: A therapeutic window opens. Nat. Rev. Drug Discov. 10, 671–684 (2011).*

CrossRefPubMedGoogle Scholar

113. ↵

1. I. Ben-Sahra,
2. J. J. Howell,
3. J. M. Asara,
4. B. D. Manning

, *Stimulation of de novo pyrimidine synthesis by growth signaling through mTOR and S6K1. Science 339, 1323–1328 (2013).*

Abstract/FREE Full TextGoogle Scholar

114. ↵

1. M. P. Murphy

, *How mitochondria produce reactive oxygen species. Biochem. J. 417, 1–13 (2009).*

Abstract/FREE Full TextGoogle Scholar

115. ↵

1. M. D. Brand

, *The sites and topology of mitochondrial superoxide production. Exp. Gerontol. 45, 466–472 (2010).*

CrossRefPubMedWeb of ScienceGoogle Scholar

116. ↵

1. S. G. Rhee,
2. H. A. Woo,
3. I. S. Kil,
4. S. H. Bae

, *Peroxiredoxin functions as a peroxidase and a regulator and sensor of local peroxides. J. Biol. Chem. 287, 4403–4410 (2012).*

Abstract/FREE Full TextGoogle Scholar

117. ↵

1. A. G. Cox,
2. C. C. Winterbourn,
3. M. B. Hampton

, *Mitochondrial peroxiredoxin involvement in antioxidant defence and redox signalling. Biochem. J. 425, 313–325 (2010).*

Abstract/FREE Full TextGoogle Scholar

118. ↵

1. M. P. Murphy

, *Mitochondrial thiols in antioxidant protection and redox signaling: Distinct roles for glutathionylation and other thiol modifications. Antioxid. Redox Signal. 16, 476–495 (2012).*

CrossRefPubMedWeb of ScienceGoogle Scholar

119. ↵

1. T. Finkel

, *From sulfenylation to sulfhydration: What a thiolate needs to tolerate. Sci. Signal. 5, pe10 (2012).*

Abstract/FREE Full TextGoogle Scholar

120. ↵

1. E. C. Cheung,
2. P. Lee,
3. F. Ceteci,
4. C. Nixon,
5. K. Blyth,
6. O. J. Sansom,
7. K. H. Vousden

, *Opposing effects of TIGAR- and RAC1-derived ROS on Wnt-driven proliferation in the mouse intestine. Genes Dev. 30, 52–63 (2016).*

Abstract/FREE Full TextGoogle Scholar

121. 1. K. Irani,
2. Y. Xia,
3. J. L. Zweier,
4. S. J. Sollott,
5. C. J. Der,
6. E. R. Fearon,
7. M. Sundaresan,
8. T. Finkel,
9. P. J. Goldschmidt-Clermont

, *Mitogenic signaling mediated by oxidants in Ras-transformed fibroblasts. Science 275, 1649–1652 (1997).*

Abstract/FREE Full TextGoogle Scholar

122. ↵

1. N. S. Chandel,
2. E. Maltepe,
3. E. Goldwasser,
4. C. E. Mathieu,
5. M. C. Simon,
6. P. T. Schumacker

, *Mitochondrial reactive oxygen species trigger hypoxia-induced transcription. Proc. Natl. Acad. Sci. U.S.A. 95, 11715–11720 (1998).*

Abstract/FREE Full TextGoogle Scholar

123. ↵

1. A. L. Orr,
2. L. Vargas,
3. C. N. Turk,
4. J. E. Baaten,
5. J. T. Matzen,
6. V. J. Dardov,
7. S. J. Attle,
8. J. Li,
9. D. C. Quackenbush,
10. R. L. S. Goncalves,
11. I. V. Perevoshchikova,
12. H. M. Petrassi,
13. S. L. Meeusen,
14. E. K. Ainscow,

15. M. D. Brand

, *Suppressors of superoxide production from mitochondrial complex III. Nat. Chem. Biol. 11, 834–836 (2015).*

CrossRefPubMedGoogle Scholar

124. ↵

1. P. E. Porporato,
2. V. L. Payen,
3. J. Pérez-Escuredo,
4. C. J. De Saedeleer,
5. P. Danhier,
6. T. Copetti,
7. S. Dhup,
8. M. Tardy,
9. T. Vazeille,
10. C. Bouzin,
11. O. Feron,
12. C. Michiels,
13. B. Gallez,
14. P. Sonveaux

, *A mitochondrial switch promotes tumor metastasis. Cell Rep. 8, 754–766 (2014).*

CrossRefPubMedGoogle Scholar

125. ↵

1. J. M. Munson,
2. L. Fried,
3. S. A. Rowson,
4. M. Y. Bonner,
5. L. Karumbaiah,
6. B. Diaz,
7. S. A. Courtneidge,
8. U. G. Knaus,
9. D. J. Brat,
10. J. L. Arbiser,
11. R. V. Bellamkonda

, *Anti-invasive adjuvant therapy with imipramine blue enhances chemotherapeutic efficacy against glioma. Sci. Transl. Med. 4, 127ra36 (2012).*

Abstract/FREE Full TextGoogle Scholar

126. ↵

1. N. S. Chandel,
2. D. A. Tuveson

, *The promise and perils of antioxidants for cancer patients. N. Engl. J. Med. 371, 177–178 (2014).*

CrossRefPubMedWeb of ScienceGoogle Scholar

127. ↵

1. M. C. Jaramillo,
2. D. D. Zhang

, *The emerging role of the Nrf2–Keap1 signaling pathway in cancer. Genes Dev. 27, 2179–2191 (2013).*

Abstract/FREE Full TextGoogle Scholar

128. ↵

1. G. M. DeNicola,
2. F. A. Karreth,
3. T. J. Humpton,
4. A. Gopinathan,
5. C. Wei,
6. K. Frese,

7. D. Mangal,
8. K. H. Yu,
9. C. J. Yeo,
10. E. S. Calhoun,
11. F. Scrimieri,
12. J. M. Winter,
13. R. H. Hruban,
14. C. Iacobuzio-Donahue,
15. S. E. Kern,
16. I. A. Blair,
17. D. A. Tuveson

, *Oncogene-induced Nrf2 transcription promotes ROS detoxification and tumorigenesis. Nature 475, 106–109 (2011).*

CrossRefPubMedWeb of ScienceGoogle Scholar

129. ↵

1. G. M. DeNicola,
2. P.-H. Chen,
3. E. Mullarky,
4. J. A. Sudderth,
5. Z. Hu,
6. D. Wu,
7. H. Tang,
8. Y. Xie,
9. J. M. Asara,
10. K. E. Huffman,
11. I. I. Wistuba,
12. J. D. Minna,
13. R. J. DeBerardinis,
14. L. C. Cantley

, *NRF2 regulates serine biosynthesis in non–small cell lung cancer. Nat. Genet. 47, 1475–1481 (2015).*

CrossRefPubMedGoogle Scholar

130. ↵

1. J. Ye,
2. J. Fan,
3. S. Venneti,
4. Y.-W. Wan,
5. B. R. Pawel,
6. J. Zhang,
7. L. W. S. Finley,
8. C. Lu,
9. T. Lindsten,
10. J. R. Cross,
11. G. Qing,
12. Z. Liu,
13. M. C. Simon,
14. J. D. Rabinowitz,
15. C. B. Thompson

, *Serine catabolism regulates mitochondrial redox control during hypoxia. Cancer Discov. 4, 1406–1417 (2014).*

Abstract/FREE Full TextGoogle Scholar

131. ↵

1. I. S. Harris,
2. A. E. Treloar,
3. S. Inoue,
4. M. Sasaki,

5. C. Gorrini,
6. K. C. Lee,
7. K. Y. Yung,
8. D. Brenner,
9. C. B. Knobbe-Thomsen,
10. M. A. Cox,
11. A. Elia,
12. T. Berger,
13. D. W. Cescon,
14. A. Adeoye,
15. A. Brüstle,
16. S. D. Molyneux,
17. J. M. Mason,
18. W. Y. Li,
19. K. Yamamoto,
20. A. Wakeham,
21. H. K. Berman,
22. R. Khokha,
23. S. J. Done,
24. T. J. Kavanagh,
25. C.-W. Lam,
26. T. W. Mak

, *Glutathione and thioredoxin antioxidant pathways synergize to drive cancer initiation and progression. Cancer Cell 27, 211–222 (2015).*

CrossRefPubMedGoogle Scholar

132. ↵

1. D. J. Garama,
2. T. J. Harris,
3. C. L. White,
4. F. J. Rossello,
5. M. Abdul-Hay,
6. D. J. Gough,
7. D. E. Levy

, *A synthetic lethal interaction between glutathione synthesis and mitochondrial reactive oxygen species provides a tumor-specific vulnerability dependent on STAT3. Mol. Cell. Biol. 35, 3646–3656 (2015).*

Abstract/FREE Full TextGoogle Scholar

133. ↵

1. C. Gorrini,
2. I. S. Harris,
3. T. W. Mak

, *Modulation of oxidative stress as an anticancer strategy. Nat. Rev. Drug Discov. 12, 931–947 (2013).*

CrossRefPubMedGoogle Scholar

134. ↵

1. Y. Saito,
2. R. H. Chapple,
3. A. Lin,
4. A. Kitano,
5. D. Nakada

, *AMPK protects leukemia-initiating cells in myeloid leukemias from metabolic stress in the bone marrow. Cell Stem Cell 17, 585–596 (2015).*

PubMedGoogle Scholar

135. ↵

1. S.-M. Jeon,
2. N. S. Chandel,
3. N. Hay

, *AMPK regulates NADPH homeostasis to promote tumour cell survival during energy stress. Nature 485, 661–665 (2012).*

CrossRefPubMedWeb of ScienceGoogle Scholar

136. ↵

1. V. S. LeBleu,
2. J. T. O'Connell,
3. K. N. Gonzalez Herrera,
4. H. Wikman,
5. K. Pantel,
6. M. C. Haigis,
7. F. M. de Carvalho,
8. A. Damascena,
9. L. T. Domingos Chinen,
10. R. M. Rocha,
11. J. M. Asara,
12. R. Kalluri

, *PGC-1α mediates mitochondrial biogenesis and oxidative phosphorylation in cancer cells to promote metastasis. Nat. Cell Biol. 16, 992–1003 (2014).*

CrossRefPubMedGoogle Scholar

137. ↵

1. Z. T. Schafer,
2. A. R. Grassian,
3. L. Song,
4. Z. Jiang,
5. Z. Gerhart-Hines,
6. H. Y. Irie,
7. S. Gao,
8. P. Puigserver,
9. J. S. Brugge

, *Antioxidant and oncogene rescue of metabolic defects caused by loss of matrix attachment. Nature 461, 109–113 (2009).*

CrossRefPubMedWeb of ScienceGoogle Scholar

138. ↵

1. L. Jiang,
2. A. A. Shestov,
3. P. Swain,
4. C. Yang,
5. S. J. Parker,
6. Q. A. Wang,
7. L. S. Terada,
8. N. D. Adams,
9. M. T. McCabe,
10. B. Pietrak,
11. S. Schmidt,
12. C. M. Metallo,
13. B. P. Dranka,
14. B. Schwartz,
15. R. J. DeBerardinis

, *Reductive carboxylation supports redox homeostasis during anchorage-independent growth. Nature 532, 255–258 (2016).*

PubMedGoogle Scholar

139. ↵

1. A. Erez,
2. R. J. DeBerardinis

, *Metabolic dysregulation in monogenic disorders and cancer—Finding method in madness. Nat. Rev. Cancer 15, 440–448 (2015).*

CrossRefPubMedGoogle Scholar

140. ↵

1. E. L. Pearce,
2. M. C. Poffenberger,
3. C.-H. Chang,
4. R. G. Jones

, *Fueling immunity: Insights into metabolism and lymphocyte function. Science 342, 1242454 (2013).*

Abstract/FREE Full TextGoogle Scholar

141. ↵

1. K. Ito,
2. T. Suda

, *Metabolic requirements for the maintenance of self-renewing stem cells. Nat. Rev. Mol. Cell Biol. 15, 243–256 (2014).*

CrossRefPubMedGoogle Scholar

142. ↵

1. H. Shim,
2. C. Dolde,
3. B. C. Lewis,
4. C.-S. Wu,
5. G. Dang,
6. R. A. Jungmann,
7. R. Dalla-Favera,
8. C. V. Dang

, *c-Myc transactivation of LDH-A: Implications for tumor metabolism and growth. Proc. Natl. Acad. Sci. U.S.A. 94, 6658–6663 (1997).*

Abstract/FREE Full TextGoogle Scholar

143. ↵

1. V. R. Fantin,
2. J. St-Pierre,
3. P. Leder

, *Attenuation of LDH-A expression uncovers a link between glycolysis, mitochondrial physiology, and tumor maintenance. Cancer Cell 9, 425–434 (2006).*

CrossRefPubMedWeb of ScienceGoogle Scholar

144. ↵

1. A. Le,
2. C. R. Cooper,
3. A. M. Gouw,
4. R. Dinavahi,
5. A. Maitra,
6. L. M. Deck,
7. R. E. Royer,
8. D. L. Vander Jagt,
9. G. L. Semenza,
10. C. V. Dang

, *Inhibition of lactate dehydrogenase A induces oxidative stress and inhibits tumor progression. Proc. Natl. Acad. Sci. U.S.A. 107, 2037–2042 (2010).*

Abstract/FREE Full TextGoogle Scholar

145. ↵

1. H. Xie,
2. J.-. Hanai,

3. J.-G. Ren,
4. L. Kats,
5. K. Burgess,
6. P. Bhargava,
7. S. Signoretti,
8. J. Billiard,
9. K. J. Duffy,
10. A. Grant,
11. X. Wang,
12. P. K. Lorkiewicz,
13. S. Schatzman,
14. M. Bousamra II.,
15. A. N. Lane,
16. R. M. Higashi,
17. T. W. M. Fan,
18. P. P. Pandolfi,
19. V. P. Sukhatme,
20. P. Seth

, *Targeting lactate dehydrogenase-a inhibits tumorigenesis and tumor progression in mouse models of lung cancer and impacts tumor-initiating cells. Cell Metab. 19, 795–809 (2014).*

CrossRefPubMedWeb of ScienceGoogle Scholar

146. ↵

1. Y.-H. Wang,
2. W. J. Israelsen,
3. D. Lee,
4. V. W. C. Yu,
5. N. T. Jeanson,
6. C. B. Clish,
7. L. C. Cantley,
8. M. G. Vander Heiden,
9. D. T. Scadden

, *Cell-state-specific metabolic dependency in hematopoiesis and leukemogenesis. Cell 158, 1309–1323 (2014).*

CrossRefPubMedGoogle Scholar

147. ↵

1. R. Haas,
2. J. Smith,
3. V. Rocher-Ros,
4. S. Nadkarni,
5. T. Montero-Melendez,
6. F. D'Acquisto,
7. E. J. Bland,
8. M. Bombardieri,
9. C. Pitzalis,
10. M. Perretti,
11. F. M. Marelli-Berg,
12. C. Mauro

, *Lactate regulates metabolic and pro-inflammatory circuits in control of T cell migration and effector functions. PLOS Biol. 13, e1002202 (2015).*

CrossRefPubMedGoogle Scholar

148. ↵

1. O. R. Colegio,
2. N.-Q. Chu,
3. A. L. Szabo,

4. T. Chu,
5. A. M. Rhebergen,
6. V. Jairam,
7. N. Cyrus,
8. C. E. Brokowski,
9. S. C. Eisenbarth,
10. G. M. Phillips,
11. G. W. Cline,
12. A. J. Phillips,
13. R. Medzhitov

, *Functional polarization of tumour-associated macrophages by tumour-derived lactic acid. Nature 513, 559–563 (2014).*

CrossRefPubMedWeb of ScienceGoogle Scholar

149. ↵

1. O. D. K. Maddocks,
2. C. R. Berkers,
3. S. M. Mason,
4. L. Zheng,
5. K. Blyth,
6. E. Gottlieb,
7. K. H. Vousden

, *Serine starvation induces stress and p53-dependent metabolic remodelling in cancer cells. Nature 493, 542–546 (2013).*

CrossRefPubMedWeb of ScienceGoogle Scholar

150. ↵

1. C. F. Labuschagne,
2. N. J. F. van den Broek,
3. G. M. Mackay,
4. K. H. Vousden,
5. O. D. K. Maddocks

, *Serine, but not glycine, supports one-carbon metabolism and proliferation of cancer cells. Cell Rep. 7, 1248–1258 (2014).*

CrossRefPubMedWeb of ScienceGoogle Scholar

151. ↵

1. D. Kim,
2. B. P. Fiske,
3. K. Birsoy,
4. E. Freinkman,
5. K. Kami,
6. R. L. Possemato,
7. Y. Chudnovsky,
8. M. E. Pacold,
9. W. W. Chen,
10. J. R. Cantor,
11. L. M. Shelton,
12. D. Y. Gui,
13. M. Kwon,
14. S. H. Ramkissoon,
15. K. L. Ligon,
16. S. W. Kang,
17. M. Snuderl,
18. M. G. Vander Heiden,
19. D. M. Sabatini

, *SHMT2 drives glioma cell survival in ischaemia but imposes a dependence on glycine clearance. Nature 520, 363–367 (2015).*

CrossRefPubMedGoogle Scholar

152. ↵

1. R. Nilsson,
2. M. Jain,
3. N. Madhusudhan,
4. N. G. Sheppard,
5. L. Strittmatter,
6. C. Kampf,
7. J. Huang,
8. A. Asplund,
9. V. K. Mootha

, *Metabolic enzyme expression highlights a key role for MTHFD2 and the mitochondrial folate pathway in cancer. Nat. Commun. 5, 3128 (2014).*

CrossRefPubMedGoogle Scholar

153. ↵

1. S. E. Weinberg,
2. N. S. Chandel

, *Targeting mitochondria metabolism for cancer therapy. Nat. Chem. Biol. 11, 9–15 (2015).*

CrossRefPubMedGoogle Scholar

154. ↵

1. J. M. M. Evans,
2. L. A. Donnelly,
3. A. M. Emslie-Smith,
4. D. R. Alessi,
5. A. D. Morris

, *Metformin and reduced risk of cancer in diabetic patients. BMJ 330, 1304–1305 (2005).*

FREE Full TextGoogle Scholar

155. ↵

1. M. Buzzai,
2. R. G. Jones,
3. R. K. Amaravadi,
4. J. J. Lum,
5. R. J. DeBerardinis,
6. F. Zhao,
7. B. Viollet,
8. C. B. Thompson

, *Systemic treatment with the antidiabetic drug metformin selectively impairs p53-deficient tumor cell growth. Cancer Res. 67, 6745–6752 (2007).*

Abstract/FREE Full TextGoogle Scholar

156. 1. R. M. Memmott,
2. J. R. Mercado,
3. C. R. Maier,
4. S. Kawabata,
5. S. D. Fox,
6. P. A. Dennis

, *Metformin prevents tobacco carcinogen–induced lung tumorigenesis. Cancer Prev. Res. 3, 1066–1076 (2010).*

Abstract/FREE Full TextGoogle Scholar

157. ↵

1. A. Tomimoto,
2. H. Endo,
3. M. Sugiyama,
4. T. Fujisawa,
5. K. Hosono,
6. H. Takahashi,

7. N. Nakajima,
8. Y. Nagashima,
9. K. Wada,
10. H. Nakagama,
11. A. Nakajima

, *Metformin suppresses intestinal polyp growth in Apc$^{Min/+}$ mice. Cancer Sci. 99, 2136–2141 (2008).*

CrossRefPubMedGoogle Scholar

158. ↵

1. H. R. Bridges,
2. A. J. Y. Jones,
3. M. N. Pollak,
4. J. Hirst

, *Effects of metformin and other biguanides on oxidative phosphorylation in mitochondria. Biochem. J. 462, 475–487 (2014).*

Abstract/FREE Full TextGoogle Scholar

159. 1. M.-Y. El-Mir,
2. V. Nogueira,
3. E. Fontaine,
4. N. Avéret,
5. M. Rigoulet,
6. X. Leverve

, *Dimethylbiguanide inhibits cell respiration via an indirect effect targeted on the respiratory chain complex I. J. Biol. Chem. 275, 223–228 (2000).*

Abstract/FREE Full TextGoogle Scholar

160. ↵

1. M. R. Owen,
2. E. Doran,
3. A. P. Halestrap

, *Evidence that metformin exerts its anti-diabetic effects through inhibition of complex 1 of the mitochondrial respiratory chain. Biochem. J. 348 (Pt. 3), 607–614 (2000).*

AbstractGoogle Scholar

161. ↵

1. W. W. Wheaton,
2. S. E. Weinberg,
3. R. B. Hamanaka,
4. S. Soberanes,
5. L. B. Sullivan,
6. E. Anso,
7. A. Glasauer,
8. E. Dufour,
9. G. M. Mutlu,
10. G. R. S. Budigner,
11. N. S. Chandel

, *Metformin inhibits mitochondrial complex I of cancer cells to reduce tumorigenesis. Elife 3, e02242 (2014).*

Abstract/FREE Full TextGoogle Scholar

162. ↵

1. T. Griss,
2. E. E. Vincent,
3. R. Egnatchik,
4. J. Chen,
5. E. H. Ma,
6. B. Faubert,
7. B. Viollet,

8. R. J. DeBerardinis,
9. R. G. Jones

, *Metformin antagonizes cancer cell proliferation by suppressing mitochondrial-dependent biosynthesis. PLOS Biol. 13, e1002309 (2015).*

CrossRefPubMedGoogle Scholar

163. ↵

1. A. Emami Riedmaier,
2. P. Fisel,
3. A. T. Nies,
4. E. Schaeffeler,
5. M. Schwab

, *Metformin and cancer: From the old medicine cabinet to pharmacological pitfalls and prospects. Trends Pharmacol. Sci. 34, 126–135 (2013).*

CrossRefGoogle Scholar

164. ↵

1. M. Pollak

, *Overcoming drug development bottlenecks with repurposing: Repurposing biguanides to target energy metabolism for cancer treatment. Nat. Med. 20, 591–593 (2014).*

CrossRefPubMedGoogle Scholar

165. ↵

1. K. Birsoy,
2. R. Possemato,
3. F. K. Lorbeer,
4. E. C. Bayraktar,
5. P. Thiru,
6. B. Yucel,
7. T. Wang,
8. W. W. Chen,
9. C. B. Clish,
10. D. M. Sabatini

, *Metabolic determinants of cancer cell sensitivity to glucose limitation and biguanides. Nature 508, 108–112 (2014).*

CrossRefPubMedWeb of ScienceGoogle Scholar

166. ↵

1. D. B. Shackelford,
2. E. Abt,
3. L. Gerken,
4. D. S. Vasquez,
5. A. Seki,
6. M. Leblanc,
7. L. Wei,
8. M. C. Fishbein,
9. J. Czernin,
10. P. S. Mischel,
11. R. J. Shaw

, *LKB1 inactivation dictates therapeutic response of non-small cell lung cancer to the metabolism drug phenformin. Cancer Cell 23, 143–158 (2013).*

CrossRefPubMedWeb of ScienceGoogle Scholar

167. ↵

1. S.-P. Gravel,
2. L. Hulea,
3. N. Toban,
4. E. Birman,

5. M.-J. Blouin,
6. M. Zakikhani,
7. Y. Zhao,
8. I. Topisirovic,
9. J. St-Pierre,
10. M. Pollak

, *Serine deprivation enhances antineoplastic activity of biguanides. Cancer Res. 74, 7521–7533 (2014).*

Abstract/FREE Full TextGoogle Scholar

168. ↵

1. G. Karsli-Uzunbas,
2. J. Y. Guo,
3. S. Price,
4. X. Teng,
5. S. V. Laddha,
6. S. Khor,
7. N. Y. Kalaany,
8. T. Jacks,
9. C. S. Chan,
10. J. D. Rabinowitz,
11. E. White

, *Autophagy is required for glucose homeostasis and lung tumor maintenance. Cancer Discov. 4, 914–927 (2014).*

Abstract/FREE Full TextGoogle Scholar

169. ↵

1. J. Son,
2. C. A. Lyssiotis,
3. H. Ying,
4. X. Wang,
5. S. Hua,
6. M. Ligorio,
7. R. M. Perera,
8. C. R. Ferrone,
9. E. Mullarky,
10. N. Shyh-Chang,
11. Y. Kang,
12. J. B. Fleming,
13. N. Bardeesy,
14. J. M. Asara,
15. M. C. Haigis,
16. R. A. DePinho,
17. L. C. Cantley,
18. A. C. Kimmelman

, *Glutamine supports pancreatic cancer growth through a KRAS-regulated metabolic pathway. Nature 496, 101–105 (2013).*

CrossRefPubMedWeb of ScienceGoogle Scholar

170. ↵

1. Y. Xiang,
2. Z. E. Stine,
3. J. Xia,
4. Y. Lu,
5. R. S. O'Connor,
6. B. J. Altman,
7. A. L. Hsieh,
8. A. M. Gouw,
9. A. G. Thomas,

10. P. Gao,
11. L. Sun,
12. L. Song,
13. B. Yan,
14. B. S. Slusher,
15. J. Zhuo,
16. L. L. Ooi,
17. C. G. L. Lee,
18. A. Mancuso,
19. A. S. McCallion,
20. A. Le,
21. M. C. Milone,
22. S. Rayport,
23. D. W. Felsher,
24. C. V. Dang

, *Targeted inhibition of tumor-specific glutaminase diminishes cell-autonomous tumorigenesis. J. Clin. Invest. 125, 2293–2306 (2015).*

CrossRefPubMedGoogle Scholar

171. ↵

1. S. A. Comerford,
2. Z. Huang,
3. X. Du,
4. Y. Wang,
5. L. Cai,
6. A. K. Witkiewicz,
7. H. Walters,
8. M. N. Tantawy,
9. A. Fu,
10. H. C. Manning,
11. J. D. Horton,
12. R. E. Hammer,
13. S. L. McKnight,
14. B. P. Tu

, *Acetate dependence of tumors. Cell 159, 1591–1602 (2014).*

CrossRefPubMedWeb of ScienceGoogle Scholar

172. ↵

1. T. Mashimo,
2. K. Pichumani,
3. V. Vemireddy,
4. K. J. Hatanpaa,
5. D. K. Singh,
6. S. Sirasanagandla,
7. S. Nannepaga,
8. S. G. Piccirillo,
9. Z. Kovacs,
10. C. Foong,
11. Z. Huang,
12. S. Barnett,
13. B. E. Mickey,
14. R. J. DeBerardinis,
15. B. P. Tu,
16. E. A. Maher,
17. R. M. Bachoo

, *Acetate is a bioenergetic substrate for human glioblastoma and brain metastases. Cell 159, 1603–1614 (2014).*

CrossRefPubMedWeb of ScienceGoogle Scholar

173. ↵

1. J. A. Engelman,
2. L. Chen,
3. X. Tan,
4. K. Crosby,
5. A. R. Guimaraes,
6. R. Upadhyay,
7. M. Maira,
8. K. McNamara,
9. S. A. Perera,
10. Y. Song,
11. L. R. Chirieac,
12. R. Kaur,
13. A. Lightbown,
14. J. Simendinger,
15. T. Li,
16. R. F. Padera,
17. C. García-Echeverría,
18. R. Weissleder,
19. U. Mahmood,
20. L. C. Cantley,
21. K.-K. Wong

, *Effective use of PI3K and MEK inhibitors to treat mutant Kras G12D and PIK3CA H1047R murine lung cancers. Nat. Med. 14, 1351–1356 (2008).*

CrossRefPubMedWeb of ScienceGoogle Scholar

174. ↵

1. D. Herranz,
2. A. Ambesi-Impiombato,
3. J. Sudderth,
4. M. Sánchez-Martín,
5. L. Belver,
6. V. Tosello,
7. L. Xu,
8. A. A. Wendorff,
9. M. Castillo,
10. J. E. Haydu,
11. J. Márquez,
12. J. M. Matés,
13. A. L. Kung,
14. S. Rayport,
15. C. Cordon-Cardo,
16. R. J. DeBerardinis,
17. A. A. Ferrando

, *Metabolic reprogramming induces resistance to anti-NOTCH1 therapies in T cell acute lymphoblastic leukemia. Nat. Med. 21, 1182–1189 (2015).*

CrossRefPubMedGoogle Scholar

175. 1. A. Viale,
2. P. Pettazzoni,
3. C. A. Lyssiotis,
4. H. Ying,
5. N. Sánchez,
6. M. Marchesini,
7. A. Carugo,
8. T. Green,

9. S. Seth,
10. V. Giuliani,
11. M. Kost-Alimova,
12. F. Muller,
13. S. Colla,
14. L. Nezi,
15. G. Genovese,
16. A. K. Deem,
17. A. Kapoor,
18. W. Yao,
19. E. Brunetto,
20. Y. Kang,
21. M. Yuan,
22. J. M. Asara,
23. Y. A. Wang,
24. T. P. Heffernan,
25. A. C. Kimmelman,
26. H. Wang,
27. J. B. Fleming,
28. L. C. Cantley,
29. R. A. DePinho,
30. G. F. Draetta

, *Oncogene ablation-resistant pancreatic cancer cells depend on mitochondrial function. Nature 514, 628–632 (2014).*

CrossRefPubMedGoogle Scholar

176. ↵

1. P. Yuan,
2. K. Ito,
3. R. Perez-Lorenzo,
4. C. Del Guzzo,
5. J. H. Lee,
6. C.-H. Shen,
7. M. W. Bosenberg,
8. M. McMahon,
9. L. C. Cantley,
10. B. Zheng

, *Phenformin enhances the therapeutic benefit of BRAF^{V600E} inhibition in melanoma. Proc. Natl. Acad. Sci. U.S.A. 110, 18226–18231 (2013).*

Abstract/FREE Full TextGoogle Scholar

177. ↵

1. A. Roesch,
2. A. Vultur,
3. I. Bogeski,
4. H. Wang,
5. K. M. Zimmermann,
6. D. Speicher,
7. C. Körbel,
8. M. W. Laschke,
9. P. A. Gimotty,
10. S. E. Philipp,
11. E. Krause,
12. S. Pätzold,
13. J. Villanueva,
14. C. Krepler,
15. M. Fukunaga-Kalabis,

16. M. Hoth,
17. B. C. Bastian,
18. T. Vogt,
19. M. Herlyn

, *Overcoming intrinsic multidrug resistance in melanoma by blocking the mitochondrial respiratory chain of slow-cycling JARID1B^{high} cells. Cancer Cell 23, 811–825 (2013).*

CrossRefPubMedGoogle Scholar

178. ↵

1. A. Janzer,
2. N. J. German,
3. K. N. Gonzalez-Herrera,
4. J. M. Asara,
5. M. C. Haigis,
6. K. Struhl

, *Metformin and phenformin deplete tricarboxylic acid cycle and glycolytic intermediates during cell transformation and NTPs in cancer stem cells. Proc. Natl. Acad. Sci. U.S.A. 111, 10574–10579 (2014).*

Abstract/FREE Full TextGoogle Scholar

179. ↵

1. J. Yun,
2. E. Mullarky,
3. C. Lu,
4. K. N. Bosch,
5. A. Kavalier,
6. K. Rivera,
7. J. Roper,
8. I. I. C. Chio,
9. E. G. Giannopoulou,
10. C. Rago,
11. A. Muley,
12. J. M. Asara,
13. J. Paik,
14. O. Elemento,
15. Z. Chen,
16. D. J. Pappin,
17. L. E. Dow,
18. N. Papadopoulos,
19. S. S. Gross,
20. L. C. Cantley

, *Vitamin C selectively kills KRAS and BRAF mutant colorectal cancer cells by targeting GAPDH. Science 350, 1391–1396 (2015).*

Abstract/FREE Full TextGoogle Scholar

180. ↵

1. Q. Chen,
2. M. G. Espey,
3. A. Y. Sun,
4. C. Pooput,
5. K. L. Kirk,
6. M. C. Krishna,
7. D. B. Khosh,
8. J. Drisko,
9. M. Levine

, *Pharmacologic doses of ascorbate act as a prooxidant and decrease growth of aggressive tumor xenografts in mice. Proc. Natl. Acad. Sci. U.S.A. 105, 11105–11109 (2008).*

Abstract/FREE Full TextGoogle Scholar

181. ↵

1. Y. Ma,
2. J. Chapman,
3. M. Levine,
4. K. Polireddy,
5. J. Drisko,
6. Q. Chen

, *High-dose parenteral ascorbate enhanced chemosensitivity of ovarian cancer and reduced toxicity of chemotherapy. Sci. Transl. Med. 6, 222ra18 (2014).*

Abstract/FREE Full TextGoogle Scholar

182. ↵

1. A. Tagde,
2. H. Singh,
3. M. H. Kang,
4. C. P. Reynolds

, *The glutathione synthesis inhibitor buthionine sulfoximine synergistically enhanced melphalan activity against preclinical models of multiple myeloma. Blood Cancer J. 4, e229 (2014).*

CrossRefPubMedGoogle Scholar

183. ↵

1. A. Glasauer,
2. L. A. Sena,
3. L. P. Diebold,
4. A. P. Mazar,
5. N. S. Chandel

, *Targeting SOD1 reduces experimental non-small-cell lung cancer. J. Clin. Invest. 124, 117–128 (2014).*

CrossRefPubMedGoogle Scholar

184. ↵

1. O. Oivares,
2. J. H. M. Däbritz,
3. A. King,
4. E. Gottlieb,
5. C. Halsey

, *Research into cancer metabolomics: Towards a clinical metamorphosis. Semin. Cell Dev. Biol. 43, 52–64 (2015).*

CrossRefPubMedGoogle Scholar

185. ↵

1. J. M. Buescher,
2. M. R. Antoniewicz,
3. L. G. Boros,
4. S. C. Burgess,
5. H. Brunengraber,
6. C. B. Clish,
7. R. J. DeBerardinis,
8. O. Feron,
9. C. Frezza,
10. B. Ghesquiere,
11. E. Gottlieb,
12. K. Hiller,
13. R. G. Jones,
14. J. J. Kamphorst,
15. R. G. Kibbey,
16. A. C. Kimmelman,
17. J. W. Locasale,

18. S. Y. Lunt,
19. O. D. K. Maddocks,
20. C. Malloy,
21. C. M. Metallo,
22. E. J. Meuillet,
23. J. Munger

, *A roadmap for interpreting ^{13}C metabolite labeling patterns from cells. Curr. Opin. Biotechnol. 34, 189–201 (2015).*

CrossRefPubMedGoogle Scholar

186. ↵

1. R. J. DeBerardinis,
2. C. B. Thompson

, *Cellular metabolism and disease: What do metabolic outliers teach us? Cell 148, 1132–1144 (2012).*

CrossRefPubMedWeb of ScienceGoogle Scholar

187. ↵

1. O. C. Andronesi,
2. O. Rapalino,
3. E. Gerstner,
4. A. Chi,
5. T. T. Batchelor,
6. D. P. Cahill,
7. A. G. Sorensen,
8. B. R. Rosen

, *Detection of oncogenic IDH1 mutations using magnetic resonance spectroscopy of 2-hydroxyglutarate. J. Clin. Invest. 123, 3659–3663 (2013).*

CrossRefPubMedWeb of ScienceGoogle Scholar

188. ↵

1. C. Lussey-Lepoutre,
2. A. Bellucci,
3. A. Morin,
4. A. Buffet,
5. L. Amar,
6. M. Janin,
7. C. Ottolenghi,
8. F. Zinzindohoué,
9. G. Autret,
10. N. Burnichon,
11. E. Robidel,
12. B. Banting,
13. S. Fontaine,
14. C.-A. Cuenod,
15. P. Benit,
16. P. Rustin,
17. P. Halimi,
18. L. Fournier,
19. A.-P. Gimenez-Roqueplo,
20. J. Favier,
21. B. Tavitian

, *In vivo detection of succinate by magnetic resonance spectroscopy as a hallmark of SDHx mutations in paraganglioma. Clin. Cancer Res. 22, 1120–1129 (2016).*

Abstract/FREE Full TextGoogle Scholar

189. ↵

1. N. M. Vacanti,

2. A. S. Divakaruni,
3. C. R. Green,
4. S. J. Parker,
5. R. R. Henry,
6. T. P. Ciaraldi,
7. A. N. Murphy,
8. C. M. Metallo

, *Regulation of substrate utilization by the mitochondrial pyruvate carrier. Mol. Cell 56, 425–435 (2014).*

CrossRefPubMedGoogle Scholar

190. 1. C. Yang,
2. B. Ko,
3. C. T. Hensley,
4. L. Jiang,
5. A. T. Wasti,
6. J. Kim,
7. J. Sudderth,
8. M. A. Calvaruso,
9. L. Lumata,
10. M. Mitsche,
11. J. Rutter,
12. M. E. Merritt,
13. R. J. DeBerardinis

, *Glutamine oxidation maintains the TCA cycle and cell survival during impaired mitochondrial pyruvate transport. Mol. Cell 56, 414–424 (2014).*

CrossRefPubMedGoogle Scholar

191. ↵

1. J. C. Schell,
2. K. A. Olson,
3. L. Jiang,
4. A. J. Hawkins,
5. J. G. Van Vranken,
6. J. Xie,
7. R. A. Egnatchik,
8. E. G. Earl,
9. R. J. DeBerardinis,
10. J. Rutter

, *A role for the mitochondrial pyruvate carrier as a repressor of the Warburg effect and colon cancer cell growth. Mol. Cell 56, 400–413 (2014).*

CrossRefPubMedGoogle Scholar

192. ↵

1. I. Marin-Valencia,
2. C. Yang,
3. T. Mashimo,
4. S. Cho,
5. H. Baek,
6. X.-L. Yang,
7. K. N. Rajagopalan,
8. M. Maddie,
9. V. Vemireddy,
10. Z. Zhao,
11. L. Cai,
12. L. Good,
13. B. P. Tu,

14. K. J. Hatanpaa,
15. B. E. Mickey,
16. J. M. Matés,
17. J. M. Pascual,
18. E. A. Maher,
19. C. R. Malloy,
20. R. J. DeBerardinis,
21. Robert M. Bachoo

, *Analysis of tumor metabolism reveals mitochondrial glucose oxidation in genetically diverse human glioblastomas in the mouse brain in vivo. Cell Metab. 15, 827–837 (2012).*

CrossRefPubMedWeb of ScienceGoogle Scholar

193. ↵

1. M. O. Yuneva,
2. T. W. M. Fan,
3. T. D. Allen,
4. R. M. Higashi,
5. D. V. Ferraris,
6. T. Tsukamoto,
7. J. M. Matés,
8. F. J. Alonso,
9. C. Wang,
10. Y. Seo,
11. X. Chen,
12. J. M. Bishop

, *The metabolic profile of tumors depends on both the responsible genetic lesion and tissue type. Cell Metab. 15, 157–170 (2012).*

CrossRefPubMedWeb of ScienceGoogle Scholar

194. ↵

1. E. A. Maher,
2. I. Marin-Valencia,
3. R. M. Bachoo,
4. T. Mashimo,
5. J. Raisanen,
6. K. J. Hatanpaa,
7. A. Jindal,
8. F. M. Jeffrey,
9. C. Choi,
10. C. Madden,
11. D. Mathews,
12. J. M. Pascual,
13. B. E. Mickey,
14. C. R. Malloy,
15. R. J. DeBerardinis

, *Metabolism of [U-13 C]glucose in human brain tumors in vivo. NMR Biomed. 25, 1234–1244 2012).*

CrossRefPubMedWeb of ScienceGoogle Scholar

195. ↵

1. K. Sellers,
2. M. P. Fox,
3. M. Bousamra II.,
4. S. P. Slone,
5. R. M. Higashi,
6. D. M. Miller,
7. Y. Wang,
8. J. Yan,

9. M. O. Yuneva,
10. R. Deshpande,
11. A. N. Lane,
12. T. W.-M. Fan

, *Pyruvate carboxylase is critical for non-small-cell lung cancer proliferation. J. Clin. Invest. 125, 687–698 (2015).*

CrossRefPubMedGoogle Scholar

196. ↵

1. T. W. M. Fan,
2. A. N. Lane,
3. R. M. Higashi,
4. M. A. Farag,
5. H. Gao,
6. M. Bousamra,
7. D. M. Miller

, *Altered regulation of metabolic pathways in human lung cancer discerned by ^{13}C stable isotope-resolved metabolomics (SIRM). Mol. Cancer 8, 41 (2009).*

CrossRefPubMedGoogle Scholar

197. ↵

1. C. T. Hensley,
2. B. Faubert,
3. Q. Yuan,
4. N. Lev-Cohain,
5. E. Jin,
6. J. Kim,
7. L. Jiang,
8. B. Ko,
9. R. Skelton,
10. L. Loudat,
11. M. Wodzak,
12. C. Klimko,
13. E. McMillan,
14. Y. Butt,
15. M. Ni,
16. D. Oliver,
17. J. Torrealba,
18. C. R. Malloy,
19. K. Kernstine,
20. R. E. Lenkinski,
21. R. J. DeBerardinis

, *Metabolic heterogeneity in human lung tumors. Cell 164, 681–694 (2016).*

CrossRefPubMedGoogle Scholar

198. ↵

1. S. M. Davidson,
2. T. Papagiannakopoulos,
3. B. A. Olenchock,
4. J. E. Heyman,
5. M. A. Keibler,
6. A. Luengo,
7. M. R. Bauer,
8. A. K. Jha,
9. J. P. O'Brien,
10. K. A. Pierce,
11. D. Y. Gui,

12. L. B. Sullivan,
13. T. M. Wasylenko,
14. L. Subbaraj,
15. C. R. Chin,
16. G. Stephanopolous,
17. B. T. Mott,
18. T. Jacks,
19. C. B. Clish,
20. M. G. Vander Heiden

, *Environment impacts the metabolic dependencies of ras-driven non-small cell lung cancer. Cell Metab. 23, 517–528 (2016).*

CrossRefPubMedGoogle Scholar

199. ↵

1. K. Birsoy,
2. T. Wang,
3. W. W. Chen,
4. E. Freikman,
5. M. Abu-Remaileh,
6. D. M. Sabatini

, *An essential role of the mitochondrial electron transport chain in cell proliferation is to enable aspartate synthesis. Cell 162, 540–551 (2015).*

CrossRefPubMedGoogle Scholar

200. ↵

1. L. B. Sullivan,
2. D. Y. Gui,
3. A. M. Hosios,
4. L. N. Bush,
5. E. Freinkman,
6. M. G. Vander Hediden

, *Supporting aspartate biosynthesis is an essential function of respiration in proliferating cells. Cell 162, 552–563 (2015).*

CrossRefPubMedGoogle Scholar

Acknowledgments: We are grateful to J. Schaffer for illustrating the figures. **Funding:** This work was supported by NIH grants RO1 CA12306708 (to N.S.C.) and RO1 CA157996 to (R.J.D.). **Author contributions:** N.S.C. wrote the abstract, bioenergetics, redox, and targeting metabolism for cancer therapy sections. R.J.D. wrote the introduction, biosynthesis, technology, **Box 1**, and conclusion sections. N.S.C. and R.J.D. both wrote the metabolic reprogramming and oncometabolites sections. **Competing interests:** R.J.D. is on the Advisory Boards of Agios Pharmaceuticals and Peloton Therapeutics. N.S.C. declares no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

29.9: IMMUNOMETABOLISM- CELLULAR METABOLISM TURNS IMMUNE REGULATOR

Róisín M. Loftus and David K. Finlay. Immunometabolism: Cellular Metabolism Turns Immune Regulator. Journal of Biological Chemistry. **MINIREVIEWS**| VOLUME 291, ISSUE 1, P1-10, JANUARY 01, 2016. https://www.jbc.org/article/S0021-92...233-5/fulltext

ABSTRACT

Immune cells are highly dynamic in terms of their growth, proliferation, and effector functions as they respond to immunological challenges. Different immune cells can adopt distinct metabolic configurations that allow the cell to balance its requirements for energy, molecular biosynthesis, and longevity. However, in addition to facilitating immune cell responses, it is now becoming clear that cellular metabolism has direct roles in regulating immune cell function. This review article describes the distinct metabolic signatures of key immune cells, explains how these metabolic setups facilitate immune function, and discusses the emerging evidence that intracellular metabolism has an integral role in controlling immune responses.

METABOLIC CHALLENGES FACING IMMUNE CELLS

During the course of an immune response, immune cells can traverse multiple tissues containing diverse conditions of nutrient and oxygen availability. Additionally, in response to activation, immune cells often dramatically change their functional activities; a lymphocyte transforms from a relatively inert cell to a cell engaging in robust growth and proliferation, often producing large amounts of effector molecules such as cytokines. These microenvironmental and functional alterations represent significant metabolic stresses that are efficiently managed by immune cells due their ability to dynamically reprogram their cellular metabolism.

INFLAMMATORY MICROENVIRONMENTS

Most normal tissue is well vascularized and replete with nutrients and oxygen. However, during an immune response, conditions in the local immune microenvironment can often be significantly less accommodating due to competition for nutrients. For example, tumor cells have a prodigious appetite for glucose and other nutrients. As a result, the microenvironment within solid tumors can become depleted of glucose, resulting in decreased rates of glycolysis in tumor infiltrating lymphocytes (1, 2, 3). Bacterial infections can also compete for nutrients with immune cells. Infection with *Staphylococcus aureus*, a common human pathogen, can result in localized tissue hypoxia due to elevated levels of oxygen consumption by the invading bacteria. As glucose is a key fuel for this bacteria, the levels of glucose available to immune cells will also be reduced (4). Viral infection can also result in a decrease in the amount of glucose that is available to infiltrating immune cells; viruses can reprogram infected cells to up-regulate glucose uptake and metabolism to facilitate viral replication (5, 6, 7). Additionally, various cells at sites of inflammation can release enzymes that consume nutrients in the local microenvironment, including arginase and indoleamine-2,3-dioxygenase, which deplete arginine and tryptophan, respectively (1). Inflammatory sites can also become hypoxic due to the pronounced influx of inflammatory cells such as neutrophils and monocytes (8).

DYNAMIC CHANGES IN CELLULAR FUNCTION

Immune activation is accompanied by substantial changes in cellular activities, such as those accompanying T cell activation. Naïve T cells are long-lived, relatively inert, exhibit low levels of cellular biosynthesis, and primarily require ATP to meet cellular demands (Fig. 1*A*). Following activation, T cells undergo substantial changes in function and engage in robust cellular growth and rapid cellular proliferation (9). Essential in supporting these cellular activities is the provision of sufficient biomolecules (amino acids, nucleotides, lipids) for the biosynthesis of new cellular components. Therefore, in activated T cells, the objectives of cellular metabolism have shifted from primarily generating ATP to the generation of sufficient ATP plus large amounts of biomolecules for the generation of biomass (10). Therefore, immune cells adapt their cellular metabolism to accommodate altered functional outputs.

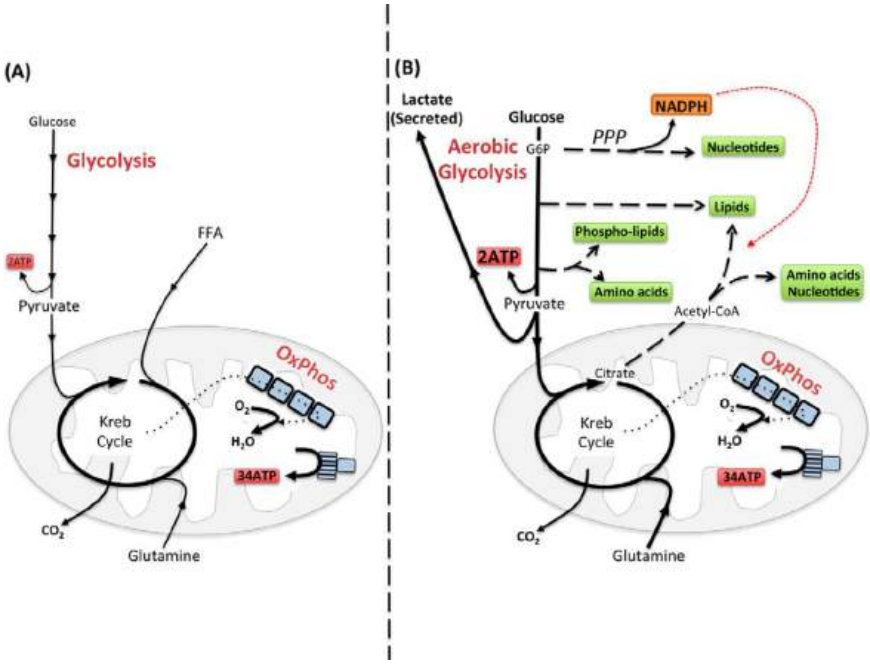


FIGURE 1.Configuring metabolism to match immune cell function. *A*, ATP is the key molecule that provides energy for cellular processes. Maintaining cellular ATP levels is essential for bioenergetic homeostasis and cell survival. Glucose, a key fuel source for mammalian cells, can be metabolized via two integrated metabolic pathways, glycolysis and OxPhos, that efficiently convert this simple sugar glucose into ATP. Glycolysis converts glucose to pyruvate through a series of enzymatic steps that occur in the cytosol, generating two molecules of ATP. Following its transportation into the mitochondria, pyruvate is further metabolized to CO_2 by the Krebs cycle, which drives OxPhos and the translocation of protons across the mitochondrial inner membrane. The resulting proton gradient drives the enzyme ATP synthase, converting ADP to ATP, generating up to 34 ATP per molecule of glucose. In addition to the breakdown of glucose via glycolysis, cells have the ability to metabolize alternative substrates, such as lipids and glutamine, which feed into the Krebs cycle and drive OxPhos. Fatty acid β-oxidation and glutaminolysis replenish the Krebs cycle intermediates acetyl-CoA and α-ketoglutarate, respectively, thereby fueling OxPhos and the efficient generation of cellular ATP. *B*, aerobic glycolysis supports biosynthetic processes of the cell as it allows the uptake of larger amounts of glucose and the maintenance of elevated glycolytic flux. Glycolytic intermediates are then diverted into various pathways for the synthesis of biomolecules that support biosynthetic processes. For instance, glucose 6-phosphate (*G6P*), generated by the first step in glycolysis, can feed into the pentose phosphate pathway (*PPP*) to support nucleotide synthesis. This pathway also generates NADPH, a cofactor that is essential for various biosynthetic processes including lipid synthesis. Glucose can also be converted into cytoplasmic acetyl-CoA via citrate in the Krebs cycle for the production of cholesterol and fatty acids for lipid synthesis. Other glycolytic intermediates can also be converted into biomolecules used for protein and lipid synthesis. During aerobic glycolysis, a significant proportion of pyruvate is also converted to lactate and secreted from the cell. Although aerobic glycolysis is an inefficient way to generate ATP (generating only two ATP molecules per glucose) due to the high rates of flux through the pathway, the rate of ATP production can be sufficient to maintain energy homeostasis even when mitochondrial ATP synthesis is impaired. Alternative fuels including glutamine feed into the Krebs cycle and can also supply biomolecules for biosynthetic processes under certain conditions.

CONFIGURING METABOLISM FOR BIOSYNTHESIS, INFLAMMATION, AND LONGEVITY

AEROBIC GLYCOLYSIS FOR CELLULAR BIOSYNTHESIS

A common feature of pro-inflammatory immune cells is that they adopt a distinct metabolic signature termed “aerobic glycolysis” to support cellular biosynthetic processes: that is, glucose metabolized to lactate in the presence of abundant oxygen (Fig. 1*B*). Aerobic glycolysis is adopted by cells engaging in robust growth and proliferation because it provides the biosynthetic precursors that are essential for the synthesis of nucleotides, amino acids, and lipids (10). Many intermediates of the glycolytic pathway act as a source of carbon that feeds into a range of biosynthetic pathways (Fig. 1*B*). Therefore, for cells engaged in aerobic glycolysis, the function of glucose is not just as a fuel to generate energy but also as a source of carbon that can be used for biosynthetic purposes (11). Hence, aerobic glycolysis provides immune cells with the components needed to facilitate proliferation and the synthesis of inflammatory molecules.

Metabolic reprogramming to aerobic glycolysis has advantages beyond enhanced biosynthetic capacity. This metabolic signature allows cells to adapt and survive as they encounter metabolically restrictive conditions, such as hypoxia. Although hypoxia prevents efficient ATP synthesis through OxPhos (Fig 2), high rates of glycolysis can generate enough ATP to maintain energy homeostasis. Glycolytic reprogramming involves increased expression of glucose transporters, especially Glut1, that facilitates elevated glucose uptake and enables immune cells to compete for glucose in nutrient restrictive environments (12). Immune cells also have a degree of metabolic plasticity in

response to limiting glucose availability. For instance, when glucose levels are low, effector T cells have the ability to adapt and increase glutamine uptake and glutaminolysis to support cellular metabolism (13).

AEROBIC GLYCOLYSIS IN ACTIVATED LYMPHOCYTES

Upon stimulation through antigen or cytokine receptors, lymphocytes increase the rates of both glycolysis and OxPhos (Fig. 2) (10). Although glucose is an essential fuel during T cell activation, glutamine is also important, and effector T cell differentiation is impaired when the supply of glutamine is disrupted (9, 14, 15). T cells that differentiate into effector subsets maintain aerobic glycolysis in response to various cytokines (16). In contrast, FoxP3$^+$ regulatory T cells (Tregs) switch to low levels of glycolysis and preferentially use oxidative metabolism (17). However, another type of regulatory T cell, FoxP3$^-$ regulatory T cells (Tr1), maintains elevated glycolysis similar to effector T cells (18). Although many of the functions of Tr1 cells overlap with those of Tregs, others are unique to Tr1 cells including granzyme/perforin-mediated cytolysis of target cells. Therefore, perhaps the distinct metabolic characteristics of these regulatory cells reflect the different mechanisms through which they regulate T cell responses. Similarly, B lymphocytes and NK cells also increase rates of glycolysis and OxPhos in response to various stimuli (19,20, 21,22). However, as metabolic analyses of B lymphocytes have all been performed using *in vitro* stimulated splenic B cells, the metabolic profile of distinct B cell subsets is currently unknown. Similarly, the metabolic signatures of distinct NK subsets, or indeed other innate lymphoid cells, also remain to be characterized.

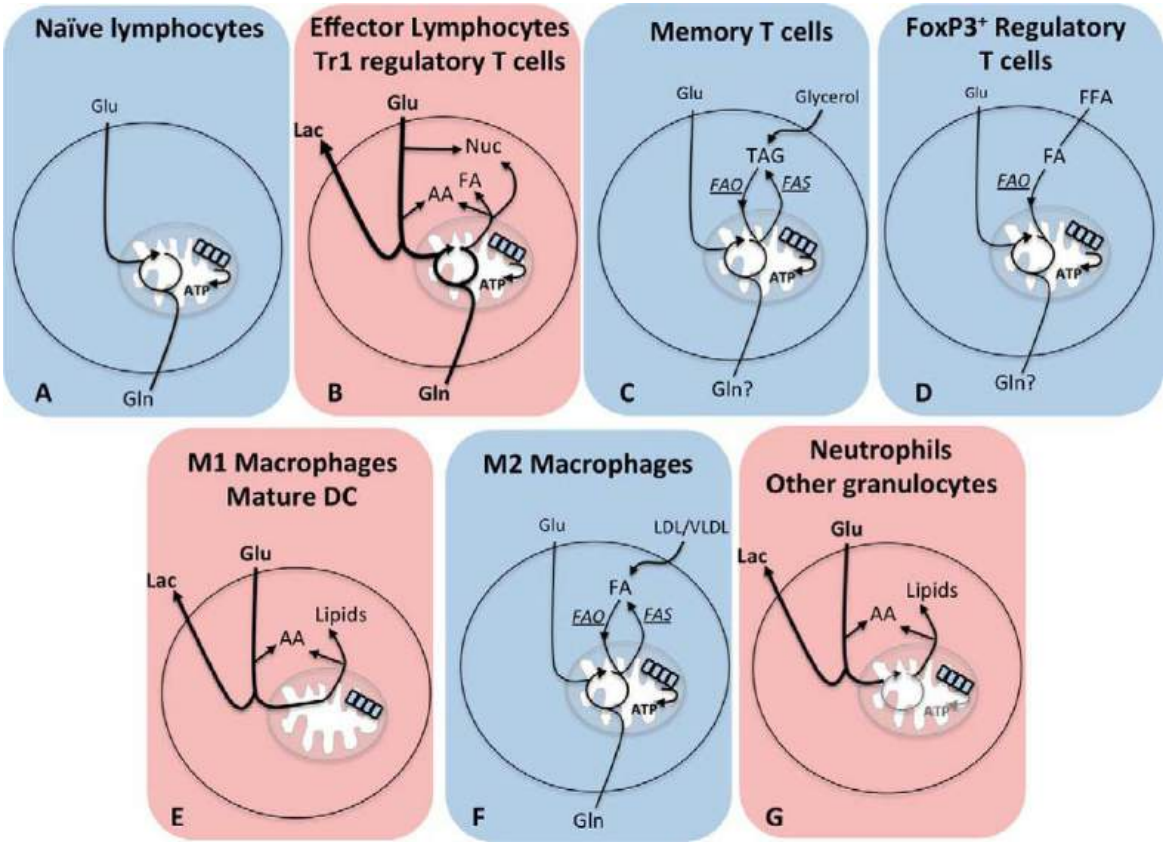


FIGURE 2.Distinct metabolic configurations of different immune cell subsets. *Blue panels* represent cells with oxidative metabolism, and *red panels* represent cells with glycolytic metabolism. *A*, naïve T cells use glucose and glutamine and OxPhos. *B*, effector lymphocytes and Tr1 regulatory T cells have high rates of both glycolysis and OxPhos, metabolize glucose to lactate, and use metabolic intermediates to support biosynthetic processes. *Nuc*, nucleotides; *FA*, fatty acids; *AA*, amino acids; *Lac*, Lactate. *C*, memory T cells use glucose to generate mitochondrial citrate, which is exported into the cytosol to support lipid synthesis. These *de novo* synthesized fatty acids are used with imported glycerol to generate and store TAGs. OxPhos is fueled by acetyl-CoA generated following β-oxidation of these TAGs. *FAO*, fatty acid oxidation; FAS, fatty acid synthesis. *D*, FoxP3$^+$ regulatory T cells use exogenously derived fatty acids metabolized by β-oxidation to support OxPhos. *E*, M1 macrophages and mature DC engage aerobic glycolysis for ATP synthesis and to support biosynthesis while also inactivating OxPhos. *F*, M2 macrophage metabolism is characterized by fatty acid β-oxidation and OxPhos. β-Oxidation is fueled by lipids that are scavenged from the external microenvironment and also by lipids generated by *de novo* fatty acid synthesis. *G*, neutrophils are highly glycolytic with few functional mitochondrial and very low rates of OxPhos.

Although the exact molecular mechanisms controlling glycolytic metabolism are not universal for all lymphocyte subsets, it is clear that mammalian target of rapamycin (mTOR) has a fundamental role (10, 23). mTOR complex 1 (mTORC1) activity is essential for the initial induction of glycolysis in T cells and is also required to maintain aerobic glycolysis in effector T cells subsets (9, 23, 24). The data also

suggest that mTORC1 has an important role for cytokine-induced glycolysis in NK cells (22). A number of transcription factors are involved in glycolytic reprogramming of T cells including both hypoxia-inducible factor 1α (HIF1α) and c-Myc (9, 25, 26). In B cells, c-Myc but not HIF1α is important for the glycolytic response (21). HIF1α and c-Myc directly bind the promoters of an array of genes, notably those of glycolytic enzymes and glucose transporters.

AEROBIC GLYCOLYSIS IN MYELOID CELLS

Unlike lymphocytes, mature myeloid cells tend to be non-proliferative and so have substantially different metabolic requirements. Activated M1 macrophages, dendritic cells (DC), and granulocytes are all highly glycolytic with little or no flux through OxPhos (27,28, 29, 30, 31,32). In activated M1 macrophages and DC, OxPhos is inactivated following inducible NOS-dependent nitric oxide production, which directly inhibits oxidative phosphorylation (33, 34). In these cells, the Krebs cycle is no longer cycling, which allows the repurposing of Krebs cycle enzymes to generate molecules that are important for proinflammatory functions (27, 33). M1 macrophages generate high levels of the Krebs cycle metabolite succinate, which can lead to increased HIF1α activity and sustained IL1β production (27). Levels of citrate are also elevated and are used to generate the antimicrobial metabolite itaconic acid that inhibits the growth of bacteria such as *Salmonella enterica* and *Mycobacterium tuberculosis* (34, 35). The metabolic changes following DC activation occur in two phases and result in a metabolic switch from fatty acid β-oxidation and OxPhos to glycolysis (33). An initial increase in glycolysis occurs within minutes of DC activation to support *de novo* lipid biosynthesis, facilitating the expansion of endoplasmic reticulum and Golgi apparatus and increasing the biosynthetic capacity that is essential for mature DC function (36). Over the course of 18 h, activated DC sustain elevated glycolysis and inactivate OxPhos (33). This metabolic shift is important in regulating DC-induced T cell responses, in part due to the fact that it impacts upon DC lifespan and thus the duration over which DC can activate T cells (37, 38). The metabolism of granulocytes is best characterized for neutrophils, which rely almost entirely on glycolysis and exhibit very low levels of OxPhos (28, 29,30, 39). Neutrophil effector functions, including the formation of neutrophil extracellular traps, require mTORC1/HIF1α signaling and glucose metabolism (29, 30, 39, 40,41). Although the metabolism of other granulocytes such as basophils and eosinophils remains poorly characterized, there is some evidence that these cells are also glycolytic and rely upon metabolic regulators such as HIF1α to maintain glycolysis and normal function (42). For instance, HIF1α accumulation upon basophil activation was shown to be required for VEGF and IL4 production (42).

OXIDATIVE CELLULAR METABOLISM IN NAÏVE LYMPHOCYTES AND MEMORY T CELLS

As previously mentioned, naïve lymphocytes are relatively inert cells with limited biosynthetic demands, and so ATP alone is relatively sufficient to sustain these cells. Given that these cells reside in well oxygenated tissues, oxidative metabolism is a consistent and efficient way to meet cellular metabolic demands. Memory cells generated during the course of an immune response share many of the same characteristics of naïve lymphocytes; they are long-lived, relatively inert cells with limited biosynthetic demands. As nothing is known regarding the metabolism of memory B cells, only memory T cells will be considered here. The key distinction between naïve and memory T cells is the rapid recall responses characteristic of memory T cells when compared with primary T cell responses. Although both naïve and memory T cells adopt oxidative metabolism, there are key differences in the metabolic configurations of these cells that contribute to rapid memory T cell recall responses. Memory T cells predominantly use fatty acid β-oxidation to generate acetyl-CoA to fuel OxPhos (43) (Fig. 2). β-Oxidation is an efficient method for generating ATP with each fatty acid molecule generating significantly more ATP (about 106 ATP/molecule of palmitate) when compared with one molecule of glucose (about 36 ATP/molecule of glucose). Indeed, fatty acid oxidation is essential for rapid memory T cell responses (43). Interestingly, these fatty acids are not taken up from the surrounding microenvironment, but rather memory T cells use glucose and glycolysis to generate citrate for *de novo* fatty acid synthesis and the generation and storage of triacylglycerides (TAGs) (44, 45). These endogenously derived TAGs are then broken down by β-oxidation in the mitochondria to generate acetyl-CoA to fuel OxPhos (45). From a bioenergetics standpoint, this would seem like an inefficient mechanism to fuel OxPhos as fatty acid synthesis utilizes both ATP and NADPH. Nonetheless, this seemingly futile cycle of fatty acid synthesis and fatty acid oxidation is important for memory T cell survival (44, 45). This approach may be taken by memory T cells, for which long term survival is of utmost importance, as glucose levels are stringently controlled in the blood, making glucose a more dependable fuel source than fatty acids, whose levels can vary in different tissues. Another advantage of this cycle of fatty acid synthesis and oxidation may be that it allows the cell to concurrently engage both glycolysis and OxPhos, thus maintaining the machinery required for rapid induction of metabolic flux through these pathways upon antigen recognition and so facilitating rapid functional responses. Indeed, memory T cells can induce rates of glycolysis much more rapidly and robustly than naïve T cells (43, 46).

OXIDATIVE CELLULAR METABOLISM IN CELLS WITH SIGNIFICANT BIOSYNTHETIC OUTPUT

$FoxP3^+$ Tregs also primarily engage in oxidative metabolism, but in contrast to naïve lymphocytes and memory T cells, $FoxP3^+$ Tregs are not inert cells and are in fact producing relatively large quantities of biomolecules (17, 47). Tregs make immunosuppressive cytokines IL10 and TGFβ and can also engage in cellular proliferation in response to IL2. In this respect, M2 macrophages are similar to $FoxP3^+$ Tregs; M2 macrophages engage in oxidative metabolism and yet have significant biosynthetic outputs. M2 macrophages have roles in tissue repair and secrete anti-inflammatory cytokines, growth factors, and factors involved in tissue remodeling (48). Tregs and M2 macrophages oxidize both glucose and fatty acids in the mitochondria to sustain OxPhos (17, 49,50, 51,52). In contrast to memory T cells, Tregs fuel β-oxidation and the Krebs cycle using exogenously derived fatty acids. Meanwhile, in M2 macrophages, there is evidence that both exogenously derived lipids scavenged from the microenvironment and *de novo* synthesized lipids fuel β-oxidation and OxPhos (52). It is likely that Tregs and M2

macrophages use glutamine metabolites to sustain cellular biosynthetic processes (Fig. 1) (53). Indeed, M2 macrophages have increased glutamine metabolism when compared with M1 macrophages (34). Additionally, given that M2 macrophages are professional scavengers of apoptotic debris, it is tempting to speculate that M2 macrophages sustain cellular biosynthesis using biomolecules scavenged from the surrounding microenvironment (48, 52).

OXIDATIVE METABOLISM SUPPORTS IMMUNE CELL LONGEVITY

Controlling the longevity of immune cells is an important aspect of a healthy immune system. For example, a long lifespan (years) is essential for naïve and memory T cells to maintain functional primary and recall T cell responses. In contrast, it is crucial that upon resolution of a viral infection, the large population of CTL undergoes apoptosis as these effector T cells have the potential to cause significant immunopathology (54). Therefore, CTL have a short lifespan of days to weeks. Similarly, differences in lifespan are apparent in different subsets of macrophages. M1 macrophages are short-lived and are a key component of the innate immune system that forms the first line of defense occurring within hours to days of an immunological challenge. In contrast, M2 macrophages are longer-lived as they have important roles within the resolution phase and in tissue repair and remodeling. Strikingly, the cellular metabolic signature of an immune cell corresponds to the longevity of the cell; aerobic glycolysis is characteristic of short-lived immune cells, whereas oxidative metabolism is characteristic of long-lived cells (Fig. 2).

It is perhaps unsurprising that OxPhos is important for longevity in immune cells given the importance of mitochondrial membrane potential in controlling the induction of apoptosis. Certainly, in activated DC, preserving OxPhos results in an increased cellular lifespan (38). Moreover, in macrophages, switching cellular metabolism from glycolysis to oxidative metabolism promotes a shift from short-lived M1 macrophages to longer-lived M2 macrophages (50). In addition, manipulating glycolytic *versus* oxidative metabolism impacts upon the formation of long-lived memory T cells; inhibiting glycolysis promotes memory T cell formation, whereas inhibiting fatty acid oxidation-dependent OxPhos represses memory T cell formation (55, 56). These reports are consistent with a number of other studies that also support the notion that promoting oxidative phosphorylation enhances cell survival and lifespan (57, 58,59). On the other hand, there are also numerous reports on a variety of cell types showing that manipulating glycolytic metabolism has profound impacts upon cellular viability (60,61, 62, 63,64). Growth factors that promote elevated levels of cellular glycolysis also have the consequence of making that cell highly dependent on continued growth factor signaling and glycolysis for survival (64). This provides an elegant mechanism for terminating effector T cell responses. For instance, glycolytic metabolism in $CD8^+$ CTL is sustained by IL2, and upon IL2 withdrawal, as will occur upon resolution of a viral infection, glycolytic metabolism is rapidly lost and the CTL will die (24, 65).

METABOLIC CONTROL OF IMMUNE CELL FUNCTION

METABOLIC ENZYMES OR REGULATORS CONTROLLING IMMUNE CELL FUNCTION

Cellular metabolism is crucial for facilitating immune cell functions, but in addition, there is emerging evidence that metabolic enzymes and regulators can also have a direct role in controlling immune cell functions. For instance, in CD4 T cells, GAPDH has been described to bind to the 3′-UTR of IFNγ and IL2 mRNA and inhibit translation (66). This function of GAPDH is perhaps unsurprising due to the numerous reports describing RNA binding activities for GAPDH over the past two decades (67, 68, 69,70). Indeed, in myeloid cells, GAPDH is a component of the IFNγ-activated inhibitor of translation (GAIT) complex that binds defined 3′-UTR elements within a family of inflammatory mRNAs and suppresses their translation (71). Importantly, GAPDH functions in glycolysis and mRNA binding are likely to be mutually exclusive so that in glycolytic cells, GAPDH is preferentially engaged in glycolysis, and thus the translation of IFNγ and IL2 mRNA is unconstrained. This mechanism provides a direct link between rates of glycolysis and the expression of important immunological effector molecules. Intriguingly, it appears that many other metabolic enzymes can bind to mRNA molecules including numerous glycolytic enzymes, Krebs cycle enzymes, and enzymes involved in other metabolic pathways (72). Although the specific mRNA transcripts that these metabolic enzymes bind to still have to be identified, this study highlights the abundant potential for cellular metabolism to directly impact upon cellular functions.

Various metabolic regulators that evolved to control cellular metabolic pathways have since acquired roles in directly controlling immune cell function. The glycolytic regulator HIF1α also promotes the expression of IL1β in M1 macrophages and programmed death ligand-1 (PD-L1), a ligand for the immune checkpoint receptor PD-1, on various myeloid cells (27, 73). The aryl-hydrocarbon receptor (AhR), which together with HIF1α controls glycolytic metabolism in Tr1 regulatory T cells, also directly regulates T cell responses. AhR promotes Th17 differentiation, while inhibiting Treg differentiation, and is required for the production of the Th17 cytokines IL17 and IL22 (74, 75, 76). Additionally, AhR is important for Tr1 regulatory T cell differentiation, directly promoting the expression of IL10 and IL21 (18, 77). The transcription factor sterol regulatory element-binding protein (Srebp), a central regulator fatty acid and cholesterol synthesis, has dual roles in controlling T cell metabolism and directly controlling genes required for immune function. $CD8^+$ T cells lacking Srebp activity fail to undergo metabolic reprogramming and blastogenesis and do not mount a functional T cell response (78). In $CD4^+$ T cells, the Srebp1c isoform is involved in Th17 differentiation and directly binds to the IL17 promotor to inhibit AhR-induced IL17 expression (79). Moreover, the Srebp1a isoform is required for pro-inflammatory functions in myeloid cells, including IL1β production, as it promotes the expression of a key component of the inflammasome, Nlrp1 (80). Therefore, there is growing evidence that multiple important regulators of cellular metabolism have additional functions in directly controlling immune responses.

METABOLITES CONTROLLING IMMUNE CELL FUNCTION

Distinct metabolic configurations will result in different levels of metabolites that can directly impact upon cellular function. It has recently been shown that the glycolytic intermediate phosphoenolpyruvate is important in sustaining T cell receptor (TCR) signaling and T cell effector functions. Phosphoenolpyruvate inhibits Ca^{2+} re-uptake into the endoplasmic reticulum, thus sustaining nuclear factor of activated T-cells (NFAT) signaling (2). Mitochondrial reactive oxygen species generated as a side product of OxPhos are also important for optimal TCR signal transduction. T cells that cannot produce mitochondrial reactive oxygen species fail to activate nuclear NFAT, produce IL2, or engage in proliferative expansion (81). In M1 macrophages, the levels of Krebs cycle metabolites are substantially altered, leading to dramatically elevated levels of succinate, the stabilization of HIF1α, and prolonged production of IL1β (27, 34). Succinate can stabilize HIF1α by inhibiting the α-ketoglutarate-dependent prolyl-hydroxylases responsible for tagging HIF1α for proteasomal degradation (27, 82, 83). Indeed, succinate can inhibit other α-ketoglutarate-dependent enzymes that can impact upon immune cells due their roles in controlling cellular epigenetics, namely TET2 DNA hydroxylates and Jumonji C (JmjC) domain-containing histone demethylases (discussed further below) (84, 85). Succinate can act as a signaling molecule that acts through the receptor SUCNR1 and can also be used as a substrate for the post-translational modification of proteins (that is, succinylation) (86). Succinate acting through SUCNR1 impacts upon DC functions and also induces DC chemotaxis to enhance DC-induced T cell responses (87). Numerous metabolic enzymes are succinylated on lysine residues, but at present, it is not clear whether this modification impacts upon the regulation of immune responses (86). Citrate levels are also elevated in M1 macrophages, and this metabolite is important for the production of various proinflammatory molecules: nitric oxide, reactive oxygen species, and prostaglandins (27, 88).

Cellular metabolites are also important substrates for various enzymes involved in the epigenetic control of gene expression via covalent modification of DNA and histones. Given that the distinct metabolic configurations that characterize immune cells result in different levels of these cellular metabolites, it follows that the epigenetic control of gene expression will differ in parallel with differences in metabolism. For example, TET family enzymes, which oxidize methylcytosine, leading to DNA demethylation, and JmjC domain-containing histone demethylases both require α-ketoglutarate as a substrate and are both inhibited by succinate (Fig. 3). Indeed, TET2 has recently been shown to regulate the expression of IFNγ, IL17a, and IL10 in Th1 and Th17 cells (89). Jmjd3 has been shown to be of particular importance in controlling gene expression in LPS-stimulated macrophages (90). Acetylation of histones is another post-translational modification that impacts on DNA structure and gene expression. Acetylation of histones by histone acetyl transferases (HATs) requires acetyl-CoA, which is supplied via the export of mitochondrial citrate (Fig. 3). Indeed, there is evidence in yeast that the concentration of acetyl-CoA is important for histone acetylation (91). Histone acetylation levels are also controlled by the rate of deacetylation. The activity of sirtuin histone deacetylases is linked to cellular metabolism as these deacetylases are sensitive to the ratio of oxidized NAD^+ to reduced NADH, which is affected by the balance of glycolysis and OxPhos (92). Oxidized NAD^+ is an essential substrate for sirtuins, whereas reduced NADH acts to inhibit sirtuin activity (Fig. 3) (93). In fact, sirtuins can also deacetylate targets other than histones, which are important in immune regulation. For example, Sirt1 deacetylates FoxP3 to inhibit Treg responses and RORγt to promote Th17 responses (94,95, 96,97). Additionally, sirtuins can also have a negative impact upon inflammatory responses, in part through inhibition of NFκB activity (98, 99). Although there are numerous studies suggesting that cellular metabolism impacts upon epigenetic programming of immune cells to affect immune cell fate and function, the best evidence of this comes from a study of trained immunity in macrophages. Cheng *et al.* (92) elegantly demonstrated that mTORC1/HIF1α-stimulated glycolysis is required for changes in the epigenome of human or murine myeloid cells that provides enhanced nonspecific protection from secondary infections. Therefore, it is clear that metabolites can impact directly on immune cell function, and it is likely that further examples of this will be revealed as the field of immunometabolism progresses.

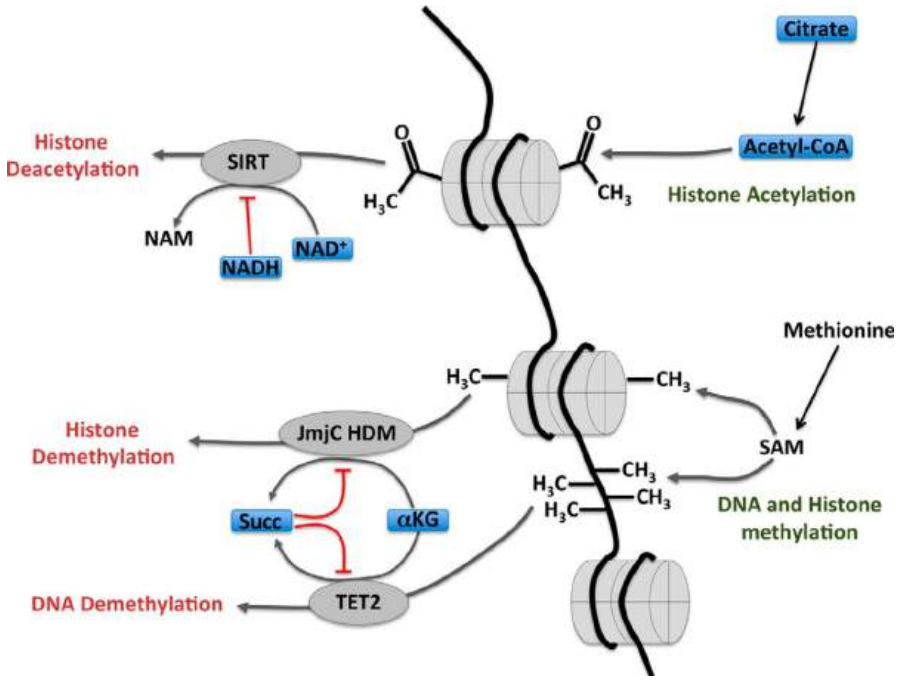


FIGURE 3.Links between cellular metabolism and epigenetic modifications. Histone deacetylation by sirtuin (SIRT) family members requires NAD^+ as a substrate, and the activity of these enzymes is inhibited by NADH. The balance of oxidized NAD^+ and reduced NADH is affected by levels of glycolysis and OxPhos. Methylation of DNA and histones is controlled by the rates of methylation and demethylation. The activities of JmjC domain-containing histone demethylases and the TET2 hydroxylase lead to histone and DNA demethylation, respectively, and can be regulated by Krebs cycle intermediates α-ketoglutarate (α-KG) and succinate (*Succ*). α-Ketoglutarate is a substrate for these enzymes, and succinate acts as an inhibitor. *NAM*, nicotinamide; *HDM*, histone demethylase; *SAM*, *S*-adenosylmethionine.

IMMUNE METABOLISM RELAYS EXTERNAL SIGNALS TO REGULATE IMMUNE CELL FUNCTION

The data now support an important role for cellular metabolism in controlling the function of immune cells. Given that metabolic regulators and pathways are acutely sensitive to external levels of nutrients, oxygen, and growth factors, cellular metabolism represents a means to relay information from the local microenvironment to modulate immune cell function accordingly. Nutrients such as glucose, glutamine, and fatty acids that directly supply metabolic pathways also regulate the activity of important regulators of immune metabolism and function including mTORC1, HIF1α, and Srebp. Other nutrients are important for providing the substrates for enzymes that impact upon immune cell function. For example, methionine, which is an essential amino acid and so must be imported into the cell, is used to generate *S*-adenosylmethionine for epigenetic methylation of DNA and histones. Although most studies have focused on how activating immune receptors affect cellular metabolism, it is now becoming apparent that ligation of inhibitory receptors also alters metabolic pathways. Recent research has demonstrated that ligation of the inhibitory receptors PD-1 and CTLA-4 expressed on human CD4 T cells has pronounced effects on cellular metabolism, inhibiting aerobic glycolysis, and in the case of PD-1, promoting fatty acid oxidation (100). These data suggest that the inhibitory actions of these receptors may be mediated, at least in part, due to changes in cellular metabolism.

FINAL COMMENTS

The emerging data now argue that metabolism has duel roles in immune cells to facilitate requirements for energy and biosynthesis and to directly regulate immune cell functions. There are likely to be numerous opportunities for novel therapeutic strategies that modulate this metabolic regulatory axis.

REFERENCES

1. Hirayama A.
   Kami K.
   Sugimoto M.
   Sugawara M.
   Toki N.
   Onozuka H.
   Kinoshita T.
   Saito N.

- Ochiai A.
- Tomita M.
- Esumi H.
- Soga T.

Quantitative metabolome profiling of colon and stomach cancer microenvironment by capillary electrophoresis time-of-flight mass spectrometry.

Cancer Res. 2009; **69**: 4918-4925View in Article

- Scopus (623)
- PubMed
- Crossref
- Google Scholar

2. - Ho P.C.
- Bihuniak J.D.
- Macintyre A.N.
- Staron M.
- Liu X.
- Amezquita R.
- Tsui Y.C.
- Cui G.
- Micevic G.
- Perales J.C.
- Kleinstein S.H.
- Abel E.D.
- Insogna K.L.
- Feske S.
- Locasale J.W.
- Bosenberg M.W.
- Rathmell J.C.
- Kaech S.M.

Phosphoenolpyruvate is a metabolic checkpoint of anti-tumor T cell responses.

Cell. 2015; **162**: 1217-1228View in Article

- Scopus (563)
- PubMed
- Abstract
- Full Text
- Full Text PDF
- Google Scholar

3. - Chang C.H.
- Qiu J.
- O'Sullivan D.
- Buck M.D.
- Noguchi T.
- Curtis J.D.
- Chen Q.
- Gindin M.
- Gubin M.M.
- van der Windt G.J.
- Tonc E.
- Schreiber R.D.
- Pearce E.J.
- Pearce E.L.

Metabolic competition in the tumor microenvironment is a driver of cancer progression.

Cell. 2015; **162**: 1229-1241View in Article

- Scopus (1015)
- PubMed
- Abstract
- Full Text
- Full Text PDF
- Google Scholar

4. Vitko N.P.
   Spahich N.A.
   Richardson A.R.

   Glycolytic dependency of high-level nitric oxide resistance and virulence in *Staphylococcus aureus*.

   MBio. 2015; **6**: e00045View in Article

   - Scopus (64)
   - PubMed
   - Crossref
   - Google Scholar

5. Ripoli M.
   D'Aprile A.
   Quarato G.
   Sarasin-Filipowicz M.
   Gouttenoire J.
   Scrima R.
   Cela O.
   Boffoli D.
   Heim M.H.
   Moradpour D.
   Capitanio N.
   Piccoli C.

   Hepatitis C virus-linked mitochondrial dysfunction promotes hypoxia-inducible factor 1 α-mediated glycolytic adaptation.

   J. Virol. 2010; **84**: 647-660View in Article

   - Scopus (0)
   - PubMed
   - Crossref
   - Google Scholar

6. Yu Y.
   Maguire T.G.
   Alwine J.C.

   Human cytomegalovirus activates glucose transporter 4 expression to increase glucose uptake during infection.

   J. Virol. 2011; **85**: 1573-1580View in Article

   - Scopus (89)
   - PubMed
   - Crossref
   - Google Scholar

7. Thai M.
   Graham N.A.
   Braas D.
   Nehil M.
   Komisopoulou E.
   Kurdistani S.K.
   McCormick F.
   Graeber T.G.
   Christofk H.R.

   Adenovirus E4ORF1-induced MYC activation promotes host cell anabolic glucose metabolism and virus replication.

   Cell Metab. 2014; **19**: 694-701View in Article

- Scopus (116)
- PubMed
- Abstract
- Full Text
- Full Text PDF
- Google Scholar

8. Taylor C.T.
   Colgan S.P.

   Hypoxia and gastrointestinal disease.

   J. Mol. Med. (Berl.). 2007; **85**: 1295-1300View in Article

   - Scopus (197)
   - PubMed
   - Crossref
   - Google Scholar

9. Wang R.
   Dillon C.P.
   Shi L.Z.
   Milasta S.
   Carter R.
   Finkelstein D.
   McCormick L.L.
   Fitzgerald P.
   Chi H.
   Munger J.
   Green D.R.

   The transcription factor Myc controls metabolic reprogramming upon T lymphocyte activation.

   Immunity. 2011; **35**: 871-882View in Article

   - Scopus (961)
   - PubMed
   - Abstract
   - Full Text
   - Full Text PDF
   - Google Scholar

10. Donnelly R.P.
    Finlay D.K.

    Glucose, glycolysis and lymphocyte responses.

    Mol. Immunol. 2015; (10.1016/j.molimm.2015.07.034)View in Article

    - Scopus (52)
    - PubMed
    - Crossref
    - Google Scholar

11. Vander Heiden M.G.
    Cantley L.C.
    Thompson C.B.

    Understanding the Warburg effect: the metabolic requirements of cell proliferation.

    Science. 2009; **324**: 1029-1033View in Article

    - Scopus (7897)
    - PubMed
    - Crossref
    - Google Scholar

12. Macintyre A.N.
    Gerriets V.A.

- Nichols A.G.
- Michalek R.D.
- Rudolph M.C.
- Deoliveira D.
- Anderson S.M.
- Abel E.D.
- Chen B.J.
- Hale L.P.
- Rathmell J.C.

The glucose transporter Glut1 is selectively essential for CD4 T cell activation and effector function.

Cell Metab. 2014; **20**: 61-72View in Article

- Scopus (461)
- PubMed
- Abstract
- Full Text
- Full Text PDF
- Google Scholar

13. - Blagih J.
- Coulombe F.
- Vincent E.E.
- Dupuy F.
- Galicia-Vázquez G.
- Yurchenko E.
- Raissi T.C.
- van der Windt G.J.
- Viollet B.
- Pearce E.L.
- Pelletier J.
- Piccirillo C.A.
- Krawczyk C.M.
- Divangahi M.
- Jones R.G.

The energy sensor AMPK regulates T cell metabolic adaptation and effector responses *in vivo*.

Immunity. 2015; **42**: 41-54View in Article

- Scopus (278)
- PubMed
- Abstract
- Full Text
- Full Text PDF
- Google Scholar

14. - Carr E.L.
- Kelman A.
- Wu G.S.
- Gopaul R.
- Senkevitch E.
- Aghvanyan A.
- Turay A.M.
- Frauwirth K.A.

Glutamine uptake and metabolism are coordinately regulated by ERK/MAPK during T lymphocyte activation.

J. Immunol. 2010; **185**: 1037-1044View in Article

- Scopus (334)
- PubMed
- Crossref

- Google Scholar

15. Sinclair L.V.
Rolf J.
Emslie E.
Shi Y.B.
Taylor P.M.
Cantrell D.A.

Control of amino-acid transport by antigen receptors coordinates the metabolic reprogramming essential for T cell differentiation.

Nat. Immunol. 2013; **14**: 500-508View in Article

- Scopus (408)
- PubMed
- Crossref
- Google Scholar

16. Gerriets V.A.
Rathmell J.C.

Metabolic pathways in T cell fate and function.

Trends Immunol. 2012; **33**: 168-173View in Article

- Scopus (207)
- PubMed
- Abstract
- Full Text
- Full Text PDF
- Google Scholar

17. Michalek R.D.
Gerriets V.A.
Jacobs S.R.
Macintyre A.N.
MacIver N.J.
Mason E.F.
Sullivan S.A.
Nichols A.G.
Rathmell J.C.

Cutting edge: distinct glycolytic and lipid oxidative metabolic programs are essential for effector and regulatory $CD4^+$ T cell subsets.

J. Immunol. 2011; **186**: 3299-3303View in Article

- Scopus (950)
- PubMed
- Crossref
- Google Scholar

18. Mascanfroni I.D.
Takenaka M.C.
Yeste A.
Patel B.
Wu Y.
Kenison J.E.
Siddiqui S.
Basso A.S.
Otterbein L.E.
Pardoll D.M.
Pan F.
Priel A.
Clish C.B.
Robson S.C.
Quintana F.J.

Metabolic control of type 1 regulatory T cell differentiation by AHR and HIF1-α.

Nat. Med. 2015; **21**: 638-646View in Article

- Scopus (218)
- PubMed
- Crossref
- Google Scholar

19. - Doughty C.A.
    - Bleiman B.F.
    - Wagner D.J.
    - Dufort F.J.
    - Mataraza J.M.
    - Roberts M.F.
    - Chiles T.C.

    Antigen receptor-mediated changes in glucose metabolism in B lymphocytes: role of phosphatidylinositol 3-kinase signaling in the glycolytic control of growth.

    Blood. 2006; **107**: 4458-4465View in Article

    - Scopus (166)
    - PubMed
    - Crossref
    - Google Scholar
20. - Dufort F.J.
    - Bleiman B.F.
    - Gumina M.R.
    - Blair D.
    - Wagner D.J.
    - Roberts M.F.
    - Abu-Amer Y.
    - Chiles T.C.

    Cutting edge: IL-4-mediated protection of primary B lymphocytes from apoptosis via Stat6-dependent regulation of glycolytic metabolism.

    J. Immunol. 2007; **179**: 4953-4957View in Article

    - PubMed
    - Crossref
    - Google Scholar
21. - Caro-Maldonado A.
    - Wang R.
    - Nichols A.G.
    - Kuraoka M.
    - Milasta S.
    - Sun L.D.
    - Gavin A.L.
    - Abel E.D.
    - Kelsoe G.
    - Green D.R.
    - Rathmell J.C.

    Metabolic reprogramming is required for antibody production that is suppressed in anergic but exaggerated in chronically BAFF-exposed B cells.

    J. Immunol. 2014; **192**: 3626-3636View in Article

    - Scopus (205)
    - PubMed
    - Crossref
    - Google Scholar
22. - Donnelly R.P.

- Loftus R.M.
- Keating S.E.
- Liou K.T.
- Biron C.A.
- Gardiner C.M.
- Finlay D.K.

mTORC1-dependent metabolic reprogramming is a prerequisite for NK cell effector function.

J. Immunol. 2014; **193**: 4477-4484View in Article

- Scopus (163)
- PubMed
- Crossref
- Google Scholar

23. - Pollizzi K.N.
- Powell J.D.

Integrating canonical and metabolic signalling programmes in the regulation of T cell responses.

Nat. Rev. Immunol. 2014; **14**: 435-446View in Article

- Scopus (229)
- PubMed
- Crossref
- Google Scholar

24. - Finlay D.K.
- Rosenzweig E.
- Sinclair L.V.
- Feijoo-Carnero C.
- Hukelmann J.L.
- Rolf J.
- Panteleyev A.A.
- Okkenhaug K.
- Cantrell D.A.

PDK1 regulation of mTOR and hypoxia-inducible factor 1 integrate metabolism and migration of $CD8^+$ T cells.

J. Exp. Med. 2012; **209**: 2441-2453View in Article

- Scopus (316)
- PubMed
- Crossref
- Google Scholar

25. - Oestreich K.J.
- Read K.A.
- Gilbertson S.E.
- Hough K.P.
- McDonald P.W.
- Krishnamoorthy V.
- Weinmann A.S.

Bcl-6 directly represses the gene program of the glycolysis pathway.

Nat. Immunol. 2014; **15**: 957-964View in Article

- Scopus (108)
- PubMed
- Crossref
- Google Scholar

26. - Shi L.Z.
- Wang R.
- Huang G.
- Vogel P.

- Neale G.
- Green D.R.
- Chi H.

HIF1α-dependent glycolytic pathway orchestrates a metabolic checkpoint for the differentiation of TH17 and Treg cells.

J. Exp. Med. 2011; **208**: 1367-1376View in Article

- Scopus (899)
- PubMed
- Crossref
- Google Scholar

27.
- Tannahill G.M.
- Curtis A.M.
- Adamik J.
- Palsson-McDermott E.M.
- McGettrick A.F.
- Goel G.
- Frezza C.
- Bernard N.J.
- Kelly B.
- Foley N.H.
- Zheng L.
- Gardet A.
- Tong Z.
- Jany S.S.
- Corr S.C.
- Haneklaus M.
- Caffrey B.E.
- Pierce K.
- Walmsley S.
- Beasley F.C.
- Cummins E.
- Nizet V.
- Whyte M.
- Taylor C.T.
- Lin H.
- Masters S.L.
- Gottlieb E.
- Kelly V.P.
- Clish C.
- Auron P.E.
- Xavier R.J.
- O'Neill L.A.

Succinate is an inflammatory signal that induces IL-1β through HIF-1α.

Nature. 2013; **496**: 238-242View in Article

- Scopus (1412)
- PubMed
- Crossref
- Google Scholar

28.
- Borregaard N.
- Herlin T.

Energy metabolism of human neutrophils during phagocytosis.

J. Clin. Invest. 1982; **70**: 550-557View in Article

- PubMed
- Crossref

- Google Scholar

29. Fossati G.
Moulding D.A.
Spiller D.G.
Moots R.J.
White M.R.
Edwards S.W.

The mitochondrial network of human neutrophils: role in chemotaxis, phagocytosis, respiratory burst activation, and commitment to apoptosis.

J. Immunol. 2003; **170**: 1964-1972View in Article

- Scopus (217)
- PubMed
- Crossref
- Google Scholar

30. Maianski N.A.
Geissler J.
Srinivasula S.M.
Alnemri E.S.
Roos D.
Kuijpers T.W.

Functional characterization of mitochondria in neutrophils: a role restricted to apoptosis.

Cell Death Differ. 2004; **11**: 143-153View in Article

- Scopus (228)
- PubMed
- Crossref
- Google Scholar

31. Newsholme P.
Curi R.
Gordon S.
Newsholme E.A.

Metabolism of glucose, glutamine, long-chain fatty acids and ketone bodies by murine macrophages.

Biochem. J. 1986; **239**: 121-125View in Article

- PubMed
- Crossref
- Google Scholar

32. Krawczyk C.M.
Holowka T.
Sun J.
Blagih J.
Amiel E.
DeBerardinis R.J.
Cross J.R.
Jung E.
Thompson C.B.
Jones R.G.
Pearce E.J.

Toll-like receptor-induced changes in glycolytic metabolism regulate dendritic cell activation.

Blood. 2010; **115**: 4742-4749View in Article

- Scopus (602)
- PubMed
- Crossref
- Google Scholar

33. Pearce E.J.
Everts B.

Dendritic cell metabolism.

Nat. Rev. Immunol. 2015; **15**: 18-29View in Article

Scopus (236)
PubMed
Crossref
Google Scholar

34. Jha A.K.
Huang S.C.
Sergushichev A.
Lampropoulou V.
Ivanova Y.
Loginicheva E.
Chmielewski K.
Stewart K.M.
Ashall J.
Everts B.
Pearce E.J.
Driggers E.M.
Artyomov M.N.

Network integration of parallel metabolic and transcriptional data reveals metabolic modules that regulate macrophage polarization.

Immunity. 2015; **42**: 419-430View in Article

Scopus (642)
PubMed
Abstract
Full Text
Full Text PDF
Google Scholar

35. Michelucci A.
Cordes T.
Ghelfi J.
Pailot A.
Reiling N.
Goldmann O.
Binz T.
Wegner A.
Tallam A.
Rausell A.
Buttini M.
Linster C.L.
Medina E.
Balling R.
Hiller K.

Immune-responsive gene 1 protein links metabolism to immunity by catalyzing itaconic acid production.

Proc. Natl. Acad. Sci. U.S.A. 2013; **110**: 7820-7825View in Article

Scopus (373)
PubMed
Crossref
Google Scholar

36. Everts B.
Amiel E.
Huang S.C.

- Smith A.M.
- Chang C.H.
- Lam W.Y.
- Redmann V.
- Freitas T.C.
- Blagih J.
- van der Windt G.J.
- Artyomov M.N.
- Jones R.G.
- Pearce E.L.
- Pearce E.J.

TLR-driven early glycolytic reprogramming via the kinases TBK1-IKKvarepsilon supports the anabolic demands of dendritic cell activation.

Nat. Immunol. 2014; **15**: 323-332View in Article

- Scopus (474)
- PubMed
- Crossref
- Google Scholar

37.
- Amiel E.
- Everts B.
- Freitas T.C.
- King I.L.
- Curtis J.D.
- Pearce E.L.
- Pearce E.J.

Inhibition of mechanistic target of rapamycin promotes dendritic cell activation and enhances therapeutic autologous vaccination in mice.

J. Immunol. 2012; **189**: 2151-2158View in Article

- Scopus (98)
- PubMed
- Crossref
- Google Scholar

38.
- Amiel E.
- Everts B.
- Fritz D.
- Beauchamp S.
- Ge B.
- Pearce E.L.
- Pearce E.J.

Mechanistic target of rapamycin inhibition extends cellular lifespan in dendritic cells by preserving mitochondrial function.

J. Immunol. 2014; **193**: 2821-2830View in Article

- Scopus (65)
- PubMed
- Crossref
- Google Scholar

39.
- Rodríguez-Espinosa O.
- Rojas-Espinosa O.
- Moreno-Altamirano M.M.
- López-Villegas E.O.
- Sánchez-García F.J.

Metabolic requirements for neutrophil extracellular traps formation.

Immunology. 2015; **145**: 213-224View in Article

- Scopus (105)

- PubMed
- Crossref
- Google Scholar

40. Azevedo E.P.
- Rochael N.C.
- Guimarães-Costa A.B.
- de Souza-Vieira T.S.
- Ganilho J.
- Saraiva E.M.
- Palhano F.L.
- Foguel D.

A metabolic shift toward pentose phosphate pathway is necessary for amyloid fibril- and phorbol 12-myristate 13-acetate-induced neutrophil extracellular trap (NET) formation.

J. Biol. Chem. 2015; **290**: 22174-22183View in Article

- Scopus (71)
- PubMed
- Abstract
- Full Text
- Full Text PDF
- Google Scholar

41. McInturff A.M.
- Cody M.J.
- Elliott E.A.
- Glenn J.W.
- Rowley J.W.
- Rondina M.T.
- Yost C.C.

Mammalian target of rapamycin regulates neutrophil extracellular trap formation via induction of hypoxia-inducible factor 1 α.

Blood. 2012; **120**: 3118-3125View in Article

- Scopus (0)
- PubMed
- Crossref
- Google Scholar

42. Sumbayev V.V.
- Nicholas S.A.
- Streatfield C.L.
- Gibbs B.F.

Involvement of hypoxia-inducible factor-1 HiF(1α) in IgE-mediated primary human basophil responses.

Eur. J. Immunol. 2009; **39**: 3511-3519View in Article

- Scopus (0)
- PubMed
- Crossref
- Google Scholar

43. van der Windt G.J.
- O'Sullivan D.
- Everts B.
- Huang S.C.
- Buck M.D.
- Curtis J.D.
- Chang C.H.
- Smith A.M.
- Ai T.
- Faubert B.

- Jones R.G.
- Pearce E.J.
- Pearce E.L.

CD8 memory T cells have a bioenergetic advantage that underlies their rapid recall ability.

Proc. Natl. Acad. Sci. U.S.A. 2013; **110**: 14336-14341View in Article

- Scopus (270)
- PubMed
- Crossref
- Google Scholar

44.
- Cui G.
- Staron M.M.
- Gray S.M.
- Ho P.C.
- Amezquita R.A.
- Wu J.
- Kaech S.M.

IL-7-induced glycerol transport and TAG synthesis promotes memory $CD8^+$ T cell longevity.

Cell. 2015; **161**: 750-761View in Article

- Scopus (160)
- PubMed
- Abstract
- Full Text
- Full Text PDF
- Google Scholar

45.
- O'Sullivan D.
- van der Windt G.J.
- Huang S.C.
- Curtis J.D.
- Chang C.H.
- Buck M.D.
- Qiu J.
- Smith A.M.
- Lam W.Y.
- DiPlato L.M.
- Hsu F.F.
- Birnbaum M.J.
- Pearce E.J.
- Pearce E.L.

Memory $CD8^+$ T cells use cell-intrinsic lipolysis to support the metabolic programming necessary for development.

Immunity. 2014; **41**: 75-88View in Article

- Scopus (359)
- PubMed
- Abstract
- Full Text
- Full Text PDF
- Google Scholar

46.
- Gubser P.M.
- Bantug G.R.
- Razik L.
- Fischer M.
- Dimeloe S.
- Hoenger G.
- Durovic B.

- Jauch A.
- Hess C.

Rapid effector function of memory CD8$^+$ T cells requires an immediate-early glycolytic switch.

Nat. Immunol. 2013; **14**: 1064-1072View in Article

- Scopus (240)
- PubMed
- Crossref
- Google Scholar

47.
- MacIver N.J.
- Michalek R.D.
- Rathmell J.C.

Metabolic regulation of T lymphocytes.

Annu. Rev. Immunol. 2013; **31**: 259-283View in Article

- Scopus (615)
- PubMed
- Crossref
- Google Scholar

48.
- Mantovani A.
- Biswas S.K.
- Galdiero M.R.
- Sica A.
- Locati M.

Macrophage plasticity and polarization in tissue repair and remodelling.

J. Pathol. 2013; **229**: 176-185View in Article

- Scopus (1126)
- PubMed
- Crossref
- Google Scholar

49.
- Gerriets V.A.
- Kishton R.J.
- Nichols A.G.
- Macintyre A.N.
- Inoue M.
- Ilkayeva O.
- Winter P.S.
- Liu X.
- Priyadharshini B.
- Slawinska M.E.
- Haeberli L.
- Huck C.
- Turka L.A.
- Wood K.C.
- Hale L.P.
- Smith P.A.
- Schneider M.A.
- MacIver N.J.
- Locasale J.W.
- Newgard C.B.
- Shinohara M.L.
- Rathmell J.C.

Metabolic programming and PDHK1 control CD4$^+$ T cell subsets and inflammation.

J. Clin. Invest. 2015; **125**: 194-207View in Article

- Scopus (300)

- PubMed
- Crossref
- Google Scholar

50. Tan Z.
Xie N.
Cui H.
Moellering D.R.
Abraham E.
Thannickal V.J.
Liu G.

Pyruvate dehydrogenase kinase 1 participates in macrophage polarization via regulating glucose metabolism.

J. Immunol. 2015; **194**: 6082-6089View in Article

- Scopus (128)
- PubMed
- Crossref
- Google Scholar

51. Vats D.
Mukundan L.
Odegaard J.I.
Zhang L.
Smith K.L.
Morel C.R.
Wagner R.A.
Greaves D.R.
Murray P.J.
Chawla A.

Oxidative metabolism and PGC-1β attenuate macrophage-mediated inflammation.

Cell Metab. 2006; **4**: 13-24View in Article

- Scopus (710)
- PubMed
- Abstract
- Full Text
- Full Text PDF
- Google Scholar

52. Huang S.C.
Everts B.
Ivanova Y.
O'Sullivan D.
Nascimento M.
Smith A.M.
Beatty W.
Love-Gregory L.
Lam W.Y.
O'Neill C.M.
Yan C.
Du H.
Abumrad N.A.
Urban Jr., J.F.
Artyomov M.N.
Pearce E.L.
Pearce E.J.

Cell-intrinsic lysosomal lipolysis is essential for alternative activation of macrophages.

Nat. Immunol. 2014; **15**: 846-855View in Article

- Scopus (472)
- PubMed
- Crossref
- Google Scholar

53. Maciolek J.A.
    Pasternak J.A.
    Wilson H.L.

    Metabolism of activated T lymphocytes.

    Curr. Opin. Immunol. 2014; **27**: 60-74View in Article

    - Scopus (98)
    - PubMed
    - Crossref
    - Google Scholar

54. Peiris J.S.
    Hui K.P.
    Yen H.L.

    Host response to influenza virus: protection versus immunopathology.

    Curr. Opin. Immunol. 2010; **22**: 475-481View in Article

    - Scopus (111)
    - PubMed
    - Crossref
    - Google Scholar

55. Pearce E.L.
    Walsh M.C.
    Cejas P.J.
    Harms G.M.
    Shen H.
    Wang L.S.
    Jones R.G.
    Choi Y.

    Enhancing CD8 T-cell memory by modulating fatty acid metabolism.

    Nature. 2009; **460**: 103-107View in Article

    - Scopus (885)
    - PubMed
    - Crossref
    - Google Scholar

56. Sukumar M.
    Liu J.
    Ji Y.
    Subramanian M.
    Crompton J.G.
    Yu Z.
    Roychoudhuri R.
    Palmer D.C.
    Muranski P.
    Karoly E.D.
    Mohney R.P.
    Klebanoff C.A.
    Lal A.
    Finkel T.
    Restifo N.P.
    Gattinoni L.

    Inhibiting glycolytic metabolism enhances $CD8^+$ T cell memory and antitumor function.

J. Clin. Invest. 2013; **123**: 4479-4488View in Article

- Scopus (399)
- PubMed
- Crossref
- Google Scholar

57. - Six E.
    - Lagresle-Peyrou C.
    - Susini S.
    - De Chappedelaine C.
    - Sigrist N.
    - Sadek H.
    - Chouteau M.
    - Cagnard N.
    - Fontenay M.
    - Hermine O.
    - Chomienne C.
    - Reynier P.
    - Fischer A.
    - André-Schmutz I.
    - Gueguen N.
    - Cavazzana M.

    AK2 deficiency compromises the mitochondrial energy metabolism required for differentiation of human neutrophil and lymphoid lineages.

    Cell Death Dis. 2015; **6**: e1856View in Article

    - Scopus (34)
    - PubMed
    - Crossref
    - Google Scholar

58. - Rivadeneira D.B.
    - Caino M.C.
    - Seo J.H.
    - Angelin A.
    - Wallace D.C.
    - Languino L.R.
    - Altieri D.C.

    Survivin promotes oxidative phosphorylation, subcellular mitochondrial repositioning, and tumor cell invasion.

    Sci. Signal. 2015; **8**: ra80View in Article

    - Scopus (46)
    - PubMed
    - Crossref
    - Google Scholar

59. - Maryanovich M.
    - Zaltsman Y.
    - Ruggiero A.
    - Goldman A.
    - Shachnai L.
    - Zaidman S.L.
    - Porat Z.
    - Golan K.
    - Lapidot T.
    - Gross A.

    An MTCH2 pathway repressing mitochondria metabolism regulates haematopoietic stem cell fate.

    Nat. Commun. 2015; **6**: 7901View in Article

- Scopus (99)
- PubMed
- Crossref
- Google Scholar

60. Leunda-Casi A.
    Genicot G.
    Donnay I.
    Pampfer S.
    De Hertogh R.

    Increased cell death in mouse blastocysts exposed to high d-glucose *in vitro*: implications of an oxidative stress and alterations in glucose metabolism.

    Diabetologia. 2002; **45**: 571-579View in Article

    - Scopus (0)
    - PubMed
    - Crossref
    - Google Scholar

61. MacFarlane M.
    Robinson G.L.
    Cain K.

    Glucose—a sweet way to die: metabolic switching modulates tumor cell death.

    Cell Cycle. 2012; **11**: 3919-3925View in Article

    - Scopus (22)
    - PubMed
    - Crossref
    - Google Scholar

62. Robinson G.L.
    Dinsdale D.
    Macfarlane M.
    Cain K.

    Switching from aerobic glycolysis to oxidative phosphorylation modulates the sensitivity of mantle cell lymphoma cells to TRAIL.

    Oncogene. 2012; **31**: 4996-5006View in Article

    - Scopus (37)
    - PubMed
    - Crossref
    - Google Scholar

63. Shim H.
    Chun Y.S.
    Lewis B.C.
    Dang C.V.

    A unique glucose-dependent apoptotic pathway induced by c-Myc.

    Proc. Natl. Acad. Sci. U.S.A. 1998; **95**: 1511-1516View in Article

    - Scopus (241)
    - PubMed
    - Crossref
    - Google Scholar

64. Vander Heiden M.G.
    Plas D.R.
    Rathmell J.C.
    Fox C.J.
    Harris M.H.
    Thompson C.B.

    Growth factors can influence cell growth and survival through effects on glucose metabolism.

Mol. Cell. Biol. 2001; **21**: 5899-5912View in Article

- Scopus (400)
- PubMed
- Crossref
- Google Scholar

65. - Lenardo M.
    - Chan K.M.
    - Hornung F.
    - McFarland H.
    - Siegel R.
    - Wang J.
    - Zheng L.

    Mature T lymphocyte apoptosis: immune regulation in a dynamic and unpredictable antigenic environment.

    Annu. Rev. Immunol. 1999; **17**: 221-253View in Article

    - Scopus (0)
    - PubMed
    - Crossref
    - Google Scholar
66. - Chang C.H.
    - Curtis J.D.
    - Maggi Jr., L.B.
    - Faubert B.
    - Villarino A.V.
    - O'Sullivan D.
    - Huang S.C.
    - van der Windt G.J.
    - Blagih J.
    - Qiu J.
    - Weber J.D.
    - Pearce E.J.
    - Jones R.G.
    - Pearce E.L.

    Posttranscriptional control of T cell effector function by aerobic glycolysis.

    Cell. 2013; **153**: 1239-1251View in Article

    - Scopus (934)
    - PubMed
    - Abstract
    - Full Text
    - Full Text PDF
    - Google Scholar
67. - Zang W.Q.
    - Fieno A.M.
    - Grant R.A.
    - Yen T.S.

    Identification of glyceraldehyde-3-phosphate dehydrogenase as a cellular protein that binds to the hepatitis B virus posttranscriptional regulatory element.

    Virology. 1998; **248**: 46-52View in Article

    - Scopus (0)
    - PubMed
    - Crossref
    - Google Scholar
68. - Schultz D.E.
    - Hardin C.C.

- Lemon S.M.

Specific interaction of glyceraldehyde 3-phosphate dehydrogenase with the 5′-nontranslated RNA of hepatitis A virus.

J. Biol. Chem. 1996; **271**: 14134-14142View in Article

- Scopus (120)
- PubMed
- Abstract
- Full Text
- Full Text PDF
- Google Scholar

69. - Nagy E.
- Rigby W.F.

Glyceraldehyde-3-phosphate dehydrogenase selectively binds AU-rich RNA in the NAD^+-binding region (Rossmann fold).

J. Biol. Chem. 1995; **270**: 2755-2763View in Article

- Scopus (291)
- PubMed
- Abstract
- Full Text
- Full Text PDF
- Google Scholar

70. - De B.P.
- Gupta S.
- Zhao H.
- Drazba J.A.
- Banerjee A.K.

Specific interaction *in vitro* and *in vivo* of glyceraldehyde-3-phosphate dehydrogenase and LA protein with cis-acting RNAs of human parainfluenza virus type 3.

J. Biol. Chem. 1996; **271**: 24728-24735View in Article

- Scopus (0)
- PubMed
- Abstract
- Full Text
- Full Text PDF
- Google Scholar

71. - Mukhopadhyay R.
- Jia J.
- Arif A.
- Ray P.S.
- Fox P.L.

The GAIT system: a gatekeeper of inflammatory gene expression.

Trends Biochem. Sci. 2009; **34**: 324-331View in Article

- Scopus (147)
- PubMed
- Abstract
- Full Text
- Full Text PDF
- Google Scholar

72. - Castello A.
- Fischer B.
- Eichelbaum K.
- Horos R.
- Beckmann B.M.
- Strein C.

- Davey N.E.
- Humphreys D.T.
- Preiss T.
- Steinmetz L.M.
- Krijgsveld J.
- Hentze M.W.

Insights into RNA biology from an atlas of mammalian mRNA-binding proteins.

Cell. 2012; **149**: 1393-1406View in Article

- Scopus (1072)
- PubMed
- Abstract
- Full Text
- Full Text PDF
- Google Scholar

73.
- Noman M.Z.
- Desantis G.
- Janji B.
- Hasmim M.
- Karray S.
- Dessen P.
- Bronte V.
- Chouaib S.

PD-L1 is a novel direct target of HIF-1α, and its blockade under hypoxia enhanced MDSC-mediated T cell activation.

J. Exp. Med. 2014; **211**: 781-790View in Article

- Scopus (788)
- PubMed
- Crossref
- Google Scholar

74.
- Kimura A.
- Naka T.
- Nohara K.
- Fujii-Kuriyama Y.
- Kishimoto T.

Aryl hydrocarbon receptor regulates Stat1 activation and participates in the development of Th17 cells.

Proc. Natl. Acad. Sci. U.S.A. 2008; **105**: 9721-9726View in Article

- Scopus (375)
- PubMed
- Crossref
- Google Scholar

75.
- Veldhoen M.
- Hirota K.
- Westendorf A.M.
- Buer J.
- Dumoutier L.
- Renauld J.C.
- Stockinger B.

The aryl hydrocarbon receptor links TH17-cell-mediated autoimmunity to environmental toxins.

Nature. 2008; **453**: 106-109View in Article

- Scopus (1146)
- PubMed
- Crossref
- Google Scholar

76. Quintana F.J.
Basso A.S.
Iglesias A.H.
Korn T.
Farez M.F.
Bettelli E.
Caccamo M.
Oukka M.
Weiner H.L.

Control of T_{reg} and T_H17 cell differentiation by the aryl hydrocarbon receptor.

Nature. 2008; **453**: 65-71View in Article

Scopus (1151)
PubMed
Crossref
Google Scholar

77. Apetoh L.
Quintana F.J.
Pot C.
Joller N.
Xiao S.
Kumar D.
Burns E.J.
Sherr D.H.
Weiner H.L.
Kuchroo V.K.

The aryl hydrocarbon receptor interacts with c-Maf to promote the differentiation of type 1 regulatory T cells induced by IL-27.

Nat. Immunol. 2010; **11**: 854-861View in Article

Scopus (460)
PubMed
Crossref
Google Scholar

78. Kidani Y.
Elsaesser H.
Hock M.B.
Vergnes L.
Williams K.J.
Argus J.P.
Marbois B.N.
Komisopoulou E.
Wilson E.B.
Osborne T.F.
Graeber T.G.
Reue K.
Brooks D.G.
Bensinger S.J.

Sterol regulatory element-binding proteins are essential for the metabolic programming of effector T cells and adaptive immunity.

Nat. Immunol. 2013; **14**: 489-499View in Article

Scopus (207)
PubMed
Crossref
Google Scholar

79. Cui G.
Qin X.

- Wu L.
- Zhang Y.
- Sheng X.
- Yu Q.
- Sheng H.
- Xi B.
- Zhang J.Z.
- Zang Y.Q.

Liver X receptor (LXR) mediates negative regulation of mouse and human Th17 differentiation.

J. Clin. Invest. 2011; **121**: 658-670View in Article

- Scopus (171)
- PubMed
- Crossref
- Google Scholar

80.
- Im S.S.
- Yousef L.
- Blaschitz C.
- Liu J.Z.
- Edwards R.A.
- Young S.G.
- Raffatellu M.
- Osborne T.F.

Linking lipid metabolism to the innate immune response in macrophages through sterol regulatory element binding protein-1a.

Cell Metab. 2011; **13**: 540-549View in Article

- Scopus (179)
- PubMed
- Abstract
- Full Text
- Full Text PDF
- Google Scholar

81.
- Sena L.A.
- Li S.
- Jairaman A.
- Prakriya M.
- Ezponda T.
- Hildeman D.A.
- Wang C.R.
- Schumacker P.T.
- Licht J.D.
- Perlman H.
- Bryce P.J.
- Chandel N.S.

Mitochondria are required for antigen-specific T cell activation through reactive oxygen species signaling.

Immunity. 2013; **38**: 225-236View in Article

- Scopus (542)
- PubMed
- Abstract
- Full Text
- Full Text PDF
- Google Scholar

82.
- Selak M.A.
- Armour S.M.
- MacKenzie E.D.

- Boulahbel H.
- Watson D.G.
- Mansfield K.D.
- Pan Y.
- Simon M.C.
- Thompson C.B.
- Gottlieb E.

Succinate links TCA cycle dysfunction to oncogenesis by inhibiting HIF-α prolyl hydroxylase.

Cancer Cell. 2005; **7**: 77-85View in Article

- Scopus (1275)
- PubMed
- Abstract
- Full Text
- Full Text PDF
- Google Scholar

83.
- Koivunen P.
- Hirsilä M.
- Remes A.M.
- Hassinen I.E.
- Kivirikko K.I.
- Myllyharju J.

Inhibition of hypoxia-inducible factor (HIF) hydroxylases by citric acid cycle intermediates: possible links between cell metabolism and stabilization of HIF.

J. Biol. Chem. 2007; **282**: 4524-4532View in Article

- Scopus (343)
- PubMed
- Abstract
- Full Text
- Full Text PDF
- Google Scholar

84.
- Xiao M.
- Yang H.
- Xu W.
- Ma S.
- Lin H.
- Zhu H.
- Liu L.
- Liu Y.
- Yang C.
- Xu Y.
- Zhao S.
- Ye D.
- Xiong Y.
- Guan K.L.

Inhibition of α-KG-dependent histone and DNA demethylases by fumarate and succinate that are accumulated in mutations of FH and SDH tumor suppressors.

Genes Dev. 2012; **26**: 1326-1338View in Article

- Scopus (516)
- PubMed
- Crossref
- Google Scholar

85.
- Smith E.H.
- Janknecht R.

- Maher 3rd, L.J.

Succinate inhibition of α-ketoglutarate-dependent enzymes in a yeast model of paraganglioma.

Hum. Mol. Genet. 2007; **16**: 3136-3148View in Article

- Scopus (119)
- PubMed
- Crossref
- Google Scholar

86. - Mills E.
    - O'Neill L.A.

Succinate: a metabolic signal in inflammation.

Trends Cell Biol. 2014; **24**: 313-320View in Article

- Scopus (243)
- PubMed
- Abstract
- Full Text
- Full Text PDF
- Google Scholar

87. - Rubic T.
    - Lametschwandtner G.
    - Jost S.
    - Hinteregger S.
    - Kund J.
    - Carballido-Perrig N.
    - Schwärzler C.
    - Junt T.
    - Voshol H.
    - Meingassner J.G.
    - Mao X.
    - Werner G.
    - Rot A.
    - Carballido J.M.

Triggering the succinate receptor GPR91 on dendritic cells enhances immunity.

Nat. Immunol. 2008; **9**: 1261-1269View in Article

- Scopus (217)
- PubMed
- Crossref
- Google Scholar

88. - Infantino V.
    - Convertini P.
    - Cucci L.
    - Panaro M.A.
    - Di Noia M.A.
    - Calvello R.
    - Palmieri F.
    - Iacobazzi V.

The mitochondrial citrate carrier: a new player in inflammation.

Biochem. J. 2011; **438**: 433-436View in Article

- Scopus (193)
- PubMed
- Crossref
- Google Scholar

89. - Ichiyama K.

- Chen T.
- Wang X.
- Yan X.
- Kim B.S.
- Tanaka S.
- Ndiaye-Lobry D.
- Deng Y.
- Zou Y.
- Zheng P.
- Tian Q.
- Aifantis I.
- Wei L.
- Dong C.

The methylcytosine dioxygenase Tet2 promotes DNA demethylation and activation of cytokine gene expression in T cells.

Immunity. 2015; **42**: 613-626View in Article

- Scopus (164)
- PubMed
- Abstract
- Full Text
- Full Text PDF
- Google Scholar

90. - De Santa F.
- Narang V.
- Yap Z.H.
- Tusi B.K.
- Burgold T.
- Austenaa L.
- Bucci G.
- Caganova M.
- Notarbartolo S.
- Casola S.
- Testa G.
- Sung W.K.
- Wei C.L.
- Natoli G.

Jmjd3 contributes to the control of gene expression in LPS-activated macrophages.

EMBO J. 2009; **28**: 3341-3352View in Article

- Scopus (289)
- PubMed
- Crossref
- Google Scholar

91. - Cai L.
- Sutter B.M.
- Li B.
- Tu B.P.

Acetyl-CoA induces cell growth and proliferation by promoting the acetylation of histones at growth genes.

Mol. Cell. 2011; **42**: 426-437View in Article

- Scopus (390)
- PubMed
- Abstract
- Full Text
- Full Text PDF
- Google Scholar

92. Cheng S.C.
Quintin J.
Cramer R.A.
Shepardson K.M.
Saeed S.
Kumar V.
Giamarellos-Bourboulis E.J.
Martens J.H.
Rao N.A.
Aghajanirefah A.
Manjeri G.R.
Li Y.
Ifrim D.C.
Arts R.J.
van der Veer B.M.
Deen P.M.
Logie C.
O'Neill L.A.
Willems P.
van de Veerdonk F.L.
van der Meer J.W.
Ng A.
Joosten L.A.
Wijmenga C.
Stunnenberg H.G.
Xavier R.J.
Netea M.G.

mTOR- and HIF-1α-mediated aerobic glycolysis as metabolic basis for trained immunity.

Science. 2014; **345**: 1250684View in Article

Scopus (657)
PubMed
Crossref
Google Scholar

93. Lin S.J.
Ford E.
Haigis M.
Liszt G.
Guarente L.

Calorie restriction extends yeast life span by lowering the level of NADH.

Genes Dev. 2004; **18**: 12-16View in Article

Scopus (511)
PubMed
Crossref
Google Scholar

94. Kwon H.S.
Lim H.W.
Wu J.
Schnölzer M.
Verdin E.
Ott M.

Three novel acetylation sites in the Foxp3 transcription factor regulate the suppressive activity of regulatory T cells.

J. Immunol. 2012; **188**: 2712-2721View in Article

Scopus (92)

- PubMed
- Crossref
- Google Scholar

95. Beier U.H.
Wang L.
Bhatti T.R.
Liu Y.
Han R.
Ge G.
Hancock W.W.

Sirtuin-1 targeting promotes Foxp3^{+} T-regulatory cell function and prolongs allograft survival.

Mol. Cell. Biol. 2011; **31**: 1022-1029View in Article

- Scopus (141)
- PubMed
- Crossref
- Google Scholar

96. van Loosdregt J.
Vercoulen Y.
Guichelaar T.
Gent Y.Y.
Beekman J.M.
van Beekum O.
Brenkman A.B.
Hijnen D.J.
Mutis T.
Kalkhoven E.
Prakken B.J.
Coffer P.J.

Regulation of Treg functionality by acetylation-mediated Foxp3 protein stabilization.

Blood. 2010; **115**: 965-974View in Article

- Scopus (274)
- PubMed
- Crossref
- Google Scholar

97. Lim H.W.
Kang S.G.
Ryu J.K.
Schilling B.
Fei M.
Lee I.S.
Kehasse A.
Shirakawa K.
Yokoyama M.
Schnölzer M.
Kasler H.G.
Kwon H.S.
Gibson B.W.
Sato H.
Akassoglou K.
Xiao C.
Littman D.R.
Ott M.
Verdin E.

SIRT1 deacetylates RORγt and enhances Th17 cell generation.

J. Exp. Med. 2015; **212**: 607-617View in Article

- Scopus (83)
- PubMed
- Crossref
- Google Scholar

98. - Yeung F.
    - Hoberg J.E.
    - Ramsey C.S.
    - Keller M.D.
    - Jones D.R.
    - Frye R.A.
    - Mayo M.W.

    Modulation of NF-κB-dependent transcription and cell survival by the SIRT1 deacetylase.

    EMBO J. 2004; **23**: 2369-2380View in Article

    - Scopus (1856)
    - PubMed
    - Crossref
    - Google Scholar
99. - Zhang J.
    - Lee S.M.
    - Shannon S.
    - Gao B.
    - Chen W.
    - Chen A.
    - Divekar R.
    - McBurney M.W.
    - Braley-Mullen H.
    - Zaghouani H.
    - Fang D.

    The type III histone deacetylase Sirt1 is essential for maintenance of T cell tolerance in mice.

    J. Clin. Invest. 2009; **119**: 3048-3058View in Article

    - Scopus (195)
    - PubMed
    - Crossref
    - Google Scholar
100. - Patsoukis N.
    - Bardhan K.
    - Chatterjee P.
    - Sari D.
    - Liu B.
    - Bell L.N.
    - Karoly E.D.
    - Freeman G.J.
    - Petkova V.
    - Seth P.
    - Li L.
    - Boussiotis V.A.

    PD-1 alters T-cell metabolic reprogramming by inhibiting glycolysis and promoting lipolysis and fatty acid oxidation.

    Nat. Commun. 2015; **6**: 6692View in Article

    - Scopus (412)
    - PubMed
    - Crossref
    - Google Scholar

29.10: ETHANOL METABOLISM- THE GOOD, THE BAD, AND THE UGLY

David F.WilsonFranz M.Matschinsky. Medical Hypotheses, Volume 140, July 2020, 109638
https://doi.org/10.1016/j.mehy.2020.109638. Under a Creative Commons license

ABSTRACT

Throughout the world, ethanol is both an important commercial commodity and a source of major medical and social problems. Ethanol readily passes through biological membranes and distributes throughout the body. It is oxidized, first to acetaldehyde and then to acetate, and finally by the citric acid cycle in virtually all tissues. The oxidation of ethanol is irreversible and unregulated, making the rate dependent only on local concentration and enzyme activity. This unregulated input of reducing equivalents increases reduction of both cytoplasmic and intramitochondrial NAD and, through the latter, cellular energy state {[ATP]/([ADP][Pi])}. In brain, this increase in energy state stimulates dopaminergic neural activity signalling reward and a sense of well being, while suppressing glutamatergic neural activity signalling anxiety and unease. These positive responses to ethanol ingestion are important to social alcohol consumption. Importantly, decreased free [AMP] decreases AMP-dependent protein kinase (AMPK) activity, an important regulator of cellular energy metabolism. Oxidation of substrates used for energy metabolism in the absence of ethanol is down regulated to accommodate the input from ethanol. In liver, chronic ethanol metabolism results in fatty liver and general metabolic dysfunction. In brain, transport of other oxidizable metabolites through the blood-brain barrier and the enzymes for their oxidation are both down regulated. For exposures of short duration, ethanol induced regulatory changes are rapid and reversible, recovering completely when the concentrations of ethanol and acetate fall again. Longer periods of ethanol exposure and associated chronic suppression of AMPK activity activates regulatory mechanisms, including gene expression, that operate over longer time scales, both in onset and reversal. If chronic alcohol consumption is abruptly ended, metabolism is no longer able to respond rapidly enough to compensate. Glutamatergic neural activity adapts to chronic dysregulation of glutamate metabolism and suppression of glutamatergic neural activity by increasing excitatory and decreasing inhibitory amino acid receptors. A point is reached (ethanol dependence) where withdrawal of ethanol results in significant metabolic energy depletion in neurons and other brain cells as well as hyperexcitation of the glutamatergic system. The extent and regional specificity of energy depletion in the brain, combined with hyperactivity of the glutamatergic neuronal system, largely determines the severity of withdrawal symptoms.

OVERVIEW: ETHANOL AND HUMAN METABOLISM

In nature, ethanol is rarely found in foodstuffs at levels where consumption can raise blood concentrations to the levels reached by humans during social drinking. When it does occur, such as when fruit ferments either on the tree or after falling to the ground, exposure to that food source is brief. As a result, the amount of ethanol consumed is a tiny fraction of that consumed by even modest social drinkers. Ethanol is a “rogue” nutrient for which there was no evolutionary pressure for its regulation. Ethanol dependence became a problem only after humans developed methods for preservation of food, one of which was fermentation. In order to provide a framework for understanding the importance of ethanol metabolism and its effect on metabolic homeostasis in humans, Table 1 summarizes some aspects of ethanol consumption important to understanding its impact. Dettling et al. [25] measured the rate of elimination of ethanol by humans following ingestion of 0.9 g/kg body weight over a period of 2 h, an amount that resulted in maximal blood alcohol levels near 0.08 g/dl (legal limit for driving in US). After consumption stopped, ethanol disappearance, measured by decrease in blood levels, occurred at a constant rate (zero order) of about 0.016 g/dl/h for men and 0.018 g/dl/h for women. Translated into kilocalories (cal), for an 80 kg man this is 73 cal/h while for 60 kg women it is 55 cal/h. A 70 kg (155 lb) person expends about 102 cal/h working on a computer, 120 cal/h sitting in meetings, 130 cal/h during desk work or sitting in class, 186 cal/h doing police work, bartending or waitressing (Table 2). Maintenance diets for average adults are 2000 cal/day for females (83 cal/h) and 2600 cal/day (108 cal/h) for males. As such, ethanol becomes responsible for more than 50% of each individual’s energy metabolism until it is eliminated. Elimination requires about 5hr after consumption stops (7h total) and throughout that time, ethanol oxidation is the largest carbon source for energy metabolism. Further increase in alcohol consumption, either acutely or chronically, extends the time before blood alcohol levels return to near zero but not the rate at which it is removed. For ethanol dependent individuals, blood alcohol levels remain significant throughout each day. This means ethanol contributes about 1700 cal/day (80 kg male) and 1300 cal/day (60 kg female), and many metabolic and nutrient deficiencies may develop [66], [69], [62], [63], [64].

Table 1. Ethanol metabolism and the human diet (social drinking).

0.9 g/kg (male)	Blood max 0.084 g/dl	Elimination 0.016 g/dl/h	time 7 h#; 5 h*
0.84 g/kg (female)	Blood max 0.082 g/dl	Elimination 0.018 g/dl/h	Time 6.6 h#; 4.6hr*
80 kg/male (176 lb)	Total, 72 g (5drinks)	511 calories	73 calories/h
60 kg/female (132 lb)	Total, 50 g (3.6 drinks)	355 calories	54 calories/h

Measurements of the rate of elimination of ethanol from the body in human volunteers [25]. Ethanol ingestion followed a typical breakfast and consisted of consuming drinks of the subject’s choice. The number of drinks was calculated for 14 g of pure ethanol per drink. The amount consumed raised blood alcohol to near 0.08 mg/dl (17 mM, legal limit for driving in US) by the end of the 2 h period for

consumption. After consumption stopped, ethanol disappeared from the blood at a nearly linear (zero order) rate of about 0.016 g/dl/h for men and 0.018 g/dl/h for women. The times required for blood alcohol to fall to near zero, h# and h*, were measured from beginning and end of consumption, respectively.

Table 2. Metabolic energy utilization by a 70 kg (155 lb) person.

Sitting (computer work)	102 calories/h
Sitting in meetings	120 calories/h
Desk work or sitting in class	130 calories/h
Police work, bartending, waitressing	186 calories/h

As noted in Table 1, for a 70 kg person ethanol contributes about 63 calories/h while the total hourly caloric expenditure is shown in the table. Assuming a drinker is awake and at activities similar to sitting in class both while and after drinking, they would be expected to expend about 130 calories/h. In that case, ethanol would provide about 50% of whole body energy metabolism until the blood alcohol fell below about 0.006 g/dl (below the K_M for ADH1 of 1.4 mM).

Oxidation of ethanol occurs through three enzymes, alcohol dehydrogenase (ADH), catalase, and P450 (CYP2E1), and they all produce acetaldehyde. In the present paper, focus is on the first two because they are responsible for oxidation of most of the consumed ethanol. Ethanol readily passes through most biological membranes, and the concentration in all water spaces, including the brain, approximates that in blood plasma. Thus, tissues and cells throughout the body are nearly equally exposed to ethanol and its rate of metabolism is determined by local enzyme content. The reactions of the three ethanol oxidizing systems and the K_M of each for ethanol are:

1. Alcohol dehydrogenase; ADH1, K_M = 1.4 mM [12]:

(1)Ethanol + NAD^+ → Acetaldehyde + NADH + H^+

2. Catalase; K_M = 12 mM [103]:

(2)Ethanol + H_2O_2 → Acetaldehyde + 2 H_2O

3. Cytochrome P450 2E1 (CYP2E1); K_M = 8–10 mM (12):

(3)Ethanol + O_2 + NADPH → $NADP^+$ + acetaldehyde + acetate

The equilibrium constant for the ADH reaction at pH 7.0 is approximately 10^{-4} [1], [6], [83]. In the cytoplasm, where the reaction occurs, the [NAD^+]/[NADH] is greater than 100 and acetaldehyde concentrations remain less than 20 μM [83]. The [NAD^+)/(NADH) is high enough that the ratio of the forward to reverse reactions of ADH is about 10 and the net reaction is strongly toward acetaldehyde (and NADH) formation. The reactions of catalase and CYP2E1 are irreversible under all physiological conditions. Acetaldehyde is chemically reactive and can react nonenzymatically with other cellular components to form products that are metabolically active and/or cytotoxic. The concentration of acetaldehyde is kept low through rapid removal by aldehyde dehydrogenase (ALDH) which is widely distributed in tissues [12], [127], [128]. ALDH oxidizes acetaldehyde to acetate, passing the reducing equivalents to NAD^+:(4)Acetaldehyde + NAD^+ → Acetate + NADH + H^+

ALDH2 is localized in the mitochondrial matrix, has a high affinity for acetaldehyde (K_M = 1.3 μM in brain cortex [40], and high specific activity. The equilibrium constant for this reaction at pH 7.0 is approximately 4×10^9, and it is irreversible under physiological conditions [1]. As a result of ALDH activity, systemic levels of acetaldehyde are kept low [40]. Acetaldehyde is a chemically reactive compound with significant toxicity and, for individuals who express an inactive form of ALDH2, ethanol consumption can have side effects ranging from unpleasant sensations to serious illness [12], [60]. Neither ethanol oxidation (ADH, catalase, CYP2E1) nor acetaldehyde oxidation (ALDH) is subject to significant regulation and increasing alcohol concentration reduces both cytoplasmic (by ADH) and mitochondrial (by ALDH2) NAD pools. Because ethanol oxidation is not regulated, local (individual cell) rates of oxidation are cell specific, determined by the activity and degree of saturation of the responsible enzyme(s) in each cell.

Acetate, the major product of ethanol oxidation, has roles both in energy metabolism and in regulation of metabolism. Acetate is converted to acetylCoA by acetylCoA synthetase:(5)acetate + CoASH + ATP → acetylCoA + AMP + pyrophosphate

In mitochondria, acetylCoA primarily enters the citric acid cycle and is oxidized, but at the cellular level is also used to acetylate proteins, altering the activity of individual enzymes and, through histone acetylation, gene expression [74]. The regulatory role of protein acetylation, although important, is outside the scope of this paper and will not be further discussed.

In summary: Ethanol readily distributes throughout the body and can be a major fraction of caloric utilization. Oxidation of ethanol and its products, acetaldehyde and acetate, is irreversible and unregulated. This “pushes” energy metabolism and can increase reduction of both cytoplasmic and mitochondrial NAD, thereby increasing cellular energy state. The rate of ethanol oxidation, and its associated metabolic disturbance, is dependent on ethanol concentration and local (cellular) enzyme content.

ETHANOL OXIDATION IN LIVER AND ITS METABOLIC CONSEQUENCES

Most of consumed alcohol is oxidized in liver by ADH1 to acetate [52], [67]. In the process, there are profound alterations in hepatic metabolism (Fig. 1, including: excessive reduction of cytoplasmic and mitochondrial NAD pools, inhibition of gluconeogenesis and fatty acid oxidation, increased [ATP]/([ADP]$_f$[Pi]), and decrease in [AMP]$_f$. The subscript f is used to emphasis that the concentrations of ADP and AMP relevant to metabolic regulation are free concentrations, not total amounts in the cell. Lundquist and coworkers [67] reported that, in healthy humans infused iv with ethanol at rates sufficient to result in blood concentrations near 26 mM (0.12 g/dL) oxidation of ethanol to acetate accounted for most of hepatic oxygen consumption. Although some of the acetate produced in liver was further oxidized through the citric acid cycle, most (>75%) was released into the blood. The rate of NADH production from ethanol oxidation to acetate is high enough that flux through the citric acid cycle (CAC), and therefore CO_2 production by the liver, is substantially decreased [52]. The lactate/pyruvate ratio is increased, indicating reduction of cytosolic NAD, as is the β-hydroxybutyrate/acetoacetate ratio, indicating increased reduction of intramitochondrial NAD [29], [52], [67]. Increase in intramitochondrial NAD reduction with little or no change in oxygen concentration or the rate of oxygen consumption [97] is consistent with increase in cellular energy state {[ATP]/([ADP]$_f$[Pi])} [115], [118]. In liver, increase in energy state suppresses the activity of pyruvate dehydrogenase (increased acetylCoA, [NADH]/[NAD$^+$], and [ATP]/[ADP]$_f$), activates pyruvate carboxylase (high acetylCoA), decreases [AMP]$_f$, lowering AMP-dependent protein kinase (AMPK) activity [61], [98], [105], [124]. Decreased AMPK activity increases acetylCoA carboxylase activity [19], [30], [38], [39], [42], [124] and the concentration of malonylCoA [19]. MalonylCoA inhibits carnitine palmitoyl transferase 1 [30], [73] suppressing uptake and oxidation of long chain fatty acids by mitochondria. Fatty acid synthesis is activated and this, combined with excess fatty acids taken up by the liver, results in excess fatty acids being made into triglycerides. Interestingly, Galli and coworkers [32] reported that expression of ADH1 in HeLa cells was sufficient to result in ethanol induced fat accumulation, consistent with ethanol oxidation induced inhibition of AMPK being responsible for cellular lipid accumulation.

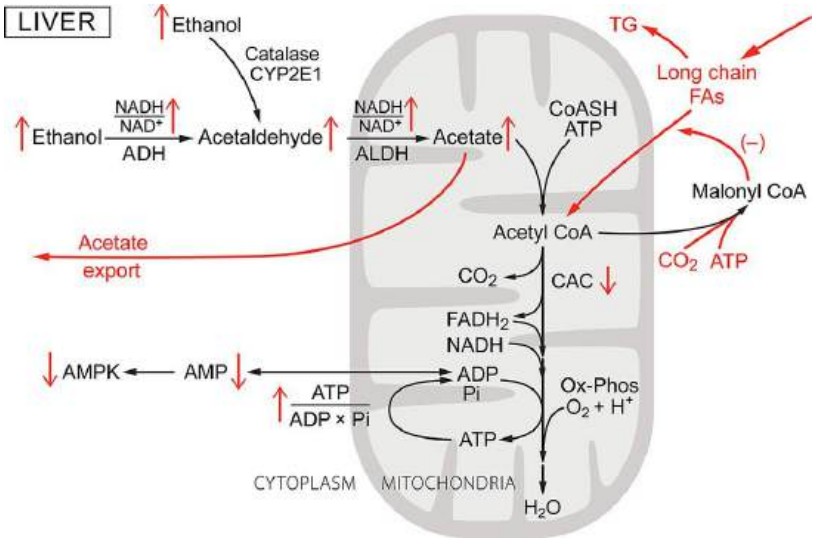


Fig. 1. **Ethanol metabolism in the liver: a general schematic of its metabolism and effect on liver metabolism**. As the ethanol concentration in blood increases it enters all tissues and is oxidized, mostly through alcohol dehydrogenase 1 (ADH1) in the liver. Whether a particular metabolic parameter is increased or decreased is indicated by an up or down arrow. The metabolic steps characteristic to liver are shown in red. ADH is cytoplasmic and as ethanol concentration increases ethanol oxidation increases cytoplasmic NADH/NAD$^+$ (lactate/pyruvate) ratio. Acetaldehyde, the product of ethanol oxidation, enters the mitochondria and is oxidized to acetate by mitochondrial acetaldehyde dehydrogenase (ALDH2), increasing intramitochondrial NADH/NAD$^+$ (β-hydroxybutyrate/acetoacetate) ratio. The acetate formed is mostly exported into the blood. Both the ADH and ALDH reactions are irreversible and the rate of ethanol oxidation by ADH1 is determined by enzyme content and degree of saturation (K_M = 1.4 mM). At blood alcohol greater than about 0.01 g/dl (2.2 mM), the rates are high enough that hepatic citric acid (CAC) activity, as measured by CO_2 production, is markedly suppressed. Oxygen consumption (Ox-Phos) is slightly decreased due to the higher ATP/O ratio when oxidizing NADH than for long chain fatty acids. The energy state is increased, decreasing the concentration of free AMP ([AMP]$_f$) and thereby the activity of AMP-dependent protein kinase (AMPK). MalonyCoA is increased and this inhibits acylcarnitine transferase 1, inhibiting long chain fatty acid transport into the mitochondria and subsequent oxidation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Acetate, the primary product of ethanol oxidation by the liver, is a short (shortest) chain fatty acid and oxidation of short chain fatty acids is not subject to the regulation imposed on oxidation of long chain fatty acids. Acetate produced by the liver results in a large increase in acetate concentration in the blood, to about 1 mM [44], [78]. The rate at which acetate is taken up, and oxidized, by other tissues is

dependent on concentration in the blood. In both humans and rats, if blood acetate concentrations are maintained near 1 mM or higher, this results in under consumption of long chain fatty acids and activation of fatty acid synthesis [19], [30], [62], [64], [95], [122]. Since ethanol and acetate are only two carbons long, their products cannot be used for net synthesis of glucose or for anaplerotic support of the citric acid cycle or net glucose synthesis. The lactate/pyruvate ratio is increased, in large part due to decrease in pyruvate [52], and this limits the activity of pyruvate carboxylase, suppressing gluconeogenesis [29], [52], [53]. Blood glucose concentrations decrease [53], but only slightly because the set point for glucose homeostasis in only slightly decreased [72] and glucose consumption by peripheral tissue and brain is decreased.

AMP-DEPENDENT PROTEIN KINASE AND THE METABOLIC RESPONSES TO ETHANOL OXIDATION.

Essentially all cells have adenylate kinase and the activity is sufficient to maintain the reaction:(1)ATP + AMP = 2 ADPnear equilibrium. The equilibrium constant is approximately 1.0, and [ATP] is maintained nearly constant, so $[AMP]_f$ decreases as the square of the decrease in free $[ADP]_f$. This makes $[AMP]_f$ a very sensitive measure of the energy state [51], [115], [118]. As a result, energy metabolism has evolved with $[AMP]_f$ as a core regulatory parameter in setting metabolic homeostasis. It is notable that metabolic homeostasis in eukaryotes has an $[AMP]_f$ “set point” and cellular metabolism operates over a narrow range of $[AMP]_f$ [115], [118]. Central to maintaining metabolic homeostasis is AMP-dependent protein kinase (AMPK) [30], [38], [39], [42], [49], [118]. Although $[AMP]_f$ also contributes to the regulation of many other enzymes and regulatory pathways, the importance of AMPK is shown schematically in Fig. 2A. This includes modulation not only of complex cellular functions, such as protein synthesis, autophagy, mitophagy, and gene expression, but also of individual enzymes and regulatory proteins (Fig. 2B). Most discussions of metabolic regulation involving AMPK focus on the conditions where $[AMP]_f$ increases, as occurs in exercise, hypoxia, or inhibition of oxidative phosphorylation, and AMPK activity increases. Activation of AMPK is designed to increase ATP production by increasing catabolic ATP production and inhibiting anabolic ATP consumption. Ethanol ingestion, however, decreases $[AMP]_f$ and thereby AMPK activity, suppressing catabolic metabolism and enhancing anabolic metabolism. Indeed, treatment with 5-aminoimidazole-4-carboxamide-1-β-D-ribofuranoside (AICAR), which activates AMPK, has been reported to provide significant protection from alcohol induced fatty liver in rats [105]. Because oxidation of ethanol and acetate is cell specific and largely unregulated, the metabolic consequences differ among cell types and overall metabolic integration disrupted.

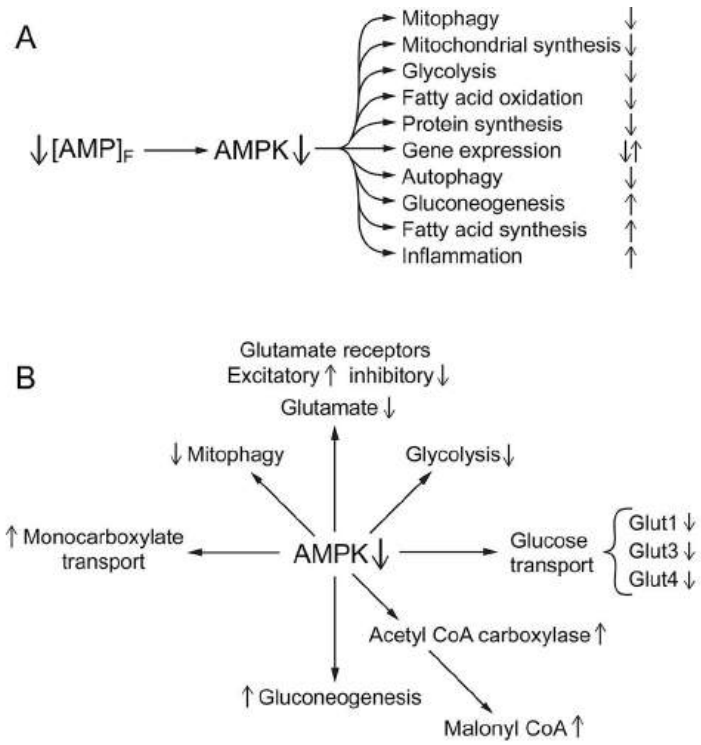


Fig. 2. **A, B. The central role of AMPK in the metabolic responses to ethanol consumption**. Ethanol oxidation, by increasing intramitochondrial [NADH]/[NAD^+] and cellular energy state, decreases the concentration of free AMP ($[AMP]_f$). This suppresses phosphorylation of AMPK and thereby its activity. AMPK has a central role in maintaining metabolic homeostasis (for reviews see [38], [39], [49]). Fig. 2A indicates some of the wide range of metabolism related processes that respond to alterations in AMPK activity. Some of these are expected to increase and others decrease in response to the ethanol induced decrease in AMPK activity with the direction indicated by up or down arrows on the right. Fig. 2B indicates a few effects of decreased [AMPK] on specific enzymes/processes that are noted in the text. When ethanol is consumed, its high caloric content and lack of regulation results in wide spread disruption of metabolism. Central to this disruption is suppression of "normal" $[AMP]_f$ and AMPK activity.

ETHANOL/ACETATE METABOLISM AND ITS EFFECTS ON MUSCLE

Although liver and brain are the primary tissues of concern in when considering the consequences of ethanol consumption, other tissues are also affected. This includes cardiac and skeletal muscle. Although neither has significant ADH activity, they have high acetylCoA synthetase and acetylCoA carboxylase and rely heavily on fatty acid oxidation for energy metabolism. This is particularly true when only moderately active (heart) or during rest and after endurance exercise (skeletal muscle) (Fig. 3). Chen et al. [16] reported that chronic ethanol feeding in rats suppresses expression of AMPK, myocyte enhancer factor 2, and glucose transporter 4 (Glut4) in myocardium. The effects of ethanol on muscle metabolism would be expected to be indirect, arising from elevated blood acetate. This is consistent with the report by Kiviluoma [48] that addition of acetate to the perfusate of isolated rat hearts increases incorporation of fatty acids into myocardial lipids. In addition, adding acetate at mM concentrations to the perfusate of isolated rat hearts increases reduction of intramitochondrial NAD and energy state relative to glucose perfusion [48], [99]. Similar observations have been reported for rat skeletal muscle *in vivo* by Bertocci et al. [11]. When lactate and acetate consumption were measured in resting skeletal muscle, acetate provided most of the acetylCoA (65%) oxidized by the citric acid cycle (CAC). In contracting muscle, the fraction provided by acetate decreased, but only to 43%. Thus, in heart under moderate work load and in resting skeletal muscle, acetate taken up from the blood can be oxidized at rates sufficient to maintain energy metabolism. Acetate displaces (down regulates) oxidation of the physiologically preferred substrate, long chain fatty acids. Putman et al. [82] measured metabolites and pyruvate dehydrogenase activity in muscle during acetate infusion in humans. The reported changes included increase in acetylCoA/CoASH and acetyl-carnitine, as well as decrease in active pyruvate dehydrogenase (PDHa). Uncontrolled production of acetylCoA from acetate in mitochondria and the resultant increase in acetylCoA/CoA, if sustained, would deplete intramitochondrial CoASH. CoASH is required for citric acid cycle. Synthesis of acetyl-carnitine by carnitine acetyl transferase allows export of excess active acetyl groups to the cytoplasm but at the expense of intramitochondrial carnitine. This process appears to be self limiting due to depletion of intramitochondrial CoASH for acetylCoA synthase combined with decreased oxalacetate for citrate synthase limits input of acetylCoA to the rate of oxidation through the citric acid cycle. Evidence supports the view that, following ethanol ingestion, elevated blood acetate becomes a major energy source for resting and moderately active muscle, consistent with associated suppression of

AMPK activity [16], [48], [99]. Interestingly, tissues from alcohol dependent humans show dysfunctions in protein synthesis, mitochondrial content, and morphology, consistent with chronic suppression of AMPK activity (Fig. 2A).

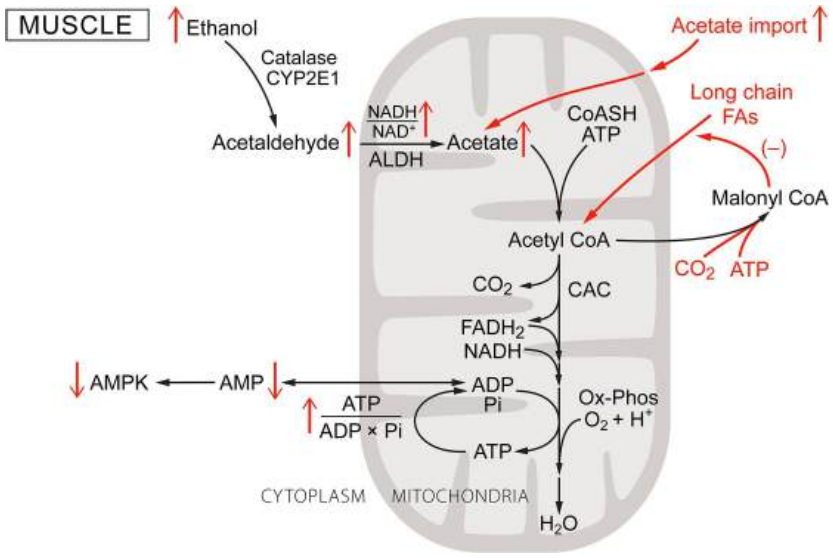


Fig. 3. **Ethanol metabolism in muscle: A schematic of metabolism in muscle, which does not have alcohol dehydrogenase**. Muscle, due to the lack of ethanol dehydrogenase, may oxidize some ethanol through catalase and/or CYP2E1, but these rates are low. Acetate is, however, taken up from the blood, activated by acetylCoA synthetase, and oxidized through the CAC and oxidative phosphorylation. This uncontrolled input of acetylCoA increases the acetylCoA/CoASH ratio and energy state, decreasing $[AMP]_f$ and thereby AMPK activity. Decreased AMPK activity suppresses long chain fatty acid oxidation.

ETHANOL METABOLISM IN BRAIN AND ITS METABOLIC CONSEQUENCES: A. GENERAL CONSIDERATIONS

Brain is a complex organ with several types of specialized cells, each with its own metabolic requirements and supporting complement of enzymes and metabolite transporters. Ethanol produces major metabolic disturbances in brain metabolism (shown schematically in Fig. 4) including: decreased glucose uptake and metabolism, increased monocarboxylate uptake and metabolism, stimulation of dopaminergic neural activity, suppression of glutamatergic neural activity, and disruption of glutamate metabolism. Catalase, ADH, and P450 systems for oxidizing ethanol are all present in brain, and brain actively takes up acetate from the blood. Determining the role of ADH, however, is complicated by the presence of multiple different forms of alcohol dehydrogenase as well as heterogeneous distributions among cells and tissues [5], [12], [13], [47]. Initial reports by Raskin and Sokoloff [84] of low, but significant, rates of oxidation of ethanol by ADH in subfractions of brain tissue were largely discounted because of the low activity and concerns about the method of analysis [126], [127]. The enzyme properties reported, however, were consistent with the presence of small amounts of ADH1. Initial immunohistochemical data focused on ADH3, an alcohol dehydrogenase that is widely distributed in brain but has negligible activity for ethanol [27], [126], leading some to conclude that metabolism of ethanol by ADH in brain was not significant. Later measurements using antibodies grown to purified liver ADH indicated that while the total amount of enzyme are low, it is concentrated in individual neurons of cerebral cortex, hypothalamus, infundibular stalk of pituitary, and Purkinje cells of cerebellum. This distribution correlates with known sites of ethanol toxicity (13. 47). More recently, selective antibodies for ADH1 (K_M = 1.4 mM) and ADH4 (K_M = 37 mM) have been used [27]. Again the total amounts were low but concentrated in a small fraction of the brain cells. Martinez et al [70] examined tissue sections from several regions of adult rat brain by *in situ* hybridization to detect expression of genes encoding ADH1. ADH1 mRNA was found in granular and Purkinje cell layers of cerebellum, in pyramidal and granule cells of hippocampal formation, and in some types of cells in cerebral cortex. ADH4 expression was detected in Purkinje cells, pyramidal and granule cells of the hippocampal formation, and pyramidal cells of cerebral cortex. Substantial levels of both ADH1 and ADH4 mRNAs were detected cells of the CNS epithelial and vascular tissues: leptomeninges, choroidplexus, ependymocytes of ventricle walls, and endothelium. Thus, current evidence is consistent with oxidation of ethanol by ADH capable of significantly affecting brain function in subpopulations of cells in which ADH1 and/or ADH4 are expressed.

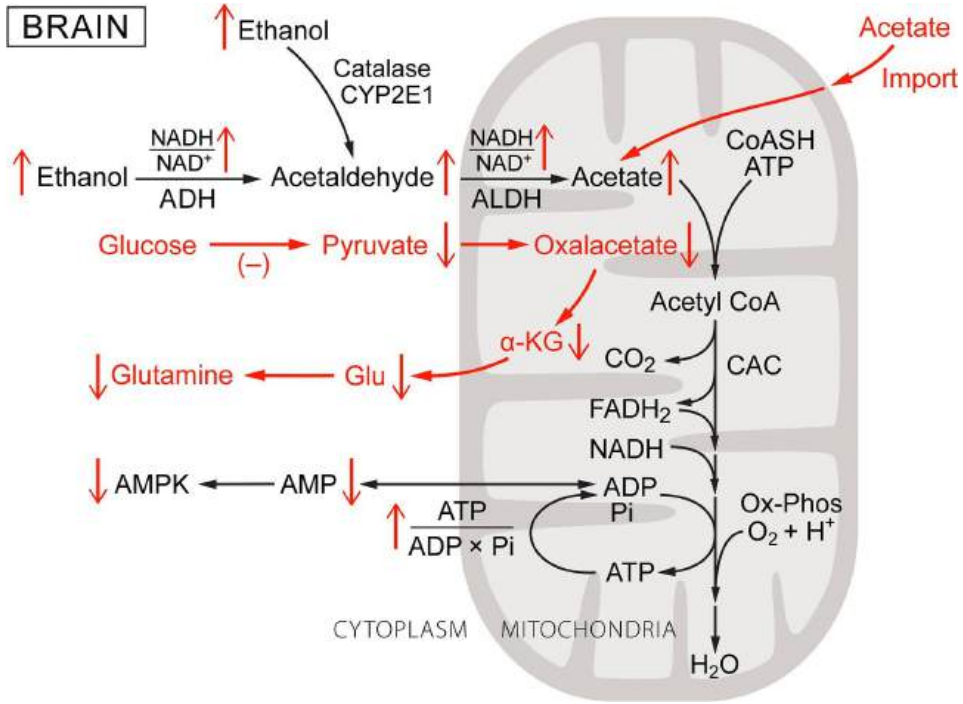


Fig. 4. **Ethanol metabolism in the brain: a schematic of its metabolism and the changes induced with respect to normal brain metabolism.** As was the case for Fig. 1, Fig. 2, Fig. 3, increase or decrease in each metabolic parameter is indicated by an up or down arrow. Metabolic steps characteristic brain cells and important to the effects of ethanol are shown in red. Heterogeneity in cell types and cell metabolism in brain precludes a general characterization but cells can be roughly separated into those with ADH1, ADH4 and those without this activity. Those without ADH1 and/or ADH4 oxidize ethanol by catalase or cytochrome P450 (CYP2E1) to acetaldehyde without reducing NAD. The amounts of ADH1 and ADH4 in brain are small but concentrated in a small fraction of the total number of cells. In cells with ADH1 and/or ADH4, ethanol oxidation increases the cytoplasmic NADH/NAD^+. All brain cells have mitochondrial aldehyde dehydrogenase and produce acetaldehyde, if not by ADH, by catalase and CYP2E1. Oxidation of acetaldehyde and/or acetate taken up from blood augments that produced within the cell, increasing intramitochondrial NADH/NAD^+. Brain cells, notably astrocytes, normally consume glucose and produce pyruvate for anaplerosis through pyruvate carboxylase. Ethanol and acetate catabolism, however, down regulates glucose consumption by decreasing both glucose transport and glycolysis. Decreased pyruvate production from glucose, combined with increased lactate/pyruvate, limits the rate of pyruvate carboxylation and thereby *de novo* synthesis of glutamate. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

ETHANOL AND ACETATE METABOLISM IN BRAIN: B. RATES AND CONTRIBUTIONS TO TOTAL ENERGY PRODUCTION

Ethanol, acetaldehyde, and acetate have been shown to be oxidized in brain and the combined rates are high enough to disrupt glucose and amino acid metabolism. Wang and coworkers [111] studied the rate of ethanol oxidation in rat brain cortex *in vivo by* ^{13}C MRI. Using a combination of 2-^{13}C labeled ethanol and 1,2-^{13}C labeled acetate, they reported significant rates of incorporation of carbon from ethanol into glutamate and glutamine. The authors calculated ethanol oxidation accounted for 12% and 20% cortical oxidative metabolism in ethanol naïve and ethanol pre-treated rats. The measurements were specifically of ethanol oxidation and did not include oxidation of acetate taken up from the blood. Measurements of acetate metabolism in rat brain *in vivo* using ^{13}C MRI [21], [81], [107] show it is readily transported through the blood-brain barrier and metabolized. As acetate concentration in blood was raised to 2–3 mM, the rate of metabolism by the brain saturated. At saturation, enrichment of glutamine and glutamate at the C4 position were consistent with most of the

acetylCoA entering the citric acid cycle in astrocytes arising from acetate with the remaining coming from unlabeled glucose. The models used in the interpretation assumed, however, that acetate metabolism occurred almost entirely in astrocytes with negligible contribution by neurons. This may not be appropriate [87] particularly as the neurons contain substantial levels of acetylCoA synthase. As noted earlier, when ethanol is ingested, acetate is produced in the liver and released into the blood. Blood acetate levels rise rapidly to about 1 mM or more [44], [78], and ethanol oxidation in brain is augmented by oxidation of acetate taken up from the blood. When these are combined, ethanol becomes a major source of acetylCoA for production of ATP used for ATP production in many brain cells. Recent reports include evidence that in mice ethanol consumption leads to rapid increase in histone acetylation in the brain with the acetyl moiety arising from ethanol [74]. This acetylation is reported to alter gene expression in hippocampus and to affect cellular mechanisms related to learning. The level of acetylCoA and the acetylCoA/CoASH ratio are important regulators of protein acetylation and are coupled to cellular energy state and AMPK activity.

ETHANOL EXPOSURE PHASE 1: ACUTE EFFECTS OF ETHANOL INGESTION ON METABOLISM

As noted above, alcohol ingestion results in rapid changes in the metabolism of brain and other tissues. Following ingestion, oxidation of ethanol takes precedence over other metabolites used for energy metabolism and the latter are down regulated to accommodate the input from ethanol. The effects are cell and tissue specific and depend on how rapidly ethanol is oxidized and/or acetate is taken up through the monocarboxylate transporter. Liver, due to its large size, immediate exposure to the ingested alcohol from the portal circulation, and high content of ADH1 dominates systemic ethanol disposal. Effects of acute ethanol ingestion on liver include rapid increase in metabolic energy state, production and export of acetate, suppression of long chain fatty acid oxidation and synthesis, increased β-hydroxybutyrate/acetoacetate ratios, and under production of glucose by gluconeogenesis. The response of other tissues to ethanol exposure is heterogeneous. In pancreas, increased energy state due to ethanol/acetate metabolism slightly lowers the glucose homeostasis set point [116], [117], [119]. In most other tissues, input of acetylCoA from ethanol and acetate largely replaces acetylCoA from long chain fatty acids and amino acids for energy metabolism [31], [107], [110]. The oxidation of ethanol and/or acetate induces an increase in energy state and associated decrease in $[AMP]_f$ and AMPK activity [19], [30], [31], [38], [39]. Increase in blood acetate in rats to 300 μM, for example, decreased AMPK phosphorylation in hypothalmus by about 50% [31]. Decreased AMPK activity suppresses GLUT1, GLUT3, and GLUT4, the dominant glucose transporters in brain [31], [37], [75], [119]. Handa et al. [37] isolated plasma membranes and low density microsomal fractions from rat brain 4 h after their treatment with ethanol (3 g/kg body weight) and reported immunoblot measured protein decreased by 17% and 71% for GLUT1 and 54% and 21% for GLUT3, respectively. In addition, immunofluorescence imaging showed decreased GLUT1 in choroid plexus and cortical microvessels. This is consistent with reports by Muneer and coworkers [75] that chronic ethanol treatment suppresses glucose transport and GLUT1 protein expression in microvessels of the blood-brain barrier of rats and by Frost et al. [31] that it decreased GLUT1 in the plasma membrane of brain cells in primary culture. Accompanying ethanol induced decrease in glucose transporters is increase in monocarboxylate transporter [65]. Pyruvate is needed not only for anaplerotic maintenance of citric acid cycle activity but also for *de novo* synthesis of glutamate in astrocytes through the activity of pyruvate carboxylase [125]. Increased uptake of lactate from the blood [65] would normally provide increased pyruvate through oxidation by cytoplasmic NAD^+. For cells with ADH1 and/or ADH4, however, oxidation of ethanol in the cytoplasm results in increased reduction of the cytoplasmic NAD pool and an increase in the lactate/pyruvate ratio [52]. In these and other cells, acetaldehyde and acetate oxidation in the mitochondria also increases reduction of intramitochondrial NAD, suppressing aspartate-glutamate shuttle activity. Although intracellular lactate concentrations in brain increase during ethanol oxidation, this is due to elevated systemic lactate/pyruvate and pyruvate concentrations actually decrease. Intracellular pyruvate concentrations are normally less than the K_M of pyruvate carboxylase, which is near 400 μM [92], making the rate of pyruvate carboxylation strongly dependent on intracellular pyruvate concentration [52]. Ethanol induced decrease in intracellular pyruvate suppresses *de novo* synthesis of oxalacetate, and thereby of glutamate, in astrocytes and this disrupts glutamate homeostasis in astrocytes and glutamatergic neurons.

ETHANOL EXPOSURE PHASE 2: METABOLIC CONSEQUENCES OF ALCOHOL INGESTION IN GREATER AMOUNTS AND OVER INTERMEDIATE PERIODS OF TIME (DAYS, WEEKS)

Ethanol dependence arises after significant periods of nearly continuous alcohol consumption, typically months to years. This is not an all or nothing phenomenon, however, and there is a transition period during which withdrawal of alcohol results in a progressively greater sense of unease, anxiety, and irritability that can be overcome by consuming alcohol. The shortest time of exposure that generates significant ethanol dependence may be binge drinking, in which substantial amounts of ethanol are consumed over periods of one or more days. A few hours after drinking stops, ethanol levels approach zero and there is a period that often includes severe headaches and other unpleasant after effects, referred to as a hangover. The colloquial antidote for a hangover is “hair of the dog that bit you”, i.e. drinking some of the same alcoholic beverage consumed during the binge in order to alleviate the hangover. This antidote acutely relieves many of the symptoms, consistent with hangovers arising in part through deficiency in oxidizable substrates for energy metabolism in brain. The extended time and extent of alcohol exposure during a binge allows participation of slower regulatory processes, including altered rates of protein turnover and gene expression. When the levels of alcohol and acetate fall too quickly for regulation to fully compensate by switching back to the usual (non-ethanol) energy sources, the deficiency can be overcome by providing some alcohol. This temporarily relieves the deficiency in energy

metabolism, and thereby the hangover, as well as providing additional time for metabolic regulation to compensate for lack of ethanol/acetate.

ETHANOL EXPOSURE PHASE 3: ETHANOL DEPENDENCE RESULTING FROM HEAVY ALCOHOL CONSUMPTION OVER PERIODS OF MONTHS AND YEARS

Individuals with alcohol dependence consume a substantial fraction of their daily caloric intake as ethanol, typically near 50% of the total calories [5], [25]. This large intake, spread throughout the day and sustained for weeks to months, results significant (mM) levels of ethanol being in the blood most of the time. Diets of individuals with ethanol dependence are often deficient in essential nutrients, including vitamins, minerals, unsaturated fatty acids, and amino acids. In addition, ethanol is only 2 carbons long and can not be used for *de novo* synthesis of glucose, glutamate and other metabolites critical to long term metabolic homeostasis. In brain, chronic suppression of the oxidation of glucose and other (non-ethanol related) metabolites results in depletion of their associated transporters and enzymes. Conversely, the capacity for metabolism of ethanol and acetate increases [43], [65], [107]. If ethanol consumption ends abruptly, ethanol and acetate levels in the body fall to near zero within a few hours. The oxidizable metabolites used to support energy metabolism in non-ethanol dependent animals have been chronically down regulated and can no longer increase sufficiently to maintain the cellular energy state.

As noted above, for individuals who are ethanol dependent, ethanol normally provides 50% or more of the daily caloric intake. Glucose consumption in brain is substantially decreased by ethanol consumption [107], [108], [110], [112]. Oxidation of one mole of glucose provides slightly more ATP (38 ATP/mole) than does oxidation of two moles of ethanol. Oxidation of alcohol to acetaldehyde by ADH and of acetaldehyde to acetate by ALDH each contributes one NADH for oxidative phosphorylation (6 ATP/ethanol) and oxidation of acetate through the citric acid cycle provides an additional 3 NADH and 1 $FADH_2$ (total of 17 ATP/ethanol, 11 ATP/acetate). Even after correction for the ATP required for acetate activation (acetylCoA synthetase), oxidation of 2 mol of ethanol provides 30 mol of ATP, about 80% of that supplied by oxidation of 1 mol of glucose. Volkow et al. [108] measured brain glucose metabolism in twenty healthy control subjects using positron emission tomography (PET) and fluorodeoxyglucose after administration of placebo or either 0.25 g/kg, or 0.5 g/kg of ethanol over a 40 min period. Both doses of ethanol significantly decreased whole-brain glucose metabolism (10% and 23% respectively). Similar decreases have been reported by other authors, i.e. 26% for 0.75 g/kg [112], 25% for approximately 0.6 g/kg iv [91]. Volkow et al. [107] reported that the response to ethanol is regionally specific; whereas 0.25 g/kg predominantly reduced glucose metabolism in cortical regions, 0.5 g/kg reduced metabolism in cortical as well as subcortical regions (i.e. cerebellum, mesencephalon, basal ganglia and thalamus). As noted earlier, suppression of brain glucose consumption is in response to combination of oxidation of ethanol within the brain and of acetate taken up from the blood [107], [108], [110], [111]. Indeed, in heavy drinkers studied during sobriety, acetate metabolism in the occipital cortex is reported to be increased [43].

Persistent reliance on acetate for energy metabolism of alcoholics during early alcohol detoxification may explain why the decreased brain glucose consumption observed in alcoholics mostly disappears within the first 2 weeks of detoxification [109]. It can be inferred that for cells in brain for which ethanol (and/or acetate) oxidation is responsible for substantial fraction of their ATP production, sudden withdrawal of this important support of energy metabolism results in cellular malfunction and thereby abnormal brain activity. In ethanol dependent mice, withdrawal symptoms correlate with decrease in blood alcohol and acetate, symptoms maximizing as the concentrations approach zero [85]. Derr and coworkers [22], [23], [24] reported that, for ethanol dependent rats, providing acetate during withdrawal significantly alleviated, and providing a mixture of butyrate, lactate, and β-hydroxybutyrate fully suppressed [22],the tremulous portion of their withdrawal symptoms. It did not affect the handling induced convulsions. This is consistent with the tremulous, but not the convulsive, part of the withdrawal symptoms arising from a deficiency in energy metabolism needed to support neuronal activity in the brain.

NEUROLOGICAL RESPONSES TO ETHANOL CONSUMPTION: A PROLOG

Our paper focuses on metabolism of ethanol and the impact of ethanol consumption on metabolism as a whole. This is an integral and important part of understanding how alcohol influences human biochemistry, physiology, and social behavior. In brain, we focus on effects of ethanol on metabolism of neurons and astrocytes, where changes in energy state directly influences not only all aspects of neurotransmitter function (release, receptor sensitivity, transport, and metabolism) but also induces long term alterations in the underlying cellular machinery (through gene expression, etc). It is recognized, however, that connections between cellular metabolism and neuronal function, and from neural function to behavior, are very complex. Others have made great effort to understand how ethanol alters neuronal function; receptor affinities, ion channel conductance, synaptic connections etc, but many of the particulars remain poorly understood. There is general agreement that ethanol induces substantial changes in neural function but the relative contribution of metabolism and concentration remains under discussion. Due to the complexity of the neurological consequences of ethanol consumption we will address only those we consider to be primarily responses to ethanol and acetate metabolism. There is, for example, an extensive literature on the role of ethanol metabolism and of AMPK activity in modulation of neuronal function (see as examples, [3], [93], [94], [123]). We emphasize the role of energy metabolism and $[AMP]_f$ in order to illustrate not just their importance in the response to ethanol but also their central role in coordinating metabolic processes essential to biological existence as we know it.

NEUROLOGICAL RESPONSES TO ETHANOL CONSUMPTION A: EARLY ALTERATIONS IN THE DOPAMINERGIC SYSTEM AND POTENTIATION OF CONSUMPTION

Ethanol induces dopamine release and increased activity in the striatum as indicated by functional imaging [91]. Dopaminergic neurons that respond to nutrients, such as glucose, involve metabolism of that nutrient. Increase in nutrient concentration causes an increase in energy state and modulates ion channel (such as the K_{ATP} channel) conductance to initiate the dopaminergic signaling cascade. Ethanol, in contrast to other nutrients, has access to neurons behind the blood-brain barrier and this augments the usual nutrient reward circuitry. Dopaminergic projections go to the nucleus accumbens, which has a crucial role in the reward system of human brain. These dopaminergic effects have been attributed to inhibition of GABAergic interneurons of the ventral tegmental area (VTA) by alcohol. Ethanol-induced dopamine release in nucleus accumbens has been reported to be antagonized by administration of an opiate receptor antagonist [34], suggesting inhibition of GABAergic VTA interneurons are at least partly mediated by opioidergic afferents to these neurons. It is reasonable to suggest ethanol induced dopamine release is in response to increased energy state (nutrient sensing) in GABAergic and dopaminergic neurons.

Considerable attention has been paid to acetaldehyde in relation to the dopaminergic response. Inhibition of catalase activity has been reported to prevent ethanol induced dopaminergic signaling in nucleus accumbens [46]. The authors attributed this effect to decrease in acetaldehyde concentration, but inhibiting catalase alters metabolism in many ways, of which acetaldehyde concentration is only one. Follow-up studies have reported that, in rats bred for ethanol preference, administration of a lentiviral vector coding for aldehyde dehydrogenase-2 (ALDH2) into the ventral tegmental area decreased long term ethanol consumption. This decrease was observed, however, only for ethanol naïve rats and not rats that had consumed ethanol on a 24-hour basis for 81 days [46]. Rivera-Meza et al. [86] reported that treating naïve UChB alcohol preferring rats with N-(1,3-benzodioxol-5-ylmethyl)-2,6-dichlorobenzamide, a drug that increased ALDH2 activity in brain by 3 fold, markedly lowered the amount of alcohol they consumed. The authors suggested that activation of ALDH2 in brain decreased acetaldehyde concentrations and this was responsible for inhibiting both acquisition and maintenance of chronic ethanol intake by alcohol-preferring rats. It has also been suggested that acetaldehyde produced from ethanol reacts with dopamine to form salsolinol and that salsolinol is an agonist for μ-opioid receptors, contributing to the reward response to alcohol [10]. There is, however, little evidence for a role for salsolinol in ethanol consumption and metabolism in humans [56], [57] and the contribution of acetaldehyde concentration remains to be established. In humans, expression of inactive aldehyde dehydrogenase 2 (ALDH2) is protective against alcohol use disorders [12], [60]. This decrease in ethanol consumption, however, has been attributed to the unpleasant side effects of systemic increased acetaldehyde. Lack of ALDH2 also suppresses oxidation of acetaldehyde in mitochondria, disrupting ethanol induced increase energy state, particularly in cells with ADH1 and the reward response to ethanol consumption.

Studies in mice by Tabakoff et al. [101] indicated that ethanol tolerance and physical dependence involve different mechanisms. The authors used intraventricular injection of 6-OH dopamine to deplete dopamine and norepinephrine in the brain. Injected prior to chronic ethanol exposure, this prevented development of tolerance to the hypnotic and hypothermic effects of ethanol without significantly affecting development of physical dependence as measured by withdrawal symptoms (convulsions etc). Once tolerance was established, however, 6-OH dopamine injection did not alleviate that tolerance. It is reasonable to conclude that dopaminergic and noradrenergic systems contribute substantially to the positive physiological response to ethanol and to development of alcohol tolerance, but not to alcohol dependence or withdrawal symptoms.

NEUROLOGICAL RESPONSES TO ETHANOL CONSUMPTION B: EARLY EFFECTS ON GLUTAMATE AND GABA METABOLISM

Chronic ethanol intake in rodents, primates, and humans results in physical dependence and it is widely considered that dependence involves excitatory glutamate and inhibitory GABA neurons. Ethanol oxidation is particularly disruptive of glutamate and GABA metabolism and handling in brain. It is hypothesized that metabolism of astrocytes and neurons is coordinated through glutamate/glutamine cycling [90] and/or lactate transfer [9]. Our version of that metabolic coordination in the absence of ethanol is shown schematically in Fig. 5. Astrocytes have high glycolytic activity and glycogen storage as well as high pyruvate carboxylase activity whereas neurons are oxidative and have high pyruvate dehydrogenase activity. Pyruvate carboxylase in astrocytes is responsible for *de novo* synthesis of oxaloacetate (and thereby glutamate and glutamine) from pyruvate provided by glycolysis (anaplerosis). Pyruvate carboxylase provides α-ketoglutarate (α-KG) for synthesis of glutamate by aspartate/α-KG aminotransferase. Glutamate is then condensed with ammonia (glutamine synthetase) to form glutamine which is exported and taken up by neurons. In neurons, the glutamine is deaminated (glutaminase) and used as both a neurotransmitter and energy source. In order to be used for ATP production, glutamate is first deaminated by either glutamate/oxalacetate aminotransferase [41] or glutamate dehydrogenase to α-KG which enters the CAC. The CAC, however, does not carry out net oxidation of substrates with chain lengths of 4 or more carbons [50]. α-KG is oxidized to malate in the CAC and malate is exported to the cytoplasm where malic enzyme oxidatively decarboxylates malate to pyruvate. In neurons, pyruvate dehydrogenase is active and pyruvate from either malate or lactate is oxidized to acetylCoA and then to CO_2 and water by CAC plus Ox-Phos. Glutamate released at the synapse during neural activity is taken up partly by synaptic glutamate transporters [26], [120] while glutamate that diffuses out of the synapse into the perisynaptic space is taken up by astrocytes. Glutamate consumed for energy metabolism in neurons is replaced by *de novo* synthesis from pyruvate in astrocytes. Metabolism of ethanol and/or its metabolites inhibits both glucose uptake and glycolytic production of pyruvate through increased energy state (decreased $[AMP]_f$) and increased NADH. Ethanol induced increase in systemic lactate/pyruvate also

suppresses intracellular pyruvate concentrations, despite an increase in intracellular lactate through increased monocarboxylate transport. Intracellular pyruvate concentrations are typically below the K_M for pyruvate carboxylase (about 400 μM), and decrease would be expected to suppress astrocytic *de novo* glutamate synthesis. Extracellular glutamate levels in the synapse also decrease due to increase in energy state through its effect on the energy dependence of synaptosomal glutamate transport [26]. The net early effects of ethanol on glutamatergic neural activity include decreased glutamate release in response to the increased energy state in neurons and increased levels of inhibitory neurotransmitter GABA [33]. These contribute to the calming effects of ethanol through suppression of glutamatergic anxiety producing neural activity.

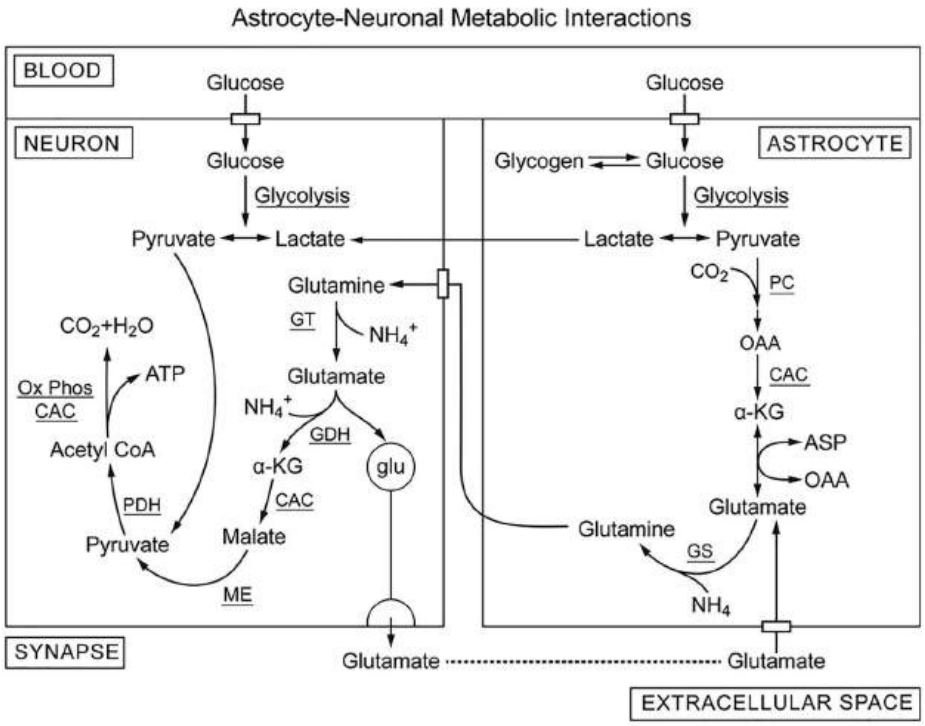


Fig. 5. **Our version of the glutamate/glutamine cycle that helps to support glutamatergic neural activity (no ethanol)**. The astrocytes are primarily glycolytic, using puruvate from glycolysis for *de novo* synthesis of oxalacetate by pyruvate carboxylase whereas glutamatergic neurons are more oxidative and consume lactate and glutamate for energy metabolism. Pyruvate carboxylase is activated when acetylCoA and the acetylCoA/CoASH ratio are high, conditions for which pyruvate dehydrogenase activity is suppressed. In addition, the CAC does not catalyze net oxidation of intermediates with 4 or more carbon atoms. As a result, astroctes take up glucose and produce pyruvate by glycolysis for net synthesis of oxalacetate. Oxalacetate is converted to citrate and then oxidized to α-ketoglutarate (α-KG) in the CAC. The α-KG is transaminated to form glutamate and then condensed with ammonia glutamine synthetase (GS). Glutamine released from astrocytes is taken up by neurons which reverse this process. Glutamine is deaminated by glutaminase and then oxidatively dehydrogenated to α-KG by glutamate dehydrogenase (GDH). α-KG is oxidized to malate in the CAC and exported to the cytoplasm where it is converted to pyruvate by malic enzyme (ME). In neurons, pyruvate from lactate and glutamate is oxidized by pyruvate dehydrogenase and provides acetylCo-A for energy metabolism. Most of the glutamate released during neural activity is taken up by glutamate transporters in the synapse but some diffuses out to the surrounding perisynaptic space, is taken up by astrocytes and recycled. *De novo* synthesis of glutamate in astrocytes replaces that used by neurons for energy metabolism, maintaining stable glutamate/glutamine levels in support of glutamatergic neuronal function.

NEUROLOGICAL RESPONSES TO ETHANOL CONSUMPTION: C. LONG TERM ALTERATIONS IN GLUTAMATERGIC NEURAL SYSTEM AND ALCOHOL WITHDRAWAL SYNDROME

It has been proposed that neurophysiological and pathological effects of ethanol are mediated, to an important extent, by glutamatergic neurons [35], [36], [79], [96], [106]. Biochemical and electrophysiological studies have reported that chronic ethanol treatment increases the number of excitatory N-methyl-D-aspartate (NMDA) receptor-ionophore complexes in hippocampus, a brain area associated with ethanol withdrawal seizure activity [35], [36], [96]. Withdrawal symptoms vary, but in mice removal of ethanol from the diet induces withdrawal-associated tremors and handling-induced seizures [35], [36], [85]. The withdrawal symptoms are reported to correlate with decrease in blood alcohol and acetate levels, with the most severe seizure activity occurring when these concentrations fall to near zero. In mice, the NMDA receptor levels return to control in 24 h and the mice no longer have handling induced seizures. Treatment with NMDA during

withdrawal exacerbated handling induced withdrawal seizures, while administration of MK-801, a noncompetitive NMDA receptor antagonist, decreased the occurrence and severity of seizures [35]. Treatment with MK-801 did not, in alcohol dependent mice, suppress the tremors which occur during withdrawal [35], [36]. Thus, following chronic ethanol treatment the glutamate receptors, particularly NMDA receptors, appear to mediate handling induced seizures, but not tremors, associated with ethanol withdrawal. Blocking NMDA receptors by memantine, a moderate affinity but clinically approved NMDA receptor antagonist, has been shown to reduce withdrawal induced-seizures and neurotoxicity in mice [100]. In clinical studies, memantine has been reported to attenuate cue-induced craving for alcohol and withdrawal symptoms associated with alcohol consumption [54], [55].

In rodents, exposure to alcohol has been reported to induce a significant increase in the expression and synaptic localization of α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptors (AMPA receptors) in brain regions with reward circuitry [15], [17]. AMPA receptors are ionotropic transmembrane receptors for glutamate that mediate fast synaptic transmission in the CNS. Infusion of an AMPA receptor inhibitor into the dorsomedial striatum exhibited promising results in reducing alcohol consumption in rats [113] and an AMPA/kainate receptor antagonist, topiramate, is reported to suppress withdrawal symptoms in humans [55]. Another drug, acamprosate, known to affect two neurotransmitter systems, GABA (as an agonist) and glutamate (as an NMDA receptor agonist and mGluR5 antagonist), has been reported to be effective in increasing complete abstinence rate as well as cumulative abstinence duration in several long-term placebo-controlled trials in alcohol-dependent patients [59], [80], [88]. In a large clinical trial involving 1383 patients in nine possible treatment groups, however, acamprosate neither alone nor with naltrexone or combined behavioral intervention showed statistically significant reduction in alcohol consumption over placebo [4]. Despite the inconsistencies in findings, overall advantageous pharmacological effects of acamprosate on alcohol consumption are such that it is approved for treatment of alcohol use disorders in Europe and USA.

Understanding how glutamate signaling may contribute to addiction needs to account for multiple pools of intra- and extra-cellular glutamate. Glutamate concentrations within glutamatergic synapses are determined by balance of vesicular release and energy dependent glutamate reuptake whereas that in the perisynaptic space is determined by diffusion from the synapse and extrasynaptic release/uptake [28], [104]. Perisynaptic glutamate regulates neurotransmission by stimulating group II glutamate receptors (mGluRs) in the presynaptic area [7], [121]. These are presynaptic receptors capable of inhibiting vesicular release [8], [18], [89]. Thus, presynaptic receptors permit cross talk between the two pools and altered nonvesicular glutamate release may contribute to pathological glutamate signaling linked to addiction [110]. As indicated above, under normal physiological conditions, plasma acetate concentration is low but with ethanol ingestion it rapidly increases to about 1 mM. Acetate is readily taken up and metabolized by brain [20], [81], [114], largely by astrocytes. Calculated rates of utilization of acetate from the blood were reported to be 15–25% of total glucose consumption in non-stimulated tissue and to range from 28 to 115% of estimated rates of glucose oxidation in astrocytes [20], [114]. During intoxication, blood acetate concentrations are higher for alcoholics than for controls [78]. This would facilitate metabolism of acetate by glial cells, as shown for laboratory animals chronically exposed to alcohol [107], [111].

Madayag and coworkers [68] reported that repeated N-acetylcysteine administration alters plasticity-dependent effects of cocaine. The authors provided evidence that the ability of N-acetylcysteine to regulate drug-seeking behavior results from actions on cystine-glutamate exchanger (system x_c^-) and is associated with suppression of the large release of glutamate that occurs in accumbens during drug seeking. *In vitro* measurements have been reported to show that cystine induced activation of cystine/glutamate exchange system x_c^- results in release of glial glutamate and stimulation of mGlu2 receptors [45]. This inhibits neuronal glutamate release. In cocaine addiction, decrease in reinstated cocaine seeking elicited by activating x_c^- with N-acetylcysteine is prevented by pretreatment with an mGlu2/3 antagonist [58]. Pharmacological treatment with N-acetylcysteine is reported to modulate responses to stress and depressive-like behaviors by increasing x_c^- expression and to indirectly activate mGlu2 receptors [76]. The fact that pharmacological blockade of the x_c^- system also blocked N-acetylcysteine effects on mGlu2 receptors further supports linkage of mGlu2 regulation in promoting resilience to stress. A consistent biomarker of mood-related behaviors associated with electrophysiological changes in the dentate gyrus is reduction in the mGlu2 receptors that regulate vesicular release of glutamate at synapses [77].

ETHANOL METABOLISM AND ASSOCIATED ALCOHOL WITHDRAWAL SYNDROME: A SYNOPSIS

Ethanol, through its role as a major but essentially unregulated metabolite for energy metabolism, causes a truly metabolic disease. An initial positive sense of well being is induced by “pushing” energy metabolism and increasing energy state to above its homeostatic set point. The positive sense of well being is part of the neurological signaling system that has evolved to guide adaptation to the environment. Positive reward feelings associated with adequate food, warmth etc. are enhanced and negative feelings associated with hunger, anxiety, etc are suppressed. This leads to a euphoric state with an enhanced sense of being well fed and unstressed. Short exposures are reversible and cause little long term injury, but as ethanol exposure continues there is progressive disruption of critical regulatory mechanisms that maintain metabolic homeostasis. The extent of disruption is dependent on individual tissues and their metabolic requirements with many other contributing parameters (including genetic differences, diet, etc). A key aspect of ethanol metabolism appears to be the decrease in $[AMP]_f$ and attendant decrease in AMPK activity. Both chronic decrease in AMPK activity and alcohol dependence are reported to cause decreased activity of oxidative phosphorylation, abnormally shaped mitochondria and other cellular morphological and enzymatic changes [14], [102]. In addition to the debilitating effects of prolonged ethanol consumption, it is necessary to deal with alcohol withdrawal syndrome. As described earlier, this is a complex response to stopping ethanol consumption and correlates with disappearance of ethanol

and its metabolic products from the body. The symptoms can be severe and life threatening but do not arise from a single cause. Due to the multiplicity in causes, a single drug is unlikely to fully alleviate the withdrawal symptoms. Effective treatment is likely to require simultaneous treatment of multiple causes. Identifiable primary metabolic contributions to alcohol withdrawal symptoms and treatments that should substantially suppress the symptoms include:

1). Correct the energy deficiency in specific brain cells (neurons and astrocytes). Derr and coworkers [22] reported this can be addressed by providing a metabolite "cocktail" (butyrate, lactate, and β-hydroxybutyrate) for energy metabolism and anaplerotic support of the citric acid cycle. Derr and Derr [23] quantified the tremulous and rigidity symptoms of alcohol withdrawal in rats and concluded that feeding β-hydroxybutyrate suppressed the tremulous, but not the rigidity, symptoms. Other suitable nutrients include branched chain amino acids and branched short chain fatty acids, as these also rapidly cross the blood-brain barrier. Importantly, each can provide both acetylCoA for oxidative metabolism and pyruvate for anaplerotic synthesis of glutamate. Given the importance of carnitine in acetate metabolism, supplementing with carnitine and/or acetyl-L-carnitine may also be helpful [71].

2A). Stabilize glutamate distribution and metabolism in the brain. Treatment with N-acetylcysteine has been reported to modulate responses to stress and depressive-like behaviors in several ways. These include increasing expression of glutamate-cystine antiporter (system x_c^-), increasing the concentrations of cysteine and cystine, and indirectly by activating mGlu2 receptors [2], [76]. This helps to rebalance glutamate homeostasis and glutamatergic neuronal/astrocyte interactions.

2B). Pharmacological compensation for increased sensitivity of the glutamatergic neural system. As noted above, chronic ethanol consumption induced suppression of glutamatergic activity results in increase in excitatory (NMDA, AMPA) and decrease in inhibitory (mGluR) glutamate receptors. The contribution of enhanced glutamatergic sensitivity and neurotoxicity of excessive NMDA receptor activity to ethanol withdrawal symptoms can be suppressed by antagonists for excitatory glutamate receptors. Among the glutamate receptor antagonists currently clinically approved or in trials are acamprosate, memantine, and MK-801. Suppression of glutamatergic sensitivity through the critical withdrawal period has been reported to decrease the severity of symptoms [55]. Once energy and glutamate metabolism have normalized, receptor numbers should return to near normal.

Alcohol dependence is not just a metabolic disease, and there are substantial genetic, psychological, and social contributions. Identifying and characterizing the metabolic consequences of ethanol ingestion and of alcohol dependence is important to understanding the disease and how to more effectively treat alcohol dependence. This is, however, just one piece of a complex medical and social problem.

CONFLICT OF INTEREST STATEMENT

Neither DFW nor FMM have any financial and personal relationships that could influence the work presented in this manuscript.

ACKNOWLEDGEMENT

The authors are indebted to Dr. Joseph J. Higgins for insightful and memorable discussions concerning ethanol metabolism and its regulation.

REFERENCES

[1]

R.A. Alberty

Thermodynamics of reactions of nicotinamide adenine dinucleotide and nicotinamide adenine dinucleotide phosphate

Arch Biochem Biophys, 307 (1) (1993), pp. 8-14

ArticleDownload PDFView Record in ScopusGoogle Scholar

[2]

P. Albrecht, J. Lewerenz, S. Dittmer, R. Noack, P. Maher, A. Methner

Mechanisms of oxidative glutamate toxicity: the glutamate/cystine antiporter system x_c^- as a neuroprotective drug target

CNS Neurol Disord: Drug Targets, 9 (3) (2010), pp. 373-382

CrossRefView Record in ScopusGoogle Scholar

[3]

A.L. Alhadeff, N. Goldstein, O. Park, M.L. Klima, A. Vargas, J.N. Betley

Natural and drug rewards engage distinct pathways that converge on coordinated hypothalamic and reward circuits

Neuron, 103 (2019), pp. 891-908

View Record in ScopusGoogle Scholar

[4]

R.F. Anton, S.S. O'Malley, D.A. Ciraulo, R.A. Cisler, D. Couper, D.M. Donovan, *et al.*

Combined pharmacotherapies and behavioral interventions for alcohol dependence: the COMBINE study: a randomized controlled trial

JAMA, 295 (17) (2006), pp. 2003-2017

CrossRefView Record in ScopusGoogle Scholar

[5]

P. Bach, E. Zois, S. Vollstädt-Klein, M. Kirsch, S. Hoffmann, A. Jorde, *et al.*

Association of the alcohol dehydrogenase gene polymorphism rs1789891 with gray matter brain volume, alcohol consumption, alcohol craving and relapse risk

Addict Biol, 24 (2017), pp. 110-120

Google Scholar

[6]

C.-I. Bäcklin

The equilibrium constant of the system ethanol, aldehyde, DPN, DPNH and H^+

Acta Chem Scand, 12 (1958), pp. 1279-1285

CrossRefView Record in ScopusGoogle Scholar

[7]

D.A. Baker, A. Madayag, L.V. Kristiansen, J.H. Meador-Woodruff, V. Haroutunian, I. Raju

Contribution of cystine-glutamate antiporters to the psychotomimetic effects of phencyclidine

Neuropsychopharm, 33 (7) (2008), pp. 1760-1772, 10.1038/sj.npp.1301532

CrossRefView Record in ScopusGoogle Scholar

[8]

A. Baskys, R.C. Malenka

Agonists at metabotropic glutamate receptors presynaptically inhibit EPSCs in neonatal rat hippocampus

J Physiol, 444 (1991), pp. 687-701

CrossRefView Record in ScopusGoogle Scholar

[9]

M. Bèlanger, I. Allaman, P.J. Magistretti

Brain energy metabolism: Focus on astrocyte-neuron metabolic cooperation

Cell Metab, 14 (2011), pp. 724-738

ArticleDownload PDFView Record in ScopusGoogle Scholar

[10]

P. Berríos-Cárcamo, M.E. Quintanilla, M. Herrera-Marschitz, V. Vasiliou, G. Zapata-Torres, M. Rivera-Meza

Racemic salsolinol and its enantiomers act as agonists of the m-opioid receptor by activating the Gi protein-adenylate cyclase pathway

Front Behav Neurosci, 10 (2016), p. 253, 10.3389/fnbeh.2016.00253

Google Scholar

[11]

L.A. Bertocci, J.G. Jones, C.R. Malloy, R.G. Victor, G.D. Thomas

Oxidation of lactate and acetate in rat skeletal muscle: analysis by 13C nuclear magnetic resonance spectroscopy

J Appl Physiol, 83 (1) (1997), pp. 32-39

CrossRefView Record in ScopusGoogle Scholar

[12]

W.F. Bosron, T. Ehrig, T.K. Li

Genetic factors in alcohol metabolism and alcoholism

Semin Liver Dis, 13 (1993), pp. 126-135

PMID: 8337601

CrossRefView Record in ScopusGoogle Scholar

[13]
R. Bűhler, D. Pestalozzi, M. Hess, J.-P. von Wartburg
Immunohistochemical localization of alcohol dehydrogenase in human kidney, endocrine organs, and brain
Pharmacol Biochem Behav, 18 (Suppl 1) (1983), pp. 55-59
ArticleDownload PDFView Record in ScopusGoogle Scholar
[14]
A.L. Bujak, J.D. Crane, J.S. Lally, R.J. Ford, S.J. Kang, I.A. Rebalka, *et al.*
AMPK activation of muscle autophagy prevents fasting induced hypoglycemia and myopathy during aging
Cell Metab, 21 (2015), pp. 883-890
ArticleDownload PDFView Record in ScopusGoogle Scholar
[15]
L.J. Chandler, D. Norwood, G. Sutton
Chronic ethanol upregulates NMDA and AMPA, but not kainate receptor subunit proteins in rat primary cortical cultures
Alcohol Clin Exp Res, 23 (2) (1999), pp. 363-370
View Record in ScopusGoogle Scholar
[16]
I.-Y. Chen, F.-R. Wang, X.-L. Sun, J. Zhou, L. Gao, Y.-L. Jiao, *et al.*
Chronic ethanol feeding impairs AMPK and MEF2 expression and is associated with GLUT4 decrease in rat myocardium
Expt Mol Med, 42 (3) (2010), pp. 205-215
CrossRefView Record in ScopusGoogle Scholar
[17]
D.T. Christian, N.J. Alexander, M.R. Diaz, S. Robinson, B.A. McCool
Chronic intermittent ethanol and withdrawal differentially modulate basolateral amygdala AMPA-type glutamate receptor function and trafficking
Neuropharm, 62 (7) (2012), pp. 2429-2438, 10.1016/j.neuropharm.2012.02.017
Google Scholar
[18]
A.J. Cochilla, A.S. Simon
Metabotropic glutamate receptor-mediated control of neurotransmitter release
Neuron, 20 (1998), pp. 1007-1016
ArticleDownload PDFView Record in ScopusGoogle Scholar
[19]
D.W. Crabb, S. Liangpunsakul
Alcohol and lipid metabolism
J Gastroenterol Hepatol, 3 (2006), pp. S56-S60
CrossRefView Record in ScopusGoogle Scholar
[20]
N.F. Cruz, A. Lasater, H.R. Zielke, G.A. Dienel
Activation of astrocytes in brain of conscious rats during acoustic stimulation: acetate utilization in working brain
J Neurochem, 92 (2005), pp. 934-947
CrossRefView Record in ScopusGoogle Scholar
[21]
D.K. Deelchand, A.A. Shestov, D.M. Koski, K. Uğurbil, P.-G. Henry
Acetate transport and utilization in the rat brain
J Neurochem, 109 (Suppl 1) (2009), pp. 46-54, 10.1111/j.1471-4159.2009.05895.x
CrossRefView Record in ScopusGoogle Scholar
[22]

R.F. Derr

The ethanol withdrawal syndrome: a consequence of lack of substrate for a cerebral Krebs-cycle

J Theor Biol, 106 (1984), pp. 375-381

ArticleDownload PDFView Record in ScopusGoogle Scholar

[23]

R.F. Derr, M.I. Derr

Separation of the tremorous and muscular rigidity components of the ethanol withdrawal syndrome in the rat

Physiol Behav, 38 (1986), pp. 1-3

ArticleDownload PDFView Record in ScopusGoogle Scholar

[24]

R.F. Derr, K. Draves, M. Derr

Abatement by acetate of an ethanol withdrawal syndrome

Life Sci, 29 (1981), pp. 1787-1790

ArticleDownload PDFView Record in ScopusGoogle Scholar

[25]

A. Dettling, F. Fischer, S. Böhler, F. Ulrich, G. Skopp, M. Graw, *et al.*

Ethanol elimination rates in men and women in consideration of the calculated liver weight

Alcohol, 41 (2007), pp. 415-420

ArticleDownload PDFView Record in ScopusGoogle Scholar

[26]

M. Erecińska, D. Wontorsky, D.F. Wilson

Aspartate transport in synaptosomes from rat brain

J Biol Chem, 258 (1983), pp. 9069-9077

ArticleDownload PDFView Record in ScopusGoogle Scholar

[27]

J. Farrès, A. Moreno, B. Crosas, J.M. Peralba, A. Allali-Hassani, L. Hjelmqvist, *et al.*

Alcohol dehydrogenase of class IV (ss-ADH) from human stomach. cDNA sequence and structure/function relationships

Eur J Biochem, 224 (1994), pp. 549-557

CrossRefView Record in ScopusGoogle Scholar

[28]

D.E. Featherstone

Intercellular glutamate signaling in the nervous system and beyond

ACS Chem Neurosci, 1 (2010), pp. 4-12, 10.1021/cn900006n

CrossRefView Record in ScopusGoogle Scholar

[29]

O.A. Forsander

Influence of the metabolism of ethanol on the lactate/pyruvate ratio of rat-liver slices

Biochem J, 98 (1966), pp. 244-247

CrossRefView Record in ScopusGoogle Scholar

[30]

D.W. Foster

Malonyl-CoA: the regulator of fatty acid synthesis and oxidation

J Clin Invest, 122 (6) (2012), pp. 1958-1959, 10.1172/JCI63967

CrossRefView Record in ScopusGoogle Scholar

[31]

Gary Frost, Michelle L. Sleeth, Meliz Sahuri-Arisoylu, Blanca Lizarbe, Sebastian Cerdan, Leigh Brody, Jelena Anastasovska, Samar Ghourab, Mohammed Hankir, Shuai Zhang, David Carling, Jonathan R. Swann, Glenn Gibson, Alexander Viardot, Douglas Morrison,

E.L. Thomas, Jimmy D. Bell

The short-chain fatty acid acetate reduces appetite via a central homeostatic mechanism

Nat Commun, 5 (1) (2014), 10.1038/ncomms4611

Google Scholar

[32]

A. Galli, D. Price, D. Crabb

High-level expression of rat class I alcohol dehydrogenase is sufficient for ethanol-induced fat accumulation in transduced HeLa cells

Hepatology, 29 (1999), pp. 1164-1170

View Record in ScopusGoogle Scholar

[33]

G.L. Gessa, R. Agabio, M.A.M. Carai, C. Lobina, M. Pani, R. Reali, *et al.*

Mechanism of the antialcohol effect of gamma-hydroxybutyric acid

Alcohol, 20 (2000), pp. 271-276

ArticleDownload PDFView Record in ScopusGoogle Scholar

[34]

R.A. Gonzales, F. Weiss

Suppression of ethanol-reinforced behavior by naltrexone is associated with attenuation of the ethanol-induced increase in dialysate dopamine levels in the nucleus accumbens

J Neurosci, 18 (24) (1998), pp. 10663-10671

CrossRefView Record in ScopusGoogle Scholar

[35]

K.A. Grant, P. Valverius, M. Hudspith, B. Tabakoff

Ethanol withdrawal seizures and the NMDA receptor complex

Eur J Phamacol, 176 (1990), pp. 289-296

ArticleDownload PDFView Record in ScopusGoogle Scholar

[36]

K. Gulya, K.A. Grant, P. Valverius, P.L. Hoffman, B. Tabakoff

Brain regional specificity and time course of changes in the NMDA receptor-ionophore complex during ethanol withdrawal

Brain Res, 547 (1991), pp. 129-134

View Record in ScopusGoogle Scholar

[37]

R.K. Handa, M.R. DeJoseph, L.D. Singh, R.A. Hawkins, S.P. Singh

Glucose transporters and glucose utilization in rat brain after acute ethanol administration

Metab Brain Dis, 15 (2000), pp. 211-222

View Record in ScopusGoogle Scholar

[38]

D.G. Hardie

Keeping the home fires burning: AMP-activated protein kinase

J R Soc Interface, 15 (2018), p. 20170774, 10.1098/rsif.2017.0774

CrossRefView Record in ScopusGoogle Scholar

[39]

D.G. Hardie, B.E. Schaffer, A. Brunet

AMPK: an energy-sensing pathway with multiple inputs and outputs

Trends Cell Biol, 26 (2016), pp. 190-201, 10.1016/j.tcb.2015.10.013

ArticleDownload PDFView Record in ScopusGoogle Scholar

[40]

I.E. Hassinen, M.A. Härkönen, R.H. Ylikahri

Metabolic effects of acetaldehyde in the intact rat brain cortex and its subcellular fractions

Brain Res, 70 (1974), pp. 301-312

ArticleDownload PDFView Record in ScopusGoogle Scholar

[41]

L. Hertz, D.L. Rothman

Glutamine-glutamate cycle flux is similar in cultured astrocytes and brain and both glutamate production and oxidation are mainly catalyzed by aspartate aminotransferase

Biology, 6 (17) (2017), pp. 1-21, 10.3390/biology6010017

CrossRefView Record in ScopusGoogle Scholar

[42]

S. Herzig, R.J. Shaw

AMPK: guardian of metabolism and mitochondrial homeostasis

Nat Rev Mol Cell Biol, 19 (2) (2018), pp. 121-135, 10.1038/nrm.2017.95

CrossRefView Record in ScopusGoogle Scholar

[43]

L. Jiang, B.I. Gulanski, H.M. De Feyter, S.A. Weinzimer, B. Pittman, E. Guidone, *et al.*

Increased brain uptake and oxidation of acetate in heavy drinkers

J Clin Invest, 123 (4) (2013), pp. 1605-1614, 10.1172/JCI65153

View Record in ScopusGoogle Scholar

[44]

A. Juhlin-Dannfelt

Ethanol effects of substrate utilization by the human brain

Scand J Clin Lab Invest, 37 (5) (1977), pp. 443-449

ISSN: 0036-5513, 0036-5513

Google Scholar

[45]

P.W. Kalivas

The glutamate homeostasis hypothesis of addiction

Nat Rev Neurosci, 10 (2009), pp. 561-572

CrossRefView Record in ScopusGoogle Scholar

[46]

E. Karahanian, M. Rivera-Meza, L. Tampier, M.E. Quintanilla, M. Herrera-Marschitz, Y. Israel

Long-term inhibition of ethanol intake by the administration of an aldehyde dehydrogenase-2 (ALDH2)-coding lentiviral vector into the ventral tegmental area of rats

Addict Biol, 20 (2015), pp. 336-344, 10.1111/adb.12130

CrossRefView Record in ScopusGoogle Scholar

[47]

J.T. Kerr, D.S. Maxwell, D.W. Crabb

Immunocytochemistry of alcohol dehydrogenase in the rat central nervous system

Alcoholisn: Clin Exptl Res, 13 (6) (1989), pp. 730-736

CrossRefView Record in ScopusGoogle Scholar

[48]

K. Kiviluoma, I. Hassinen

Role of acetaldehyde and acetate in the development of ethanol-induced cardiac lipidosis, studied in isolated perfused rat hearts

Alcoholism: Clin Exptl Res, 1 (2) (1982), pp. 169-175

Google Scholar

[49]
R. Kjøbsted, J.R. Hingst, J. Fentz, M. Foretz, M.-N. Sanz, C. Pehmøller, *et al.*
AMPK in skeletal muscle function and metabolism
FASEB J, 32 (2018), pp. 1741-1777, 10.1096/fj.201700442R
CrossRefView Record in ScopusGoogle Scholar
[50]
H.A. Krebs
The citric acid cycle
Nobel Lecture (1953)
Google Scholar
[51]
H.A. Krebs
The Croonian lecture. Gluconeogenesis
Proc R Soc Lond B, 159 (1964), pp. 545-564, 10.1098/rspb.1964.0019)
View Record in ScopusGoogle Scholar
[52]
H.A. Krebs, R.A. Freedland, M. Stubbs
Inhibition of hepatic gluconeogenesis by ethanol
Biochem J, 112 (1969), pp. 117-124
CrossRefView Record in ScopusGoogle Scholar
[53]
R.A. Kreisberg, W.C. Owen, A.M. Siegal
Ethanol-induced hyperlacticacidemia: inhibition of lactate utilization
J Clin Invest, 50 (1971), pp. 166-174
CrossRefView Record in ScopusGoogle Scholar
[54]
E.M. Krupitsky, O. Neznanova, D. Masalov, A.M. Burakov, T. Didenko, T. Romanova, *et al.*
Effect of Memantine on cue-induced alcohol craving in recovering alcohol-dependent patients
Am J Psychiatry, 164 (3) (2007), pp. 519-523
View Record in ScopusGoogle Scholar
[55]
E.M. Krupitsky, A.A. Rudenko, A.M. Burakov, T.Y. Slavina, A.A. Grinenko, B. Pittman, *et al.*
Antiglutamatergic strategies for ethanol detoxification: comparison with placebo and diazepam
Alcohol Clin Exp Res, 31 (2007), pp. 604-611
[PubMed: 17374039]
View Record in ScopusGoogle Scholar
[56]
M. Kurnik-Łucka, P. Panula, A. Bugajski, K. Gil
Salsolinol: an unintelligible and double-faced molecule—lessons learned from in vivo and in vitro experiments
Neurotox Res, 33 (2018), pp. 485-514, 10.1007/s12640-017-9818-6
CrossRefView Record in ScopusGoogle Scholar
[57]
J. Lee, V.A. Ramchandani, K. Hamazaki, E.A. Engleman, W.J. McBride, T.K. Li, *et al.*
A critical evaluation of influence of ethanol and diet on salsolinol enantiomers in humans and rats
Alcohol Clin Exp Res, 34 (2010), pp. 242-250
View Record in ScopusGoogle Scholar
[58]

J. Lewerenz, S.J. Hewett, Y. Huang, M. Lambros, P.W. Gout, P.W. Kalivas, *et al.*

The cystine/glutamate antiporter system X_c^- in health and disease: from molecular mechanisms to novel therapeutic opportunities

Antiox Redox Signal, 18 (2013), pp. 522-555

CrossRefView Record in ScopusGoogle Scholar

[59]

J.P. Lhuintre, N. Moore, G. Tran, L. Steru, S. Langrenon, M. Daoqust, *et al.*

Acamprosate appears to decrease alcohol intake in weaned alcoholics

Alcohol Alcohol, 25 (6) (1990), pp. 613-622

CrossRefView Record in ScopusGoogle Scholar

[60]

D. Li, H. Zhao, J. Gelernter

Strong association of the alcohol dehydrogenase 1B gene (*ADH1B*) with alcohol dependence and alcohol-induced medical diseases

Biol Psychiatry, 70 (6) (2011), pp. 504-512, 10.1016/j.biopsych.2011.02.024

ArticleDownload PDFView Record in ScopusGoogle Scholar

[61]

S. Liangpunsakul, S.-E. Wou, Y. Zeng, R.A. Ross, H.N. Jayaram, D.W. Crabb

Effect of ethanol on hydrogen peroxide-induced AMPK phosphorylation

Am J Physiol Gastrointest Liver Physiol, 295 (6) (2008), pp. G1173-G1181

CrossRefView Record in ScopusGoogle Scholar

[62]

C.S. Lieber

Relationships between nutrition, alcohol use, and liver disease

Alcohol Res Health, 27 (3) (2003), pp. 220-231

View Record in ScopusGoogle Scholar

[63]

C.S. Lieber, L.M. DeCarli

An experimental model of alcohol feeding and liver injury in the baboon

J Med Primatol, 3 (1974), pp. 153-163

CrossRefView Record in ScopusGoogle Scholar

[64]

C.S. Lieber, D.P. Jones, L.M. Decarli

Effects of prolonged ethanol intake: production of fatty liver despite adequate diets

J Clin Invest, 44 (1965), pp. 1009-1021

CrossRefView Record in ScopusGoogle Scholar

[65]

D. Lindberg, A.M.C. Ho, L. Peyton, D.-S. Choi

Chronic ethanol exposure disrupts lactate and glucose homeostasis and induces dysfunction of the astrocyte-neuron lactate shuttle in the brain

Alcoholism: Clin Exp Res, 43 (9) (2019), pp. 1838-1847

CrossRefView Record in ScopusGoogle Scholar

[66]

C. Loguercio, Blanco F Del Vecchio, V. De Girolamo, D. Disalvo, G. Nardi, A. Parente, *et al.*

Ethanol consumption, amino acid and glutathione blood levels in patients with and without chronic liver disease

Alcohol Clin Exp Res, 23 (11) (1999), pp. 1780-1784

View Record in ScopusGoogle Scholar

[67]
F. Lundquist, N. Tygstrup, K. Winkler, K. Mellemgaard, S. Munck-Petersen
Ethanol metabolism and production of free acetate in the human liver
J Clin Invest, 41 (5) (1962), pp. 955-961
CrossRefView Record in ScopusGoogle Scholar
[68]
A. Madayag, D. Lobner, K.S. Kau, J.R. Mantsch, O. Abdulhameed, M. Hearing, *et al.*
Repeated N-acetylcysteine administration alters plasticity-dependent effects of cocaine
J Neurosci, 27 (2007), pp. 13968-13976
View Record in ScopusGoogle Scholar
[69]
S. Mandayam, M.M. Jamal, T.R. Morgan
Epidemiology of alcoholic liver disease
Semin Liver Dis, 24 (2004), pp. 217-232
View Record in ScopusGoogle Scholar
[70]
S.E. Martínez, J. Vaglenova, J. Sabriá, M.C. Martínez, J. Farrès, X. Parès
Distribution of alcohol dehydrogenase mRNA in the rat central nervous system: Consequences for brain ethanol and retinoid metabolism
Eur J Biochem, 268 (2001), pp. 5045-5056
View Record in ScopusGoogle Scholar
[71]
G. Martinotti, D. Reina, M. Di Nicola, S. Andreoli, D. Tedeschi, I. Ortolani, *et al.*
Acetyl-L-carnitine for alcohol craving and relapse prevention in anhedonic alcoholics: a randomized, double-blind, placebo-controlled pilot trial
Alcohol Alcohol, 45 (5) (2010), pp. 449-455
CrossRefView Record in ScopusGoogle Scholar
[72]
F.M. Matschinsky, D.F. Wilson
The central role of glucokinase in glucose homeostasis: a perspective 50 years after demonstrating the presence of the enzyme in islets of Langerhans
Frontiers Physiol, 10 (148) (2019), p. 2019, 10.3389/fphys.2019.00148. eCollection
Google Scholar
[73]
J.D. McGarry, G.F. Leatherman, W. Daniel, D.W. Foster
Carnitine palmitoyltransferase 1: The site of inhibition of hepatic fatty acid oxidation by malonyl-CoA
J Biol Chem, 253 (12) (1978), pp. 4128-4136
ArticleDownload PDFView Record in ScopusGoogle Scholar
[74]
P. Mews, G. Egervari, R. Nativio, S. Sidoli, G. Donahue, S.I. Lombroso, *et al.*
Alcohol metabolism contributes to brain histone acetylation
Nature, 574 (2019), pp. 717-721
CrossRefView Record in ScopusGoogle Scholar
[75]
P.M.A. Muneer, S. Alikunju, A.M. Szlachetka, J. Haorah
Inhibitory effects of alcohol on glucose transport across the blood–brain barrier leads to neurodegeneration: preventive role of acetyl-L-carnitine

Psychopharmacology, 214 (2011), pp. 707-718
CrossRefView Record in ScopusGoogle Scholar

[76]
C. Nasca, B. Bigio, D. Zelli, P. de Angelis, T. Lau, M. Okamoto, *et al.*
Role of the astroglial glutamate exchanger xCT in ventral hippocampus in resilience to stress
Neuron, 96 (2017), pp. 402-413
View Record in ScopusGoogle Scholar

[77]
C. Nasca, D. Zelli, B. Bigio, S. Piccinin, S. Scaccianoce, R. Nisticó, *et al.*
Stress dynamically regulates behavior and glutamatergic gene expression in hippocampus by opening a window of epigenetic plasticity
PNAS, 112 (2015), pp. 14960-14965
CrossRefGoogle Scholar

[78]
H. Nuutinen, K. Lindros, P. Hekali, M. Salaspuro
Elevated blood acetate as indicator of fast ethanol elimination in chronic alcoholics
Alcohol, 2 (1985), pp. 623-626
ArticleDownload PDFView Record in ScopusGoogle Scholar

[79]
M.F. Olive, R.M. Cleva, P.W. Kalivas, R.J. Malcolm
Glutamatergic medications for the treatment of drug and behavioral addictions
Pharmacol Biochem Behav, 100 (4) (2012), pp. 801-810, 10.1016/j.pbb.2011.04.015
ArticleDownload PDFView Record in ScopusGoogle Scholar

[80]
F.M. Paille, J.D. Guelfi, A.C. Perkins, R.J. Royer, L. Steru, P. Parot
Double-blind randomized multicentre trial of acamprosate in maintaining abstinence from alcohol
Alcohol Alcohol, 30 (2) (1995), pp. 239-247
View Record in ScopusGoogle Scholar

[81]
A.B. Patel, R.A. de Graaf, D.L. Rothman, K.L. Behar, G.F. Mason
Evaluation of cerebral acetate transport and metabolic rates in the rat brain in vivo using 1 H-[13C]-NMR
J Cerebral Blood Flow Metab, 30 (2010), pp. 1200-1213
CrossRefView Record in ScopusGoogle Scholar

[82]
C.T. Putman, L.L. Spriet, E. Hultman, D.J. Dyck, G.J. Heigenhauser
Skeletal muscle pyruvate dehydrogenase activity during acetate infusion in humans
Am J Physiol, 268 (1995), pp. E1007-E1017
CrossRefView Record in ScopusGoogle Scholar

[83]
M.E. Quintanilla, Y. Israel, A. Sapag, L. Tampier
The UChA and UChB rat lines: metabolic and genetic differences influencing ethanol intake
Addict Biol, 11 (2006), pp. 310-323, 10.1111/j.1369-1600.2006.00030.x
CrossRefView Record in ScopusGoogle Scholar

[84]
N.H. Raskin, L. Sokoloff
Alcohol dehydrogenase activity in rat brain and liver
J Neurochem, 17 (1970), pp. 1677-1687

CrossRefView Record in ScopusGoogle Scholar

[85]

R.F. Ritzmann, B. Tabakoff

Body temperature in mice: a quantitative measure of alcohol tolerance and physical dependence

J Pharmacol Exp Ther, 199 (1976), pp. 158-170

View Record in ScopusGoogle Scholar

[86]

M. Rivera-Meza, D. Vásquez, M.E. Quintanilla, D. Lagos, B. Rojas, M. Herrera-Marschitz, *et al.*

Activation of mitochondrial aldehyde dehydrogenase (ALDH2) by ALDA-1 reduces both the acquisition and maintenance of ethanol intake in rats: A dual mechanism?

Neuropharm, 146 (2019), pp. 175-183

ArticleDownload PDFView Record in ScopusGoogle Scholar

[87]

B.D. Rowlands, M. Klugmann, C.D. Rae

Acetate metabolism does not reflect astrocytic activity, contributes directly to GABA synthesis, and is increased by silent information regulator 1 activation

J Neurochem, 140 (6) (2017), pp. 903-918, 10.1111/jnc.13916

Epub 2017 Jan 23

CrossRefView Record in ScopusGoogle Scholar

[88]

H. Sass, M. Soyka, K. Mann, W. Zieglgänsberger

Relapse prevention by acamprosate. Results from a placebo-controlled study on alcohol dependence

Arch Gen Psychiatry, 53 (8) (1996), pp. 673-680

CrossRefView Record in ScopusGoogle Scholar

[89]

D.D. Schoepp

Unveiling the functions of presynaptic metabotropic glutamate receptors in the central nervous system

J Pharm Exptl Ther, 299 (2001), pp. 12-20

View Record in ScopusGoogle Scholar

[90]

A. Schousboe, S. Scafidi, L.K. Bak, H.S. Waagepetersen, M.C. McKenna

Glutamate metabolism in the brain focusing on astrocytes

Adv Neurobiol, 11 (2014), pp. 13-30, 10.1007/978-3-319-08894-5_2

CrossRefView Record in ScopusGoogle Scholar

[91]

M. Schreckenberger, R. Amberg, A. Scheurich, M. Lochmann, W. Tichy, A. Klega, *et al.*

Acute alcohol effects on neuronal and attentional processing: striatal reward system and inhibitory sensory interactions under acute ethanol challenge

Neuropsychopharmacology, 29 (2004), pp. 1527-1537

View Record in ScopusGoogle Scholar

[92]

M.C. Scrutton, M.F. Utter

Pyruvate carboxylase III. Some physical and chemical properties of the highly purified enzyme

J Biol Chem, 240 (1) (1965), pp. 1-9

ArticleDownload PDFView Record in ScopusGoogle Scholar

[93]

K.-Z. Shen, Y.-N. Wu, A.C. Munhall, S.W. Johnson

AMP Kinase regulates ligand-gated K-ATP channels in substantia nigra dopamine neurons

Neuroscience, 330 (2016), pp. 219-228

ArticleDownload PDFView Record in ScopusGoogle Scholar

[94]

K.-Z. Shen, V. Yakhnitsa, A.C. Munhall, S.W. Johnson

AMP kinase regulates K-ATP currents evoked by NMDA receptor stimulation in rat subthalamic nucleus neurons

Neuroscience, 274 (2014), pp. 138-152, 10.1016/j.neuroscience.2014.05.031

ArticleDownload PDFView Record in ScopusGoogle Scholar

[95]

G.I. Smith, A.E. Jeukendrup, D. Ball

Sodium acetate induces a metabolic alkalosis but not the increase in fatty acid oxidation observed following bicarbonate ingestion in humans

J Nutr, 137 (2007), pp. 1750-1756

CrossRefView Record in ScopusGoogle Scholar

[96]

L.D. Snell, B. Tabakoff, P.L. Hoffman

Radioligand binding to the N-methyl-D-aspartate receptor/ionophore complex: alterations by ethanol in vitro and by chronic in vivo ethanol ingestion

Brain Res, 602 (1993), pp. 91-98

ArticleDownload PDFView Record in ScopusGoogle Scholar

[97]

K. Souda, T. Kawasaki, T. Yoshimi

Effects of acute and chronic ethanol administration on hepatic hemodynamics and hepatic oxygen consumption in awake rats

J Hepatology, 24 (1996), pp. 101-108

ArticleDownload PDFCrossRefView Record in ScopusGoogle Scholar

[98]

M.P. Srinivasan, K.K. Bhopale, S.M. Amer, J. Wan, L. Kaphalia, G.S. Ansari, *et al.*

Linking Dysregulated AMPK signaling and ER stress in ethanol-induced liver injury in hepatic alcohol dehydrogenase deficient deer mice

Biomolecules, 9 (560) (2019), pp. 1-17, 10.3390/biom9100560

CrossRefGoogle Scholar

[99]

J.W. Starnes, D.F. Wilson, M. Erecinska

Substrate dependence of metabolic state and coronary flow in perfused rat heart

Am J Physiol, 249 (1985), pp. H799-H806

View Record in ScopusGoogle Scholar

[100]

T.D. Stepanyan, J.M. Farook, A. Kowalski, E. Kaplan, S. Barron, J.M. Littleton

Alcohol withdrawal-induced hippocampal neurotoxicity in vitro and seizures in vivo are both reduced by memantine

Alcoholism: Clin Exptl Res, 32 (2008), pp. 2128-2135

CrossRefView Record in ScopusGoogle Scholar

[101]

B. Tabakoff, R.F. Ritzmann

The effects of 6-hydroxydopamine on tolerance to and dependence on ethanol

J Pharmacol Exp Ther, 203 (1977), pp. 319-331

View Record in ScopusGoogle Scholar

[102]

M.R. Teli, C.P. Day, A.D. Burt, M.K. Bennett, O.F.W. James

Determinants of progression to cirrhosis or fibrosis in pure alcoholic fatty liver

Lancet, 346 (1995), pp. 987-990

ArticleDownload PDFView Record in ScopusGoogle Scholar

[103]

R.G. Thurman, H.G. Ley, R. Scholz

Hepatic microsomal ethanol oxidation. Hydrogen peroxide formation and the role of catalase. Euro

J Biochem, 25 (3) (1972), pp. 420-430

[PubMed: 4402915]

CrossRefView Record in ScopusGoogle Scholar

[104]

W. Timmerman, B.H. Westerink

Brain microdialysis of GABA and glutamate: what does it signify?

Synapse, 27 (3) (1997), pp. 242-261

View Record in ScopusGoogle Scholar

[105]

K. Tomita, G. Tamiya, S. Ando, N. Kitamura, H. Koizumi, S. Kato, *et al.*

AICAR, an AMPK activator, has protective effects on alcohol-induced fatty liver in rats

Alcohol Clin Exp Res, 29 (12) (2005), pp. 240S-245S

CrossRefGoogle Scholar

[106]

G.E. Tsai, P. Ragan, R. Chang, S. Chen, V.M.I. Linnoila, J.T. Coyle

Increased glutamatergic neurotransmission and oxidative stress after alcohol withdrawal

Am J Psychiatry, 155 (1998), pp. 726-732

View Record in ScopusGoogle Scholar

[107]

N.D. Volkow, S.W. Kim, G.-J. Wang, D. Alexoff, J. Logan, L. Muench, *et al.*

Acute alcohol intoxication decreases glucose metabolism but increases acetate uptake in the human brain

NeuroImage, 64 (2013), pp. 277-283

ArticleDownload PDFView Record in ScopusGoogle Scholar

[108]

N.D. Volkow, G.-J. Wang, D. Franceschi, J.S. Fowler, P.K. Thanos, L. Maynard, *et al.*

Low doses of alcohol substantially decrease glucose metabolism in the human brain

NeuroImage, 29 (2006), pp. 295-301

ArticleDownload PDFView Record in ScopusGoogle Scholar

[109]

N.D. Volkow, G.-J. Wang, R. Hitzemann, J.S. Fowler, J.E. Overall, G. Burr, *et al.*

Recovery of brain glucose metabolism in detoxified alcoholics

J Psychiatry, 151 (1994), pp. 178-183

View Record in ScopusGoogle Scholar

[110]

N.D. Volkow, G.-J. Wang, E.S. Kojori, J.S. Fowler, H. Benveniste, D. Tomasi

Alcohol decreases baseline brain glucose metabolism more in heavy drinkers than controls but has no effect on stimulation-induced metabolic increases

J Neurosci, 35 (7) (2015), pp. 3248-3255

View Record in ScopusGoogle Scholar

[111]

J. Wang, H. Du, L. Jiang, X. Ma, R.A. de Graaf, K.L. Behar, *et al.*

Oxidation of ethanol in the rat brain and effects associated with chronic ethanol exposure

Proc Natl Acad Sci, 110 (35) (2013), pp. 14444-14449

CrossRefView Record in ScopusGoogle Scholar

[112]

G.J. Wang, N.D. Volkow, D. Franceschi, J.S. Fowler, P.K. Thanos, N. Scherbaum, *et al.*

Regional brain metabolism during alcohol intoxication

Alcohol Clin Exp Res, 24 (2000), pp. 822-829

View Record in ScopusGoogle Scholar

[113]

J. Wang, S.B. Hamida, E. Darcq, W. Zhu, S.L. Gibb, M.F. Lanfranco, *et al.*

Ethanol-mediated facilitation of AMPA receptor function in the dorsomedial striatum: Implications for alcohol drinking behavior

J Neurosci, 32 (43) (2012), pp. 15124-15132

View Record in ScopusGoogle Scholar

[114]

R.A. Waniewski, D.L. Martin

Preferential utilization of acetate by astrocytes is attributable to transport

J Neurosci, 18 (1998), pp. 5225-5233

CrossRefView Record in ScopusGoogle Scholar

[115]

D.F. Wilson

Oxidative phosphorylation: regulation and role in cellular and tissue metabolism

J Physiol, 595 (2017), pp. 7023-7038, 10.1113/JP273839

CrossRefView Record in ScopusGoogle Scholar

[116]

D.F. Wilson, A.T.J. Cember, F.M. Matschinsky

The thermodynamic basis of glucose-stimulated insulin release: a model of the core mechanism

Physiol Rep, 5 (12) (2017), 10.14814/phy2.13327

pii: e13327 PMID: 28655753

Google Scholar

[117]

D.F. Wilson, A.T.J. Cember, F.M. Matschinsky

Glutamate dehydrogenase: Role in regulating metabolism and insulin release in pancreatic β-cells

J Appl Physiol, 125 (2) (2018), pp. 419-428

doi: 0.1152/japplphysiol.01077.2017. Epub 2018 Apr 12.PMID:29648519

CrossRefView Record in ScopusGoogle Scholar

[118]

D.F. Wilson, F.M. Matchinsky

Metabolic homeostasis: Oxidative phosphorylation and the metabolic requirements of higher plants and animals

J Appl Physiol, 125 (2019), pp. 1183-1192

CrossRefGoogle Scholar

[119]

D.F. Wilson, F.M. Matschinsky

Hyperbaric oxygen toxicity in brain: a case of hyperoxia induced hypoglycemic brain syndrome

Med Hypotheses, 132 (2019), Article 109375, 10.1016/j.mehy.2019.109375

PMID: 31454640

ArticleDownload PDFView Record in ScopusGoogle Scholar

[120]

D.F. Wilson, A. Pastuszko

Transport of cysteate by synaptosomes isolated from rat brain: Evidence that it utilizes the same transporter as aspartate, glutamate and cysteine sulfinate

J Neurochem, 47 (1986), pp. 1091-1097

View Record in ScopusGoogle Scholar

[121]

Z.-X. Xi, S. Ramamoorthy, D.A. Baker, H. Shen, D.J. Samuvel, P.W. Kalivas

Modulation of group II metabotropic glutamate receptor signaling by chronic cocaine

J Pharm Exptl Ther, 303 (2002), pp. 608-615

View Record in ScopusGoogle Scholar

[122]

T. Yamamoto, M. Yamakawa

Effect of acetate administration on rats with chronic uremia

Artif Organs, 11 (1987), pp. 208-213

CrossRefView Record in ScopusGoogle Scholar

[123]

W. Yang, A.C. Munhall, S.W. Johnson

AMP-activated protein kinase slows D2 dopamine autoreceptor desensitization in substantia nigra neurons

Neuropharmacology, 158 (2019), Article 107705

ArticleDownload PDFView Record in ScopusGoogle Scholar

[124]

M. You, M. Matsumoto, C.M. Pacold, W.K. Cho, D.W. Crabb

The role of AMP-activated protein kinase in the action of ethanol in the liver

Gastroenterology, 127 (2004), pp. 1798-1808

ArticleDownload PDFView Record in ScopusGoogle Scholar

[125]

A.C. Yu, J. Drejer, L. Hertz, A. Schousboe

Pyruvate carboxylase activity in primary cultures of astrocytes and neurons

J Neurochem, 41 (1983), pp. 1484-1487

CrossRefView Record in ScopusGoogle Scholar

[126]

S.M. Zimatkin, R.A. Deitrich

Ethanol metabolism in the brain

Addiction Biol, 2 (1997), pp. 387-399

View Record in ScopusGoogle Scholar

[127]

S.M. Zimatkin, S.P. Pronko, V. Vasiliou, F.J. Gonzalez, R.A. Deitrich

Enzymatic mechanisms of ethanol oxidation in the brain

Alcoholism: Clin Exptl Res, 30 (2006), pp. 1500-1505

PMID: 16930212

CrossRefView Record in ScopusGoogle Scholar

[128]

S.M. Zimatkin, U.K. Rout, M. Koivusalo, R. Bühler, K.O. Lindros

Regional distribution of low-Km mitochondrial aldehyde dehydrogenase in the rat central nervous system

Alcohol Clin Exp Res, 16 (6) (1992), pp. 1162-1167

29.11: GUT MICROBIOTA-DERIVED METABOLITES AS CENTRAL REGULATORS IN METABOLIC DISORDERS

ABSTRACT

Metabolic disorders represent a growing worldwide health challenge due to their dramatically increasing prevalence. The gut microbiota is a crucial actor that can interact with the host by the production of a diverse reservoir of metabolites, from exogenous dietary substrates or endogenous host compounds. Metabolic disorders are associated with alterations in the composition and function of the gut microbiota. Specific classes of microbiota-derived metabolites, notably bile acids, short-chain fatty acids, branched-chain amino acids, trimethylamine N-oxide, tryptophan and indole derivatives, have been implicated in the pathogenesis of metabolic disorders. This review aims to define the key classes of microbiota-derived metabolites that are altered in metabolic diseases and their role in pathogenesis. They represent potential biomarkers for early diagnosis and prognosis as well as promising targets for the development of novel therapeutic tools for metabolic disorders.

KEY MESSAGES

- Metabolic disorders, a growing worldwide health challenge, are associated with alterations in the composition and function of the gut microbiota.
- Microbial metabolites are key factors in host-microbiota cross-talk.
- Specific classes of microbiota-derived metabolites, notably bile acids, short-chain fatty acids, branched-chain amino acids, trimethylamine N-oxide, tryptophan and indole derivatives, have been strongly implicated in the pathogenesis of metabolic disorders.
- Gut microbiota-derived metabolites represent potential biomarkers for the early diagnosis and show promise for identifying targets for the development of novel therapeutic tools for metabolic disorders.

INTRODUCTION

The human intestine harbours a complex and diverse system of mutualistic microorganisms, consisting of bacteria, fungi, viruses, archaea and protozoa. This rich ecosystem contributes to a large number of physiological functions: fermentation of indigestible dietary components and vitamin synthesis, defenses against pathogens, host immune system maturation and maintenance of gut barrier function. Thus, this central regulator, sometimes qualified as the ‘second brain’, plays a significant role in maintaining host physiology and homeostasis. All the species interconnected in the gut produce an extremely diverse reservoir of metabolites from exogenous dietary components and/or endogenous compounds generated by microorganisms and the host. Notably, while food is generally examined for calories and macronutrients and micronutrients, microbial metabolism (and even human enzymes) recognises food molecules and transforms them into metabolites. These microbial metabolites are key actors in host-microbiota cross-talk. The beneficial or detrimental effect of specific microbiota-derived metabolites depends on the context and the host state, suggesting the primordial nature of the symbiotic microbiota in ensuring optimal health in humans.

With the widespread westernisation of lifestyles, alteration of the gut microbiota composition and functions has become a worldwide phenomenon. Despite the difficulty to distinct a direct causal relationship and an association between dysbiosis and diseases, several lines of evidence demonstrate that the alteration of the gut microbiota is involved in the pathogenesis of multiple diseases affecting the GI tract, such as IBD or colorectal cancer, as well as many non-digestive systems. Metabolic disorders have been recognised to be massively impacted by gut microbiota. In the last two decades, increasing calorie intake and decreasing levels of physical activity have contributed to a progression in the prevalence of metabolic disorders. Metabolic disorders represent a group of disorders with the clustering of various inter-related pathological conditions combining obesity, non-alcoholic steatohepatitis (NASH), dyslipidaemia, glucose intolerance, insulin resistance, hypertension and diabetes that, when occurring together, strongly increase the incidence of cardiovascular diseases and mortality. Deciphering the mechanisms of host-intestinal microbiota interactions represents a major public health challenge in the development of new preventive or curative therapeutic strategies. In the present review, we will focus on the results from the most significant studies dealing with the role of microbiota-derived metabolites in metabolic disorders.

DISRUPTED EQUILIBRIUM OF THE GUT MICROBIOME-HOST INTERACTIONS IN METABOLIC DISORDERS

The gut microbiota plays a crucial role in maintaining the physiological functions of the host. A disruption of the fragile host-microbiota interaction equilibrium can play a role in the onset of several metabolic diseases. The gut microbiota can interact with the host by producing metabolites, which are small molecules (<1500 Da) representing intermediates or end-products of microbial metabolism. These metabolites can derive directly from bacteria or the transformation of dietary or host-derived substrates.

GUT MICROBIOTA INCRIMINATION

The implication of the gut microbiota in the regulation of host metabolic balance has been demonstrated in the last decade. Studies conducted both in animal models and humans revealed a significant role of the gut microbiota in the pathogenesis of metabolic disorders, strongly influenced by diet and lifestyle modifications.

EVIDENCE FROM ANIMAL EXPERIMENTS

The gut microbiota modulates energy expenditure and homeostasis in several animal models, including germ-free mice (GF mice) and genetically induced mice with obesity (ob/ob mice). GF mice are protected against obesity in a Western diet setting. Independent of daily food intake, Bäckhed *et al* reported a 60% increase in body fat, hepatic triglycerides and insulin resistance in conventionalised adult GF mice compared with GF mice, notably due to better absorption of monosaccharides. Interestingly, the transfer of gut microbiota from ob/ob mice to GF mice results in a significant increase in body weight and fat mass compared with colonisation with a lean microbiota, showing a causal relationship. The gut microbiota composition is unique to each individual. Caecal microbiota transplantation, from two mice with different responses to high-fat diet (HFD), into GF mice leads to the transmission of the donor's responder (RR) or non-responder (NR) phenotype. The gut microbiota of severely hyperglycaemic RR mice is enriched in Firmicutes, whereas NR is dominated by Bacteroidetes and Actinobacteria. Moreover, the transplantation of faecal microbiota from human twin pairs, discordant for obesity, into GF mice led to the acquisition of lean and obese phenotypes according to the donor. This phenotype transmission is strongly diet-dependent and notably favoured by a low-fat diet enriched in vegetables and fruits and thus enriched in fibre. The effect of the gut microbiota seems to occur even before birth, as the maternal gut microbiota, through short-chain fatty acid (SCFAs), triggers embryonic GPR41 and GPR43 and influences prenatal development of neural, enteroendocrine and pancreatic systems of the offspring to maintain postnatal energy homeostasis and eventually prevent metabolic disorder development.

Overall, these animal studies demonstrate the tight interconnection between diet and the gut microbiome in the pathogenesis of metabolic disorders as well as in its vertical transmissibility.

EVIDENCE FROM HUMAN STUDIES

Alterations in the gut microbiome composition and functions are associated with various traits observed in metabolic disorders. Although there are some conflicting results, the obesity-associated gut microbiota has been characterised by a decline in Bacteroidetes and a compensatory expansion of the Firmicutes phylum and by a reduction in microbial diversity and richness. There is notably a negative correlation between the severity of metabolic markers and the richness of the gut microbiota. Individuals with low microbiota gene content present more adiposity, insulin resistance and dyslipidaemia than high bacterial richness populations. Even in severe obesity conditions, those with diminished gut microbiota richness have a more severe metabolic condition.

In patients with diabetes, the higher proximity of the altered microbiota to epithelial cells could promote pro-inflammatory signals, contributing to the development of aggravated metabolic alterations. In humans, faecal microbiota transplantation (FMT) demonstrated some positive but moderate effects in patients with metabolic syndrome traits, proving the involvement of the gut microbiota in the pathogenesis and its potential therapeutic role.However, the efficiency of FMT in improving metabolic amelioration was dependent on the recipient gut microbiota profile, with low baseline richness promoting gut microbiota engraftment.

GUT MICROBIOTA-DERIVED METABOLITE IMPLICATIONS IN METABOLIC DISEASES

THE GUT METABOLOME

Metabolomics, which consists of the study of the small molecules present in any type of biological sample, has proven to be helpful in enriching the knowledge on microbiota-host interactions. Several hundred faecal or serum metabolites have been associated with clinical features associated with metabolic disorders. Moreover, a combination of metagenomics and metabolomics was used to elucidate the associations between gut microbiota imbalances and metabolic disturbances. This field is still in its infancy and, for some metabolites, it remains difficult to determine whether they are fully microbiota-derived or if other sources are involved, including diet or the host itself.

Metagenome and metabolome studies led to the discovery of new associations between microbial-derived metabolites and metabolic syndrome, but additional arguments are needed to establish a potential causality link. Notably, the decreased abundance of *Bacteroides thetaiotaomicron*, a glutamate fermenting commensal, in subjects with obesity is inversely correlated with serum glutamate. Furthermore, positive correlations between insulin resistance and microbial functions are driven mainly by a few species, such as *Prevotella copri* and *Bacteroides vulgatus*, suggesting that they may directly impact host metabolism. Metabolomics studies in plasma, saliva or urine identified different biochemical classes of metabolites that may be altered in metabolic disorders in association with gut microbiota perturbations. Dysregulation of lipolysis, fatty acid oxidation and aminogenesis and ketogenesis, as well as changes in the levels of triglycerides, phospholipids and trimethylamine N-oxide (TMAO) are described in samples from humans with metabolic disorders, and more recently, imidazole propionate (IMP) was discovered as being involved in insulin resistance. Shotgun metagenomics data suggest that hepatic steatosis and metabolic alterations are associated with dysregulated aromatic and branched-chain amino acid (BCAA) metabolism. The dysregulation of SCFA and bile acid (BA) metabolism are also associated with metabolic diseases, including obesity, type 2 diabetes mellitus and non-alcoholic fatty liver diseases.

BILE ACIDS

BAs are small molecules synthesised in hepatocytes from cholesterol. The primary BAs chenodeoxycholic acid (CDCA) and cholic acid (CA), conjugated to glycine or taurine, are essential for lipid/vitamin digestion and absorption. Ninety-five per cent of them are reabsorbed actively from the terminal ileum and are recycled in the liver (enterohepatic circulation). Primary BAs are also transformed into secondary BAs and deconjugated by gut microbiota. They can be either passively reabsorbed to reenter the circulating BA pool or excreted in the faeces as shown in Figure 1 below.

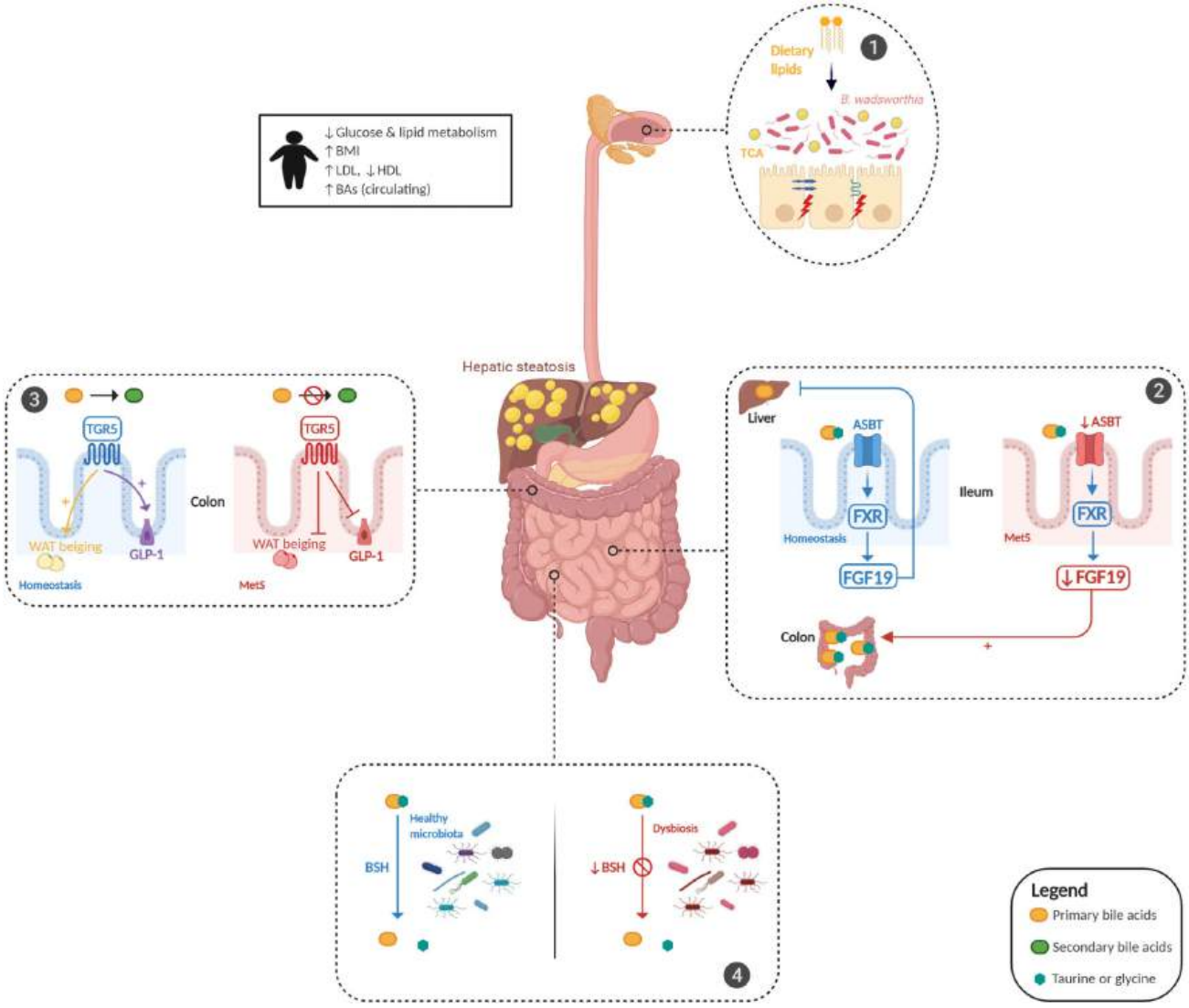


Figure 1

Bile acid (BA) dysmetabolism in metabolic syndrome. BA metabolism is altered in patients with metabolic syndrome (MetS) and is associated with hepatic steatosis and glucose and lipid dysmetabolism. Dietary animal fat consumption promotes taurocholic acid (TCA) production, which favours the proliferation of sulfite-reducing bacteria, *Bilophila wadsworthia*, leading to an increase in intestinal permeability and inflammation (panel 1). Gut microbiota alterations induce an impairment in the ileal absorption of BAs, which occurs normally via the apical-sodium BA transporter (ASBT). This induces a decrease in the expression of nuclear Farnesoid-X receptor (FXR) and fibroblast growth factor 19 (FGF19) in intestinal epithelial cells and the abundance of colonic primary conjugated BAs (panel 2). Gut microbiota dysfunction leads to a decreased transformation of primary conjugated BAs to secondary BAs in the colon, leading to defective activation of Takeda-G-protein-receptor-5 (TGR5). The effect of TGR5 activation on the increase in glucagon-like peptide 1 (GLP-1) and white adipose tissue (WAT) browning was thus inhibited (panel 3). Gut microbiota alterations impair bile salt hydrolase (BSH) activity, leading to primary conjugated BA accumulation in the colon (panel 4). BMI, body mass index; HDL, high-density lipoprotein; LDL, low-density lipoprotein.

Increased total circulating BA levels in individuals with obesity positively correlate with body mass index and serum triglycerides in patients with hyperlipidaemia. BAs regulate their synthesis through FGF19/FGF15, but they also have metabolic effects through their

receptors Farnesoid-X receptor (FXR) and Takeda-G-protein-receptor-5 (TGR5). Activation of FXR and TGR5 (1) promotes glycogen synthesis and insulin sensitivity in the liver; (2) increases insulin secretion by the pancreas; (3) facilitates energy expenditure, especially in the liver, brown adipose tissue and muscles (browning); (4) favours thermogenesis, resulting in a decrease in body weight and (5) mediates satiety in the brain. BAs also impact lipid metabolism, especially by exerting profound effects on triacylglycerol. The perturbations of the intestinal microbiota composition in metabolic disorders strongly impact BA metabolism, especially characterised by a failure to metabolise primary BAs, thus leading to their accumulation. Indeed, an increase in primary CDCA levels induces a decrease in very low-density lipoprotein production and plasma triglyceride concentrations. Short-term antibiotic supplementation in mice induces a decrease in secondary BA-producing bacteria and a reduction in hepatic deoxycholic acid (DCA) and lithocholic acid concentrations as well as serum triglyceride levels, suggesting that secondary BAs can act as regulators to maintain metabolic host homeostasis. Moreover, this alteration in the primary to secondary BA pool in metabolic disorders might play a role in the observed low-grade intestinal inflammation, as conjugated primary BAs exhibit pro-inflammatory effects on intestinal epithelial cells. Conversely, secondary BAs have anti-inflammatory properties. In addition, Parséus *et al* showed that the promoting effect of the gut microbiome on obesity and hepatic steatosis is dependent on the FXR pathway.However, the FXR-dependent role of secondary BAs in the regulation of glucose and lipid metabolism is debated and might be context-dependent. The accumulation of hepatic lipids, triglycerides and cholesterol has been observed in FXR-deficient mice on a normal chow diet, while in HFD-fed mice or an obese background, FXR deficiency improves glucose homeostasis and decreases body weight, possibly a consequence of different basal gut microbiota. The effects of FXR in the pathogenesis of metabolic disorders are also likely to be different from one tissue to the other, as demonstrated by studies in conditional knockout mice. FXR induces the transcription of fibroblast growth factor 19 (FGF19) in intestinal epithelial cells, which reach the liver and inhibit BA synthesis in a feedback loop. Mice overexpressing FGF19 exhibit increased metabolic activity and energy expenditure by increasing brown adipose tissue and decreasing liver expression of acetyl coenzyme A carboxylase 2, thus leading to protection against HFD-induced metabolic injury. Gut microbiota perturbations induce impairment in the ileal absorption of BAs, which normally occurs via the apical-sodium bile acid transporter, resulting in decreased expression of FXR and FGF19 and an imbalance of BAs, notably characterised by an increase in colonic primary conjugated BAs. Transgenic mice overexpressing TGR5 exhibit improved glucose tolerance with increased secretion of glucagon-like peptide 1 (GLP-1) and insulin. This BA-TGR5 axis elicits beige remodelling in subcutaneous white adipose tissue and may contribute to improvement in whole-body energy homeostasis. The alteration of gut microbiota-dependent BA metabolism, through qualitative (primary vs secondary and conjugated vs deconjugated BAs) or quantitative modification of the BA pool, is likely to participate in the pathogenesis of metabolic disorders. Moreover, BAs have an important impact on intestinal epithelium function. Primary BAs, such as CA and CDCA, and some secondary deconjugated BAs, such as DCA, increase epithelial permeability through the phosphorylation of occludin in intestinal Caco-2 cells. Some correlations have been observed between BA levels and intestinal permeability in mouse models. The effect of the BA-microbiota dialogue is massively impacted by diet. High consumption of animal fat promotes taurocholic acid production, leading to a shift in microbiota composition with a bloom of sulfite-reducing microorganisms such as *Bilophila wadsworthia* and to increased susceptibility to colitis in IL-$10^{-/-}$ mice and more severe liver steatosis, barrier dysfunction and glucose metabolism alteration in HFD-fed mice. Moreover, bile salt hydrolase (BSH) activity, which is responsible for BA deconjugation in the normal gut microbiota, is impaired in metabolic disorders and likely plays a role in the accumulation of primary conjugated BAs in the colon of these patients. In mouse models, correcting BSH defects by the administration of BSH-overexpressing *Escherichia coli* improved lipid metabolism, homeostasis and circadian rhythm in the liver and GI tract, resulting in protection against metabolic disorders.

SHORT-CHAIN FATTY ACIDS

SCFAs, such as butyrate, propionate and acetate, are end-products of microbial fermentation implicated in a multitude of physiological functions. SCFAs participate in the maintenance of intestinal mucosa integrity, improve glucose and lipid metabolism, control energy expenditure and regulate the immune system and inflammatory responses as shown in Figure 2 below. They act through different mechanisms, including specific G protein-coupled receptor family (GPCR) and epigenetic effects.

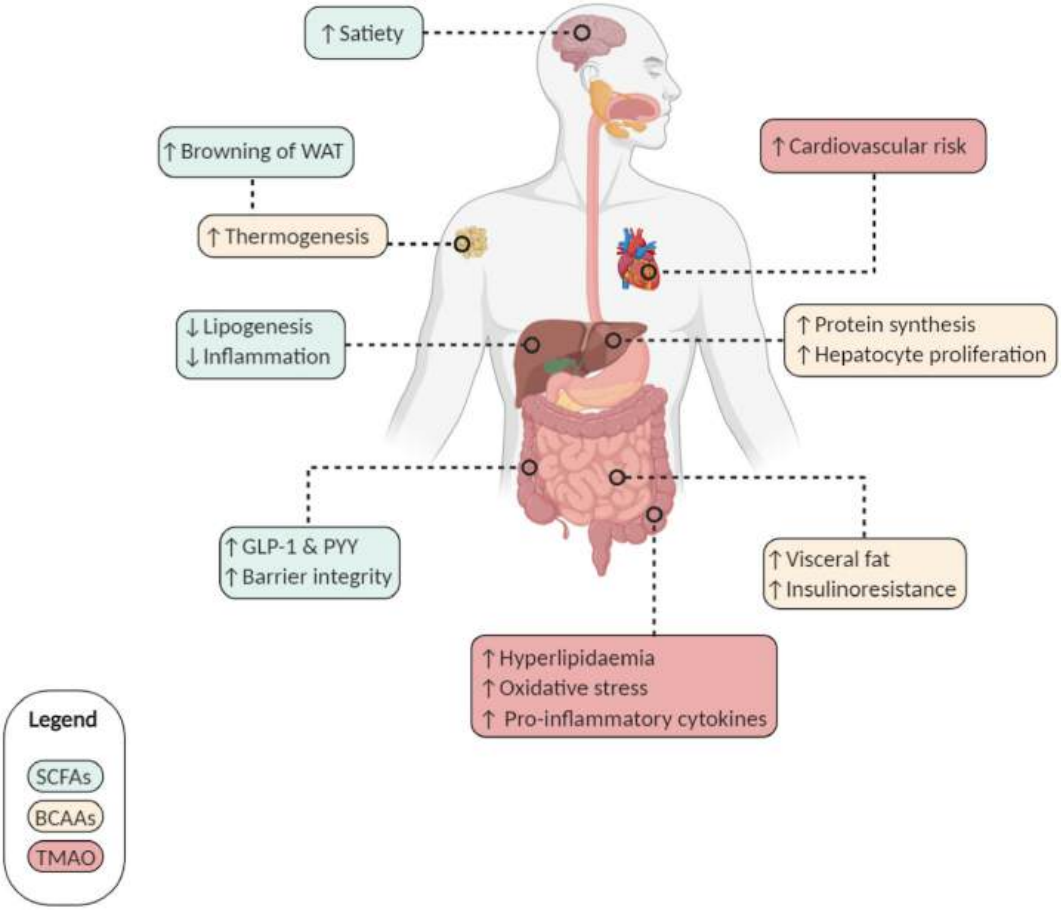


Figure 2

Short-chain fatty acids (SCFAs), branched-chain amino acids (BCAAs) and Trimethylamine N-oxide (TMAO): relevant effects for metabolic syndrome on the host. Microbiota-derived metabolites mediate diverse effects on host metabolism. SCFAs (green frame): (i) increase satiety and browning of white adipose tissue (WAT); (ii) induce a decrease in lipogenesis and associated inflammation; (iii) increase the secretion of glucagon-like peptide 1 (GLP-1) and peptide YY (PYY) and (iv) participate in the maintenance of intestinal barrier integrity. BCAAs (yellow frame): (i) increase thermogenesis, protein synthesis and hepatocyte proliferation but (ii) are also associated with insulin resistance and visceral fat accumulation. TMAO (red frame): increases cardiovascular risks by inducing hyperlipidaemia, oxidative stress and pro-inflammatory cytokines.

The amount of SCFA-producing bacteria and SCFAs is reduced in faecal samples of dysmetabolic mice and in humans with obesity and diabetes. In rodents with diabetes and obesity, supplementation with SCFAs improves the metabolic phenotype by increasing energy expenditure, glucose tolerance and homeostasis. Adding back fermentable fibres (inulin) to an HFD seems to be enough to protect against metabolic alterations. In humans, SCFA administration (inulin-propionate ester, acetate or propionate) stimulates the production of GLP-1 and PYY, leading to a reduction in weight gain. The protective effects of SCFAs on metabolic alterations might occur as early as in utero. In mice, high-fibre diet-induced propionate from the maternal microbiota crosses the placenta and confers resistance to obesity in offspring through the SCFA-GPCR axis.

BRANCHED-CHAIN AMINO ACIDS

The most abundant BCAAs, valine, isoleucine and leucine, are essential amino acids synthesised by plants, fungi and bacteria, particularly by members of the gut microbiota. They play a critical role in maintaining homeostasis in mammals by regulating protein synthesis, glucose and lipid metabolism, insulin resistance, hepatocyte proliferation and immunity. BCAA catabolism is essential in brown adipose tissue (BAT) to control thermogenesis. It occurs in mitochondria via SLC25A44 transporters and contributes to an improvement in metabolic status. Moreover, supplementation of mice with a mixture of BCAAs promotes a healthy microbiota with an increase in *Akkermansia* and *Bifidobacterium* and a decrease in Enterobacteriaceae. However, the potential positive effects of BCAAs are controversial. Elevated systemic BCAA levels are associated with obesity and diabetes, probably a consequence of the 20% increased consumption of calories over the last 50 years. In genetically obese mice (ob/ob mice), BCAA accumulation induces insulin resistance. The gut microbiota is a modulator of BCAA levels, as it can both produce and use BCAAs. *Prevotella copri* and *B. vulgatus* are potent producers of BCAAs, and their

amounts correlate positively with BCAA levels and insulin resistance. In parallel, a reduced abundance of bacteria able to take up BCAAs, such as *Butyrivibrio crossotus* and *Eubacterium siraeum*, occurs in patients with insulin resistance. Further studies are needed to more precisely elucidate the effects of BCAAs in the pathogenesis of metabolic disorders.

TRIMETHYLAMINE N-OXIDE

The gut microbiota can metabolise choline and L-carnitine from dietary sources (eg, red meat, eggs and fish) to produce trimethylamine (TMA). This gut microbiota-derived TMA is then absorbed and reaches the liver where it is converted into TMAO through the enzymatic activity of hepatic flavin monooxygenases 3.

In humans, the level of TMAO increases in patients with diabetes or at risk of diabetes and in obesity. Increasing evidence demonstrates that the gut microbiota-dependent metabolite TMAO is also associated with a higher risk of developing cardiovascular disease and kidney failure. In mice, dietary supplementation with TMAO, carnitine or choline alters the caecal microbial composition, leading to TMA/TMAO production that increases the atherosclerosis risk. This effect is dependent on the gut microbiota, as it is lost in antibiotic-treated mice. Moreover, transferring the gut microbiota of high-TMAO mice recapitulates atherosclerosis susceptibility in recipient low-TMAO mice. Importantly, the role of the gut microbiota in the production of TMAO from TMA has also been demonstrated in humans. Overall, in metabolic disorders, the altered microbiota associated with an increased intake of choline and L-carnitine from dietary sources leads to an increase in plasma levels of TMAO, which is directly involved in the pathogenesis of metabolic disease comorbidities and particularly cardiovascular disorders. However, detailed investigations are needed in populations from different countries to understand the interaction between food consumption patterns, TMAO production and cardiovascular risks.

TRYPTOPHAN AND INDOLE-DERIVATIVE METABOLITES

Tryptophan is an essential aromatic amino acid acquired through common diet sources, including oats, poultry, fish, milk and cheese. In addition to its role in protein synthesis, tryptophan is a precursor for crucial metabolites. Dietary tryptophan can follow two main pathways in host cells, namely, the kynurenine and serotonin routes. The third pathway implicates gut microorganisms in the direct metabolism of tryptophan into several molecules, such as indole and its derivatives, with some of them acting as aryl hydrocarbon receptor (AhR) ligands, as shown in Figure 3 below.

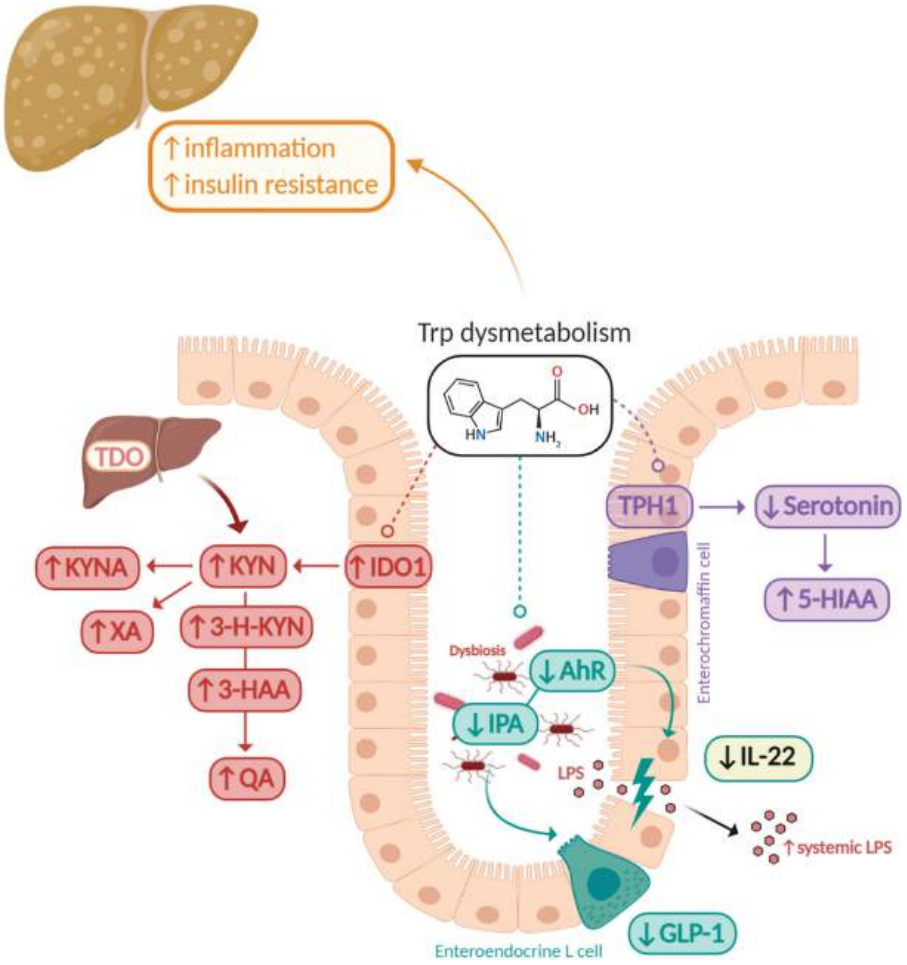


Figure 3

Tryptophan metabolism alterations in metabolic syndrome. Tryptophan dysmetabolism is associated with liver inflammation, steatosis and insulin resistance. In metabolic syndrome (MetS), the inflammatory state is associated with kynurenine (KYN) production through the activation of indoleamine 2,3-dioxygenase 1 (IDO1). This leads to an increase in kynurenine-derived metabolites, such as kynurenic acid (KYNA), xanthurenic acid (XA), 3-hydroxykynurenine (3-H-KYN), 3-hydroxyanthranilic acid (3-HAA) and quinolinic acid (QA). In parallel, the gut microbiota presents a defect in the production of aryl hydrocarbon receptor (AhR) ligands such as indole-3-propionic acid (IPA). The incretin hormone glucagon-like peptide 1 (GLP-1) secretion from intestinal enteroendocrine L cells and interleukin (IL)-22 production are decreased, altering gut permeability and promoting lipopolysaccharide (LPS) translocation. Serotonin (5-HT) biosynthesis from intestinal enterochromaffin cells is also reduced in the context of MetS due to a decrease in the production of microbiota-derived metabolites inducing the production of host 5-HT.

We have identified in a previous study, in both preclinical and clinical settings, that metabolic disorders are characterised by a reduced capacity of the microbiota to metabolise tryptophan into AhR agonists. Defective activation of the AhR pathway leads to decreased production of GLP-1 and IL-22, which contribute to intestinal permeability and lipopolysaccharide (LPS) translocation, resulting in inflammation, insulin resistance and liver steatosis. In this context, treatment with AhR agonists or administration of *Lactobacillus reuteri*, which naturally produces AhR ligands, can reverse metabolic dysfunction. Similarly, indole prevents LPS-induced alterations of cholesterol metabolism and alleviates liver inflammation in mice. Moreover, exploring human jejunum samples from patients with severe obesity led to the observation that a low AhR tone correlated with a high inflammatory score. Interestingly, the use of the AhR ligand is able to prevent damage to barrier integrity and inflammation in Caco-2/TC7 cells.

We and others also showed strong activation of the kynurenine pathway in metabolic diseases. Genetic or pharmacological approaches inhibiting the activity of indoleamine 2,3-dioxygenase (IDO), the rate-limiting enzyme of the kynurenine pathway, are protective against HFD-induced obesity and metabolic alterations. The mechanism is likely to be mediated by the microbiota and AhR. The increased amount of available tryptophan, due to the inactivation of IDO, can be converted by the microbiota in AhR agonists. Conversely, in obesity, the overactivation of IDO, associated with an increase in plasma levels of downstream metabolites such as kynurenic acid, xanthurenic acid, 3-hydroxykynurenine, 3-hydroxyanthranilic acid and quinolinic acid, decreases the tryptophan pool, which is less available for the production of AhR agonists by the microbiota. The third pathway of tryptophan metabolism, serotonin (5-HT), is also involved, as it affects feeding behaviour and satiety and is thus important for obesity development. The gut microbiota, and primarily indigenous spore-forming bacteria, represent an essential modulator of the intestinal production of 5-HT in enterochromaffin cells that represents >80% of the whole body 5-HT synthesis. These effects are notably mediated by SCFAs and BAs. Mice deficient for the production of peripheral serotonin are protected from HFD-induced obesity. Mechanistically, 5-HT inhibits brown adipose tissue thermogenesis, thus leading to fat accumulation. Human data support these results, as elevated plasma levels of 5-hydroxyindole-3-acetic acid, an end-product of serotonin metabolism, are increased in patients with metabolic disorders.

IMIDAZOLE PROPIONATE

Exploring the interaction between food intake, gut microbiota and derived metabolites might be of interest to discover metabolites impacting metabolic health. As such, it was recently shown that IMP, a metabolite produced by histidine utilisation of gut microbiota, was enhanced in type 2 diabetes and associated with insulin resistance. In the liver, IMP appeared to affect the insulin signalling pathway via mammalian target of rapamycin complex 1 (mTORC1). The examination of IMP in large human cohorts also links it with metabolic health and lifestyle. IMP was elevated in subjects with prediabetes and diabetes in the MetaCardis cohort and in subjects with low bacterial gene richness and Bacteroides 2 enterotype in this cohort. Associations between IMP levels and markers of low-grade inflammation were also identified. Importantly, relationships were observed between serum IMP levels and unhealthy diet measured by dietary quality scores emphasising the importance of nutrition in this context. Thus, this study confirms that in type 2 diabetes, the intestinal microbiota may is switched towards IMP production, which can impact host inflammation and metabolism.

THERAPEUTIC RELEVANCE

The mechanistic links between gut microbiota-derived metabolites and metabolic disorders make these interactions a promising therapeutic target in these complex diseases.

LESSONS FROM FAECAL MICROBIOTA TRANSPLANTATION

FMT is a drastic strategy to modify the gut microbiome. It is highly effective in the treatment of recurring *Clostridioides difficile* infections and has been evaluated in small trials in metabolic syndrome and obesity. The clinical efficacy of this strategy is so far mild, with mostly some positive effects on insulin sensitivity in subgroups of patients. However, these studies had several limitations, including small size and limited duration of intervention. Nevertheless, they provide relevant information to identify the critical molecules involved in biological effects. Following successful FMT, both the microbiota composition and metabolomics, such as BA and SCFA profiles, can be restored. In patients with obesity, FMT can induce engraftment of the butyrate-producing and bile-hydrolysing genus *Faecalibacterium*, leading to a restoration of the BA profile and microbiota BSH activity. FMT increases the relative abundance of SCFA-producing bacteria such as *Roseburia intestinalis* and the protective strain *Akkermansia muciniphila*, with a possible role in the improvement in insulin sensitivity through regulation of GLP-1. *A. muciniphila* supplementation alone improves metabolic parameters in overweight/obese insulin-resistant

volunteers characterised by better insulin sensitivity and a reduction in plasma total cholesterol levels and fat mass. In mice, *A. muciniphila* promotes the production of SCFAs and the restoration of HFD-induced alterations in tryptophan metabolism. These data highlight the key family of microbiota-derived metabolites with potential therapeutic effects.

SYNTHETIC AGONISTS OF BILE ACID RECEPTORS

Given their potential benefits in metabolic diseases, BAs and synthetic FXR and TGR5 agonists are currently under development in the metabolic field. Preclinical trials based on in vitro and in vivo studies identified potent synthetic FXR and TGR5 agonists, which are currently being investigated in phase II or III clinical trials. Due to the regulatory roles of FXR and TGR5 receptors on glucose and lipid metabolism, multiple specific agonists have been designed. Obeticholic acid (OCA), one of the best-characterised FXR agonists, protects the liver from damage in mice with a reduction in hepatic steatosis and inflammation and is currently being evaluated in a phase III trial in patients with NASH. The synthetic FXR agonist GW4064 improves hyperglycaemic and hyperlipidaemia in mice with diabetes and is able to correct BA dysmetabolism and alleviate liver toxicity in rodents with short bowel. The intestine-restricted FXR agonist fexaramine can also promote adipose tissue browning and GLP-1 secretion in wild type (WT) and leptin receptor-deficient diabetic mice. Finally, a TGR5 agonist ameliorated insulin resistance and glucose homeostasis in mice with diabetes by the cyclic AMP/protein kinase A pathway in skeletal muscles.

SHORT-CHAIN FATTY ACID AND BRANCHED-CHAIN AMINO ACID TREATMENT

Dietary supplementation with fermentable fibres, such as inulin in HFD-fed mice or inulin-propionate ester in overweight humans, protects against metabolic disturbances by restoring the gut microbial composition and the action of the IL-22-mediated axis.Oral SCFA treatment in obese mice can modulate lipid synthesis and insulin receptors by upregulating peroxisome proliferator-activated receptor-γ. It also improves intestinal barrier functions with a lower serum LPS concentration. SCFAs exert their beneficial effects partly through specific G-protein-coupled receptors, and their activation by specific agonists is an attractive strategy in the treatment of MetS. GPR40/FFA1, GPR41/FFA3, GPR43/FFA2 and GPR120/FFA4 agonists induce protection against diet-induced obesity in mice through the improvement in insulin, GLP-1 and incretin secretion and anti-inflammatory effects. In addition, a link between dietary BCAAs and energy balance was noted in animals with obesity, and reducing the proportion of dietary BCAAs was associated with a restoration of metabolic health.

CONCLUDING REMARKS

Gut microbiota-derived metabolites have a central role in the physiology and physiopathology of metabolic disorders. The microbial metabolites described above, specifically BAs, SCFAs, BCAAs, TMAO, tryptophan and indole derivatives, are implicated in the pathogenesis of these complex disorders and represent potential biomarkers for the early diagnosis and prognosis of these diseases. Moreover, microbiota-derived metabolites and their host receptors, possibly in combination with dietary intervention, represent promising targets for the development of novel therapeutic tools for metabolic disorders.

ACKNOWLEDGMENTS

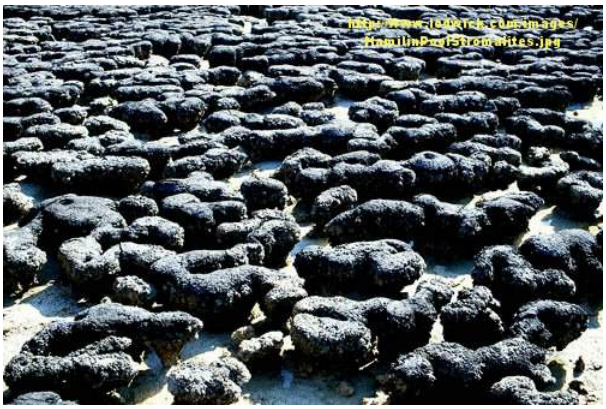
The authors would like to thank BioRender, for its revolutionised tool to create custom scientific figures (https://biorender.com/)

30: Abiotic Origins of Life

30.0.1: The Start of Life

This book began with the notion that understandings derived from the study of simple molecules can be applied to complex biological macromolecules and systems. We developed an understanding of the structural, thermodynamic, and kinetic properties of the "simplest" biomolecules, including single chain amphiphiles like fatty acids, and double chain ones like phospholipids, and how these properties could explain the propensity of these molecules to form complex lipid aggregates (micelles and bilayers). We extended these ideas to the process of protein folding and the assembly of biological complexity. Is there something intrinsic to the property of molecules such that their localization together in the right microenvironment could lead to a "living cell"? How did life originate? That is the topic of this last capstone chapter.

Defining life is actually quite difficult. Here is a list of requirements that seem reasonable, but other have noted that this list would exclude the mule. Life can self-replicate, self-sustain, evolve, respond to environmental changes, and die. The earliest known fossils (stromatolites from cyanobacteria) are approximately 3.5 billion years old.



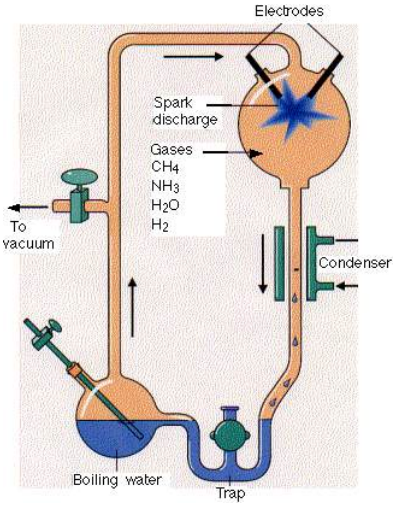
We have just finished studying the complex interactions involved in cell signaling. How could they have evolved? Consider the central dogma of biology. In the present biological world, proteins (DNA polymerase, RNA polymerase, transcription factors) are necessary for DNA synthesis, replication and gene transcription. But you need DNA to encode the proteins. This "chicken vs egg" dilemma has been addressed when it was realized that RNA can both carry genetic information as well as enzymatic activity (even at the level of the ribosome used for protein synthesis).

https://clockwise.software/blog/solv...d-egg-problem/ for new picture

30.0.1: Abiotic Synthesis of Amino Acids and Peptides

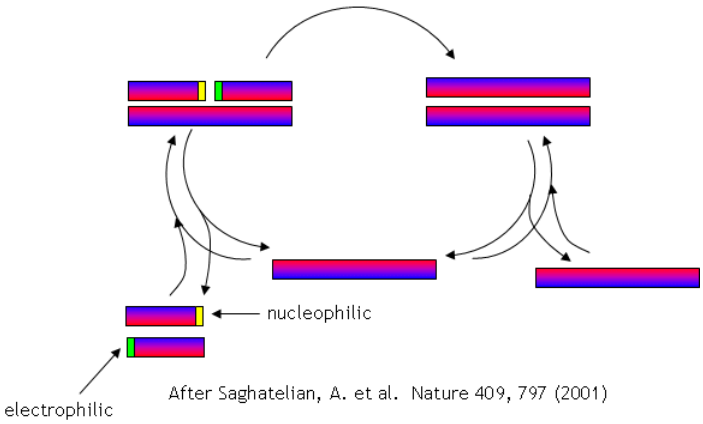
Much work has been done to determine if the building blocks for present biological molecules could have been synthesized early in Earth's history. Amino acids and fatty acids have been found in meteors suggesting the possibility. Earth's early atmosphere would have had little oxygen, so most components should have been reduced. It probably consisted of methane, ammonia, hydrogen and water similar to the atmospheres of other planets in our solar system. The composition of the early atmosphere is still contentious. In 1953 (the same year that Watson and Crick published the structure of double-stranded

DNA), Stanley Miller showed that electric discharges (to simulate lightening) in a reducing atmosphere over a "simulated sea" produced many amino acids. Up to 11 different amino acids have been produced in this fashion along with purines and pyrimidines (these required concentrated reaction mixtures) necessary for nucleic acids. Adenine can be produced just through the reaction of hydrogen cyanide and ammonia in an aqueous solution. Other nucleic acid bases can be made with hydrogen cyanide, cyanogen (C2N2) and cyanoacetylene (HC3N).



http://www.hencoup.com/Photo%20Stanley%20Miller.jpg

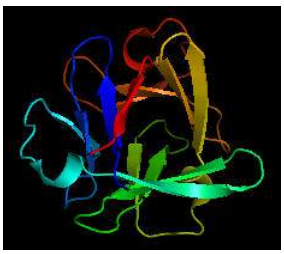
No complex polymers arise through these reactions. However, in 2004, Lehman, Orgel and Ghadiri were able to show that in the presence of carbon disulfide, a gas discharged from volcanoes, homo- and hetero-peptides were produced. Amphiphilic peptides can even catalyze their own formation from peptide fragments, if the fragments are activated. The fragments would bind to the larger "template" peptide through nonpolar actions of the side chains which are oriented along one face of the helical axes. If the fragments bind in a fashion in which the electrophilic end is adjacent to the nucleophilic end of the other peptide fragment, condensation of the two peptide fragments results. The larger template peptide acts as a template (effectively as an "enzyme") in orienting the two fragments for chemical reaction and effectively increasing their local concentration. The reaction of the bound fragments is essentially intramolecular. The reaction even proceeds with amplification of homochirality.



Could the prebiotic amino acids have polymerized into a protein that could fold in a fashion similar to modern proteins? That question has recently been addressed by Longo et al (2013). They asked the question whether the amino acids found in Miller-type prebiotic synthesis mixture and in comets/meteors (Ala, Asp, Glu, Gly, Ile, Leu, Pro, Ser, Thr and Val), a restricted set (10) compared to the present 20 naturally-occurring amino acid, could form a polymer that could fold. Notice that this reduced ensemble of amino acids lacks aromatic and basic amino acids. These proteins would be acidic with a low pI and may have trouble, given the lack of nonpolar aromatic amino acids, in forming a buried hydrophobic core which stabilize proteins. Nevertheless

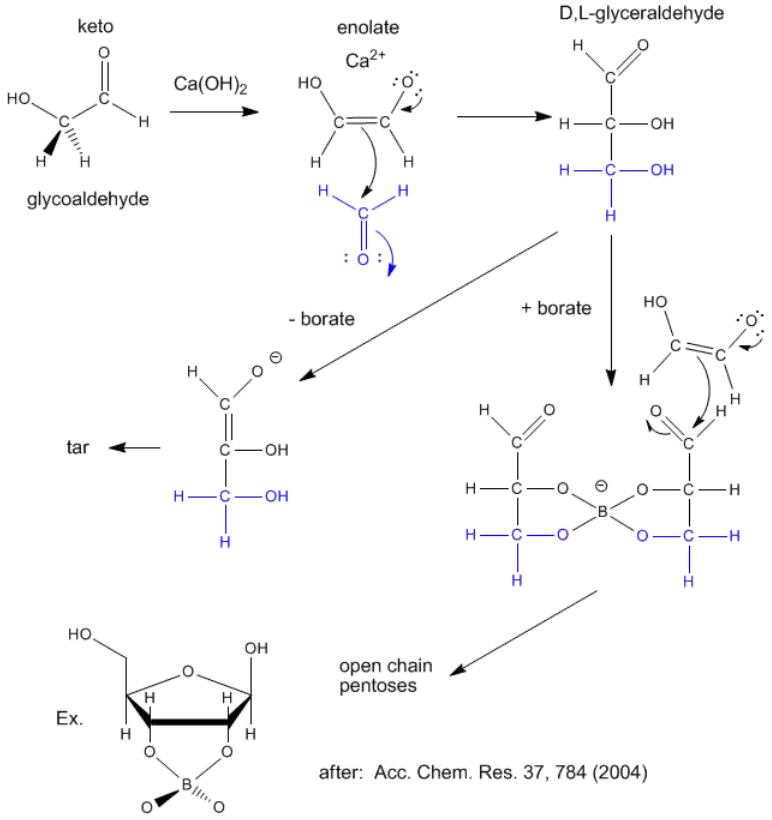
Longo et al were able to synthesize a protein with a slightly expanded set of amino acids (12, including Asn and Gln, with 70% prebiotic amino acids). The structure of one of the proteins, PV2, is shown below. The protein was more stable in 2 M NaCl (compared to 0.1 M) in which it showed a cooperative thermal denaturation with a melting point near 650C using differential scanning calorimetry. The protein had properties similar to those from halophilic organisms that thrive in high salt. These properties include low pIs and high negative charge density, which allows cation-protein interactions in the high salt environment, and lower stability in low salt environments. Earlier oceans were saltier. Halophiles are an example of extremeophiles which are highly represented in archea. Although most halophiles are aerobic, some are anerobic. Perhaps life arose in high salt environments.

Figure: Structure of PV2 protein comprised of a reduced alphabet of mainly prebiotic amino acids.



30.0.1: **Abiotic Synthesis of Sugars**

Sugars are required for present energy production but also as a part of the backbone (ribose, deoxyribose) of present genetic material. Many sugars can be synthesized in prebiotic conditions, using carbon based molecules with oxygen, such as glycoaldehyde and formaldehyde (both found in interstellar gases), as shown in the figure below. Glycoaldehyde was recently found in star forming regions of the Milky Way where planets are likely to form. The presence of borate, which stabilizes vicinal OHs on sugars, is required for the production of sugars instead of tars.



RNA molecules containing sugars such as threose, aldopentopyraonses and hexopyranose can also form stable secondary structures like helices. (Remember, RNA probably preceded DNA as the genetic carrier of information given that is also has enzymatic activity). Is there something special about ribose that made it selected over other sugars for nucleic acids, especially since it is

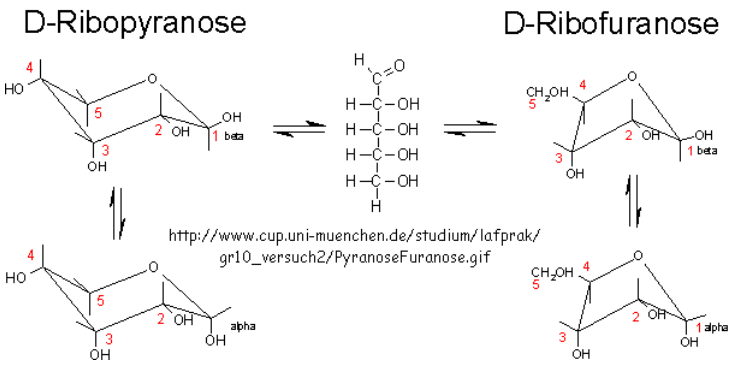
found in low abundancy in the products in synthesis reactions conducted under prebiotic conditions? One probable reason is it unusually high (compared to other sugars) permeability coefficient through vesicles made of phospholipids or single chain fatty acids, as shown below.

Table 1. Solute permeability coefficients (P_s) across fatty acid and phospholipid membranes

		Membrane composition					
	Solute	Myristoleate*	Palmitoleate*	Oleate*	Oleate†	POPC*	DPPC/C/PA‡
	Glycerol§			490			
4C	Erythritol§	28	14	12	4.8	21	
4C	DL-Threitol	50	23	22	11	40	
	Adonitol	2.9	1.7	0.91	0.38		
	Arabitol	2.1	1.4	0.63	0.39		
	Xylitol	5.0	3.0	1.8	0.95		
	Dulcitol				0.013		
	Mannitol	0.051	0.015	0.0058	0.0058		
	Sorbitol	0.18	0.071	0.026	0.027		
5C	Arabinose	3.5	2.1	1.1	0.49	2.8	1.1
5C	Lyxose	5.7	3.2	1.9	0.84	4.5	1.2
5C	Ribose	31	15	11	2.9	20	9
5C	D-Xylose	3.4	1.9	0.98	0.51	2.8	1.0
5C	L-Xylose	3.3	1.9	0.94		2.7	
5C	Ribulose			31			
6C	Galactose	0.060	0.024	0.011	0.0086		
6C	Glucose§	0.047	0.017	0.0071	0.0050		
6C	Mannose	0.10	0.035	0.018	0.014		
6C	Fructose§	1.1	0.68				
6C	L-Sorbose	0.88	0.56				

Values reported in 10^{-8} cm/s. Chiral solutes are D-enantiomers unless otherwise noted.
*Experiments performed at 23°C in 1 buffer, 0.5 M solute.
†Conditions as above but with 0.1 M solute.
‡Conditions as above, values are relative initial slopes only (D-xylose set to 1.0).
§Literature P_s values across lecithin membranes: glycerol, 540×10^{-8} cm/s (33); erythritol, 75×10^{-8} cm/s (34); glucose, 0.003×10^{-8} cm/s (35); and fructose, 0.04×10^{-8} cm/s (35).

In general, the greater the number of carbon atoms, the smaller the permeability. However, the table above clearly shows large differences in permeability for sugar isomers with the same number of C atoms, and the difference is not affected by the lipid composition. Ribose has markedly elevated permeability compared to other 5C sugars (as do erythritol and threitol among 4 C sugar alcohols). What is so unique about ribose? 20% of the sugar is in the furanose form. Rate constants for ring opening of furanoses are elevated, suggesting greater flexibility. The α-furanose anomer is amphiphilic in that one face is hydrophobic and the other hydrophilic. All of these may promote ribose permeability.

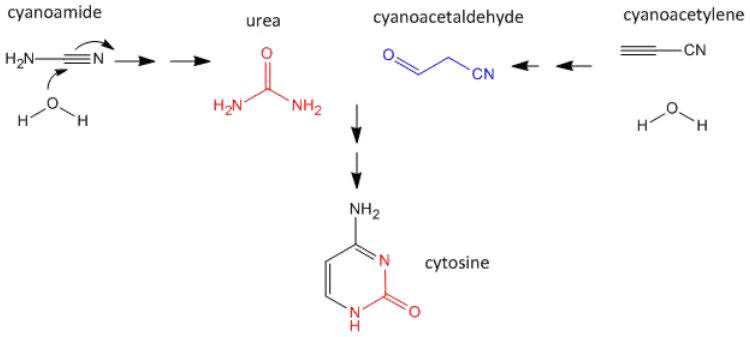


EXTERNAL Jmol: a-D-ribofuranose alpha-D-ribopyranose at

30.0.1: **Abiotic Synthesis of Nucleobases**

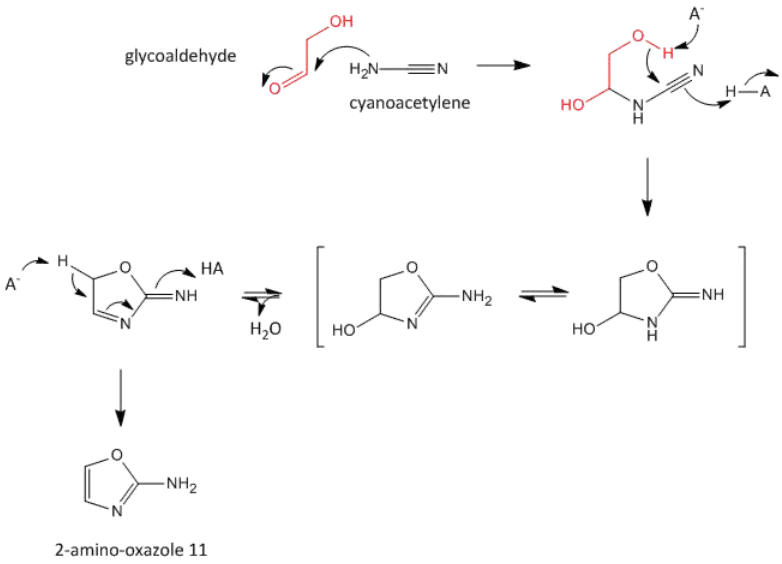
As mentioned above, nucleobases can be made under abiotic conditions with appropriate sources of carbon molecules with nitrogen. These include hydrogen cyanide, cyanoamide (NH2CN) (along with ammonia), hydrogen cyanide, cyanogen (C2N2) and cyanoacetylene (HC3N). Cyanoamide and cyanoacetylene can both react with water to form urea and cyanoacetaldehyde, respectively. The latter two can condense to form cytosine as shown below.

Figure: Abiotic Synthesis of Cytosine



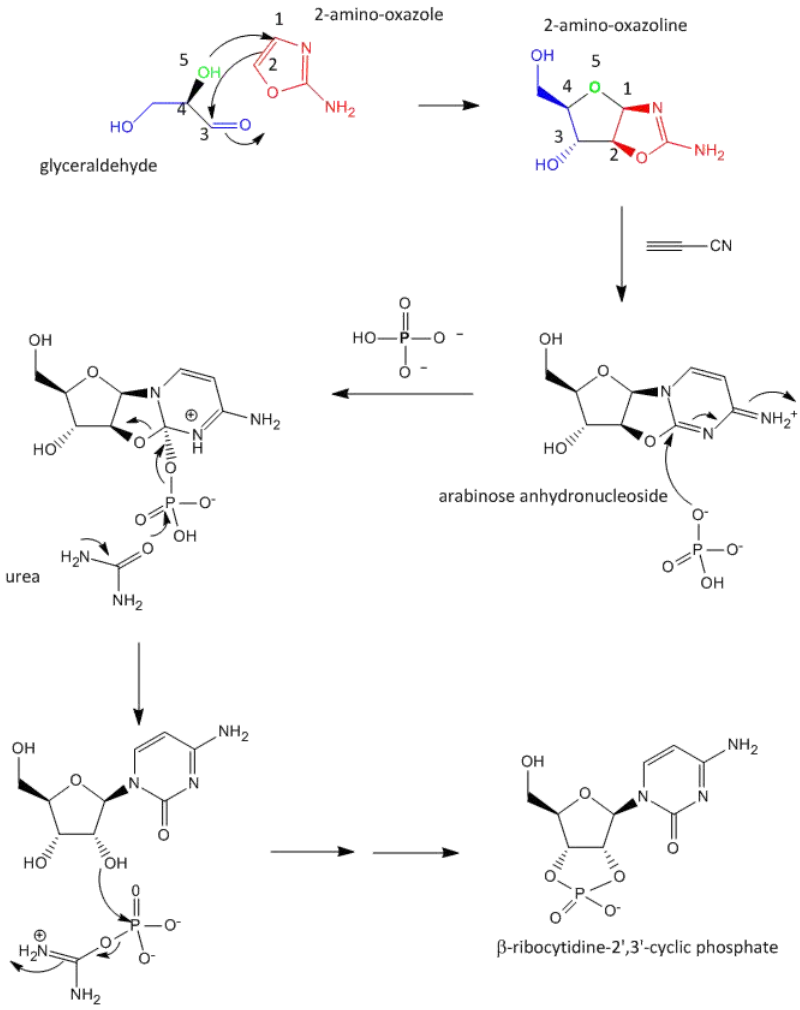
A major problem that has plagued prebiotic researchers is how the product of the carbon-oxygen compound (sugar) links to the product of the carbon-nitrogen compound (nucleobase) to form the nucleoside. Recent research by Powner et al has offered an innovative solution. Instead of forming the sugar and base in separate reaction, and then linking them covalently, the combined molecule could be synthesized in a single set of linked reactions. Inorganic phosphate serves as both a general acid/base catalyst (HA/A- in the figure below) in these new reactions in the formation of an important intermediate, 2-amino-oxazole, and as a nucleophilic catalyst.

Figure: Abiotic Synthesis of 2-amino-oxazole



Glyceraldehyde, 2-amino-oxazole, and inorganic phosphate can react to form a ribocytidine phosphate. Possible reaction mechanisms are shown below.

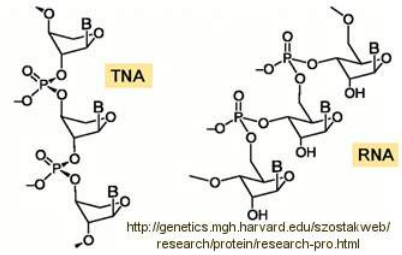
Figure: Abiotic synthesis of a ribocytidine phosphate without condensation of a preformed ribose and cytosine



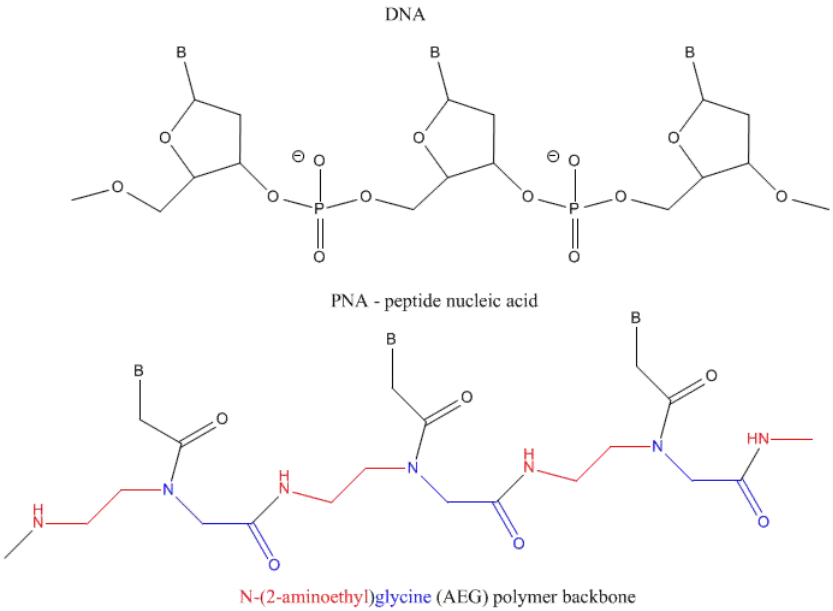
30.0.1: **Abiotic Synthesis of Genetic Polymers**

Abiological synthesis of polymer precursors is a long way from creating genetic polymers like RNA and DNA. These genetic polymers have one property that at first glance seems not conducive to a genetic molecule. Both are polyanions, which must be packed into a cell and folded onto itself to form the classic dsDNA helix and many different RNA structures. This problem is solved to some degree by the presence of counterions that help mask the charge on the negative backbone of the nucleic acids. The presence of phosphate in the phospodiester backbone linkage does confer an important advantage over other possible links (carboxylic acid esters, amides and anhydrides). The electrophilic phosphorous atom is hindered from nucleophilic attack by the negative O attached to the phosphorous. Also, the phosphorous is sp3 hybridized compared to the sp2 hybridization of the electrophilic carbon atom in anhydrides, esters, and amides, and hence is less accessible to nucleophilic attack. Most people now believe that RNA, which can act both as an enzyme and genetic template, preceded DNA as the genetic carrier. The evolution of DNA as the primary genetic carrier required an enzyme to convert ribose to deoxyribose. This would make the nucleic acid less likely to cleave at the phosphodiester bond with the replacement of a nucleophilic 2' OH with an H, and make the genetic molecule more stable. Other types of genetic carriers might have preceded the RNA world, especially if the monomer required could be more readily synthesized from abiological sources. One such alternative are threose nucleic acids (TNA). Synthetics ssTNA can base pair with either RNA, DNA, or itself to form duplexes.

EXTERNAL http://genetics.mgh.harvard.edu/szostakweb/research/protein/research-pro.html



Other possible candidate include peptide nucleic acids (PNA). These can also form double stranded structures with DNA, RNA, or PNA single strands. They were initially designed to bind to dsDNA in the major grove forming a triple-stranded structure. Binding could alter DNA activity, possibly by inhibiting transcription, for example. The structure of a single-stranded PNA is shown. Note that the backbone, a polymer of N-(2-aminoethyl)glycine (AEG) which can be made in prebiotic soups, is not charged, making it easier to bind to dsDNA. AEG polymerizes at 100oC to form the backbone.

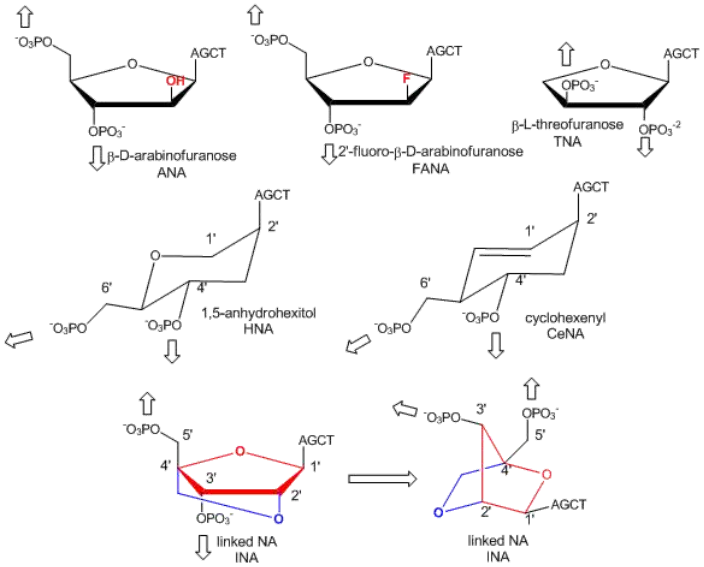


after Gamesh, K. & Nielsen, P. Peptide Nucleic Acids: Analogs and Derivatives. Current Organic Chemistry. 4, 931 (2000)

In addition to changing the backbone, additional bases other than A, C, T, G, and U can be accommodated into dsDNA and ssRNA molecules (Brenner, 2004)

In a recent extension, Pinheiro et al have shown that 6 different foreign backbone architectures can produce xeno-nucleic acids (XNAs) that can be replicated by engineered polymerases which make XNAs from a complementary DNA strand, and a polymerase that can make a complementary copy of DNA from an XNA. XNAs can also be evolved as aptamers to bind specific target molecules. The investigators replaced the deoxyribose and ribose backbone sugar with xenoanalogs (congeners) including 1,5-anhydrohexitol (HNAs), cyclohexene (CeNA), 2'-O,4'-C-methylene-b-D ribose (locked nucleic acids - LNA), L-arabinose (LNA), 2'-fluoro-L-arabinose (FANAs) and threose (TNAs) as shown in the figure below.

Figure: Xeno-nucleic acid sugar congeners

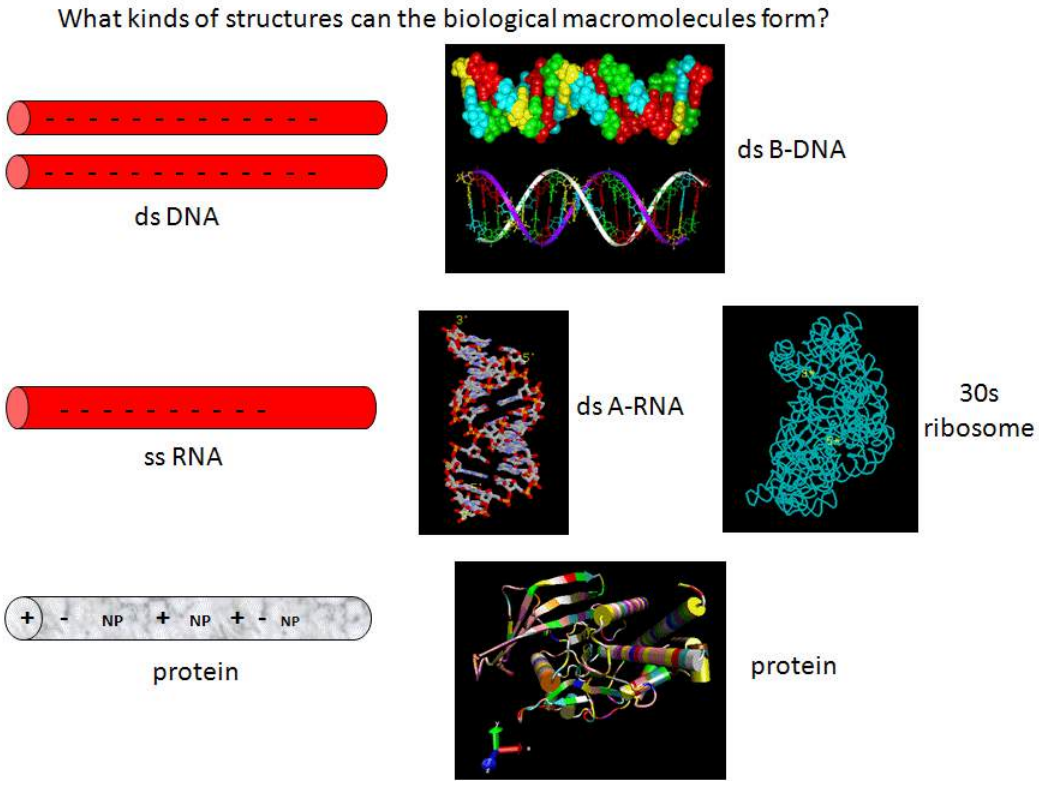


Polymers of these XNA can bind to complementary RNA and DNA and as such act as nuclease-resistant inhibitors of translation and transcription.

Von Kiedrowski, in an experiment similar to the self-replication of peptides described above, has shown that a single stranded 14 mer DNA strand, when immobilized on a surface, can serve as a template for the binding of complementary 7 mers and their conversion to 14 mers. When released by base, this process can occur with exponential growth of the complementary 14 mers. (von Kiedrowski Nature, 396, Nov 1998). Ferris has shown that if the clay montmorillonite is added to an aqueous solution of diadensosine pyrophosphate, polymerization occurs to produce 10 mers which are 85% linked in a 5' to 3' direction.

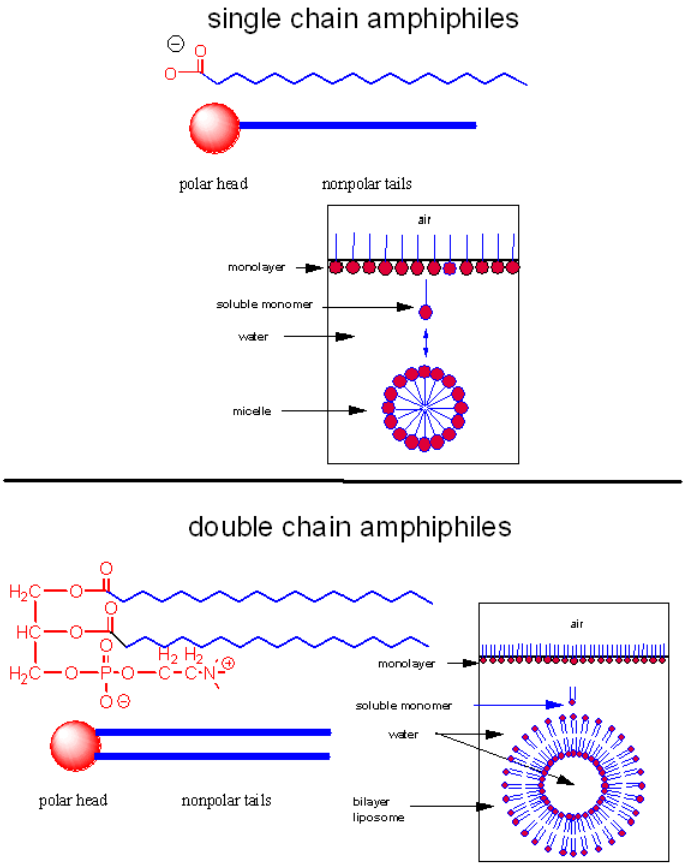
Polyanions as Carriers of Genetic Information

There are other reasons why polyanions are useful genetic molecules, other than their resistance to nucleophilic attack. The biological form of DNA is a large double stranded polyanionic polymer, in contrast to RNA which is a single-stranded polyanion polymer and protein which are polymers with varying combination of anionic, cationic, and hydrophobic properties. Even with counterions, it would be difficult to fold DNA into complicated and compact 3D structures as occurs for proteins, given the large electrostatic repulsions among the charged phosphates. Rather it forms a elongated double stranded rod, not unlike the rod-like structure of proteins denatured with sodium dodecyl sulfate (used in SDS PAGE gels). The elonged rod-shaped structure of ds-DNA is critical for the molecule which is the main carrier of our genetic information since mutations in the bases (leading to a switch in base pairs) causes no change in the overall structure of dsDNA. This enables evolutionary changes in the genetic material to produce new functionalities. A single change an amino acid of a protein, however, can cause a large change in the structure of a whole protein, a feature unacceptable for a carrier of genetic information. RNA structure effectively lies between that of DNA and proteins. Since it has less charge density than dsDNA, it can actually form dsRNA helices, so it can carry genetic information, as well as form complex 3D shapes necessary for its activity as an ribozyme. Perhaps more importantly, steric interference prevents ribose in RNA from adopting the 2'endo conformation, and allows only the 3'endo form, precluding the occurrences of extended ds-B-RNA helices.

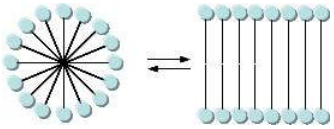


30.0.1: **The Lipid World**

Let's assume that abiological precursors would react to form polymer-like molecules that might be complex enough to fold to structures that would allow binding, catalysis, and rudimentary replication. All this would be worthless unless they could be sequestered in a small volume which would limit diffusion and increase their local concentration. What is required is a membrane structure. Amphiphilic molecules, like lipids, with which we started this book, would be prime candidates since they spontaneously assembly to form micelles and bilayers, as shown in the review diagram below.



As mentioned in Chapter 1, alternative lipid phases are possible. Bilayers can also be formed from single chain amphiphiles, such as certain fatty acids, as illustrated in the equilibrium shown below. This occurs more readily at pH values close to the the pKa of the fatty acid, at which the fatty acids are not all deprotonated with full maximal negative charges. Single chain amphiphiles like fatty acids, which were more likely to formed in abiotic conditions, have been found in meteorites.



Clay surfaces, which have been shown to facilitate the formation of nucleic acids polymers, can also promote the conversion of fatty acid micelles to bilayers (Szostak). One such clay surface, montmorillonite, whose structure is shown below, promotes bilayer formation. It has an empirical formula of $Na_{0.2}Ca_{0.1}Al_2Si_4O_{10}(OH)_2(H_2O)_{10}$.

EXTERNAL Chime Molecule Modeling: montmorillonite EXTERNAL: montmorillonite

The effect of montmorillonite on vesicle formation can be shown by simple measurements of turbidity with time. Microscopy of fluorophore-encapsulated vesicles also shows encapsulated montmorillonite. The fatty acids presumably absorb to the cation layer of the clay particles. Time studies using light scattering also indicate that the vesicles grow in the presence of fatty acid micelles. To differentiate between the formation of new vesicles and the increase in size of pre-existing vesicles (which couldn't be done by simple light scattering without separation of the vesicles), investigators used two different fluorescent molecules to label fatty acid vesicles. The two probes were selected such that if the two probes came in close contact, energy transfer from the excited state of one fluorophore to the other fluorophore could occur, an example of fluorescence resonance energy transfer (FRET). FRET is observed when emission of the second dye occurs after excitation of the first dye, at a wavelength outside of the excitation

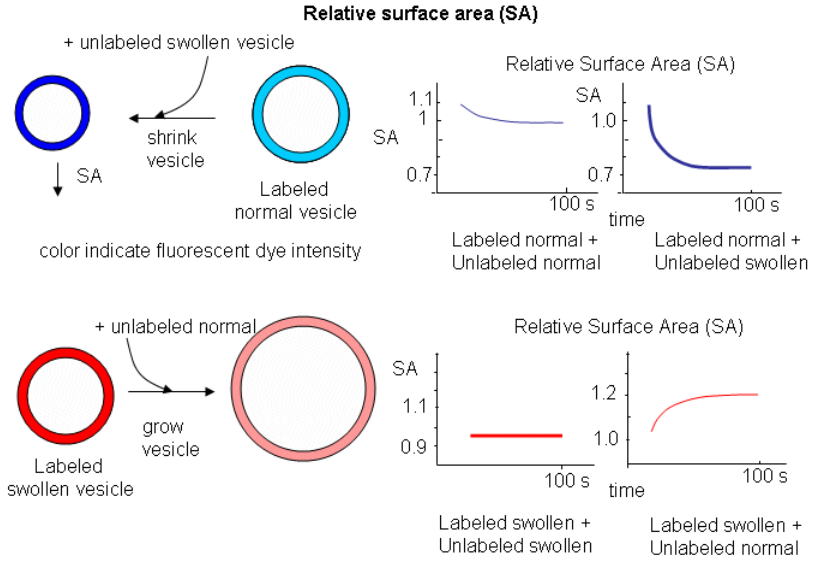
wavelength of the second dye. If unlabeled vesicles were added to either labeled vesicles, no changes in FRET were observed, suggesting that the dyes did not move between vesicles. If fatty acid micelles were added, a decrease in FRET was observed, suggesting that new fatty acids were transferred to the doubly-labeled vesicles, effectively diluting the dye concentrations in the bilayer and their relative proximity, both which would decrease FRET. Most of the new fatty acid was incorporated into pre-existing vesicles which grew. The vesicles could also divide if extruded through a small pore. Later we will see that the energy to grow the vesicles can derive in part from a transmembrane proton concentration collapse. Division of vesicles might be promoted by bilayer assymetries associated with addition of substances to the outer leaflet, causing membrane distortion.

30.0.1: **Protocells**

At some point, early genetic material must have been encapsulated in a membranous vesicles. Would new properties emerge from this mixture that might have a competitive (evolutionary) advantage over either component alone, and thus be a step on the way to the formation of a "living" cell? The answer appears to be yes. Chen et al. have incorporated RNA into fatty acid vesicles with interest effects. They asked the question as whether those vesicles could grow at the expense of vesicles without encapsulated RNA. RNA, with a high charge density and its associated counter ions would create osmotic stress on the vesicles membranes. To relieve that stress they could acquire fatty acids from other fatty acid vesicles (or fatty acid micelles), increasing their surface area, and concomitantly reducing tension in the membrane.

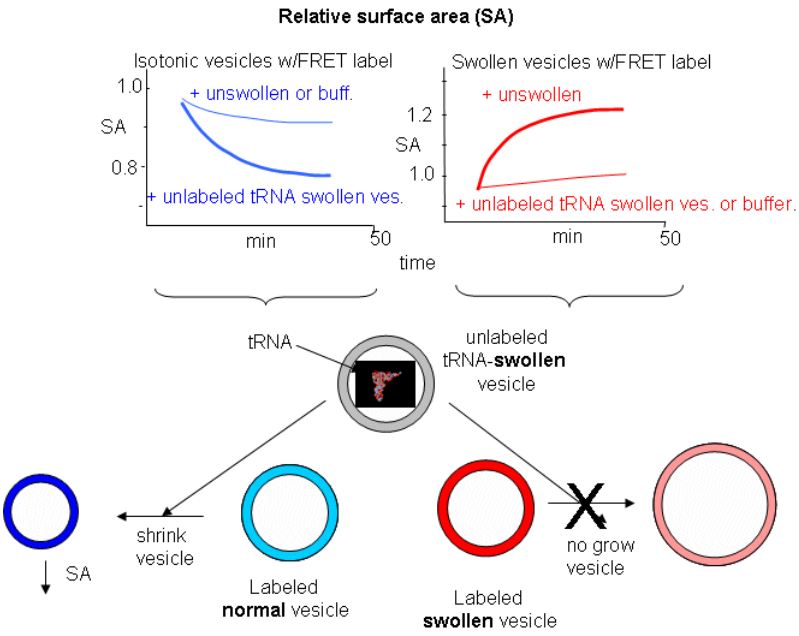
Oleic acids vesicles were first placed under stress by encapsulating 1 M sucrose in the vesicle and then diluting it in hypotonic media. Water would enter and swell the vesicle (but without bursting and resealing, as evident from control experiments). Then they prepared stressed and unstressed oleic acid vesicles in the presence of two nonpolar flurophores, NBD-PE (excitation at 430 nm, emission at 530 nm) and Rh-DHPE (emission at 586 nm). These fluorophores were chosen for fluorescence resonance energy transfer measurements. If the membrane vesicles changed size, the FRET signal would change, based on the relative concentration and proximity of the dual fluorophores. If the separation between probe molecules increased, the FRET signal would decrease. Conversely, if the vesicle shrunk, the FRET signal would increase.

The results showing the effect of adding unlabeled swollen vesicles to labeled normal vesicles, and labeled swollen vesicles to unlabeled normal are shown below. The surface area of normal labeled vesicles decreased by about 25% when unlabeled swollen vesicles were added, but not when unlabeled normal vesicles were added. Labeled swollen vesicles increased 25% in size only if mixed with unlabeled normal vesicles, not with unlabeled swollen vesicles. Hence swollen vesicles win the competition and "steal" lipid from normal vesicles.



After Chen et al. Science 305, 1474 (2004)

Now what about vesicles swollen with encapsulated RNA? RNA, with its associated charge and charged counter ions also placed an osmotic stress on the vesicles. FRET labels (the two fluorophores) were place in vesicles without RNA. Fatty acids were removed from isotonic labeled vesicles in the presence of unlabeled tRNA swollen vesicles (left panel below). Labeled vesicles swollen with glycerol took fatty acids from unswollen vesicles (without tRNA), but not from tRNA swollen vesicles, as both were swollen so no net drive to reduce swelling by lipid exchange was present.

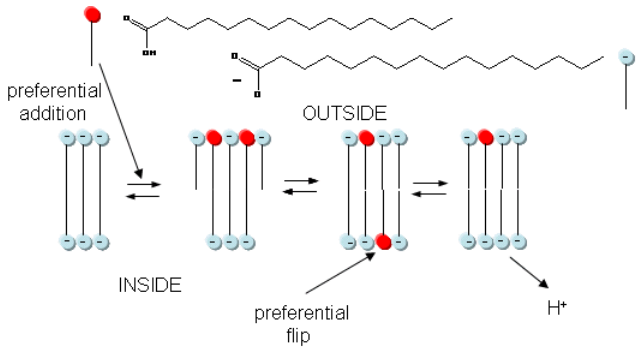


After Chen et al. Science 305, 1474 (2004)

These results show the vesicles with encapsulated RNA have a competitive (evolutionary) advantage over normal vesicles. This data suggests that having a polyanion as the source of genetic material is actually advantageous to the protocell. In addition the move in modern membranes to phospholipids with esterified fatty acids (instead of free ones) may actually have stabilized membranes, given the movement of free fatty acids to different membranes.

30.0.1: Energy Transduction in Protocells

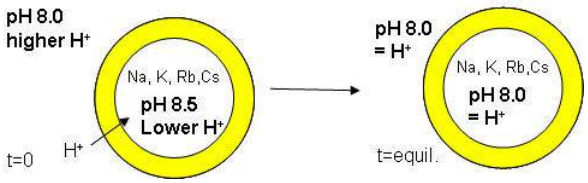
In addition to a genetic macromolecule and a semipermeable membrane, a source of energy to drive intracellular processes must be present. A common source of free energy used in many cells to drive unfavorable processes is a proton gradient, whose formation in modern cells can be coupled to energy input from oxidation, ATP cleavage, light, or the collapse of another gradient. Could a proton gradient be formed in protocells? It can, quite easily, when coupled to the growth of fatty acid vesicles. If a fatty acid vesicle is to grow, more fatty acid must be added to the outer leaflet. The protonated, uncharged form of the fatty acid would preferentially be added, since it would lead to less electrostatic repulsion between adjacent head groups. The protonated, uncharged form of the fatty acid would also be most likely to flip to the inner leaflet to minimize stress asymmetries in the leaflets. Once in the inner leaflet, it could deprotonate to form H+(aq) in the inside of the membrane, creating a transmembrane proton gradient and transmembrane potential. The energy released on growth of the membrane is partly captured in the formation of a proton gradient, as shown in the figure below.



Generation of proton gradient on growth of fatty acid vesicle

The proton gradient would soon inhibit its own formation since further movement of protons into the cell would be attenuated by the positive transmembrane potential unless metal ions inside moved outside. In addition, the gradient would collapse after growth stopped. The investigators made fatty acids vesicles in the presence of pH 8.5 buffers whose pH was adjusted with an alkali metal

hydroxide. The external pH was reduced to 8.0, resulting in a 0.5 pH unit proton concentration gradient. (Changes in intravesicular pH were measured with pH-sensitive fluorophore, HPTS.) Inward movement of protons down a concentration gradient, as shown in the figure below, would occur with time, collapsing the imposed concentration gradient.



With fatty acid vesicles, this artificial pH gradient collapsed quickly, suggesting the vesicle permeability to protons was high. The rate was too high for simple flip-flop diffusion. Inward movement of protons appeared to be facilitated by outward movement of the M+ ions. The rate of decay of the proton gradient was exponential, and the resulting first order rate constant was easily determined. A graph of the rate constant for pH gradient collapse vs unsolvated ionic radius of M+ decreased with increasing radius (i.e. kNa > kK > KRb > KCs, suggesting that the pH gradient would be more stable if large, impermeable or otherwise trapped cations were encapsulated. When vesicles were made with encapsulated Arg+, the imposed pH gradient did not collapse for hours. If oleic acid micelles were added to oleic acid vesicles with encapsulated Arg+, with no artificial pH gradient induced across the membrane, the vesicle grew with concomitant movement of protons into the vesicle, producing a pH gradient of 0.3 within seconds.

These experiments show that membrane growth and energy storage could be coupled, and the right composition of encapsulated material could lead to a stable transmembrane pH gradient, a source of energy to drive biological processes. It even suggests that a charge polyanion would be beneficial as a genetic carrier.

30.0.1: Hydrothermal Vents or Primordial Soup

The case for the origin of life in deep sea hydrothermal vents and not in a primordial "Campbell's" soup has been argued convincingly by Lane et al (2010). What's needed for life are reasonably complex molecules and an energy source to drive unfavorable reactions. It's the latter on which that Lane et al focus. In an early primordial world that was low in oxygen, exergonic oxidation reaction of organic molecules would provide little energy. This can be surmised from the low energy yield (compared to aerobic respiration) achieved in present day glycolytic (fermentative) pathways from all major domains of life, archaea, bacteria, and eukaryotes.

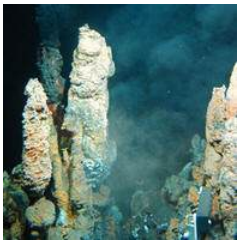
> Background: Based on rRNA sequences, a primordial cell evolved into two different types of cells, one that became bacteria, and another that split further into archaea (single cells, similar to bacteria) and eukaryotic cells (complex cells with internal organelles that eventually formed multicellular organisms). Bacterial and archaea are collected called prokaryotic cells.

In addition, these pathways required the evolution of up to dozen different enzymes to produce their relatively meager energy yield which ultimately depends on the oxidation of an organic molecule by another organic molecule instead of by a powerful oxidant like dioxygen. An anisotropic arrangement of molecules in a concentrated soup could lead to transient chemical potential fluctuations but these would be inefficient and impermanent sources of energy. Effectively the primordial soup would be at equilibrium and hardly expected to provide the energy for synthesis of RNA enzymes and replicators. UV light leads to photo-damage and photolysis not replication of complex molecules. What is needed is a way to drive the synthesis of molecule with high chemical potential energy (like sulfur esters and phosphoanhydrides) compared to their lytic products. These could then provide an energy sources to drive ATP synthesis, for example.

A detailed look a the bioworld shows that the earliest organisms used energy from the collapse of the proton gradient (chemisomotic principle elucidated by Peter Mitchell). All present autotrophs (organisms that can fix CO2 and form complex organic molecules) and many heterotrophs (use complex organic molecules of other organisms for fuel) use redox complexes in membranes coupled to membrane gradients. These complexes would take reduced molecules and pass electrons from them to oxidizing agents (electron acceptors), including O2, CO2, and Fe3+ to form H2O, CH4, and Fe2+. Fermentors also use ATPase membrane enzymes to transport nutrients. Yet genomic analysis of bacteria and archaea show that enzymes involved in fermentation differ significantly, suggesting that they evolved separately towards a convergent function. Structure in common include DNA, RNA, ribosomes and membrane ATPases, which Lane et al suggest were in a the Last Universal Common Cell (LUCA).

All autotrophs produce their energy source by fixing CO2 using either H2 directly or indirectly using H2O and H2S. All of the are available in nonhydrothermal deep sea vents. Volcanic vents, however, are extremely hot (not optimal for organic molecule synthesis), very acidic, and lack hydrogen gas. A different type of nonvolcanic vent, an alkaline hydrothermal one, might produce more conducive as a site of the origin of life. In these vents, water chemically reacts with minerals in the crust (such as olivine) leading to their hydroxylation and subsequent fracture, with promotes more water entry into the crust. It has been reported that there is more water found as hydroxylated minerals in the crust, that there is liquid water in the oceans. These processes result in temperatures up to 200 degrees Celsius and release of hydrogen gas into a moderately alkaline vents into the sea water at temperatures more conducive (70 degrees C) to the origin of life.

Figure: Alkaline Vent



30.0.1: Fixing CO_2

Of the five different pathways known to fix CO2, all require ATP except one. That one is present in both methanogen, which produce methane from CO2 and H2 and in the acetogens, which produce acetate (CH3CO2-) in the form of acetylCoA. The simpler reactions of forming acetic acid and methane are shown below:

2CO2 + 4H2 --> CH3CO2H + 2H2O.

CO2 + 4H2 --> CH4 + 2H2O.

The DG0 values for these reaction (calculated using DG0f for gas phase H2, CO2 and CH4 and liquid acetic acid and water are -75 and -131 kJ, respectively, at 250C, showing that they are thermodynamically favored. Making AcetylCoA, a "high" energy molecule compared to its hydrolysis products (as is ATP) from acetic acid and CoASH, a would require energy input. A proton gradient is the likely source.

Some bacteria and Achaea cells (primordial or present) use the reductive acetylCoA pathway, also known as the Wood-Ljungdahl pathway, to form, in a noncyclic process, acetyl CoA from CO2 and at the same time makes ATP. This process is paid for by a proton gradient. This has been described by Shock as "a free lunch you get paid to eat". The energetics of the present acetyl CoA pathway based on the overall reaction below show an approximate DG0 value of -59 kJ/mol which can drive ATP synthesis.

2CO2 + 4H2 + CoASH --> CH3COSCoA + 3H2O.

The concentration of carbon dioxide in the primordial ocean was 1000 times higher than now. Vents produced large amounts of methane and hydrogen gas. There was little oxygen and hence lots of Fe2+. The enzymes involved in this acetyl-CoA pathway of carbon fixation have FeS clusters. It has also been shown that bubbles (which are really membrane bound spaces) of FeS and NiS can be made in deep sea vents. These could not only encapsulate precursor molecules but also serve as catalysts. Vents also can catalyze the fixation of nitrogen (to ammonia) and laboratory studied show that FeS can catalyze the conversion of formate (found in vents) into pyrimidines and purines. The studies of present methanogens and vent chemistry suggest that the critical ingredients and conditions for development of the first biological cells probably occurred in the vents.

To produce polymers, an energy source and monomers must exists. Concentration gradients found in simulations of vents produce million fold concentrated molecules. The transient heating and cooling of any double-stranded nucleic acids could lead to concentration amplification by a PCR like strand separation followed by reannealing. In addition, these vent regions possess a powerful, reoccurring energy source, a pH gradient, as the alkaline vented material entered the acidic oceans that exists with high CO2 concentrations, creating a gradient across an inorganic membrane. This is startlingly analogous to the pH gradient across membranes (acidic outside, alkaline inside) driven by the membrane complexes in the mitochondria and bacteria. Lane et al argue that the existence of membrane proton gradients as an energy source in all cells (eukaryotes, bacteria, and archaea) and in chloroplasts, mitochondria, corroborate their hypothesis. Bacteria and archaea share homologous ATPases and electron carriers (ferredoxins, quinones, and cytochromes). These similarities contrast to the differences in enzyme structures in fermentative pathways. Arguments that proton pumps evolved to pump proteins (and reduce pH gradients) can't explain their ubiquitous

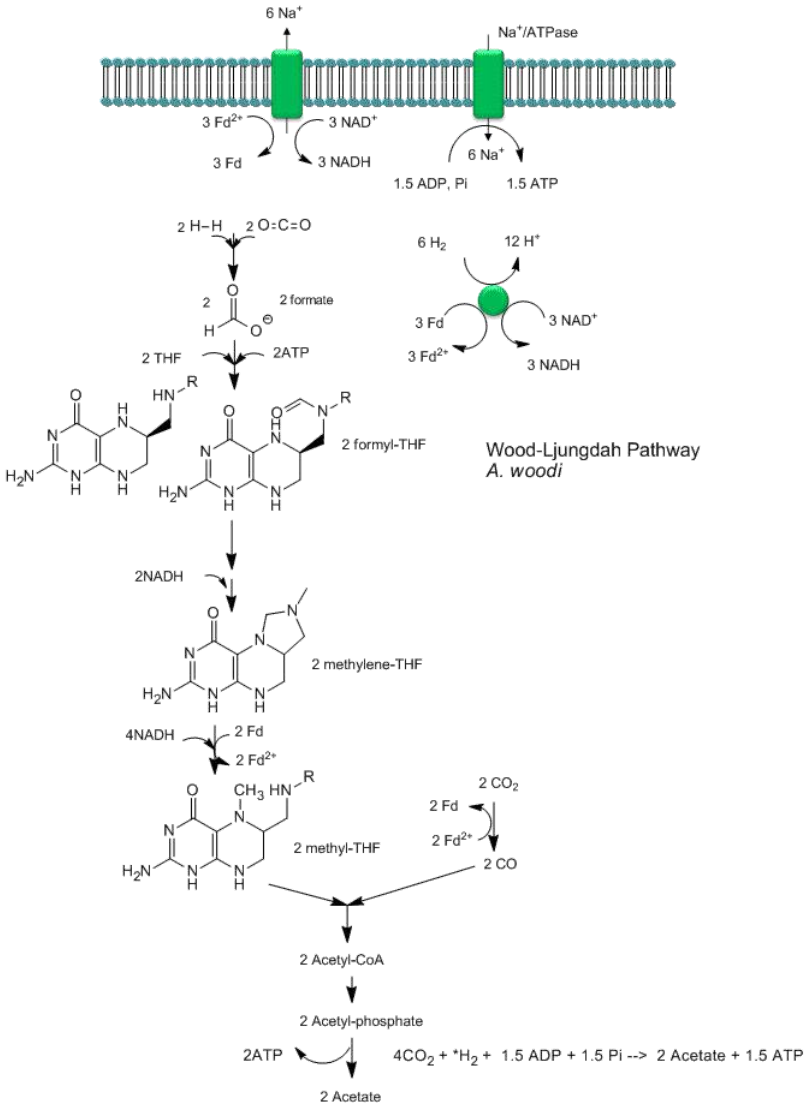
presence even in organisms not subjected to low pH. Hence the ubiquity of proton pumps supports the conjecture that they arose from the first protocells, possible comprised of inorganic walls and ultimately with amphiphilic molecules synthesized from precursors.

Creationists would argue that it would be impossible to evolve a structure with the complexity of membrane ATPase (which serve to collapse a pH gradient as the power the synthesis of molecules with large negative DG0 of hydrolysis). Lane et al propose that the earliest cells evolved ATPase like molecules in alkaline vents where pH gradients analogous to those in cells today arose. They envision cell-like columns lined by FeS membrane like structure with alkaline conditions inside and acid conditions outside. Nonpolar or amphiphilic molecule would line the inside of the cells/columns. A ATPase-like system could then take advantage of the pH gradient which constantly replenishes itself. If structures as complicated as ribosomes evolved from a subsequent RNA world, surely ATPase-like molecules could also. Other chemistry might have evolved earlier to utilize the energy source provided by the pH gradient.

If life originated in the vents, it would need an energy source to leave the vents. Presumably it would have evolved one to utilized pH gradient to replace the one it left in the alkaline vents. The substrate level phosphorylation of glycolysis that requires ATP input to make ATP would not provide the energy source needed. Cells that left would have had to produce their own proton gradient. Perhaps all the was needed initially was concerted conformational changes in proteins that upon exposure of a different pH changed their shape inducing pKa shifts in adjacent proton donors/acceptors leading to vectorial discharge of protons across a membrane. Perhaps the method described above in protocells was sufficient.

Recent analyses by Poehlein et a show that CO2 reduction (fixation) can be coupled to the production of a sodium ion gradient, which could collapse to drive ATP synthesis. Analysis of the genome of a gram positive bacteria, Acetobacterium woodii, an acetogen, shows the it has an ancient pathway for production of acetyl-CoA that can, in an anabolic fashion form biomass or in a catabolic fashion be cleaved to acetate with the production of ATP. It does not require classic electron carriers like ubiquinone or cytochrome C linked to protein gradient formation to drive ATP synthesis. Rather it has only a ferredoxin:NAD+ oxioreductase which couples oxidation to the formation of a sodium ion gradient, which collapses through an sodium ion transporter/ATP synthase to drive ATP synthase. A plausible reaction scheme based on genomic analysis is shown below:

Figure: Acetyl-CoA Synthase and Acetogenesis



30.0.1: **The Role of Fe/S Center**

Let's return to the chicken and egg dilemma one more time. What is needed for biological polymer formation are monomeric precursors, an energy source, and a way to compartmentalized them all. We discuss how monomeric precursors could form, but wouldn't it be far better if even the synthesis of precursors could be catalyzed? One source of catalysis mostly absent from the "bioorganic" abiotic chemistry in the above discussion is the transition metals. Transition metals can form complexes. Ligands containing lone pairs on O, N, and S atoms can donate them to transition metals ions, which can hold up to 18 electrons in s, p, and d orbitals. Hence as many as 9 lone pairs on ligand molecules (which are often multidentate) could be accommodated around the transition metal ion. Many present small molecule metabolites and their abiotic precursors (H2O, CO, CO2, NH3 and thiols) bind cations as mono- or polydentate donors of electrons. Hence transition metal ions would have a thermodynamic tendencies to be bound in complexes.

Bound ligands that contain potentially ionizable hydrogens could become deprotonated and made better nucleophiles for reactions. Hence the transition state metal ion, acting with the complex, becomes a catalyst as it decreases the pKa of a bound ligand (such as water). In addition, since transition metals ions can have multiple charge and oxidation states, they can easily act as redox centers in the oxidation/reduction of bound ligands that were redox active. Given the relative anoxic conditions of the early oceans, Fe2+ would predominant. It could easily be oxidized to Fe3+ as it reduced a bound ligand. Highly charged transition states would withdraw electron density from bound ligands leading to their possible oxidation.

Metals obviously still play a strong role in catalysis, both indirectly in promoting correct protein folding and directly in stabilizing charge in both the transition state and intermediates in chemical reaction pathways. FeS clusters are of significant importance.

Their biosynthesis involves removal by an active site Cys in a desulfurase enzyme of a sulfur from a free amino acid Cys followed by its transfer to an Fe in a growing FeS cluster in a FeS scaffold protein, which then transfers the cluster to an acceptor protein where it acts as a cofactor. FeS clusters can adopt a variety of stoichiometries and shapes, as well as redox states for the participating Fe ions. The continuing importance of FeS clusters in all cells, their involvement in not only redox enzymes in which electron transfer is facilitated by delocalization of electrons over both Fe and S centers, but also in coupled electron/proton transport in mitochondrial electron transport, Fe storage (ferrodoxins), and in regulation of enzyme activity and gene expression, suggests that they were of primordial importance in the evolution of life. T

hey are often found at substrate binding sites of FeS enzymes involved in both redox and nonredox catalysis. A ligand can bind to a particular Fe in the cluster, activating it for hydration or dehydrogenation reactions. Fe 4 of the FeS cluster in the TCA enzyme aconitase can have a coordination numbers of 4, 5, or 6 as it binds water, hydroxide or substrate. It acts to both decrease electron density in the transition state and to change the pKa of bound water as the enzyme catalyzes an isomerization of tricarboxylic acids (citric and isocitric acid) through an elimination/addition reaction with water. In another example it can bind S-adenosylmethionine through its amine and carboxylate groups, which activates the molecule for cleavage and radical formation. In some cases metals other than Fe (Ni for example) are incorporated into the cluster. FeS effects on transcription factors involves facilitation of optimal structure for DNA binding. FeS and FeNi centers in proteins are similar in structure tp FeS units in minerals like greigite and presumably to FeS structure formed when H2S and S2- react with Fe2+ (present in abundance in the early ocean) and other metals in vents Metal sulfides participate in reduction of both CO and CO2. For example the synthesis of CH3SH from CO2 and H2S is catalyzed by "inorganic" FeS.

30.0.1: **The Minimal Genome**

This question is being addressed by eliminating "unnecessary" gene from simple bacteria. Cells placed in a rich nutrient broth with essential lipids, vitamins, and amino acids would need fewer genes than those placed in a more nutrient-poor medium. Bacteria cells like Mycoplasma genetalium, that live within "nutrient rich" eukaryotic cell, have been genetically manipulated to delete unnecessary genes. Based on knockout studies, it may be possible for the cell to survive with only 300-350 genes. Bacillus subtilis has approximately 4100 genes. Estimates have been made that it could survive with as few as 271 genes.

more to be added

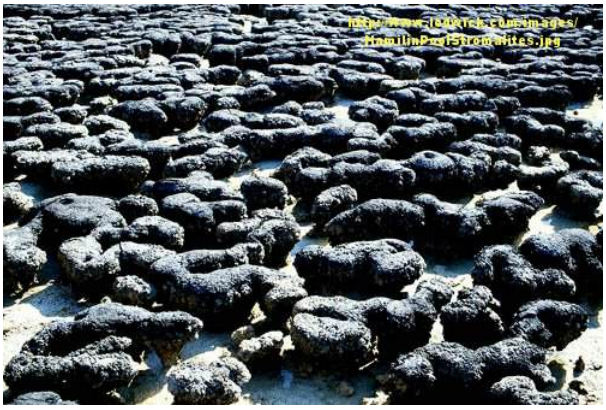
30.1: Abiotic Origins of Life

"Ugly giant bags of mostly water" -
The crystal life form describing humans on Star Trek Next Generation, Home Soil Episode.

30.1.0.1: **The Start of Life**

This book began with the notion that understandings derived from the study of simple molecules can be applied to complex biological macromolecules and systems. We developed an understanding of the structural, thermodynamic, and kinetic properties of the "simplest" biomolecules, including single chain amphiphiles like fatty acids, and double chain ones like phospholipids, and how these properties could explain the propensity of these molecules to form complex lipid aggregates (micelles and bilayers). We extended these ideas to the process of protein folding and the assembly of biological complexity. Is there something intrinsic to the property of molecules such that their localization together in the right microenvironment could lead to a "living cell"? How did life originate? That is the topic of this last capstone chapter.

Defining life is actually quite difficult. Here is a list of requirements that seem reasonable, but other have noted that this list would exclude the mule. Life can self-replicate, self-sustain, evolve, respond to environmental changes, and die. The earliest known fossils (stromatolites from cyanobacteria) are approximately 3.5 billion years old.



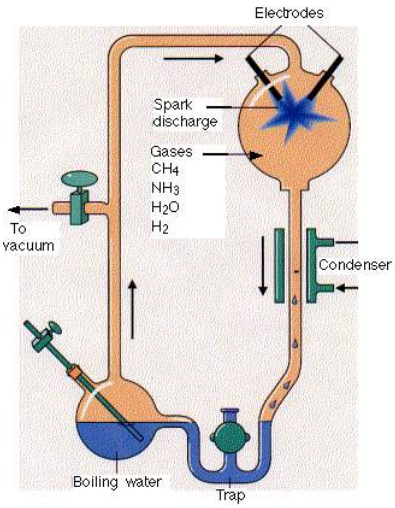
We have just finished studying the complex interactions involved in cell signaling. How could they have evolved? Consider the central dogma of biology. In the present biological world, proteins (DNA polymerase, RNA polymerase, transcription factors) are necessary for DNA synthesis, replication and gene transcription. But you need DNA to encode the proteins. This "chicken vs egg" dilemma has been addressed when it was realized that RNA can both carry genetic information as well as enzymatic activity (even at the level of the ribosome used for protein synthesis).

https://clockwise.software/blog/solv...d-egg-problem/ for new picture

30.1.0.1: Abiotic Synthesis of Amino Acids and Peptides

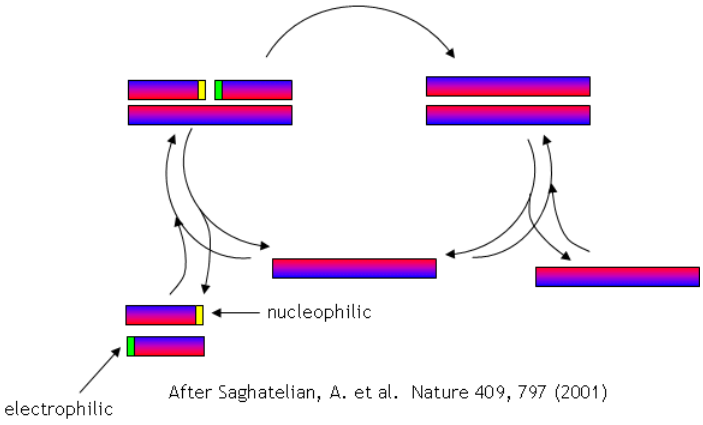
Much work has been done to determine if the building blocks for present biological molecules could have been synthesized early in Earth's history. Amino acids and fatty acids have been found in meteors suggesting the possibility. Earth's early atmosphere would have had little oxygen, so most components should have been reduced. It probably consisted of methane, ammonia, hydrogen and water similar to the

atmospheres of other planets in our solar system. The composition of the early atmosphere is still contentious. In 1953 (the same year that Watson and Crick published the structure of double-stranded DNA), Stanley Miller showed that electric discharges (to simulate lightening) in a reducing atmosphere over a "simulated sea" produced many amino acids. Up to 11 different amino acids have been produced in this fashion along with purines and pyrimidines (these required concentrated reaction mixtures) necessary for nucleic acids. Adenine can be produced just through the reaction of hydrogen cyanide and ammonia in an aqueous solution. Other nucleic acid bases can be made with hydrogen cyanide, cyanogen (C2N2) and cyanoacetylene (HC3N).



http://www.hencoup.com/Photo%20Stanley%20Miller.jpg

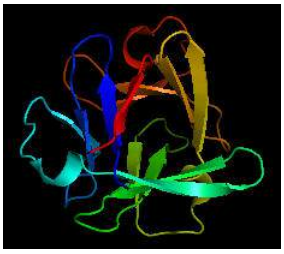
No complex polymers arise through these reactions. However, in 2004, Lehman, Orgel and Ghadiri were able to show that in the presence of carbon disulfide, a gas discharged from volcanoes, homo- and hetero-peptides were produced. Amphiphilic peptides can even catalyze their own formation from peptide fragments, if the fragments are activated. The fragments would bind to the larger "template" peptide through nonpolar actions of the side chains which are oriented along one face of the helical axes. If the fragments bind in a fashion in which the electrophilic end is adjacent to the nucleophilic end of the other peptide fragment, condensation of the two peptide fragments results. The larger template peptide acts as a template (effectively as an "enzyme") in orienting the two fragments for chemical reaction and effectively increasing their local concentration. The reaction of the bound fragments is essentially intramolecular. The reaction even proceeds with amplification of homochirality.



Could the prebiotic amino acids have polymerized into a protein that could fold in a fashion similar to modern proteins? That question has recently been addressed by Longo et al (2013). They asked the question whether the amino acids found in Miller-type prebiotic synthesis mixture and in comets/meteors (Ala, Asp, Glu, Gly, Ile, Leu, Pro, Ser, Thr and Val), a restricted set (10)

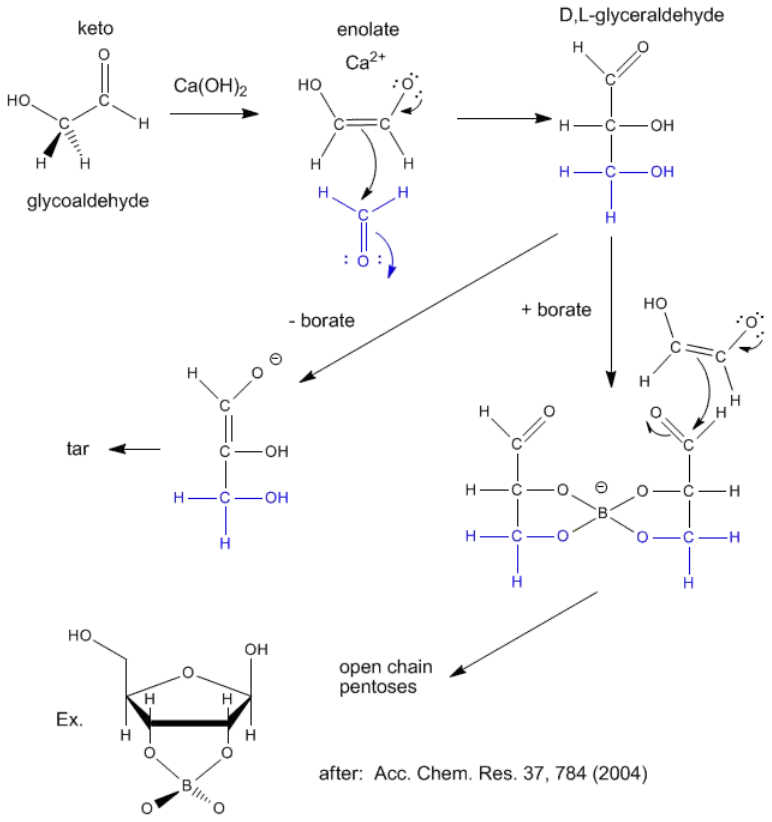
compared to the present 20 naturally-occurring amino acid, could form a polymer that could fold. Notice that this reduced ensemble of amino acids lacks aromatic and basic amino acids. These proteins would be acidic with a low pI and may have trouble, given the lack of nonpolar aromatic amino acids, in forming a buried hydrophobic core which stabilize proteins. Nevertheless Longo et al were able to synthesize a protein with a slightly expanded set of amino acids (12, including Asn and Gln, with 70% prebiotic amino acids). The structure of one of the proteins, PV2, is shown below. The protein was more stable in 2 M NaCl (compared to 0.1 M) in which it showed a cooperative thermal denaturation with a melting point near 650C using differential scanning calorimetry. The protein had properties similar to those from halophilic organisms that thrive in high salt. These properties include low pIs and high negative charge density, which allows cation-protein interactions in the high salt environment, and lower stability in low salt environments. Earlier oceans were saltier. Halophiles are an example of extremeophiles which are highly represented in archea. Although most halophiles are aerobic, some are anerobic. Perhaps life arose in high salt environments.

Figure: Structure of PV2 protein comprised of a reduced alphabet of mainly prebiotic amino acids.



30.1.0.1: Abiotic Synthesis of Sugars

Sugars are required for present energy production but also as a part of the backbone (ribose, deoxyribose) of present genetic material. Many sugars can be synthesized in prebiotic conditions, using carbon based molecules with oxygen, such as glycoaldehyde and formaldehyde (both found in interstellar gases), as shown in the figure below. Glycoaldehyde was recently found in star forming regions of the Milky Way where planets are likely to form. The presence of borate, which stabilizes vicinal OHs on sugars, is required for the production of sugars instead of tars.



RNA molecules containing sugars such as threose, aldopentopyraonses and hexopyranose can also form stable secondary structures like helices. (Remember, RNA probably preceded DNA as the genetic carrier of information given that is also has enzymatic activity). Is there something special about ribose that made it selected over other sugars for nucleic acids, especially since it is found in low abundancy in the products in synthesis reactions conducted under prebiotic conditions? One probable reason is it unusually high (compared to other sugars) permeability coefficient through vesicles made of phospholipids or single chain fatty acids, as shown below.

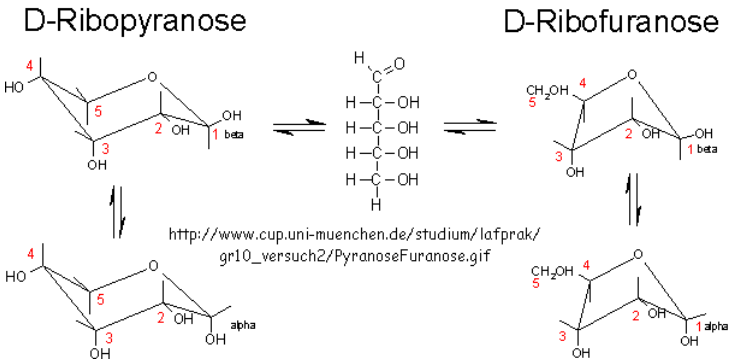
Table 1. Solute permeability coefficients (P_s) across fatty acid and phospholipid membranes

	Solute	Membrane composition: Myristoleate*	Palmitoleate*	Oleate*	Oleate†	POPC*	DPPC/C/PA‡
	Glycerol§			490			
4C	Erythritol§	28	14	12	4.8	21	
4C	DL-Threitol	50	23	22	11	40	
	Adonitol	2.9	1.7	0.91	0.38		
	Arabitol	2.1	1.4	0.63	0.39		
	Xylitol	5.0	3.0	1.8	0.95		
	Dulcitol				0.013		
	Mannitol	0.051	0.015	0.0058	0.0058		
	Sorbitol	0.18	0.071	0.026	0.027		
5C	Arabinose	3.5	2.1	1.1	0.49	2.8	1.1
5C	Lyxose	5.7	3.2	1.9	0.84	4.5	1.2
5C	Ribose	31	15	11	2.9	20	9
5C	D-Xylose	3.4	1.9	0.98	0.51	2.8	1.0
5C	L-Xylose	3.3	1.9	0.94		2.7	
5C	Ribulose			31			
6C	Galactose	0.060	0.024	0.011	0.0086		
6C	Glucose§	0.047	0.017	0.0071	0.0050		
6C	Mannose	0.10	0.035	0.018	0.014		
6C	Fructose§	1.1	0.68				
6C	L-Sorbose	0.88	0.56				

Values reported in 10^{-8} cm/s. Chiral solutes are D-enantiomers unless otherwise noted.
*Experiments performed at 23°C in 1 buffer, 0.5 M solute.
†Conditions as above but with 0.1 M solute.
‡Conditions as above, values are relative initial slopes only (D-xylose set to 1.0).
§Literature P_s values across lecithin membranes: glycerol, 540 × 10^{-8} cm/s (33); erythritol, 75 × 10^{-8} cm/s (34); glucose, 0.003 × 10^{-8} cm/s (35); and fructose, 0.04 × 10^{-8} cm/s (35).

PNAS | April 26, 2005 | vol. 102 | no. 17 | 6005

In general, the greater the number of carbon atoms, the smaller the permeability. However, the table above clearly shows large differences in permeability for sugar isomers with the same number of C atoms, and the difference is not affected by the lipid composition. Ribose has markedly elevated permeability compared to other 5C sugars (as do erythritol and threitol among 4 C sugar alcohols). What is so unique about ribose? 20% of the sugar is in the furanose form. Rate constants for ring opening of furanoses are elevated, suggesting greater flexibility. The α-furanose anomer is amphiphilic in that one face is hydrophobic and the other hydrophilic. All of these may promote ribose permeability.

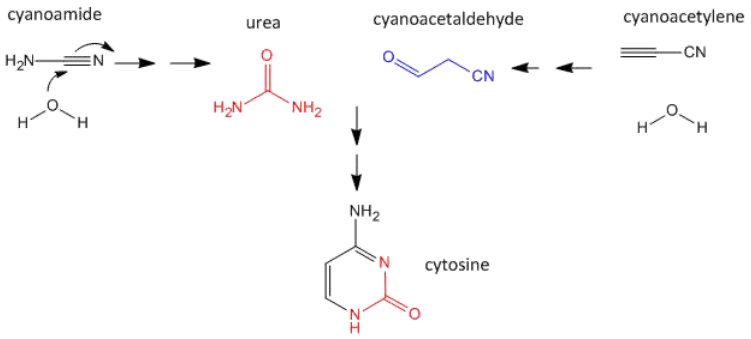


EXTERNAL Jmol: a-D-ribofuranose alpha-D-ribopyranose at

30.1.0.1: Abiotic Synthesis of Nucleobases

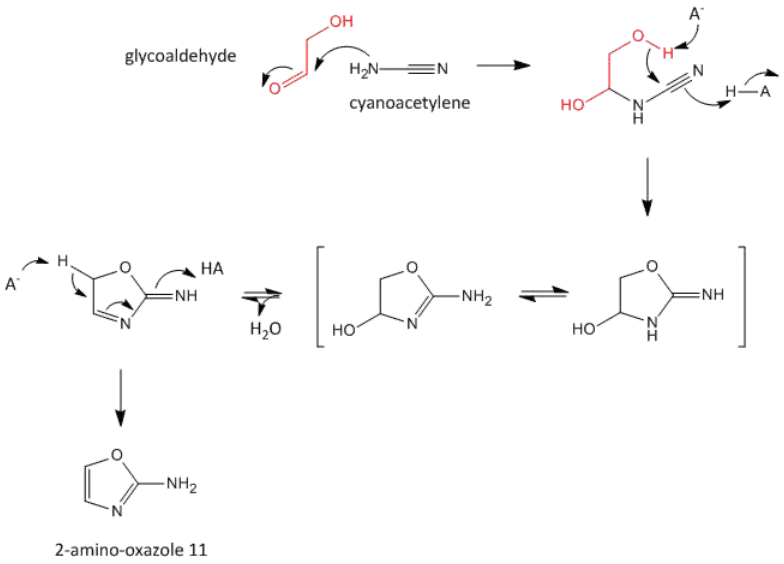
As mentioned above, nucleobases can be made under abiotic conditions with appropriate sources of carbon molecules with nitrogen. These include hydrogen cyanide, cyanoamide (NH2CN) (along with ammonia), hydrogen cyanide, cyanogen (C2N2) and cyanoacetylene (HC3N). Cyanoamide and cyanoacetylene can both react with water to form urea and cyanoacetaldehyde, respectively. The latter two can condense to form cytosine as shown below.

Figure: Abiotic Synthesis of Cytosine



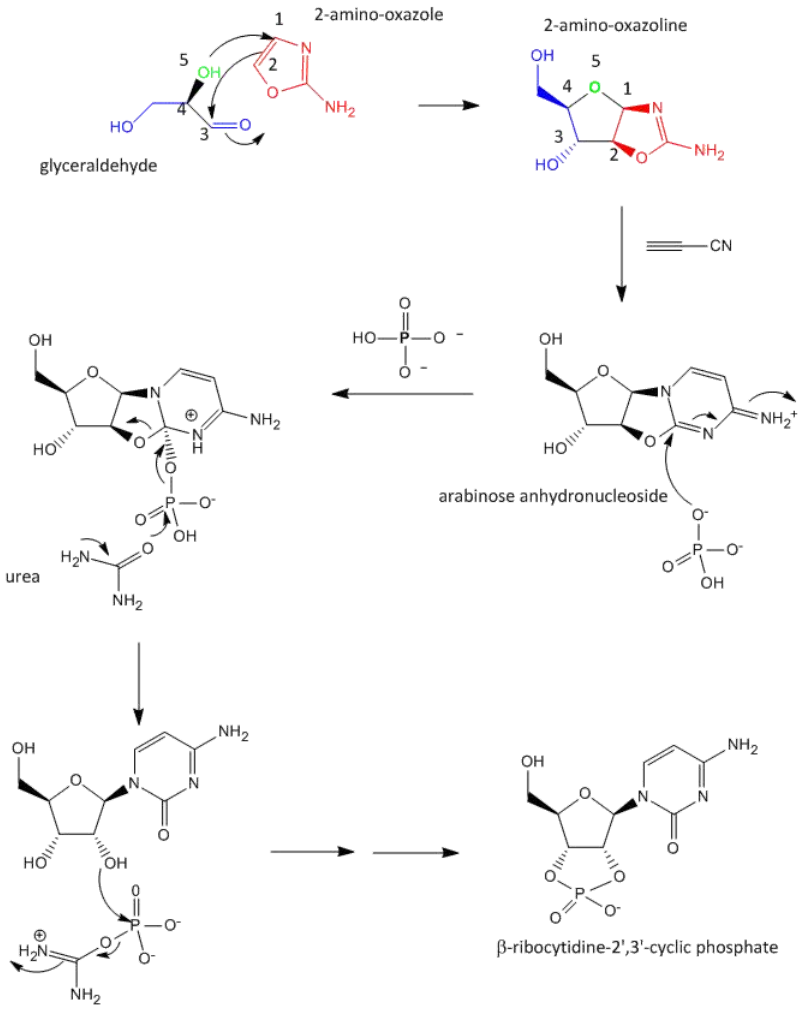
A major problem that has plagued prebiotic researchers is how the product of the carbon-oxygen compound (sugar) links to the product of the carbon-nitrogen compound (nucleobase) to form the nucleoside. Recent research by Powner et al has offered an innovative solution. Instead of forming the sugar and base in separate reaction, and then linking them covalently, the combined molecule could be synthesized in a single set of linked reactions. Inorganic phosphate serves as both a general acid/base catalyst (HA/A- in the figure below) in these new reactions in the formation of an important intermediate, 2-amino-oxazole, and as a nucleophilic catalyst.

Figure: Abiotic Synthesis of 2-amino-oxazole



Glyceraldehyde, 2-amino-oxazole, and inorganic phosphate can react to form a ribocytidine phosphate. Possible reaction mechanisms are shown below.

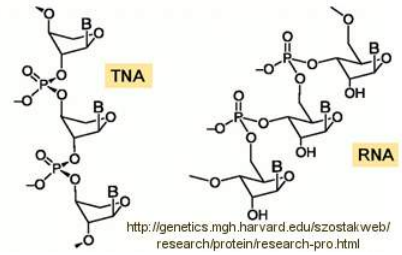
Figure: Abiotic synthesis of a ribocytidine phosphate without condensation of a preformed ribose and cytosine



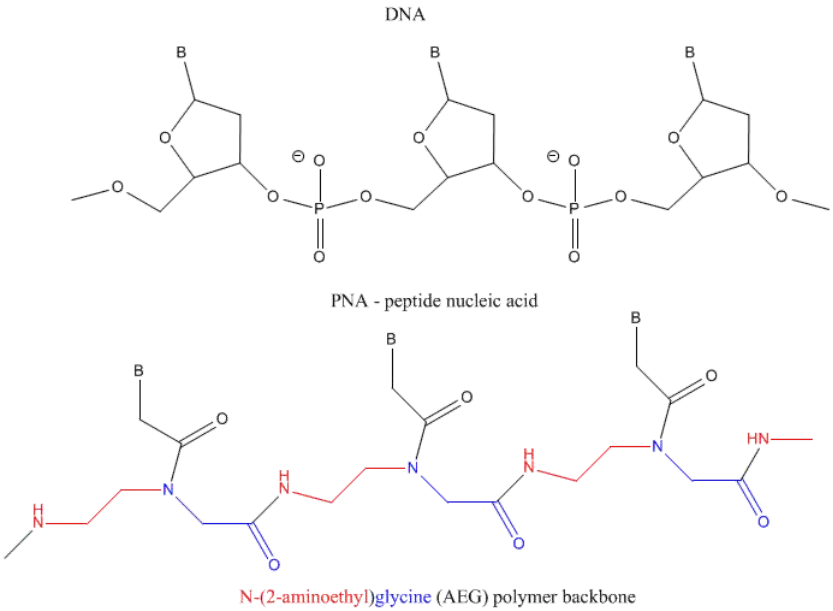
30.1.0.1: Abiotic Synthesis of Genetic Polymers

Abiological synthesis of polymer precursors is a long way from creating genetic polymers like RNA and DNA. These genetic polymers have one property that at first glance seems not conducive to a genetic molecule. Both are polyanions, which must be packed into a cell and folded onto itself to form the classic dsDNA helix and many different RNA structures. This problem is solved to some degree by the presence of counterions that help mask the charge on the negative backbone of the nucleic acids. The presence of phosphate in the phospodiester backbone linkage does confer an important advantage over other possible links (carboxylic acid esters, amides and anhydrides). The electrophilic phosphorous atom is hindered from nucleophilic attack by the negative O attached to the phosphorous. Also, the phosphorous is sp3 hybridized compared to the sp2 hybridization of the electrophilic carbon atom in anhydrides, esters, and amides, and hence is less accessible to nucleophilic attack. Most people now believe that RNA, which can act both as an enzyme and genetic template, preceded DNA as the genetic carrier. The evolution of DNA as the primary genetic carrier required an enzyme to convert ribose to deoxyribose. This would make the nucleic acid less likely to cleave at the phosphodiester bond with the replacement of a nucleophilic 2' OH with an H, and make the genetic molecule more stable. Other types of genetic carriers might have preceded the RNA world, especially if the monomer required could be more readily synthesized from abiological sources. One such alternative are threose nucleic acids (TNA). Synthetics ssTNA can base pair with either RNA, DNA, or itself to form duplexes.

EXTERNAL http://genetics.mgh.harvard.edu/szostakweb/research/protein/research-pro.html



Other possible candidate include peptide nucleic acids (PNA). These can also form double stranded structures with DNA, RNA, or PNA single strands. They were initially designed to bind to dsDNA in the major grove forming a triple-stranded structure. Binding could alter DNA activity, possibly by inhibiting transcription, for example. The structure of a single-stranded PNA is shown. Note that the backbone, a polymer of N-(2-aminoethyl)glycine (AEG) which can be made in prebiotic soups, is not charged, making it easier to bind to dsDNA. AEG polymerizes at 100oC to form the backbone.

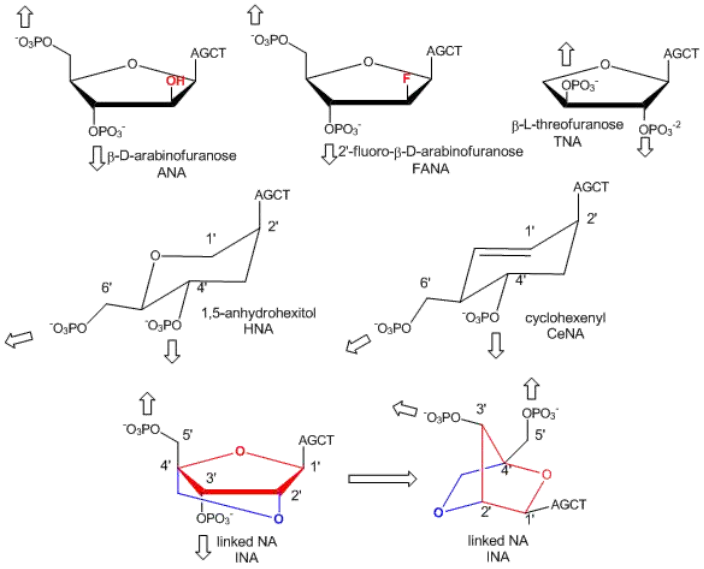


after Gamesh, K. & Nielsen, P. Peptide Nucleic Acids: Analogs and Derivatives. Current Organic Chemistry. 4, 931 (2000)

In addition to changing the backbone, additional bases other than A, C, T, G, and U can be accommodated into dsDNA and ssRNA molecules (Brenner, 2004)

In a recent extension, Pinheiro et al have shown that 6 different foreign backbone architectures can produce xeno-nucleic acids (XNAs) that can be replicated by engineered polymerases which make XNAs from a complementary DNA strand, and a polymerase that can make a complementary copy of DNA from an XNA. XNAs can also be evolved as aptamers to bind specific target molecules. The investigators replaced the deoxyribose and ribose backbone sugar with xenoanalogs (congeners) including 1,5-anhydrohexitol (HNAs), cyclohexene (CeNA), 2'-O,4'-C-methylene-b-D ribose (locked nucleic acids - LNA), L-arabinose (LNA), 2'-fluoro-L-arabinose (FANAs) and threose (TNAs) as shown in the figure below.

Figure: Xeno-nucleic acid sugar congeners



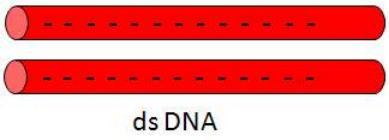
Polymers of these XNA can bind to complementary RNA and DNA and as such act as nuclease-resistant inhibitors of translation and transcription.

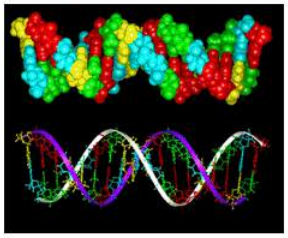
Von Kiedrowski, in an experiment similar to the self-replication of peptides described above, has shown that a single stranded 14 mer DNA strand, when immobilized on a surface, can serve as a template for the binding of complementary 7 mers and their conversion to 14 mers. When released by base, this process can occur with exponential growth of the complementary 14 mers. (von Kiedrowski Nature, 396, Nov 1998). Ferris has shown that if the clay montmorillonite is added to an aqueous solution of diadensosine pyrophosphate, polymerization occurs to produce 10 mers which are 85% linked in a 5' to 3' direction.

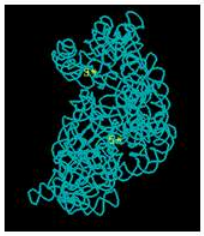
30.1.0.1: Polyanions as Carriers of Genetic Information

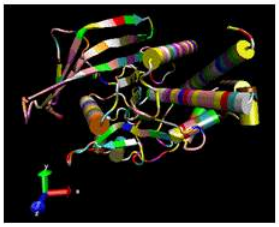
There are other reasons why polyanions are useful genetic molecules, other than their resistance to nucleophilic attack. The biological form of DNA is a large double stranded polyanionic polymer, in contrast to RNA which is a single-stranded polyanion polymer and protein which are polymers with varying combination of anionic, cationic, and hydrophobic properties. Even with counterions, it would be difficult to fold DNA into complicated and compact 3D structures as occurs for proteins, given the large electrostatic repulsions among the charged phosphates. Rather it forms a elongated double stranded rod, not unlike the rod-like structure of proteins denatured with sodium dodecyl sulfate (used in SDS PAGE gels). The elonged rod-shaped structure of ds-DNA is critical for the molecule which is the main carrier of our genetic information since mutations in the bases (leading to a switch in base pairs) causes no change in the overall structure of dsDNA. This enables evolutionary changes in the genetic material to produce new functionalities. A single change an amino acid of a protein, however, can cause a large change in the structure of a whole protein, a feature unacceptable for a carrier of genetic information. RNA structure effectively lies between that of DNA and proteins. Since it has less charge density than dsDNA, it can actually form dsRNA helices, so it can carry genetic information, as well as form complex 3D shapes necessary for its activity as an ribozyme. Perhaps more importantly, steric interference prevents ribose in RNA from adopting the 2'endo conformation, and allows only the 3'endo form, precluding the occurrences of extended ds-B-RNA helices.

What kinds of structures can the biological macromolecules form?



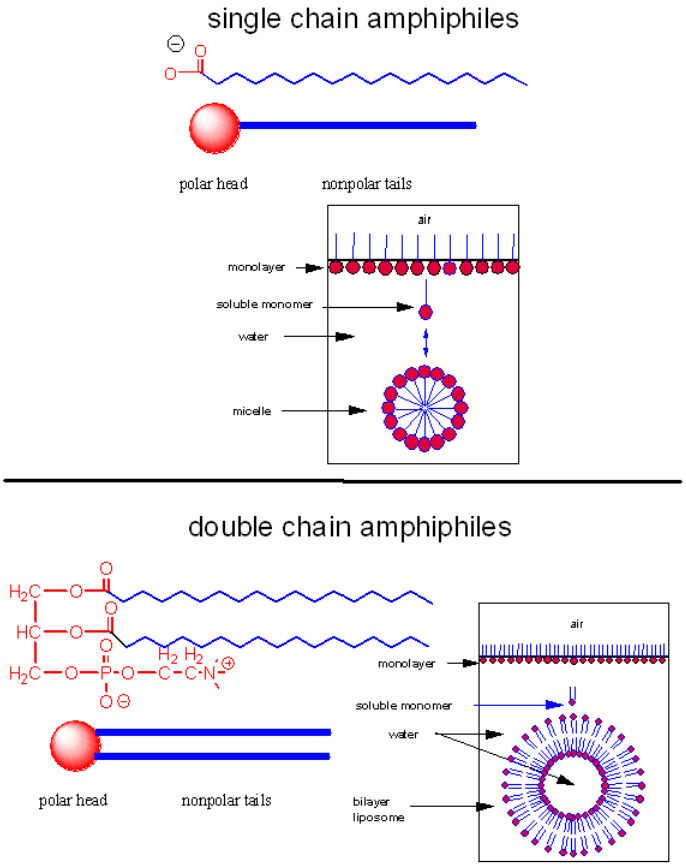




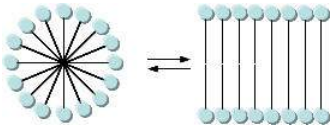


30.1.0.1: The Lipid World

Let's assume that abiological precursors would react to form polymer-like molecules that might be complex enough to fold to structures that would allow binding, catalysis, and rudimentary replication. All this would be worthless unless they could be sequestered in a small volume which would limit diffusion and increase their local concentration. What is required is a membrane structure. Amphiphilic molecules, like lipids, with which we started this book, would be prime candidates since they spontaneously assembly to form micelles and bilayers, as shown in the review diagram below.



As mentioned in Chapter 1, alternative lipid phases are possible. Bilayers can also be formed from single chain amphiphiles, such as certain fatty acids, as illustrated in the equilibrium shown below. This occurs more readily at pH values close to the the pKa of the fatty acid, at which the fatty acids are not all deprotonated with full maximal negative charges. Single chain amphiphiles like fatty acids, which were more likely to formed in abiotic conditions, have been found in meteorites.



Clay surfaces, which have been shown to facilitate the formation of nucleic acids polymers, can also promote the conversion of fatty acid micelles to bilayers (Szostak). One such clay surface, montmorillonite, whose structure is shown below, promotes bilayer formation. It has an empirical formula of $Na_{0.2}Ca_{0.1}Al_2Si_4O_{10}(OH)_2(H_2O)_{10}$.

EXTERNAL Chime Molecule Modeling: montmorillonite EXTERNAL: montmorillonite

The effect of montmorillonite on vesicle formation can be shown by simple measurements of turbidity with time. Microscopy of fluorophore-encapsulated vesicles also shows encapsulated montmorillonite. The fatty acids presumably absorb to the cation layer of the clay particles. Time studies using light scattering also indicate that the vesicles grow in the presence of fatty acid micelles. To differentiate between the formation of new vesicles and the increase in size of pre-existing vesicles (which couldn't be done by simple light scattering without separation of the vesicles), investigators used two different fluorescent molecules to label fatty acid vesicles. The two probes were selected such that if the two probes came in close contact, energy transfer from the excited state of one fluorophore to the other fluorophore could occur, an example of fluorescence resonance energy transfer (FRET). FRET is observed when emission of the second dye occurs after excitation of the first dye, at a wavelength outside of the excitation

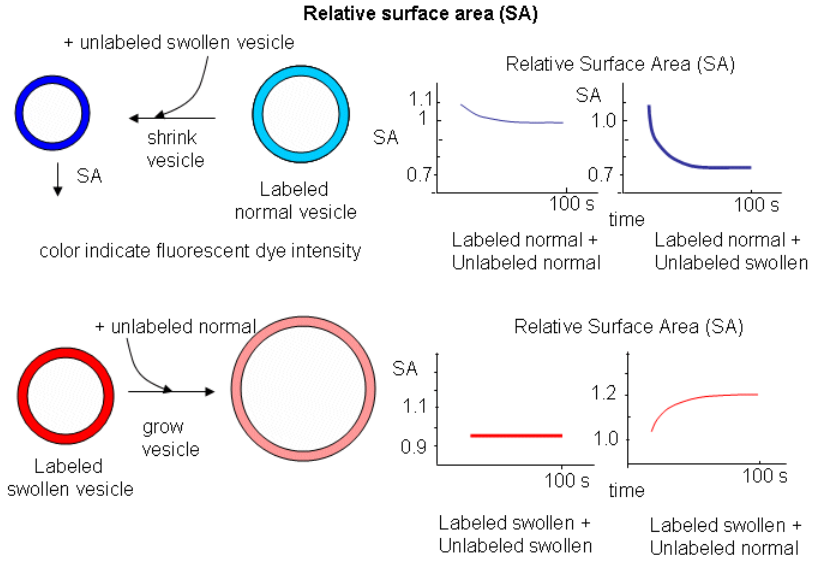
wavelength of the second dye. If unlabeled vesicles were added to either labeled vesicles, no changes in FRET were observed, suggesting that the dyes did not move between vesicles. If fatty acid micelles were added, a decrease in FRET was observed, suggesting that new fatty acids were transferred to the doubly-labeled vesicles, effectively diluting the dye concentrations in the bilayer and their relative proximity, both which would decrease FRET. Most of the new fatty acid was incorporated into pre-existing vesicles which grew. The vesicles could also divide if extruded through a small pore. Later we will see that the energy to grow the vesicles can derive in part from a transmembrane proton concentration collapse. Division of vesicles might be promoted by bilayer assymetries associated with addition of substances to the outer leaflet, causing membrane distortion.

30.1.0.1: Protocells

At some point, early genetic material must have been encapsulated in a membranous vesicles. Would new properties emerge from this mixture that might have a competitive (evolutionary) advantage over either component alone, and thus be a step on the way to the formation of a "living" cell? The answer appears to be yes. Chen et al. have incorporated RNA into fatty acid vesicles with interest effects. They asked the question as whether those vesicles could grow at the expense of vesicles without encapsulated RNA. RNA, with a high charge density and its associated counter ions would create osmotic stress on the vesicles membranes. To relieve that stress they could acquire fatty acids from other fatty acid vesicles (or fatty acid micelles), increasing their surface area, and concomitantly reducing tension in the membrane.

Oleic acids vesicles were first placed under stress by encapsulating 1 M sucrose in the vesicle and then diluting it in hypotonic media. Water would enter and swell the vesicle (but without bursting and resealing, as evident from control experiments). Then they prepared stressed and unstressed oleic acid vesicles in the presence of two nonpolar flurophores, NBD-PE (excitation at 430 nm, emission at 530 nm) and Rh-DHPE (emission at 586 nm). These fluorophores were chosen for fluorescence resonance energy transfer measurements. If the membrane vesicles changed size, the FRET signal would change, based on the relative concentration and proximity of the dual fluorophores. If the separation between probe molecules increased, the FRET signal would decrease. Conversely, if the vesicle shrunk, the FRET signal would increase.

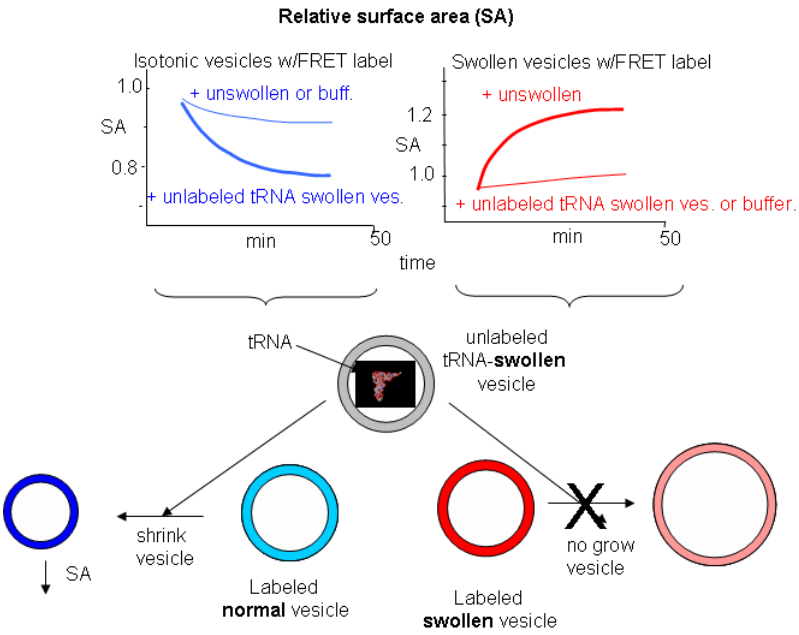
The results showing the effect of adding unlabeled swollen vesicles to labeled normal vesicles, and labeled swollen vesicles to unlabeled normal are shown below. The surface area of normal labeled vesicles decreased by about 25% when unlabeled swollen vesicles were added, but not when unlabeled normal vesicles were added. Labeled swollen vesicles increased 25% in size only if mixed with unlabeled normal vesicles, not with unlabeled swollen vesicles. Hence swollen vesicles win the competition and "steal" lipid from normal vesicles.



After Chen et al. Science 305, 1474 (2004)

Now what about vesicles swollen with encapsulated RNA? RNA, with its associated charge and charged counter ions also placed an osmotic stress on the vesicles. FRET labels (the two fluorophores) were place in vesicles without RNA. Fatty acids were removed from isotonic labeled vesicles in the presence of unlabeled tRNA swollen vesicles (left panel below). Labeled vesicles

swollen with glycerol took fatty acids from unswollen vesicles (without tRNA), but not from tRNA swollen vesicles, as both were swollen so no net drive to reduce swelling by lipid exchange was present.

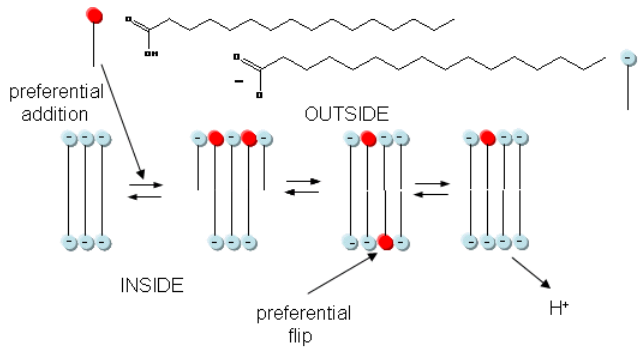


After Chen et al. Science 305, 1474 (2004)

These results show the vesicles with encapsulated RNA have a competitive (evolutionary) advantage over normal vesicles. This data suggests that having a polyanion as the source of genetic material is actually advantageous to the protocell. In addition the move in modern membranes to phospholipids with esterified fatty acids (instead of free ones) may actually have stabilized membranes, given the movement of free fatty acids to different membranes.

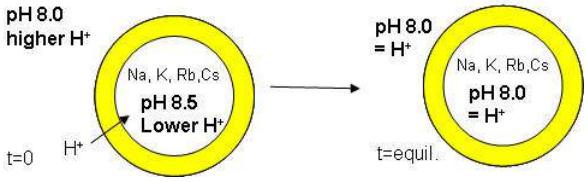
30.1.0.1: Energy Transduction in Protocells

In addition to a genetic macromolecule and a semipermeable membrane, a source of energy to drive intracellular processes must be present. A common source of free energy used in many cells to drive unfavorable processes is a proton gradient, whose formation in modern cells can be coupled to energy input from oxidation, ATP cleavage, light, or the collapse of another gradient. Could a proton gradient be formed in protocells? It can, quite easily, when coupled to the growth of fatty acid vesicles. If a fatty acid vesicle is to grow, more fatty acid must be added to the outer leaflet. The protonated, uncharged form of the fatty acid would preferentially be added, since it would lead to less electrostatic repulsion between adjacent head groups. The protonated, uncharged form of the fatty acid would also be most likely to flip to the inner leaflet to minimize stress asymmetries in the leaflets. Once in the inner leaflet, it could deprotonate to form H+(aq) in the inside of the membrane, creating a transmembrane proton gradient and transmembrane potential. The energy released on growth of the membrane is partly captured in the formation of a proton gradient, as shown in the figure below.



Generation of proton gradient on growth of fatty acid vesicle

The proton gradient would soon inhibit its own formation since further movement of protons into the cell would be attenuated by the positive transmembrane potential unless metal ions inside moved outside. In addition, the gradient would collapse after growth stopped. The investigators made fatty acids vesicles in the presence of pH 8.5 buffers whose pH was adjusted with an alkali metal hydroxide. The external pH was reduced to 8.0, resulting in a 0.5 pH unit proton concentration gradient. (Changes in intravesicular pH were measured with pH-sensitive fluorophore, HPTS.) Inward movement of protons down a concentration gradient, as shown in the figure below, would occur with time, collapsing the imposed concentration gradient.



With fatty acid vesicles, this artificial pH gradient collapsed quickly, suggesting the vesicle permeability to protons was high. The rate was too high for simple flip-flop diffusion. Inward movement of protons appeared to be facilitated by outward movement of the M+ ions. The rate of decay of the proton gradient was exponential, and the resulting first order rate constant was easily determined. A graph of the rate constant for pH gradient collapse vs unsolvated ionic radius of M+ decreased with increasing radius (i.e. kNa > kK > KRb > KCs, suggesting that the pH gradient would be more stable if large, impermeable or otherwise trapped cations were encapsulated. When vesicles were made with encapsulated Arg+, the imposed pH gradient did not collapse for hours. If oleic acid micelles were added to oleic acid vesicles with encapsulated Arg+, with no artificial pH gradient induced across the membrane, the vesicle grew with concomitant movement of protons into the vesicle, producing a pH gradient of 0.3 within seconds.

These experiments show that membrane growth and energy storage could be coupled, and the right composition of encapsulated material could lead to a stable transmembrane pH gradient, a source of energy to drive biological processes. It even suggests that a charge polyanion would be beneficial as a genetic carrier.

30.1.0.1: Hydrothermal Vents or Primoridal Soup?

The case for the origin of life in deep sea hydrothermal vents and not in a primordial "Campbell's" soup has been argued convincingly by Lane et al (2010). What's needed for life are reasonably complex molecules and an energy source to drive unfavorable reactions. It's the latter on which that Lane et al focus. In an early primordial world that was low in oxygen, exergonic oxidation reaction of organic molecules would provide little energy. This can be surmised from the low energy yield (compared to aerobic respiration) achieved in present day glycolytic (fermentative) pathways from all major domains of life, archaea, bacteria, and eukaryotes.

> Background: Based on rRNA sequences, a primordial cell evolved into two different types of cells, one that became bacteria, and another that split further into archaea (single cells, similar to bacteria) and eukaryotic cells (complex cells with internal organelles that eventually formed multicellular organisms). Bacterial and archaea are collected called prokaryotic cells.

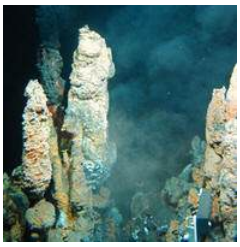
In addition, these pathways required the evolution of up to dozen different enzymes to produce their relatively meager energy yield which ultimately depends on the oxidation of an organic molecule by another organic molecule instead of by a powerful oxidant like dioxygen. An anisotropic arrangement of molecules in a concentrated soup could lead to transient chemical potential fluctuations but these would be inefficient and impermanent sources of energy. Effectively the primordial soup would be at equilibrium and hardly expected to provide the energy for synthesis of RNA enzymes and replicators. UV light leads to photo-damage and photolysis not replication of complex molecules. What is needed is a way to drive the synthesis of molecule with high chemical potential energy (like sulfur esters and phosphoanhydrides) compared to their lytic products. These could then provide an energy sources to drive ATP synthesis, for example.

A detailed look a the bioworld shows that the earliest organisms used energy from the collapse of the proton gradient (chemisomotic principle elucidated by Peter Mitchell). All present autotrophs (organisms that can fix CO2 and form complex organic molecules) and many heterotrophs (use complex organic molecules of other organisms for fuel) use redox complexes in membranes coupled to membrane gradients. These complexes would take reduced molecules and pass electrons from them to oxidizing agents (electron acceptors), including O2, CO2, and Fe3+ to form H2O, CH4, and Fe2+. Fermentors also use ATPase membrane enzymes to transport nutrients. Yet genomic analysis of bacteria and archaea show that enzymes involved in

fermentation differ significantly, suggesting that they evolved separately towards a convergent function. Structure in common include DNA, RNA, ribosomes and membrane ATPases, which Lane et al suggest were in a the Last Universal Common Cell (LUCA).

All autotrophs produce their energy source by fixing CO2 using either H2 directly or indirectly using H2O and H2S. All of the are available in nonhydrothermal deep sea vents. Volcanic vents, however, are extremely hot (not optimal for organic molecule synthesis), very acidic, and lack hydrogen gas. A different type of nonvolcanic vent, an alkaline hydrothermal one, might produce more conducive as a site of the origin of life. In these vents, water chemically reacts with minerals in the crust (such as olivine) leading to their hydroxylation and subsequent fracture, with promotes more water entry into the crust. It has been reported that there is more water found as hydroxylated minerals in the crust, that there is liquid water in the oceans. These processes result in temperatures up to 200 degrees Celsius and release of hydrogen gas into a moderately alkaline vents into the sea water at temperatures more conducive (70 degrees C) to the origin of life.

Figure: Alkaline Vent



30.1.0.1: Fixing CO_2

Of the five different pathways known to fix CO_2, all require ATP except one. That one is present in both methanogens, which produce methane from CO_2 and H_2, and in the acetogens, which produce acetate ($CH_3CO_2^-$) in the form of acetyl-CoA. The simpler reactions of forming acetic acid and methane are shown below:

$2CO_2 + 4H_2 \rightarrow CH_3CO_2H + 2H_2O$.

$CO_2 + 4H_2 \rightarrow CH_4 + 2H_2O$.

The ΔG^0 values for these reaction (calculated using ΔG^0_f for gas phase H_2, CO_2 and CH_4 and liquid acetic acid and water are -75 and -131 kJ, respectively, at 250C, showing that they are thermodynamically favored. Making AcetylCoA, a "high" energy molecule compared to its hydrolysis products (as is ATP) from acetic acid and CoASH, requires an additional source of energy to drive the reaction. A proton gradient is the likely source.

Some bacteria and Achaea cells (primordial or present) use the reductive acetyl-CoA pathway, also known as the Wood-Ljungdahl pathway, to form, in a noncyclic process, acetyl CoA from CO_2 and at the same time makes ATP. This process is paid for by a proton gradient. This has been described by Shock as "a free lunch you get paid to eat". The energetics of the present acetyl CoA pathway based on the overall reaction below show an approximate ΔG^0 value of -59 kJ/mol which can drive ATP synthesis.

$2CO_2 + 4H_2 + CoASH \rightarrow CH_3COSCoA + 3H_2O$.

The concentration of carbon dioxide in the primordial ocean was 1000 times higher than now. Vents produced large amounts of methane and hydrogen gas. There was little oxygen and hence lots of Fe^{2+}. The enzymes involved in this acetyl-CoA pathway of carbon fixation have FeS clusters. It has also been shown that bubbles (which are really membrane-bound spaces) of FeS and NiS can be made in deep-sea vents. These could not only encapsulate precursor molecules but also serve as catalysts. Vents also can catalyze the fixation of nitrogen (to ammonia) and laboratory studies show that FeS can catalyze the conversion of formate (found in vents) into pyrimidines and purines. The studies of present methanogens and vent chemistry suggest that the critical ingredients and conditions for the development of the first biological cells probably occurred in the vents.

To produce polymers, an energy source and monomers must exist. Concentration gradients found in simulations of vents produce million-fold concentrated molecules. The transient heating and cooling of any double-stranded nucleic acids could lead to concentration amplification by a PCR-like strand separation followed by reannealing. In addition, these vent regions possess a powerful, reoccurring energy source, a pH gradient, as the alkaline vented material entered the acidic oceans that exist with high CO2 concentrations, creating a gradient across an inorganic membrane. This is startlingly analogous to the pH gradient across membranes (acidic outside, alkaline inside) driven by the membrane complexes in the mitochondria and bacteria. Lane et al argue

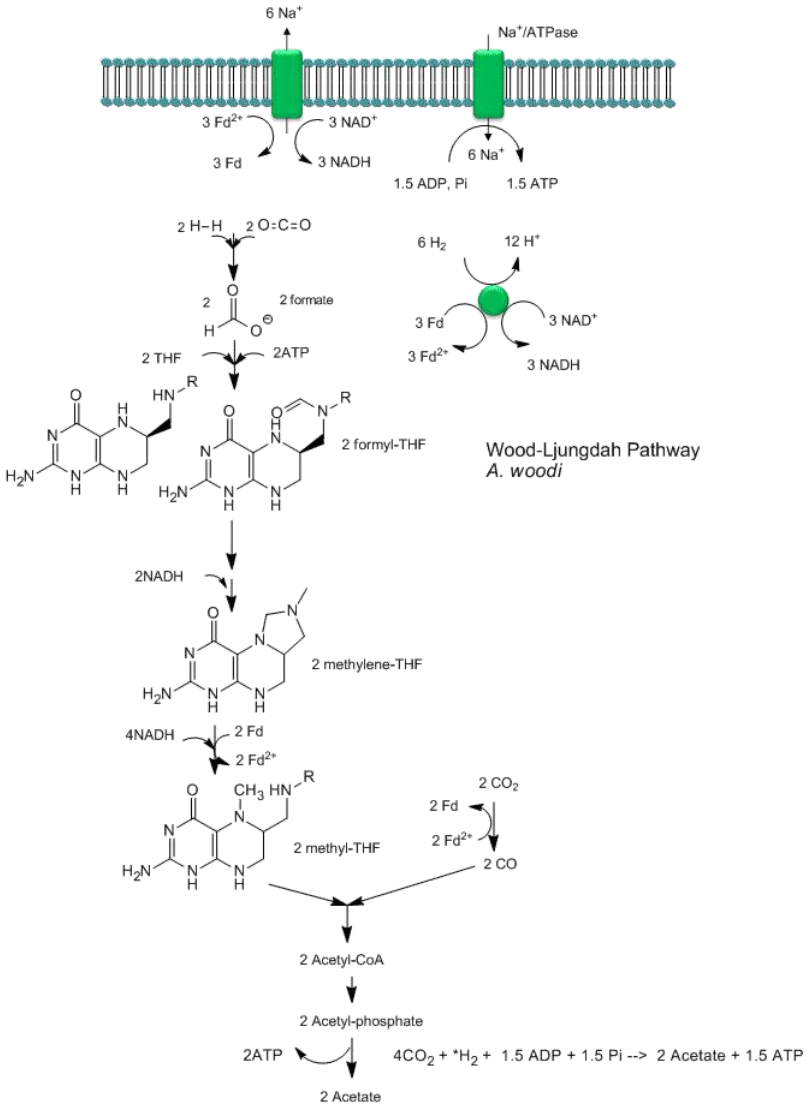
that the existence of membrane proton gradients as an energy source in all cells (eukaryotes, bacteria, and archaea) and in chloroplasts and mitochondria, corroborate their hypothesis. Bacteria and Archaea share homologous ATPases and electron carriers (ferredoxins, quinones, and cytochromes). These similarities contrast to the differences in enzyme structures in fermentative pathways. Arguments that proton pumps evolved to pump proteins (and reduce pH gradients) can't explain their ubiquitous presence even in organisms not subjected to low pH. Hence the ubiquity of proton pumps supports the conjecture that they arose from the first protocells, possibly comprised of inorganic walls and ultimately with amphiphilic molecules synthesized from precursors.

Creationists would argue that it would be impossible to evolve a structure with the complexity of membrane ATPase (which collapse a pH gradient as the power the synthesis of molecules with large negative ΔG^0 of hydrolysis). Lane et al propose that the earliest cells evolved ATPase-like molecules in alkaline vents where pH gradients analogous to those in cells today arose. They envision cell-like columns lined by FeS membrane-like structures with alkaline conditions inside and acid conditions outside. Nonpolar or amphiphilic molecules would line the inside of the cells/columns. An ATPase-like system could then take advantage of the pH gradient which constantly replenishes itself. If structures as complicated as ribosomes evolved from a subsequent RNA world, surely ATPase-like molecules could also. Other chemistry might have evolved earlier to utilize the energy source provided by the pH gradient.

If life originated in the vents, it would need an energy source to leave the vents. Presumably it would have evolved one to utilized pH gradient to replace the one it left in the alkaline vents. The substrate level phosphorylation of glycolysis that requires ATP input to make ATP would not provide the energy source needed. Cells that left would have had to produce their own proton gradient. Perhaps all the was needed initially was concerted conformational changes in proteins that upon exposure of a different pH changed their shape inducing pKa shifts in adjacent proton donors/acceptors leading to vectorial discharge of protons across a membrane. Perhaps the method described above in protocells was sufficient.

Recent analyses by Poehlein et a show that CO2 reduction (fixation) can be coupled to the production of a sodium ion gradient, which could collapse to drive ATP synthesis. Analysis of the genome of a gram positive bacteria, Acetobacterium woodii, an acetogen, shows the it has an ancient pathway for production of acetyl-CoA that can, in an anabolic fashion form biomass or in a catabolic fashion be cleaved to acetate with the production of ATP. It does not require classic electron carriers like ubiquinone or cytochrome C linked to protein gradient formation to drive ATP synthesis. Rather it has only a ferredoxin:NAD+ oxioreductase which couples oxidation to the formation of a sodium ion gradient, which collapses through an sodium ion transporter/ATP synthase to drive ATP synthase. A plausible reaction scheme based on genomic analysis is shown below:

Figure: Acetyl-CoA Synthase and Acetogenesis



30.1.0.1: The Role of Fe/S Centers

Let's return to the chicken and egg dilemma one more time. What is needed for biological polymer formation are monomeric precursors, an energy source, and a way to compartmentalized them all. We discuss how monomeric precursors could form, but wouldn't it be far better if even the synthesis of precursors could be catalyzed? One source of catalysis mostly absent from the "bioorganic" abiotic chemistry in the above discussion is the transition metals. Transition metals can form complexes. Ligands containing lone pairs on O, N, and S atoms can donate them to transition metals ions, which can hold up to 18 electrons in s, p, and d orbitals. Hence as many as 9 lone pairs on ligand molecules (which are often multidentate) could be accommodated around the transition metal ion. Many present small molecule metabolites and their abiotic precursors (H2O, CO, CO2, NH3 and thiols) bind cations as mono- or polydentate donors of electrons. Hence transition metal ions would have a thermodynamic tendencies to be bound in complexes.

Bound ligands that contain potentially ionizable hydrogens could become deprotonated and made better nucleophiles for reactions. Hence the transition state metal ion, acting with the complex, becomes a catalyst as it decreases the pKa of a bound ligand (such as water). In addition, since transition metals ions can have multiple charge and oxidation states, they can easily act as redox centers in the oxidation/reduction of bound ligands that were redox active. Given the relative anoxic conditions of the early oceans, Fe2+ would predominant. It could easily be oxidized to Fe3+ as it reduced a bound ligand. Highly charged transition states would withdraw electron density from bound ligands leading to their possible oxidation.

Metals obviously still play a strong role in catalysis, both indirectly in promoting correct protein folding and directly in stabilizing charge in both the transition state and intermediates in chemical reaction pathways. FeS clusters are of significant importance.

Their biosynthesis involves removal by an active site Cys in a desulfurase enzyme of a sulfur from a free amino acid Cys followed by its transfer to an Fe in a growing FeS cluster in a FeS scaffold protein, which then transfers the cluster to an acceptor protein where it acts as a cofactor. FeS clusters can adopt a variety of stoichiometries and shapes, as well as redox states for the participating Fe ions. The continuing importance of FeS clusters in all cells, their involvement in not only redox enzymes in which electron transfer is facilitated by delocalization of electrons over both Fe and S centers, but also in coupled electron/proton transport in mitochondrial electron transport, Fe storage (ferrodoxins), and in regulation of enzyme activity and gene expression, suggests that they were of primordial importance in the evolution of life. T

hey are often found at substrate binding sites of FeS enzymes involved in both redox and nonredox catalysis. A ligand can bind to a particular Fe in the cluster, activating it for hydration or dehydrogenation reactions. Fe 4 of the FeS cluster in the TCA enzyme aconitase can have a coordination numbers of 4, 5, or 6 as it binds water, hydroxide or substrate. It acts to both decrease electron density in the transition state and to change the pKa of bound water as the enzyme catalyzes an isomerization of tricarboxylic acids (citric and isocitric acid) through an elimination/addition reaction with water. In another example it can bind S-adenosylmethionine through its amine and carboxylate groups, which activates the molecule for cleavage and radical formation. In some cases metals other than Fe (Ni for example) are incorporated into the cluster. FeS effects on transcription factors involves facilitation of optimal structure for DNA binding. FeS and FeNi centers in proteins are similar in structure tp FeS units in minerals like greigite and presumably to FeS structure formed when H2S and S2- react with Fe2+ (present in abundance in the early ocean) and other metals in vents Metal sulfides participate in reduction of both CO and CO2. For example the synthesis of CH3SH from CO2 and H2S is catalyzed by "inorganic" FeS.

30.1.0.1: The Minimal Genome

This question is being addressed by eliminating "unnecessary" gene from simple bacteria. Cells placed in a rich nutrient broth with essential lipids, vitamins, and amino acids would need fewer genes than those placed in a more nutrient-poor medium. Bacteria cells like Mycoplasma genetalium, that live within "nutrient rich" eukaryotic cell, have been genetically manipulated to delete unnecessary genes. Based on knockout studies, it may be possible for the cell to survive with only 300-350 genes. Bacillus subtilis has approximately 4100 genes. Estimates have been made that it could survive with as few as 271 genes.

more to be added

CHAPTER OVERVIEW

31: Quantum Biochemistry - TBA

31.1: Quantum Biochemistry

31.1: Quantum Biochemistry

CHAPTER OVERVIEW

32: BIOCHEMISTRY AND CLIMATE CHANGE

32.01A: The Basics of Climate Change
32.01B: Back to the Present and Future of Climate Change
32.2: Use of Isotope Analysis in Measuring Climate Change
32.3: Climate Change - The Carbon Cycle and Carbon Chemistry
32.4: Biofuels A - Corn and Sugar Cane Ethanol
32.5: Biofuels B - Cellulosic Ethanol
32.6: Algae - an Introduction
32.7: Algae - Bioethanol production
32.8: Biodiesel, Syngas and Bioaviation fuels
32.09: Biohydrogen - An Introduction
32.10: Biohydrogen - Hydrogenases
32.11: A Warmer World: Temperature Effects On Chemical Reactions
32.12: A Warmer World: Temperature Effects On Proteins
32.13: Biochemistry, Climate Change and Human Health
32.14: Climate Change, Infectious Disease and Pandemics
32.15: Pandemic Diseases and Drug Discovery - Under Construction
32.16: Fixing Carbon Fixation
32.17: Fixing Nitrogen Fixation
32.18: Turning Trees into Plexiglass - Synthetic Biology For Production of Green Foods and Products

32.01A: THE BASICS OF CLIMATE CHANGE

Search Fundamentals of Biochemistry

LEARNING OBJECTIVES

- to demonstrate how climate has changed over geological time through the present
- to explain mechanisms, using knowledge from biology, chemistry, and physics, for climate change
- to show the central role of atmospheric CO_2 as a causative agent of past and present climate change
- to contrast the effects of anthropogenic burning of fossil fuels on climate change with causes of past climate changes
- to address arguments made by climate change skeptics

INTRODUCTION

We've known for a very long time that burning fossil fuels and releasing CO_2 into the atmosphere would warm our climate. Perhaps the first paper addressing this, Circumstances affecting the Heat of the Sun's Rays, was published in 1856, before the US Civil War, by a woman scientist, Eunice Foote. John Tyndall (of the Tyndall effect) published more comprehensively on greenhouse gases in 1859. Given the complexity of the biosphere's climate, it was not until the 1980s that climate models became sophisticated enough for scientists like James Hansen to become convinced and alarmed enough to discuss in Congressional hearings the role of anthropogenic (made by humans) CO_2 released into the atmosphere as a cause of ever-worsening global warming. The knowledge of human-induced climate change has been politicized and subjected to an orchestrated campaign of misinformation and disinformation by fossil fuel companies and their political contributors. We have delayed global actions for so long that we must act immediately and aggressively to address climate change before we reach climatic conditions that are so austere for humans that parts of the world become uninhabitable. Homo sapiens evolved in a world dominated by repetitive glaciation and deglaciation. Hans Joachim Schellnhuber, an atmospheric physicist, climatologist, and founding director of the Potsdam Institute for Climate Impact Research, has stated that humans have so affected the world that we have eliminated the possibility of the next glaciation cycle.

Many readers might not be familiar with the data and models supporting human-caused climate change and that climate scientists are almost unanimous in their support of the data and models. As in any field, however, you will find outliers who don't and whose ideas carry disproportionate weight among climate change skeptics. Hence, we provide the basic data to show the relationship between increasing atmospheric CO_2 to global warming, both drought and flooding, ocean acidification, and loss of biodiversity. We also provide supportive information that would allow users to address questions from those who question the reality of present human-induced climate change. We don't shy away from using basic physics as well since most students studying biochemistry at the level found in this book have also studied physics as well as biology. In subsequent sections, we will then address the biochemistry of climate change and its mitigation.

GREEN HOUSE EFFECT

Before the advent of the Industrial Revolution, the Earth's climate was fairly constant since the last ice age, which peaked about 22,000 years ago (YA) and ended about 12,000 YA. There have been short (on a geological time scale) periods of cooling since the end of the last ice age. Humans evolved around 200,000 years ago with modern civilizations arising about 4000 BCE so it could be said that humans are ice-age peoples (a distinctly Northern Hemisphere perspective). Humans have had the benefits of a fairly stable climate since then.

The sun's energy warms the Earth. If the Earth did not radiate back into space an equivalent amount of energy, it would slowly and continually warm. The Earth reflects energy back in the form of light. In addition, as the Earth is heated by the sun, the Earth releases heat in the form of infrared light (as do all warm objects). Earth's temperatures are stable when the sum of the energy emitted by the Earth equals the energy it receives from the sun. This is illustrated in Figure 32.1.1.

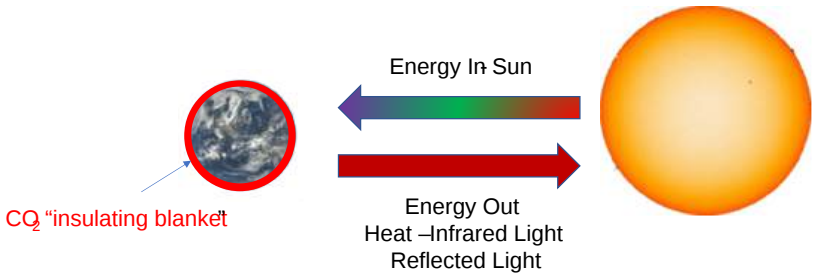


Figure 32.1.1: Energy balance for Earth as it adsorbs and emits solar radiation

Our stable climate has been enabled by fairly constant levels of atmospheric CO_2, a trace atmospheric gas, which has hovered around 280 parts per million (ppm) until the start of the Industrial Revolution in 1770. CO_2 is a greenhouse gas, which as anyone who has run an IR spectra knows, absorbs in the infrared. CO_2 in the atmosphere absorbs some of the infrared radiation released by the Earth, allowing the Earth to be warmer than in its absence. The CO_2 effectively acts as an insulating blanket. In fact, without CO_2 or other "greenhouse" gases, the Earth would be completely covered by snow.

Other greenhouse gases in the atmosphere include methane and nitrous oxide. The IR spectra of these gases are shown in Figure 32.1.2. Students who have taken organic chemistry labs and obtained IR spectra of samples always blank the instrument to remove spectral signals from both CO_2 and H_2O.

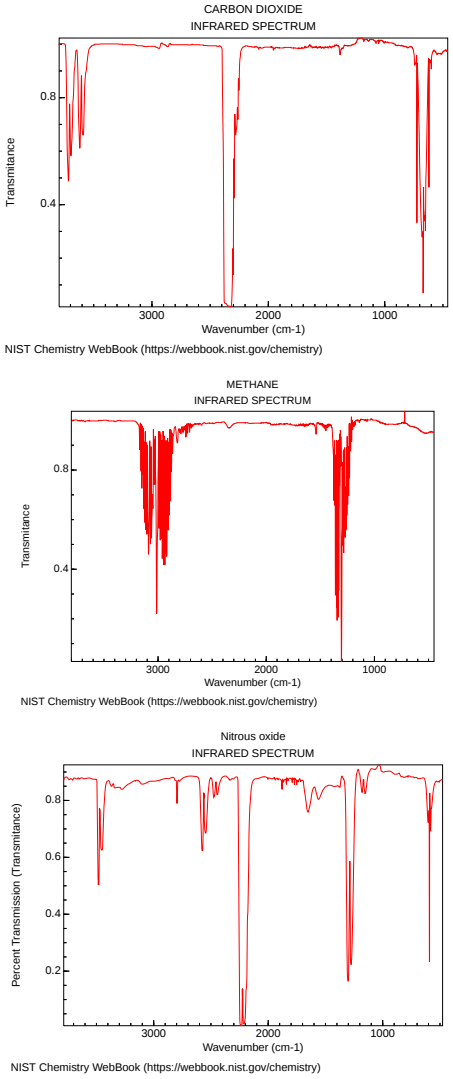


Figure 32.1.2: IR spectra of some greenhouse gases. NIST (ex: https://webbook.nist.gov/cgi/cbook.c...ndex=1#IR-SPEC)

Since the start of the Industrial Revolution, humans have been releasing into the atmosphere ever-increasing amounts of CO_2 from the burning of fossil fuels and methane from agricultural practices and natural gas production. CO_2 levels, as of November 2022 have reached 415 ppm, while methane has increased to 1900 part per billion (ppb) or 1.9 ppm. Increasing methane in the atmosphere contributes about 20% of the global warming effect of the more concentrated CO_2, given methane's intense IR absorbance spectra. It has a short half-life in the atmosphere (about 20 years) compared to that of CO_2 (hundreds of years). Nitrous oxide (N_2O) is also a powerful greenhouse gas, which also depletes ozone in the stratosphere. Our increased use of synthetic fertilizers and manure is the primary anthropogenic source of N_2O. Its emission is accelerated from poorly drained farmlands.

The net effect is that the Earth is no longer in energy balance. The net contributions of energy input and output are expressed as energy/sec/m^2 of the Earth's surface area. Hence the units are Joule/s·m^2 = Watts/m^2, where 1 Watt is 1 J/s. From around 2005 to 2019, the average **Earth Energy Imbalance (EEI)** (or budget) was about +460 x 10^{12} W or 460 terrawatts (TW) = 4.6 x 10^{14} W. Dividing this by the surface area of the planet (5.1 x 10^{14} m^2), gives an average global value of +0.90 ± 0.15 W/m^2 (or J/(m^2s)). Newer and more refined estimates show it to be +1 W/m^2 with most of the excess energy stored in the Earth's Oceans, which have a high specific heat capacity and

cover 71% of the surface area of the planet. Most of the rest of the excess energy is used to warm land masses and melt ice shields and glaciers. 1 W/m^2 might not seem like a lot but it is (pardon the pun), astronomically big.

The atomic bomb dropped on Hiroshima released about 1.8 x 10^{13} Joules (some estimates are as high as 9 x 10^{13} Joules) of energy. Assume for the sake of calculation that this energy was released in 1 second over the entire surface of the planet (5.1 x 10^{14} m^2). The earth energy imbalance for that 1 explosion was 0.035 J/m^2·s (0.176 J/m^2·s for the higher estimate). The ratios of the present imbalance (+1 W/m^2) to the distributed bomb's imbalances are 1/0.035 = 28x for the lower energy estimate bomb or 1/0.176 = 5.7 for the higher energy estimate. Hence the present imbalance, 1 J/m^{2}s from human-caused climate change, is equivalent to detonating from 6-28 Hiroshima-type atomic bombs every second. Other estimates are lower and closer to 600 bombs per **day** but the bases for the calculation are not readily available. Regardless, these values should give you a dramatic sense of the imbalance and how unsustainable it is.

Figure 32.1.3 below shows a graphical representation of the Earth's energy budget.

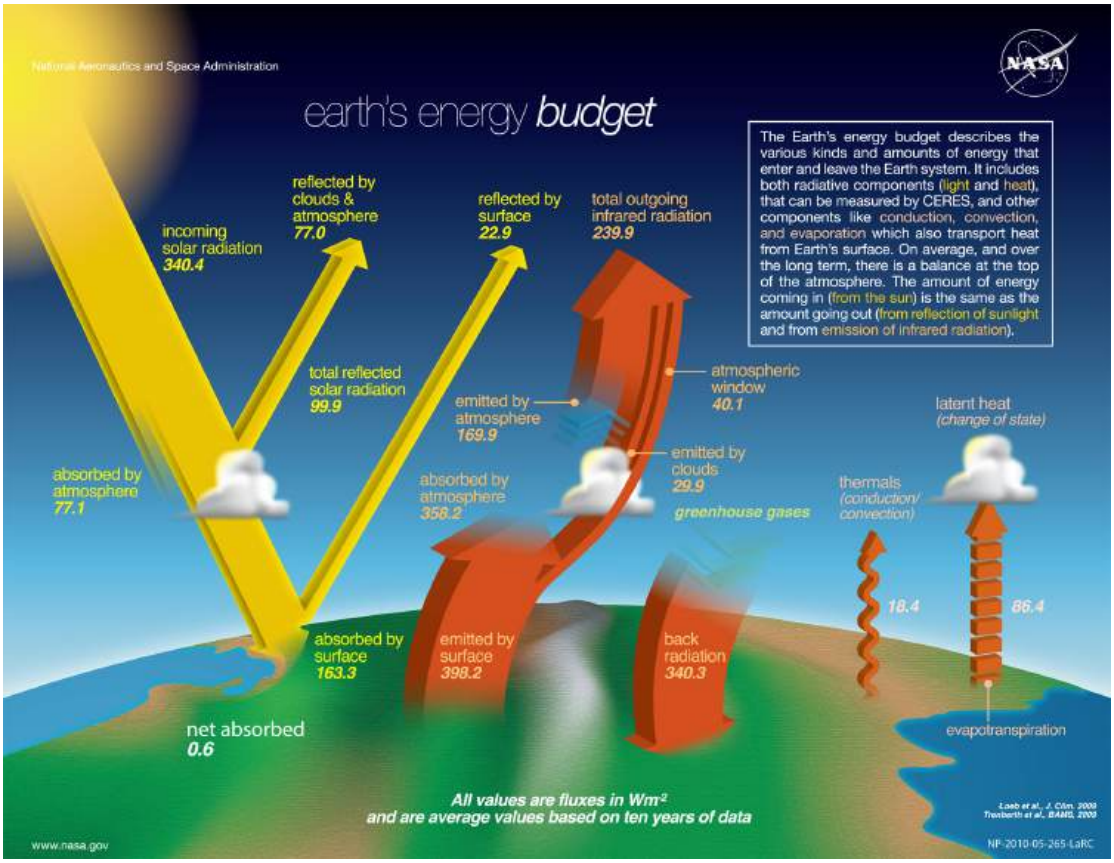


Figure 32.1.3: Earth's energy budget, with incoming and outgoing radiation (Values are shown in W/m^2). Satellite instruments (CERES) measure the reflected solar and emitted infrared radiation fluxes. **Year 2021 update**: Net absorbed energy (shown as 0.6) rose to above 1.0 W/m^2 based on independent CERES and ocean heat content measurements.(see fig.1 in Loeb et al.(2021), Geoph. Res. Let 48 (13), doi:10.1029/2021GL093047). Public Domain

GLOBAL WARMING POTENTIAL

The **global warming potential (GWP)** is used to calculate the total contribution of all emitted greenhouse gases. It is expressed in units of CO_2 equivalents. It adds the contribution of other greenhouse gases like CH_4 and nitrous oxide (N_2O), each of which has unique IR absorption spectra (as shown in Figure 2 above) and atmospheric half-lives. The IPCC uses a 100-year time frame for the calculation of the GWP, which is often abbreviated as GWP100, and uses this formula:

$$\text{CO}_2 \text{ equivalent kg} = \text{CO}_2 \text{ kg} + (\text{CH}_4 \text{ kg} \times 28) + (\text{N}_2\text{O}kg \times 265) \quad (32.1.1)$$

- CO_2 has GWP of 1 by definition since it is the reference. Its time frame in the atmosphere (100s to 1000 years) doesn't matter since it is the reference.
- CH_4 has a GWP of around 27-30 over 100 years. It reflects its higher IR absorbance but lower life-time (around 12 years).
- N_2O has a GWP of around 265-273 over a 100-year timescale. N_2O has a life-time of around 109 years.

Water is also a greenhouse gas as you can attest to on humid days and how its lack in the atmosphere in deserts leads to a large temperature drop at night. It's very different than other greenhouse gases. Its concentration varies enormously (from 40 ppm to 40000 or more) based on humidity and precipitation events, which remove it from the atmosphere. The amount of water in the atmosphere increases with increasing global temperatures, which gives rise to more intense precipitation events and also to warmer temperatures in a positive feedback loop. Its concentration in the atmosphere hence changes enormously on the time scale of hours and days, so its half-life in the atmosphere is short. In contrast, the half-life of CO_2 in the atmosphere is measured in decades to centuries.

CLIMATE CHANGES OVER THE LAST MILLION YEARS

THE SKEPTIC'S CORNER: CLIMATE CHANGE MISINFORMATION

Climate has always changed. Our present period is no different, so there is no need for action.

Indeed, the Earth has been subject to cycles of glaciation and deglaciation for hundreds of thousands of years. Luckily, we are able to determine atmospheric levels of CO_2 dating back to hundreds of thousands of years ago by measuring entrapped CO_2 in ice cores from Antarctica and Greenland. In addition, we've been able to infer the temperature over this time frame using proxies for temperature (tree rings, fossils, and more as described in section 31.2). Figure 32.1.4 shows how atmospheric CO_2 and temperature have varied over the last 800,000 years using ice core data.

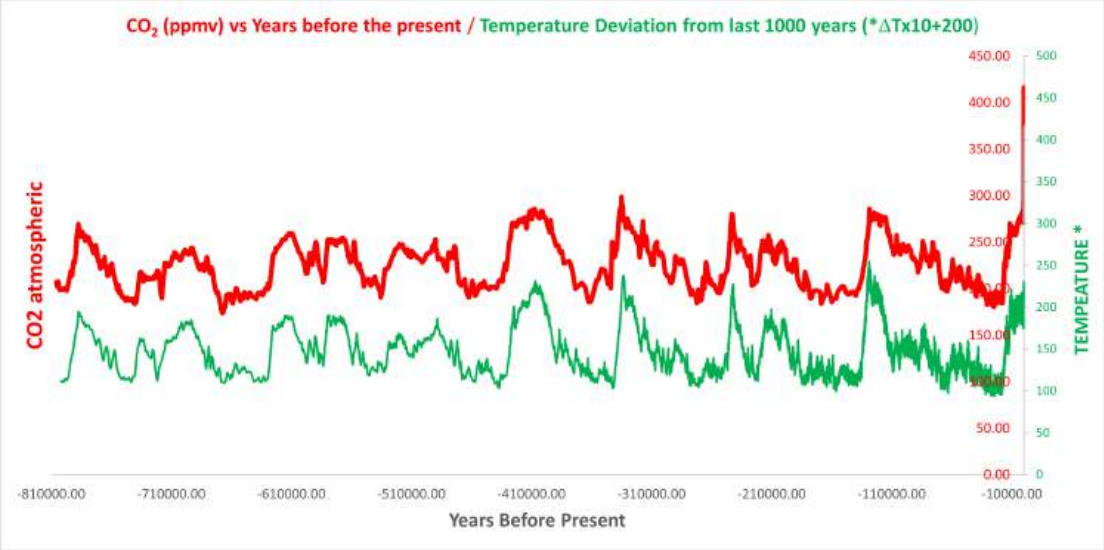
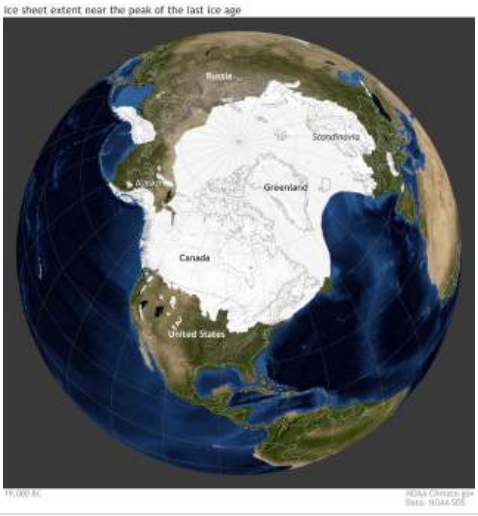


Figure 32.1.4: Atmospheric CO2 concentrations (ppm, red) and temperature deviations from the past 1000 years (green) using ice core date and temperature proxies.

Several key features of the graph should be apparent:

- Both atmospheric CO_2 and temperature change (ΔT) are periodic. So yes, it is obviously true that "climate changes" as climate change skeptics argue
- Both CO_2 and ΔT change in synchrony. An obvious question might be what changes first. Does ΔT drive CO_2 changes or vice versa? More on that in a bit.
- **The CO_2 levels in more modern times (right hand side of the graph) have soared in ways not seen in the last 800,000 years**! This change is caused by CO_2 emissions from the burning of fossil fuels.

In those 800K years, the Earth has experience cycles of glaciation/deglaciation with recurring ice ages. Figure (\PageIndex{5}\) below shows a depiction of the last ice age which peaked 21,000 years ago (left). At that time, the ice cap over New York City was about 1 mile high (right) as CO_2 was at 185 ppm!



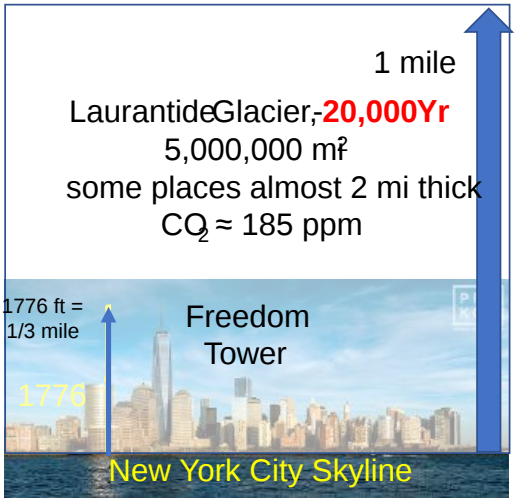


Figure 32.1.5: The Laurentide glacier during the last ice age (left) and its peak height over present-day Manhattan. https://www.climate.gov/news-feature...oming-sunlight

By 5000 BCE, the glacier had retreated to more modern levels, leaving ice over the Arctic Ocean, and over Greenland. CO_2 levels were then around 260 ppm. A change of just 100 ppm in CO_2 was sufficient to lead to the melting of the Northern Hemisphere glaciers. The image above is not "Northern Hemisphere-Centric" since the great glaciers were localized in the Northern Hemisphere in the ice ages. That's because glaciers grow over land and most of the land on the planet is in the Northern Hemisphere. (Our climate studies won't include the time when one continent - Pangea- existed.) The video below shows an animation of the Northern Hemisphere ice shield as it changed with time from 19,000 BCE to now to a projected future that assumes little action to change CO_2 emissions. Pay special attention to the graphs which show sea level changes as well.

Best estimates by Tierney et al now show that during the last ice age, the average global temperature was 6 degrees Celsius (11 F) cooler than today, which in the 20th century is 14 C (57 F). The Arctic however was much colder (about 14 C or 25 F). The group also came up with an estimate of **climate sensitivity**, the increase in temperature with increasing CO_2. That value is a **rise of 3.4 °C (6.1 °F) for a doubling in CO_2**. In 1896, Arrhenius, recognizing that CO_2 was a greenhouse gas, actually calculated that doubling atmospheric CO_2 would cause a rise of 4-5 °C. No one can say we haven't known!

A new value for climate sensitivity giving a **rise of 4.8 °C (8.6 °F) for a doubling of CO_2** was determined by Hansen et al (2023) using more precise data about temperatures and CO_2 concentrations derived from isotopic analyses of ice core samples. If this value holds, the environment needed for human life and our present society is in serious jeopardy unless we take immediate measures to stop emissions and actually cool the planet. In addition, the rate of increase in global and ocean temperatures has accelerated after 2010 and the likely culprit is an actual decrease in toxic SO_2 aerosols released by ocean vessels in the North Pacific and Atlantic Oceans and in China from the reduction of dirty coal-burning facilities. Aerosols reflect incident solar light and also increase cloud formation, both of which help cool the planet. Paradoxically, the reduction of anthropogenic aerosols, while great for cardiovascular and pulmonary health, will accelerate global warming. India would be much warmer now if not for the high aerosol concentrations from smog that covers many of its large cities and countryside.

WHAT THE SCIENCE SHOWS

Climate, CO_2 and temperature have always changed over geological time, but our present rise in anthropogenic CO_2 in such a brief time is unprecedented and has led to CO_2 levels that far exceed those during the warmer interglacial periods when Northern Hemisphere glaciers had retreated.

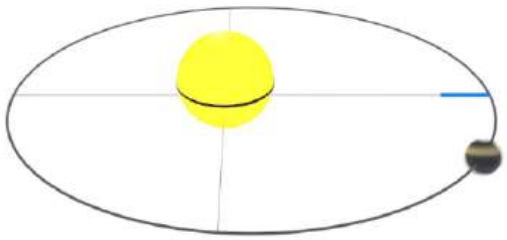
THE ICE AGES, CO_2 AND TEMPERATURE

THE SKEPTIC'S CORNER: CLIMATE CHANGE MISINFORMATION

It's not increasing CO_2 that is causing any observed increases in temperature. CO_2 is going up after temperature increases so we don't have to worry about CO_2 levels. It's just a natural process and requires no action to reduce fossil fuel use. Why reduce it if it doesn't cause global warming?

Data and models show that the global increase in temperature is driven mostly by increases in CO_2 (and not increasing temperatures driving increasing CO_2) as the predominant cause. That begs the question as to what starts the process of deglaciation. It turns out that cyclic increases in solar irradiance that increase temperatures, especially in the Northern Hemisphere, start deglaciation. A prime factor is the changes in the orbital dynamics of the Earth with respect to the sun. As you know, the orientation of the Earth's rotation axis remains generally fixed and pointed in the same orientation as the Earth rotates around the sun. This fixed orientation leads to our annual spring, summer, fall, and winter cycles on Earth. In the winter, the northern hemisphere is pointed away from the sun, leading to decreased solar irradiance per square meter in the Northern Hemisphere, causing winter there. When the Earth is on the opposite side of the sun, the axis points in the same direction but tilts towards the sun, leading to summer in the northern hemisphere. However, the orbital dynamics of the Earth do change in cyclic fashions over long periods of time. These long-term changes in the Earth's orbital shape (eccentricity), tilt (obliquity), and wobble (precession) are called the Milankovitch Cycle, and are illustrated in Figure 32.1.6. These cycles cause small temperature increases that start deglaciation. **Click on each image below** to download and view very short videos illustrating these orbital changes.

Change in eccentricity (orbital shape) - (100,000 yr cycle)



Change in obliquity (tilt) (41,000 yr cycle)

Axle precession (wobble) (26,000 yr cycle)

Figure 32.1.6: The Milankovitch cycle showing changes in the Earth's orbital dynamics with respect to the sun. https://climate.nasa.gov/news/2948/m...arths-climate/

Based on these cycles, Milankovitch calculated that recurring ice ages should occur approximately every 41000 years. Ice ages did occur at this interval from about 3 million years ago (MYA) to 1 million years ago (MYA). About 800,000 YA they lengthened to about 100,000 years, which corresponds to the Earth's eccentricity cycle. The increased duration of the cycle led to longer-lasting glaciers which moved further south in the Northern Hemisphere. One likely explanation for the increase in time between ice ages is that repeated glaciation/deglaciation eroded the bedrock in the Northern Hemisphere, converting it to **regolith** (rocks, soil, and dust). This allowed an increased velocity of movement of the glaciers to the south due to decreased frictional resistance, and thicker ice cap formation (more time to accrue ice), which required a longer time to melt. This also provided a positive feedback loop as the increased northern ice area would reflect more of the sun's energy back into space, cooling the planet. Punctuating these rhythmic orbital and ice age cycles are other events such as large volcanic eruptions, asteroid impacts, etc, that could produce minor to major changes in climate, and resulting mass extinctions.

Figure 32.1.7 shows how a combination of tilt angle, precession axis, and orbital shape at around 200 KYA (narrow rectangle across all the graphs) combined to lead to low glacial ice volume (bottom graph).

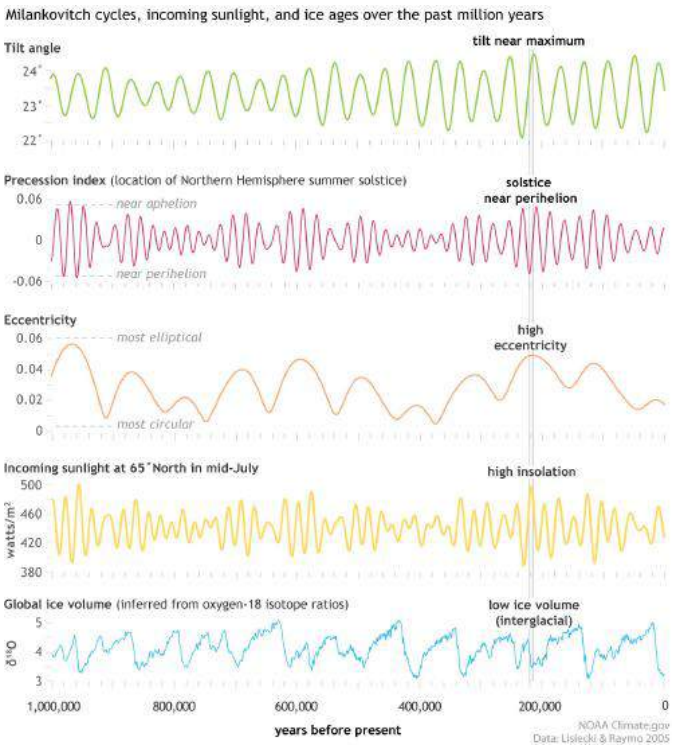


Figure 32.1.7: Milankovitch cycle contribution to ice volume over the past 1M years

If orbital changes (or **forcing**) trigger deglaciation, what is the role of increasing levels of the greenhouse gas CO_2, which clearly covary with temperature (see Fig 3)? Temperature increases derived from orbital and hence solar "forcing" seem to precede CO_2 increases for just short periods of time (perhaps 100 - 200 years). After that, CO_2 causes almost all of the global increase in temperatures during deglaciation, with CO_2 and temperature going up together. A global increase of about 0.3 C due to the Milankovitch cycle leads to greater Northern Hemisphere irradiance. This causes localized and limited melting of the Northern ice shield, leading to increases in ocean temperatures in the northern oceans. These increases slow a major ocean current (the Atlantic Meridional Overturning Circulation - AMOC) which inhibited the burial and return of cold water in tropical and southern oceans. This in turn led to a warming in the south accompanied by the release of large amounts of CO_2 stored in the oceans (see Carbon Cycle in 31.3). The release of this greenhouse gas was then responsible for most of the warming that lead to massive deglaciation. This "interhemispheric see-saw" transfer of heat from the north waters to the southern waters is key. For the far majority of the warming during glacial melting, CO_2 and temperature change synchronously.

Interpreting climate data is difficult. For example, it was found through measuring $^{15}N/^{14}N$ ratios that gases like N_2 and by extension CO_2 could rapidly diffuse through the compacting snow (firn, comprising the top 50-100 meters of the ice cap) until it became trapped in the solid ice beneath it. This would lead to the presence of "newer" CO_2 in older ice samples, and the conclusion the temperature changes preceded changes in CO_2. Corrections are made to the data to address the "apparent" time shift.

The CO_2 trapped in bubbles in the ice core samples from Antarctica reflects **global** CO_2 levels given atmospheric circulation but the temperatures measured from the same core samples (see Chapter 31.2) represent **local** (Antarctic) temperatures. Ice core samples from Greenland and ocean sediment samples from around the world are used to measure temperature at different locations over time. All of these data are required to model climate. Combined they lead us to our present interpretation of the linkage of CO_2 and temperature rise over time.

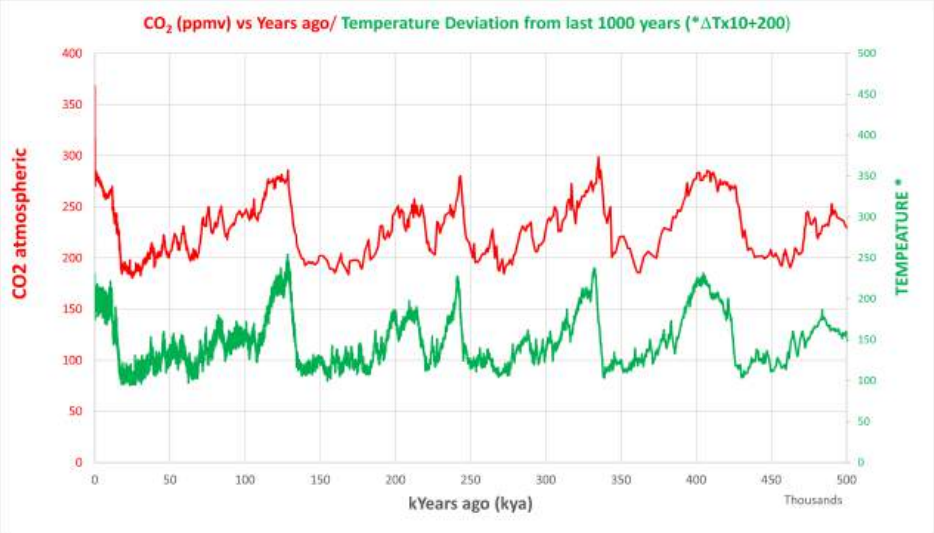
WHAT THE SCIENCE SHOWS

Increased solar irradiance on Earth arising from cyclic changes in the Earth's orbit leads to short, small temperature increases in the North Hemisphere. These lead to the **release of the greenhouse gas CO_2 from the oceans, which causes synchronous warming of the planet and subsequent deglaciation.**

So when skeptics say that temperature increases preceded CO_2 increases, you can acknowledge they did but that the bulk of the warming is attributed to increasing CO_2 released from ocean stores which leads to synchronous temperature increases and deglaciation. Using the words of chemistry, small temperature increases from orbital forcing catalyzed the release of huge amounts of CO_2 dissolved in the ocean. In Chapter 31.3 we will explore the carbon cycle in more detail and look at how it affects CO_2 levels.

TERMINATION OF THE ICE AGES

How did the ice ages terminate? Contributions from the orbital forcing derived from the Milankovitch cycle play a part. Another factor seems to be dust derived from regolith, itself made by glacier movement as we mentioned above. How can that hypothesis be tested? By using proxies for dust, namely iron and long-chain n-alkanes (derived from plant waxes) that have been deposited in sediments. First let's look at a graph of CO_2 and temperature changes and superimpose those on iron and long-chain fatty acid levels, as shown in Figure 32.1.8.



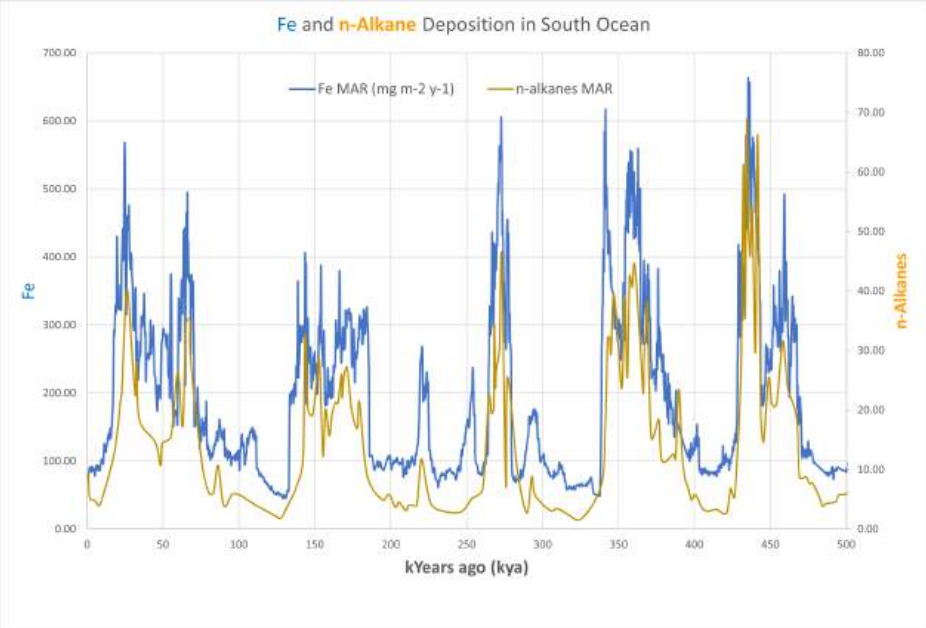


Figure 32.1.8: Fe and n-alkane deposition compared to CO_2 and temperatures in the Southern Ocean over the last 500,000 years. Note that the x axis shows time in kYA going to the right. (so the past is towards the right).

A close examination of the two vertically aligned graphs from around 120 K to 130 KYA shows that the iron and n-alkane depositions are at a minimum at the same that CO_2 and temperature are peaking! What explains this negative correlation? It depends on the intimate connection of the biosphere with the nonbiological world (an arbitrary distinction).

Iron and n-alkanes are circulated and delivered in dust. The long-chain alkanes, highly abundant in waxes and enriched in odd carbon number chains, were presumably derived from leaf waxes which prevent water loss from plants, especially during higher temperatures. Dust deposits were first observed in geological time in the switch from the warmer Pliocene (5.3 to 2.6 MYA) to the Pleistocene (2.6 MYA to 11.7KYA, see Fig 8 below). During the warmer Pliocene, the difference in global and atmospheric temperatures was lower, and with this

smaller temperature gradient, winds that could globally transfer dust would be diminished. Also, the warmer Plicoene (5.3 to 2.6 MYA) would have more rain, which would have removed dust from the global circulation.

As temperatures cooled in the Pleistocene (2.6 MYA to 11.7KYA), cycles of glaciation would produce more dust-containing regolith (rocks, soil, and dust), which would be dispersed through stronger global winds from higher temperature gradients and less rain. Dust contains carbon (for example long chain fatty acids) and perhaps more importantly iron, which is needed for oceanic phytoplankton growth. Without Fe, the uptake of CO_2 by phytoplankton (primary production) would not occur, leading to increased CO_2 in the atmosphere. Stronger regional atmospheric winds would lead to increased upwelling of nutrients as well as deep ocean CO_2. The CO_2 would enter the atmosphere more readily in the absence of dust deposition of iron.

In summary:

- High CO_2 and high temperature (lower global temperature gradients, more rain) are associated with low dust, as measured with the proxies Fe and n-alkanes). Low dust leads to low deposition of Fe and n-alkanes in the ocean, which decreases phytoplankton **primary production**, the fixing of CO_2 into biomass), leading to increased CO_2 movement from the ocean to the atmosphere, increasing temperature. This is an example of a positive feedback loop (higher temperatures leading to higher temperatures.
- Low CO_2 and low temperature (higher global temperature gradients, stronger winds, less rain) are associated with high dust with Fe and n-alkanes deposition. This increases phytoplankton primary production and decreases CO_2 movement from the ocean to the atmosphere, in a negative feedback loop.

By the end of a glacial deposition cycle, dust, blown by stronger winds from higher temperature gradients, was increasingly deposited on the ice sheets. Along with leading to more heat absorption by the sheets, it would also decrease their reflectivity (albedo). Both effects would promote ice sheet melting. Also, a cooler planet during glacial maximum had less precipitation, which along with lower CO_2, would lead to more plant and tree death, increasing soil erosion and desertification, both effects which would have increased dust production and its deposition on ice sheets. Then when CO_2 rose to 280 ppm, plant life renewed itself, and dust levels dropped.

CLIMATE CHANGE FROM 66 MILLION YEARS AGO TO NOW

Antarctic ice core data are now available for the past 2 M years. Ocean sediment data can be used to go back even further in time to 66 million years ago (MYA) just before the dinosaurs died after the massive asteroid impact forming the Chicxulub ***crater*** buried underneath the Yucatán Peninsula in Mexico. A brief review of geological eras, periods and epochs is shown below in Figure 32.1.9

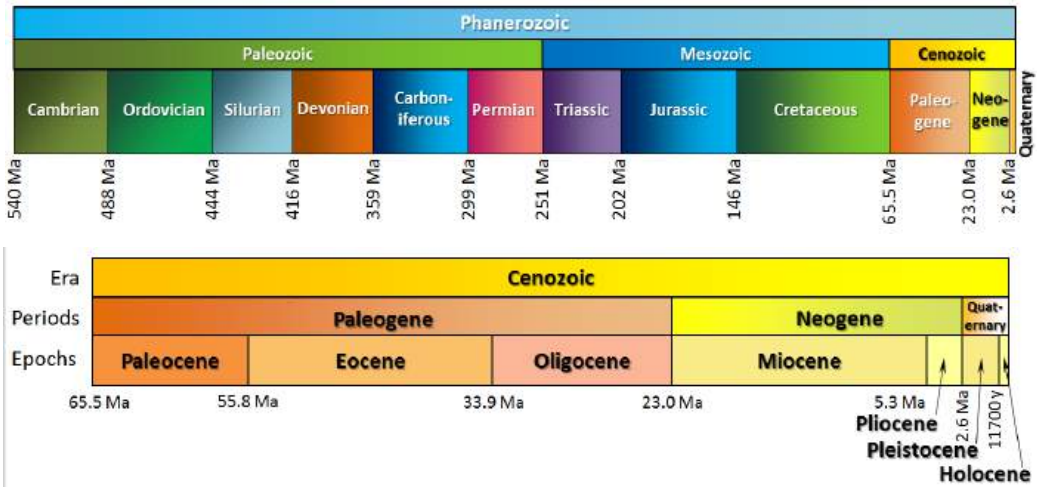


Figure 32.1.9: Geological Era, Periods and Epochs

CO_2 levels and associated temperatures derived from ocean sediment cores going back to 66 MYA are shown in Figure 32.1.10.

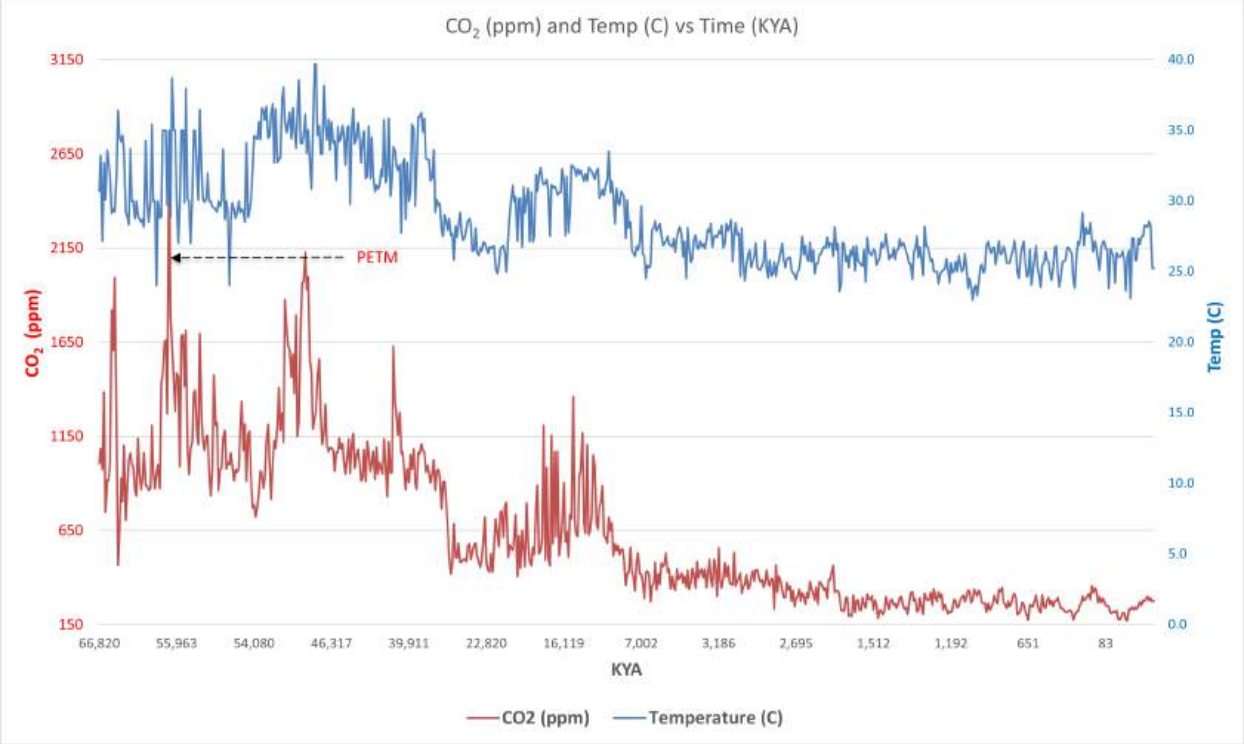


Figure 32.1.10 CO_2 levels (red) and temperatures (blue) derived from ocean sediment cores going back to 66 MYA = 66,000 KYA . Data from Rae et al. Annual review of Earth and planetary sciences, 49, 2021

Note again the parallel rise and fall of CO_2 and temperature. Eventually, they fall further in the Pliocene (5.3 to 2.6 MYA) and Pleistocene (2.6 MYA to 11.7KYA) epochs with cyclic glacier/interglacial periods we've discussed above. It wasn't until the late Miocene (10 to 6 MYA) that Northern hemisphere glaciation started and both poles of the planet had glacial sheets.

The time frame shown in Figure 9 encompasses the Cenozoic era (65 MYA when the dinosaurs died to about now). CO_2 levels were much higher than today in the greenhouse Paleocene and Eocene eras but decreased to about 500 ppm in the Oligocene (34 MYA). An almost stepwise drop in CO_2 and temperature occurred in the **Eocene to Oligocene transition (EOT)**, about 33 MYA. Data shows the development of large ice sheets appearing on Antarctica at this time. Before the EOT (33 MYA), Antarctica was ice-free, as shown in the recreation in Figure 32.1.11.

Figure 32.1.11: Reconstruction of the West Antarctic mid-Cretaceous temperate rainforest. Image credit: J. McKay / Alfred-Wegener-Institut / CC-BY 4.0. https://www.sci.news/othersciences/p...ing-09921.html

Proxy data for temperatures show that the transition was most likely caused by a decrease in CO_2 and some orbital forcing was probably involved. Present models still struggle to explain the EOT (33 MYA) transition, but it is clear that both CO_2 and temperature decreased. Where did the CO_2 go? Most assuredly into the oceans.

To understand that, we have to understand a bit about the carbon cycle, which we will discuss more fully in the next chapter section. Let's briefly discuss the role of atmospheric CO_2 and its interaction with the ocean. The main gases in the atmosphere, N_2 and O_2, are found in very low concentrations in the ocean since they are nonpolar and generally unreactive. CO_2 is also a nonpolar trace gas, but in contrast, it

can readily react with water to form HCO_3^- and CO_3^{-2}, which are found in great abundance in ocean reserves. Hence the **ocean chemistry of CO_2 determines in large part the levels of atmospheric CO_2**. The coupled reactions of CO_2 are shown below.

$$CO_2(\text{ g, atm}) \leftrightarrow CO_2(\text{aq, ocean}) \tag{32.1.2}$$

$$CO_2(\text{aq, ocean }) + H_2O(\text{I, ocean }) \leftrightarrow H_3O^+(\text{aq}) + HCO_3^-(\text{aq}) \tag{32.1.3}$$

$$H_2O(\text{I}) + HCO_3^-(\text{aq}) \leftrightarrow H_3O^+(\text{aq}) + {CO_3}^{2-}(\text{aq, sparingly soluble }) \tag{32.1.4}$$

This chemistry helps determine the pH of the ocean. Figure 32.1.12 shows atmospheric levels of CO_2 and ocean pH over the last 66 million years.

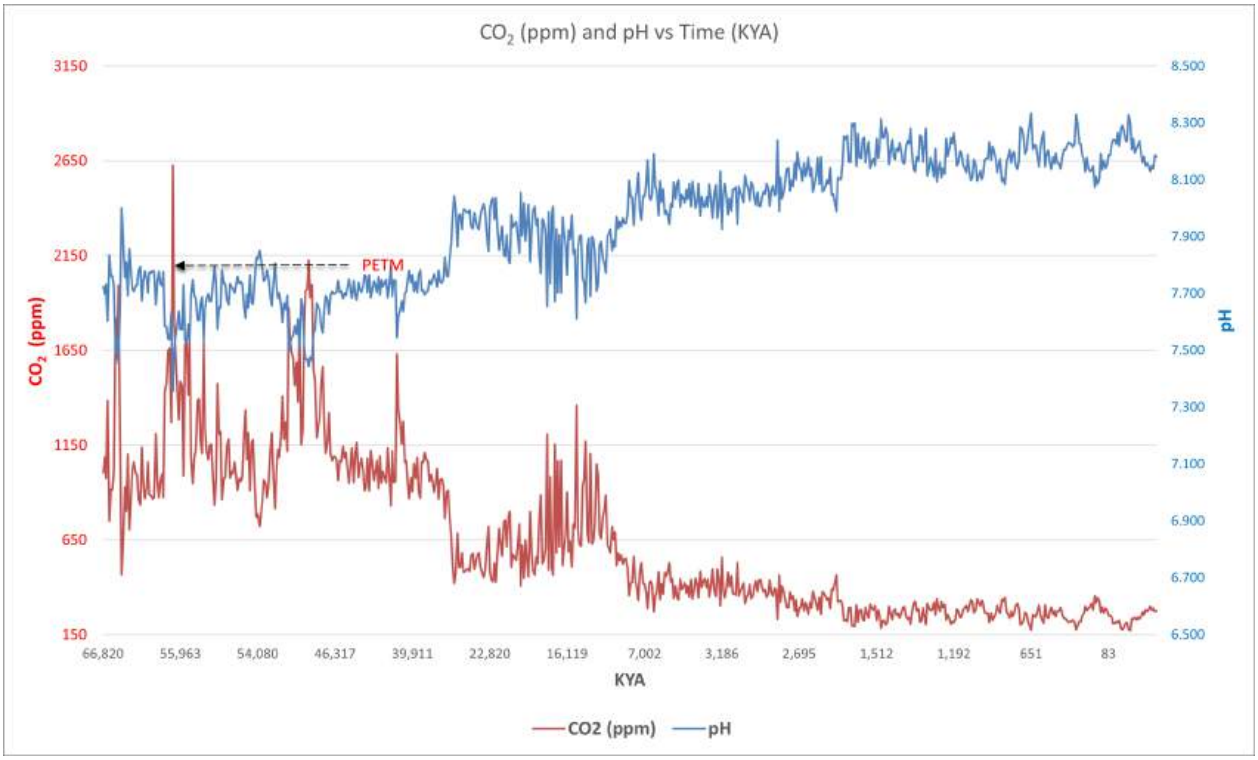


Figure 32.1.12: Atmospheric levels of CO_2 and ocean pH over the last 66 million years

Before the EOT at 34 MYA, atmospheric CO_2 levels were higher and ocean pH levels lower (around 7.7). After the EOT (33 MYA), atmospheric CO_2 is much lower and ocean pH is higher (more basic, 7.9 rising to 8.1). What happened to the CO_2 is a bit unclear. Atmospheric CO_2 decreased by moving into the oceans but wouldn't that have lowered the pH based on the chemical equations presented above? It would have but it turns out that the ocean alkalinity is determined not just by H_3O^+ produced by the equations above, but by the dissolved inorganic carbon ions, HCO_3^- (aq) and CO_3^{2-} (aq), which are conjugate bases. Increased HCO_3^- (aq) and SiO_4^{-2} (aq) from weathering solid carbonates and silicates that entered the oceans would raise the pH of the oceans.

A little review of introductory chemistry helps here.

Let's take bicarbonate, the weak conjugate base of the weak acid carbonic acid. HCO_3^- can act as both an acid and base.

Rx 1: Acts as an acid: HCO_3^- (aq) + H_2O (l) ↔ H_3O^+(aq) + CO_3^{2-} (aq)

$$K_{a2} = \frac{\left[H_3O^+\right]\left[CO_3^{2-}\right]}{\left[HCO_3^-\right]} = 4.7 \times 10^{-11} \tag{32.1.5}$$

Rx 2: Acts as a base: HCO_3^- (aq) + H_2O (l) ↔ H_2CO_3 (aq) + OH^- (aq)

$$K_{b2} = \frac{\left[H_2CO_3\right]\left[OH^-\right]}{\left[HCO_3^-\right]} = 2.2 \times 10^{-8} \tag{32.1.6}$$

The equilibrium constant for the reaction of HCO_3^- as a base is much larger so bicarbonate is a stronger base than acid.

Whatever the mechanism of the CO_2 drawdown, it led to decreasing temperatures in the EOT transition. Increased alkalinity of the ocean would also consume H_3O^+, increasing ocean pH.

A summary of planetary temperatures across geological time is shown in Figure 32.1.13.

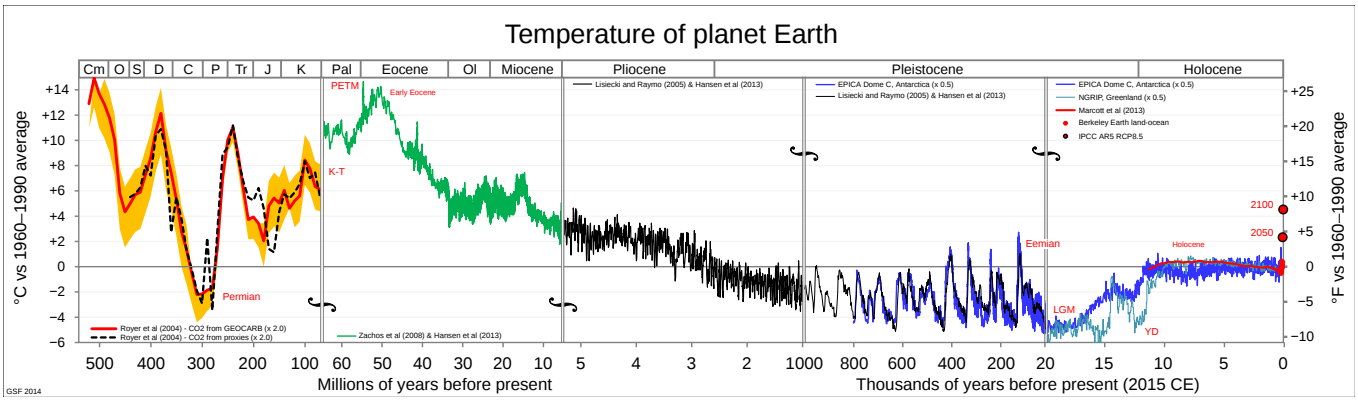


Figure 32.1.13: Temperature of Earth over 500 million years. https://commons.wikimedia.org/wiki/F...alaeotemps.png. (Excel available).
Creative Commons Attribution-Share Alike 3.0 Unported

There are several key features to note. The last time CO_2 was as high as today (415 ppm) was about 3 million years ago. Repetitive cycles of glaciation/deglaciation are obvious in the Pleistocene (2.6 MYA to 11.7KYA).

At around 55 MYA, a spike in temperatures of about 5^0 - 9^0F (or even more) on an already warmer planet occurred over about a 100K year timeframe. This is known as the **Paleocene/Eocene thermal maximum (PETM)**. The poles warmed to near 70 ^{0}F and alligators and palm trees were found there. The warming also led to the spread of tropic rainforests from the equator, allowing the evolution and proliferation of new plant species including flowering plants and an incredible biodiversity of insects, bird,s and animals that relied on them. Flowering plants produce fruits, which helped drive the evolution of mammals and the first true primates, including the tiny Teilhardina magnoliana. It may have resembled the picture below. (https://en.wikipedia.org/wiki/Teilhardina). Bigs eyes and hands would allow them to find fruit in the tropical forests.

The temperature increase was caused by a dramatic spike in CO_2 from deep-sea volcanoes and vents, which also led to a dramatic drop in ocean pH as measured by the loss of deep-sea $CaCO_3$ (chalk). Methane hydrates were also released. These environmental and biosphere changes are visually evident in geological deep-sea sediment records as shown in Figure 32.1.14.

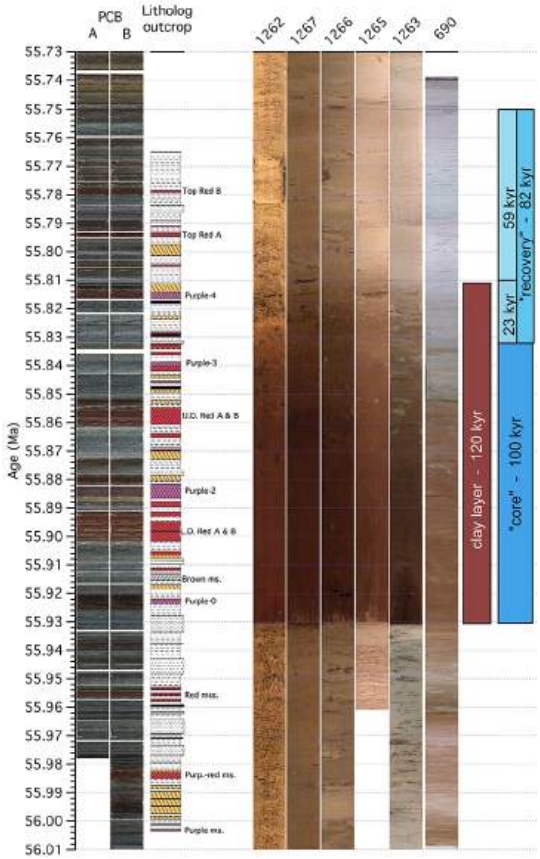


Figure 32.1.14: Overview for the Paleocene–Eocene Thermal Maximum (PETM, 55.5 MYA) data from deep-sea records and the terrestrial Polecat Bench
(PCB) drill core against age. Westerhold et al. Clim. Past, 14, 303–319, 2018. https://doi.org/10.5194/cp-14-303-2018. Creative Commons Attribution 3.0 License

Sediment cores were taken at various sites (1262, 1267, 1266, 1265, 1263, and 690) that are aligned from left to right according to the water depth from deep to shallow. Note at 55.93 million years ago, at the start of the PETM, there was a sharp transition from light brown/gray which is enriched in chalk, to dark brown enriched in clay. Ocean acidification dissolved the chalk. It took over 100,000 years to recover.

This very short time frame is called the **Paleocene/Eocene thermal maximum (PETM**, 55.5 MYA), which shows very quick spikes (on the geological time scale) can and do occur. Approximately 1.5 petagrams (10^{15}) of CO_2 were released annually during the PETM. Now we are releasing about 25 petagram per year. Our present rate of warming is much greater than the rate of warming during the PETM (55.5 MYA). The best candidates for the source of CO_2 release that caused the PETM are volcanoes, the oceans, and the permafrost. In addition, methane hydrates (a solid form of methane found in low- temperature high-pressure waters) might also be another factor.

The PETM and time after allowed for optimal evolution of primates. In fact that time might be called the Age of the Primates (analogous to the Age of the Dinosaurs). Figure 13 shows that the arrival of Oligocene was accompanied by much lower temperatures resulting from CO_2 drawdown. This decimated the habitat for the primates except for in the tropic rainforests. Primates disappeared completely from North America but thrived in Africa. Tectonic forces caused the generation of the African Rift valley, characterized by diverse geology and a less homogenous and more fractured landscape which again provided new challenges but also new habitats for the evolution of primates and the eventual appearance of Homo Sapiens.

This chapter has been divided into two sections. Got to Chapter 01B: Back to the Present and Future of Climate Change to see where we are and where we are headed!

KEY POINTS - BETA VERSION FROM CHAT.OPENAI

1. Climate change is the long-term change in the average weather patterns on Earth.
2. The primary cause of climate change is the burning of fossil fuels, which releases large amounts of greenhouse gases, such as carbon dioxide (CO2), into the atmosphere.
3. Greenhouse gases trap heat in the atmosphere, causing the Earth's temperature to rise. This is known as the greenhouse effect.
4. The most significant contributor to climate change is CO2, which is released when fossil fuels are burned. Other significant contributors include methane, nitrous oxide, and fluorinated gases.
5. Climate change has a wide range of impacts on the Earth's systems, including rising sea levels, changes in precipitation patterns, increased frequency and intensity of extreme weather events, and disruptions to ecosystems.
6. The global temperature has already risen by 1 degree Celsius (1.8 degrees Fahrenheit) since the pre-industrial era, with most of the warming occurring in the last few decades.
7. The Intergovernmental Panel on Climate Change (IPCC) has stated that limiting global warming to 1.5 degrees Celsius (2.7 degrees Fahrenheit) above pre-industrial levels could significantly reduce the risks and impacts of climate change.
8. Reducing greenhouse gas emissions is essential in order to slow or stop climate change. This can be achieved through a combination of actions, such as increasing the use of renewable energy sources, improving energy efficiency, and reducing deforestation.

32.01B: BACK TO THE PRESENT AND FUTURE OF CLIMATE CHANGE

Search Fundamentals of Biochemistry

LEARNING OBJECTIVES

- to demonstrate how climate has changed over geological time through the present
- to explain mechanisms, using knowledge from biology, chemistry, and physics, for climate change
- to show the central role of atmospheric CO_2 as a causative agent of past and present climate change
- to contrast the effects of anthropogenic burning of fossil fuels on climate change with causes of past climate changes
- to address arguments made by climate change skeptics

BACK TO THE RECENT PAST AND PRESENT

Let's return to more recent human history and the anthropogenic forcing of our climate. Figure 32.1.14 shows an interactive graph of atmospheric CO_2 over more recent times. **Zoom into the steep rise in CO_2 starts which around 1760 with the industrial revolution.**

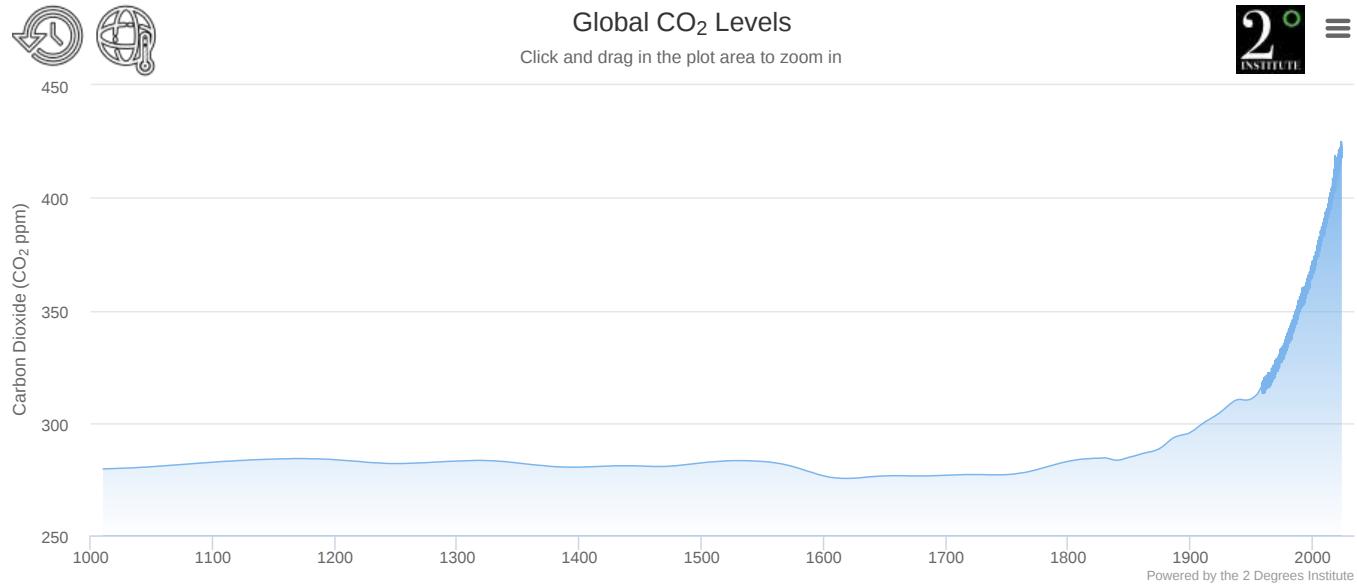


Figure 32.1.15: Interactive graph of atmospheric CO_2 vs time over the last 1000 years. Historical CO_2 record from the Law Dome DE08, DE08-2, and DSS ice cores. Credits: D.M. Etheridge, L.P. Steele, R.L. Langenfelds, R.J. Francey and the Division of Atmospheric Research, CSIRO, Aspendale, Victoria, Australia. 2 Degrees Institute. https://www.2degreesinstitute.org/. Also, https://www.co2levels.org/

Orbital mechanics **cannot** explain the warming of the planet in the brief (in geological terms) times since the industrial revolution. Neither can volcanic activity, changes in solar activity, changes in land use (for example deforestation that slows down photosynthesis and CO_2 removal from the atmosphere), or even aerosols released on burning fossil fuels (which would actually decrease global temperature due to increased reflection of sunlight). Click on Figure 32.1.15 to see an animated explanation of which factors best explains the global temperature increase since 1880. The results are clear: It's us!

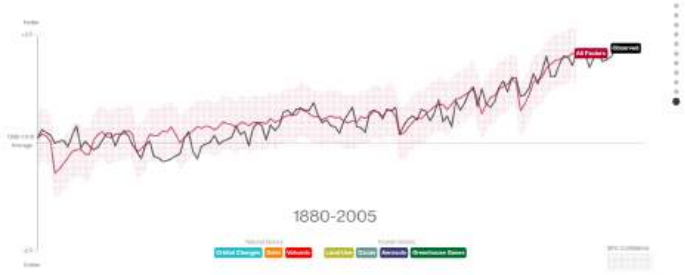


Figure 32.1.15: Factors contributing to global warming based on NASA data and NASA's Goddard Institute for Space Studies (GISS) climate models.

Unfortunately, other greenhouse gases have risen as well since 1975, as shown in Figure 32.1.16.

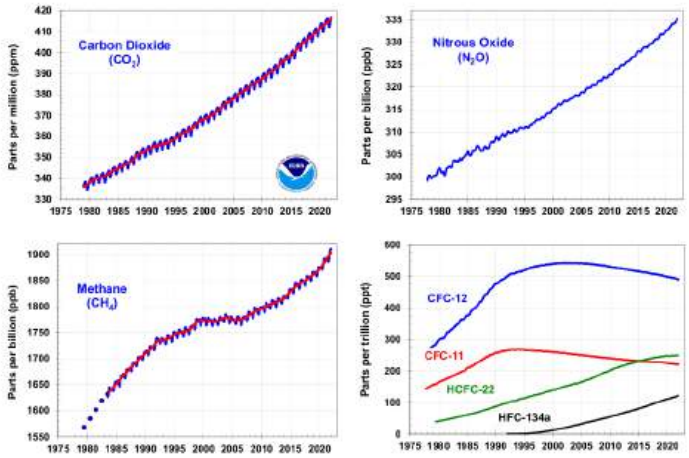


Figure 32.1.16: Rise in greenhouse gases since 1975. https://www.co2.Earth/annual-ghg-index-aggi

Take special note of the zig-zag nature of the CO_2 curve. The curve dips a bit in the summer when CO_2 is actively removed by plants in the Northern hemisphere. Increasing methane now accounts for up to 20% of the warming observed. **Figure 32.1.17 shows as interactive graph a very worrisome rise of atmospheric methane with time.**

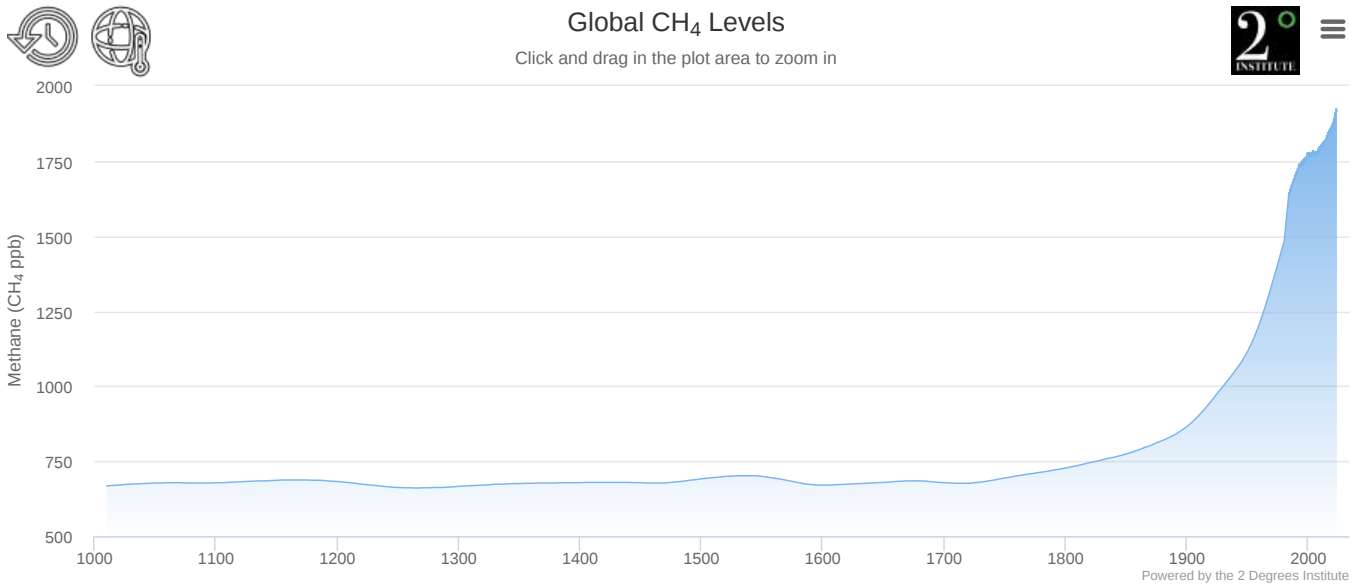


Figure 32.1.17: Interactive graph of atmospheric CH_4 vs time over the last 1000 years. 2 Degrees Institute. https://www.2degreesinstitute.org/

Present-day warming unequivocally is caused by humans' burning of fossil fuels.

PAST CLIMATE ANOMALIES

THE SKEPTIC'S CORNER: CLIMATE CHANGE MISINFORMATION

In recorded human history there have been other times of climate change, so we shouldn't worry about the present time! Look at the Little Ice Ages!

Several dramatic but short-lived (in geological time) climate changes have punctuated recorded human history. Let's look at two, mostly to equip you to address climate skeptics. They also show the sensitivity of our climate to subtle changes.

The Little Ice Ages

Actual and proxy temperature records show a mild period in Europe from around 950-1100 followed by colder weather, especially from 1450 to 1850. The latter period is called the "Little Ice Ages" although there was no significant expansion of the North Hemisphere ice shield. It was especially cold worldwide in 1816 when much of the world experienced a "year without summer". The effect in 1816 has a clear cause, the explosion of the volcano Mount Tambora in Indonesia on April 10, 1815.

But in addition to this identifiable influence in 1816, there was a cool period reported for the northern hemisphere from about 1800 to 1820 that started earlier than the Tambora eruption. Also, a low period of the sun's irradiance, called the Dalton Minimum, occurs from 1790-1860. Proxies for solar activity in the 1600s also show small solar irradiance drops, as we will discuss below. The dip in global average temperatures following the Medieval warm period, is shown in Figure 32.1.18.

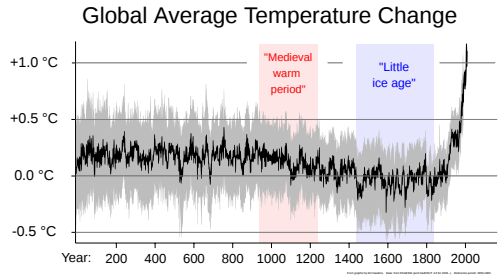


Figure 32.1.18: Dip in global average temperatures following the Medieval warm period, By RCraig09 - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/inde...curid=87832845

Modern climate changes have been captured in literature and art. One example is a painting showing "Ice Fairs" on the Thames in London, shown in Figure 32.1.19.

Figure 32.1.19: https://commons.wikimedia.org/wiki/F...enell.jpgdfdfd

Many factors probably contributed to the Little Ice Ages including a drop in solar irradiance. A newer explanation has also been proposed. Marine records show that the water near Greenland and the Nordic seas were warmer, caused by a strengthening of the Atlantic Meridional Overturning Circulation (AMOC). This would have caused the loss of Arctic ice in the late 1300s and 1400s, cooling the water and diluting its salinity, since ice when it crystallized with a tetrahedral hydrogen-bonded coordination of water, excludes salt. This would have collapsed the AMOC and its transfer of heat to the northern waters, leading to rapid and prolonged cooling. An analogous strengthening of the AMOC was observed between 1960 and 1980, which was attributed to a long-duration high-pressure system over Greenland. A similar event might have occurred to kick-start the Little Ice Ages. Tree rings show evidence of higher solar irradiance before the Little Ice Ages, which may be associated with the initial strengthening of the AMOC.

The Little Ice Ages also affected China and may account in part for a crop failure in 1644, the year in which the Ming Dynasty fell. There was also an Arctic hurricane in 1588 that helped destroy the Spanish Armada. The Great Fire in London in 1666 was preceded by a very dry summer that followed an exceptionally cold winter. Food production were severely disrupted, which might have led to significant social change in Europe and elsewhere, much as the Plague in Europe shattered societal and cultural norms.

The explosion of Mount Tambora, in present-day Indonesia, in 1816 greatly exacerbated the effects of cooling. The ash and SO_2 aerosols block solar irradiance, Droughts, floods, cholera epidemics, famine, and migration from Europe to the US and from East to West arose in part from this event.

One of the worst times to be alive: 536

Historians report that in 536 AD, parts of Europe, the Middle East, and Asia experienced 24 hours of darkness for up to 18 months. Summer temperatures plummeted. Famines occurred for a few years after. It snowed in China in the summer. The worst effects were in the Northern Hemisphere but the effects were world-wide. It was probably the most pronounced cooling in the last 2000 years. To make matters worse, a pandemic erupted around 541 that spread from southern Asia to northern Europe. It had a huge effect on the Byzantine Empire and has been called the Justinian (bubonic) Plague after the Byzantine emperor. Crop failures, an expansion of trade, and an influx of rodents derived from the cold temperatures could have led to and also exacerbated the plague.

This second and severe example of cooling was shorter-lived in a geological time frame. Temperatures fell in the summer about 1.5-2.5^{0}C. A "smoking gun" has been linked to this cooling, a volcanic explosion in Iceland. In addition, another eruption occurred in 540, which dropped the temperature another 1.5-2.5^{0}C, and in 547. The combined effects of climate change and the plague led to a significant economic fall in Europe. Signs of airborne lead in the ice in 640, arising from silver mining, suggest a recovery of economic growth. You should ask yourself how the modern world with cope with such an occurrence.

WHAT THE SCIENCE SHOWS

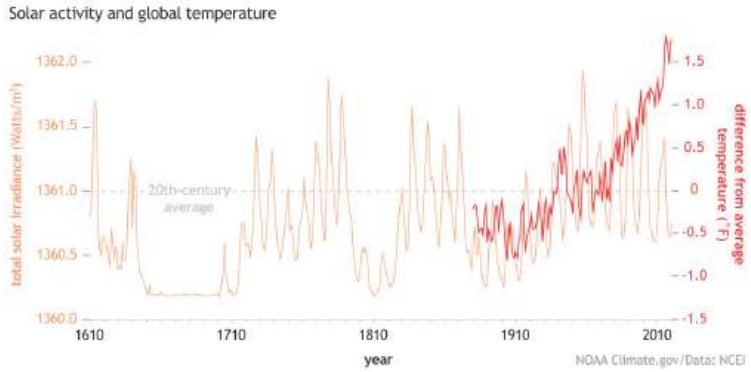
The Little Ice Ages and the climate changes preceding and after the Justinian plague had multiple causes, including volcanic eruptions, small changes in solar irradiance, and changes in the North Atlantic ocean currents and associated weather patterns. These short-term climate changes had disastrous effects on people's lives and the economic health of societies. Predicted future warming arising from CO_2 emitted from fossil fuel use (and other greenhouse causes) would bring far worse immediate and potentially irreversible consequences. It is incumbent on us as people who know the causes of climate change to act with due diligence and speed to avert the worst climate futures.

SOLAR ACTIVITY AND CLIMATE CHANGE

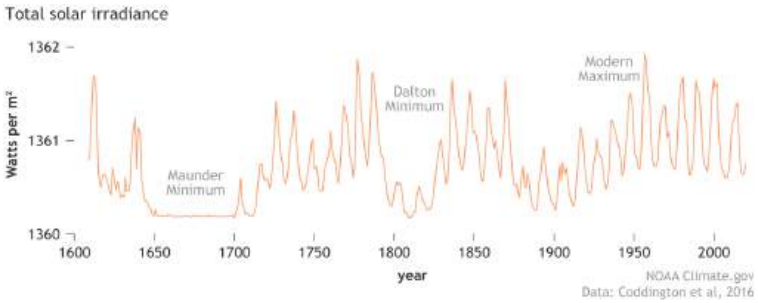
THE SKEPTIC'S CORNER: CLIMATE CHANGE MISINFORMATION

It's not increasing CO_2 that causes any observed increases in temperature. The sun's activity is changing. It always has and always will. There's nothing we can do about it.

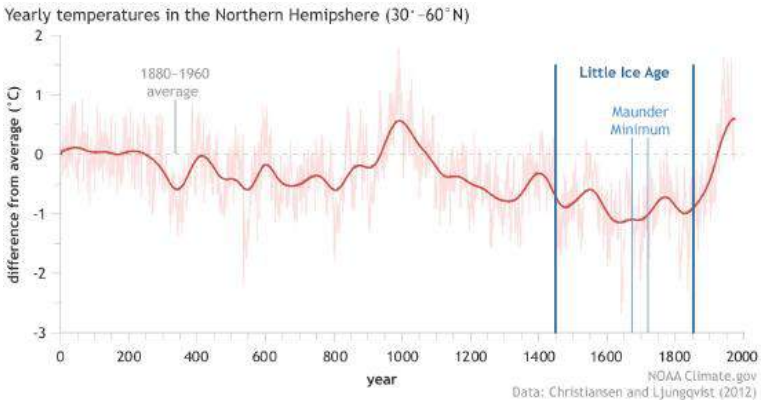
We have discussed how the orbital forcing of the climate kick-started each of the recurring ice ages in the Pleistocene. Some effects of the change in solar activity independent of the sun's orbit have been noted above. Specifically, we have shown that it cannot account for present warming. We present a series of graphs from the NOAA (National Oceanic and Atmospheric Administration) in the collective Figure 32.1.20 below to show the actual change in solar activity over recent times. Comments are shown at the bottom of each graph.



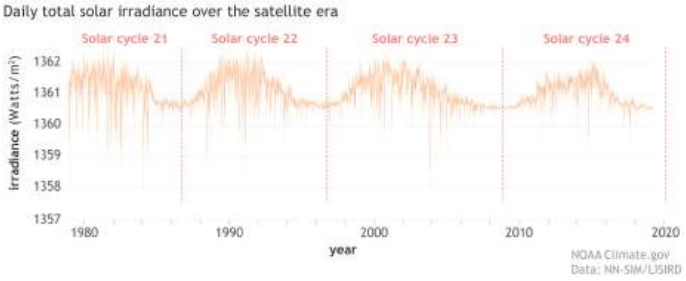
(Above) The maximal % spread from the lowest to the highest is very small. Such a small change shouldn't have such dramatic effects on climate unless it is sustained, as it was from around 1630-1700. Hence this decline in solar activity probably played some part in part of the Little Ice Ages. The regular rise and fall (spikes) are associated with the 11-year sunspot cycle activity. Note that the rise in average temperature since 1910 (shown in red) cannot be accounted for by change sin solar activity



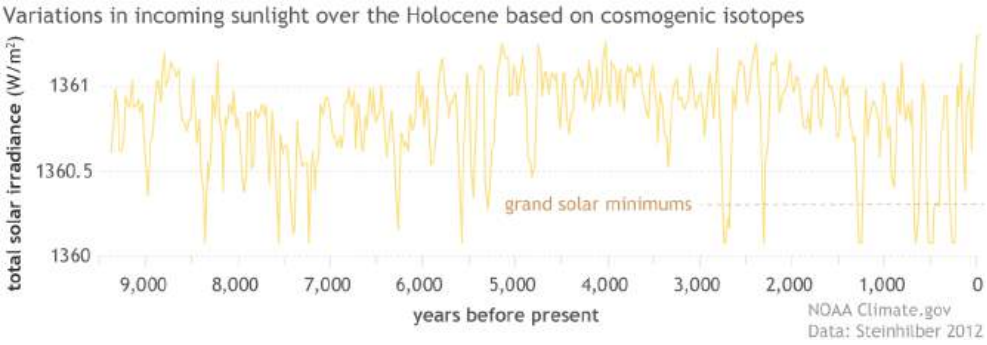
The above graph shows that the irradiance decreased by about 0.06% (although other values have been reported as high as 0.22%) during the Maunder Minimum, which occurred in the Little Ice Ages. The average decrease in terrestrial temperatures was 1.0-2^{0}C.



The graph above shows yearly average temperatures in the Northern Hemisphere. The dark red line shows the average change. Note that the averages are clearly lower in the Little Ice Age with the lowest values and lowest spike temperatures close to and in the Maunder Minimum.



(Above) The 11-year repeat of sunspot activity and resulting solar irradiance is clearly seen in the graphs. In 2020, a low in activity occurred, yet 2020 was the **second warmest year** on record since 1880.



This graph does not show the effects of climate forcing due to orbitals changes. Rather it shows that solar activity has not changed significantly for the 10000 years prior to 0 CE.

Figures 32.1.20: Changing in solar activity in recent geologic time.

This would be true if not for the massive amount of CO_2, approximately 1.5 trillion tons, injected into the atmosphere since the industrial revolution from the use of fossil fuels. Not all of that is still in the atmosphere, of course, but enough to raise CO_2 to levels not seen for 3 million years. Based on the relationship between CO_2 and temperature across the ice ages, science can predict when conditions might exist to initiate and propagate the next ice age. The data arising from these models, illustrated in Figure 32.1.21, show how much incoming solar radiation (insolation) must arrive at the Earth (watts/m^2) to trigger the next ice age.

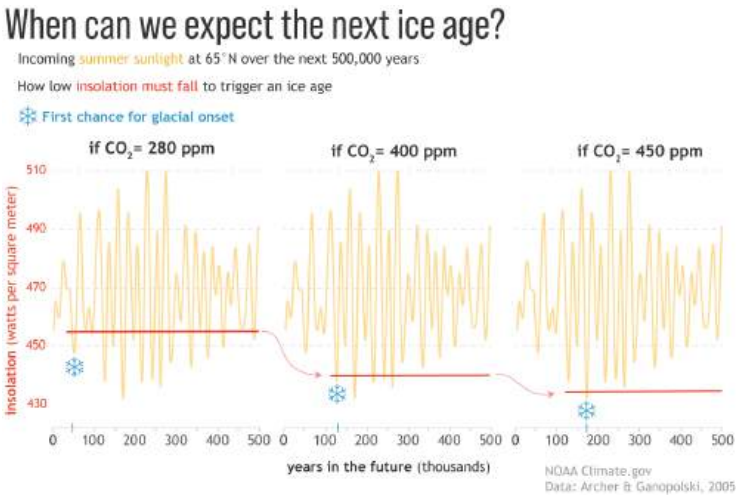


Figure 32.1.21: Incoming solar radiation required to trigger the next ice age.

As shown in the left side of the figure, if CO_2 were 280 ppm (typical of peaks in past interglacial periods), it would take repetitive drops of insolation below the threshold of about 455 watts/m^2 (red line) to start glaciation. As of November 2022, we are at 415 ppm and rising. If it rises to 450 ppm, as it assuredly will in the absence of carbon capture, it would require much less insolation, since the greenhouse effect of the higher CO_2 would warm the atmosphere The right graph shows there is little chance of another ice age in the absence of large and sustained volcanic activity or asteroid impact that would lead to blocking of solar radiation.

WHAT SCIENCE SAYS

Changes in solar irradiance (not changes in Earth's orbital dynamics) **cannot** account for warming since the Industrial Revolution. They have contributed to short-term (on a geological time scale) cooling during the Little Ice Ages.

Summary of Climate Change Causes and Effects Since 900 AC

Figure 32.1.22 shows a great summary of possible contributions to temperature change over the last 1000 years. Note again that present-day warming can only be attributed to greenhouse gases (GHG). One panel shows changes in land use. This has caused a temperature drop since 1800. That effect is caused by deforestation and other land cover changes, which leads to more reflection of incident solar radiation back into space. This effect is increased in the winter if the changed land is snow- covered. Deforestation would also decrease CO_2 capture (photosynthesis) by plants, which would raise the temperature. That component has been added to the GHG panel.

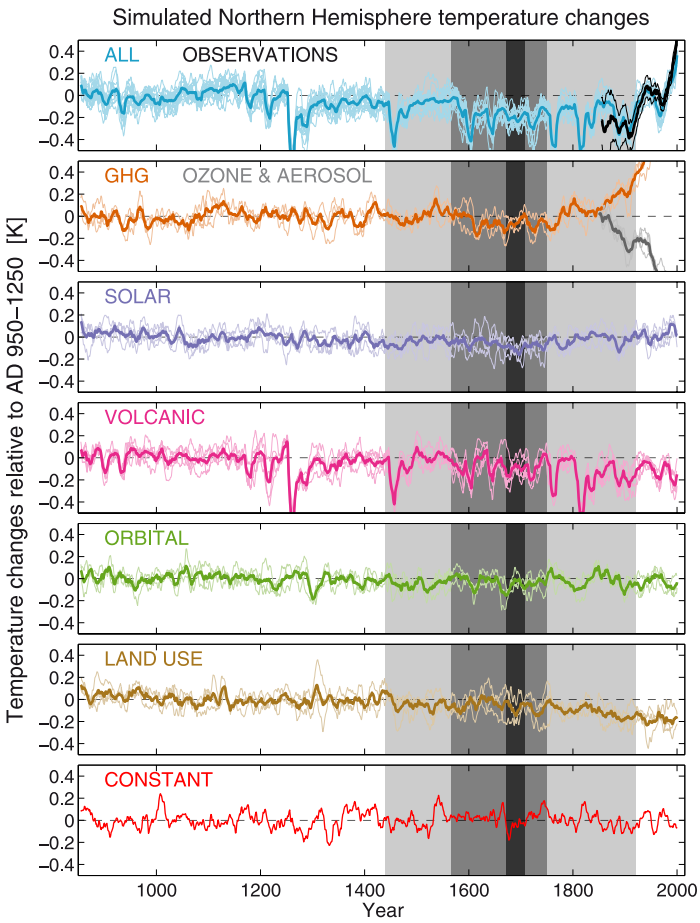


Figure 32.1.22: Simulated northern hemisphere temperature changes, smoothed with an 11 year running mean, relative to the period AD 950–1250. Owens et al. J. Space Weather Space Clim. 2017, 7, https://doi.org/10.1051/swsc/2017034. Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0).

The black line in the top panel shows the observed instrumental northern hemisphere temperature variations with their associated uncertainties (Morice et al., 2012), which match the simulations well. The bottom panel shows a simulation with no changes to the radiative forcings. This quantifies the magnitude of natural internal variability in the simulations in the absence of changes in forcings. Note periodic dip but short time frame dip in temperature due to volcanic activity. Clearly, warming since the Industrial Revolution is due to emissions from the use of fossil fuels.

CLIMATE JUSTICE: THE EMITTERS AND THE AFFECTED

THE SKEPTIC'S CORNER: CLIMATE CHANGE MISINFORMATION

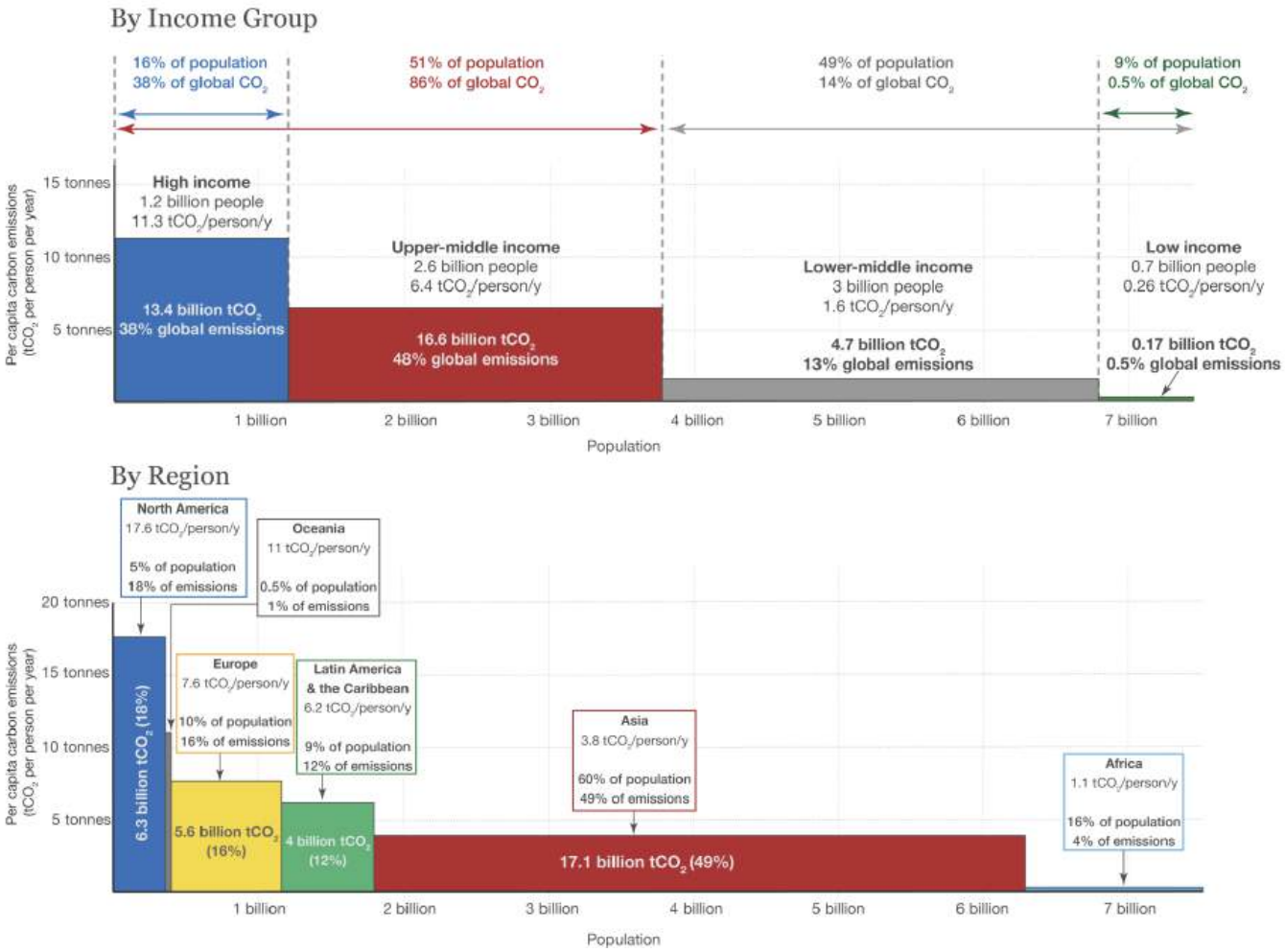
Why should we make changes to reduce fossil fuel emissions when China is the biggest emitter of CO_2!

We present a series of graphs in Figure 32.1.24 below, taken from CO_2 Emissions - Our World in Data to show which countries have emitted the most CO_2 in the past and now. In a just world, those countries which have emitted the most should move swiftly and forcefully not only to decrease emissions but to aid other countries' transitions to clean fuels and to help them with climate change mitigation and adaptation. We don't wish to demonize the fossil fuel industry and those who work in it. The use of fossil fuels, which are high energy, high density, and cheap fuel (because of historically massive subsidies) has lifted millions if not billions out of poverty over time. We had no alternative to fossil fuels until recently. Most did not realize how significantly fossil fuel use would affect our present and future climate and the health of not only humans but the entire biosphere. (Yet there is evidence that.) We can't just stop the use of fossil fuels without inflicting great economic pain on those who can least afford it. In order to help those who are currently suffering and who will suffer most in the future, as well as to help ourselves, our children and out grandchildren, we must move away from the use of fossil fuels as soon as possible.

Above: The dip in total world emissions in 2020 was due to the COVID pandemic. Unfortunately, the rise has resumed. Note that China is now the biggest net emitter but the US and EU emissions are dropping. India is on the rise and if they follow a similar economic path as China, which they need to lift many out of poverty, it will come with a huge cost in CO_2 emission unless they can jumpstart their conversion to clean fuels. The world needs to help.

Above: Although China is the largest net emitter, **the US and Australia are the biggest emitters per person**, although that is dropping.

The US still leads the world in the total amount of CO_2 emitted since the industrial revolution. We also have the greatest GDP. Pakistan suffered tragical flooding, exacerbated by climate change, in 2022. Up to a 1/3 of the country was under water. In a just world, the biggest emitters would aid the rest of the world.



Above: Inequality is clearly evident in this graph as the wealthiest people (high and upper-middle income) collectively contribute 86% of CO_2 emissions

Figure 32.1.24: CO_2 emissions by country and income since the industrial revolution. Our World in Data. Creative Commons BY license.

WHAT SCIENCE SAYS

The United States has emitted the most CO_2 since the beginning of the industrial revolution and the most per capita. China is not even close.

FUTURE PROJECTIONS

We know the science, and we know the consequences if we choose not to act or act in ways insufficient to meet the challenges of climate change. It is one of the most difficult challenges we have faced as a species. It requires sacrifice and united action for the common good. The benefits of our choice are mostly in the future and for future generations.

The Intergovernmental Panel on Climate Change (IPCC), a body composed of leading climate scientists and experts, has defined several different **Relative Concentration Pathways (RCPs)** leading to different emissions and different climate futures. Where we end up depends on economic, social, and political choices. The IPCC initially designated four pathways, RCP 2.6, 4.5, 6, and 8.5, with higher numbers associated with higher temperatures and CO_2 levels. Each assumes a starting value and estimated emissions (which depend on technology, politics, economics, etc). RCP 8.5 assumes extra radiative forces (heat energy/(m^2s)) by 2100 equal to 8.5 J/(s m^2) or 8.5 watts/m^2. This worst-case scenario assumes business as usual with no interventions to reduce our emissions, a totally unlikely scenario given present actions (including the rapid rise of clean energy). The RCP 2.6 scenario assumes that the peak radiative forcing would be 3 watts/m^2 which would decline through very strong governmental and economic actions to 2.5 by 2030-2040. Table 32.1.1 below shows the four RCP scenarios with projected ending CO_2 equivalents (which include other greenhouse gases) and temperature increases.

RCP (W/m^2)	Timeframe	$CO_{2\ atm}$ equivalents (ppm)	Temp increase (°C/°F)	Description
8.5	in 2100	1370	4.9/8.8	rising
6.0	after 2100	850	3/5.4	stabilizing without overshoot
4.5	after 2100	650	2.4/4.3	stabilizing without overshoot
2.6	decline from 3 before 2100	490	1.5/2.7	peak and decline

Translating the projected CO_2 equivalents in the atmosphere into associated temperature increases requires a high-quality value for climate sensitivity (rise in temperate/rise in CO_2). Figure 32.1.25 shows the likely increase in temperatures for the four different scenarios.

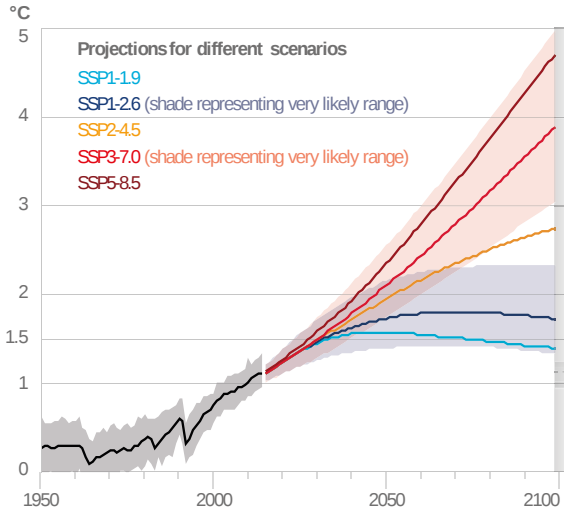


Figure 32.1.25 Projected temperature rises for 4 RCP scenarios. Pörtneret al. *Climate Change 2022: Impacts, Adaptation, and Vulnerability.* Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. doi:10.1017/9781009325844.001.

The scenarios in Figure 25 are labeled SSP#-## with the second number ## representing the RCP number. The IPCC 6th report issued in 2021 changed from using RCP scenarios to Shared Socioeconomic Pathways (SSPs) scenarios which are based on possible social and economic developments that would pose different challenges to reduce future temperature increases and hence different strategies for mitigation and adaptation. The SSP scenarios are consistent with the RCP scenarios but use a more enhanced socio-economic and political framework for their construction. The mitigation strategies are based on the RCP forcing levels. The SSP scenarios are described below. They start with SSP1, which leads to a world that has adapted well and moved away from fossil fuels, to SSP5, which assumes a continued and high reliance on fossil fuels.

2023 - NEW CLIMATE SENSITIVITY CALCULATIONS

We mentioned in 31.1A that a new value for climate sensitivity giving a **rise of 4.8 °C (8.6 °F) for a doubling of CO_2** has been determined by Hansen et al (2023). This value has startling implications. If this new value stands, the IPCC/UN call to keep warming to 1.5 ^{0}C above the pre-industrial revolution average is now **impossible**. We are already there or will be there in a few years. It is highly likely that we will not make the changes required to keep warming below 2°C before 2050. The reasons for our inaction are politicians who continue to support the fossil fuel industry and who block significant legislation to act now, religious institutions and leaders who don't make climate action a sacred duty, and the fossil fuel industry that continues to expand fossil fuel exploration and mislead the public as to the causes of climate change and action to address it. Those who say we can't afford to act never mention the much greater cost of inaction!

SSP1: **Sustainability – Taking the Green Road (Low challenges to mitigation and adaptation)**
The world shifts gradually, but pervasively, toward a more sustainable path, emphasizing more inclusive development that respects perceived environmental boundaries. Management of the global commons slowly improves, educational and health investments accelerate the demographic transition, and the emphasis on economic growth shifts toward a broader emphasis on human well-being. Driven by an increasing commitment to achieving development goals, inequality is reduced both across and within countries. Consumption is oriented toward low material growth and lower resource and energy intensity.

SSP2: **Middle of the Road (Medium challenges to mitigation and adaptation)**
The world follows a path in which social, economic, and technological trends do not shift markedly from historical patterns. Development and income growth proceeds unevenly, with some countries making relatively good progress while others fall short of expectations. Global and national institutions work toward but make slow progress in achieving sustainable development goals. Environmental systems experience degradation, although there are some improvements and overall the intensity of resource and energy use declines. Global population growth is moderate and levels off in the second half of the century. Income inequality persists or improves only slowly and challenges to reducing vulnerability to societal and environmental changes remain.

SSP3: **Regional Rivalry – A Rocky Road (High challenges to mitigation and adaptation)**
A resurgent nationalism, concerns about competitiveness and security, and regional conflicts push countries to increasingly focus on domestic or, at most, regional issues. Policies shift over time to become increasingly oriented toward national and regional security issues. Countries focus on achieving energy and food security goals within their own regions at the expense of broader-based development. Investments in education and technological development decline. Economic development is slow, consumption is material-intensive, and inequalities persist or worsen over time. Population growth is low in industrialized and high in developing countries. A low international priority for addressing environmental concerns leads to strong environmental degradation in some regions.

SSP4: **Inequality – A Road Divided (Low challenges to mitigation, high challenges to adaptation)**
Highly unequal investments in human capital, combined with increasing disparities in economic opportunity and political power, lead to increasing inequalities and stratification both across and within countries. Over time, a gap widens between an internationally-connected society that contributes to knowledge- and capital-intensive sectors of the global economy, and a fragmented collection of lower-income, poorly educated societies that work in a labor-intensive, low-tech economy. Social cohesion degrades and conflict and unrest become increasingly common. Technology development is high in the high-tech economy and sectors. The globally connected energy sector diversifies, with investments in both carbon-intensive fuels like coal and unconventional oil, but also low-carbon energy sources. Environmental policies focus on local issues around middle and high income areas.

SSP5: **Fossil-fueled Development – Taking the Highway (High challenges to mitigation, low challenges to adaptation)**
This world places increasing faith in competitive markets, innovation and participatory societies to produce rapid technological progress and development of human capital as the path to sustainable development. Global markets are increasingly integrated. There are also strong investments in health, education, and institutions to enhance human and social capital. At the same time, the push for economic and social development is coupled with the exploitation of abundant fossil fuel resources and the adoption of resource and energy-intensive lifestyles around the world. All these factors lead to rapid growth of the global economy, while global population peaks and declines in the 21st century. Local environmental problems like air pollution are successfully managed. There is faith in the ability to effectively manage social and ecological systems, including by geo-engineering if necessary.

The projected increases in emitted CO_2 (Gigatons/yr) and other greenhouse gases over the next 80 years for each SSP scenario are shown in Figure 32.1.26. The second number in the SSP label is the RCP scenario number based on radiative forcing listed in Table 1 above.

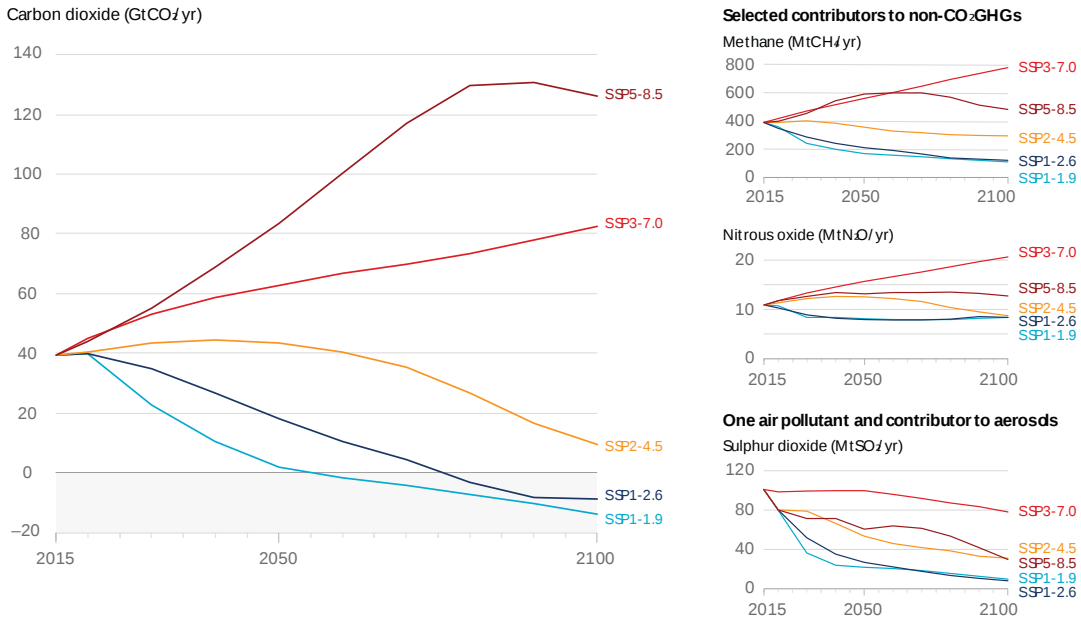


Figure 32.1.26: Projected increases in greenhouse gases under different SSP (RCP) scenarios. Masson-Delmotte et al. Climate Change 2021: The Physical Science Basis. Contribution of Working Group I. to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 3−32, doi:10.1017/9781009157896.001.

Note the welcome decline in SO_2 which causes acid rain as well as aerosols. This shows that under all SSP scenarios, we are moving to clean up our air (in this case reducing SO_2 from burning sulfur-enriched coal or capturing SO_2 before it enters the atmosphere). Paradoxically and unfortunately, decreasing aerosols leads to increasing temperatures due to lower reflectance of incident solar irradiation.

Our final figure, Figure 32.1.27, shows how each greenhouse gas and SO_2 are projected to change in 2081-2100, compared to 1850-1900 levels, for each of the SSP scenarios.

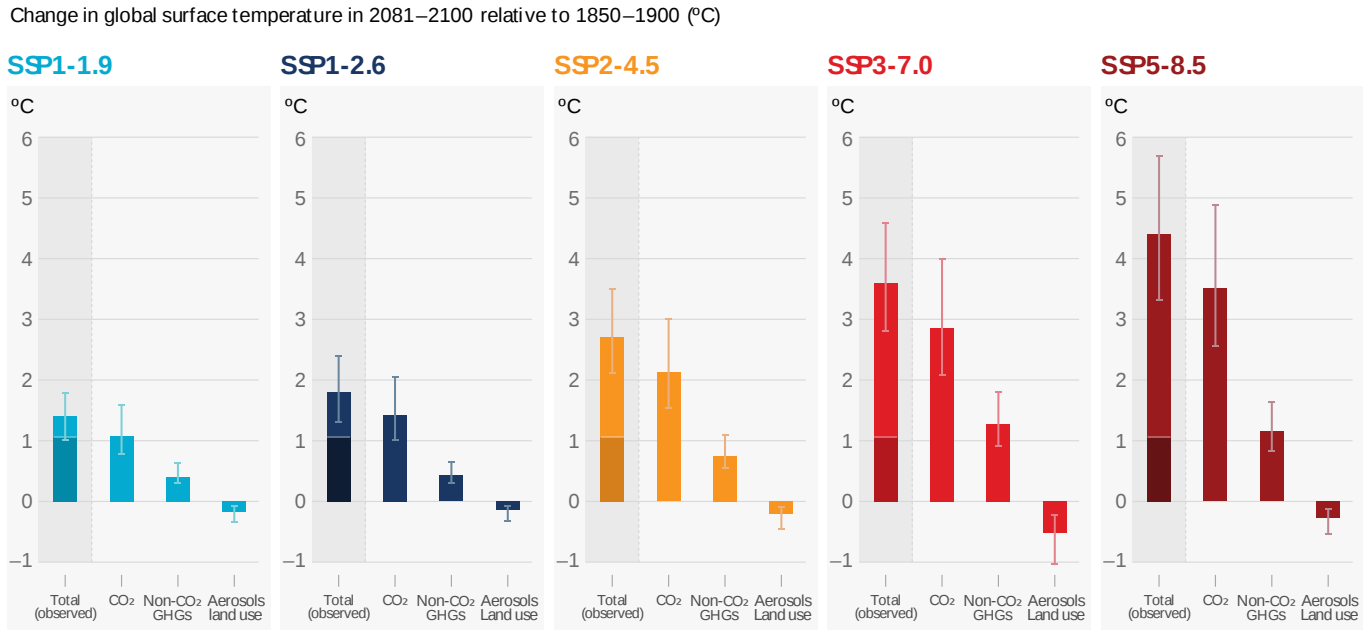


Figure 32.1.26: Projected changes in greenhouse gas es and SO_2 in 2081-2100 compared to 1850-1900 levels for different SSP scenarios. Masson-Delmotte et al, ibid.

We discussed in Chapter 32.1: A that the Earth's Energy Imbalance (EEI) rose from about +0.6 Watts/m^2 in 2000 to +0.9 from around 2005 to 2019, and to +1.0 in around 2021. The trend is unfortunately upward, as shown in Figure 32.1.27.

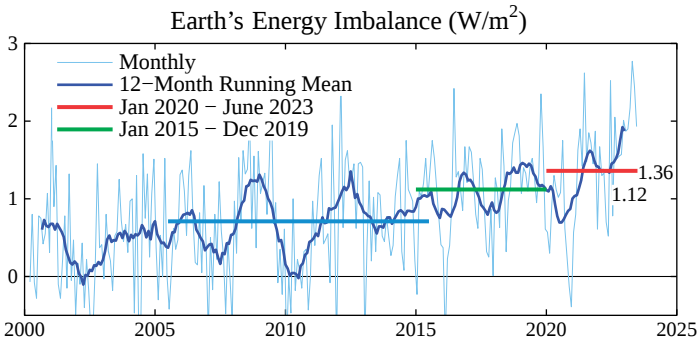


Figure 32.1.27: Earth's Energy Imbalance since 2000. 12-month running-mean of Earth's energy imbalance from CERES satellite data normalized to 0.71W/m2 mean for July 2005–June 2015 (blue bar). James E Hansen, Makiko Sato, Leon Simons, Larissa S Nazarenko, Isabelle Sangha, Pushker Kharecha, James C Zachos, Karina von Schuckmann, Norman G Loeb, Matthew B Osman, Qinjian Jin, George Tselioudis, Eunbi Jeong, Andrew Lacis, Reto Ruedy, Gary Russell, Junji Cao, Jing Li, Global warming in the pipeline, *Oxford Open Climate Change*, Volume 3, Issue 1, 2023, kgad008, https://doi.org/10.1093/oxfclm/kgad008. Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/)

The projected value of 1.36 (red bar) for the 2020s bodes poorly for our ability to reverse the effects of climate change. The very large change that started around 2015 probably stems from the **decrease** in atmospheric sulfates (required by legislation) in emissions from ships in the northern Atlantic and Pacific oceans and from coal burning in China. These decreases in atmospheric aerosols, which cause cardiovascular, pulmonary, and cancer, paradoxically increase global warming. Sulfate aerosols reflect solar radiation. They also form clouds which do the same. Both cool the planet. In fact, aerosols that contribute to some of the worst pollution in the world in cities in South Asia have helped keep temperatures there from rising to dangerous levels.

These last two chapter sections clearly show that humans have been engaged in large-scale geoengineering of the planet and its biosphere. It is likely that we will have to devise new methods to bring the temperature of the Earth down. One might be large scale production of salt-aerosols over the oceans that would reflect incident sunlight and increase cloud formation.

As all the data presented in this chapter shows, our climate fate will depend on the choices we make individually and collectively as societies.

KEY POINTS - BETA VERSION FROM CHAT.OPENAI

1. Climate change is the long-term change in the average weather patterns on Earth.
2. The primary cause of climate change is the burning of fossil fuels, which releases large amounts of greenhouse gases, such as carbon dioxide (CO2), into the atmosphere.
3. Greenhouse gases trap heat in the atmosphere, causing the Earth's temperature to rise. This is known as the greenhouse effect.
4. The most significant contributor to climate change is CO2, which is released when fossil fuels are burned. Other significant contributors include methane, nitrous oxide, and fluorinated gases.
5. Climate change has a wide range of impacts on the Earth's systems, including rising sea levels, changes in precipitation patterns, increased frequency and intensity of extreme weather events, and disruptions to ecosystems.
6. The global temperature has already risen by 1 degree Celsius (1.8 degrees Fahrenheit) since the pre-industrial era, with most of the warming occurring in the last few decades.
7. The Intergovernmental Panel on Climate Change (IPCC) has stated that limiting global warming to 1.5 degrees Celsius (2.7 degrees Fahrenheit) above pre-industrial levels could significantly reduce the risks and impacts of climate change.
8. Reducing greenhouse gas emissions is essential in order to slow or stop climate change. This can be achieved through a combination of actions, such as increasing the use of renewable energy sources, improving energy efficiency, and reducing deforestation.

32.2: USE OF ISOTOPE ANALYSIS IN MEASURING CLIMATE CHANGE

Search Fundamentals of Biochemistry

In this chapter section, we will explore how we are able to reconstruct CO_2 and temperature values across millions of years time. It is truly a remarkable, if not awe-inspiring achievement that shows the intellectual rigor and endurance scientists employ to obtain, interpret, and model data. The graphs of CO_2 vs temperature across such a vast swath of time shown in the previous chapter section were obtained through analyses of isotopes of oxygen and carbon found in chemical species in ice and ocean floor sediments. The interpretation of the isotope data requires an understanding of the link between individual and linked chemical and biochemical reactions. So we have to take a deep dive into isotopes and their use.

WHY STUDY ISOTOPE EFFECTS

- Isotopes and their effects, critical in understanding both structure and activity in biochemistry, are key in climate science.
- Both equilibrium and nonequilibrium reactions and processes apply to isotope partitioning into water and biomolecules, in ways similar to linked biochemical reactions and pathways.
- The study and application of isotope effects can integrate and expand learning from previous science courses

In more global terms

- The strength and quality of our data and models that support explanations in biochemistry and climate analyzes must be explored.
- Scientific and ethical rigor in pursuit of knowledge, explanations, and solutions as we pursue complex fields like biochemistry and climate change is essential if we are to place trust in our experts and faith in their findings.

ABSOLUTE AND PROXY MEASUREMENTS FOR CO_2 AND TEMPERATURE

It's simple to see how CO_2 levels over the last 800,000 years have been determined from ice cores since ancient air is actually trapped in bubbles in them. The bubbles can be liberated by melting/shattering and analyzed. However, how can we infer temperature changes or actual temperature from ice core samples and both CO_2 and temperatures from ocean sediment cores? Scientists use "proxies" to determine temperature as modern thermometers and temperature scales were invented only recently (early- and mid-1700s by Fahrenheit and Celsius). These proxies include tree rings, growth bands in coral, pollens found in core samples, and calcite shells from marine organisms found in lake and ocean sediments. The organisms include certain types of algae, phytoplankton like dinoflagellates and diatoms, and foraminferas (single-celled protozoans with shells).

From a more chemical perspective, the analyses of the isotopic compositions of ice, using the isotope ratios of $^{18}O/^{16}O$ in water, and of marine sediments, using the ratio of $^{18}O/^{16}O$ and $^{13}C/^{12}C$ in carbonate-containing shells, have proved critical in determining temperatures back millions of years ago. Even molecules like components of leaf waxes (as we discussed in Chapter 31.1) can be used. Isotope analyses of diatoms (with silicate shells) in lake sediments are also useful.

The analysis of proxies is quite complicated since many factors contribute to the measurement derived from proxy use. Take tree rings as an example. The width of a given ring depends not only on temperature but precipitation. Calibrations for each proxy must be made using alternative data such as rings from a variety of other trees. Proxy data from ice cores go back a few million years, while data from marine sediments back as far as 100 million years ago. Isotope analysis in rocks formed from marine or land sediments can go back billions of years. Proxy data (tree rings) taken from recent times can be compared to actual temperatures measurement from the same time. The calibration relationships can then be used for samples from the past. Alternatively, proxy data can be taken across many different places and temperatures to develop calibration constants, an approach useful for pollen analysis. Organisms could also be cultured in different temperatures, nutrient, and CO_2 conditions to calibrate past data. From a statistical sense, it is best to combine multiple proxies to develop past temperature records. Proxy data sets are available for over 10,000 sites around the world. Figure $32.2.x$ below gives a link to sites compiled by Carbon Brief where databases with information about each study is available.

Figure 32.2.1: Map of over 10,000 proxy data studies and sites. Robert McSweeney and Zeke Hausfather and Tom Prater. https://interactive.carbonbrief.org/...-distant-past/

Most people might care little about climate change millions of years ago. The value in understanding climate change so far in the past is, in part, to build confidence in the data, methods of analysis, and climate model to better understand the relationship between CO_2 and temperature. Some, particularly the PAGES 2k Consortium, focus their attention on the last 2000 years of the Common Era. Figure 32.2.2 shows global mean temperatures obtained from proxy data (yrs 0 - 2000) and direct observations (through the use of thermometer and satellite measurements) since around 1850.

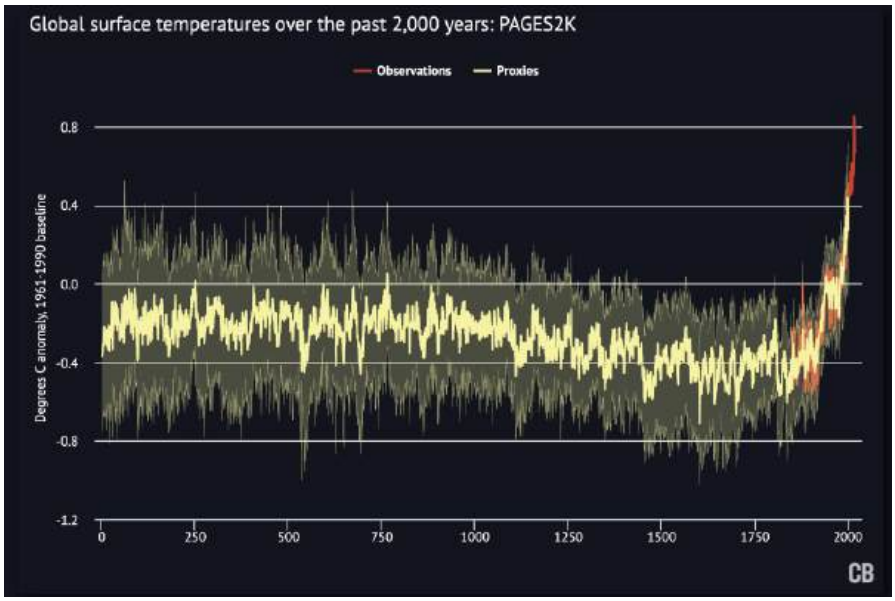


Figure 32.2.2: Global mean surface temperature reconstruction (yellow line) and uncertainties (yellow range) for the years 0-2000 period from the PAGES 2k Consortium along with observations from Cowtan and Way from 1850-2017. Data available in the NOAA Paleoclimate Archive.

Note the overlap of proxy measurements and direct observations of temperatures since around 1850.

OCEAN MICROORGANISMS

Students of biochemistry come from many backgrounds and all do not have a strong biology background. To help with that, and to develop a sense of wonder about the microorganisms that inhabit the oceans and play such a key part in the biosphere, let's look at a few relevant to this chapter. For those with a "chemistry-centric" background, these descriptions are probably new.

Plankton

The word plankton derives from a Greek word meaning drifter or wanderer. There are two main types. One is **zooplankton,** which are not plants, but rather microscopic animals and protozoans. They are heterotrophs that don't synthesize their own food. Most have calcite shells. The other is **phytoplankton,** which are autotrophic plants that use photosynthesis for food production. Hence they are carbon "capturers", which are key players in the carbon cycle in maintaining atmospheric CO_2 and in producing O_2. The phytoplankton broadly include algae

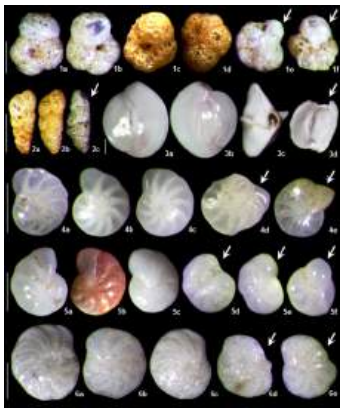
(protists), cyanobacteria (also known as blue-green algae), and dinoflagellates which also fit into other groups. Here are some examples. Also remember that protists are eukaryotic organisms that are not animals, plants, or fungi.

Table 32.2.1 below lists types and examples of plankton.

Type of plankton	Examples
Zooplankton (heterotrophs)	benthic foraminiferans, which live mostly at sea bottoms and in sediment and capture carbon indirectly through the carbon cycle though the use of CO_3^- in their shells. planktonic foraminifera, which live near the surface but are found buried in ocean sediments after their death dinoflagellates that don't photosynthesize are small animals including tiny fish and crustaceans such as krill and jellyfish. radiolarians, single-cell protozoans that have calcium silicate shells
Phytoplankton (autotrophs, primary producers)	diatoms photosynthesizing dinoflagellates blue-green algae which are prokaryotic bacteria green algae, a photosynthetic eukaryotic protist. some foraminifera that live near the surface watera and can photosynthesize

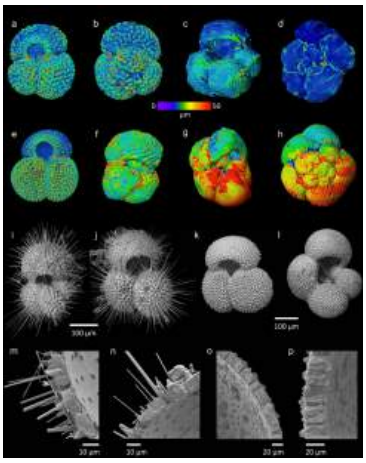
Zooplankton

Existing shells in sea sediments from foraminifera have been critical in dating studies and determining CO_2 and temperatures over millions of years. There are two major types which include benthic foraminifera (which live at the sea bottom and in sediment) and a smaller group of planktonic foraminifera which live near the surface. They are heterotrophs, but some, on ingestion of small autotrophic phytoplankton, can retain and sequester their chloroplasts, which can engage in photosynthesis for a period of time. Figures $32.2.3-5$ below show examples of zooplankton that have been important in climate studies.



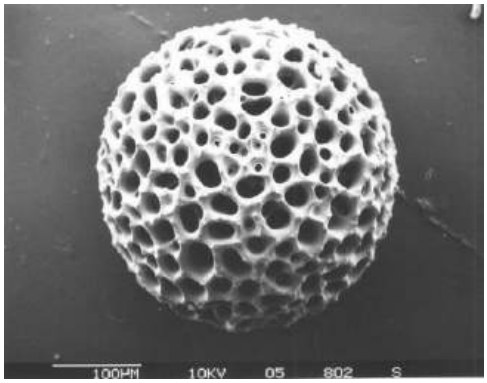
Figures 32.2.3 **above: Benthic foraminifera:**

Live mostly at sea bottoms and in sediment); capture carbon indirectly through the carbon cycle through the incorporation of CO_3^- into their shells. Living benthic foraminifera in the Bohai Sea, showing normal specimens and abnormal individuals (indicated by arrows).



Figures 32.2.4 **above: Planktonic foraminifera**

Live near the surface but are found buried in ocean sediments after their death. (**a**–**h**) Nano-CT scan of planktonic foraminifera specimens with color map of test thickness, warm colors indicating areas of relatively thicker shell; (**a**,**b**) *Globigerinoides ruber* (*Tara*), (**c**) *Globigerina bulloides* (*Tara*), (**d**) *Neogloboquadrina dutertrei* (*Tara*), (**e**) *G. ruber* (*Challenger*), (**f**) *Trilobatus trilobus* (*Challenger*), (**g**) *N. acostaensis* (*Challenger*), (**h**) *N. dutertrei* (*Challenger*); (**i**–**p**) SEM images of selected planktonic foraminifera specimens; (**i**) *T. trilobus* (*Tara*), (**j**) *G. ruber* (*Tara*), (**k**) *G. ruber* (*Challenger*), (**l**) *G. bulloides* (*Challenger*), (**m**,**n**) *G. ruber* test cracked to reveal wall texture (*Tara*), (**o**,**p**) *G. ruber* test cracked to reveal wall texture (*Challenger*).



Figures 32.2.5 **Above: Radiolaria .**single-cell protists that secrete silica

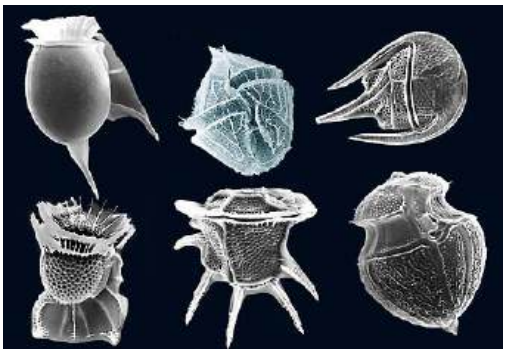
Figures $32.2.3-5$: Examples of zooplankton. Benthic foraminifera: https://commons.wikimedia.org/wiki/F...raminifera.p; Planktonic foraminifera: Creative Commons Attribution 4.0 International License. Fox, L., Stukins, S., Hill, T. *et al.* Quantifying the Effect of Anthropogenic Climate Change on Calcifying Plankton. *Sci Rep* **10**, 1620 (2020). https://doi.org/10.1038/s41598-020-58501-w. http://creativecommons.org/licenses/by/4.0/.; Radilaria: https://commons.wikimedia.org/wiki/F...ria-sp2_hg.jpg

Phytoplankton

Phytoplankton are microscopic plants, and as such, engage in photosynthesis, capture CO_2, and produce O_2. Hence they are primary autotrophs. We will consider three types, diatoms, photosynthetic dinoflagellates and coccolithophores. Diatoms and photosynthetic dinoflagellates are the major ones and are prey for the zooplankton. They are described in Figures $32.2.6-8$ below.

Figures 32.2.6 **Above: diatoms**

Single-celled eukaryotic algae surrounded by a silica shell (test). These can reach 1 mm in diameter and can form an assortment of shapes. Some can form multicellular chains. They engage in high efficiency photosynthesis and resulting carbohydrate synthesis. They are found in coastal and cold waters with lots of nutrients.



Figures 32.2.7 **Above: photosynthetic dinoflagellates**

Algae with a single shell. They are smaller than diatoms. Most have two flagella for motion. They have a cellulose shell, which degrade on death. Hence, they don't have shells that enter the sediment. Some are nonphotosynthetic and are considered zooplankton.

Figures 32.2.8 **Above: coccolithophores**

Coccolithus pelagicus; coccosphere. These are very small single cell algae, which form interlinked calcium carbonate circulate plates that cover the surface.

Figures 32.2.6 – 8 : Some phytoplankton. Diatoms: https://commons.wikimedia.org/wiki/C...le:Diatom2.jpg; photosynthetic dinoflagellates: https://commons.wikimedia.org/wiki/F...lagellates.jpg; coccolithophores: https://commons.wikimedia.org/wiki/F..._pelagicus.jpg

Along the coast in summer, nutrient-rich upwelling of water occurs which can lead to explosive growth of dinoflagellates, causing the water to become red-gold (often called a red tide). Some species in these blooms produce neurotoxins such as saxitoxin (inhibitor of sodium channels), which can produce paralytic shellfish poisoning if shellfish from the bloom area are eaten, and brevitoxin (stimulate voltage-gated sodium channels in nerve and muscle).

Ice cores from Antarctica and Greenland can extend to over 3.4 km (2.1 miles) in depth and yield direct information on CO_2 and indirect measurements of temperature. The oldest continuous ice core records extend to 130,000 years in Greenland, and 800,000 years in Antarctica. Data going back 2 million years is available using discontinuous cores. Concentrations of trapped CO_2 as a function of time, and the temperature of each layer can be determined. Figure 32.2.9 shows a section of an ice core from the West Antarctic Ice Sheet Divide (WAIS Divide).

Figure 32.2.9: *The dark band in this ice core from the West Antarctic Ice Sheet Divide (WAIS Divide) is a layer of volcanic ash that settled on the ice sheet approximately 21,000 years ago. Credit: Heidi Roop, NSFhttps://icecores.org/about-ice-cores*

The remains of plankton shells described above are found in cores from sea sediments. Analyses of shells, especially from foraminifera, have provided climate data going back tens of millions of years ago. For both ice and ocean sediment cores, isotope analyses have been the key to obtaining CO_2 and temperature data.

ISOTOPE ANALYSES

In analyzing ice and ocean sediment cores, three things are needed: the age of the layer, a direct or indirect measure of the atmospheric CO_2 at the time the layer was deposited, and an indirect measurement of the temperature at the time of deposition. As shown in the figure above, ice core samples have rings, similar to trees, that can be used to count backward in time. The rings get harder to distinguish the further back you go. Figure 3 above shows a visible dark band deposited by volcanoes 21,000 years ago. Ultimately, isotopic analyses of H_2O and CO_2 in ice samples and of carbonates in minerals and deposited microfossils in ocean samples are critical in determining past CO_2 and temperature values.

Most readers are familiar with ^{14}C radioisotope dating and ^{13}C-NMR. Metabolic pathways have been elucidated using ^{2}H (deuterium), ^{3}H (tritium), ^{13}C, and ^{14}C to label specific atoms in substrates and follow their flow into products. These same isotopes have been used in kinetic experiments to determine enzyme reaction mechanisms. We will explore isotopes in some detail in this section.

Use of unstable radioactive isotopes

The use of ^{14}C radioisotope dating is limited in climate analyses given its short half live ($t_{1/2}$ = 5730 years). In contrast to most isotopes made in stellar nucleosynthesis or by the radioactive decay of a precursor radioactive elements to an isotope of another element, ^{14}C is made on a continual basis in the atmosphere when high energy neutrons (n) from solar radiation react with atmospheric nitrogen (N). The neutron kicks out a proton to form ^{14}C as shown in the nuclear reaction below.

$$n + {}^{14}_{7}\mathrm{N} \rightarrow {}^{14}_{6}\mathrm{C} + p \tag{32.2.1}$$

^{14}C becomes oxidized to form $^{14}CO_2$ which can then enter the carbon cycle and enter the organic carbon pool through uptake by photosynthetic organisms and organisms that consume them. It can also form inorganic bicarbonates and carbonates, which could enter into shells.

All living things take in ^{14}C until their death, after which ^{14}C decays through the conversion of a neutron to a proton, a beta particle (electron) and an antineutrino, forming stable ^{14}N. Hence the amount of ^{14}C in dead organisms or their remains diminishes with a $t_{1/2}$ = 5730 years, in a process not affected by temperature or pressure. ^{14}C dating can be used in samples dating back about 55,000 years, a time span representing 9.6 half-lives. Only 0.13% of the original ^{14}C would be left. Data by this method give the age of death of the organism.

Carbon-14 dating depends on the assumption that its amount in the environment is constant, but the burning of fossil fuels and detonation of nuclear weapons has altered its amount (see box below). Changes in solar activity and resulting changes in high-energy neutrons also affect the amount of carbon 14. Also given the relatively short time frame used in ^{14}C dating, differences in CO_2 based on its sequestration and circulation in the oceans are also factors. 61% of the Northern Hemisphere is covered by oceans compared to 81% of the Southern Hemisphere. Books of calibration factors can be used to control for these effects. The calibrations are based on tree rings, lake and ocean sediments, corals and stalagmites data which allows dating back to 55,000 years ago.

NUCLEAR WEAPONS, FOSSIL FUELS AND ^{14}C DATING

In the 1950s up to 1962, nuclear weapons were tested in the air, doubling the amount of ^{14}C in the air. This spike has been taken up into organisms and into the ocean. Also since then, the amount of CO_2 from the burning of fossil fuels has gone up dramatically. This source does not contain ^{14}C as it derives from fossils that long ago decayed. The effects canceled each other in 2021. Since 2021 a lot more CO_2 from fossil fuels has been added so the net effect is now lower levels of ^{14}C equivalent to preindustrial time. It will continue to lessen until well after we stop using fossil fuels. By 2050 the levels might be equivalent to those in the Middle Ages. This, and the human-made stoppage of the next ice age glaciation cycle is yet another warning to us about our effects on the entire biosphere.

The decay of other "unstable" radioactive isotopes is used for dating samples and determining their age of burial:

- ^{39}Ar , an extremely rare isotope ($t_{1/2}$ =269 yr), has been used to date ice cores from the Tibetan Plateau over the last 1,300 years.
- ^{40}K ($t_{1/2}$ =1.25 billion yr) decays to ^{40}Ar (stable), so their ratios can be used to determine how much time has passed since magma solidification into rock, based on rates of diffusion of the resulting stable ^{40}Ar .
- The ratio of $^{26}Al/^{10}Be$ in buried samples is used in dating analyses. The two isotopes are rare and produced in a fixed ratio (6.75/1) when they are formed in surface quartz by solar radiation (much like the formation of ^{14}C). When buried through geological processes, there is no further production of the isotope, but fortunately (for those who measure age of burial), they decay with different half lives ($t_{1/2}$ = 717,000 yr for ^{26}Al and $t_{1/2}$ = 1.39 million yr ^{10}Be)
- The ratios of $^{21}Ne/^{26}Al$ and $^{21}Ne/^{10}Be$ can be used. ^{21}Ne is a stable isotope and these ratios are independent of the $^{26}Al/^{10}Be$ rate.
- Uranium isotopes are widely used in age measures on the long time scale. ^{238}U ($t_{1/2}$ =4.45 billion yr) is converted to ^{206}Pb (stable) and ^{235}U ($t_{1/2}$ =704 million yr) to ^{207}Pb (stable) by parallel decay routes which allow for multiple types of dating measurement.

Use of stable isotopes

Most of the data and graphs of CO_2 and temperature vs time (years ago) presented in Chapter 31.1 were determined by using stable isotopes that do not decay. Much of the data is based on either the ratio of the stable isotope pairs of oxygen ($^{18}O/^{16}O$) or carbon ($^{13}C/^{12}C$) in buried ice or ocean sediment cores. These isotopes have also been used to infer the temperature or temperature change when targets were buried in ice core or ocean sediments. Temperatures at the time of deposition of water in the ice layers are often inferred from $^{18}O/^{16}O$ ratios in the ice layers.

ICE CORE $^{18}O/^{16}O$ ANALYSES

The oceans are huge and generally homogenous reservoirs that can give clues to long-term changes in climate. Short-term climate change would have limited effects on the oceans. The $^{18}O/^{16}O$ ratios in Greenland and Antarctic ice cores have allowed dating and temperature reconstruction over geological time since the ratio is determined by the $^{18}O/^{16}O$ in the liquid oceans at the time of ice formation.

The % natural abundance of ^{18}O (0.205 %) and ^{16}O (99.757 %) gives a ratio of the two isotopes of 0.0021, which is so small that the exact ratio is inconvenient for routine use. Rather, a comparison of the ratios in a target sample vs a universal reference, the **$\delta^{18}O$ value**, is determined using a mass spectrometer. The $\delta^{18}O$ value is calculated by the following equation:

$$\delta^{18}O = \left[\frac{\left(\frac{18}{16}O\right)_{\text{sample}}}{\left(\frac{18}{16}\text{O}\right)_{\text{reference}}} - 1 \right] * 1000 \qquad (32.2.2)$$

Similar δ values are determined for D/H ratios (δ^2H) and for $^{13}C/^{12}C$ ($\delta^{13}C$)

- The reference for $\delta^{18}O$ calculations is the Standard Mean Ocean Water (SMOW or V-SNOW)
- The δ^2D reference value is also based on SMOW or V-SNOW
- The standard for the analogous $\delta^{13}C$ value is the Cretaceous Peedee Belemnite (an extinct order of squid-like cephalopods with an internal cone skeleton) sample from the Peedee belemnite (PDB) formation in South Carolina, USA. This standard is no longer available so an alternative, NBS 19, a carbonate material, is used in a new V-PDB (Vienna-PDB) scale.

Table 32.2.4 shows the ratio of the isotope abundancies in nature and in the references.

Element ratio	Ratio of natural abundance	Reference ratio
$^{18}O/^{16}O$	0.205/99.757 = 0.00205	0.0020052 (SMOW or V-SMOW)
$^{13}C/^{12}C$	1.1/98.9 =0.0111	0.011238 (PDB or V-PDB)
$^{2}H/^{1}H$ (D/H)	0.0156/99.9844 = 0.000156	0.00015576

Note that the ratios of the standards, and likewise of the samples, are small. Also note that in the equation for $\delta^{18}O$, the bracketed term is multiplied by 1000. If multiplied by 100, the value for $\delta^{18}O$ would be a percentage. Instead, it's multiplied by 1000 to convert it to **permill** (per mil or ‰) or parts per thousand (just like percent is parts per 100). Hence 1‰ is 1 part per 1000 or 0.1%.

Equations can be just collections of letters with little intuitive meaning, or they can be deconstructed by the user to make intuitive sense. To help understand this equation, which most readers have likely never encountered, given its importance in climate studies, let's look at 3 sets

of conditions. If ...

- ^{18}O is enriched in the sample compared to the reference, then $(^{18}O/^{16}O)_{sample}/(^{18}O/^{16}O)_{reference}$ is >1, so subtracting 1 from it makes the bracketed term +, along with $\delta^{18}O$;
- ^{18}O in the sample is equal to that in the reference, then $(^{18}O/^{16}O)_{sample}/(^{18}O/^{16}O)_{reference}$ =1, so the bracketed term = 0, and $\delta^{18}O = 0$;
- ^{18}O in depleted in the sample compared to the reference, then $(^{18}O/^{16}O)_{sample}/(^{18}O/^{16}O)_{reference}$ <1, so the bracketed term is -, along with $\delta^{18}O$.

In summary, a sample with a higher $^{18}O/^{16}O$ ratio (enriched in heavier isotope) than the SMOW reference will have a positive (+) δ value. If the $^{18}O/^{16}O$ ratio of the substance is lower (depleted in heavier isotope) than the SMOW reference, the δ value will be negative (-). The δ values of SMOW (O and H isotopes) and PDB (C isotopes) are zero as they are compared to themselves.

Now let's apply this to the analysis of ice and ocean $\delta^{18}O$ values and see how delta values are used as a proxy for temperature or change in temperature at deposition.

The ice shields come from snow which comes from water evaporated from the oceans. Water can have multiple isotopic compositions, but from the % abundancies, the most likely ones are $H_2{}^{16}O$ and $H_2{}^{18}O$, which is heavier than the light form, $H_2{}^{16}O$. $H_2{}^{16}O$ evaporates more readily from the mid-latitudes of the oceans, and when it reaches the poles, condenses to form snow and eventually ice enriched in $H_2{}^{16}O$. Urey showed that the vapor pressure of $H_2{}^{18}O$ is about 1% less that that of $H_2{}^{16}O$ between 46.35°C and 11.25°C.

In addition, the $H_2{}^{18}O$ that evaporates at lower-latitudes is more likely to condense and be removed in rain, leaving the southern oceans enriched in $H_2{}^{18}O$. This effect is actually quite significant as water, in the form of ice, in Greenland and Antarctica has about 5% less $H_2{}^{18}O$ than water at 20^0C from midlatitudes, making the $\delta^{18}O$ for water a proxy for temperature **but even better as measures of ice volume and from that sea levels**.

Now consider the Ice Ages, when oceans were enriched in $H_2{}^{18}O$. On glacial melting of $H_2{}^{16}O$-enriched ice, with the melting flowing into the oceans, the $H_2{}^{18}O$ would get diluted with $H_2{}^{16}O$. At the same time, the salinity of the ocean would decrease since ice condenses without oceans salts. These differences in $H_2{}^{18}O/H_2{}^{16}O$ values in ice core samples are converted to $\delta^{18}O$ values, which can be positive or negative, as shown in Figure 32.2.10.

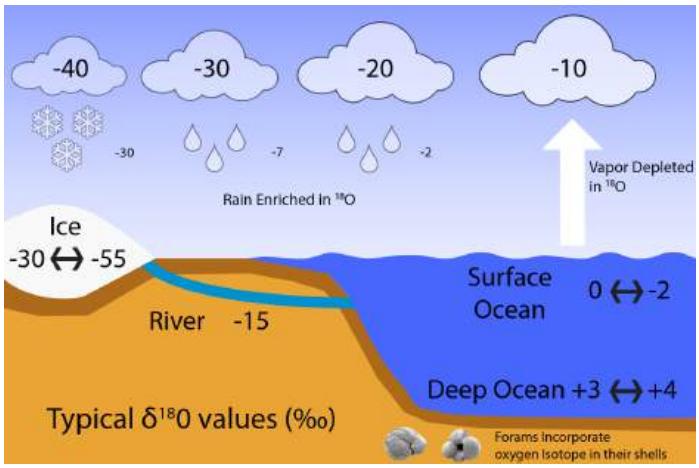


Figure 32.2.10: Typical $\delta^{18}O$ values (in permil). Andreas Schmittner. https://eng.libretexts.org/Bookshelv...A_Paleoclimate

Surface ocean water has $\delta^{18}O$ values of around zero. Due to fractionation during evaporation, less heavy isotopes make it into the air, which leads to negative delta values of around -10 ‰ for the evaporated water vapor. Condensation prefers the heavy isotopes, as described above. In this example, the first precipitation thus has a $\delta^{18}O$ value of about -2 ‰,(more positive than the first vapor). The remaining water vapor will be further depleted in ^{18}O relative to ^{16}O and its $\delta^{18}O$ value become morr negative (-20 ‰). Any subsequent precipitation event further depletes ^{18}O. This process is known as Rayleigh distillation and leads to very low $\delta^{18}O$ values of less than -30 ‰ for snow falling onto ice sheets. Thus, ice has very negative $\delta^{18}O$ of between -30 and -55 ‰. Deep ocean values today are about +3 to +4 ‰. During the last glacial maximum, as more water was locked up in ice sheets, the remaining ocean water became heavier in $\delta^{18}O$ by about 2 ‰. We know this, as we explain below, because foraminifera build their calcium carbonate ($CaCO_3$) shells using the surrounding sea water. Thus they incorporate the oxygen isotopic composition of the water into their shells which are then preserved in the sediments and can be measured in the lab. Bralower and David Bice. https://www.e-education.psu.edu/earth103/node/5. Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License(link is external).

In summary, in cold conditions, Greenland and Antarctic ice is enriched in ^{16}O since $H_2{}^{16}O$ preferentially evaporates and then condenses and freezes into ice a the poles. In addition, as we will see below, deep sea Foraminifera shells contain more ^{18}O in shells since water is enriched $H_2{}^{18}O$ in cold conditiond, since less evaporates. Hence $\delta^{18}O$ becomes more positive.

OCEAN SEDIMENT CORE $^{18}O/^{16}O$ ANALYSES

In the last section, we looked at $\delta^{18}O$ values in ice core layers and their use as proxies for land and ocean temperatures as more fundamentally as measures of ice volume., which affects sea levels as well as ocean salinity. It's amazing what we can surmise about past climate based on the fact that $H_2{}^{16}O$ evaporates more readily than $H_2{}^{18}O$ and that $H_2{}^{18}O$ that did evaporate condenses at low and mid-latitudes more readily than $H_2{}^{16}O$. These two factors lead to the enrichment of $H_2{}^{16}O$ in polar ice and the enrichment of $H_2{}^{18}O$ in low and mid-latitude ocean water. Remember that evaporation and condensation are physical reactions involving a change in state.

Yet ice cores go back only so far in geological time. To go back further, scientists analyze shells buried in ocean sediments. More specifically, they analyze the isotopic ratio of $^{18}O/^{16}O$ in calcium carbonate (calcite, $CaCO_3$) in buried shells of organisms like Foraminifera. (Calcite is a stable anhydrous form of $CaCO_3$, but under high pressure it can form different calcite phases.) Now we are dealing with a new atom, C, in the carbonates, so the interpretation of $\delta^{18}O$ of buried carbonate depends on many factors compared to than $\delta^{18}O$ values of solid and liquid water. It must include these factors

- the chemical reactions of inorganic carbonate formation (precipitation) and dissolution (compared to the physical reactions of water evaporation and condensation)
- the biological formation of $CaCO_3$ in shells
- an understanding of the carbon cycle
- the temperature at which the $CaCO_3$ was deposited
- the salinity at deposition as the formation of $CaCO_3$ from its ions as the overall ionic strength of the medium in which $CaCO_3$ formed would influence the thermodynamics and kinetics of how the separate ions approach each to form the solid
- equilibrium and kinetic controls of the precipitation reactions.

To truly understand biochemistry, we must include both biological and chemical aspects. Biochemistry requires a synthesis of knowledge from many disciplines, including introductory chemistry (where precipitation reactions were likely covered for the first time) and analytic chemistry (which delves more deeply into precipitation reactions). Hence we don't apologize for bringing back your previous knowledge of precipitation reactions. At the same time, our description of the use of $\delta^{18}O$ values in buried ocean sediments is very simplified.

We need to consider two different reactions to understand $\delta^{18}O$ values in calcite Foraminifera shells. The first describes how ^{18}O gets into CO_3^{2-} in the first place. The second describes seemingly simple reactions for the formation of $CaCO_3$ from its ions.

ENRICHMENT OF $CACO_3$ WITH ^{18}O

The incorporation or **fractionation** of ^{18}O from $H_2{}^{18}O$ into calcite shells can be described most easily by the reaction below.

$$\text{Fractionation Reaction: } (1/3)\ CaC^{16}O_3 + H_2{}^{18}O \rightarrow (1/3)\ CaC^{18}O_3 + H_2{}^{16}O$$

In this reaction, the source of ^{18}O comes from the most abundant and likely sources, $H_2{}^{18}O$. A simplified reaction mechanism is shown in Figure 32.2.11.

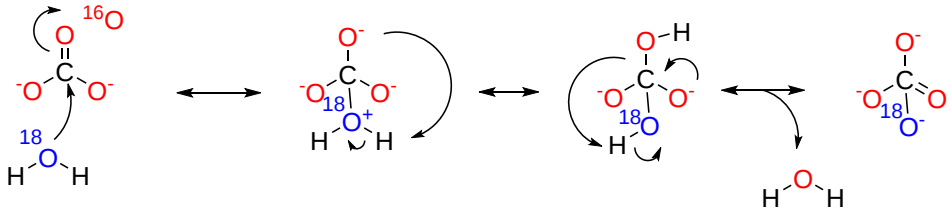


Figure 32.2.11: Incorporation of ^{18}O into carbonate from $H_2{}^{18}O$

More broadly, there would be an exchange of isotopes in the entire **dissolved inorganic pool (DIC)** = $CO_2(aq)$ + H_2CO_3 + HCO_3^- (bicarbonate) + CO_3^{2-} (carbonate) with H_2O.

FORMATION OF $CACO_3$

Two reactions in general describe calcite formation and its growth:

$$Ca^{2+}(aq) + CO_3^{2-}(aq) \leftrightarrow CaCO_3 \tag{32.2.3}$$

and

$$Ca^{2+}(aq) + HCO_3^-(aq) \leftrightarrow CaCO_3 + H^+ \tag{32.2.4}$$

These show the uptake of ^{18}O into $CaCO_3$ should also include a consideration of HCO_3^-.

In a simple and environmentally-controlled in the lab, reaction 2 above can be considered in equilibrium and defined by a K_{sp} value (as you learned in introductory chemistry courses.

In Chapter 4.12, we saw the relationships between K_{eq}, ΔG^0, ΔH^0, DS^0 and temperature. There is an inverse relationship between K_{eq} and temperature.

$$\begin{aligned} \Delta G^0 &= \Delta H^0 - T\Delta S^0 = -RT\ln K_{eq} \\ \ln K_{eq} &= -\frac{\Delta H^0 - T\Delta S^0}{RT} \\ \ln K_{eq} &= -\frac{\Delta H^0}{RT} + \frac{\Delta S^0}{R} \end{aligned} \tag{32.2.5}$$

Assuming that the formation of $CaCO_3$ is in equilibrium, then you would expect that $CaCO_3$ would be enriched in ^{18}O, and have a $+\delta^{18}O$ value. Why? Urey showed that under equilibrium conditions, calcite is enriched in $C^{18}O_3^{2-}$ probably because of lower vibrational energy of the heavier form of carbonate, which favors stabiity and formation of the solid. In addition, the incorporation of ^{18}O is even more pronounced in climatically colder water, which as we have seen, has a $+\delta^{18}O$ value.

Hence in cold periods with large ice shields, $\delta^{18}O$ values from shells of foraminifera living both in the illuminated upper ocean (planktic foraminifera, which engage in photosynthesis) and deep sea benthic foraminifera, are more positive. On ice shield melting, as the $\delta^{18}O$ values of water become more negative, so do the values of $\delta^{18}O$ values of the foraminifera. Benthic foraminifera give a global temperature estimate as deep waters are more homogeneous. Planktic foraminifera $\delta^{18}O$ values are proxies for more local temperatures as they are in a more changing, less mixed environment, and are more affected by evaporation and precipitation.

Yet the formation of $CaCO_3$ in shells in many cases is **not** in equilibrium and is in part determined by the concentration of the reactants, the rate of diffusion of ions into and out of the growing calcite shell, which would also depend on salinity (affecting electrostatic attractions of the ions to the growing crystal), the pH (which affects the ratio of CO_3^{2-} and HCO_3^-) and biological effects (from the mechanism by which shells are formed which at some point may involve HCO_3^- transporters). It also depends on the of transfer of carbonate within the dissolved inorganic carbon pool (DIC). The reaction has been shown to be in equilibrium in some species of foraminifera but not in others.

The term **fractionation** is often used in the isotope and climate literature. Using water as an example, it describes the ratio of heavy to light isotopes of O that partition into the liquid, solid, and gas phases of water. The fractionation factor determines $\delta^{18}O$ value of water. Likewise, there is a fractionation process that determines the partitioning of ^{18}O from water into CO_3^{2-} and $CaCO_3$ during the precipitation of calcite. The fractionation factor α shows how the ratio of the isotopes changes in either a physical (such as a phase transition) or chemical process. It is the factor by which the abundance ratio of two isotopes will change during a chemical reaction or a physical process.

The formation of calcite from HCO_3^- in controlled studies shows that $CaCO_3$ has different oxygen isotope concentrations depending on the initial concentrations of reactants. The size of a shell can also affect the $\delta^{18}O$ of additionally deposited $CaCO_3$. These ideas support the notion that the fractionation of isotopes in $CaCO_3$ occurs through both equilibrium fractionation and kinetic fractionation.

Several different **theoretical "paleotemperature" equations** have been developed to show how temperature T is related to $\delta^{18}O$ in calcite during equilibrium conditions. One theoretical quadratic equation is shown below.

$$\begin{aligned} T &= 16.9 - 4.38\,(\delta_c - A) + 0.10(\delta_c - A)^2 \\ A &= \delta_w - 0.27\% \end{aligned}$$

where T is the temperature in ^oC, δc and δw are the $\delta^{18}O$ of calcite and water, respectively. The 0.27% is just an adjustment factor to convert from the PDB to the VPDB reference values. The equation applies to sea water of normal salinity and freshwater.

A more recent calibration equation for the formation of calcite from HCO_3^- (done with bubbling of the reaction mixture with N_2 for a variety of temperatures) was derived by Kim and O'Neil in 1997. The equation was derived from carefully controlled laboratory studies that apply under equilibrium conditions and is shown below. The Kim and O'Neil equation shows the relationship between the fractionation factor alpha (α) of $^{18}0/^{16}0$ between inorganically precipitated $CaCO_3$ and H_20 as a function of the temperature, and is shown below.

$$1000 \ln \alpha\,(\text{ Calcite- } H_2O) = 18.03\left(10^3 T^{-1}\right) - 32.42 \tag{32.2.6}$$

Alpha is the fractionation factor, and T is in Kelvin. Note: An update of this equation to conform to IUPAC conventions gives 10^3 ln α = 18.04 x 1000 / T - 32.18)

The oxygen isotope fractionation factor alpha between two substances A and B is defined as

$$\alpha = \left({}^{18}O/{}^{16}O\right)_A / \left({}^{18}O/{}^{16}O\right)_B \tag{32.2.7}$$

The left hand side of the equation (1000xlnα) is used for convenience and its relationship to $\delta^{18}O$ values (which are expressed per %o), similar in a way to the use of pKa = -log[K_A] instead of K_A.

Here is an alternative form of the Kim and O'Neil equation expressed in quadratic form.

$$T(^\circ C) = 16.1 - 4.64 \cdot \left(\delta^{18}O_f - \delta^{18}O_w\right) + 0.09 \cdot \left(\delta^{18}O_f - \delta^{18}O_w\right)^2 \tag{32.2.8}$$

A controlled equilibrium study using the cultured foraminifera B. marginata of different sizes at different temperatures was used to develop an experimental equation to compare with the theoretical equations described above. Figure 32.2.12 shows graphs of the empirically-determined equation (nonred lines) vs the theoretical Kim and O'Neil equation (red line).

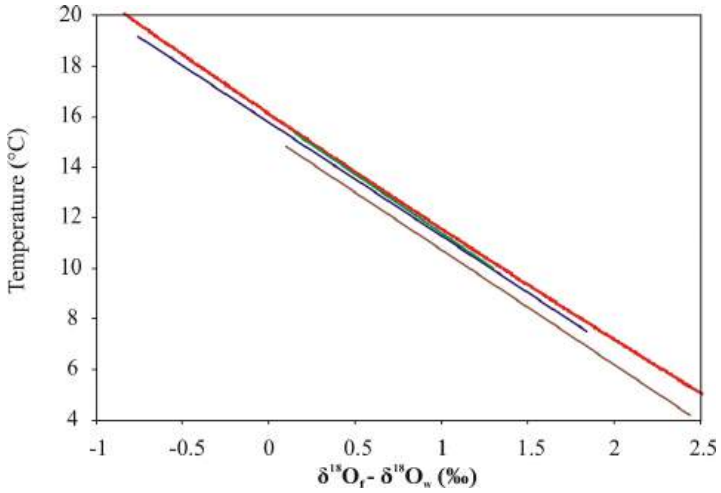


Figure 32.2.12: Comparison of experimental calibration equation with the theoretical equation for equilibrium calcite of Kim and O'Neil (1997). Barras, Christine & Duplessy, J.-C & Geslin, Emmanuelle & Michel, Elisabeth & Jorissen, Frans. (2010). Calibration of δ 18O of cultured benthic foraminiferal calcite as a function of temperature. Biogeosciences. 7. 1349-1356. 10.5194/bg-7-1349-2010. CC Attribution 3.0 License

The brown, blue and green lines represent the calibration equations of cultured B. marginata from ≤ 150, 150–200 and 200–250 µm size fractions, respectively. The quadratic equation derived from Kim and O'Neil (1997) relationship is represented by the red line.

A quick inspection of the empirical equation for different sizes of B. marginata shows the same relationships between T an dδ^{18}O values as shown in Table 32.2.5 below.

	Size fraction	Temperature range	T (°C) = b + a * ($\delta^{18}O_f$-$\delta^{18}O_w$)		R^2
			a	b	
Equation (1)	≤150 µm	4.1-14.7°C	– 4.54 (± 0.14)	15.25 (± 0.17)	0.97
Equation (2)	150-200 µm	7.9-19.3°C	– 4.49 (± 0.13)	15.73 (± 0.14)	0.98
Equation (3)	200-250 µm	10.1-14.7°C	– 4.61 (± 0.37)	16.00 (± 0.33)	0.93
Equation (4)	>250 µm	10.1-14.7°C	– 5.31 (± 0.23)	16.93 (± 0.23)	0.99

Table 32.2.5: Best fit linear plot of temperature T vs ($\delta^{18}O_f$ - $\delta^{18}O_w$) for foraminifera B. marginata vs size, where the subscript f is foraminifera and w is water.

We take this opportunity to reshow the graph that reconstructs changes in planetary temperatures over the last 66+ million years (Figure 32.2.13). The data between 66 MYA and 100,000 years ago (note the change in scale in the x-axis to allow fitting of a large time range in one graph) was obtained, in large part, from δ^{18}O values from deep ocean sediments, while the data from around 100,000 YA to the advent of modern temperature recordings were obtained mostly from δ^{18}O from ice core samples from Antarctica and Greenland. Of course, other temperature proxies, as described above, were important as well.

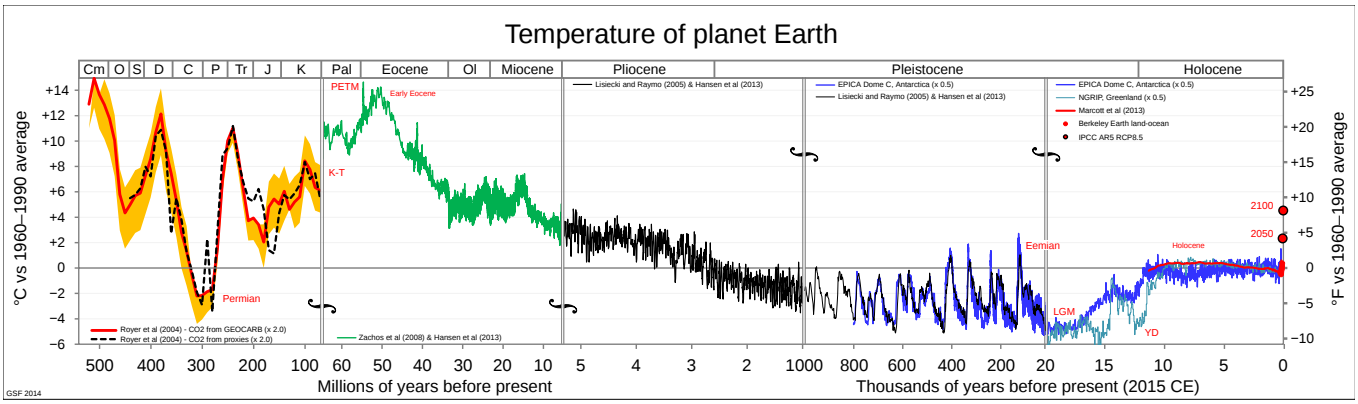


Figure 32.2.13: https://commons.wikimedia.org/wiki/F...alaeotemps.png. (Excel available). Creative Commons Attribution-Share Alike 3.0 Unported

These detailed, but hopefully understandable explanations of the relationship of temperature with $\delta^{18}O$ in foraminifera shells from ocean sediment cores were presented for reasons expressed in the beginning of Chapter 31.2:

- Isotopes and their effects, critical in understanding both structure and activity in biochemistry, are key in climate science.
- Both equilibrium and nonequilibrium reactions and processes apply to isotope partitioning into water and biomolecules, in ways similar to linked biochemical reactions and pathways.
- The study and application of isotope effects can integrate and expand learning from previous science courses

Astute readers will notice that we concentrated on $\delta^{18}O$ values (in water and carbonates) and barely mentioned $\delta^{13}C$ values for carbonate precipitations. We will discuss that in the next chapter section as we consider the carbon cycle.

KEY POINTS - BETA VERSION FROM CHAT.OPENAI

1. Isotope analysis is a technique used to measure the isotopic composition of elements in order to understand the processes and interactions that have occurred in the past, present, and future.
2. Isotopes of carbon, oxygen, and hydrogen can be used to study the effects of climate change on different Earth systems, such as the atmosphere, oceans, and biosphere.
3. Carbon isotopes can be used to study the sources and sinks of CO2 in the atmosphere, and to understand the role of different types of vegetation in the carbon cycle.
4. Oxygen isotopes can be used to study the sources and sinks of water vapor in the atmosphere, and to understand the effects of climate change on precipitation patterns.
5. Hydrogen isotopes can be used to study the sources and sinks of water vapor in the atmosphere and to understand the effects of climate change on the water cycle.
6. Isotope analysis is an important tool for understanding the complex dynamics of the Earth's climate system and for developing effective strategies to mitigate the impacts of climate change.
7. Isotope analysis can provide important information about the Earth's climate and environment and can help scientists understand the causes and impacts of climate change.

32.3: CLIMATE CHANGE - THE CARBON CYCLE AND CARBON CHEMISTRY

Search Fundamentals of Biochemistry

32.3.1: THE CARBON CYCLE

In the last chapter section, we used oxygen isotopes in ice and ocean sediment cores going back millions of years ago to address the history and mechanisms of climate change. We focused on $^{18}\mathbf{O}/^{16}\mathbf{O}$ ratios in $H_2\mathbf{O}$ and in calcite shells ($CaC\mathbf{O}_3$), and their corresponding $\delta^{18}O$ values in ice and sediment cores, to determine CO_2 and temperature over climatic history. Now it's time to talk about the other key atom, **C**, the ratio of $^{13}\mathbf{C}/^{12}\mathbf{C}$ and corresponding $\boldsymbol{\delta}^{\mathbf{13}}\mathbf{C}$ values, not only in $Ca\mathbf{C}O_3$ but also in $\mathbf{C}O_2$ and the organic molecules it transforms into through photosynthesis and the heterotrophic organisms that consume them. $^{13}\mathbf{C}$ partitions not only into inorganic carbon but also into organic molecules throughout life. Hence we need a more detailed understanding of the **carbon cycle**.

The carbon cycle is likely discussed in introductory chapters in biology textbooks, but probably never in chemistry texts. It is fundamental to an understanding of climate, its control and change, and human processes that alter it. Figure 32.3.1 shows a representation of the carbon cycle. Calculated amounts of carbon found in the **lithosphere** (the solid part of the earth), the **atmosphere** (specifically the lower part, the troposphere), and the **hydrosphere** are shown. (The **cryosphere**, the frozen ice found in Greenland and Antarctica, is not shown). The **biosphere** includes part of each of these "spheres" that harbor life. Since life has been shown to exist 10 km down in the crust, we'll refer to the entire region in the diagrams as the biosphere. Figure 1 presents carbon stores in petagrams (10^{15} g) or gigatons of carbon (GtC), as 1 petagram equals 1 billion metric tons (or approximately 1.1 billion US tons).

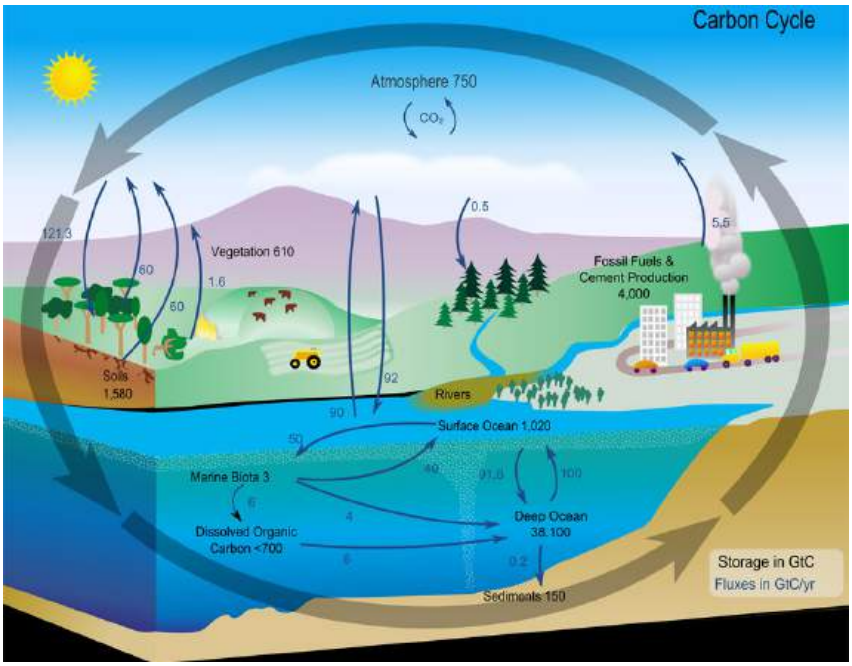


Figure 32.3.1: The carbon cycle. https://commons.wikimedia.org/wiki/F...te_diagram.svg

In addition to the total amount of carbon stored in each region, (GtC), the net changes in carbon per year as it moves into and out of reserves (GtC/yr) are shown in blue arrows with attached numbers.

The exchanges of carbon in the cycle occur at different time scales. Geologically "fast" exchange, on a time scale up to 1000s of years, occurs among the oceans, atmosphere, and land, while a slow exchange (over hundreds of thousands to millions of years) occurs in deep soils, deep ocean sediments, and rocks. We will mostly consider exchanges among the atmosphere, land, and oceans.

CO_2 in the terrestrial biosphere is removed by photosynthesis and returned by respiration by autotrophs like plants, and heterotrophs like microbes that consume soil carbon and plant remains. CO_2 in the atmosphere is also removed by ocean autotrophs like ocean phytoplankton and through partitioning into dissolved inorganic carbons (DIC) molecules like HCO_3^- and CO_3^{2-} into the oceans.

Before we probe some relevant reactions within it, let's look at the big picture and perhaps the most relevant to our climate crisis - the factors causing our increasing CO_2 atm and global warming. To do that, we must put numbers on the cycle to quantify it.

32.3.2: QUANTITATING THE CARBON CYCLE

In Chapter 31.1, we used parts per million (ppm) as a unit for expressing the amount of CO_2 in the atmosphere. Table 32.3.1 below shows how to translate the percentage (parts per 100) for each component gas in the atmosphere (with which you are familiar) into ppm.

Gas	% (parts per 100) in atm	part per million
$\mathbf{N_2}$	78.09	780,900
$\mathbf{O_2}$	20.94	209,400
Ar	0.93	9300
$\mathbf{CO_2}$	0.0415	415

Table 32.3.1: Unit conversion - % to ppm

Climate scientists use ppm instead of concentration (in molecules/m^3) since they wish to know the relative percent or ppm increase with time, which does not depend on temperature and pressure. In contrast, concentration does depend on T and P, as you will remember from the ideal gas law, PV=nRT or n/V=P/RT that you studied in introductory chemistry.

It is important to use dimensional analyses to interconvert units as well. Table 32.3.2 below shows conversion factors to switch between GtC, Gt CO_2, and ppm.

Convert from	to	conversion factor
GtC (Gigatons of carbon)	ppm CO_2	divide by 2.124
GtC (Gigatons of carbon)	PgC (Petgrams of carbon)	multiply by 1
Gt CO_2 (Gigatons of carbon)	GtC (Gigatons of C)	divide by 3.664 = 44.01/12.01)
GtC (Gigatons of carbon)	MtC	multiply by 10000

Table 32.3.2: Unit conversion - GtC and Gt CO_2

To be more technical, atmospheric CO_2 concentrations are expressed in mol fraction of CO_2 in the dry air atmosphere. The ppm for CO_2 is hence µmol CO_2 per mole of dry air.

We have to put numbers on the components of the carbon cycle to quantitatively analyze changes in its components, otherwise, we can't know what is presently happening nor will we be able to predict with some certainty the future. Stoichiometry and reaction kinetics are probably the least liked parts of chemistry for many, but they are critical in understanding climate change We have to apply them on a global scale. Two key terms are important:

Stocks or reserves: How much carbon (mass in Gigatons or petagrams) is stored in given locations in the biosphere. This allows us to understand what % of all carbon stocks are in the ocean, for example. Stocks are usually reported as gigatons of carbon (GtC), not gigatons of carbon dioxide, since many stocks (like fossil fuels) consist of mostly C and H without oxygen. As in stoichiometry calculations in introductory chemistry courses, GtC in the atmosphere can be converted to gigatons of CO_2 by using dimensional analysis.

Fluxes (rates): How much carbon is transferred from one reserve to another per year (Gigatons/yr). Climate scientists are simply applying the Law of Mass Conservation that you learned in introductory chemistry, to the entire biosphere.

Figure 32.3.2 shows the reserves/stock (GtC) for reserves and **decadal (2012-2021) average** fluxes (large and small arrows, GtC/yr) for individual or aggregated stocks.

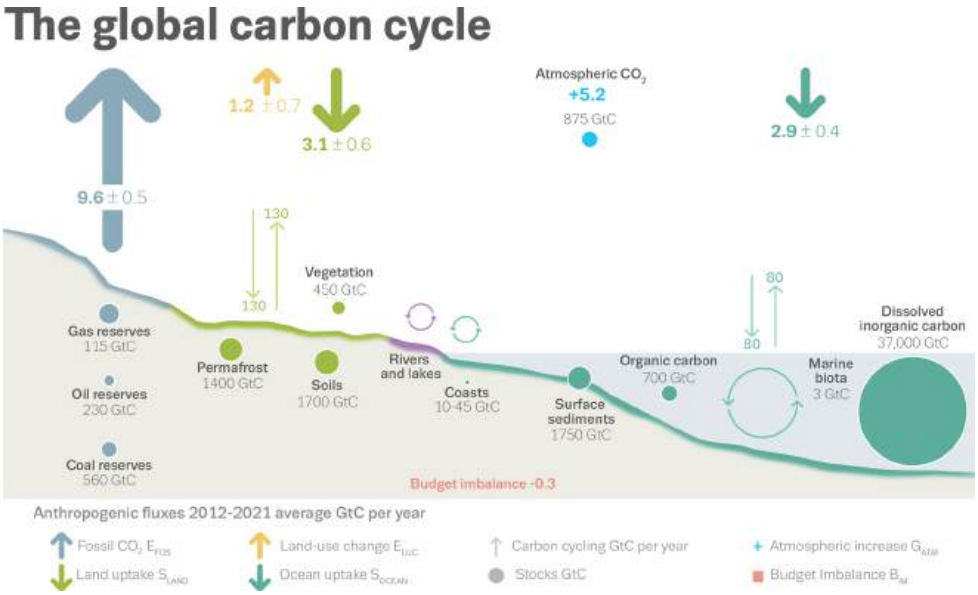


Figure 32.3.2: Schematic representation of the overall perturbation of the global carbon cycle caused by anthropogenic activities **averaged globally for the decade 2012–2021**. E represents emission and S "sink". Earth Syst. Sci. Data, 14, 4811–4900, 2022. https://doi.org/10.5194/essd-14-4811-2022. © Author(s) 2022. This work is distributed under the Creative Commons Attribution 4.0 License.

The abbreviations used are: E_{FOS} (emissions, fossil fuels), E_{LUC} (emissions land use changes - mostly deforestation), S_{LAND} (terrestrial CO_2 sink), S_{OCEAN} (ocean CO_2 sink), G_{ATM} (Growth Rate CO_2 atm), B_{IM} (carbon budget imbalance). Uncertainties are also shown except for the atmospheric CO_2 growth rate which is known precisely and accurately through modern measurements. (It's also the easiest to measure). Human (anthropogenic changes) occurs on top of the carbon cycle.

The upward arrows indicate release into the atmosphere and the downward arrows the absorption in the oceans and land. The thickness of the arrows gives a relative measure of the size of emission or absorption. The thickest arrow and highest value (9.6 GtC/yr) is for the anthropogenic emission of carbon from our use of fossil fuels. Think about that! **Humans are presently the biggest contributor to the carbon cycle**. Before the industrial revolution, human contributions were minimal.

IT'S WORSE THAN THAT!

Figure x above shows that on average, 9.6 GtC/yr was released from fossil fuel use between 2012-2021. The actual flux of anthropogenic carbon release **in 2021 was 9.9 GtC**, equivalent to 36.4 Gt CO_2.

If you add the up arrows and subtract from that sum the down ones, you get +4.8 GtC/yr. This represents the net average increase in GtC in atmosphere CO_2 per year for 2012-2021. That is close to the accurately, and precisely known value of +5.2 GtC/yr increase from the CO_2 we pour into the atmosphere through our use of fossil fuels. Hence the figure above is a bit out of balance (about 0.3 GtC too low - the B_{IM} carbon budget imbalance), but given the difficulty in calculating these values, it is remarkably close to "mass balance" as you learned in introductory chemistry classes. In general, before the industrial revolution, the sum of the fluxes leading to the addition of CO_2 into the atmosphere was equal to the sum of the fluxes that removed it. That is, the system was in a **steady state.** That is no longer the case.

Figure 32.3.3 shows a breakdown of the factors contributing to annual (left) and cumulative (right) fluxes of carbon (GtC/yr), a metric for CO_2 flux, over time since 1850.

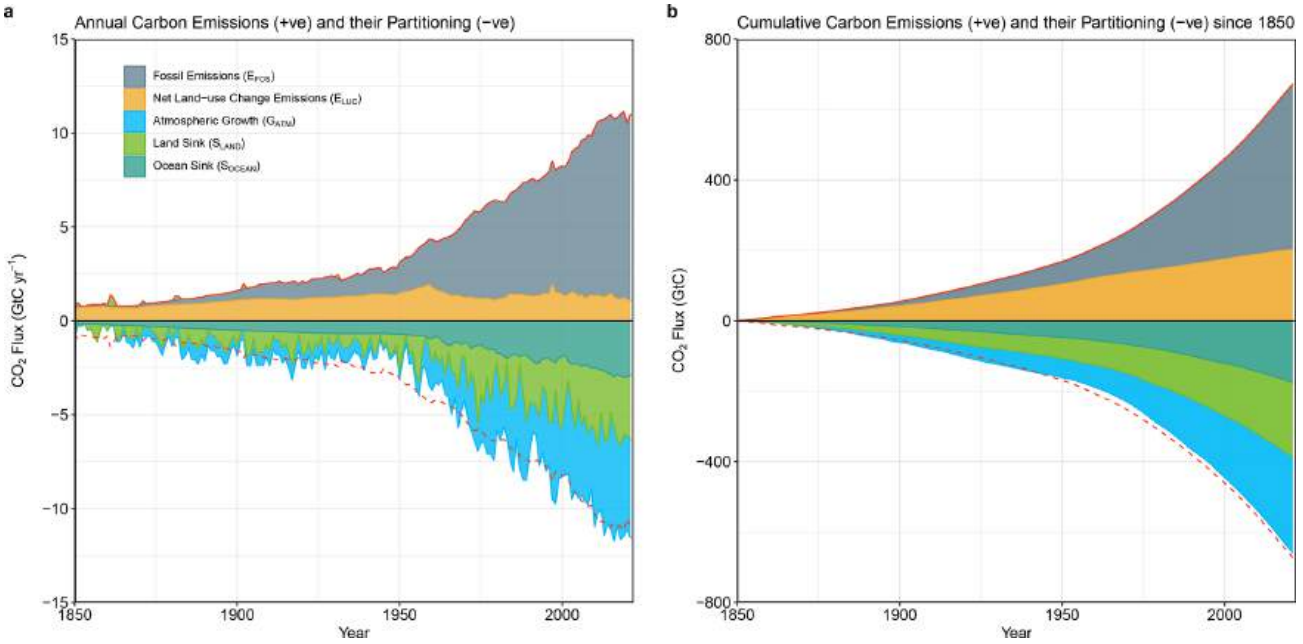


Figure 32.3.3: Combined components of the global carbon budget as a function of time for fossil CO_2 emissions (E_{FOS}, including a small sink from cement carbonation; grey) and emissions from land-use change (E_{LUC}; brown), as well as their partitioning among the atmosphere (G_{ATM}; cyan), ocean (S_{OCEAN}; blue), and land (S_{LAND}; green). Panel **(a)** shows annual estimates of each flux (GtC yr^{-1}, and panel **(b)** shows the cumulative flux (the sum of all prior annual fluxes) since the year 1850. Again, the graph shows GtC not Gt CO_2. . © Author(s), ibid

You might ask why the atmospheric growth in CO_2 (shown in green) is negative. We'll answer that question below.

Lastly, let's think about the total cumulative changes in GtC released and absorbed since 1850 (pre-US civil war and before the big release of CO_2 in modern times). Those data are shown in a bar graph in panel A of Figure 32.3.4. The bar graph in the right panel shows the mean decadal averages that are shown in Figure 2 above.

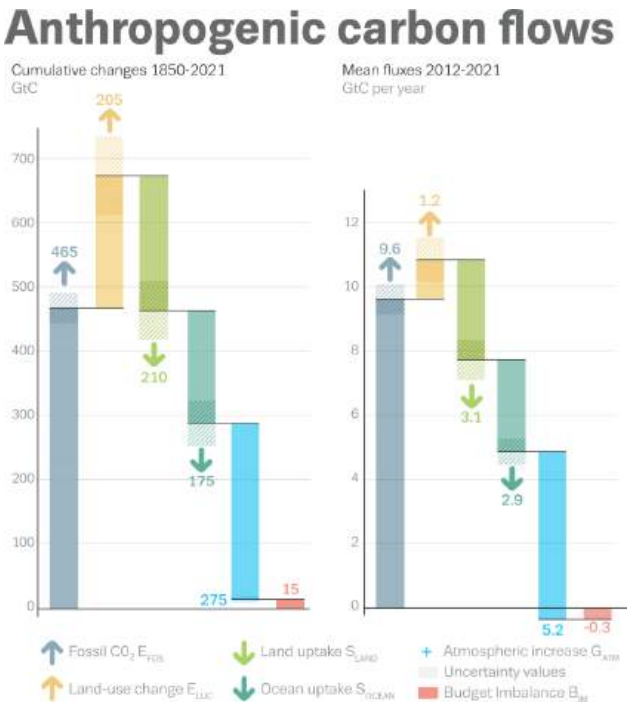


Figure 32.3.4: Total cumulative changes in GtC released and absorbed since 1850 (panel A) and mean decadal fluxes (panel B). E_{FOS} (emissions, fossil fuels), E_{LUC} (emissions land use changes - mostly deforestation), S_{LAND} (terrestrial CO_2 sink), S_{OCEAN} (ocean CO_2 sink), G_{ATM} (Growth Rate CO_2 atm), B_{IM} (carbon budget imbalance). © Author(s), ibid

The positive emission and negative absorption contributions are easy to see in the bar graph. The blue bar represents the net emission of carbon from fossil fuels and fills the gap to complete mass balance as we discussed above. It also explains the negative blue region in Figure 3. Just keep in mind that the blue net flux from fossil fuels is positive.

The cumulative contributions from fossil fuel emissions required to close the gap and fulfill mass balance is is +275 GtC, which when multiplied by the conversion factor (1ppm/2.124 GtC) translates into a 129.5 ppm increase in atmospheric CO_2 over that time. This is very close to independent measurements of a rise of 129.3 ppm (14.7-284.7) over that time.

The data from Figure 32.3.4 has been entered in the first four columns of Table 32.3.3 below.

Source	Subtype	Stock reserves (GtC)	J (Fluxes) GtC/yr (avg 2012/2021) + emission - absorption	$J=k_{app}$[stock]
Atmosphere	-	875		
Buried Fossil Fuels	Coal	560	+9.6	J=+9.6=k[905] k=0.0106
	Oil	230		
	Gas	115		
Terrestrial	Permafrost	1,400	+1.2 (Land use Δ) -3.1 (Land uptake)	J_{use}=+1.2=k_{use}[3550]; k_{use}=0.000338 J_{up}=-3.1=k_{up}[875]; k_{up}=0.00354
	Soil	1,700		
	Vegetation	450		
Oceans	Coasts	10-45	-2.9	J=-2.9=k[875] k=0.00331
	Ocean Surface Sediments	1,750		
	Organic carbon	700		
	Marine Biota	3		
	Dissolve Inorganic Carbon (DIC)	37,000		

We can use this data to develop our own crude computational model predicting future CO_2 emissions using Vcell, the program we used to produce time course (concentration vs time) graphs for both simple and coupled signal transduction reaction pathways.

Vcell can be used to calculate fluxes (J) in reaction pathways, where J is the change in concentrations of a species with time, given the initial concentration or amount of a reactant, and the rate constants affecting its production or removal. If we use the amount of carbon (GtC) in each reservoir in the biosphere and crust as a relative measure of "concentration" and the known fluxes (GtC/yr) for the transfer of carbon to and from the atmosphere as given in Table 3, we could calculate an "apparent rate constant for each flux using this equation:

$J_{stock} = k_{app}$[stock] (where stock is given in GtC).

$$J_{stock} = k_{app}[\text{stock in GtC}] \tag{32.3.1}$$

These "apparent" rate constants are needed to run the Vcell simulation that can reproduce the actual fluxes shown in Figures 2 and 4). The simulation can be run over time

A simple four-term model based on Figures 2 and 4 is shown below. Run the simulation and see how atmospheric CO_2 changes with time. This model is offered only to show how climate models are made and used, and also for fun. The graphs are valid and sound based on the input parameters, but the outputs are based on many assumptions that vastly simplify the model.

The simulation shows CO_2 atm levels peaking at about 982 **GtC** in 51 years (2073) from its average decadal (2011-2021) value of 875. That is an increase of 107 **GtC** over now (50 ppm CO_2 rise from the present 414 to 463 ppm). Over a 50-year period, this gives an average annual rise of 2.14 GtC/yr or about 1 ppm CO_2/yr. A comparison of the predicted atmospheric CO_2 (ppm) levels through 2100 for the IPCC SSP1-2.6 scenario (**blue**) and simple Vcell model (**red**) is shown in Figure 32.3.5.

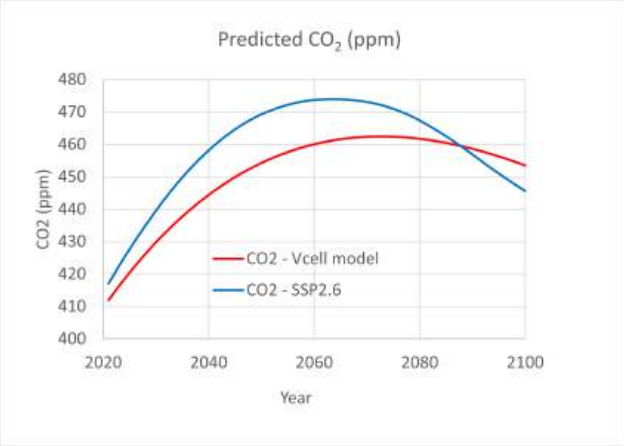


Figure 32.3.5: Predicted atmospheric CO_2 (ppm) for SSP1-2.6 scenario (**blue**) and simple Vcell model (**red**)

SSP1-2.6 data - History: Meinshausen et al. GMD 2017 (https://doi.org/10.5194/gmd-10-2057-2017); Future: Meinshausen et al., GMD, 2020 (https://doi.org/10.5194/gmd-2019-222). https://climateanalytics.org/media/g...-3571-2020.pdf. https://gmd.copernicus.org/articles/13/3571/2020/

However imperfect the Vcell model is (incorrect assumptions, lack of complexity and feedback mechanisms, etc), the results shown above are remarkably close to the projected increases in carbon dioxide in ppm described in IPCC reports for the SSP1-2.6 socioeconomic pathways, shown in the right panel (dark blue line) of Figure 32.3.6. This pathway predicts a rise of approximately 1.8^0 C in average global temperatures.

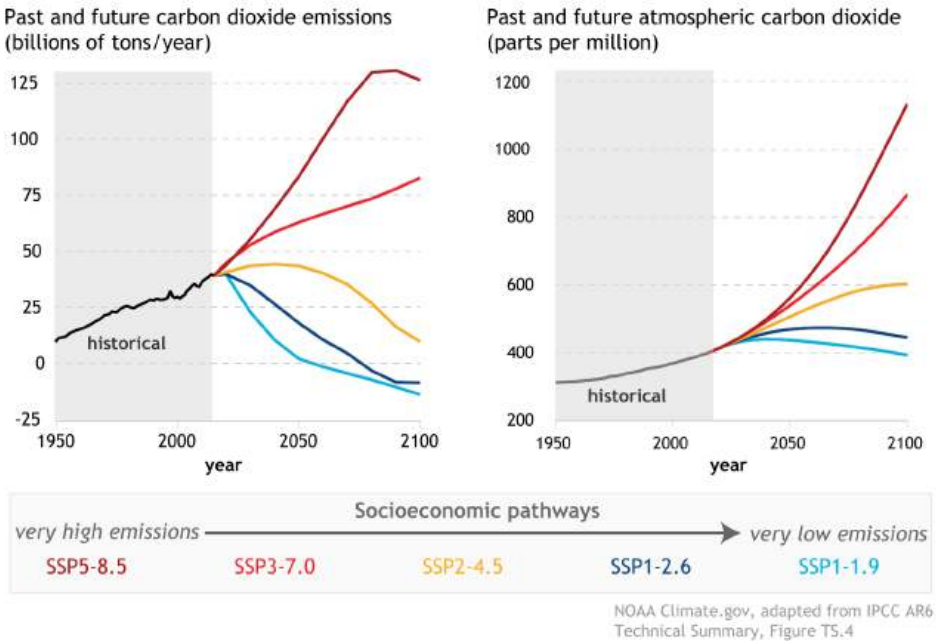


Figure 32.3.6: IPCC, 2021: Summary for Policymakers. In: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)].

Again, remember that the model is based on a ten-year average of CO_2 emissions. Think of all the other assumptions in this model (other than the stock reserves and fluxes) that would give higher or lower values of future CO_2 levels. One major one is that flux values are all held constant to allow calculations of the apparent rate constants for Vcell use. The model depletes much of the fossil fuel reserve. In addition, CO_2 emissions in 2021 were actually 9.9 GtC/yr and going up!

In addition, a change in one parameter can affect the others. For example, the net uptake of atmospheric CO_2 into the land and oceans has increased from 1960-2010, which makes sense given increased CO_2 in the air forcing additional uptake (think LaChatelier's Principle). The oceans have taken up nearly 40% of the CO_2 from fossil fuel use since the industrial revolution. If the rate of uptake decreases (i.e. if we start to saturate the uptake into oceans), CO_2 accumulation in the atmosphere would accelerate. Data also suggest that if we successfully decrease CO_2 in the atmosphere, the oceans would respond by decreasing uptake as well, which would slow the progress in reducing temperatures.

An interesting example relating atmospheric and ocean CO_2 occurred from 1990-2000 when it has been shown that the ocean acted as a weaker sink. This occurred because of a decreasing gradient (the Δ or"effective concentration differences") between atmospheric CO_2 and ocean "CO_2", which decreased the ability of the ocean to act as a sink for CO_2. You can decrease the Δ in two ways:

- decreasing the rate of entry of CO_2 into the atmosphere from fossil fuel use. There indeed was a temporary slowdown in this decade.

- paradoxically, by briefly making the ocean in a shorter term a better sink. This happened in 1991 after the eruption of Mt. Pinatubo, which led to decreased air and ocean temperatures. CO_2 is a nonpolar gas, which has higher solubility in water at lower temperatures (think about soda). This was a short-term and more minor effect than the decreased rate of fossil fuel emissions.

More complex models with more terms for emissions and absorptions of CO_2 can be made. One is shown in Figure 32.3.7. This model adds CO_2 release from the soil through respiration by microorganisms, as well as from plant respiration (CHO to CO2atm). Another term has been added for release by oceans.

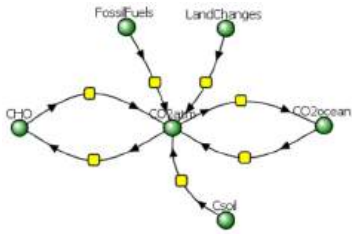


Figure 32.3.7: More complicated Vcell climate model.

Fortunately, we don't have to rely on these simple models to predict future trends in temperature and CO_2. A complex dynamic model simulator that is in accord with many different climate models is available at your fingertips. Developed at MIT and Climate Interactive, and available for free from any web browser, the EN-ROADS program allows users to change sliders for key inputs and see future predicted temperature and CO_2 levels. In accordance with RCP and SSP IPCC pathways that tie future emissions to socio-economic policies (discussed in Chapter 31.1), the program allows users to change variables such as carbon pricing and incentives to move to clean energy in transportation, building and energy supplies sectors. Access the program directly from this page by clicking the Close icon in the program window in Figure 32.3.8 below.

Figure 32.3.8: EN-ROADS global climate simulator

Here is also an external link to the En-Roads global climate simulator (Developed by Climate Interactive, the **MIT Sloan Sustainability Initiative**, and **Ventana Systems**)

Move the interactive sliders and see the effect on greenhouse gas emissions and global temperatures. Here is a link to a one-page tutorial on its use.

THE SKEPTIC'S CORNER: CLIMATE CHANGE MISINFORMATION

We should all be skeptical of models, especially ones that predict changes 80 or more years into the future. We gain confidence in a model if it accurately fits data going back in time and into the future data as well. We mentioned in Chapter 31.1 that oil company scientists knew of the likely climatic effects arising from fossil fuel emissions, but the company executives did not act on their models. Their models were startlingly accurate as shown in Figure 32.3.9 below, which shows their predictions for both CO_2 levels and the associated increases in temperature caused by them.

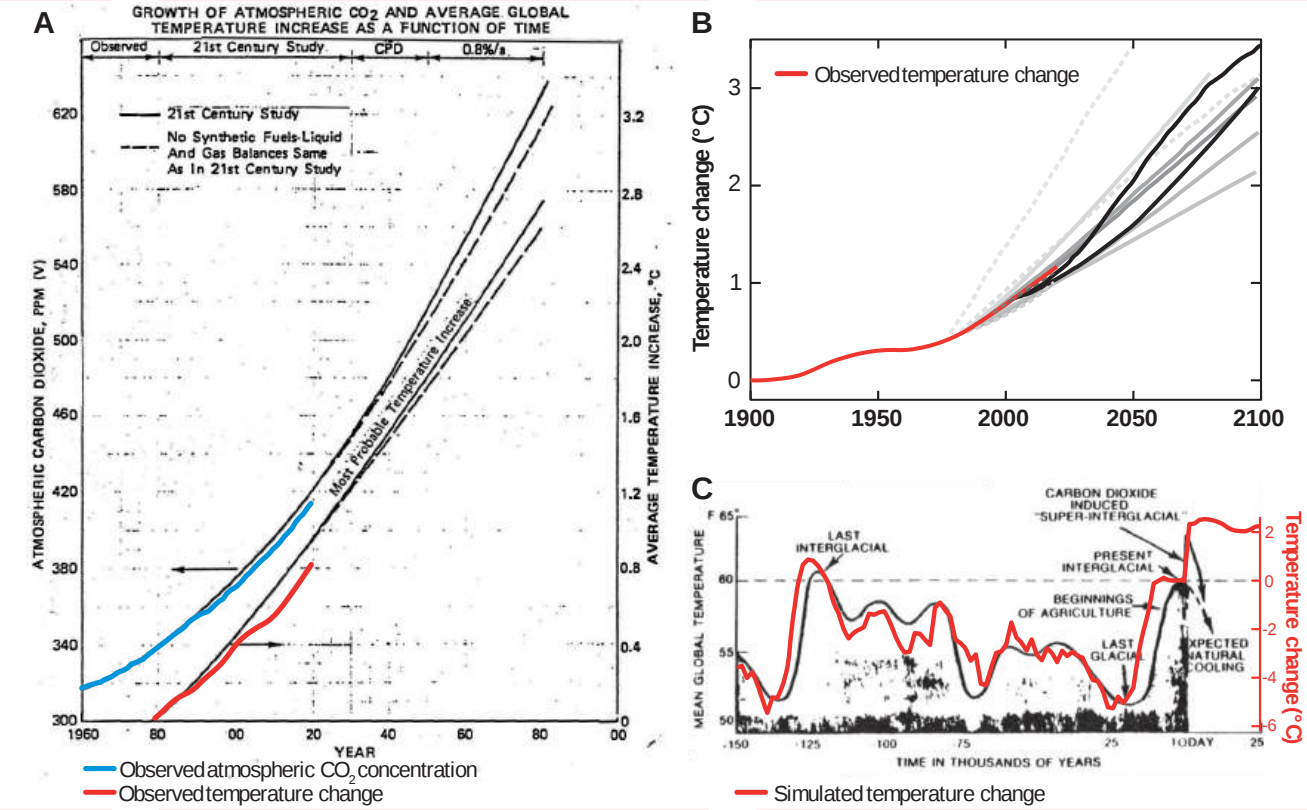


Figure 32.3.9: Historically observed temperature change (red) and atmospheric carbon dioxide concentration (blue) over time, compared against global warming projections reported by ExxonMobil scientists. Supran, G., Rahmstorf, S., and Oreskes, N. Assessing ExxonMobil's global warming projections. Science (2023). https://www.science.org/doi/abs/10.1...cience.abk0063. Reprinted with permission from AAAS. Not for reuse.

Panel (A) shows "Proprietary" 1982 Exxon-modeled projections.

Panel (B) shows a summary of projections in seven internal company memos and five peer-reviewed publications between 1977 and 2003 (gray lines).

Panel (C) shows a 1977 internally reported graph of the global warming "effect of CO2 on an interglacial scale." (A) and (B) display averaged historical temperature observations, whereas the historical temperature record in (C) is a smoothed Earth system model simulation of the last 150,000 years.

As these graphs clearly show, oil companies knew since the late 1970s, over 40 years ago, of the climatic effects of CO_2 emissions. They could even predict the temperatures since the last ice age. In the 70s, solar and wind energy were much more expensive to produce and use than now, but if we had subsidized their development back then as we have done for decades for the fossil fuel industries, our climatic situation now would be much less precarious. Figure 32.3.10 below shows worldwide fossil fuel subsidies in US $billion and in % global GDP from 2015 to 2020 and projections after that.

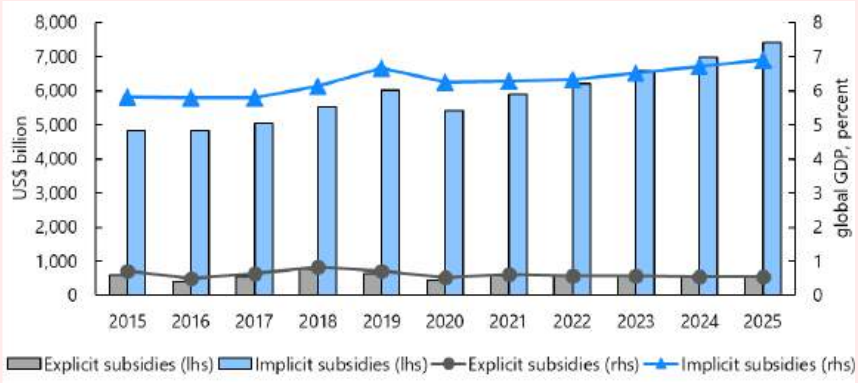


Figure 32.3.10: Worldwide subsidizes in US $billion and in % global GDP. Bar graphs are for US$biillons and the circles and triangles for % global GDP. IMF. ttps://www.imf.org/en/Publications/W...bsidies-466004

The subsidies are broken down into explicit subsidies (tax breaks or direct payments to help fossil fuel companies to fund their uncompensated costs) and implicit ones (undercharging for environmental costs of fossil fuel use that the oil companies don't pay). These latter "hidden" costs are passed down to countries, states, and individuals. In 2020, global subsidies were $5.9 trillion or 6.8% of the world's GDP. The explicit subsidies given to fossil fuel companies, about 8% of the total, amounted to $472 billion just in 2020!

"Company executives chose to publicly denigrate climate models, insist there was no scientific consensus on anthropogenic climate change, and claim the science was highly uncertain when their own scientists were telling them the opposite" (ref). They also propagated the myth that the global climate was actually cooling. This is a powerful and unsettling example of disinformation with enormous consequences.

Now that we have seen the big picture, let's look at how carbon moves through various pools of carbon-containing molecules. We have already discussed photosynthesis in great detail in Chapter 20, so we fill focus our attention more on dissolved inorganic carbons (DIC) including species such as HCO_3^- and CO_3^{2-}. Another view of the carbon cycle that includes weathering of rocks to produce silicates and bicarbonates, and the formation of shells in the ocean from HCO_3^-, CO_3^{2-} and silicates, is shown in Figure 32.3.10.

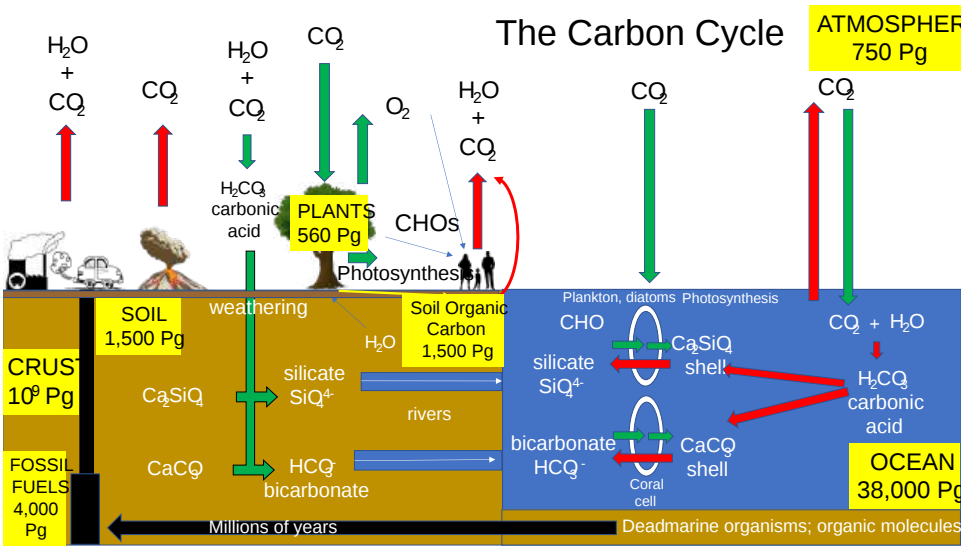


Figure 32.3.10:

Let's focus on the oceans first. The reversible movement of CO_2 from the atmosphere to the oceans, $CO_{2\ atm} \leftrightarrow CO_{2\ ocean}$, depends on the difference in the partial pressures of CO_2 (ΔpCO_2) in the atmosphere and surface waters. The reaction is also driven to the right by the removal of CO_2 (aq) as it forms carbonic acid (H_2CO_3), which then forms bicarbonate (HCO_3^-) and carbonate (CO_3^{2-}). These coupled reactions chemically buffer ocean water, thus regulating ocean pCO_2 and pH.

pCO_2 is not homogenous in ocean surface waters and varies with different conditions of current and temperature. CO_2 can be more readily released from upwellings of nutrient-rich and warm waters, especially in the tropics. In cold Northern waters and also in the Southern Ocean, where water sinks, it is taken up from the atmosphere (again CO_2 is more soluble in cold water).

As we discussed in Chapter 31.1, the **ocean chemistry of CO_2 determines in large part the levels of atmospheric CO_2.** The coupled reactions of CO_2 in the oceans are shown below.

$$CO_2(\text{ g, atm}) \leftrightarrow CO_2(\text{aq, ocean}) \tag{32.3.2}$$

$$CO_2(\text{aq, ocean}) + H_2O(\text{I, ocean}) \leftrightarrow H_3O^+(\text{aq}) + HCO_3^-(\text{aq}) \tag{32.3.3}$$

$$H_2O(\text{I}) + HCO_3^-(\text{aq}) \leftrightarrow H_3O^+(\text{aq}) + {CO_3}^{2-}(\text{aq, sparingly soluble}) \tag{32.3.4}$$

These reactions should be familiar to all chemistry students and were presented previously in Chapter 31.1 and in Chapter 2. A significant contributor to ocean bicarbonate is **weathering** of rocks like limestone. and marble, which are both forms of $CaCO_3$. The relevant reactions are shown below.

$$\begin{aligned} &CaCO_3(\text{ s}) + H_2O \leftrightarrow Ca^{2+}(\text{aq}) + {CO_3}^{2-}(\text{aq}) \\ &{CO_3}^{2-}(\text{aq}) + H_2O \leftrightarrow {HCO_3}^-(\text{aq}) + OH^-(\text{aq}) \end{aligned} \tag{32.3.5}$$

CO_2 is nonpolar and not very soluble in water. Either is CO_3^{2-} in the presence of divalent cations like Ca^{2+}. However HCO_3^- is and can be considered a "soluble" form of carbon. This soluble form from terrestrial weatherings ends up in rivers and eventually enters the ocean. It is also the form of carbonate that is transferred into cells by anion transporters for eventual shell formation. HCO_3^- is also a chief regulator of both blood and ocean pH. Weathering is slow compared to anthropogenic emissions of CO_2 from fossil fuel use, but it is nevertheless a key player in the carbon cycle and the regulation of ocean pH.

The same weathering reactions on silicate rocks lead to the transfer of silicate ions into rivers and then into the ocean, where they can be taken up by diatoms in the formation of $CaSiO_4$ shells. As the oceans take up more CO_2, they become more acidic, which leads to the equivalent of "weathering" of shells of living organisms, leading to their potential death. Silicon is directly underneath carbon in the periodic table so the following simplified reaction is analogous to those we seen with CO_2 and its inorganic ions.

$$H_4SiO_4 = SiO_2 + 2H_2O \tag{32.3.6}$$

H_4SiO_4 is silicic acid.

32.3.3: $^{13}C/^{12}C$ RATIOS IN ICE CORE AND OCEAN SEDIMENTS

We are now in the position to explore how isotopes of carbon can be used for more than radio- ^{14}C dating, which is quite limited in climate studies. ^{13}C, a stable isotope of carbon, however, is extremely useful because C-^{13}C bond dynamics are influenced by it. Reaction rates are affected by the presence of ^{13}C when C-C bonds are made or cleaved. This isotope effect leads to different $^{13}C/^{12}C$ ratios in reactants and products, and hence different $\delta^{13}C$ values.

Isotopes have a long history in the study of biochemical reactions. The k_{cat} and k_{cat}/K_M values for enzyme-catalyzed reactions can be affected if the rate-limiting step involves cleavage or the creation of a C-^{13}C, C-D (deuterium) or C-T (tritium) bond. Substrates labeled with the isotopes have similar transition state energies for the formation/cleavage of a bond involving an isotope, but the ground state vibrational energy for the isotope-substituted atom are proportionately lower, as illustrated in Figure 32.3.11.

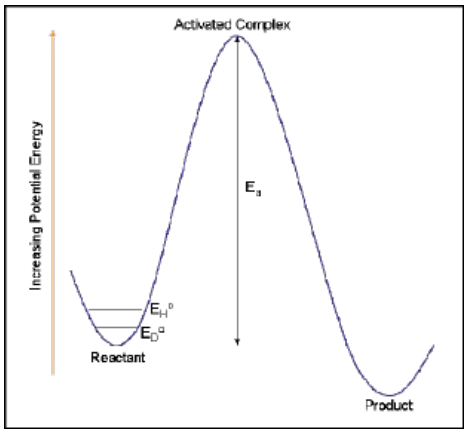


Figure 32.3.11: Kinetic Isotope Effects.

This primary **kinetic isotope effect** leads to higher activation energy for the formation/cleavage of a bond with the isotope. For C-D and C-T bond cleavages that are rate-limiting, the rates are 7X and 16X slower than the cleavage of a C-H bond, respectively. Cleavage or formation of bonds to heavy isotopes of carbon, oxygen, nitrogen, sulfur, and bromine have much smaller isotope effects (between 1.02 and 1.10). The difference in the magnitude of the kinetic isotope effect is directly related to the percentage change in mass. Large effects are seen when hydrogen is replaced with deuterium because the percentage mass change is very large (mass is being doubled). .

Hence the kinetic isotope effect is at play in carbon fixation in photosynthesis, for example. This is evidenced by the observation that the $^{13}C/^{12}C$ ratios are lower in plants than in the atmosphere, showing that $^{12}CO_2$ is preferentially "fixed" in the ribulose bisphosphate carboxylase/oxygenase reaction in plants and other photosynthetic organisms. Also, $^{12}CO_2$, a lighter molecule, has a faster rate of diffusion through the stromata, regulated pores in leaves that facilitate the passage of CO_2, O_2 and H_2O.

In Chapter 31.2, we discussed the use of $\delta^{18}O$ values in ice core and ocean core sediments for measuring past CO_2 and temperatures.

$$\delta^{18}O = \left[\frac{\left(\frac{^{18}O}{^{16}O}\right)_{\text{sample}}}{\left(\frac{^{18}O}{^{16}O}\right)_{\text{reference}}} - 1 \right] * 1000$$

$\delta^{18}O$ values for ice core water samples werer easier to interpret than $\delta^{18}O$ values for $CaCO_3$ sampls, since the deposition of ice is a simple physical process compared to the complexity of the deposition of $CaCO_3$ in ocean sediments, which depends on chemical reactions and nonequilbrium processes (as described in Chapter 31).

Climate scientists can determine and use $\delta^{13}C$ values. An analogous equation for it is shown below.

$$\delta^{13}C = \left[\frac{\left(\frac{13}{12}C\right)_{\text{sample}}}{\left(\frac{13}{12}C\right)_{\text{reference}}} - 1 \right] * 1000$$

As for using $\delta^{18}O$ in carbonate samples, using $\delta^{13}C$ is more difficult as well. The shells of ocean sediment foraminifera were made from dissolved inorganic carbon (DIC) in the ocean at the time so their $\delta^{13}C$ values reflect that. However, shell formation is not a simple equilibrium process since biological shells are formed rapidly so kinetic effects in carbonate and hence isotope fractionation are important. In addition, the biochemistry of shell formation is complicated.

In the open ocean, planktic foraminifera are perhaps the most important marine organism that forms shells given that they produce and export into the ocean about 2.9 Gt $CaCO_3$/yr. Their shells form in a process involving many metastable calcite phases. It starts with a soft template that contains Mg^{2+} and Na^+ ions which play a key role in crystallization. Growth occurs by successive additions of "chambers" to the shell. An F-actin mesh, which forms microtubular structures, leads to the formation of protective envelopes for chamber formation. The layered templates sequester and help control the mineralization of shells and separate bulk sea water for a more intracellular vs extracellular process for biomineralization. Seawater containing minerals becomes vacuolized in a process which for some foraminfera excludes a competing cation, Mg^{2+}. In addition, both Ca^{2+} and HCO_3^- transporters are required. This all combines to form an environment low in Mg^{2+} and supersaturated in Ca^{2+} and CO_3^{2-} for calcite formation. The kinetic fractionation of ^{13}C isotopes into shells is also different than for ^{18}O isotopes since the "pool" of oxygen in the oceans is much greater than carbon. Likewise, the $\delta^{13}C$ is more location-dependent that the $\delta^{18}O$.

Buried organic matter can also be studied. The $\delta^{13}C$ value for buried organic matter depends on primary productivity on land and in the oceans. As mentioned above, autotrophs preferentially take up $^{12}CO_2$. Heterotrophs that eat them also become enriched in ^{12}C. Hence organisms have negative $\delta^{13}C$ values, typically

around −25‰, with the number depending on pathways of incorporation and metabolism. Methane in hydrates in the ocean can be either biogenic, made by methanogens, for example, at low temperatures, or thermogenic, made during high-temperature reactions. Biogenic methane has a $\delta^{13}C$ of around - 60 ‰, while thermogenic methane has a value of around −40‰. Terrestrial plants have different δ^{13} values. $\delta^{13}C$ in C_4 plants range from -16 to −10‰ while for C_3 plants they range from −33 to −24‰.

Changes in $\delta^{13}C$ in ice cores and ocean sediments are used in climate studies. Sometimes it's confusing to understand the cause and effect of the changes. This following explanation for changes in the already negative values of $\delta^{13}C$ might offer help to those with a chemistry-centric view of biochemistry who struggle with mass balance outside of simple chemical equations.

Under climatic conditions, when there is an abundance of **terrestrial** plants that lock in and sequester atmospheric $^{12}CO_2$, the atmosphere becomes depleted in $^{12}CO_2$ and correspondingly enriched in $^{13}CO_2$. Hence primary production (fixing of carbon and anabolic metabolism) by photoplankton in the oceans, under robust growing conditions, would sequester more ^{13}C, causing an **increase in $\delta^{13}C$** (i.e. more positive) values for buried organic and calcite sediments.

During times of mass extinction, when terrestrial plant primary production drops precipitously, the $\delta^{13}C$ becomes more negative with the decrease in primary production and release of plant carbon, leaving more $^{12}CO_2$ in the atmosphere. This drop is called a **negative $\delta^{13}C$ excursion**. When life is robustly favored and carbon is fixed by autotrophs, and the organic carbon resulting from them is eventually buried in sedimentary rocks, the rise in $\delta^{13}C$ is called a positive **$\delta^{13}C$ excursion**.

32.3.3.1: EXAMPLES OF CLIMATIC EVENTS ACCOMPANIED BY CHANGES IN $\Delta^{13}C$.

Late Devonian period

Fossil evidence from the late Devonian, when large terrestrial plants evolved and expanded, is characterized by increases in $\delta^{13}C$.

Paleocene/Eocene Thermal Maximum

We saw in Chapter 31.1 that around 55 MYA, sediment records indicate a spike in temperatures of about 5^0 F occurring over about a 100K year timeframe. This was accompanied by a dramatic spike in CO_2 and a dramatic drop in ocean pH as measured by the loss of deep sea $CaCO_3$ (chalk). This very short time frame is called the Paleocene/Eocene thermal maximum (PETM), which shows very quick spikes (on the geological time scale) can and do occur. Sediment records for this time indicate a large negative $\delta^{13}C$ excursion, consistent with a loss of plants with their preferential uptake of $^{12}CO_2$, leading to an accompanying increase in $^{12}CO_2$ in the atmosphere.

1500-1650 CE

We examined $\delta^{18}O$ values during the Little Ice Ages in Chapter 31.2. What about $\delta^{13}C$ values? CO_2 and $\delta^{13}C$ values from 1000 to 1900 are shown in Figure 32.3.12.

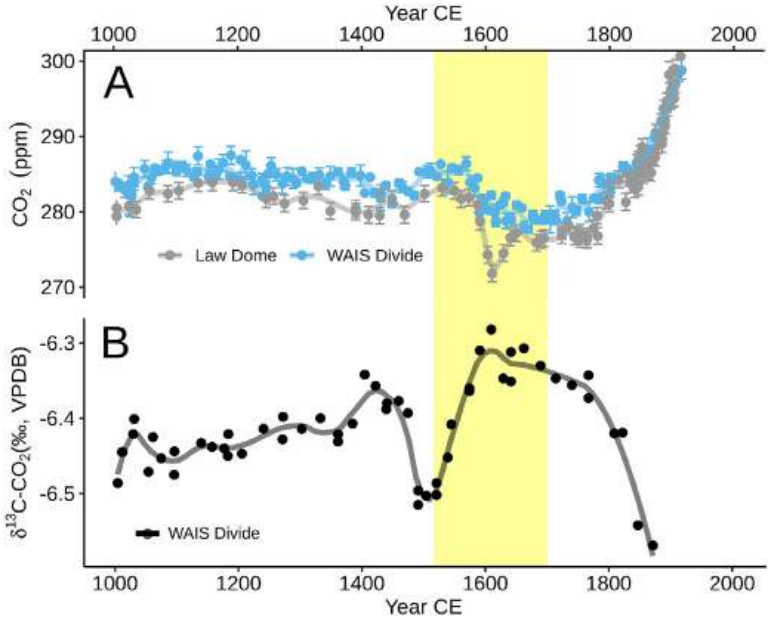


Figure 32.3.12: CO_2 and $\delta^{13}C$ values from 1000 to 1900. Koch et al. Quaternary Science Reviews, 207, 2019, 13-36. https://doi.org/10.1016/j.quascirev.2018.12.004. CC BY license (http://creativecommons.org/licenses/by/4.0/).

Panel (A) shows the CO2 concentrations recorded in two Antarctic ice cores: Law Dome (grey, MacFarling Meure et al., 2006) and West Antarctic Ice Sheet (WAIS) Divide (blue, Ahn et al., 2012).

Panel (B) shows the carbon isotopic ratios recorded in CO2 from the WAIS Divide ice core (black, Bauska et al., 2015) showing an increased terrestrial carbon uptake over the 16th century (B). The yellow box is the span of the major indigenous depopulation event (1520e1700 CE). Loess smoothed lines for visual aid.

Koch et al have strong evidence to suggest that the cooling after 1510 (area in the yellow box in the above figure) was associated with a dip in CO_2 caused by the reforestation of indigenous peoples' land in Meso and South American after epidemics of European disease killed upwards of 90% (around 55 million) of the indigenous peoples. The open and agricultural land reverted back to forests. The diseases included smallpox, measles, influenza, the bubonic plague, and eventually

malaria, diphtheria, typhus and cholera. Domesticated farm animals brought from Europe to the Americans led to most of the disease. Along with the death of so many people was a concomitant return of cleared and agricultural lands (about 56 million hectares or 212,000 mi^2) to forest and plant growth. This may have led to a 7-10 ppm drop in CO_2 in the late 1500s and early 1600s, peaking in 1601 (middle of the yellow box). This decrease in temperature was associated with a small rise (small positive excursion) in the $\delta^{13}C$ values, as $^{12}CO_2$ was preferentially removed from the atmosphere. Global surface air temperatures decreased by around 0.15ºC. This "Great Dying" of Indigenous peoples shows the power of humankind to globally alter climate in calamitous ways, even before the use of fossil fuels. The decrease in $\delta^{13}C$ values before 1500 was unexplained.

1800-the present

$\delta^{13}C$ values can also be used to unequivocally prove that the increase in CO_2 since the industrial revolution is from the burning of fossil fuels, which is of biogenic origin and hence have more negative $\delta^{13}C$ values. Figure 32.3.13 shows atmospheric CO_2 levels in ppm plotted along with $\delta^{13}C$ values. There is a perfect correlation between the rise in atmospheric CO_2 starting with the industrial revolution with the decrease in the $\delta^{13}C$ values over the same time.

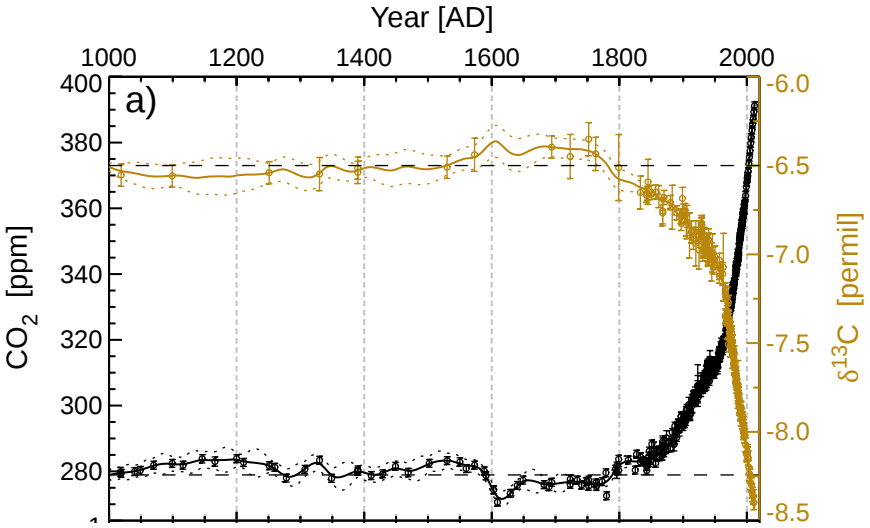


Figure 32.3.13: CO_2 concentration (black circles) and the $\delta^{13}C$ (brown circles) from 1000 to 2010. Rubino et al. Journal of Geophysical Research: Atmospheres. https://doi.org/10.1002/jgrd.50668. With permission (Copyright Clearance Center)

32.3.4: SUMMARY

In the first three sections of Chapter 31 (31.1, 31.2, and this one), we have introduced the basics of climate change, and how climate scientists obtain, analyze and interpret climate data. We emphasized the scientific rigor by which they do that and offered a detailed analysis of the use of isotopes to document past and present changes in climate, Finally, we offered models to predict and mitigate future climate changes. After reading this material, you should be enabled to discuss climate change with others from a sound and valid position. More importantly for this book, you will have a better knowledge base and understanding for the rest of the chapter sections, which will address specific topics in "biochemistry and climate change".

KEY POINTS - BETA VERSION FROM CHAT.OPENAI

1. The carbon cycle is the process by which carbon moves through the Earth's systems, including the atmosphere, oceans, and biosphere.
2. The carbon cycle is driven by the exchange of carbon between different reservoirs, such as the atmosphere, oceans, and living organisms.
3. The main processes involved in the carbon cycle include photosynthesis, respiration, and the formation and weathering of rocks.
4. Human activities, such as burning fossil fuels and deforestation, have significantly increased the amount of carbon dioxide (CO2) in the atmosphere, disrupting the natural balance of the carbon cycle.
5. The increase in atmospheric CO2 is the primary driver of climate change, as it causes the greenhouse effect, trapping heat in the atmosphere and warming the Earth's surface.
6. The ocean also plays a critical role in the carbon cycle, as it acts as a sink for CO2, absorbing about 25% of the CO2 emitted by human activities.
7. The acidification of the ocean caused by the uptake of CO2 is having a significant impact on marine ecosystems, altering the chemistry of seawater and making it more difficult for some organisms to build and maintain their shells and skeletons.
8. Understanding the carbon cycle and carbon chemistry is crucial for understanding the causes and impacts of climate change and for developing strategies to mitigate and adapt to its effects.

32.4: BIOFUELS A - CORN AND SUGAR CANE ETHANOL

32.4.1: INTRODUCTION

The world has a great need for energy. We have invested vast sums of money in finding and using fossil fuels. Fossil fuels seem to be an ideal energy source since they are highly reduced, easily stored, energy-dense, and highly abundant. Yet we now know the immense cost of their use: pollution that shortens lives and climate change. We have dramatically increased our bioethanol production from corn and sugar cane to remove our reliance on fossil fuels for transportation. Ethanol is partially oxidized as it has one oxygen atom in the two-carbon molecule. Hence, the energy released per gram is about 63% (by mass) and 70% (by volume) compared to gasoline. The energy values for various fuels are shown in Table 32.4.1 below, where $\Delta H_c°$ is the standard enthalpy of combustion.

Table 32.4.1: Energy values for a variety of fuelsData source: https://www.engineeringtoolbox.com/s...nt-d_1987.html

Name	Formula	State	-$\Delta H_c°$ *kJ/mol*	-$\Delta H_c°$ *kJ/g or MJ/kg*	-$\Delta H_c°$ *kcal/kg*
Ammonia	NH_3	gas	383	22.48	5369
Butane	C_4H_{10}	gas	2878	49.50	11823
Carbon (graphite)	C	cry	394	32.81	7836
Carbon monoxide	CO	gas	283	10.10	2413
Ethanol	C_2H_6O	liq	1367	29.67	7086
Hydrogen	H_2	gas	286	141.58	33817
Methane	CH_4	gas	891	55.51	13259
Methanol	CH_4O	liq	726	22.65	5410
Naphthalene	$C_{10}H_8$	cry	5157	40.23	9609
Octane	C_8H_{18}	liq	5470	47.87	11434
Propane	C_3H_8	gas	2220	50.33	12021
wood (red oak)				14.8	3540
coal (lignite)				15	3590
coal (anthracite)				27	4060
methyl stearate (biodiesel)	$(CH_3(CH_2)_{16}(CO)CH_3$			40	9560

Nevertheless, ethanol is readily made and is a valuable biofuel. A glance at the table suggests that H_2 would be the best possible fuel, given that it has the highest energy release per gram and contains no carbon. At present, it can't be produced at the scale needed, and it isn't easy to store and transport. The critical infrastructure for its widespread use is lacking. Yet these factors could be solved. We'll explore biohydrogen in a separate chapter section.

In theory production of ethanol from plants at first glance is carbon neutral since each carbon in the ethanol is fixed from atmospheric CO_2 during photosynthesis. Combustion of ethanol then returns the CO_2 to the atmosphere in a net zero emission fashion, as shown in the reaction below.

$$\mathbf{6CO_2}\ (g) + 6H_2O\ (l) \rightarrow C_6H_{12}O_6\ (s) + 6O_2\ \text{(photosynthesis)}$$

$$C_6H_{12}O_6\ (s) \rightarrow 2\ CH_3CH_2OH\ (l) + \mathbf{2CO_2}\ (g)\ \text{(anaerobic ethanol biosynthesis)}$$

$$2CH_3CH_2OH\ (l) + 6O_2 \rightarrow \mathbf{4CO_2}\ (g) + 6H_2O\ (g)\ \text{(combustion of ethanol)}$$

Six CO_2s in, six out! It seems simple but it's not. We'll explain more later. First, let's explore how ethanol is synthesized for its two major uses, drinking and use as a biofuel.

32.4.2: ETHANOL PRODUCTION OVERVIEW

The scale of worldwide ethanol production is quite staggering. Let's first consider the production of ethanol by yeast for alcoholic beverages. About 100 billion US gallons/yr (BGY) of beer, 7 BGY of wine, and 6 BGY of spirits are produced yearly. Assuming beer, wine and spirits are about 5%, 12%, and 40% percent ethanol by volume, respectively, the volume of actual ethanol/year made by yeast in these alcoholic beverages is about 5 BGY (beer), 0.85 BGY (wine) and 2.4 BGY (spirits). This sums to about 8.3 billion gallons of ethanol produced by these microorganisms. Compare this to fuel ethanol production each year, shown in Figure 32.4.1.

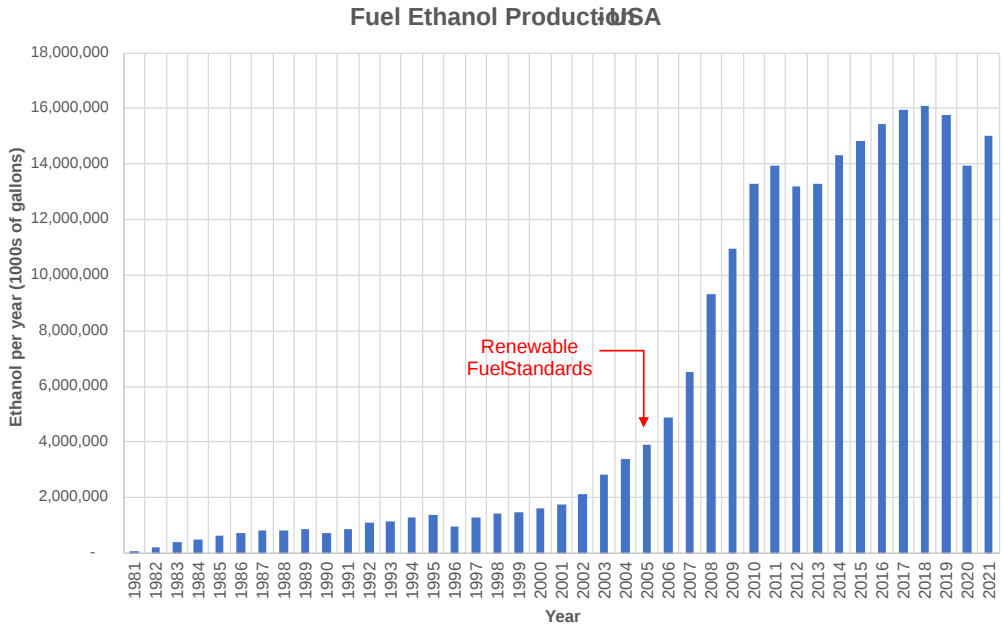


Figure 32.4.1: US Fuel Ethanol Production. Data from **U.S. Bioenergy Statistics**

Note that the y-axis is in units of 1000s gallons of ethanol. Peak US production was in 2018, when 16 billion gallons were produced, about 1/10 of the gasoline used yearly in the US. The year Renewable Fuel Standards (RFS) were introduced in the USA (2005) is also shown. This dip in 2020 is attributed to the COVID pandemic.

Combined, the US and Brazil produce about 85% of fuel ethanol, as shown below in Figure 32.4.2.

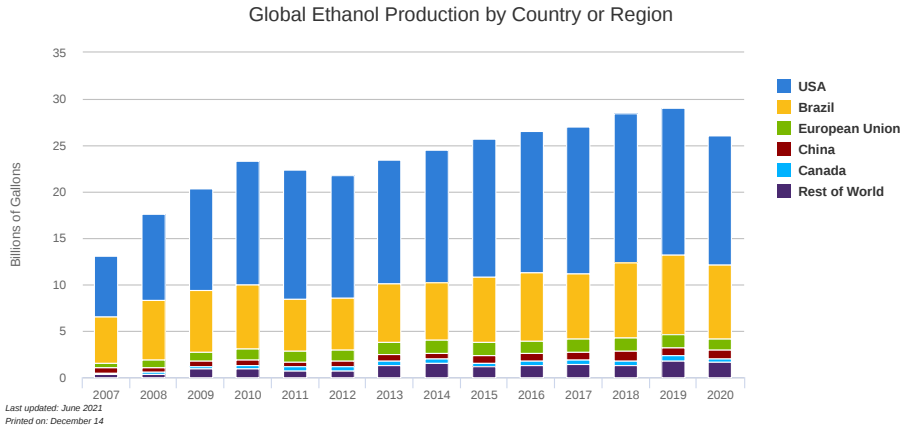


Figure 32.4.2: Fuel ethanol production (billions of gallons or BG) around the world per year. https://afdc.energy.gov/data/10331

Almost all US ethanol is made from corn, while Brazil's primary source is sugar cane.

Since the significant ramp-up of fuel ethanol around 2005, the world now produces 3x the amount of ethanol to drive our outsized vehicles than microorganisms have for our drinking. These statistics show that the world can quickly respond when it meets our needs.

32.4.3: AN OVERVIEW OF ETHANOL BIOSYNTHESIS

Whether ethanol is made for the beverage or biofuel industries, yeast play the major role, as we explored in Chapter 14.2: Fates of Pyruvate under Anaerobic Conditions- Fermentation. Yeast contains all the enzymes necessary to convert glucose (6C), made from various "feedstocks", to pyruvate (3C) through the glycolytic pathway. This is followed by the conversion of pyruvate to ethanol using two key yeast enzymes. First, pyruvate is decarboxylated to acetaldehyde by **pyruvate decarboxylase**, which uses TPP as a cofactor. Acetaldehyde is then reduced to ethanol by **ethanol dehydrogenase**, using NADH as a substrate, in a process that reforms NAD^+, allowing glycolysis to continue. These combined anaerobic reactions, known as fermentation, are shown in Figure 32.4.3.

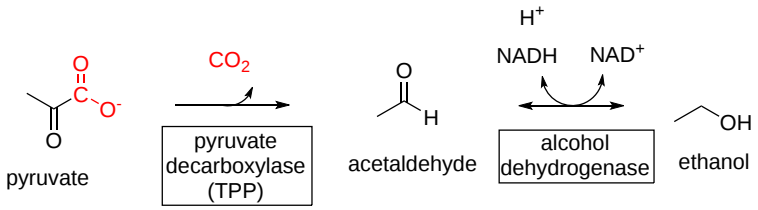


Figure 32.4.3: Summary of Ethanol Fermentation in Yeast

Yeast are facultative (not obligate) anaerobes in that they can produce energy by glycolysis and ethanol fermentation in the absence of oxygen. Of course, in the presence of oxygen, the pyruvate made from glycolysis in yeast is preferentially converted to acetyl-CoA, which enters the citric acid cycle and oxidative phosphorylation pathways to maximize ATP production. Yeast is abundant, so all that is needed is a large source of glucose.

An abundant source of glucose for bioethanol production are plants that contain starch (for example corn) or abundant simple sugars (for example sucrose in sugar cane). Starch, an α (1,4) polymer of glucose with α (1,6) branches, can easily be broken down in an industrial setting with amylases to form glucose. A significant problem with this "first" generation source of glucose is that food crops (corn, and to a lesser degree sugar cane) are used for biofuel consumption instead of for food. "Second" generation sources of glucose are crop and wood waste products that contain cellulose, a β (1,4) polymer of glucose which is found with another carbohydrate polymer lignin. A significant problem with the use of cellulose is the high chemical stability of the β (1,4) glycosidic bond. Fungi and bacteria are sources of β-glycosidases to liberate free glucose from cellulose. "Third" generation sources of glucose use algae, which does not displace cropland for bioethanol production. A "fourth" generation source of glucose are genetically engineered organisms, which might become future sources of bioethanol. Figure 32.4.4 summarizes the different generations of feedstock sources for bioethanol production.

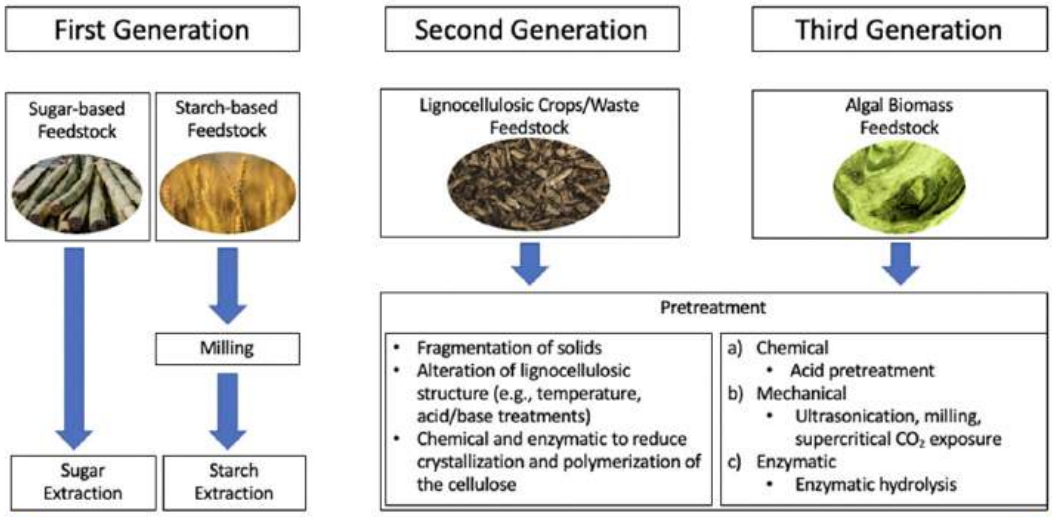


Figure 32.4.4: Generation feedstock sources for bioethanol production. Tse, T.J.; Wiens, D.J.; Reaney, M.J.T. *Fermentation* **2021**, *7*, 268. https://doi.org/10.3390/fermentation7040268. Creative Commons Attribution (CC BY) license (**https://creativecommons.org/licenses/by/4.0/**).

We will discuss the use of first generation sources, corn and sugar cane, which are used to produce most of the world's bioethanol, in this chapter section, and the other two in subsequent sections.

32.4.4: CORN BIOETHANOL

Corn is a significant source of starch, an α (1,4) polymer of glucose with α (1,6) branches. Hence glucosidases are used to hydrolyze starch to glucose. First, the dry corn is ground in a mill, breaking the outer coat of the corn kernel and increasing access to the starch. Heated water is added to form a mash or slurry. Cooking at greater than 85° C helps hydrolyze some glycosidic bonds and lowers the viscosity of the slurry. In the process of liquification, the pH is adjusted to around 6.0. Different **α-amylases** (endoglycosidases) are added, which cleave the α (1,4) glycosidic bonds to produce shorter dextrins (containing branched glucose units not cleaved by alpha-amylases), and α (1,4) linked glucose oligosaccharides of lengths from 2 glucose units (called maltose) up to 7-8. **β-amylase**, an exoamylase, is also used, which successively cleaves maltose units, Glc α (1,4)Glc, from the nonreducing ends of the chains

Alpha-amylases

A mixed-rendered structure of the human pancreatic alpha-amylase is shown below in Figure 32.4.5.

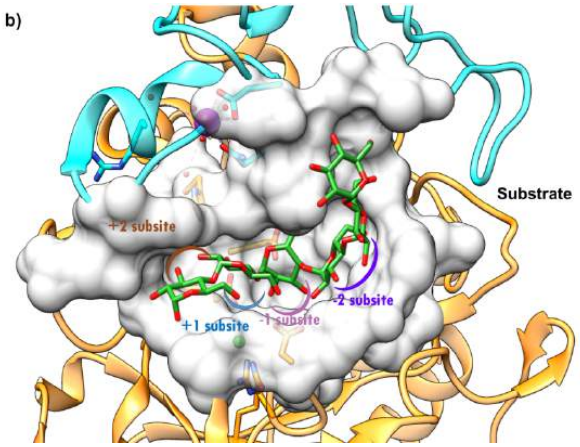
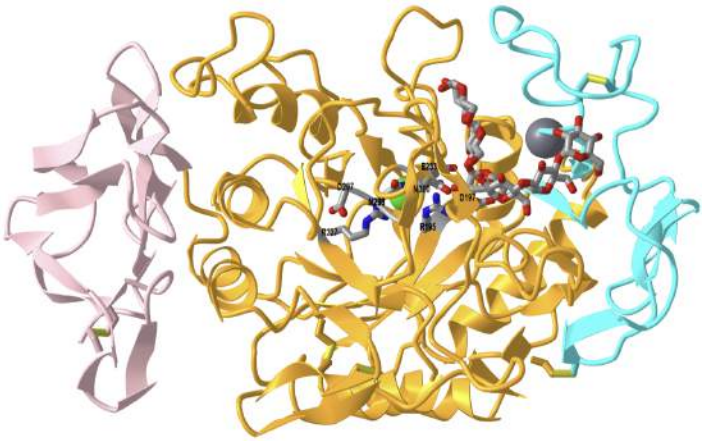


Figure 32.4.5: Surface representation of the active site of HPA (5TD4) *https://pdb101.rcsb.org/global-healt...ha-glucosidase*. CC-BY-4.0 license. Attribution: David S. Goodsell and the RCSB PDB.

The surface view highlights the deep C-shaped groove into which the substrate, in this case, octaose, is bound. Consistent with substrate numbering for proteases, the starch substrate is numbered ..-2, -1, +1, +2, with cleavage occurring between the -1 and the +1 bound alpha-glucose residues. The protein has three domains (orange, blue, and pink). This particular structure had an active site mutant (Asp300Asn, D300N). The enzyme has bound calcium and chloride ions. Ca^{2+} maintains the necessary structure, while Cl^-, bound in the C domain, is an allosteric activator.

The octaose binding site is between the A and B domains. Asp197, Glu233, and Asp300 are critical catalytic residues, with Asp 197 acting as a nucleophile to produce a glycosylated intermediate, which is hydrolyzed in the next step. Asp197 and Glu233 act as general acids/bases. We will explore in depth similar mechanisms for the action of beta-amylase (below) and cellulase (next chapter section).

Figure 32.4.6 shows an interactive iCn3D model of starch binding sites on the Human pancreatic alpha-amylase D300N variant complexed with an octaose substrate (5TD4)



NCBI iCn3D Figure 32.4.6: Starch binding sites on the Human pancreatic alpha-amylase D300N variant complexed with an octaose substrate (5TD4). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...W29jf4yAc1JEq9

The domains in the enzyme are colored-coded, as in Figure 4. Key active site residues for binding and catalysis are shown as sticks and labeled.

Beta amylase

β-amylase (also called β-1,4-maltosidase) is a key enzyme in the **saccharification** process, in which starch and cellulose are broken down into monosaccharides. β-amylase is abundant in crops (wheat, barley, soybeans, etc.) and other higher plants, as well as bacilli and fungi. It is used in making beer and caramel (malt syrup). As an exo-glycosidase, it cleaves Glc α(1,4) Glc (maltose) units from the nonreducing end of starch. It is called β-amylase since the hydrolysis proceeds with the inversion of configuration at the reducing end of the freed maltose. It can't cleave at α-1,6 branches, so if used alone, this enzyme produces free maltose and large β-limit dextrins. When fruits ripen, the enzyme cleaves starch to produce sweet maltose. It is also used in seed germination.

MALTING

Plants have to sprout, which requires energy and free sugars. Maltose is produced on activation of β-amylase during seed germination and sprouting. Although maltose is less sweet than sucrose and fructose, it is used in hard candies, given its tolerance to the heat needed in candy production. Malting of grains is accomplished by adding water to sprout them, leading to maltose and other sugars forming. This is followed by drying, with the malted grains used as sweeteners in the food industry. Malted grains are used to produce beer, whisky, some baked goods, and drinks. Barley is the most commonly malted grain used in cereals.

Huge amounts of amylases are needed for corn ethanol production, and they must withstand the conditions necessary for the industrial production of ethanol. Much effort has been devoted to finding and characterizing microbial β-amylases. We'll describe one, AmyBa, from *B. aryabhattai*. Figure 32.4.7 shows sequence similarities among various bacterial β-amylases.

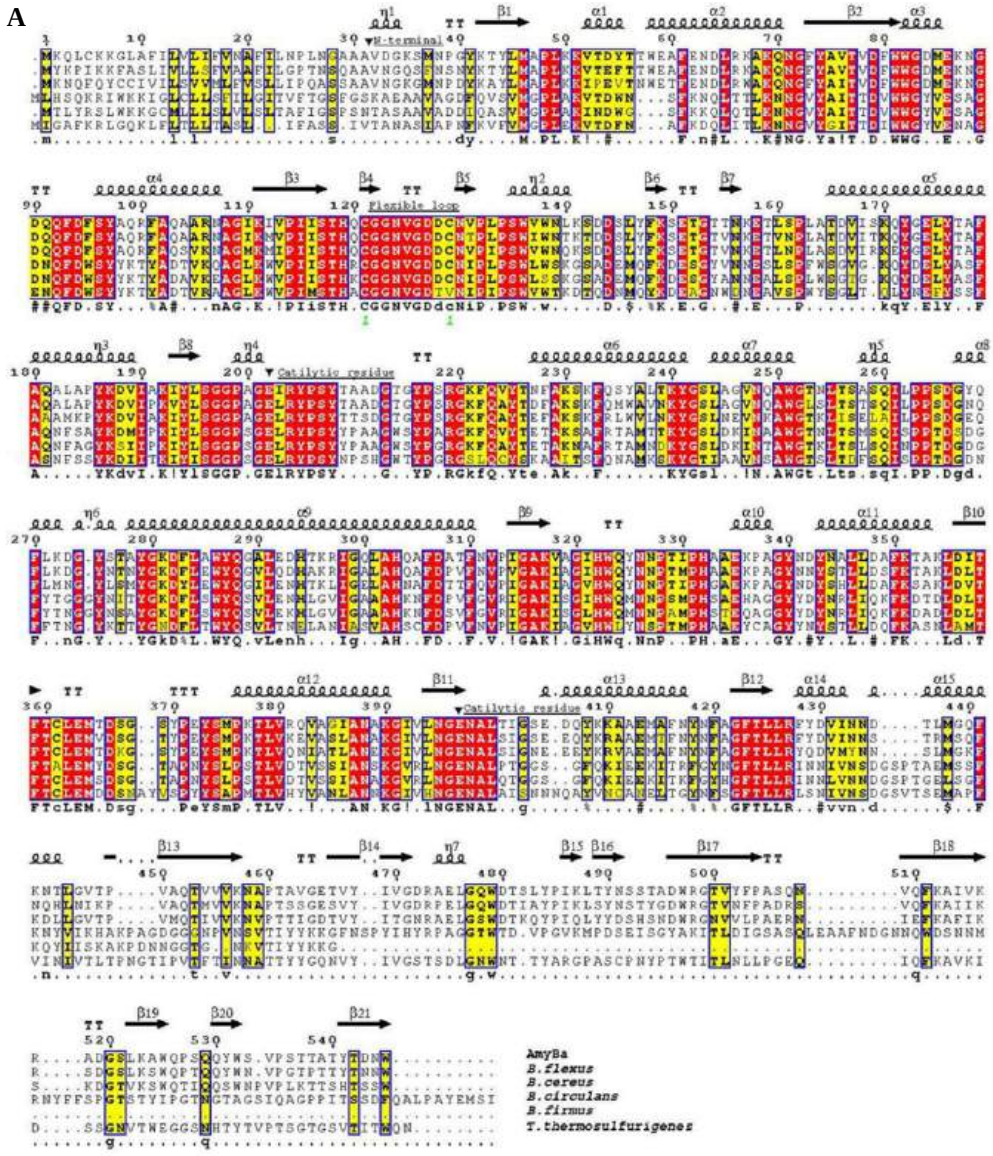


Figure 32.4.7: Sequence and structure analysis of AmyBa. . Duan, X., Zhu, Q., Zhang, X. *et al.* Expression, biochemical and structural characterization of high-specific-activity β-amylase from *Bacillus aryabhattai* GEL-09 for application in starch hydrolysis. *Microb Cell Fact* **20**, 182 (2021). https://doi.org/10.1186/s12934-021-01649-5. Creative Commons Attribution 4.0 International License. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Panel A shows multiple sequence alignments of β-amylases. The strictly conserved residues are displayed on a red background, and the highly conserved residues are shown on a yellow background. The secondary structure elements are shown for *B. cereus* β-amylase (PDB ID: 5BCA). The signal-peptide-cleavage site and two catalytic residues (E) are indicated by black triangles (black inverted triangles).

Conservation of the flexible loop motif (HXCGGNVGD) is noted. β-amylase accession numbers are as follows: *B. aryabhattai* (WP_033580731.1), *B. cereus* (P36924.2), *B. flexus* (RIV10038.1), *B. firmus* (P96513.1), *B. circulans* (P06547.1), *T. thermosulfurigenes* (P19584.1).

A comparison of the structures of *B. aryabhattai* β-amylase with soybean β-amylases is shown in Figure 32.4.8.

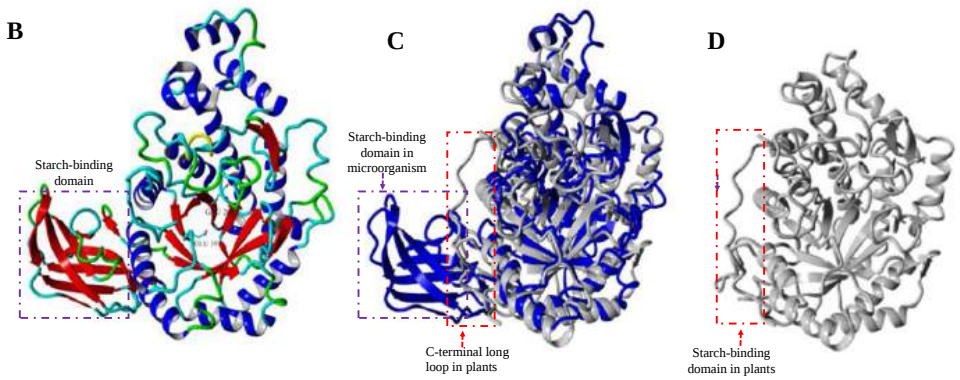
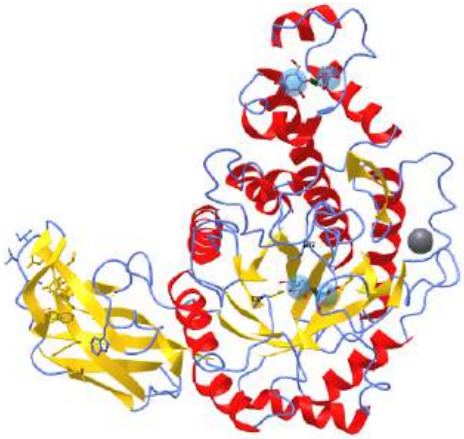


Figure 32.4.8 **B** Three-dimensional molecular model of *B. aryabhattai* β-amylase (AmyBa). **C** Superimposition of AmyBa (Blue) and soybean β-amylases (PDB ID: 1Q6C) (gray) and **D** (PDB ID: 1Q6C) (gray). The C-terminal SBD in microbial β-amylases (box, purple) and the C-terminal loop in plants (box, red). Duan, X, et al., ibid.

The AmyBa has an additional starch binding domain at the carboxy terminus (Panel B) compared to soybean β-amylases (panel D).

Since no structures of (AmyBa are publically available, we present Figure 32.4.9, which shows an interactive iCn3D model of beta-amylase from Bacillus cereus var. mycoides in complex with maltose (1J0Z)



NCBI iCn3D Figure 32.4.9: Beta-amylase from Bacillus cereus var. mycoides in complex with maltose (1J0Z). (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...S2aWbL3z3jPnW8

Similar to alpha-amylase, β-amylase has an N-terminal catalytic domain with a beta-barrel, a connecting second domain, and a third C-terminal domain which is primarily antiparallel β-sheets. Two key catalytic side chains, Glu 172 and Glu 367, are found in the beta-barrel.

In Chapter 20.6, we discussed starch synthesis (not its hydrolysis) in detail. We showed that the reaction, which uses a NDP sugar as a glycan donor, could proceed either with retention or inversion of the anomeric carbon of the donor NDP-sugar. This is illustrated for the reaction of a C1 α-NDP donor monosaccharide with a monosaccharide acceptor to produce the α(1,4) link with retention of configuration or the β(1,4) link with inversion as shown in Figure 32.4.10 below.

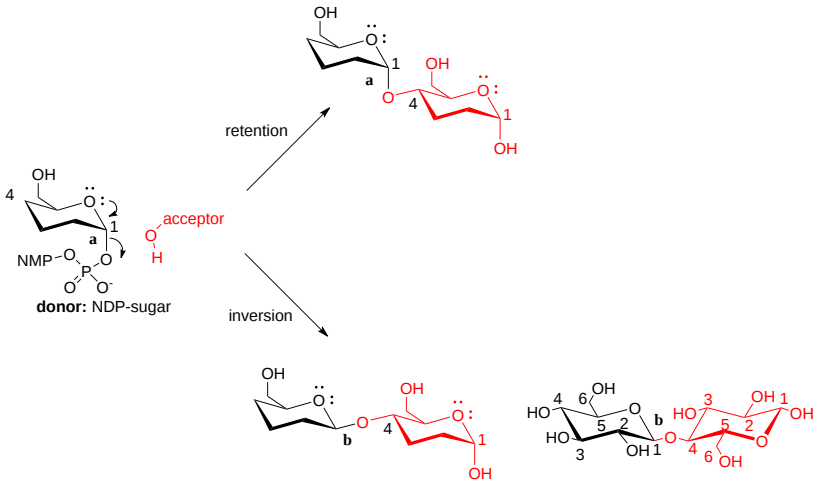


Figure 32.4.10: Reaction of a donor NDP-monosaccharide and an acceptor monosaccharide with retention or inversion of configuration at the anomeric carbon of the donor

The same stereochemical outcomes are possible for the hydrolysis of acetal bonds by glycosyl hydrolases. Alpha-amylases cleave the α (1,4) glycosidic bonds to produce shorter dextrins (containing branched glucose units not cleaved by alpha-amylases), and α (1,4) linked glucose oligosaccharides of lengths from 2 glucose units (maltose) up to 7-8. This reaction hence proceeds with the retention of configuration. In contrast, beta-amylases cleave starch to produce maltose with an inversion of configuration at the anomeric-reducing end of the maltose. We explore the chemistry of retention and inversion more in the next section on cellulase, which cleaves the β (1,4) acetal link in cellulose, but in general, reactions that proceed with inversion react in an S_N2 response, similar to the nucleophilic attack on alkyl halides. For the glycosyl transferases that proceed with inversion, the attacking nucleophile on the acceptor is made more nucleophilic by general base catalysis by a deprotonated glutamic or aspartic acid.

Figure 32.4.11 shows the results of in silicon docking studies of a small glycan, maltotetraose, to AmyBa.

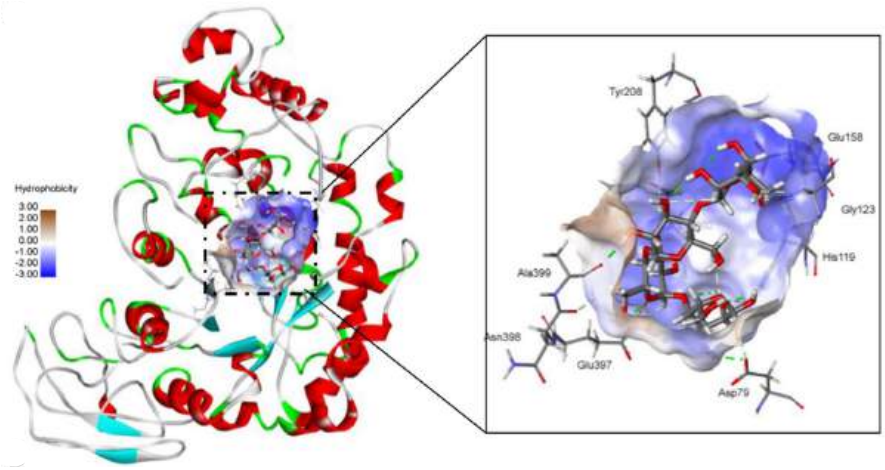


Figure 32.4.11: Molecular docking of AmyBa with maltotetraose. The overall structure and substrate binding pocket analysis of AmyBa are shown.

AmyBa displayed very high amylase activity compared to other microbial β-amylases, and its enzymatic activity was much closer to sweet potato β-amylase. Molecular dynamic and docking programs can be used to calculate binding energies for substrates. The binding energy and enzymatic activities for bacteria and sweet potato β-amylase were highly correlated, suggesting that the extensive interactions of AmyBa and maltotetraose help drive catalysis by using the energy released on binding to lower activation energies for the reaction.

32.4.4.1: SACCHARIFICATION

To enter glycolysis and fermentation, maltose must be converted to the monosaccharide glucose. The conversion of a polysaccharide to its monomers is called saccharification. To complete the conversion of starch to glucose, another enzyme, glucoamylase (also called amyloglucosidase and γ-amylase), is added. It is an exoglucosidase that cleaves both α (1,4) in amylose, amylopectin and maltose and α (1,6) branches, to form free glucose. It is a member of the glycoside hydrolase family 15 in fungi, glycoside hydrolase family 31 of human maltase-glucoamylase, and glycoside hydrolase family 97 of bacterial forms.

32.4.4.2: FERMENTATION

Glucose ($C_6H_{12}O_6$) can now enter the glycolytic pathway and continue to ethanol after conversion of pyruvate to acetaldehyde by pyruvate decarboxylase and final conversion of acetaldehyde to ethanol by alcohol dehydrogenase:

$$C_6H_{12}O_6\,(s) \rightarrow 2\,CH_3CH_2OH\,(l) + 2CO_2\,(g)\ \text{(anaerobic ethanol biosynthesis)}$$

The yeast Saccharomyces cerevisiae catalyzes this entire pathway.

The final fermentation process yields a 12-15% ethanol solution, which is distilled to form a 95% ethanol/5% water azeotrope. The water is removed by adding zeolites (molecular sieves) which can adsorb water but not ethanol.

32.4.4.3: LIFE CYCLE ANALYSIS OF BIOETHANOL: IS IT BETTER THAN FOSSIL FUELS?

We reiterate the promise of bioethanol to address global warming and climate change by presenting again the chemical equations that suggest that its use as a fuel is carbon neutral:

$$6CO_2\,(g) + 6H_2O\,(l) \rightarrow C_6H_{12}O_6\,(s) + 6O_2\ \text{(photosynthesis)}$$

$$C_6H_{12}O_6\,(s) \rightarrow 2\,CH_3CH_2OH\,(l) + 2CO_2\,(g)\ \text{(anaerobic ethanol biosynthesis)}$$

$$2CH_3CH_2OH\,(l) + 6O_2 \rightarrow 4CO_2\,(g) + 6H_2O\,(g)\ \text{(combustion of ethanol)}$$

If only these three equations, this simple model for production and use of corn bioethanol, were the only factors influencing **net** CO_2 emission on bioethanol burning, there would be no controversy about its use. Yet the actual CO_2 emissions depend on many more hidden from view by these simple equations. What is needed is a **life cycle analysis (LCA)** that determines the environmental impact (in this case, net CO_2 emissions) of corn ethanol through every phase of its existence, from cradle to grave, starting with the planting of corn to the combustion of bioethanol for transportation.

All models must be tested. It's easiest to start with the simplest model. If the data fit the model, great, you're done. If not, new, more expansive models must be developed and tested. Those vociferously supporting bioethanol use often use the simple stoichiometry evident in the three equations to state that bioethanol is carbon neutral. Most, however, would want a detailed life cycle analysis (LCA) before jumping to an immediate conclusion.

LCAs are very challenging, and data on a global scale is required. Some measurements at the worldwide scale have significant uncertainties (that don't include CO_2 in the atmosphere, however) and are estimates, at best. A recent study looked at the impact of a specific event, the adoption of the **US Renewable Fuel Standards (RFS)** that regulate biofuels in the US (which produces about half of all the world's biofuels), on CO_2 emission from the significant increase in corn plant and corn ethanol the followed the adoption of the standard. . Using LCA based on a series of economic and environmental metrics, the model shows that bioethanol is not a panacea for CO_2 emissions and may be more detrimental than fossil fuels use for vehicles.

The study calculated the **carbon intensity** changes for corn ethanol that followed after the adoption of the standards. Scientists have used other events that led to immediate changes (9/11) and 1-2 year changes (Covid pandemic) on environmental parameters like CO_2 emissions.

Carbon intensity measures how much energy-related CO_2 is emitted per dollar generated (GDP). Ideally, policies should be implemented that decrease carbon intensity. Green energy derived from both solar and wind is an example. A similar metric is **energy intensity**, the total energy production per GDP, and both are consumption-based values.

Figure 32.4.12 shows carbon intensity per GDP per country over the last 30 years (data from Our World in Data).

Figure 32.4.12: Consumption-based carbon intensity from 1990 to 2018. Our World in Data.

Generally, the world is moving to more efficient energy production, but remember that our energy consumption is still dramatically increasing.

The LCA model showed that the RFS led to these interrelated outcomes. It:

- increased corn prices by 30% and the prices of other crops by 20%
- increased US corn cultivation by 2.8 Mha (8.7%) and total cropland by 2.1 Mha (2.4%) in the years following policy enactment (2008 to 2016). (1 hectare is an area of a square with100 meters sides, equivalent to 10,000 m^2
- increased annual nationwide fertilizer use by 3 to 8%
- increased water quality degradation by 3 to 5%
- increased emissions from domestic land use changes

These all combined to lead to a carbon intensity of corn ethanol that was "no less than gasoline and likely at least 24% higher", according to the study.

The changes in the metric are visually described in Figure 32.4.13.

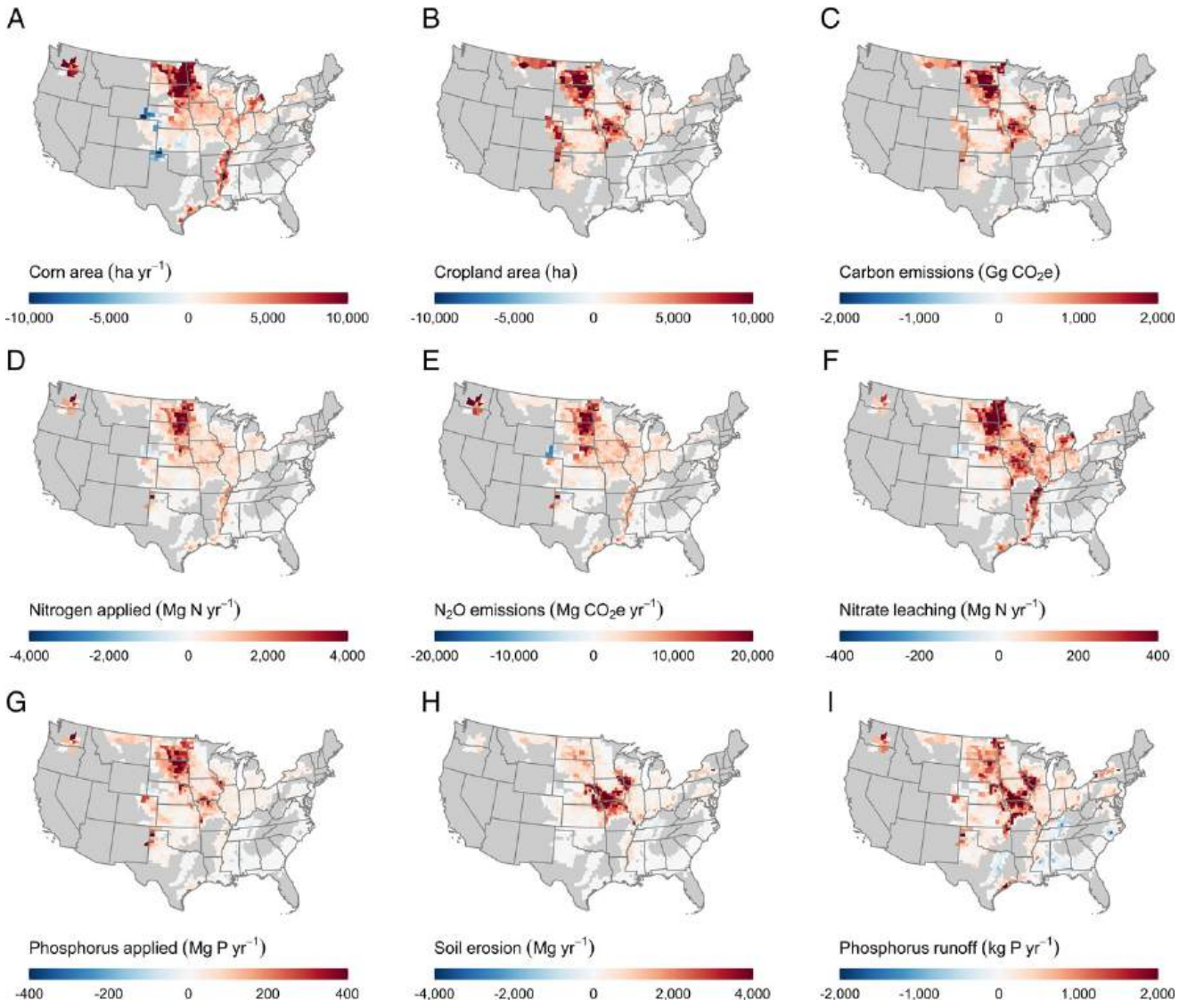


Figure 32.4.13: *Changes due to the RFS. (A) Corn planted area. (B) Cropland area. (C) Carbon emissions. (D) Nitrogen applications. (E) Nitrous oxide emissions. (F) Nitrate leaching. (G) Phosphorus applications. (H) Soil erosion. (I) Phosphorus runoff. Positive numbers indicate an increase due to the RFS. Field-level results were aggregated at the county level for enumeration and visualization.* Tyler J. Lark et al. PNAS. 119, 2022 (https://doi.org/10.1073/pnas.2101084119) Creative Commons Attribution License 4.0 (CC BY).

Land use changes include farming land that was retired or designated for conservation programs. Tilling additional land releases carbon stored in the soil. The increased farming significantly increased fertilizer production, which leads to N_2O emissions. In addition, more of the existing cropland was planted with corn. These finds contrast with a USDA **study** that shows that corn ethanol has a 39% lower corn ethanol intensity than gasoline which was stated to derive from carbon captures from the newly planted crops. However, that study did not account for emissions from land use changes.

LCA can identify aspects of production that lead to the most negative consequences, which for the sake of this chapter is greenhouse gas emissions. For example, the LCA for corn ethanol might improve if the CO_2 released on its production during anaerobic ethanol biosynthesis could be captured. Outcomes would also change if renewable energy sources were used for stages of production that require fossil fuel use.

This rigorous LCA did not address the moral question of using land that could be used to feed people to produce bioethanol for use in our ever-bigger vehicles. In addition, opponents of solar energy installations suggest that solar installs would require so much land that it would remove land for agricultural purposes. What is missing from their argument is the vast amount of land used now for corn ethanol. Farmers planted 90 million acres of corn in 2022 in the US, a land area about 90% the size of the entire state of California. 44% of that corn went to biofuels, and only 12% went to human consumption. In addition, approximately 44% percent was used to feed animals for human consumption, an inefficient and unsustainable use of crop land and resources.

32.4.5: PRODUCTION OF SUCROSE AND BIOETHANOL FROM SUGARCANE

Like corn, sugar cane, a tropical, perennial grass, is used (mainly in Brazil) to produce ethanol. Sugar cane is a C4 plant with a high ability to fix carbon. The fact that it is a perennial and does not need replanting each year makes it a more ideal feedstock than corn for bioethanol production. In 2020, sugar cane was by far the most-produced crop or livestock product in the world (1.87 billion metric tons), followed by corn (1.16 billion metric tons). The production by country for both corn and sugar cane is shown in Figure 32.4.14.

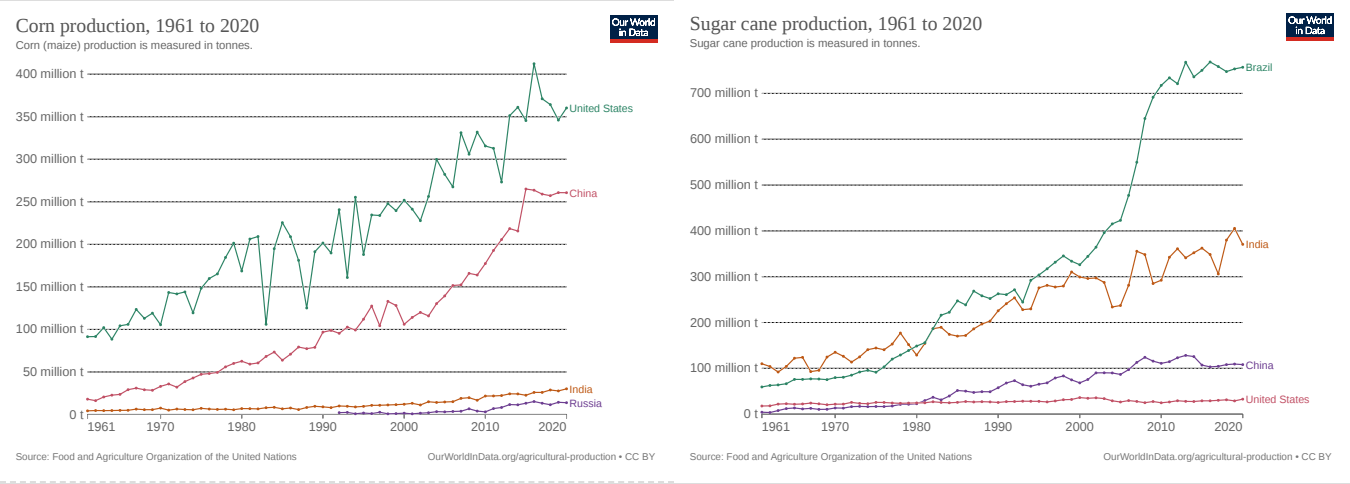


Figure 32.4.14: Corn and sugar cane production by country. Graphs from Our World in Data. https://ourworldindata.org/agricultural-production#

That sugar cane production is so high compared to grain crops that provide nutrition (not just "sweet" calories) might come as a surprise, but it shouldn't, given our addiction to sweet foods.

Sugar cane is often harvested manually in developing countries. It is then cut, milled, and mixed with water to extract the soluble sucrose (table sugar). The sugar cane components during extraction are shown below in Figure 32.4.15.

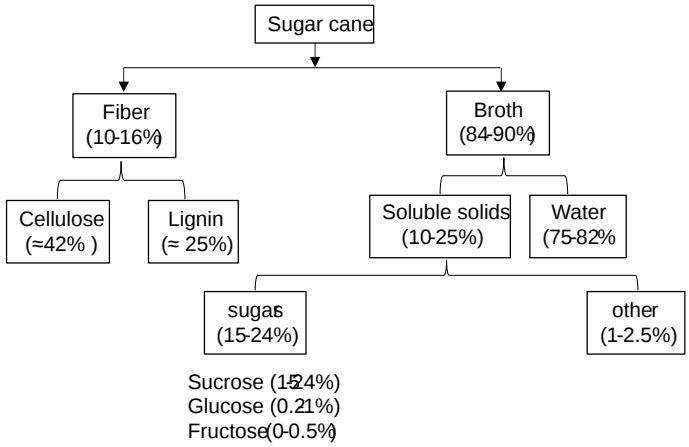


Figure 32.4.15: Components of Sugar Cane (after Larissa Canilha et al. J Biomed Biotechnol. 2012; 2012: 989572. doi: 10.1155/2012/989572

Sucrose is a nonreducing disaccharide (O-α-D-glucopyranosyl-(1,2)-β-D-fructofuranoside). Its structure is shown in Figure 32.4.16.

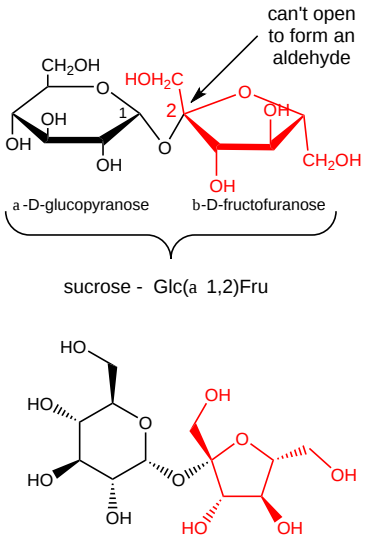


Figure 32.4.16: Structure of fructose

SUCROSE, CARAMEL AND MOLASSES

Sucrose decomposes at 186 °C (367 °F) instead of melting (a feared event for organic chemistry students who wish to record melting temperatures in the lab) to form caramel. Molasses is a very viscous liquid product from refining sugar cane or sugar beets. It is used as a sweetener with its own taste properties, and it's a component of brown sugar as well. On a sweetness scale, if sucrose is assigned a value of 100, fructose is 140, high fructose corn syrup is 120-160, and glucose is 70-80.

For bioethanol production, sucrose is degraded by the enzyme **invertase** to form monomeric glucose and fructose. Invertases are activated on the milling and liquification of the sugar cane, so if sucrose is the desired commercial product, an additional clarification step (heat to 115°C and treat with lime and sulfuric acid) is necessary to prevent hydrolytic cleavage of sucrose.

32.4.6: BIOETHANOL PRODUCTION FROM SUGAR CANE SUCROSE

Bioethanol can be made from either the fibrous lignocellulose remains of the sugar cane, called **bagasse** or from water-soluble sucrose. We will describe the production of cellulosic ethanol from field crop stalks, called stover, and leaves, straw, wood chips, and sawdust (all "waste" biomass), in Chapter 31.5. The same principles apply to bioethanol production from **bagasse**, the solid remains after the juice extraction from sugar cane. (Bagasse is often burned to provide energy for sugar cane processing).

In addition to sugar cane, sugar beets and sweet sorghum, a C4 plant similar to sugar cane, are used to produce bioethanol. As a C4 plant, sweet sorghum is very efficient at producing biomass through photosynthesis. It grows in temperate and tropical climates, has a short growing period, and is resistant to drought and cold. Its stalks have free sugars as well as lignocellulose stocks.

This chapter will focus on bioethanol production from sugar cane sucrose. Again, this is accomplished using yeast (Saccharomyces cerevisiae), which has the enzyme **invertase 2** (**beta-fructofuranosidase 2 or Saccharase**) needed to convert sucrose into sucrose fructose and glucose, which can enter glycolytic and fermentation pathways.

Invertase, shown in 1842 to invert the stereochemistry of sugars, was first isolated from yeast in 1860. It has a secreted glycosylated homooctameric form and an intracellular form, all products of the same gene. It's a member of Family 32 of the glycoside hydrolases. The structure of the Saccharomyces invertase (SInv) octamer structure is shown in Figure 32.4.17 below.

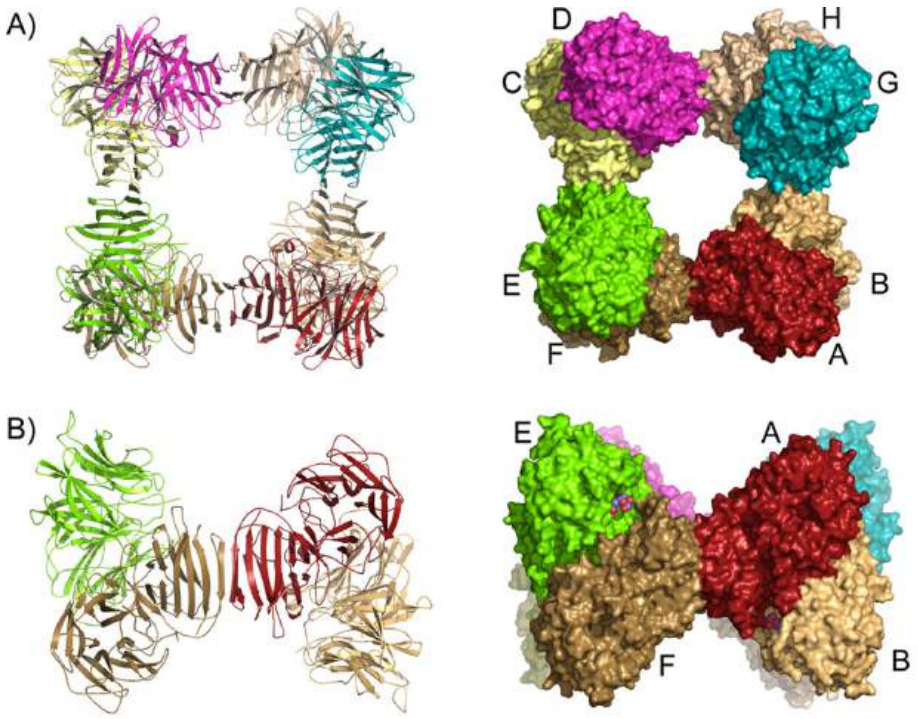


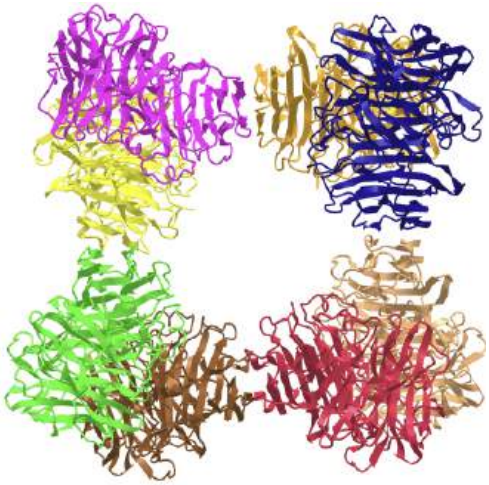
Figure 32.4.17: Structure of octameric SInv. M.Angela Sainz-Polo et al. JBC, 288, 9755-9766 (2013). DOI:https://doi.org/10.1074/jbc.M112.446435. Creative Commons Attribution (CC BY 4.0)

Panel a shows a view of the SInv octamer in ribbon (*left*) and solvent-accessible surface (*right*) representations, showing each subunit in a different color.

Panel *b shows that t*he octamer is rotated 90°, illustrating that it can be best described as a tetramer of two different kinds of dimers, AB/CD and EF/HG, which are compared by superimposing subunit F on subunit B in *c*

Even though all eight subunits in the octamer are identical (58.5K, 512 aa), the quaternary structure of the 8-mer can best be viewed as a tetramer of dimers (i.e. 4 dimers pack to form two packed tetramers giving the octamer). The AB and CD dimers pack in a "closed form" with a narrow active site pocket, allowing a glycan of 3-4 monomers. The EF and GH dimers pack in an "open form" with a wide active site pocket for longer glycans. Of course, our main interest here is in the binding of sucrose.

Figure 32.4.18 shows an interactive iCn3D model of Saccharomyces cerevisiae invertase (4EQV)



NCBI iCn3D Figure 32.4.18: Saccharomyces cerevisiae invertase (4EQV). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...ysQW8YSrvPkus9

The color coding of the subunits is the same as shown in the right top image of Panel A, Figure 16 above.

The GH32 enzymes, including invertase, have a catalytic domain consisting of a 5-bladed β-propeller, with each blade having four antiparallel beta-strands. The blades surround an active site enriched in carboxyl side chains.

The "closed" form active site of the AB and CD dimers has at its base Phe 388 and Phe 296, which provide hydrophobic interactions. The "open" form active site of the EF and GH dimers also has a salt bridge between Asp 45 and Lys 385. These are shown in Figure 32.4.19, along with a bound 1-kestose, which is a trisaccharide"sucrose analog" found in vegetables. It consists of a β-D-fructofuranose connected to β-D-fructofuranosyl and α-D-glucopyranosyl residue at the 1- and 2-positions.

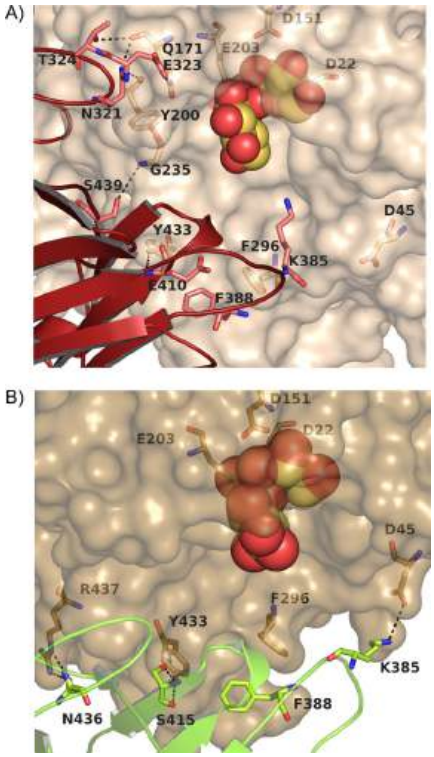


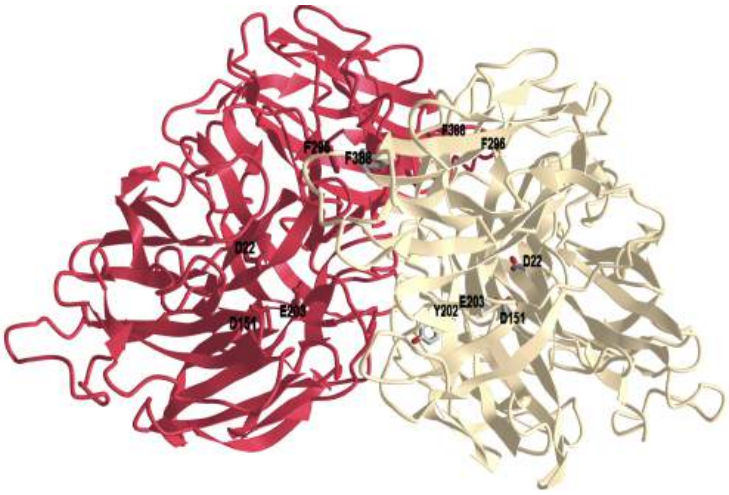
Figure 32.4.19: Dimer interface at the active site. The octameric SInv active site interfaces are detailed, keeping the same color pattern as above with one subunit being shown in ribbon representation for clarity. Angela Sainz-Polo et al. JBC, 288, 9755-9766 (2013). DOI:https://doi.org/10.1074/jbc.M112.446435. Creative Commons Attribution (CC BY 4.0)

Panel A shows that the AB/CD dimers are tightly made by interactions among both their catalytic and β-sandwich domains. Hydrophobic interactions around found at the base of the catalytic pocket through Phe-388 and Phe-296.

Panel *b*, by contrast, shows that the EF/GH dimers interact only through their β-sandwich domains. In addition, the catalytic pocket is also paved by a new salt bridge formed between Asp-45 and Lys-385 from the β-sandwich domain, which lines the cavity. A putative 1-kestose molecule is shown in a spherical representation.

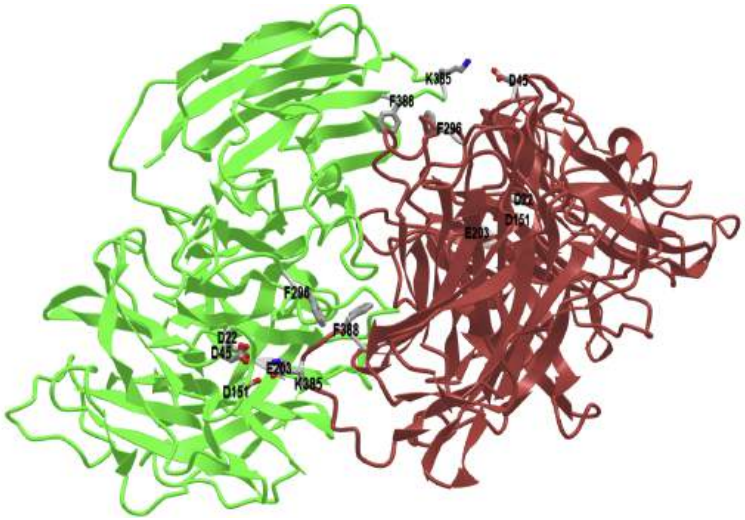
The hydrolysis of sucrose by invertase proceeds with the retention of configuration at the anomeric carbon. An active site Aspartate 22 acts as a nucleophile to form a glycosylated intermediate (fructose-Asp). This is followed by hydrolysis of the intermediate. An active site Glutamate 203 acts as a general acid/base. The fructose could also be transferred to another glycan in a transglycosylation reaction. The hydrophobic side chains Phe-388 and Phe-296 line the base of the active site pocket.

Figure 32.4.20 shows an interactive iCn3D model of the **AB dimer** of Saccharomyces cerevisiae invertase with key active site residues (4EQV)



NCBI iCn3D Figure 32.4.20: **AB dimer** of Saccharomyces cerevisiae invertase with key active site residues (4EQV). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...hX9FbxLLip1f97

Figure 32.4.21 shows an interactive iCn3D model of the **EF dimer** of Saccharomyces cerevisiae invertase with key active site residues (4EQV). It has an additional salt bridge between Asp-45 and Lys-385.



NCBI iCn3D Figure 32.4.21: **EF dimer** of Saccharomyces cerevisiae invertase with key active site residues (4EQV). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...Dh33BtT5c8GXK6

32.4.7: LIFE CYCLE ANALYSIS OF SUGAR CANE ETHANOL

Does the production of bioethanol from sugar cane lead to lower net CO_2 emissions than bioethanol produced from corn? The answer would depend on if sucrose (first generation) or lignocellulose (second generation) from bagasse is the feedstock.

A recent LCA has been performed on the first-generation (feedstock is sucrose) production of bioethanol from sugar cane in Ecuador. There is a lower cost of production from this sugar-based feedstock, which requires just grinding and the addition of yeast for fermentation. It does not require a saccharification step.

Figure 32.4.22 shows the various stages and processes used to perform LCA on the bioethanol production from sugar cane sucrose. It's presented to show the complexity of such analyses, so look at the detail only if you are especially interested.

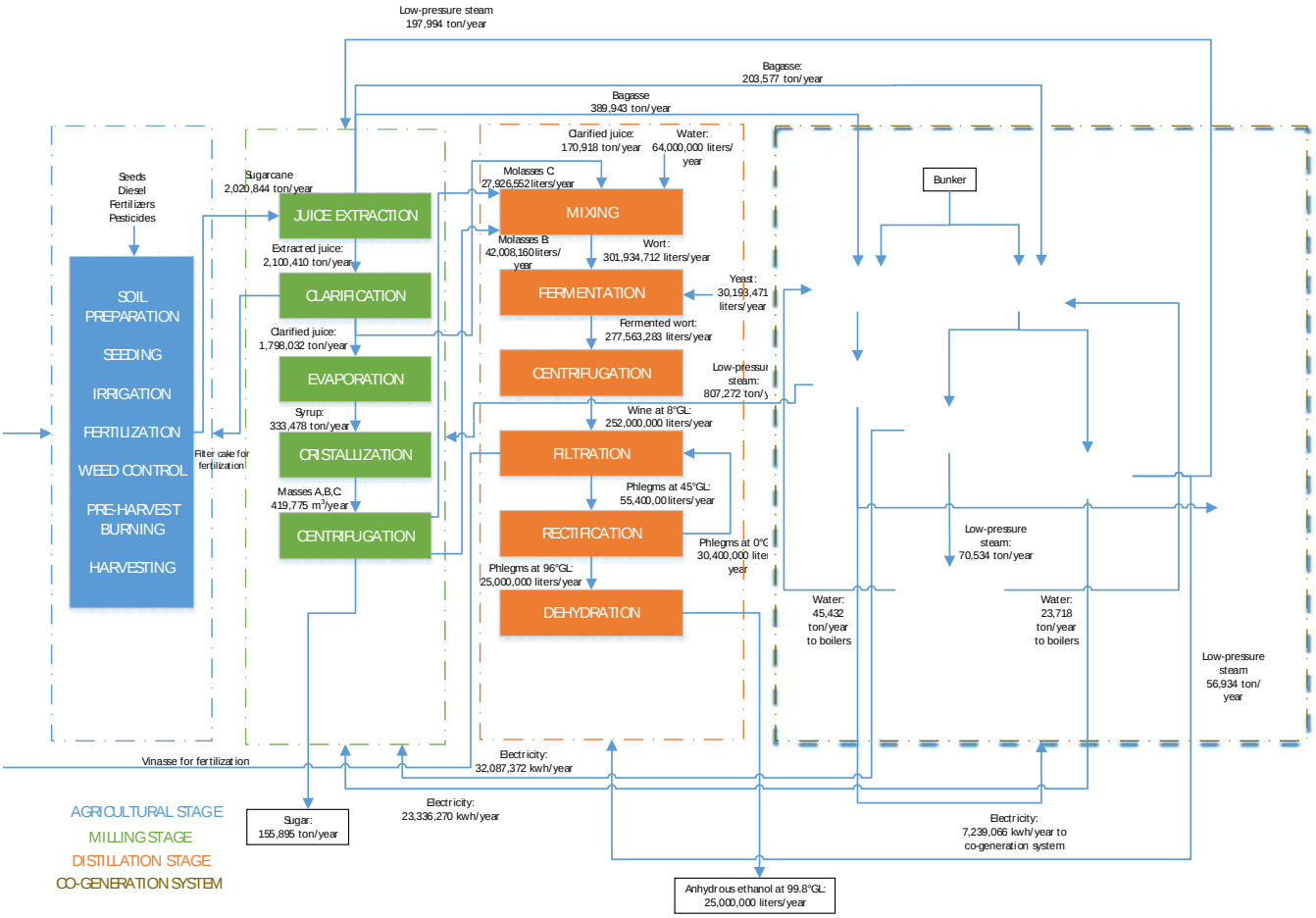


Figure 32.4.22: Anhydrous ethanol life cycle system boundaries and main product flows quantification for year 2018. Arcentales-Bastidas, D.; Silva, C.; Ramirez, A.D. The Environmental Profile of Ethanol Derived from Sugarcane in Ecuador: A Life Cycle Assessment Including the Effect of Cogeneration of Electricity in a Sugar Industrial Complex. *Energies* **2022**, *15*, 5421. https://doi.org/10.3390/en15155421. Creative Commons Attribution (CC BY) license (**https://creativecommons.org/licenses/by/4.0/**)

The study analyzed four stages for analysis, agricultural, milling, distillation and electricity generation (presumably by burning the byproduct bagasse) for impacts. They defined two functional units:

- 1 ton of sugarcane "at the farm gate" for the agricultural stage;
- 1 L of ethanol "at the plant (factory) gate".

The key results are as follows:

- The global warming potential (GWP) impact at the farm gate level was 53.6 kg of carbon dioxide equivalent (kg CO_2 equiv) per ton of sugarcane produced; This arose mostly from N_2O (34%), a potent greenhouse gas released in the process, and diesel fuel used in agricultural machinery (24%).
- The GWP for 1 L of ethanol produced at the plant gate was 0.60 kg CO_2 equiv, with the distillation phase contributing the most.

Before proceeding further, let's explain the key term, **global warming potential (GWP)**, which is widely used in LCA. It adds the contribution of other greenhouse gases like methane (CH_4) and nitrous oxide (N_2O), each of which has unique IR absorption spectra and atmospheric half-lives. The IPCC uses a 100-year time frame for the calculation of the GWP, and uses this formula:

CO_2 Equivalent kg = CO_2 kg + (CH_4 kg x 28) + (N_2O kg x 265)

$$\text{CO}_2 \text{ equivalent kg} = \text{CO}_2 \text{ kg} + (\text{CH}_4 \text{ kg} \times 28) + (\text{N}_2\text{O } kg \times 265) \tag{32.4.1}$$

- CO_2 has GWP of 1 by definition since it is the reference. Its time frame in the atmosphere (100s to 1000 years) doesn't matter since it is the reference.
- CH_4 has a GWP of around 27-30 over 100 years. It reflects its higher IR absorbance but lower lifetime (around 12 years).
- N_2O has a GWP of around 265-273 over a 100-year timescale. N_2O has a lifetime of around 109 years.

The equation can be amended by adding other greenhouse gases released in manufacturing and the use of refrigerants. These include Freon-12 (Dichlorodifluoromethane) (CFC-12, with a lifetime of 100 years and a GWP100 of 10,200, and SF_6 (used in the electricity industry to keep networks running safely and reliably) with a lifetime of 3200 years and GWP00 of 23,500!

NITROUS OXIDE - A POTENT GREENHOUSE GAS BUT NOT A LAUGHING MATTER

N_2O (laughing gas) is an overlooked source of greenhouse gases, but it leads to about 7% of the warming effect of the greenhouse gases with long life-times and a high GWP100. Agricultural practices lead to about 65% of its total emission. It is a component of the soil and atmosphere nitrogen cycle. In soil, its concentration depends on soil microbes that engage in nitrification and denitrification processes. These in turn depend on the amount of fixed nitrogen, oxygen levels and metabolically available carbon sources. The nitrification reaction, which occurs in aerated and moist soils, involves the oxidation of $NH_3 \leftrightarrow NH_4^+$ to NO_2 and NO_3^-, with some N_2O release. The major source of N_2O occurs under anaerobic conditions. These general reactions are shown below.

- Nitrification (aerobic, oxidation): $N_2 \rightarrow (NH_3 \leftrightarrow NH_4^+) \rightarrow NO_2 \rightarrow NO_3^-$
- Denitrification (anaerobic, reduction): $NO_3^- \rightarrow NO_2^- \rightarrow NO \rightarrow N_2O \rightarrow N_2$

In anaerobic sites, NO is an electron receptor during microbial respiration. N_2O is produced when there is excess nitrogen available (past the needs of plants and microorganisms), so excess use of fertilizers and manure increases its production. Nitrifying and denitrifying bacteria are most active in producing N_2O in environments with abundant N relative to assimilatory demands by other microorganisms or plants (Firestone and Davidson, 1989), as is often the case following soil amendment of fertilizers, manure, or crop residues. Physical processing of the soil (such as tillage) also affects N_2O emissions by introducing crop residues in the soil, changing soil particle size and surface area, and by changing the porosity of the soil. All of these affect soil substrate/product availability and their aqueous and gas diffusion rates.

Let's use dimensional analysis from introductory chemistry to convert the GWP from the farm gate/agricultural stage (53.6 kg CO_2 equiv/ton of sugar cane) to kg CO_2 equiv/1L of ethanol (EtOH) so we can add it the GWP from the pant gate, which is expressed in kg CO_2 equiv/L ethanol produced. The dimensional conversion is shown in Table 32.4.2 below.

53.6 kg CO_2 equiv	1 ton SC	1 L Juice	=	0.1 kg CO_2 equiv
1 ton SC	800 L juice	0.7 L EtOH		1 L EtOH produced

Table 32.4.2: Conversion of 53.6 Kg CO_2 equiv/ton of sugar cane from the farm gate (left hand column) to 0.1 Kg CO_2 equiv/1L of ethanol (EtOH).

Now add this to the reported 0.60 kg CO_2 equiv/1 L of ethanol from the plant (factor) gate and you get a total of **0.7 kg CO_2 equiv /1 L ethanol produced**.

Now use dimensional analysis from introductory chemistry to calculate how much CO_2 is actually produced on the combustion of ethanol. That value is calculated in Table 32.4.3 below.

1 L EtOH	1000 mL EtOH	0.789 g EtOH	1 mol EtOH	4 mol CO_2	44 g CO_2	1 kg CO_2	=	1.5 kg CO_2
	1L EtOH	1 mL EtOH	46 g EtOH	2 mol EtOH	1 mol CO_2	1000 g CO_2		1 L EtOH

Table 32.4.3: Total Kg CO_2 produced on combustion of 1 L of ethanol (EtOH)

The promise of bioethanol is that for every 1 C atom used to create it, 1 C atom would be released. We saw that the LCA analysis for corn ethanol in the US did **not** meet that expectation. In the Ecuadorian analysis, it appears that it **did**, since **0.7 kg CO_2 equivalents is required to produce 1 L of bioethanol** from sugar cane, but **1.5 Kg CO_2 is released on its burning**.

The LCA analysis described above reflects just the global warming potential for the use of sugar cane sucrose for bioethanol production. However, bioethanol production from sugar cane juice has other negative impacts as listed in Table 32.4.4 below.

Table 32.4.4: Impact categories included in the LCA

Impact Category	Characterization Factor	Reference Unit
Climate change	Climate change—GWP100	kg CO2eq.
Freshwater eutrophication	Freshwater eutrophication potential—FEP	kg Peq.
Marine eutrophication	Marine eutrophication potential—MEP	kg Neq.
Abiotic depletion	Metal depletion—MDP	kg Feeq.
Photo oxidant formation	Photochemical oxidant formation potential—POFP	kg NMVOCeq.
Particulate matter emissions	Particulate matter formation potential—PMFP	kg PM10eq.
Terrestrial acidification	Terrestrial acidification potential—TAP100	kg SO2eq.

KEY POINTS - BETA VERSION FROM CHAT.OPENAI

1. Biofuels are renewable energy sources derived from biomass, such as plant materials and waste.
2. Corn and sugar cane ethanol are two examples of biofuels that are produced by fermenting the sugars found in these crops.
3. Corn ethanol is typically produced by converting the starch in corn kernels into glucose, which is then fermented to produce ethanol.
4. Sugar cane ethanol is produced by crushing the cane to extract the juice, which is then fermented to produce ethanol.
5. Both corn and sugar cane ethanol are primarily used as a gasoline additive to increase octane and reduce emissions.
6. Corn ethanol is a controversial biofuel because of the high amount of energy used to produce it and the impact on food prices and the environment.
7. Sugar cane ethanol is considered to be a more sustainable biofuel because it is less energy-intensive to produce and can be produced on land not suitable for food crops.
8. However, large scale production of sugar cane ethanol has also been criticized for leading to deforestation, loss of biodiversity and displacement of local communities.
9. Cellulosic ethanol, made from non-food feedstocks like switchgrass or wood chips, is considered to be a more sustainable biofuel alternative to corn and sugar cane ethanol.

32.5: BIOFUELS B - CELLULOSIC ETHANOL

32.5.1: INTRODUCTION

In the last section, we explored how ethanol can be made from corn starch, an α(1,4) polymer of glucose with α(1,6) branches. Its production comes at a cost, however. Recent life cycle studies have shown that compared to fossil fuels, corn ethanol release as much but probably more CO_2 than from fossil fuels. In addition, is it ethically justifiable to remove so much land from food production to produce bioethanol that, at present, is worse than fossil fuels from a climatic perspective?

To address these issues, much work has been done to produce ethanol from cellulose, a β(1,4) polymer of glucose and the most abundant biomolecule in the world. Cellulose from trees, switch grasses, and "waste biomass" are prime sources of cellulose for the production of bioethanol. Waste biomass includes stover (field crop stalks and leaves), straw, wood chips and sawdust. From one ton of corn stover, about 113 gallons of ethanol can be made, close to the 124 gallons produced from corn.

Nature breaks down cellulose routinely using cellulases, enzymes found in bacteria, fungi, protozoans, plants and some animals. Ruminants and even termites obtain cellulases from microbes living within their guts. The fungal-mediated decay of dead trees requires microbial cellulases but think how slow that process is. This stems partly from the very strong β(1,4) glycosidic link connecting glucose monomers in the polymer, which in the absence of a catalyst and at neutral pH, has an estimated half-life of 5 million years. Fossilized plants have been found to have intact cellulose and chitin, a β(1,4) polymer of N-acetylglucosamine. The β(1,4) glycosidic link in cellulose is orders of magnitude more stable than the phosphodiester bond of nucleic acids and the amide link of proteins. They are, however, readily cleaved by glycosidases such as cellulases, which can increase the k_{cat}/K_M over the uncatalyzed rate by up to 10^{17} fold, even in the absence of active site metal ions to facilitate hydrolysis (reference).

Another reason for the slow decay of dead trees is the complex structure of the cell well, and in particular, the presence of a polymer called lignin, which stabilizes the cell wall and adds considerable barriers to the access of cellulose by added glycosidases. .

A final reason for cellulose's extreme stability is the "quaternary" structure of the β(1,4)-linked cellulose strands, which consists of densely packed and intertwined strands of cellulose, which limits solvent (in this case water) accessibility necessary for hydrolysis. In addition, some exposed surface planes of the packed cellulose strands are hydrophobic. This might seem startling, given the polar nature of the glucose subunits of the polymer. Let's review it here now since our goal is present climate change from a biochemical perspective! Some of this material has been presented in previous chapters, but we will reuse it here so this chapter section can stand alone.

32.5.2: A REVIEW: THE PLANT CELL WALL

(See Chapter 7.3 for more details.) There are about 35 different types of plant cells, and each may have a different cell wall depending on the local needs of a given cell. Cells synthesize thin cell wall that extends and stay thin as the cell grows. Figure 32.5.1 shows the primary cell wall of plants. The primary cell wall contains cellulose microfibrils (no surprise) and two other polymers, pectin and hemicellulose. The middle lamella consisting of pectins, is somewhat analogous to the extracellular matrix.

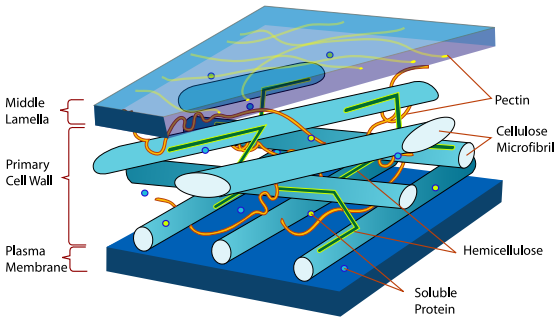


Figure 32.5.1: Primary cell wall of plants. https://commons.wikimedia.org/wiki/F...diagram-en.svg

After cell growth, the cell often synthesizes a secondary cell wall thicker than the first for extra rigidity. Since the enzymatic machinery for its synthesis is in the cytoplasm and the cell membrane, it is deposited between the cell membrane and the primary cell wall, as shown in the animated image in Figure 32.5.2.

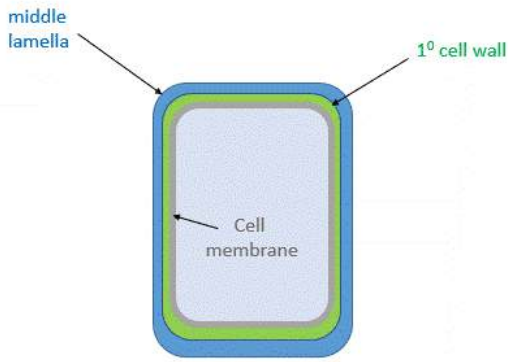


Figure 32.5.2: Primary and secondary cell wall of plants

Figure 32.5.3 shows a structural representation of both the primary and secondary cell wall.

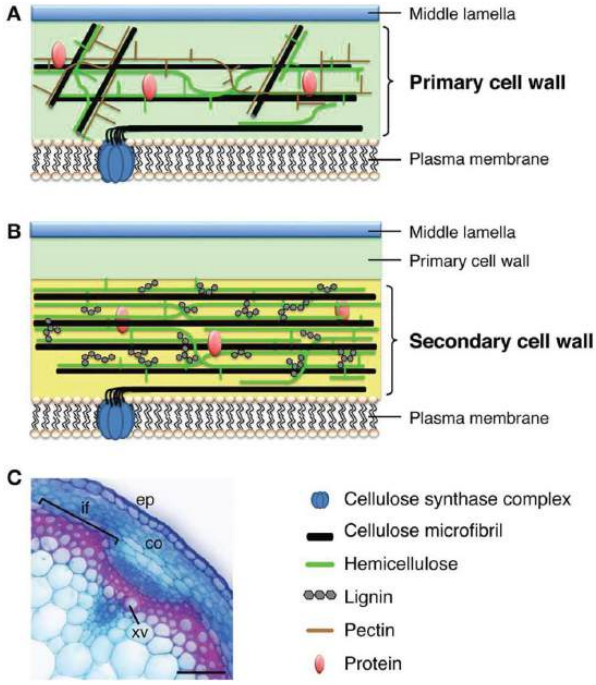


Figure 32.5.3: structural representation of both the primary and secondary cell wall of plants. Nakano Yoshimi et al. Frontiers in Plant Science (6), 288 (2015) https://www.frontiersin.org/article/...015.00288/full. Creative Commons AttributionLicense (CC BY).

The middle lamella, which contains pectins, **lignins** and some proteins, helps "glue together" the primary cell walls of surrounding plants.

Primary Cell Wall:

The main component of the primary plant wall is the homopolymer cellulose (40% -60% mass) in which the glucose monomers are linked β(1→4)-linked into strands that collect into microfibrils through hydrogen bond interactions. Two other groups of polymers, **hemicellulose,** and **pectin,** make up the plant cell wall.

Hemicellulose can make up to 20-40% by the mass These polymers have β(1,4) backbones of glucose, mannose, or xylose (called xyloglucans, xylans, mannans, galactomannans, glucomannans, and galactoglucomanannans along with some β(1,3 and 1,4)-glucans. The most abundant hemicellulose in higher plants higher plants are the xyloglucans and have a cellulose backbone linked at O6 to α-D-xylose. **Pectin** consists of linked galacturonic acids forming homogalacturonans, rhamnogalacturonans, and rhamnogalacturonans II (RGII) [12] [13]. Homogalacturonans (α1→4) linked D-GalA make up more than 50% of the pectin. Figure 32.5.4 shows some of the structures. The are generally branched, shorter than cellulose chains, and can often crystallize.

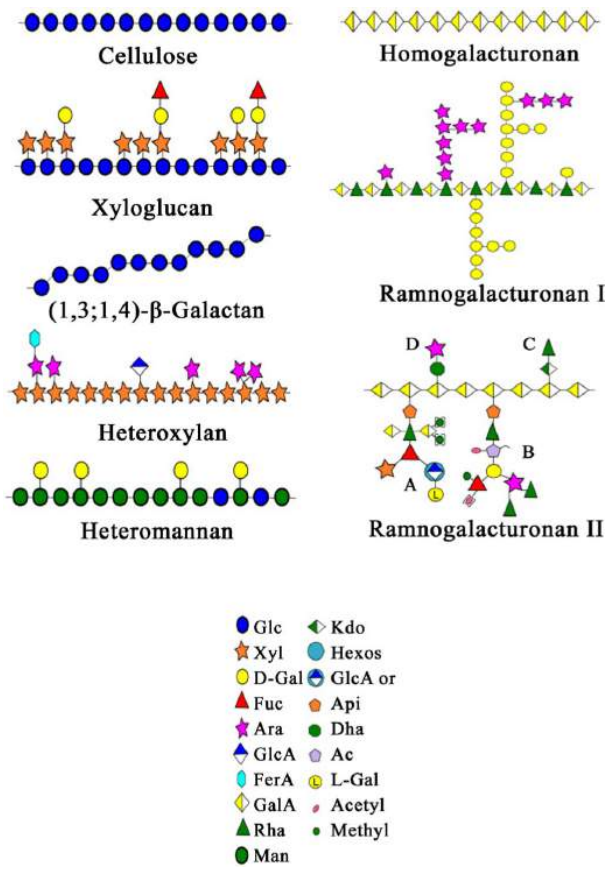


Figure 32.5.4: Variant of the cell wall components of a plant. Costa and Plazanet. Advances in Biological Chemistry 06(03):70-105. DOI: 10.4236/abc.2016.63008. License CC BY 4.0

Secondary Cell Wall

The structure of the secondary cell wall depends on the function and environment of the cell. It contains cellulose fibers, hemicellulose, and a new polymer, **lignin**. It is abundant in xylem vessels and fiber cells of woody plants. It gives the plant extra stability and new functions, including the transport of fluids within the plant through channels. The proportion of cellulose in the secondary cell wall is higher than in the primary cell wall and is less hydrated than in the primary cell wall. Given the relative volume of the secondary and primary cell walls inferred from Fig 2, most of the tree-derived cellulose for bioethanol production comes from the secondary cell wall. Switch grasses, a perennial plant, are also valuable sources of cellulose (**32–45% wt**) and hemicellulose (21–31% wt) but also have significant amounts of lignin (12–28% wt). In summary, the secondary cell wall, formed after the cell stop growing, accounts for most of the carbohydrate biomass of plants.

Glycosidases, mostly α- and β-amylases, are needed to convert corn-derived starch into glucose for fermentation and ethanol production. Likewise, cellulases are needed to degrade cellulose into glucose for cellulosic-ethanol production. However, it is a much more complex process since most of the cellulose is in the secondary cell wall. The lignin barrier in the walls protects cellulose from accessibility to cellulases, even after chemical and thermal pre-processing. In addition, xylans, which can make up 30% of the mass of the secondary cell wall, also inhibit cellulose degradation.

A thermochemical process can convert cellulose to the synthetic gases CO and H_2, which can be used as reactants to form ethanol. We'll discuss the biochemical process using pretreatment and enzymatic hydrolysis to make cellulosic ethanol. Lignin can be recovered and used to provide energy for the industrial-scale synthesis of cellulosic ethanol.

Let's explore the barriers posed by lignin and how they can be surmounted to facilitate access to cellulose and the liberation of glucose for cellulosic ethanol production.

32.5.3: LIGNIN STRUCTURE AND REACTIVITY

Lignins, which can make up to 25% of the biomass weight of secondary walls, are made from phenylalanine derivatives but more directly from cinnamic acid. This derives from is made from phenylalanine which is hydroxylated and converted through other steps to hydroxycinnamic alcohols called monolignols, as shown in Figure 32.5.5. Three typical monomers, p-coumaryl, coniferyl, and sinapyl alcohols, can polymerize into lignins, with their units in the polymer (P) named hydroxyphenyl, guaiacyl and syringly, respectively.

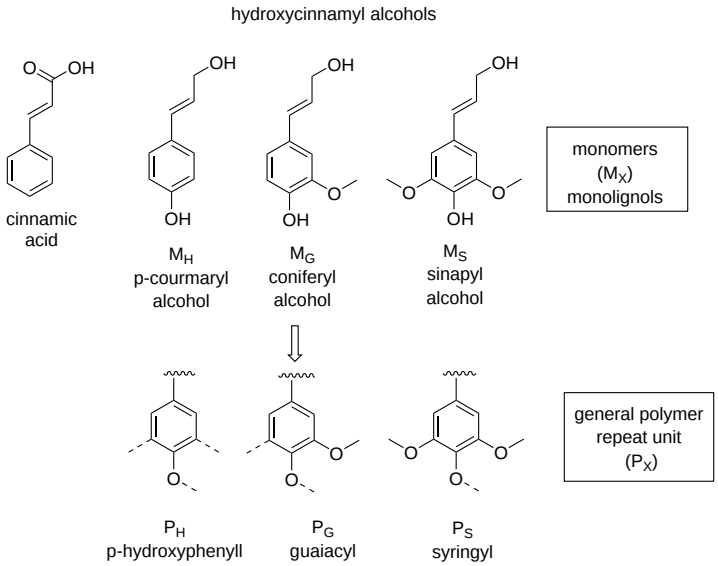


Figure 32.5.5: Monolignols and their polymers

Lignols are activated phenolic compounds, which form phenoxide free radicals (catalyzed by enzymes called peroxidases), which can attack a second lignol to form covalent dimers. Reaction mechanisms for the dimerization of the M_S sinapyl alcohol free radical are shown as an example in Figure 32.5.6.

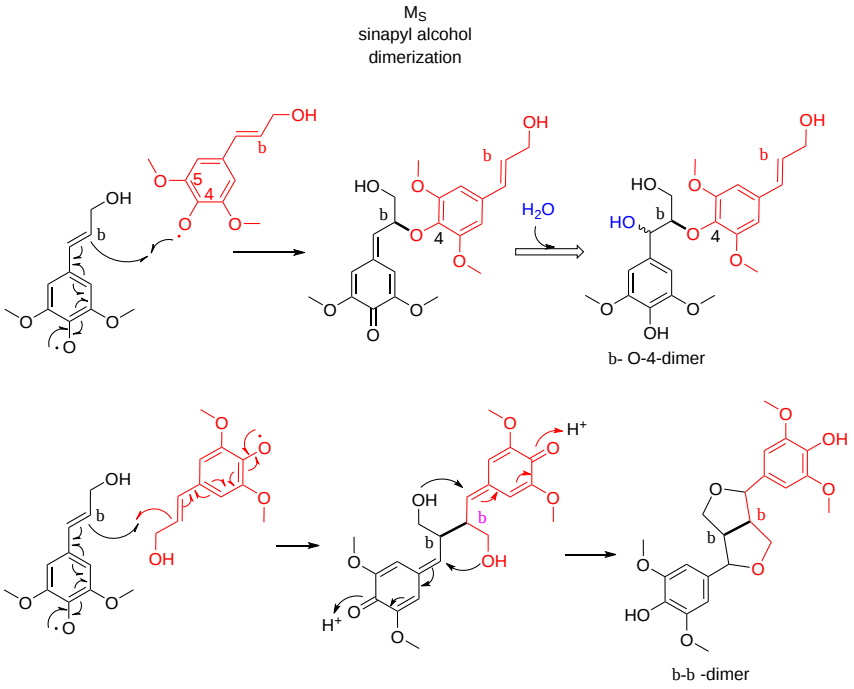


Figure 32.5.6: Dimers of lignols

Now imagine this polymerization continuing through the formation of more phenolic free radicals and coupling at a myriad of sites to form a large covalent lignin polymer. Figure 32.5.7 shows one example of a larger lignin.

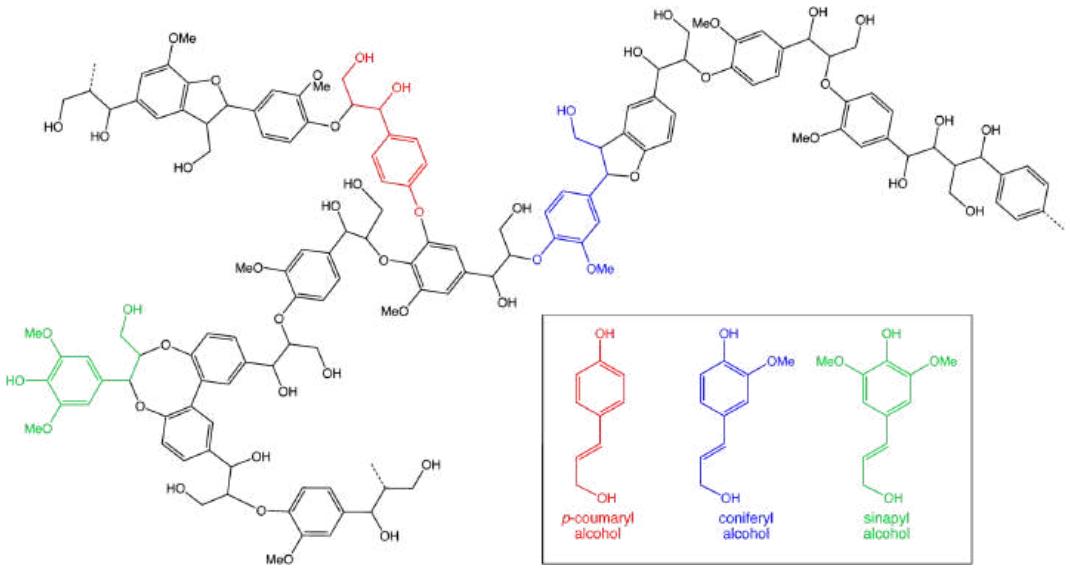


Figure 32.5.7: A larger lignin. https://commons.wikimedia.org/wiki/C...ile:Lignin.png . By Smokefoot - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=16022799

Lignin strengthens the cell wall and further stabilizes the already unreactive cellulose fibers. Let's look at a specific example - using corn stover (CS) as a cellulose source - of how pretreatment of the biomass source with a chemical treatment followed by the addition of a bacterial strain *Pandoraea* sp. B-6 (B-6) isolated from long, narrow strips of bamboo (slips). Bamboo is a type of woody grass that grows rapidly. These bacteria produce two extracellular lignin-degrading enzymes, manganese peroxidase (MnP) and laccase (Lac). Laccase (Lac) is a multi-copper oxidase that uses O_2 as an oxidizing agent in the degradation of the syringyl, guaiacyl and *p*-hydroxyphenyl monomers in lignins. MnP has similar properties. These and other enzymes can lead to the depolymerization of lignin and degradation of lignin-derived aromatic compounds

The adddition of the B-6 bacteria (a source of MnP and Lac) to milled corn stover (CS) did not increase the rate of lignin degradation unless the corn stover was preincubated with a tetrahydrofuran–water (THF–H_2O) with 0.5 wt% sulfuric acid and heated to 150 ºC. This led to the erosion of the corn stover, allowing access to the bacterial enzyme. The untreated and pre-treated CS surface, along with a diagram showing access of Lac and MnP to the lignin, is shown in Figure 32.5.8.

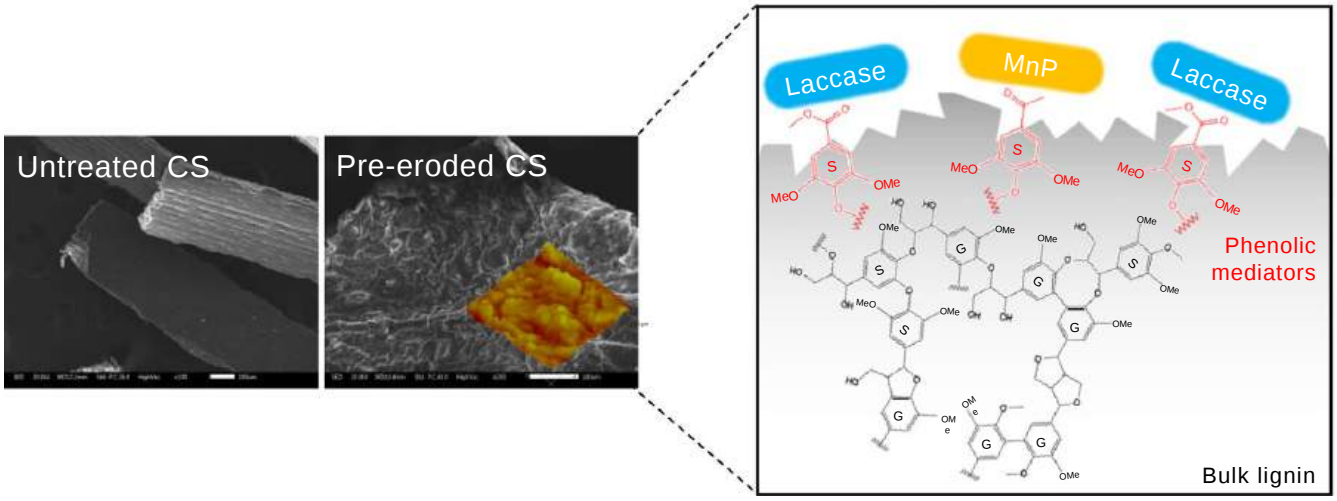


Figure 32.5.8: Untreated and pre-treated CS surface, and Lac and MnP interaction with lignins. Zhuo, S., Yan, X., Liu, D. *et al.* Use of bacteria for improving the lignocellulose biorefinery process: importance of pre-erosion. *Biotechnol Biofuels* **11**, 146 (2018). https://doi.org/10.1186/s13068-018-1...068-018-1146-4. Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/)

In addition to restricting access of cellulase to cellulose, cellulase can also nonspecifically adsorb to lignin and its pretreated forms since the lignin derivatives present a more hydrophobic surface that promotes cellulase interactions. Some plant laccases are involved in lignin biosynthesis, whereas in bacteria and fungi, they may be involved in lignin degradation

Of course, fungi, which are prime degraders of dead biomass in forests, are also sources of enzymes for lignin degradation. For example, species of white rot fungi produce manganese peroxidase (MnP), lignin peroxidase (LiP), versatile peroxidase (VP), and laccase (Lac). They work through forming reactive lignin-derived aromatic free radicals (similar to those produced in lignin synthesis), leading to breaking ether bonds, aromatic ring cleavage and removal of methoxy groups from the substrate in a process called delignification. Pretreatment of the biomass increases yields higher amounts of available cellulose. Fungi, however, grow slowly, and the rate of delignification is still low. In addition, they also have hydrolytic enzymes that decrease the yield of cellulose. That is why bacterial sources like B6 are sought for delignification.

As this is a biochemistry textbook, let's explore the structure and function of fungal laccase. The enzyme can bind a large variety of hydroxylated- and methyoxy-aromatic compounds as substrates, so its active site must be adaptable and likely dynamic. Structural analyses, in-silico docking experiments, and molecular dynamics simulations have been performed with the laccase (TvL) from the fungus *Trametes versicolor.*

The enzyme has four copper ions in a T1 Cu site and a tri-nuclear Cu cluster (T2 Cu, T3α Cu and T3β Cu) at a T2/T3 site. As the mechanism involves free radical intermediates with O_2 as an oxidant and substrate, 4 electrons are passed in single electron steps to the T1 Cu, then to the other three coppers, and finally to O_2 to form two water molecules as products. The amino acid side chain ligands for the four copper ions are shown for white rot fungi laccase from Trametes Versicolor in Figure 32.5.9.

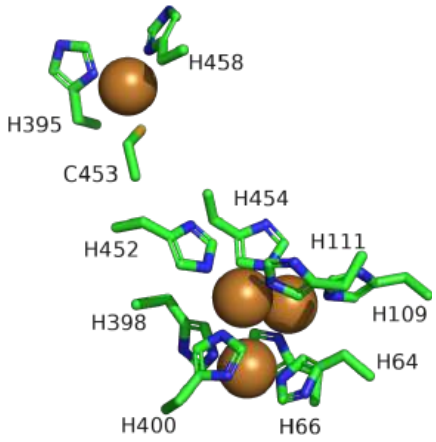


Figure 32.5.9: T1 Cu (top left) and the trinuclear Cu cluster (T2, T3α and T3β) and their ligands for Trametes Versicolor laccase (TvL, pdb: 1GYC)

Fungal laccases are extracellular proteins with about 550 amino acids arranged in three cupredoxin-like, beta-barrel domains. The T1 Cu is close to the surface and is found in domain 3, while the other copper ions are buried at the interface to domains 1 and 3. Figure 32.5.10 shows an interactive iCn3D model of Laccase from the Fungus *Trametes Versicolor* (1GYC)

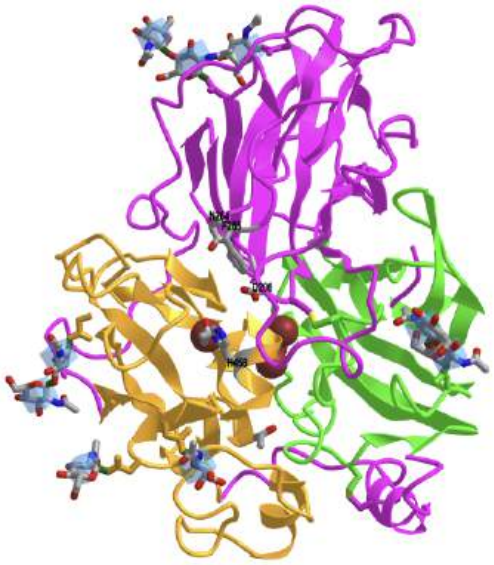


Figure 32.5.10: Laccase from the Fungus Trametes Versicolor (1GYC). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...cSYir86P9nxSW6

Domain 1 is green, domain 2 magenta and domain 3, which contains the single T1 Cu, orange. The protein is glycosylated, as shown in the blue glycan cube cartoons representing N-acetylglucosmine. Key substrate binding and catalytic side chains are shown in sticks and labeled; Asp 206 is a critical residue involved in substrate binding.

The binding interactions of TvL with a wide variety of aromatic substrates are shown schematically in Figure 32.5.11.

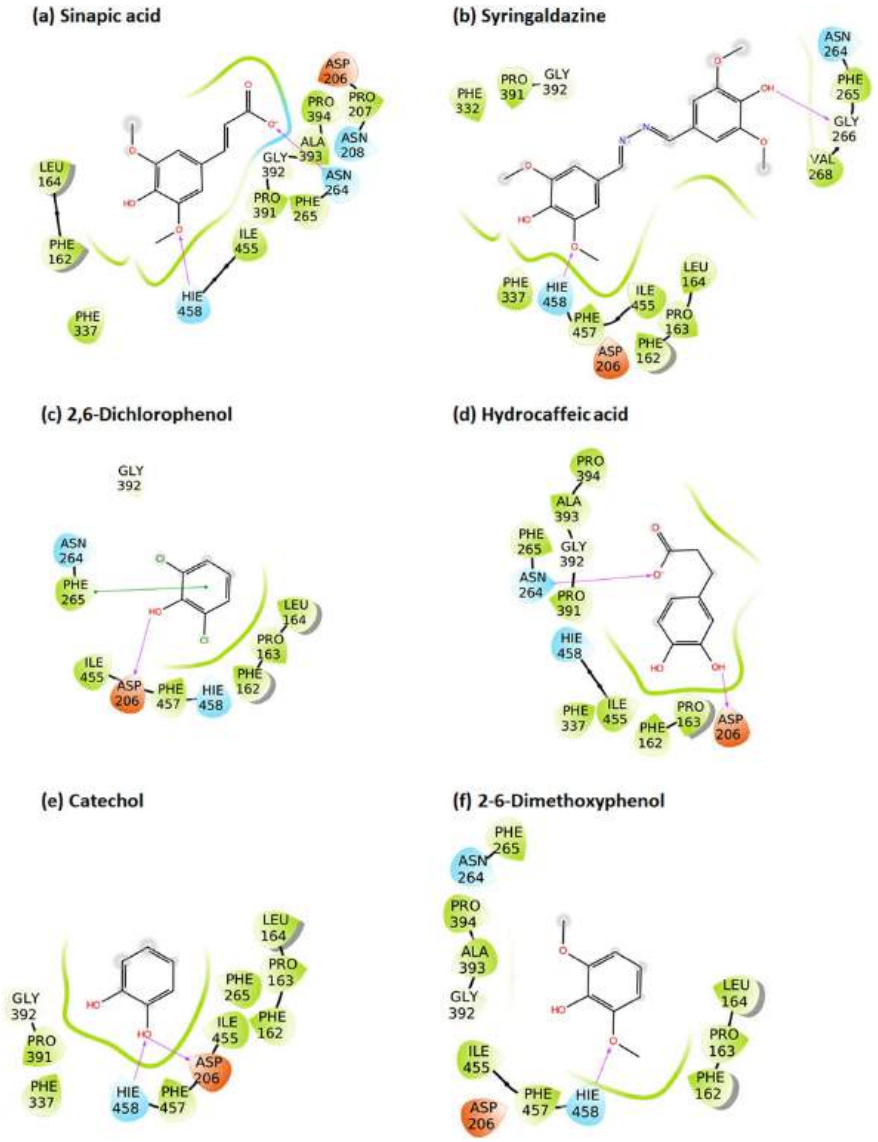


Figure 32.5.11: Binding modes of representative compounds for TvL. Mehra, R., Muschiol, J., Meyer, A.S. *et al.* A structural-chemical explanation of fungal laccase activity. *Sci Rep* **8**, 17285 (2018). https://doi.org/10.1038/s41598-018-35633-8. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/

Most substrates interact with the highly conserved His-458 (blue color, ligand for the Ti Cu)and Asp-206 (orange color) residues and form hydrogen bonds, salt bridges, or π-π stacking interactions with them. Asn-264 (blue) and Phe-265 (green) form important hydrogen bonding and π-π stacking interactions with substrates. The green color highlights nonpolar side chains. Ligands bind near the Ti Cu in domain 3 to initiate electron transfer. The active site of TvL must be dynamic to bind the various-sized ligands shown above. Molecular dynamics simulations, shown in Figure 32.5.12, support this.

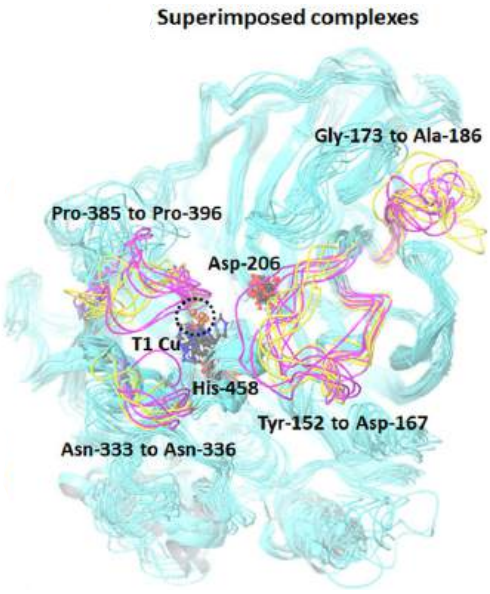


Figure 32.5.12: Display of molecular dynamics simulations showing the loop regions of TvL (magenta colored) and another laccase, CuL (yellow colored), and high levels of fluctuations. Mehra, R., ibid

32.5.4: BREAKING DOWN CELLULOSE

We just explained how the lignin barrier could be degraded so that cellulase can access cellulose. As we described above, that also poses a difficult challenge given the stability of inaccessibility of the glucosidic bonds in cellulose. The inaccessibility of "naked" cellulose fibers stems partly from the tight binding of cellulose strands into crystal lattices. Multiple crystal forms of cellulose, called polymorphs, can form. Plant cellulose has two predominant polymorphs, cellulose Iβ and Iα. Their structures are shown below in Figure 32.5.13.

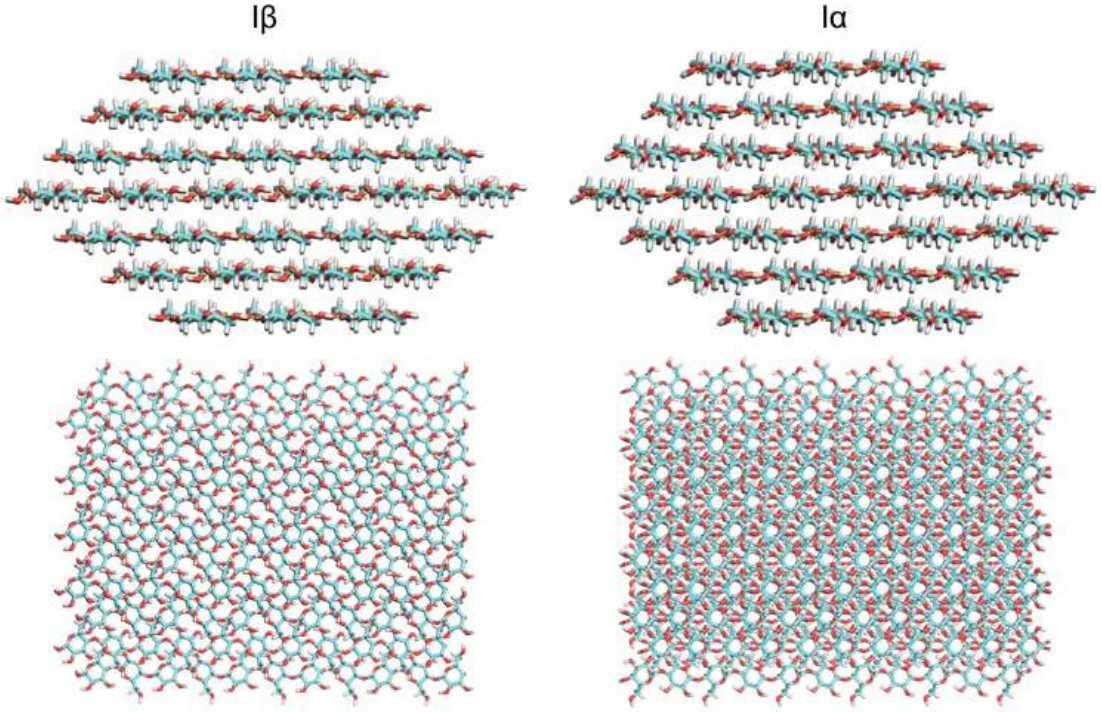


Figure 32.5.13: Natural and synthetic cellulose polymorphs. Christina M. Payne et al. *Chem. Rev.* 2015, 115, 3, 1308–1448 (2015), https://doi.org/10.1021/cr500351c. Open access through a Creative Commons public use license.

They both form hydrogen bonds within a layer, with the main differences resulting from interlayer stacking. There are no hydrogen bonds between layers. You might find that surprising at first glance until you remember that all the OH groups in the lowest energy chair form of the glucose are equatorial, which allows intralayer hydrogen bonding. The interactions between layers predominantly arise from Van der

Waals interactions, specifically induced dipole-induced dipole interactions. The hydrophobic planes, arising from axial H atoms projecting above and above each planar layer of the cellulose fibers, can be readily seen in Figure 32.5.14.

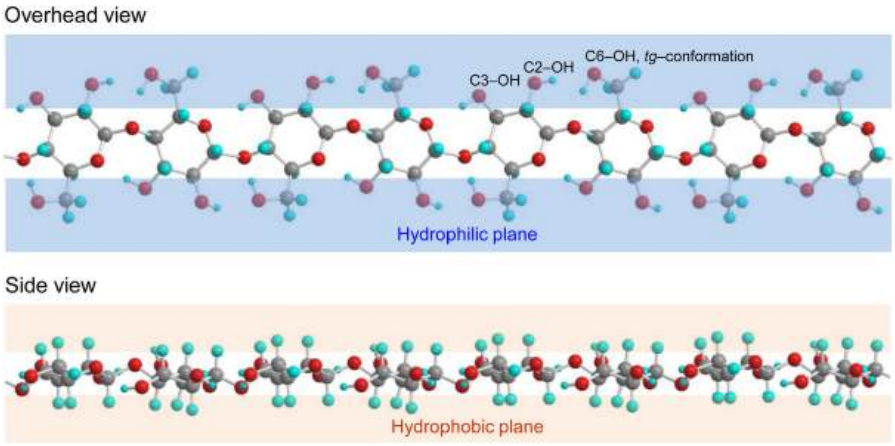


Figure 32.5.14: Hydrophobic planes arising from axial H atoms projecting above and above each planar layer of the cellulose fibers. Akira Isogai et al. Progress in Polymer Science, 86 (2018), https://doi.org/10.1016/j.progpolymsci.2018.07.007. Creative Commons license

Now we can explore the structure of cellulases and how they bind to and cleaves cellulose.

Cellulases, which cleave β(1,4) glycosidic bonds in cellulose, are members of a family of enzymes that go by many names, including **glycosidases**, or more recently, **glycoside hydrolases (GH)**. The Carbohydrate Active Enzymes (CAZypedia) has over 128 glycoside hydrolase (GH) family web pages with enzymes that form hemiacetals on the cleavage of glycosidic bonds. The fungal cellulases that work on cellulose are found in GH families 5, 6, 7, 12 and 45.

There are many types of secreted or cell-surface cellulases, including endoglucanases, exoglucanases (example is cellobiohydrolases (CBHs), and β -glucosidase (BG)). We will focus on **cellobiohydrolases (CBHs)**, the most studied one, which cleaves a 2-glucose unit (cellobiose) from either end of cellulose as it proceeds (processes) along the chain. Fungal and bacterial CBHs can work on crystalline cellulose as well. The resulting cellobiose is further cleaved by β-glucosidases. Ruminants and even termites obtain cellulases from microbes living within their guts. The enzyme has a "tunnel" between two surface loops which interacts with and processively cleaves cellulose.

As mentioned above, fungi are the major degraders of biomass, and are critical in the carbon cycle. Some fungi(brown-rot) use the Fenton reaction (Chapter 13.3) to produce the very reactive hydroxyl free radical (·OH) which causes biomass degradation. Filamentous fungi (like white and soft rot like *T. reesei*) use enzymes. The T. reesei cellulase called cellobiose hydrolase 1 is more recently named *Tr*Cel7A as it is in the GH 7A family.

Cellulase mechanism.

We have previously described the mechanisms of polysaccharide synthesis (Chapter 20.3), so we will discuss in less detail the mechanism of the very similar reaction of cellulose degradation by cellulase. Two general mechanisms are possible, one leading to the retention of configuration at the resulting hemiacetal end of the cellobiose and another that inverts the configuration. These mechanisms are shown in Figure 32.5.15 for glucose in alpha-linkage at the anomeric carbon (not the beta-linkage found in cellulose).

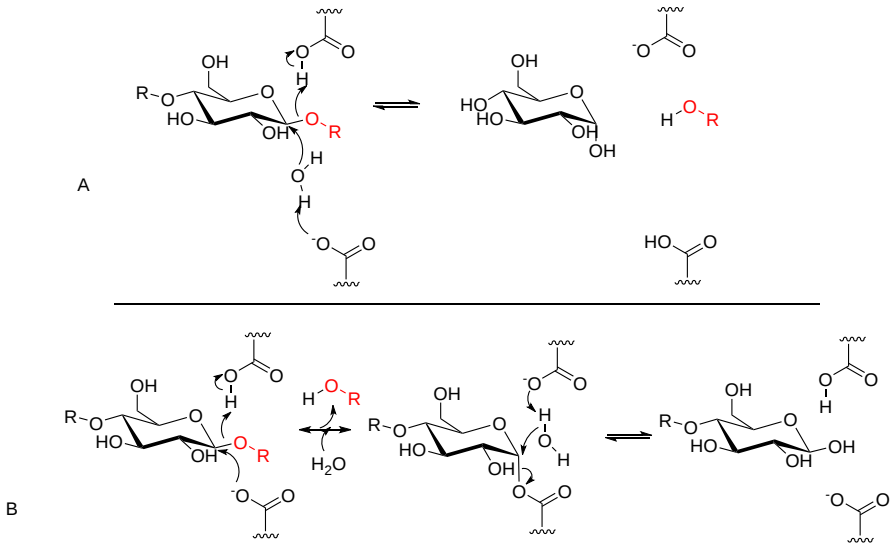


Figure 32.5.15: Two Primary Catalytic Mechanisms of GHs. After Payne et al. *Chem. Rev.* 2015, 115, 3, 1308–1448. https://doi.org/10.1021/cr500351c.

Scheme (A) shows the reaction that inverts the configuration. Water acts as a nucleophile in a S_N2 type of reaction, with catalytic assistance by two proximal carboxylic acid side chains acting as general acids and bases. This results in an inversion of the stereochemistry at the anomeric carbon.

Scheme (B) proceeds with the retention of configuration as two different nucleophilic attacks occur. In the first, an active site carboxylate forms a covalent acetal intermediate with the anomeric carbon. The carboxylate hence acts as a nucleophilic catalyst. Water, acting as a nucleophile, then attacks to form the hemiacetal with the expulsion of the carboxylate leaving group. As we discussed in Chapter 20.3 (section on glycosyl transferases), other variants of these mechanisms would include a S_N1 reaction or one with an oxocarbenium-like transition state.

The CBH1 (family 7) has a long tunnel for binding cellulose. The CBH1 (*Tr*Cel7A) cellulose catalytic site spans at least 9 glucose monomers (n-7, n-6,...,n-1,n+1, n+2) with cleavage typically of a cellobiose from the reducing end (between n-1 and n+1). The structure of the *Tr*Cel7A glycoside hydrolase (cellobiose hydrolase) with a small bound cellulose is shown in Figure 32.5.16.

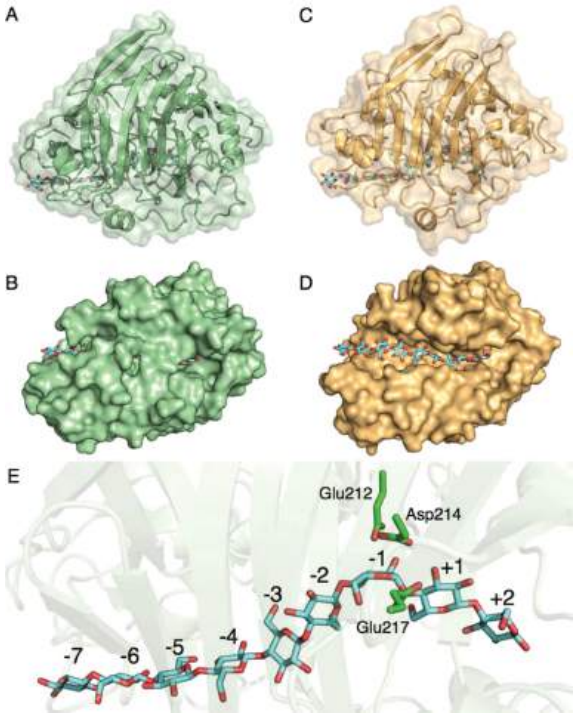


Figure 32.5.16: Crystal structure of the first GH7 CBH and EG. Payne et al. *Chem. Rev.* 2015, 115, 3, 1308–1448. https://doi.org/10.1021/cr500351c Open access through Creative Commons public use license.

The ligand from the *Tr*Cel7A Michaelis complex (PDB code 4C4C (441)) is shown in all panels. (A) CBH *Tr*Cel7A CD (PDB code 1CEL (172)) view from side, exhibiting the β sandwich structure that is characteristic of GH7 enzymes. *Tr*Cel7A was the first GH7 structure solved and is the best-characterized member of GH7. (B) *Tr*Cel7A view from bottom showing the more closed substrate binding "tunnel". (C) EG *F. oxysporum* Cel7B (PDB code 1OVW (174)) view from side. (D) *Fo*Cel7B view from the bottom showing the more open binding "groove". (E) *Tr*Cel7A Michaelis complex (PDB code 4C4C (441)) shows the standard numbering of the substrate binding sites (catalytic residues shown in green for reference). A cellulose chain enters from the −7 site. Hydrolysis occurs between the −1 and +1 sites. The +1/+2 sites are termed the "product sites".

Active site carboxylates (E212, D214, and E217) are shown near the -1/+1 cleavage site in Figure 32.5.17. Glu 217 is covalently attached to the -1 glucose, supporting the retaining mechanism illustrated in Fig 15 above.

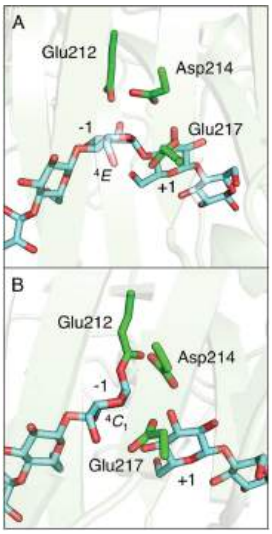


Figure 32.5.17: Michaelis complex and glycosyl-enzyme intermediate of *Tr*Cel7A. Payne et al. ibid.

Panel (A) shows the *Tr*Cel7A Michaelis complex (PDB code 4C4C (441)).

Panel (B) shows a *Tr*Cel7A glycosyl-enzyme intermediate (PDB code 4C4D (441)) with a covalent bond between the nucleophile and the broken cellooligomer chain. There is an approximate 30° rotation of the E217 nucleophile during glycosylation.

Figure 32.5.18 shows a more detailed view of the first step (glycosylation f Glu 217) for the *Hypocrea jecorina* GH Family 7 cellobiohydrolase Cel7A

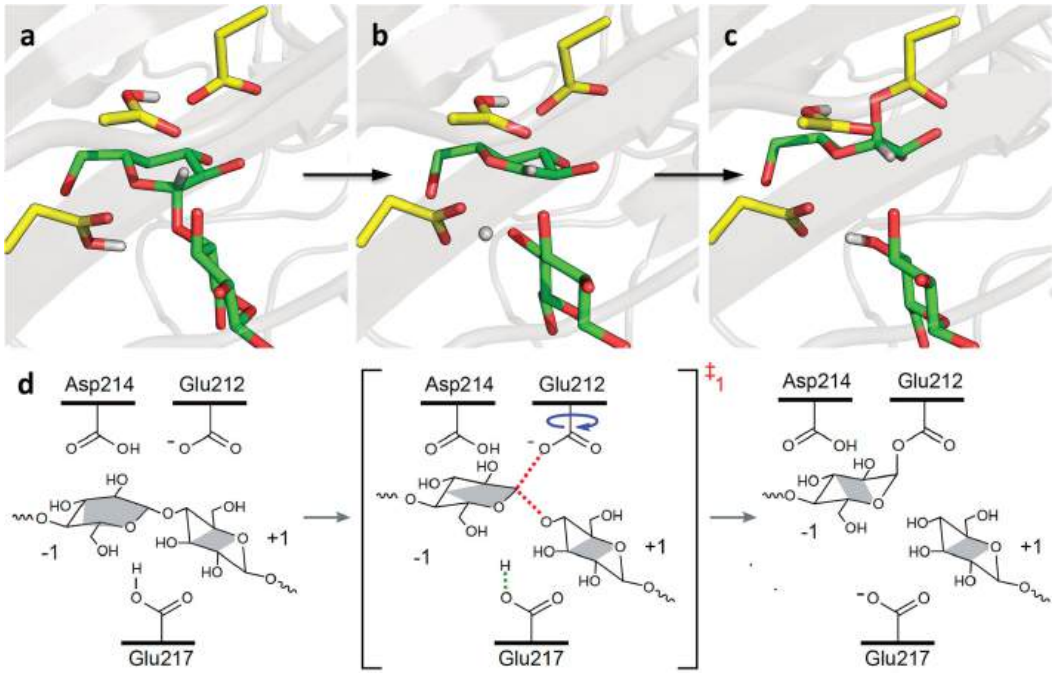


Figure 32.5.18: Figure 2. Glycosylation step for *Hypocrea jecorina* GH Family 7 cellobiohydrolase Cel7A. Knott, Brandon C. et al. - J. Am. Chem. Soc.329 (2013) https://doi.org/10.1021/ja410291u. Open access article published under an ACS AuthorChoice License

Panel (a) shows a snapshot of the reactant the conformation from a representative AS trajectory (with the substrate in green and catalytic residues in yellow) for the glycosylation step. The proton resides on the acid residue, Glu217.

Panel (b) shows a representative snapshot of the transition state. The −1 glucopyranose ring now adopts a different conformation.

Panel (c) shows the product of the glycosylation reaction.

Panel (d) shows a schematic view of the overall glycosylation reaction with the collective variables identified by LM colored at the transition state. The best three-component RC identified by LM includes the forming/breaking bonds involving the anomeric carbon, the breaking bond between Glu217 and its proton, and the orientation of the nucleophile Glu212.

Figure 32.5.19 shows the corresponding deglycosylation (of Glu 217) step.

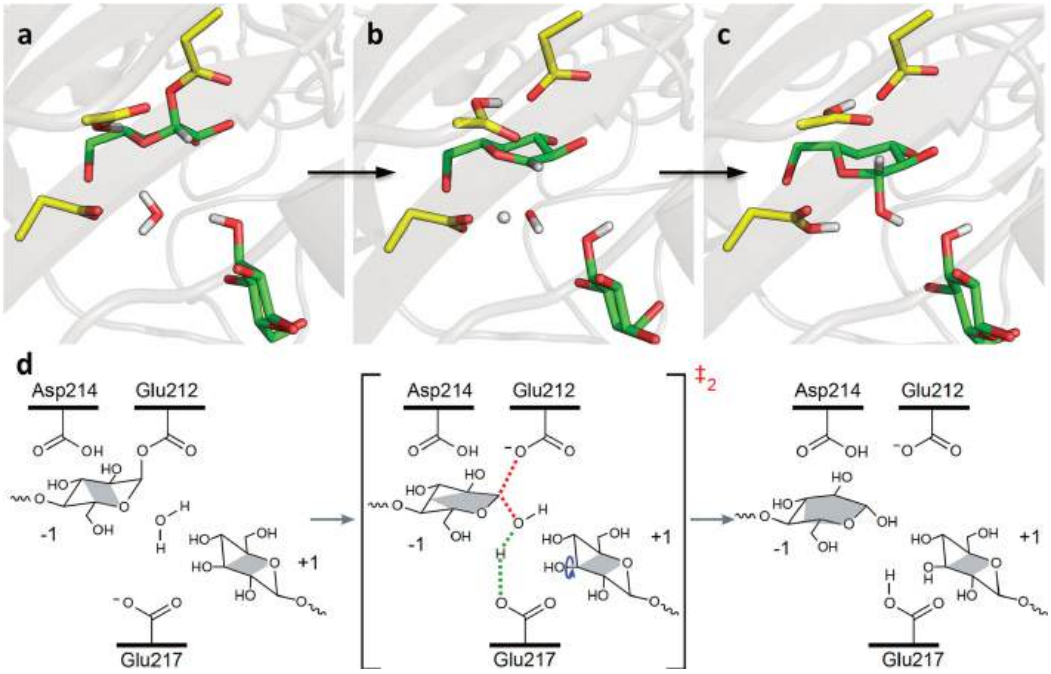


Figure 32.5.19: Figure 4. Deglycosylation step results. Knott, Brandon C. et al, ibid

Panel (a) shows a snapshot of the reactant conformation from a representative AS trajectory (with the substrate in green and catalytic residues in yellow) for the deglycosylation step. The covalent glycosyl–enzyme bond is intact, and the cellobiose product is in primed GEI mode.

Panel (b) shows a representative snapshot of the transition state. Note the distorted conformation of the −1 sugar, as the nucleophilic water molecule is ripped apart.

Panel (c) shows a snapshot of the product in which the glycosyl-enzyme bond has been broken, and the catalytic residues have been regenerated.

Panel (d) shows a schematic view of the overall deglycosylation reaction with the collective variables identified by LM colored at the transition state. The best three-component RC identified by LM includes the forming/breaking bonds involving the anomeric carbon, the forming/breaking bonds involving the transferring proton, and the orientation of the C3 hydroxyl of the +1 sugar.

32.5.5: BINDING OF CELLULASE TO CELLULOSE FIBERS AND LIGNIN

Many glycoside hydrolases contain distinct carbohydrate binding domains/modules (CBD/CBM) and catalytic domains (CD). For example, TrCel7A can be cleaved by the protease papain into a 56K domain with catalytic activity on small substrates but not large cellulose one and a smaller 10K (C terminal) domain that itself is glycosylated and which binds to the hydrophobic surface of cellulose crystals.

Many GHs, in addition, have **linkers** connecting the catalytic domain (CD) and the carbohydrate module (CBM), which add different functions to the enzymes. The linkers vary in size and amino acid sequence. Linkers in fungi tend to be long and N- and O-glycosylated, affecting binding/catalysis. The linkers can also be intrinsically disordered, which adds dynamic complexity to their effects.

The actual cellulose binding site on cellulase has been determined by solution NMR using a synthetic 36 amino acids protein fragment from the C-terminal domain of *Trichoderma reesei* Cel7A (the "carbohydrate binding module or CBM"). The amino acids involved in the binding of cellohexaose (6-mer) were determined by perturbation of the 2D NMR structure on binding cellohexaose. As we mentioned above, cellulase also binds lignin, decreasing their catalytic efficiency towards cellulase. Results of NMR binding studies of the *Tr*Cel7A carbohydrate binding module with cellohexaose and lignins from Japanese cedar (C-MWL) and *Eucalyptus globulus* (E-MWL) are shown in Figure 32.5.20.

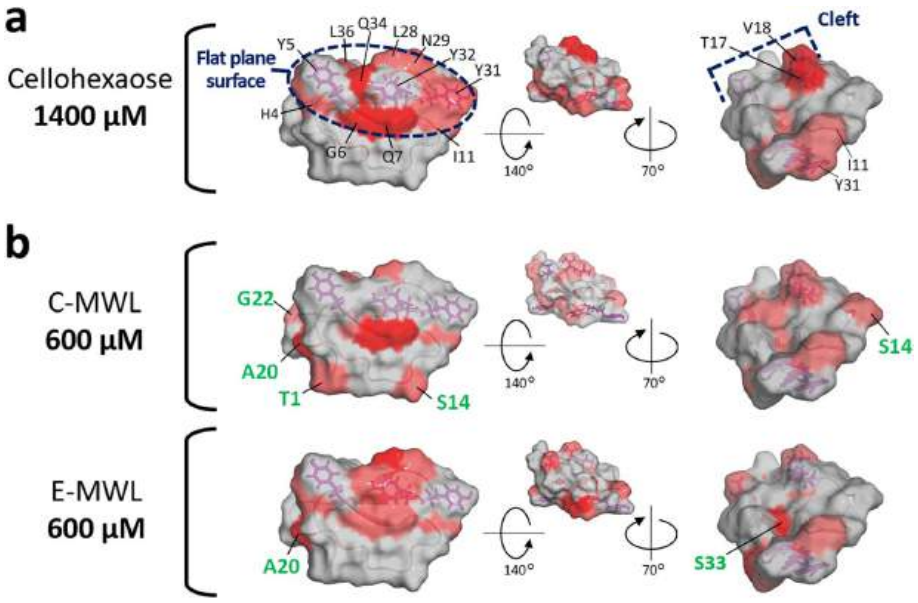
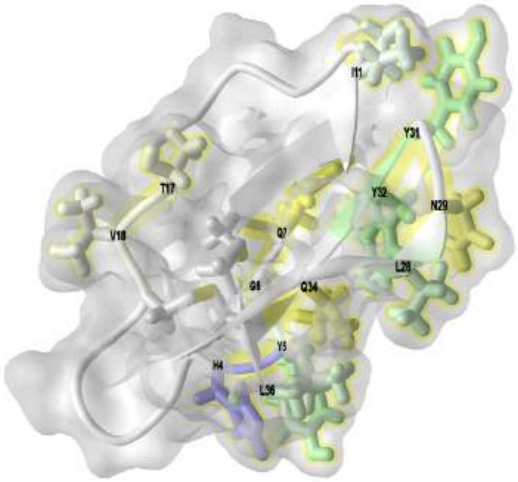


Figure 32.5.20: Comparison of interaction property between cellohexaose and MWLs. Tokunaga, Y., Nagata, T., Suetomi, T. *et al.* NMR Analysis on Molecular Interaction of Lignin with Amino Acid Residues of Carbohydrate-Binding Module from *Trichoderma reesei* Cel7A. *Sci Rep* **9**, 1977 (2019). https://doi.org/10.1038/s41598-018-38410-9. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Panel (**a**) shows cellohexaose specifically bound to the flat plane surface and cleft. The flat plane surface is defined by a triplet tyrosine (Y5, Y31, Y32) and H4, G6, Q7, I11, L28, N29, Q34, L36**.**

Panel (**b**) shows both MWLs bound to multiple binding sites, some of which are included in the flat plane surface and cleft even in low concentrations of titrant. These non-specific binding sites are labeled green.

Figure 32.5.21 shows an interactive iCn3D model of the C-terminal cellulose-binding module of cellobiohydrolase I from Trichoderma reesei (2CBH).



NCBI iCn3D Figure 32.5.21: C-terminal cellulose-binding module of cellobiohydrolase I from Trichoderma reesei (2CBH). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...Yva8BVMQCGdUZ7

This synthetic carbohydrate binding module (CBM) from the C-terminal domain of cellobiohydrolase I consists of 36 residues. NMR was used to determine the structure of the CBM. 2D NMR was used to determine the amino acids interacting with cellulose. The interacting side chains are shown as sticks underneath the molecular surface (gray). The side chains are colored according to hydrophobicity, with green followed by yellow being most hydrophobic. The numbering system refers to the 36 amino acid synthetic peptide, not the native protein. This model clearly shows this domain binds to the hydrophobic face of the cellulose microcrystal.

The *Tr*Cel7A CBM has serine and threonine side chains that are glycosylated and affect binding. The interaction of the CBM with the nonpolar cellulose surface is shown in Figure 32.5.22.

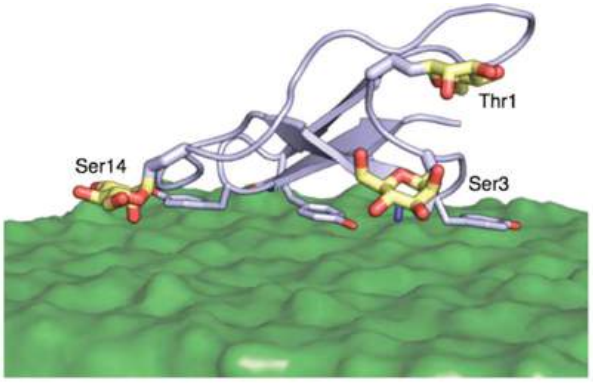


Figure 32.5.22: Glycosylated *Tr*Cel7A CBM on the hydrophobic surface of cellulose. Payne et al., ibid.

Note that the aromatic groups of the triplet tyrosines, Y5, Y31, and Y32 (not labeled), are coplanar with the cellulose surface.

As mentioned above, linkers that connect the C-terminal carbohydrate binding module (CBM) and the catalytic domain (CD) can be of different lengths and sequences and are also N- and O-glycosylated. Figure 32.5.23 shows the interactions of the glycosylated linkers with the cellulose fibers.

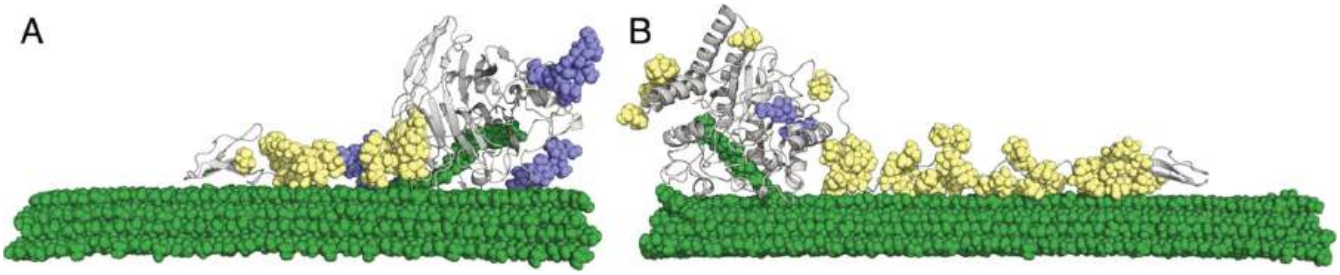


Figure 32.5.23: Molecular snapshots of *Tr*Cel7A and *Tr*Cel6A wherein the linker binds to the cellulose surface from microsecond-long MD simulations. Payne et al., ibid.

These computational predictions of cellulose linkers enhancing binding of CBMs to the cellulose surface were corroborated experimentally via binding isotherm measurements. *N*-glycosylation and *O*-glycosylation are shown in blue and yellow. The glycans attached to the enzyme significantly enhance the binding of cellulase to the cellulose fibers. Payne et al. *Chem. Rev.* 2015, 115, 3, 1308–1448. https://doi.org/10.1021/cr500351c Open access through Creative Commons public use license

A pictorial view of the hydrolytic cleavage of cellobiose from cellulose fibers is shown in Figure 32.5.24.

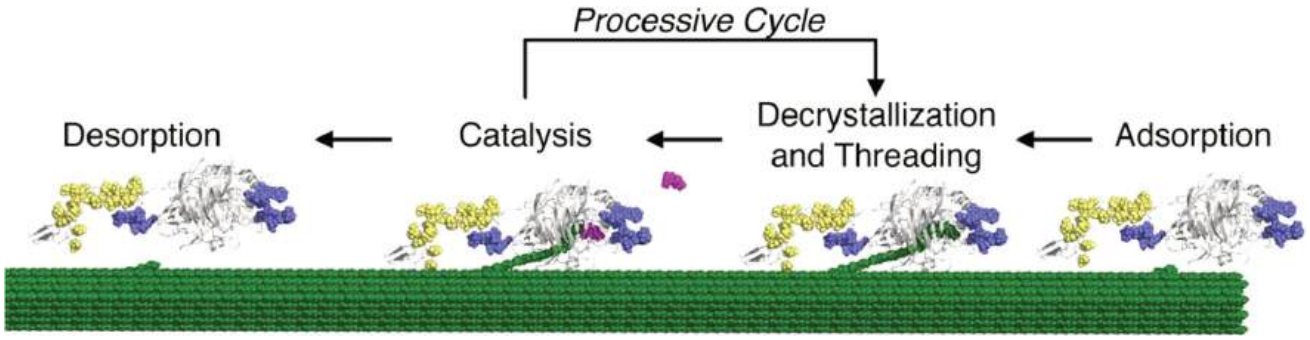
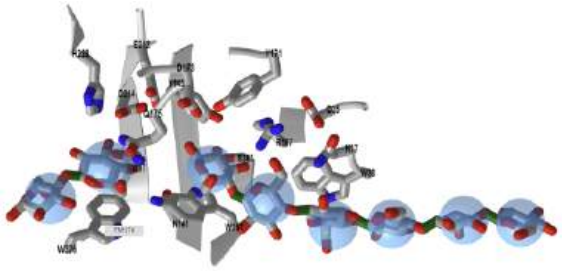


Figure 32.5.24: Complete processive cycle of a GH7 CBH. *Tr*Cel7A is shown with its CD, linker, and CBM in gray "cartoon" representation. Payne et al., ibid.

N-glycosylation and *O*-glycosylation are shown in blue and yellow, respectively. The cellulose surface is green, and the released cellobiose product magenta. Following the CBM and CD adsorption to the substrate and initial chain threading, *Tr*Cel7A processively cleaves cellobiose from a cellulose chain end. The "Processive Cycle" includes chain processivity, hydrolysis, and product expulsion (Figure 35). This processive cycle repeatedly occurs until the enzyme desorbs from the cellulose surface.

Figure 32.5.25 shows an interactive iCn3D model of cellulose bound to cellobiohydrolase I from Trichoderma reesei (7CEL)



NCBI iCn3D Figure 32.5.25: Cellulose bound to cellobiohydrolase I from Trichoderma reesei (7CEL). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...hfJFWg6hmJAxg7

KEY POINTS - BETA VERSION FROM CHAT.OPENAI

1. Cellulosic ethanol is a biofuel that is produced from the cellulose, hemicellulose, and lignin in plant material.
2. Unlike first-generation biofuels like corn and sugar cane ethanol, which are produced from sugars and starches, cellulosic ethanol can be produced from a wide range of plant material, including agricultural waste, wood chips, and switchgrass.
3. Cellulosic ethanol is considered a "second-generation" biofuel because it addresses many of the limitations of first-generation biofuels, including the competition with food crops for land and resources.
4. Cellulosic ethanol production involves a two-step process: first, the plant material is broken down into sugars through a process known as pretreatment, and then the sugars are fermented to produce ethanol.
5. The most common pretreatment methods include acid hydrolysis, ammonia fiber expansion, and steam explosion.
6. Cellulosic ethanol has a higher energy balance than first-generation biofuels and lower greenhouse gas emissions, making it a more sustainable option for biofuel production.

7. However, the technology for cellulosic ethanol production is still in the early stages of development, and the cost of production remains high.
8. Research is ongoing to improve the efficiency and cost-effectiveness of cellulosic ethanol production and to find ways to make it a viable alternative to fossil fuels.

32.6: ALGAE - AN INTRODUCTION

Search Fundamentals of Biochemistry

In the previous chapter sections, we discussed bioethanol production from plant starches (first generation) and lignocellulosic plants (second generation) from biomass waste such as stover and sugar cane. Each had its challenges. Now we will consider the third generation production of bioethanol from algae which has significant potential for minimizing damage to the environment. Below in Figure 32.6.1 is a summary of bioethanol production from each generation feedstock.

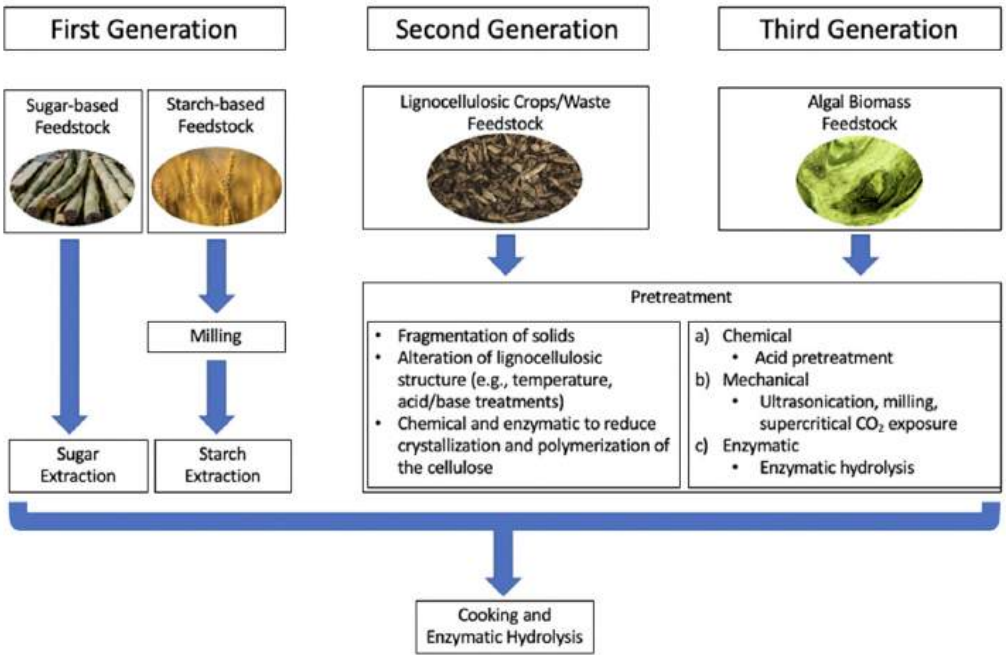


Figure 32.6.x: **Figure 1.** General flowchart of bioethanol production, comparing the pre-fermentation processing of feedstocks for the first three generations of bioethanol production. The blue highlighted area provides an example of a value-added process that can enhance the value of bioethanol production. Tse, T.J.; Wiens, D.J.; Reaney, M.J.T. Production of Bioethanol—A Review of Factors Affecting Ethanol Yield. *Fermentation* **2021**, *7*, 268. https://doi.org/10.3390/fermentation7040268. Creative Commons Attribution (CC BY) license (**https://creativecommons.org/licenses/by/4.0/**)

There are so many types of algae that it can be daunting to read about them. Some are single cells, some can form filaments and colonies, and some are multicellular with different cell types. Some are prokaryotes; some are eukaryotes.

The Algae Database suggests that they are best defined as "oxygenic photosynthesizers other than embryophyte land plants". Naming and classification are generally based on botanical names ending as follows: phyla = -phyta, classes = -phyceae, order = -ales, and family = -aceae. The main groups are blue-green algae (cyanobacteria), rhodophytes, phaeophytes, chlorophytes, euglenophytes, charophytes, diatoms, dinoflagellates, and cryptophytes. To make it more complicated, some animal species have been categorized as algae and have zoological name endings (-zoa, -ea, -ida, -idae). We apologize in advance for any errors and inconsistencies in the description of algae in the chapter section and would ask that you contact us with corrections..

Before we discuss biofuel production from algae, we will review the different algae types (much as we did for zooplankton and phytoplankton). Broadly, algae can be divided into microalgae (seen with a microscope) and macroalgae (seen with the eye). There are almost 170,00 species of algae listed in the Algaebase.

Here is a summary of their properties. They can

- fix carbon and produce food by photosynthesis. As such, they are primary producers.
- be tiny (microalgae) or large and visible (macroalgae, also known as seaweed). Kelp, which can form large underwater forests, is a macroalgae.
- be unicellular, form colonies or filaments, or larger multicellular structures
- attach to objects or float freely.

Some who study algae (phycologists) often consider any organism with chlorophyll but without the stems, roots, leaves, flowers, and vessels of plants to be algae.

The naming and classification of algae are confusing, so let's start with a broad review of their classification.

Life can be divided into domains, kingdoms, phyla, and additional subcategories. Carl Woese proposed three domains, Archaea, Bacteria, and Eukarya, in 1990, based on analyses of ribosomal RNA sequences. These domains are further classified into kingdoms - Archaebacteria, (Eu)bacteria, Protista, Fungi, Plantae, and Animalia. A seventh kingdom was added in 1981 by Thomas Cavalier-Smith, who divided Protista was two kingdoms, Protista (unicellular eukaryotes like some protozoa and some molds) and a new kingdom, Chromista (uni- or multicellular eukaryotes such as algae, diatoms, and some protozoans). Both Protista and Chromista have organisms with chlorophyll, and both also have heterotropic organisms. Newer classifications based on additional biochemical data may yet be proposed. Table 32.6.1 below shows a summary of the domains and kingdoms of life.

Domains and Kingdoms							
Domain	Bacteria	Archaea	Eukarya				
Kingdom	(Eu)bacteria	Archaebacteria	Plantae	Animale	Fungi	Protista	**Chromista**

Table 32.6.1: Domains and kingdoms of life.

All life arose from the **last universal common ancestor, LUCA**, as shown below in Figure 32.6.2.

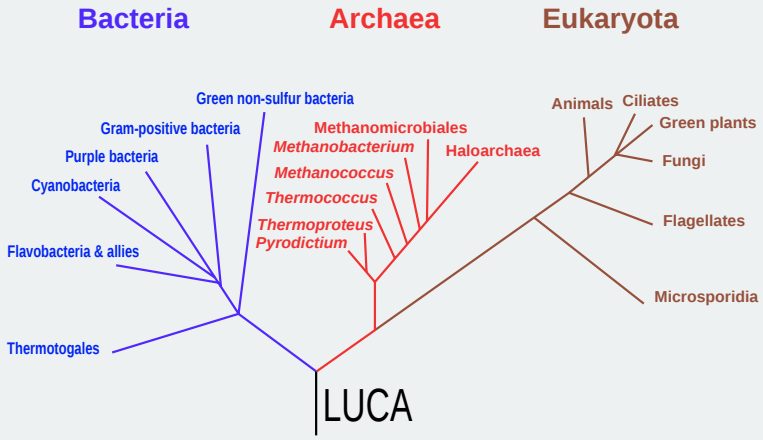


Figure 32.6.2: Phylogenetic tree linking all major groups of living organisms to the LUCAhttps://upload.wikimedia.org/wikiped..._1990_LUCA.svg

Their biological classification of algae (which illustrates how widely spread they are among different domains, kingdoms, and phyla, shown below in Figure 32.6.3

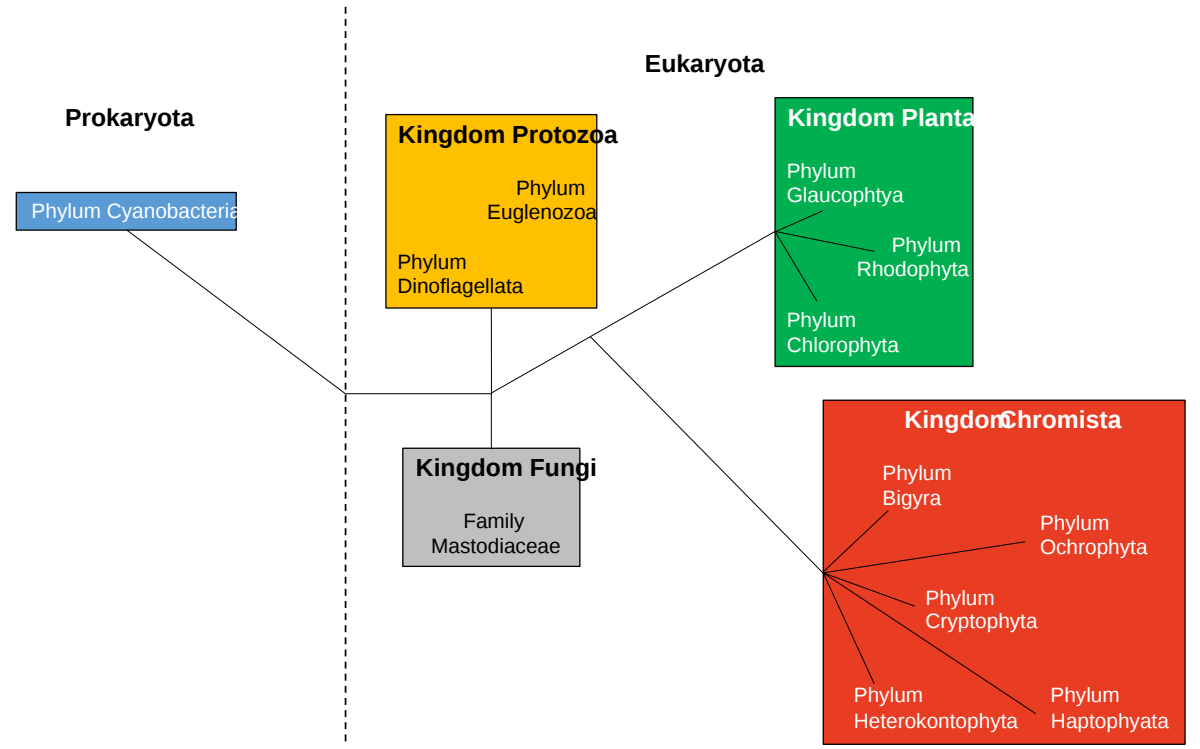


Figure 32.6.3: Distribution of algae among groups in the Tree of Life as recognized by the ITIS and Species 2000 (and ife.org) in 2011. The deep classification of algae is the subject of great debate, and even the higher clades have been discussed and revised recently. Adapted from Verdelho Vieira, V.; Cadoret, J.-P.; Acien, F.G.; Benemann, J. Clarification of Most Relevant Concepts Related to the Microalgae Production Sector. *Processes* **2022**, *10*, 175. https://doi.org/10.3390/pr10010175. Creative Commons Attribution (CC BY) license (**https://creativecommons.org/licenses/by/4.0/**).

32.6.1: MICROALGAE

Microalgae are single-cell organisms that can form filaments and colonies. This group has one prokaryotic member, **cyanobacteria**, also known as **blue-green algae**. As a prokaryote, it does not have mitochondria or chloroplasts. The rest of the microalgae are eukaryotic and include the phyla Chlorophyta, Rhodophyta, Glaucophyta, Cryptophyta, Euglenozoa, Cercozoa, Heterokontophyta, Haptophyta, and Miozoa (Myzozoa). Another simpler organizational system divides microalgae into four categories:

- cyanophyta (blue-green algae/cyanobacteria)
- pyrrophyta (dinoflagellates and cryptomonads, and can be yellowish-green to golden-brown)
- chrysophyta (diatoms, heterokonts and golden brown algae)
- chlorophyta (microscopic green algae). This term also applies to macroalgae (see below).

Green algae are really diverse and are now broken up into two phyla, chlorophyta and charophyta with a combined 17 classes. AlgaeBase dynamic species counts shows that there are about 4,500 species of **Chlorophyta** including those that live on land, in freshwater, and those that are considered macroalgae seaweeds. There are about 2500 species of **Charophyta** that are entirely freshwater.

Within the microalgae are found green, brown, and red algae. (Unfortunately for students of algae, there are also green, brown, and red macroscopic algae as well). Microalgae are much more efficient at photosynthesis than land plants. The diatoms are an example of microalgae and are the precursor of the "fossil" fuel oil deposits. Spirogyra is a unicellular green alga that forms long filaments (colonies) up to 0.1 mm in length, so it loosk like a multicellular organism.

One class of microalgae, green algae, arose when a microbe acquired a cyanobacterium, which allowed photosynthesis. Green algae eventually evolved into higher plants, and a similar process led to red algae. Brown algae and diatoms, dinoflagellates, and euglenoids, other types of algae, arose from the uptake of red or green algae into other eukaryotic host cells.

In general, green microalgae, which absorb red wavelengths, are found on the surface. The red microalgae, which absorb green, and blue wavelengths, are at lower levels. The brown microalgae are generally found in between these water layers. Figure 32.6.x below shows the light penetration spectrum in water.

LIGHT IN THE WATER

The energy of light photons is given by $E=h\nu=hc/\lambda$, where ν if the frequency of the light and λ is the wavelength. As highly energetic x-rays penetrate matter, visible light can penetrate water. Water absorbs incident radiation, with the lower energy, higher wavelength photons of red light absorbed more readily in the top layers while the blue light penetrates farther into the water. The depths of penetration of light in the open ocean (left) and coastal (right) waters are shown in Figure 32.6.4.

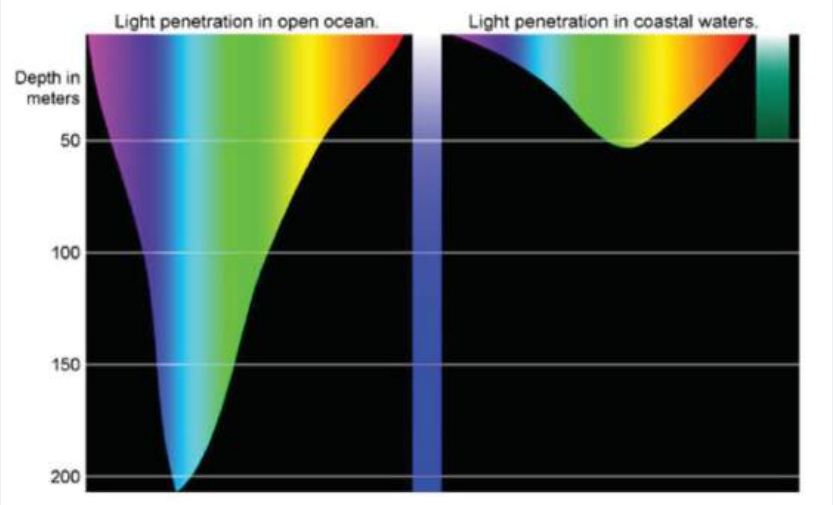


Figure 32.6.4: The light penetration spectrum in water as a function of color. https://oceanexplorer.noaa.gov/edu/m...fact-sheet.pdf

When you are underwater, everything seems blue since red light is preferentially absorbed, leaving photons enriched in blue light reaching our eyes. Blue-enriched light reaches our eyes as it reflects off of objects. Clear water in the open ocean also appears blue since there are fewer particles like phytoplankton from which light scatters back to our eyes. Rayleigh (or elastic light scattering) depends on $1/\lambda^6$, so low wavelength light scatters most from particles. Sunsets and sunrises appear red when light passes through more of the atmosphere and the blue light is scattered from atmospheric particles before reaching our eyes. Coastal waters have more sediment, algae, and microscopic organisms like plankton that can scatter light. These waters appear more green-blue, since red light is significantly absorbed by the microorganisms and blue light is more scattered.

Microalgae exist as single cells or can form multicellular filaments and colonies. They proliferate in the presence of simple nutrients and can produce large amounts of polysaccharides for industrial bioethanol production or fatty acids/triacylglycerols for biodiesel production. They are, as mentioned above, the source of underground oil deposits. Both freshwater (like *Chlorella* and *Haematococcus*) and ocean microalgae (*Dunaliella*, *Phaeodactylum*, and *Tetraselmis)* can be used for the production of biofuels. The cyanobacteria *Spirulina* sp. is commonly used for commercial purposes. It and Synechococcus have large amounts of glycogen that could be used for bioethanol production. Their triacylglycerols are low, so they can't be used for biodiesel (which we will discuss in a future section of the chapter).

Most are familiar with algae blooms in freshwater lakes (and even in saltwater environments). Cyanobacteria (blue-green algae) are the main culprit. Microcystin, a potentially lethal toxin that targets serine/threonine protein phosphatases, is released from some algae blooms. Figure 32.6.5 shows a microalgae bloom in Lake Erie, a shallow freshwater lake polluted with agricultural runoff, making it an excellent site for cyanobacterial blooms.

Figure 32.6.5: Microalgae bloom in Lake Erie, October 2011. https://commons.wikimedia.org/wiki/F..._Lake_Erie.jpg

Red tides are another type of toxic algae blooms that occurs in coastal water. Along US coast they are caused by dinoflagellates and one diatom, both phytoplankton and types of microalgae. Around the Gulf of Mexico, the main cause of red tides is the microalgae *Karenia brevis*. They release large amounts of brevotoxin, a polycyclic ether that binds to and activates voltage-gated Na+ channels in nerve and muscle. Hence it is a potentially deadly neuortoxin. The frequency of red tides is increasing around the world. Two images of red tides are shown in Figure 32.6.6.

Figure 32.6.6: Red Tide. License from Shutterstock

Microalgae are sources of glycan polymers that can be used for bioethanol production; some are excellent candidates for biodiesel fuel, which is made from triacylglycerol reserves. Table 32.6.2 below shows a classification of microalgae and Cyanobacteria along with their characteristic pigments (that impart their distinctive colors) and their energy reserves that could be used for biofuel production.

Phylum	Class	Pigments	Reserve	Habitat
Cyanobacteria	Cyanophyceae	Chl a, β-carotene, flavacene, Echinenone isozea-, zea-, myxo-, oscillaxanthin APC, C-PC, C-PE	Starch (granule) and glycogen	Marine Freshwater Terrestrial
Euglenophyta	Euglenophyceae	Some colorless Chl a, b, diadinoxanthin	Paramylon Ergosterol	Marine Freshwater Terrestrial
Heterokontophyta/ Ochrophyta	Xanthophyceae Eustigmatophyceae	Chl a and c, β-carotene, heteroxanthin, diadinoxanthin (++)	Oil Leucosin Ergosterol	Marine Freshwater Terrestrial
Miozoa	Dinophyceae	Chl a, c, β-carotene, diadinoxanthin, dinoxanthin, peridinins	Starch Lipids	Marine Freshwater
Heterokontophyta/ Ochrophyta	Chrysophycea	Chl a, c, β-carotene, Fuco-, Diato-, diadinoxanthin	Chrysolaminarin Fucosterol Porifasterol	Marine Freshwater
Haptophyta	Coccolithophyceae Pavlovophyceae Rappephyceae	Chl a, c, β-carotdnd, Fuco-, Diato-, diadinoxanthin	Chrysolaminarin Fucosterol Porifasterol	Marine Freshwater
Bacillariophyta (Diatoms)	Bacillariophyceae	Chl a, c, β-carotdnd, Fuco-, Diato-, diadinoxanthin	Chrysolaminarin Oil	Marine Freshwater Terrestrial
Cryptophyta	Cryptophyceae	Chl a, c, Biliproteins, α-carotene, Allo-, Croco-, Monado-xanthin	Starch (granule) Oil Carbohydrates	Marine Freshwater

Table 32.6.2: Classification of microalgae and Cyanobacteria. Hachicha, R.; Elleuch, F.; Ben Hlima, H.; Dubessay, P.; de Baynast, H.; Delattre, C.; Pierre, G.; Hachicha, R.; Abdelkafi, S.; Michaud, P.; Fendri, I. Biomolecules from Microalgae and Cyanobacteria: Applications and Market Survey. *Appl. Sci.* **2022**, *12*, 1924. https://doi.org/10.3390/app12041924. Creative Commons Attribution (CC BY) license (**https://creativecommons.org/licenses/by/4.0/**).

Cyanobacteria and red microalgae store glycogen and floridean starch (a hybrid between starch and glycogen) respectively, while green microalgae accumulate amylopectin-like polysaccharides.

32.6.2: MACROALGAE

As mentioned above, macroalgae, known as seaweeds, do not have **roots**, stems, leaves, or flowers. Kelp in underwater forests appear to have roots, stems and leaves, but they are not plants. They have analogous structures such as holdfast, stipes, and blades which serve the functions of roots, stems and leaves found in plants. As do the microalgae, variants of macroalgae include red (**Rhodophyta**), green (**Chlorophyta,** also a type of microalgae), and brown (**Phaeophyta**) algae. It grows up to 30 times as quickly as land-based groups. In addition, it does not have lignin. They grow much faster than terrestrial plants and can form lots of biomass for commercial processing in much less area than land groups. They can be grown at low cost in sea farms without adding nutrients or pesticides. Hence they are ideal for both food and biofuel production. Brown and red algae have many present commercial uses. They have lots of carbohydrates for potential bioethanol production and triacylglycerols for biodiesel. Their carbohydrate composition includes mannitols and cell wall constituents which could also be used for fermentation. In contrast to present petroleum sources, the biodiesel from macroalgae does not contain sulfur.

Red macroalgae

(some material below for red and green macroalgae from https://bio.libretexts.org/Bookshelv...nd_Green_Algae)

The red algae are almost exclusively marine, and some are unicellular, but most are multicellular. They have true chloroplasts with two membranes (no remnant peptidoglycan) containing chlorophyll. Like the cyanobacteria, they use phycobilins as antenna pigments, **phycoerythrin** (which makes them red), and **phycocyanin**. Red pigment allows the red algae to photosynthesize at deeper depths than the green or brown algae, harnessing more of the blue light waves that penetrate deeper into the water column. Unlike green algae and plants, red algae store carbohydrates as Floridean starch, which has glucose in α(1,4) linkages and occasional α(1,6) linkages, similar to

amylopectin. **Agar**, the base for culturing bacteria and other microorganisms, is extracted from a red alga. Multicellular forms can be filamentous, leafy, sheet-like, coralloid, or even crust-like. Some examples are shown in Figure 32.6.7.

Figure 32.6.7: These images show multicellular red algae, which can range from filamentous (first image) to "leafy" (second image, left) to sheet-like (second image, right). The red color is due to an abundance of the red pigment phycoerythrin, which gives this group reddish chloroplasts. First image by Melissa Ha CC-BY-NC. Second image by Maria Morrow CC-BY-NC. Right image: https://commons.wikimedia.org/wiki/F...ed_algae_3.jpg

Green macroalgae

These algae exhibit great diversity of form and function. Similar to red algae, green algae can be unicellular or multicellular. Many unicellular species form colonies and some green algae exist as large, multinucleate, single cells. Green algae primarily inhabit freshwater and damp soil and are a common component of plankton. They have chloroplasts and the photosynthetic pigments chlorophyll a and b, carotene, and xanthophylls. Examples include *Chlamydomonas*, *Chlorella*, *Pediastrum*, *Netrium*, *Hydrodictyon*, *Acetabularia*, *Ulva*, *and Spirogyra*. Lichens are a symbiotic combination of fungi and green algae.

The nature of the evolutionary relationships between green algae is still debatable. As of 2019, genetic data supports splitting the green algae into two major lineages: **chlorophytes** and **streptophytes**. The green algae exhibit similar features to the land plants, particularly in chloroplast structure. They have chlorophyll a and b, have lost phycobilins but gained carotenoids, and store carbohydrates as **starch inside plastids**. Green algae are an important source of food for many aquatic animals. Two types of green macroalgae are shown in Figure 32.6.8.

Figure 32.6.8:Two types of green macroalgae

left: Figure 5.3.3.125.3.3.12: *Trentepohlia* is a genus of green algae found in terrestrial environments. It forms fluffy orange colonies on trees and is a photobiont in many lichens. One might not know they were looking at a green algae, due to the orange pigmentation. However, green algae have carotenoids. These terrestrial green algae produce an abundance of carotenoids, perhaps for protection from sun damage. Photo by Scott Loarie, CC0.

Right: fresh water green algae. https://upload.wikimedia.org/wikiped...reen_Algae.jpg

Green macroalgal blooms (called green tides) can also occur (just as blooms from the microalgae cyanobacteria). Green tides in the Yellow Sea (between China and Korea) are the largest known. A particularly large one is shown in Figure 32.6.9:

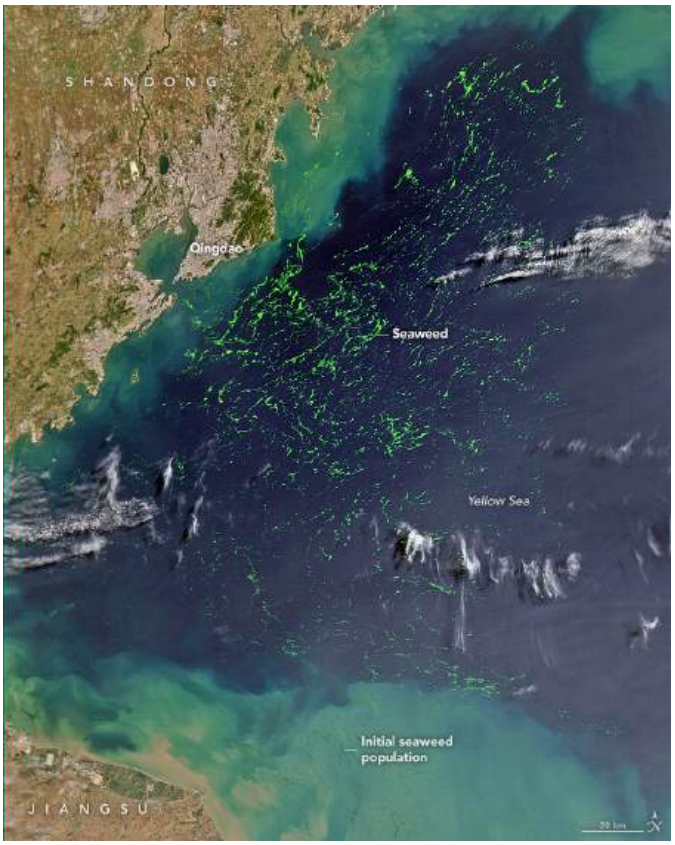


Figure 32.6.9: Green Seaweed in the Yellow Sea , June 2021. https://earthobservatory.nasa.gov/im...the-yellow-sea

This particular green time was from nontoxic green macroalgae, *Ulva prolifera*. It is often called sea lettuce as it is edible. Blooms can affect the local ecosystem and lead to hypoxic zones as they decay.

Brown macroalgae (also known as Kelp)

(some material below for brown macroalgae from https://bio.libretexts.org/Bookshelv...3A_Brown_Algae)

Macroscopic brown algae arose when a heterotrophic eukaryote merged with a unicellular photosynthetic eukaryotic red algae chloroplasts. The red alga degenerated into a chloroplast, this time with four membranes -- the engulfing membrane from the oomycete, the red alga's plasma membrane, and the two membranes of the original chloroplast within the red alga. The chloroplast has lost one of these membranes in many groups derived from secondary endosymbiosis. Figure 32.6.10 shows this secondary endosymbiosis event.

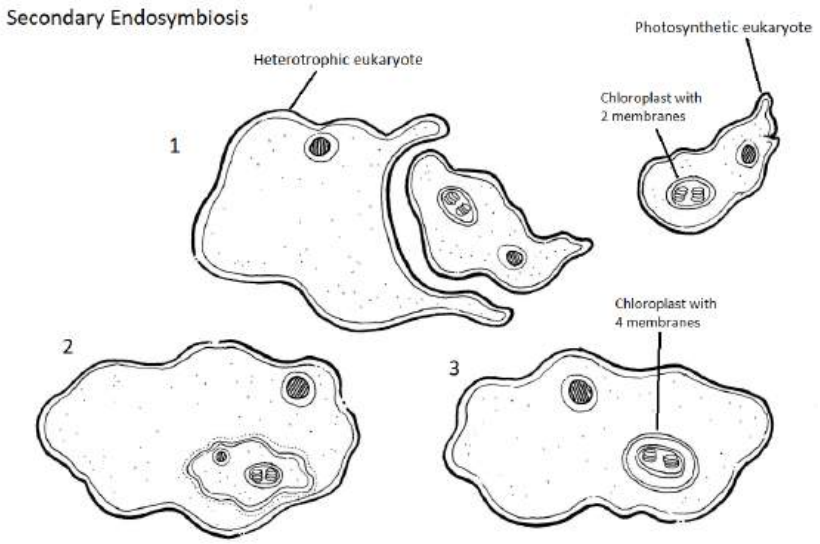


Figure 32.6.10: In the diagram above, we see a unicellular photosynthetic eukaryote with a 2-membrane chloroplast. In step one, this organism is engulfed by a heterotrophic eukaryote. In step two, we see the photosynthetic organism inside the heterotrophic organism. In step three, the original photosynthetic organism has been reduced to a chloroplast with four membranes. Artwork by Nikki Harris CC-BY-NC with added labels by Maria Morrow.

Brown algae are brown due to the large amounts of carotenoids they produce, primarily one called **fucoxanthin**. These organisms are exclusively **multicellular** and can get so large that they require special conductive cells to transport photosynthates from their blades down to the rest of their tissues. These conductive cells are called **trumpet hyphae** and have sieve plates and resemble sieve tubes found in flowering plants.

Much like *Saprolegnia*, the body of an alga is called a thallus because it is not differentiated into specialized tissues. The general morphology of a brown alga includes a **holdfast, stipe, gas bladder**(s), and **blade**(s). Figure 32.6.11 shows a diagram of kelp structure.

32.6.2.1:

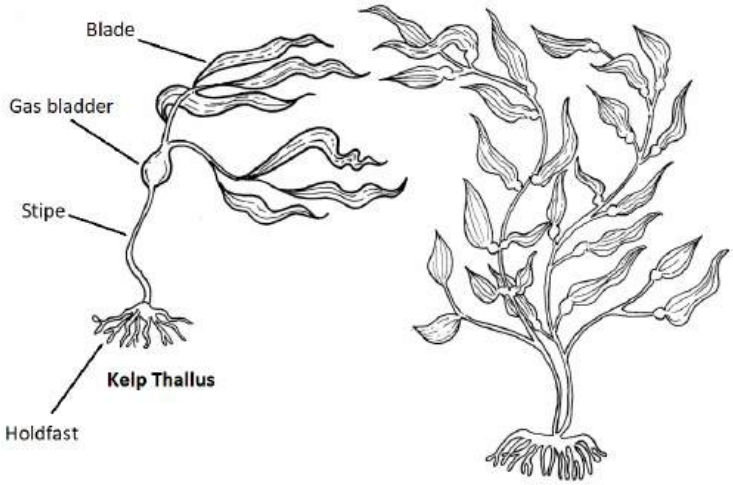


Figure 32.6.11: A diagram of kelp structure. In the diagram above, there are two kelp thalli. The one on the left side is labeled. At the bottom of the thallus is a network of root-like projections that make up the holdfast. The stem-like structure that travels up from the holdfast is the stipe, which terminates in an inflated gas bladder. There are several leaf-like structures attached to the gas bladder. These are blades. The thallus on the right has all these components, but in a slightly different arrangement. Can you find them? Artwork by Nikki Harris CC-BY-NC with added labels by Maria Morrow.

Figure 32.6.12 shows a beautiful image of a kelp forest.

Figure 32.6.12: Kelp forest. https://commons.wikimedia.org/wiki/F...1115735%29.jpg

Figure 32.6.13 shows another type of a brown macroalgae.

Figure 32.6.12: Underwater "roses" of a brown algae Padina pavonica, commonly known as the peacock's tail (Israel). https://commons.wikimedia.org/wiki/F...l_(Israel).jpg

Now we are in a position to discuss algae as a source of biofuels and nutritive foods in the next chapter section.

KEY POINTS - BETA VERSION FROM CHAT.OPENAI

1. Algae are a diverse group of aquatic organisms that range from simple, single-celled organisms to complex, multicellular forms.
2. Algae are photosynthetic organisms and they play a critical role in the Earth's carbon cycle by converting carbon dioxide into oxygen.
3. Algae are found in a wide variety of environments, including freshwater, marine, and terrestrial ecosystems.
4. Algae have been used for a long time as a source of food, medicine and as a fertilizer.
5. Algae are also of great interest as a source of biofuel due to their high growth rates, ability to grow in non-arable land and their ability to produce large amounts of lipids (oils) that can be converted into biofuels.
6. Algae are a valuable resource for the production of biofuels, as they are able to produce large amounts of lipids and carbohydrates that can be converted into biofuels.
7. Algae-based biofuels have the potential to be more sustainable than biofuels produced from terrestrial crops, as they can be grown in non-arable land and do not compete with food crops for resources.
8. Algae can be grown in different systems, including open ponds, closed photobioreactors, and hybrid systems.
9. Algae can be grown using a variety of inputs, such as sunlight, CO2, and nutrients, and can produce a wide range of biofuels, including bioethanol, biodiesel, and biomethane.
10. Algae-based biofuels are still in the early stages of development, and research is ongoing to improve the efficiency and cost-effectiveness of algae cultivation and biofuel production.
11. Algae also have other potential uses, such as for the production of food and feed, as well as for bioremediation and carbon sequestration.

32.7: ALGAE - BIOETHANOL PRODUCTION

Search Fundamentals of Biochemistry

LEARNING OBJECTIVES

- Understand the role that algae play in the production of bioethanol.
- Explore the process of converting algae into bioethanol.
- Learn about the potential benefits and drawbacks of using algae for bioethanol production.
- Evaluate the sustainability of algae-based bioethanol production.
- Analyze the potential of algae as a renewable source of bioethanol and its potential impact on climate change.

Now we are in a position to discuss algae as a source of biofuels and foods. Algae can be used to produce bioethanol from the fermentation of algal polysaccharides (starch, cellulose, and other unique polysaccharides found in them). In addition, triacylglycerols can be converted to biodiesel (which we will discuss in a different chapter). Since we explored bioethanol production from starch and cellulose in previous sections, we'll focus mainly on the unique carbohydrates found in algae and how they can be converted into monomers for glycolytic fermentation by yeast to ethanol.

Both microalgae and macroalgae contain starch reserves (in the cytoplasm and, in some cases, in the chloroplast) and cellulose (in cell walls). They can be grown quickly in aquatic environments, and the key molecules are harvested and processed for use in yeast fermentation by yeast and also by bacteria such as *Zymomonas mobilis* to produce bioethanol. This enzyme can use glucose, fructose, and sucrose as substrates.

STARCH FROM MICROALGAE

Green microalgae have been used as a source of both starch and cellulose. Some species of microalgae have very high starch/glucose composition by mass. Examples include Chlamydomonas reinhaedtil (60%), Chlorococcum humicola (33%), and Chlorella vulgaris (50%). Pretreatment of the microalgae biomass includes liquefaction using alpha-amylases followed by the addition of amyloglucosidase to produce glucose monomers (saccharification) or hydrolysis of glucosidic bonds using acid (sulfuric acid) or base (sodium hydroxide) pretreatment at elevated temperatures.

As discussed previously, starch consists of **amylose**, an unbranched glucose polymer with α(1,4) glucosidic links, and **amylopectin**, which contains α(1,6) branches. Algae starch (for example, in Chlorophyta, Cryptophyta, and Dinophyta) is found in the cytoplasm or chloroplast. A particular type of starch, Floridean (the primary energy storage molecule in the red algae Rhodophyceae), is also found in the cytoplasm. A generic structure of branched starches is shown below in Figure 32.7.1.

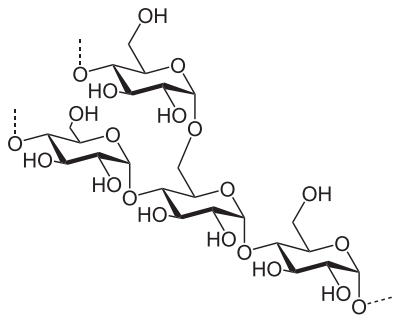


Figure 32.7.1: Branched starches (https://en.wikipedia.org/wiki/Floridean_starch.

Amylopectin has α(1,6) branches every 25-30 glucose units, while animal glycogen, another α(1,4) glucose polymer, has shorter branches every 8-12 glucose units. Floridean starch has branch lengths between these repeat values but is closer to amylopectin. β-amylases can cleave Floridean starch to mainly form glucose and maltose, while mild acid hydrolysis can lead to isomaltose formation.

Other unique glycans are potential sources of glucose for bioethanol production. These are described below.

STARCH-LIKE MOLECULES IN MICROALGAE

Given algae diversity, it should be no surprise that some have alternative, starch-like glycans for their energy reserves that could also be used for bioethanol production.

Laminarin (or Laminaran):

Laminaran is a linear polymer of glucose with β(1,3) glycosidic links with β(1,6)-branches at a ratio of 3:1. It is used for energy storage in brown algae. It can be hydrolyzed by laminarinase, which cleaves β(1,3) glucosidic bonds. It is a linear polysaccharide with a

β(1→3):β(1→6) ratio of 3:1. Its structure is shown in Figure 32.7.2 below.

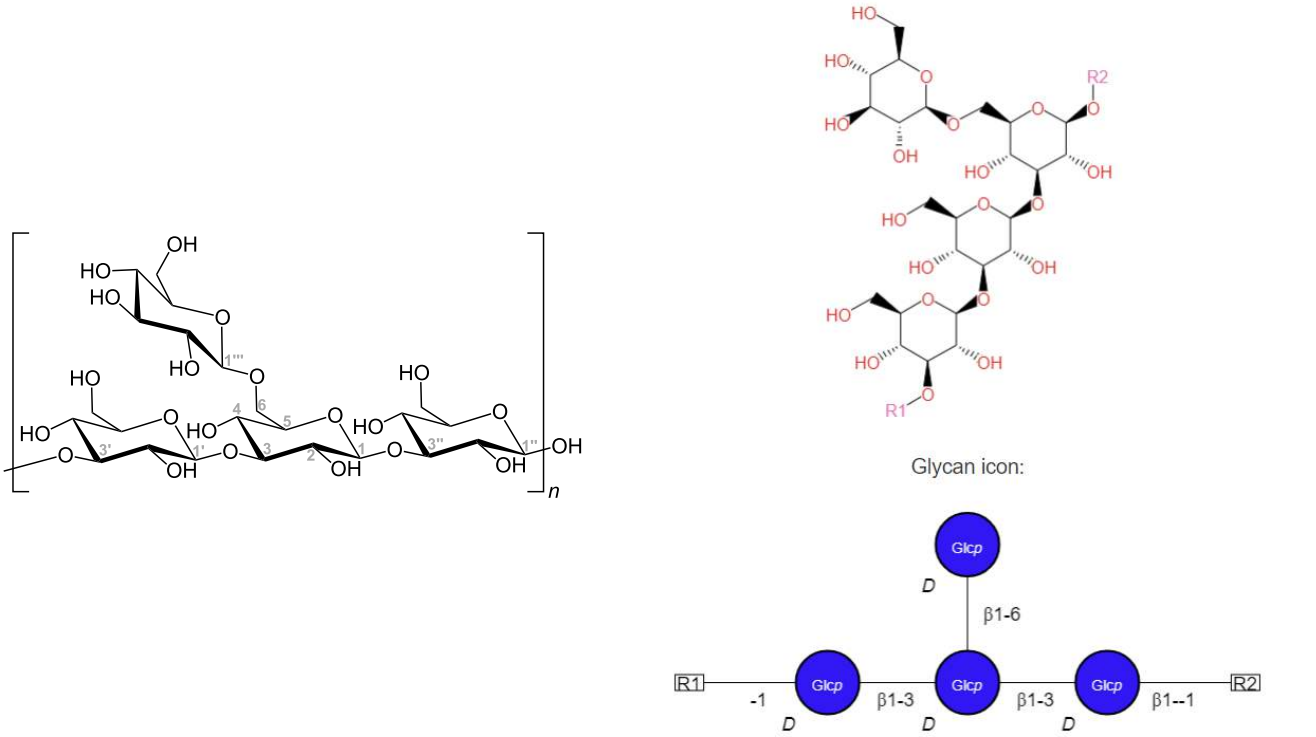


Figure 32.7.2: Structure of laminarin. Left - https://commons.wikimedia.org/wiki/F...ructure_V1.svg; Right - https://biocyc.org/compound?orgid=ME...3602#tab=STRUC

Chrysolaminarin

This glycan is a linear polymer of glucose monomers linked through β(1,3) glycosidic bonds with some β(1,6) linkages. It is found in Haptophyceae, Bacillariophyceae, and Chrysophyceae, which include diatoms. Its generic structure is shown in Figure 32.7.3 below.

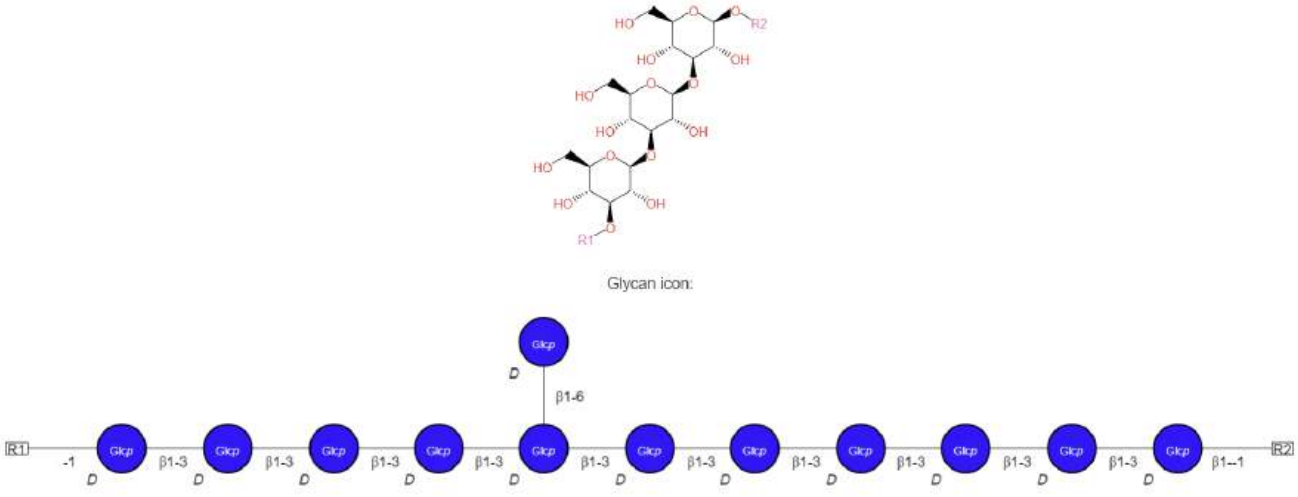


Figure 32.7.3: Structure of Chrysolaminarin. https://metacyc.org/compound?orgid=M...arin#tab=STRUC

The ratio of branching is 11/1, as indicated in the figure. Algae contain enzymes that can cleave the β(1,3) and β(1,4) links, which can be cleaved by acid hydrolysis.

Paramylon

Paramylon is a linear polymer of glucose monomers linked through β(1,3) glycosidic bonds. It is found in Euglenophyceae, Xanthophyceae, and Prymnesiophyta. A β-1,3 glucanase from *Euglena gracilis* (Euglenozoa) cleaves the β(1,3) link. Its structure is shown in Figure 32.7.4 below.

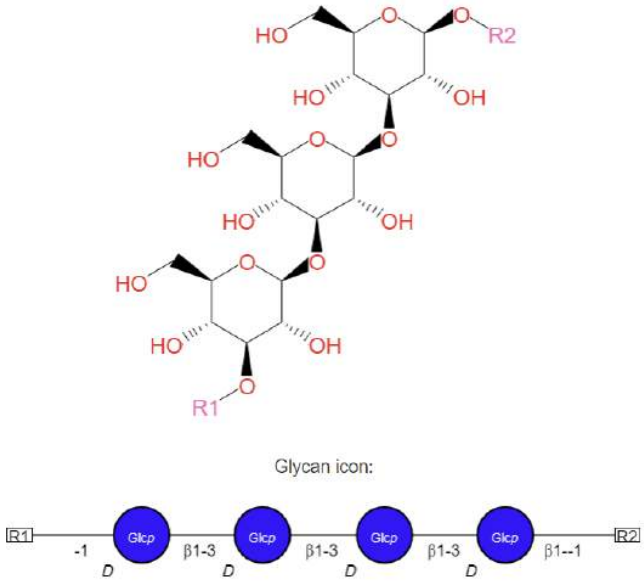


Figure 32.7.4: Paramylon. https://biocyc.org/compound?orgid=ME...ylon#tab=STRUC

ALGAE CELL WALLS

In a previous chapter, we discussed the use of lignocellulosic feedstocks from plant cell walls to produce bioethanol. Both microalgae and macroalgae (red, green, and brown) have cellulose, a β(1,4) glucose polymer, in their cell walls. It is found in Chlorophyta, Dinophyta, Phaeophyta, Prymnesiophyceae, Rhodophyceae, and Xanthophyceae. There is little or much less lignin in algae and apparently none in macroalgae. In addition, there is less hemicellulose. These characteristics make algae excellent candidates for feedstocks for third-generation bioethanol production. There are problems to overcome as well. For example, the cell wall of G*laucocystis nostochinearum* is almost 90% crystalline microfibrils, which adds to its physical and chemical stability. In addition, there are other glycans in some algae's extracellular matrix/cell wall, which can make it difficult to extract and use cellulose. Extracted cellulose can also be used as a feedstock for the chemical industry as well for the synthesis of bioplastics, a topic we will consider in another chapter sections.

Algae cellulose synthesis is performed by membrane-bound cellulose synthase terminal complexes (TCs), with the geometry of the cellulose microfibrils determined by the geometry of the TCs, as shown in Figure 32.7.5. The TCs are arranged in a hexagon with C6 symmetry, and they can form hexameric macrofibrils.

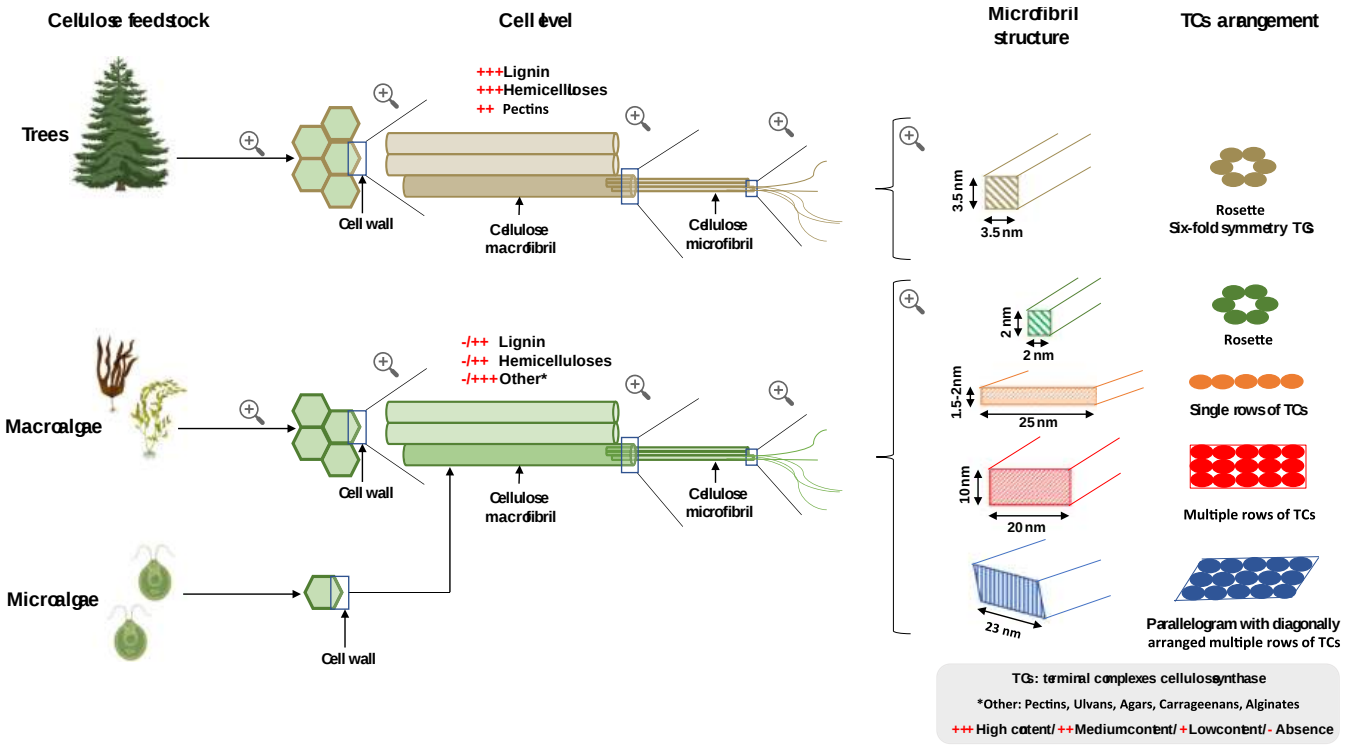


Figure 32.7.5: Organization and morphology of cellulose synthesizing terminal complexes (TCs) in different organisms. Wahlström, N. *et al. Cellulose* **27**, 3707–3725 (2020). https://doi.org/10.1007/s10570-020-03029-5. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Table 32.7.1 below shows the different algae taxa's major polymers in the cell wall and extracellular matrix.

Algal taxa	Crystalline polysaccharides	Hemicelluloses	Matrix polysaccharides
Chlorophyta (green algae)	Cellulose	Xyloglucans, xylans, mannans, glucuronan, (1 → 3)-β-glucan, (1 → 3), (1 → 4)-β-glucan	Ulvans, pectins
Charophyceae (green algae)	Cellulose	Xyloglucans, xylans, mannans, (1 → 3)-β-glucan, (1 → 3),(1 → 4)-β-glucan	Pectins
Phaeophyceae (brown algae)	Cellulose	Sulfated xylofucoglucan, sulfated xylofucoglucuronan, (1 → 3)-β-glucan	Alginates, fucoidans
Rhodophyta (red algae)	Cellulose, (1 → 4)-β-mannan, (1 → 4)-β-xylan, (1 → 3)-β-xylan	Xylans, mannans, glucomannans, sulfated (1 → 3),(1 → 4)-β-glucan, (1 → 3),(1 → 4)-β-xylan	Agars, carrageenans, porphyran
Dinophyta	Cellulose	–	–

Table 32.7.1: Major polymers found in the cell wall and extracellular matrix of different algae taxa. Enio Zanchetta et al., Algal Research, 56 (2021). https://doi.org/10.1016/j.algal.2021.102288.

The percent composition of cellulose, hemicellulose and lignin for bioethanol production is important. Table 32.7.2 shows the % composition of these polymers based on **total dry weight**. (Note: The values for Chlorophyta/Ulvophyceae are highlighted in yellow in both Tables 2 and 3 for convenience of comparision.

Empty Cell	Phylum/class	Strain	Cellulose [%]	Hemicellulose [%]	Lignin [%]
Microalgae	–	A mix of microalgae & cyanobacteria from the wastewater treatment plant	7.1	16.3	1.5
	Chlorophyta/Trebouxiophyceae	*Chlorella vulgaris*	10–47.5	n.d.	n.d.
	Ochrophyta/Eustigmatophyceae	*Nannochloropsis gaditana*	25	det.	n.d.
Macroalgae	Chlorophyta/Ulvophyceae	*Cladophora glomerata*	21.6	det.	n.d.
	Chlorophyta/Ulvophyceae	*Ulva lactuca*	12.4 6.0 16.6	n.d. 12.2 32.5	n.d. 9.8 1.5
	Chlorophyta/Ulvophyceae	*Ulva prolifera*	19.4	14.4	9.4
	Chlorophyta/Ulvophyceae	*Ulva pertusa*	6.7	16.8	n.d.
	Chlorophyta/Ulvophyceae	*Ulva* sp.	40.7	7.1	7.9
	Ochrophyta/Phaeophyceae	*Cystosphaera jacquinottii*	4.6	6.1	19
	Ochrophyta/Phaeophyceae	*Fucus vesiculosus*, *Laminaria digitata*	8	n.d.	n.d.
	Rhodophyta/Florideophyceae	*Gelidium elegans*	17.2	29.5	4.5
	Tracheophyta/Monocots	*Posidonia oceanica*	32.5 31.4–40.0 40.0	23.3 21.8–25.7 20.8	28.2 29.3–29.8 29.8
	Tracheophyta/Monocots	*Posidonia australis*	20.2	11.7	14.5

Table 32.7.2: Cellulose, hemicellulose, and lignin content **in total dry weight** basis of algal feedstock. n.d.: not determined, det.: detected (either directly or indirectly). Zanchetta et al, ibid.

Table 32.7.3 below shows the % composition **in the cell wall** (instead of total biomass) for each of the three polymers.

Empty Cell	Phylum/class	Strain	Cellulose [%]	Hemicellulose [%]	Lignin [%]
Microalgae	Chlorophyta/Trebouxiophyceae	*Chlorella pyrenoidosa*	15.4	31	n.d.
	Ochrophyta/Eustigmatophyceae	*Nannochloropsis gaditana*	75	det.	n.d.
	Charophyta/Zygnematophyceae	*Staurastrum* sp.	72	4.0	1.2–5.6
Macroalgae	Chlorophyta/Ulvophyceae	*Valonia ventricosa*	75	det.	abs.
	Chlorophyta/Ulvophyceae	*Cladophora rupestris*	28.5	abs.	n.d.
	Chlorophyta/Ulvophyceae	*Ulva lactuca*	19	det.	n.d.
	Chlorophyta/Ulvophyceae	*Chaetomorpha melagonium*	41	det.	n.d.
	Chlorophyta/Ulvophyceae	*Enteromorpha* sp.	21	det.	n.d.
	Ochrophyta/Phaeophyceae	*Fucus serratus*	13.5	det.	n.d.
	Ochrophyta/Phaeophyceae	*Laminaria digitata*	20	det.	n.d.
	Ochrophyta/Phaeophyceae	*Laminaria saccharina*	18	det.	n.d.
	Ochrophyta/Phaeophyceae	*Halidrys siliquosa*	14	det.	n.d.
	Ochrophyta/Phaeophyceae	*Himanthalia lorea*	8	det.	n.d.
	Rodophyta/Florideophyceae	*Ptilota plumosa*	24	det.	n.d.
	Rodophyta/Florideophyceae	*Rhodymenia palmata*	7	det.	n.d.

Table 32.7.3: Cellulose, hemicellulose, and lignin content **in the cell wall** of algal feedstock. n.d.: not determined, det.: detected (either directly or indirectly), abs.: absent. Zanchetta et al, ibid.

Note the complete absence of lignin in the cell walls of macroalgae, but the previous table does show some presence in some macroalgae as a whole.

OTHER CELL WALL-ASSOCIATED GLYCANS FOR BIOETHANOL PRODUCTION.

Macroalgae offer an abundant source of other glycans, three of which we describe below. Each requires unique sets of enzymes to convert them to free glucose for fermentation and bioethanol production.

Alginates

These are among the most abundant biopolymers in the world and are the prime carbohydrate in brown seaweeds, where they can reach up to 40% of the dry mass. It is found in the **cell walls** of macroalgae. They are already used for food or other commercial uses. For example, they are used as thickening agents in food and are also used in the textile industries.

The alginate (high in brown algae) is a linear polymer of 1,4-β-D-mannuronic acid (M) and 1,4 α-L-guluronic acid (G) monomers, with stretches (blocks) of pure G, pure M, and mixed MG. Representative structures are shown in Figure 32.7.6 below.

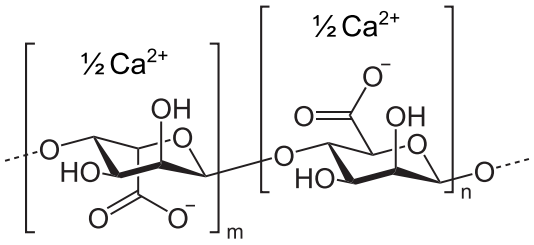


Figure 32.7.6: Left - https://commons.wikimedia.org/wiki/F...e_skeletal.svg; Right - https://biocyc.org/META/NEW-IMAGE?ty...bject=ALGINATE

Agar/Agarose

Agar is abundant in red seaweed. Agar, used in labs, is a mixture of agarose and agaropectin. It acts as a **support for the cell wall** and detaches with boiling. Agarose is a linear polymer of a disaccharide repeating unit of D-galactose and 3,6-anhydro-L-galactopyranose linked by α(1,3) and β(1->4) glycosidic bonds. It has many uses in the laboratory (agarose for chromatography and electrophoresis while agar for cell culture). Its structure is shown in Figure 32.7.7 below.

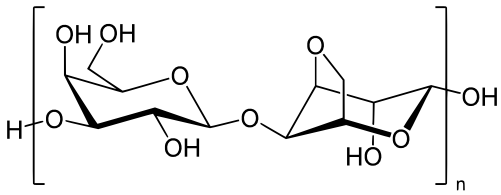


Figure 32.7.7: Agarose. Left - https://commons.wikimedia.org/wiki/F...e_polymere.svg; Right - https://biocyc.org/compound?orgid=ME...oses#tab=STRUC

Carrageenans

These are linear sulfated glycans found **in the cell walls** of red seaweeds (*Rhodophyta*). They have widespread use in the food industry and are used for thickening. They are similar to glycosaminoglycans (GAGs). Their repeating motif is a disaccharide of an α(1,3)-linked D-galactose and a β(1,4)-linked D-galactose. As with GAGs, their degree of sulfation can vary from 15-40%. Different subtypes have a different number of sulfates in the repeating disaccharide unit. For example, the κ form has one, iota (ι) has two, and λ has 3. Representative structures are shown in Figure 32.7.8 below.

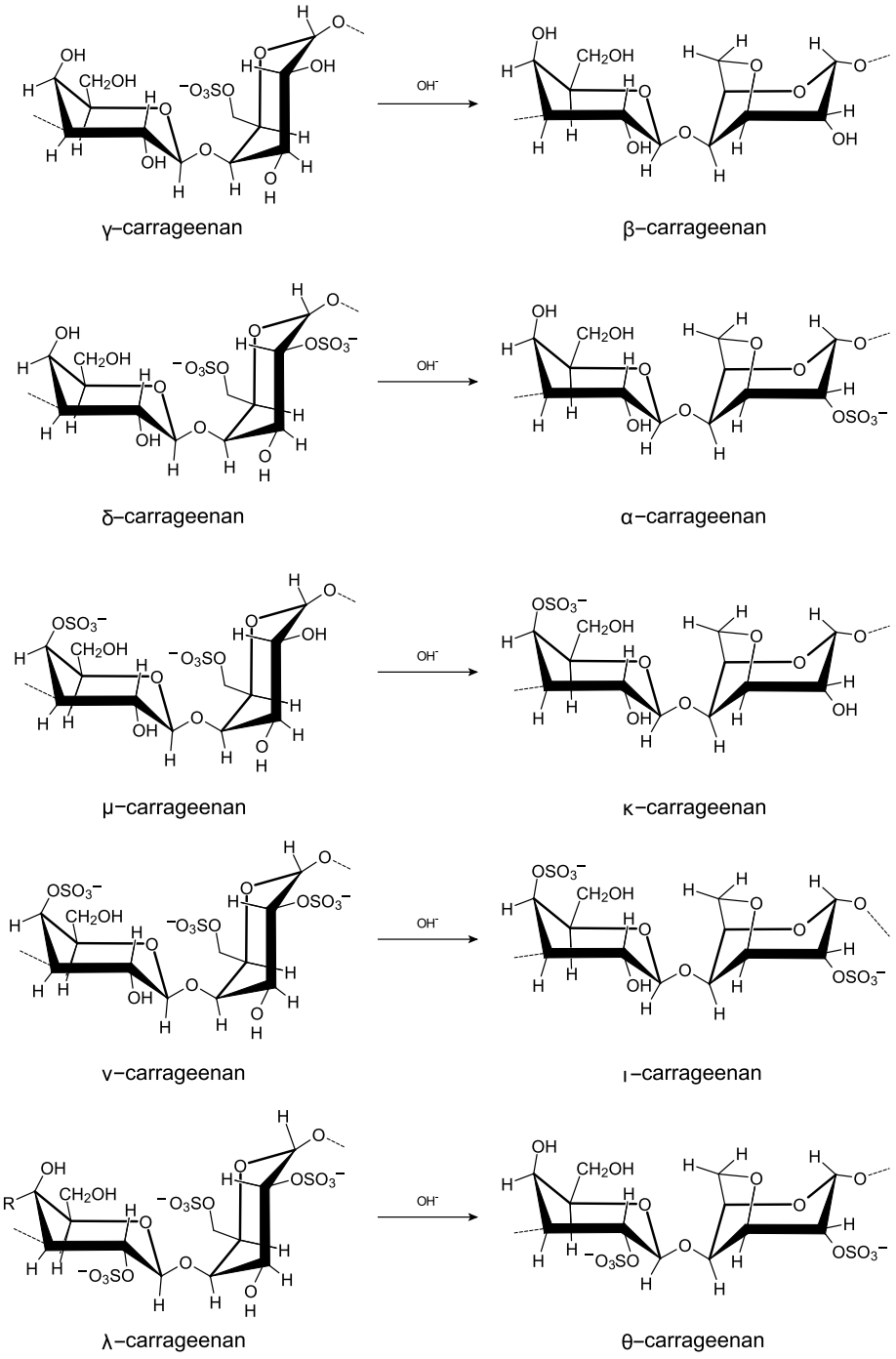


Figure 32.7.8: Carrageenans. https://upload.wikimedia.org/wikiped...enan_types.svg

Variants include the carrageenoses, in which the α(1,3)-linked galactose has a 3,6-anhydro bridge. The carrageenoses are often just called carrageenan. Their structure is shown in Figure 32.7.9.

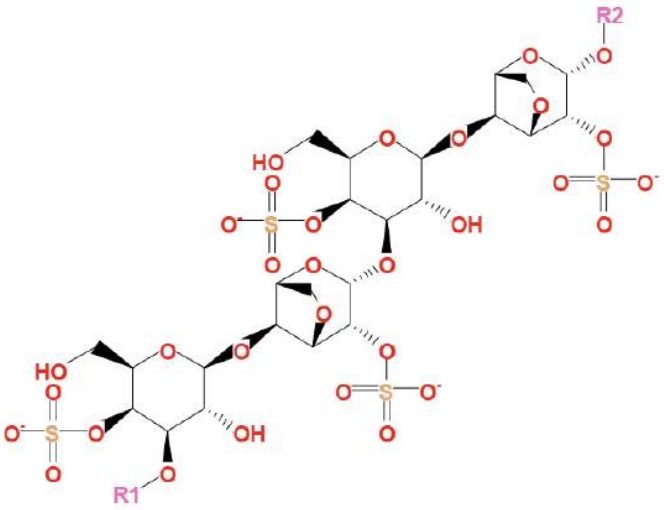


Figure 32.7.9: https://biocyc.org/compound?orgid=ME...enan#tab=STRUC

The 6-anhydro-D-galactose is not fermentable, so microorganisms that can contain enzymes that degrade carrageenan must also be used.

BIOETHANOL PRODUCTION

Now we can put all of this together and make bioethanol from algae. Figure 32.7.10 presents an overview of the entire process.

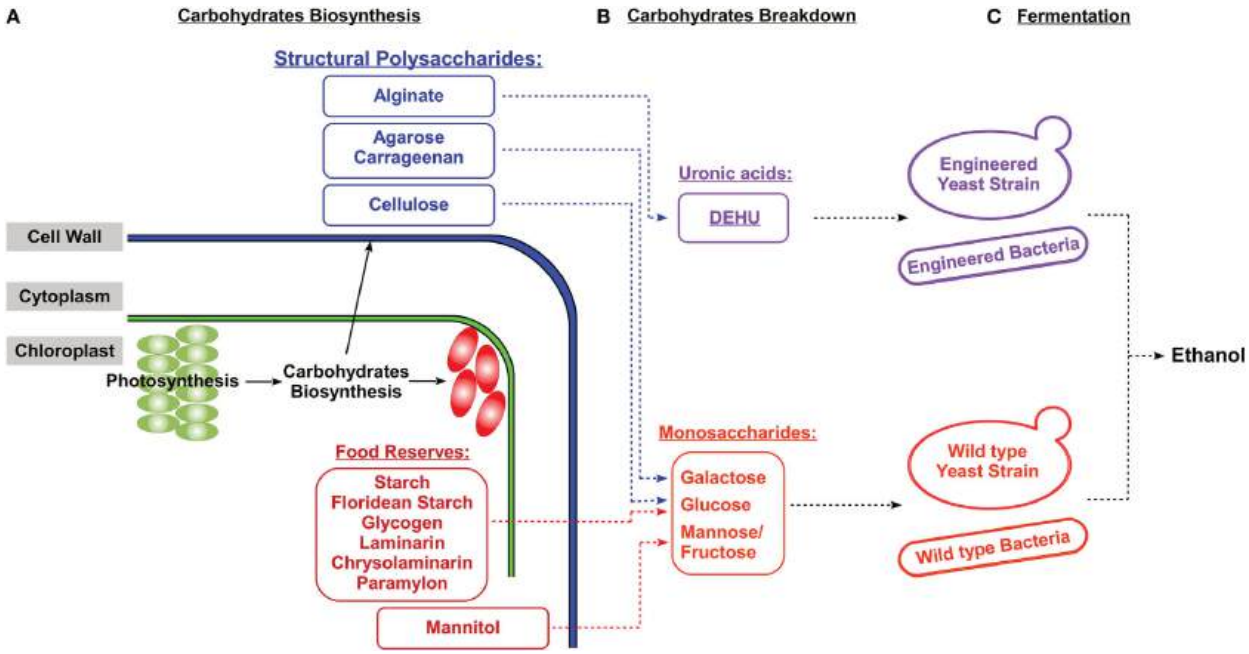


Figure 32.7.10: Overview of ethanol production from major algal carbohydrates. Qusai Al Abdallah et al., Front. Energy Res., 04 November 2016. https://doi.org/10.3389/fenrg.2016.00036. Creative Commons Attribution License (CC BY).

Panel **(A)** shows algae store simple sugars in the form of simple and complex food reserves and as structural polysaccharides.

Panel **(B)** shows how food reserves and structural polysaccharides are degraded into their basic monosaccharides and uronic acids.

Panel **(C)** shows the final fermentation into ethanol using microbial wild-type strains or their genetically engineered counterpartsDEHU, 4-deoxy-l-erythro-5-hexoseulose uronic acid.

Figure 32.7.11 below shows a schematic diagram for converting the algae feedstocks to glucose using key glycan-cleaving enzymes. Note that the colors and shapes are **not** those recommended in the SNFG standard. Those standards show glucose and galactose as blue and yellow circles, respectively.

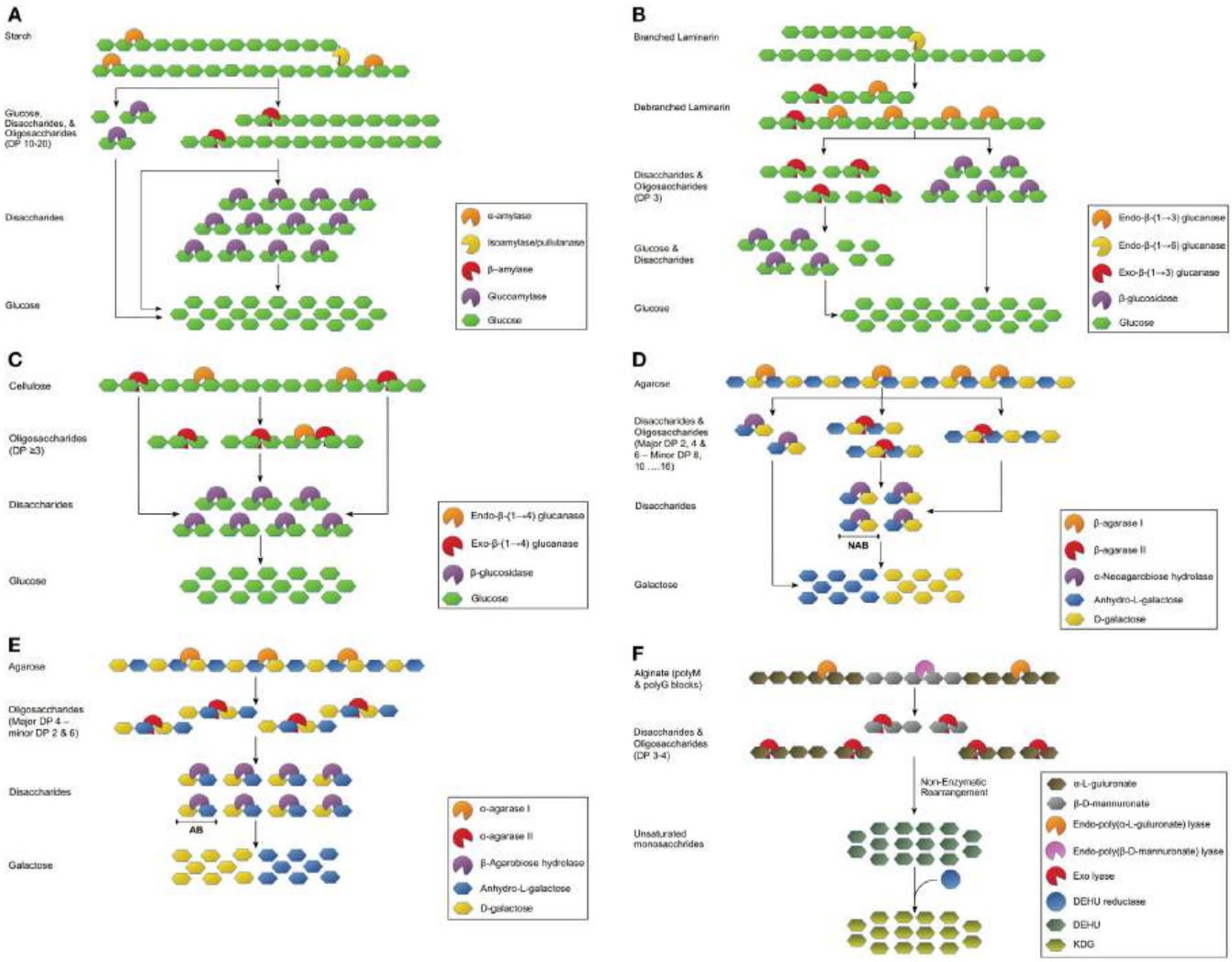


Figure 32.7.11: **SCHEMATIC DIAGRAMS FOR THE ENZYMATIC HYDROLYSIS OF ALGAL POLYSACCHARIDES. (A)** Starch, Floridean starch, and glycogen, **(B)** laminarin, chrysolaminarin, and paramylon, **(C)** cellulose, **(D)** agarose by β-agarases, **(E)** agarose by α-agarases, and **(F)** alginate. DP, degree of polymerization; NAB, neoagarobiose; AB, agarobiose; DEHU, 4-deoxy-l-erythro-5-hexoseulose uronic acid; KDG, 2-keto-3-deoxy-gluconate; M, β-d-mannuronate; G, α-l-guluronaten.

Macroalgae offer enormous advantages for producing food, bioethanol, and intermediates for conversion into bioplastics. Ocean macroalgae don't require irrigation or pesticides, or fertilizers! Figure 32.7.12 below shows rope and raft ocean cultivation of macroalgae.

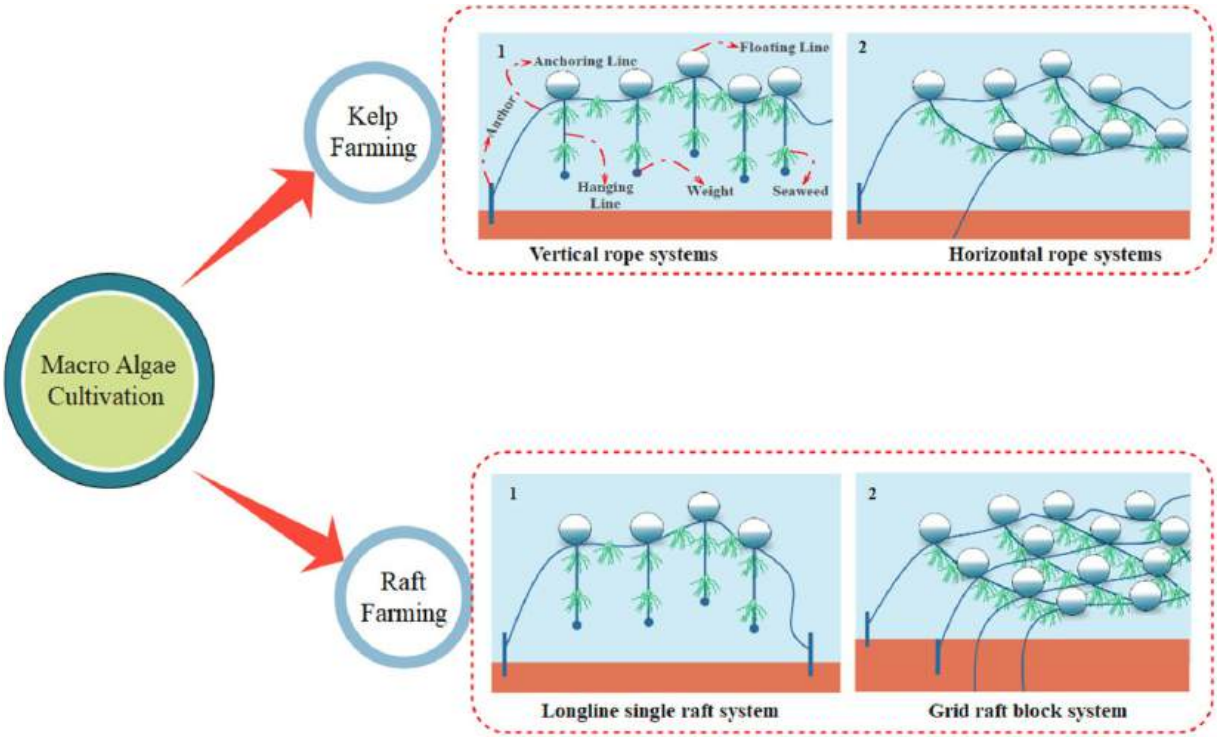


Figure 32.7.12: Macroalgae cultivation systems. Godvin Sharmila V et al. **Bioengineered,** Volume 12, 2021. https://doi.org/10.1080/21655979.2021.1996019. Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/)

Figure 32.7.13 below summarizes how macroalgae (as well as microalgae) can be used to produce a variety of biofuels in addition to bioethanol. The algae cell can be chemically processed to produce hydrogen, methane, and syngas for fuels or combusted (as in the use of wood and coal) to provide heat and electricity. Alternatively, the post-harvest algae can be fractionated to produce carbohydrates for alcohol production, straight vegetable oils (SVO - pure triacylglycerols) for fuel, and fatty acid esters (biodiesel). Of course, they can be used directly as food and for food products.

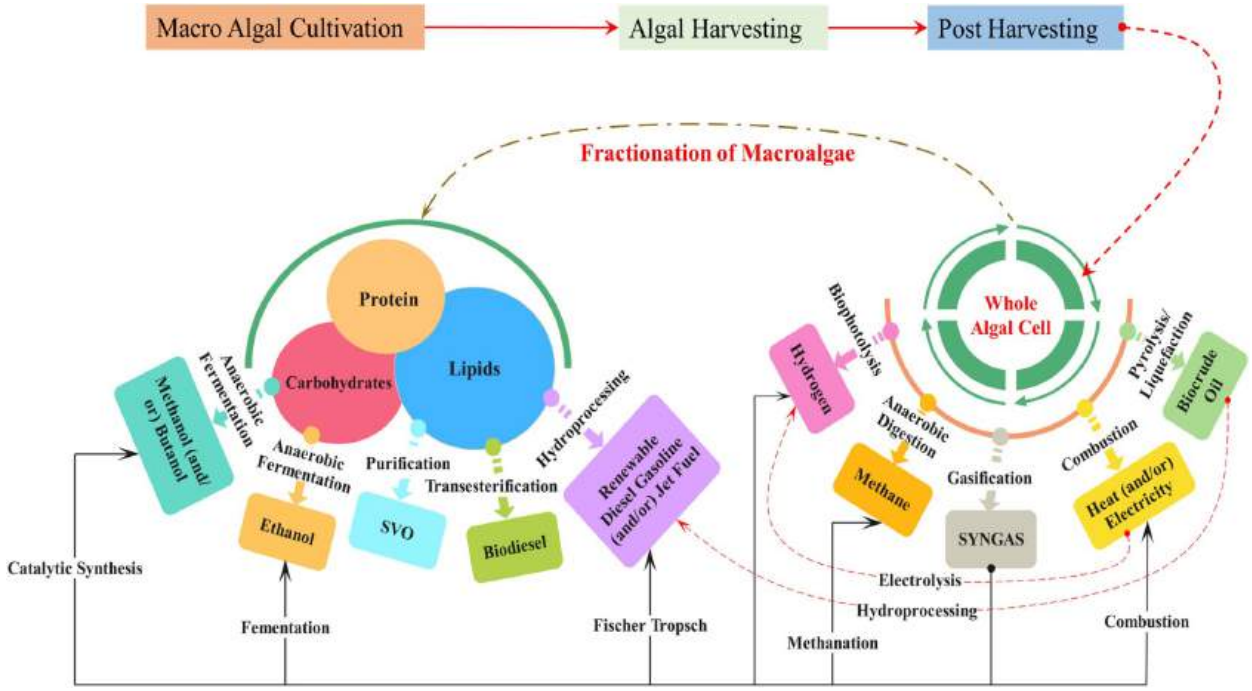


Figure 32.7.13: Macroalgal biofuel refinery. Godvin Sharmila V et al., ibid.

Table 32.7.4 below details the species and methods presently used to produce biofuel from macroalgae.

Type of Biofuel	Species Type	Pretreatment methods or Conversion techniques	Pretreatment or conversion technique conditions	Biofuel Yield or production potential
Biodiesel	*Ulva fasciata*	Catalytic transesterification	Molar ratio of methanol: oil – 9:1; Time – 6 hours; Temperature – 80-100°C	88%
	Chaetomorpha antennina	Transesterification	Chloroform-ethanol solvent- 1:20 (w/v)	2.1 mL/10 $g_{biomass}$
	Gracilaria corticata	Transesterification	Hexane-ether solvent – 1:20 (w/v)	2 mL/10 $g_{biomass}$
	Ulva intestinalis	Transesterification	-	32.3 mg/g dw
	Enteromorpha compressa	Base transesterification	Base – 1% NaOH, Methanol–oil ratio – 9:1, Temperature – 60°CTime – 70 min	90.6%
Bioethanol	*Chaetomorpha linum*	Wet oxidation method	Temperature – 200°C	44 g ethanol/100 g glucan
	Saccharina japonica	Low acid pretreatment	Acid – 0.06% (w/w) sulfuric acidTemperature – 170°CTime – 15 min	6.65 g/L
	Saccharinajaponica	Thermal acid hydrolysis	Acid – 40 mM H_2SO_4Temperature – 121°CTime – 60 min	7.7 g/L
	Laminaria digitata	Oven drying	Temperature – 70°C,Time −72 h	13.6 ± 0.2 µL/g DS
	Ulva linza	Mild acid hydrolysis	Acid condition – 3% H2SO4	12.01%
Biohydrogen	*Laminaria japonica*	Microwave	Temperature – 160°C,Time – 30 min	15.8 mL/g TS
	Laminaria japonica	Ultrasonic	Sonication frequency – 20 kHz	23.56 ± 4.5 mL/g
	Laminaria digitate	Hydrothermal	Temperature – 140°CTime – 20 min	44.0 ± 1.2 mL/g VS
	Chaetomorpha antennina	Surfactant-aided microwave pretreatment	Microwave power – 0.36 kW Ammonium dodecyl sulfate – 0.0035	74.5 mL/g COD
	Ulva reticulate	Microwave-H_2O_2 alkali pretreatment	Microwave power – 0.36 kWH_2O_2 dosage – 24 mg H_2O_2/g biomasspH – 10	87.5 mL H_2/g COD
Biomethane	*Palmaria palmata*	Anaerobic digestion	Semi-continuous anaerobic digestion	320 mL CH_4/g VS
	Chaetomorpha antennina	Ozone disperser pretreatment	Disperser g force – 1,613 g, Treatment time – 30 min, Ozone dosage – 0.00049 g O_3/g TS	0.20 g COD/g COD
	Chaetomorpha antennina	Thermo-chemo disperser	Disperser g-force of 1613 g, Temperature – 80°C, NaOH – 1 N, pH – 11	215 mL/g VS
	Laminaria digitata	Heat	Temperature – 104°C	Dried biomass – 97.66 m^3 CH_4/t fresh biomass – 67.24 m^3 CH_4/t
	Laminaria digitata	Oven drying	Temperature – 70°C,Time −72 h	235.4 ± 14.1 mL/gVS

Type of Biofuel	Species Type	Pretreatment methods or Conversion techniques	Pretreatment or conversion technique conditions	Biofuel Yield or production potential
Bio-oil	*Saccharina japonica*	Fixed bed reactor pyrolysis	Temperature – 450°C	47% conversion
	Ulva lactuca	Microwave pyrolysis	Temperature – 500°C	18.4 wt.%
	Porphyra tenera	Packed tube reactor pyrolysis	Temperature – 500°C	47.4 wt.%
	Laminaria japonica	Packed tube reactor pyrolysis	Temperature – 500°C	45.8 wt.%
	Undaria pinnatifida	Packed tube reactor pyrolysis	Temperature – 500°C	37.5 wt.%

Table 32.7.4: Biofuel production from macroalgae. Godvin Sharmila V et al., ibid.

We will discuss in another chapter the production of biodiesel from algae.

KEY POINTS - BETA VERSION FROM CHAT.OPENAI

1. Algae can be used to produce bioethanol, a biofuel that is similar to traditional ethanol produced from corn or sugar cane.
2. Algae-based bioethanol production involves the cultivation of algae, followed by the extraction of sugars and the fermentation of these sugars to produce ethanol.
3. Algae have a high growth rate and high sugar content, making them a potential source of bioethanol.
4. The process of bioethanol production from algae is being commercialized, and research is ongoing to improve the efficiency and cost-effectiveness of the process.
5. Algae-based bioethanol production has several advantages over traditional bioethanol production methods, such as the ability to grow algae in non-arable land and the ability to produce bioethanol from CO2 and sunlight.
6. The algae-based bioethanol production process, is considered more sustainable and environmentally friendly than traditional bioethanol production methods, as it does not compete with food crops for resources and consumes CO2 during production.

32.8: BIODIESEL, SYNGAS AND BIOAVIATION FUELS

Search Fundamentals of Biochemistry

Diesel fuel was used in 1900 in an engine designed by Rudolf Diesel. The fuel was peanut oil. That might have worked fine in 1900, but a century later, the demands for diesel fuel are met not by peanut oil, a legume, but by oil. Gasoline (petrol) consists of molecules containing 5-12 carbons compared to diesel at 12-20. Both are obtained through fractional distillation of oil (petroleum). Diesel has a higher boiling and melting point and releases more energy per liter (36.9 vs. 33.7 MJ) than gasoline. Regular gasoline contains about 17% n-alkanes, 32% branched alkanes, 5% cycloalkanes, 2% alkenes (olefins), and 30% aromatics. High octane gas can contain around n- and branched alkanes, with the rest from alkenes. Diesel fuel contains about 75% saturated hydrocarbons and 25% aromatics, including alkylbenzenes and napthalenes. In a diesel engine, ignition occurs on compression of the fuel and air mixture and doesn't require a spark. They use glow plugs which provide heat but not spark.

BIODIESEL

Oils (triacylglycerol) produced from biomass can be converted to diesel fuel. In the following sections, we will discuss the synthesis of gases and liquid fuels from nonpetroleum sources such as biomass through the creation of the synthetic gases H_2 and CO (collectively called syngas) and their condensation into liquid fuels using the Fischer-Tropsch reaction. This section will limit our discussion to using fats to create biodiesel. Triacylglycerols for biodiesel production can come from plants, animals, algae, and even waste oils from the food industry. Because biodiesel fuel is composed of carbon-based molecules (often fatty acid esters) with high melting and boiling point ranges which hamper its utility in cold climates, it is often blended with regular diesel (for example, a 20% **b**lend called **B**20). Still, it can be used at 100% (B100). Biodiesel enriched in unsaturated fatty acids has lower melting points and fewer problems in cold weather.

As will the production of bioethanol from lignocellulosic food stocks, biodiesel production has evolved through multiple generations, as shown in Figure 32.8.1 below.

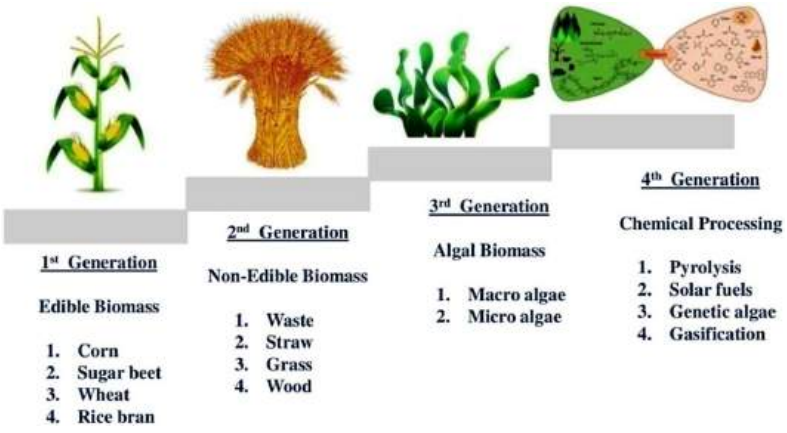


Figure 32.8.1: Palani Vignesh et al., Oil Gas Sci. Technol. – Rev. IFP Energies nouvelles, 76 (2021) 6. DOI: https://doi.org/10.2516/ogst/2020088. Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0)

Table 32.8.1 below shows the advantages and disadvantages of each generation of biodiesel.

Biodiesel generation	Advantages	Disadvantages
1st generation biodiesel	1. Low emission of greenhouse gas. 2. Easy and low-cost technology for conversion.	1. Yield is inadequate to meet the demand. 2. Causes food shortages. 3. High land footprint.
2nd generation biodiesel	1. Using food waste as a feedstock. 2. Use of non-agricultural land to grow a limited amount of crops.	1. Costly pre-treatment. 2. Sophisticated technology is used to transform biomass into fuel.
3rd generation biodiesel	1. Simple to grow algae. 2. No competition for the use of food crops; wastewater, and seawater can be used.	1. More resource usage for algae cultivation. 2. Low lipid level or biomass accumulation in algae.
4th generation biodiesel	1. High biomass and production yield. 2. More capability to eliminate CO_2.	1. The cost of the bio-reactor is higher. 2. At the early stage of research, a high investment is needed.

Table 32.8.1: Advantages and disadvantages of various biodiesel generations. Palani Vignesh et al., ibid

Triacylglycerol feedstocks for biodiesel are often converted to methyl or ethyl esters through an alcoholysis or transesterification reaction, as shown in Figure 32.8.2 below.

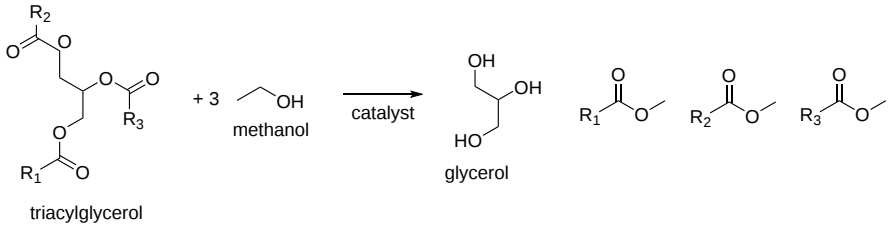


Figure 32.8.2: Methanolysis/transestiferation of triacylglycerol to produce methyl-fatty acid ester for biofuels

This reaction is simply a base-catalyzed cleavage of the ester bonds in the triacylglycerol. Alternatively, vegetable oil can be treated at high pressure and temperature in a hydrogenation reaction to produce a variant called "renewable diesel".

Figure 32.8.3 below shows feedstocks and processing for first-generation biodiesel production.

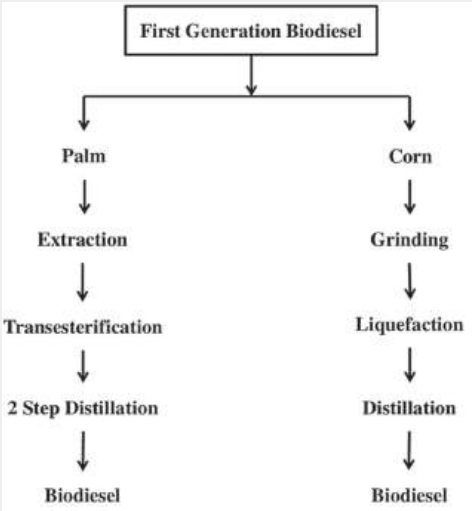
Feedstock: Processing Method	1st Gen Processing
Waste cooking oil: esterification/transesterification Food crops: extraction/transesterification Organic oils: hydrolysis, distillation Animal fat: hydrolysis, fermentation Bioethanol/butanol: chemical synthesis	

Figure 32.8.3: Feedstocks and processing for first-generation biodiesel production. Palani Vignesh et al., ibid

Figure 32.8.4 below shows feedstocks and processing for second-generation biodiesel production

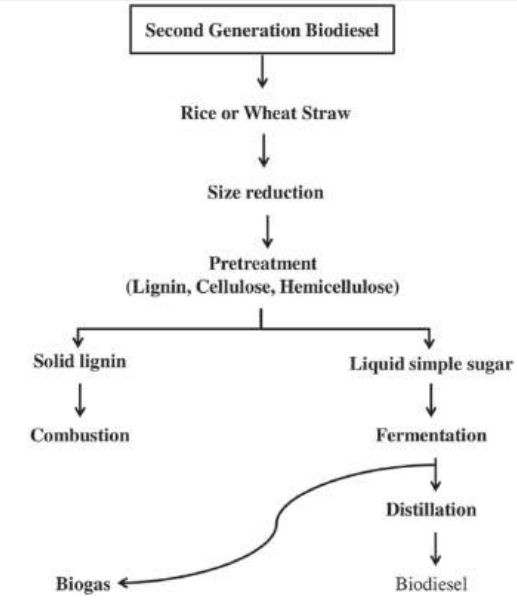
Feedstock: Processing Method	2nd Gen Processing
Cellulose: advanced fermentation Hemicellulose: hydrolysis Lignin: gasification Tannins: biological synthesis Vegetable oil/animal fats: hydrogenation	

Figure 32.8.4 below shows feedstocks and processing for second-generation biodiesel production. Palani Vignesh et al., ibid

Finally, Figure 32.8.5 below shows the processing steps for 3rd and 4th generation biodiesel from algae.

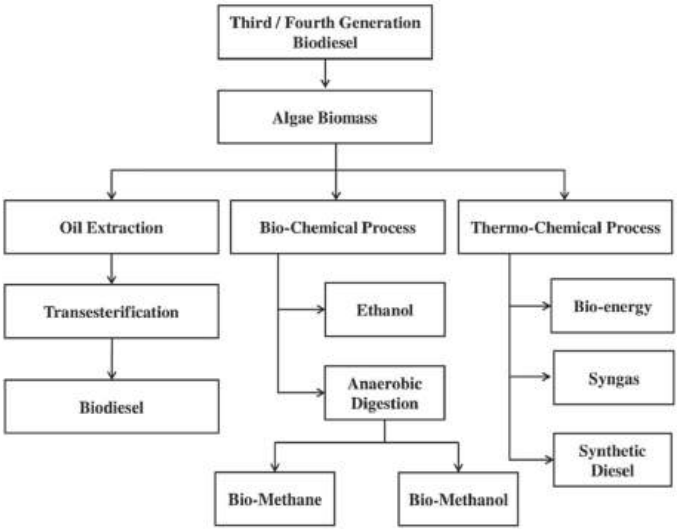


Figure 32.8.5: Processing steps for 3rd and 4th generation biodiesel from algae. Palani Vignesh et al., ibid

Let's consider 3rd and 4th generation biodiesel production using algae. First, algae can be used to produce many different products that can be used for biofuels and chemical feedstocks. These are reviewed in Figure 32.8.6 below.

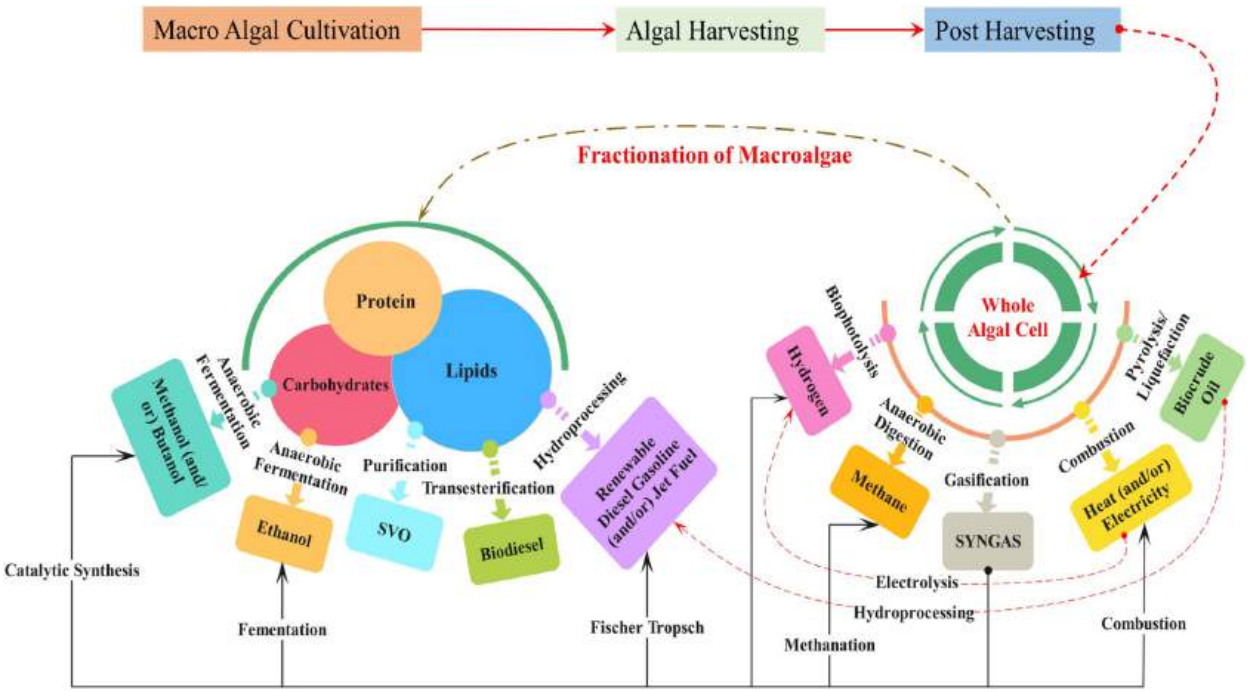


Figure 32.8.6: Macroalgal biofuel refinery; Godvin Sharmila V et al., Bioengineered. 2021 Dec;12(2):9216-9238. doi: 10.1080/21655979.2021.1996019. PMID: 34709971; PMCID: PMC8809944. Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/)

To maximize biodiesel production from algae (3rd and 4th generation), steps in the anabolic pathways for fatty acid and triacylglycerol synthesis could be genetically modified. Below is an overview of triacylglycerol synthesis in microalgae in Figure 32.8.7.

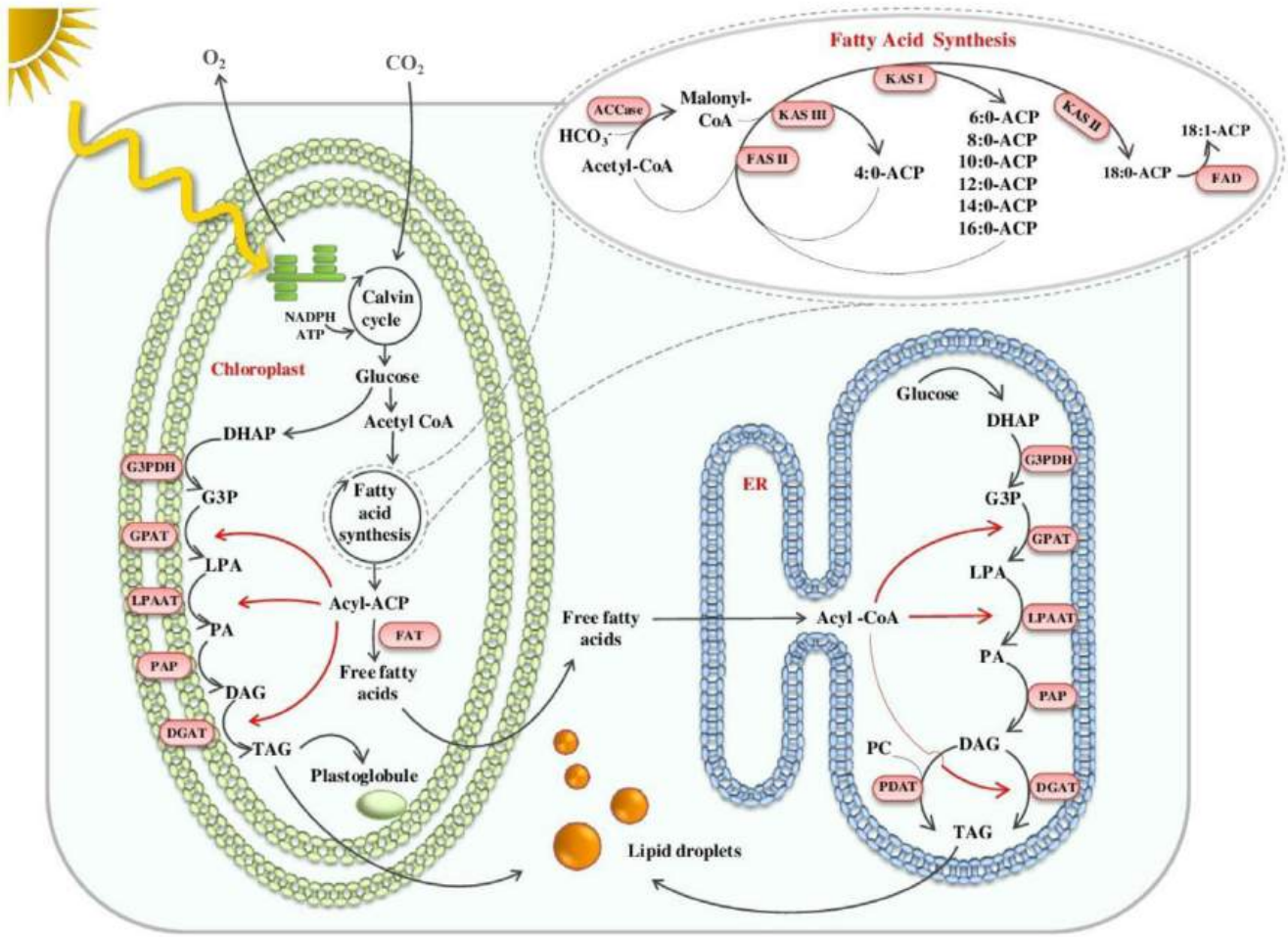


Figure 32.8.7: Schematic illustration of TAG synthesis in microalgae. NADPH; Nicotinamide adenine dinucleotide phosphate; ATP, Adenosine Triphosphate; DHAP, dihydroxyacetone phosphate; G3P/G3pDH, Glyceraldehyde 3-phosphate / G3P dehydrogenase; GPAT, Glycerol 3-phosphate acyltransferase; PA/LPA/LPAAT/PAP, Phosphatidic acid/Lyso-PA/LPA acyltransferase/PA phosphatase; DAG/DGAT, di-Acylglycerol/ DAG acyltransferase; FAT, Fatty acyl-ACP thioesterase; ACP, Acyl-carrier protein; ER, Endoplasmic reticulum; PC, Phosphatidylcholine; PDAT, Phospholipid: DGAT; ACCase, Acetyl-CoA carboxylase; FAS, Fatty acid synthase; KAS, 3-ketoacyl-ACP synthase; FAD, Flavin adenine dinucleotide. Sharma PK, et al. *Front. Mar. Sci.* 5:382, 2018. doi: 10.3389/fmars.2018.00382. .

Each step in the combined pathways are sites for optimization, as shown in Figure 32.8.8 below.

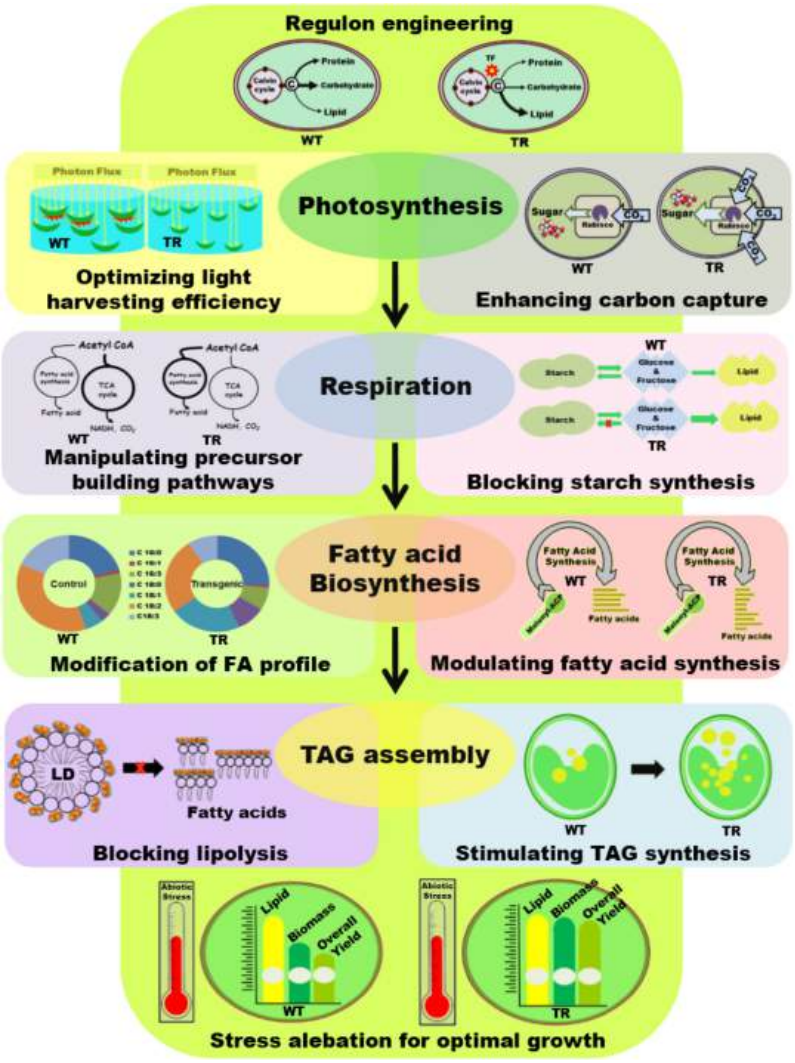


Figure 32.8.8: Schematic illustration of different genetic engineering strategies applied in microalgae for biodiesel application. WT, Wild type cells; TR, Transgenic cells; TF, Transcription factor; TCA, Tricarboxylic acid cycle; NADH, Nicotinamide adenine dinucleotide; FA, Fatty acid; LD, Lipid droplet. Sharma PK, et al., ibid

Life cycle analyses for 3rd generation biodiesel production indicate they would lead to a net decrease in CO_2 emissions, but most appear incomplete in their analyses.

SYNTHETIC GAS (SYNGAS)

As an alternative to using fossil fuels as an energy stock to power our vehicles and as feedstock for chemical production, what if biomass could produce "gasoline-like" fuel for these purposes? We have already discussed the production of bioethanol from 1st (plant starch), 2nd (lignocellulose) and 3rd (algal) generation feedstocks. It is routinely added to gasoline to upward of 15%. It is also found in E85 (or flex fuel), a gasoline blend containing 50% to 80% ethanol.

Instead of producing ethanol through the fermentation of glucose, wood could be incompletely burned to create **"synthetic gases"** (CO and H_2), called syngas, which could be further burned in vehicles to power them or converted through chemical processes (Fisher-Tropsh reaction) to liquid organic fuels.

Indeed, when fossil fuels were lacking, wood was used to create syngas power vehicles. Up to a million cars were powered by wood gas in Europe during World War II. A bus powered by wood gas (syngas) generated by a gasifier on a trailer is shown in Figure 32.8.9 below.

Figure 32.8.9: Bus power by wood gas c. 1943 in Leeds, England. By Ministry of Information Photo Division Photographer, Smith Norman? - http://media.iwm.org.uk/iwm/mediaLib.../large.jpgThis is photograph D 15675 from the collections of the Imperial War Museums., Public Domain, https://commons.wikimedia.org/w/inde...curid=24364067

The syngas emitted was cleaned up somewhat to remove tars and soot/ash particles by passing through charcoal before entering the vehicle through a tube. Tars with polycyclic aromatic hydrocarbons and methane could be lowered if wood or coals were first converted to char before use in a process called **pyrolysis** (heating to high temperatures in the relative absence of air).

The gases, derived from the incomplete combustion of the wood, have CO and H_2 in various proportions depending on the temperature of burning and the source (wood, coal). A general and very simplified reaction for the incomplete combustion reaction is:

$$\text{Carbon feedstock } + \text{ air } \rightarrow CO + H_2 + CH_4 + CO_2 + H_2O + N_2 \tag{32.8.1}$$

Of course, the reaction is not clean, and many organic side products are produced.

The relative ratios of CO and H_2 produced can be changed by the addition of water in a second reaction called the **water gas shift (WGS) reaction** (as water shifts the ratio of CO to H_2):

$$CO + H_2O \leftrightarrow CO_2 + H_2 \quad (\Delta H = -41.2\ \text{kJ/mol.}) \tag{32.8.2}$$

If run in reverse (rWGS) and at high temperatures, the water shift reaction would be a way to capture carbon. The H_2 could come from the electrolysis of water

$$2H_2O(I) \rightarrow 2H_2(\ g) + O_2(\ g) \quad (\Delta H = 286\ \text{kJ/mol}) \tag{32.8.3}$$

The electrocatalytic reduction of CO_2 and H_2O to produce syngas is shown in 32.8.10.

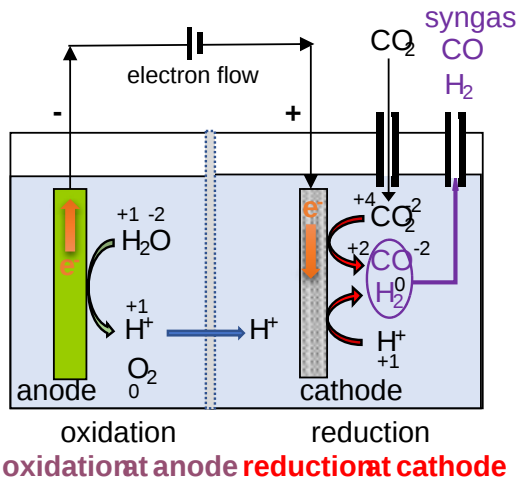


Figure 32.8.10: Syngas Generation by electrocatalytic reduction of CO_2 and H_2O. (after Kang Cheng et al. Advances in Catalysis, 60 (2017). https://doi.org/10.1016/bs.acat.2017.09.003

The oxidation numbers of each element are shown in the diagram. The cathode acts as a catalyst for the reaction. The reaction requires a power source, so this process is greener if electricity derived from green energy sources (wind/solar) is used.

ELECTROCATALYTIC REDUCTION OF CO_2 FOR SMALL MOLECULE FUEL AND CHEMICAL FEEDSTOCKS

Electrocatalytic CO_2 reduction (ECR) offers the potential to capture CO_2 before it is emitted into the air and convert it to small alkanes, alcohols, and acids for fuels (for example, methanol and ethanol) and chemical synthesis (for example, CO and formate). Again this would require a clean source of electricity to power these endergonic reactions. Table 32.8.2 below shows the standard reduction potentials for a variety of half-reactions that could be coupled to form the main ECR products.

Products	Acid		Base	
	Equation	*E*(V)	Equation	*E*(V)
Hydrogen	$2H^+ + 2e^- \rightarrow H_2$	0.000	$2H_2O + 2e^- \rightarrow H_2 + 2OH^-$	−0.828
Carbon monoxide	$CO_2 + 2H^+ + 2e^- \rightarrow CO + H_2O$	−0.104	$CO_2 + H_2O + 2e^- \rightarrow CO + 2OH^-$	−0.932
Methane	$CO_2 + 8H^+ + 8e^- \rightarrow CH_4 + 2H_2O$	0.169	$CO_2 + 6H_2O + 8e^- \rightarrow CH_4 + 8OH^-$	−0.659
Methanol	$CO_2 + 6H^+ + 6e^- \rightarrow CH_3OH + H_2O$	0.016	$CO_2 + 5H_2O + 6e^- \rightarrow CH_3OH + 6OH^-$	−0.812
Formic acid/formate	$CO_2 + 2H^+ + 2e^- \rightarrow HCOOH$	−0.171	$CO_2 + H_2O + 2e^- \rightarrow HCOO^- + OH^-$	−0.639
Ethylene	$2CO_2 + 12H^+ + 12e^- \rightarrow C_2H_4 + 4H_2O$	0.085	$2CO_2 + 8H_2O + 12e^- \rightarrow C_2H_4 + 12OH^-$	−0.743
Ethane	$2CO_2 + 14H^+ + 14e^- \rightarrow C_2H_6 + 4H_2O$	0.144	$2CO_2 + 10H_2O + 14e^- \rightarrow C_2H_6 + 14OH^-$	−0.685
Ethanol	$2CO_2 + 12H^+ + 12e^- \rightarrow CH_3CH_2OH + 3H_2O$	0.084	$2CO_2 + 9H_2O + 12e^- \rightarrow CH_3CH_2OH + 12OH^-$	−0.744
Acetic acid/acetate	$2CO_2 + 8H^+ + 8e^- \rightarrow CH_3COOH + 2H_2O$	0.098	$2CO_2 + 5H_2O + 8e^- \rightarrow CH_3COO^- + 7OH^-$	−0.653
n-Propanol	$3CO_2 + 18H^+ + 18e^- \rightarrow CH_3CH_2CH_2OH + 5H_2O$	0.095	$3CO_2 + 13H_2O + 18e^- \rightarrow CH_3CH_2CH_2OH + 18OH^-$	−0.733

Table 32.8.2: Lei Fan et al. Science Advances. 21 Feb 2020, Vol 6, DOI: 10.1126/sciadv.aay3111. Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).

The reactions are complex and require CO_2 to adsorb onto the electrocatalytic surface of the cathode. Here is a possible reaction pathway for the conversion of CO_2 to methane:

$$CO_2 \rightarrow *COOH \rightarrow *CO \rightarrow *CHO \rightarrow *CH_2O \rightarrow *CH_3O \rightarrow CH_4 + *O \rightarrow CH_4 + *OH \rightarrow CH_4 + H_2O \qquad (32.8.4)$$

The catalysts employed are heterogeneous (i.e., not solution phase) and are typically organometallic transition metal structures. Possible reaction pathways to produce small CO_2 electrochemical reduction products are shown in Figure 32.8.11 below.

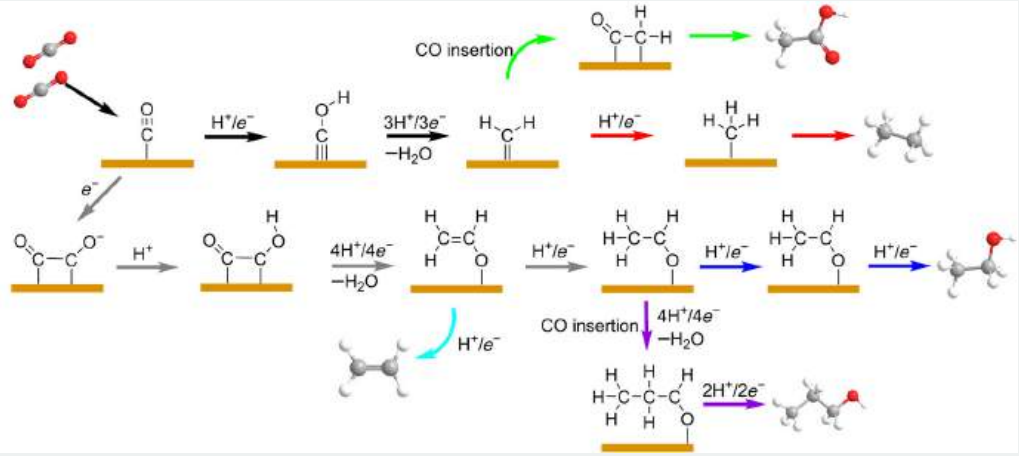


Figure 32.8.11: Possible reaction pathways to produce small CO_2 electrochemical reduction products. Lei Fan et al., ibid

This technology is early in development and will require the development of more robust catalysts and cells before it becomes commercially viable.

The synthesis of syngas (typically described as a mixture of CO, CO_2, and H_2) is widespread now, is used in various processes, and is used to make many products, including hydrocarbons for fuel and oxygen-containing derivatives, including methanol and ethanol. CO and H_2 react in the Fischer-Tropsch reaction (described below) to produce alkanes and alkenes.

$$CO + 2H_2 \rightarrow (CH_2) + H_2O \quad \Delta H = -165\ kJ/mol \tag{32.8.5}$$

where CH_2 is a methylene repeat in longer-chain alkanes.

The CO_2 produced in syngas can be somewhat selectively removed by adsorption onto a CaO catalyst, as shown in Figure 32.8.12 below.

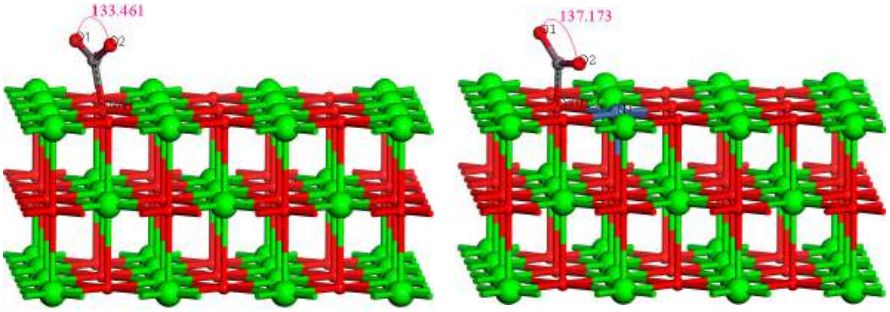


Figure 32.8.12: Adsorption configurations of CO_2 on the surfaces of CaO-based catalysts at 650 °C: (**a**) CO_2 adsorption on CaO (100) surface; (**b**) CO_2 adsorption on 10 wt % Ni/CaO (100) surface. Green, red, gray, and purple balls represent Ca, O, C, and Ni atoms. Zhao, B. et al. *Catalysts* **2019**, *9*, 757. https://doi.org/10.3390/catal9090757. Creative Commons Attribution (CC BY) license (**http://creativecommons.org/licenses/by/4.0/**)

The adsorption energies of CO, CH_4, and H_2 are so small that they have little effect on CO_2 adsorption.

At present, the easiest way to make syngas is to react natural gas (methane) and steam under very high temperatures (up to 1000ºC) over a Ni catalyst (a process called steam reforming). This results in a high H_2/CO ratio of about 3. It can be done with liquified natural gas using Ni-ZrO_2-CeO_2-La_2O_3 catalyst. This process inherently would do little to decrease CO_2 emissions. A **gasification** method converts coal, lignocellulosic biomass, and waste to syngas, with an H_2/CO ratio of <1 for coal and about 0.6-1 for biomass. The electrocatalytic method described above has an H_2/CO ratio of 0-2, depending on the nature of the cathodic catalyst. Syngas can also be made by the partial oxidation of methane. For the subsequent reactions (Fisher-Tropsch), the optimal H_2/CO is about 2. In the gasification of lignocellulosic biomass, the water gas shift (WGS) is used to increase the H_2/CO ratio. This requires lots of water and also produces CO_2 as a product. The water shift reaction use catalysts such as Co–Mo–Al_2O_3, Fe_2O_3–Cr_2O_3, and Cu–ZnO–Al_2O_3) for coal gasification.

Syngas can be used to make small molecule energy and chemical feedstocks, such as ethanol (as well as liquid alkanes and alkenes). Given the many different types of products, it is often essential to selectively make and purify products for commercial use. One method for ethanol production is shown in Figure 32.8.13 below.

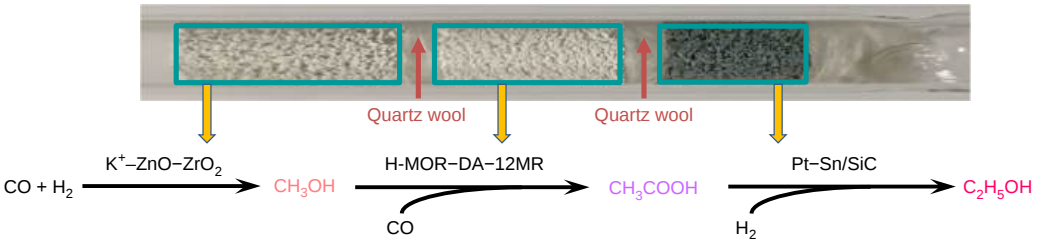


Figure 32.8.13: Conversion of syngas to ethanol proceeds through a tandem mechanism via methanol and acetic acid intermediates using a variety of sequentially positioned catalysts. Kang, J., He, S., Zhou, W. *et al.* Single-pass transformation of syngas into ethanol with high selectivity by triple tandem catalysis. *Nat Commun* **11**, 827 (2020). https://doi.org/10.1038/s41467-020-14672-8. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

H-MOR is a zeolite, which is a microporous, crystalline structure made of aluminosilicate.

In a steam reforming reaction, bioethanol could be converted to CO and H_2 as well, as shown in the equation below.

$$C_2H_5OH + H_2O(+ \text{ heat }) \rightarrow 2CO + 4H_2 \tag{32.8.6}$$

Even with the ability to capture CO_2 from the synthesis of syngas, a central issue of concern is whether the production of syngas and syngas-derived fuels and their use is associated with lower net CO_2 emissions. A life cycle analysis would be necessary to determine that.

FISCHER-TROPSCH SYNTHESIS (FTS) OF FUELS

Fischer was head of the Kaiser-Wilhelm Institute for Coal Research in Germany at the start of World War I. Germany had abundant coal but needed oil for the war, so his efforts were redirected toward that end. Fischer and Tropsch developed the water shift reaction discussed above. They deployed new cobalt catalysts to produce oil which ultimately covered 25% of car fuel and 10% of the German military fuel needs in World War II. Large amounts were also made in South African during the Apartheid regime as well.

The Fischer–Tropsch synthesis (FTS) is a polymerization-like reaction that is used to convert gas-to-liquids (GTL), coal-to-liquids (CTL), or biomass-to-liquids (BTL) fuels. It starts with syngas (H_2 and CO) produced from the gasification of coal/biomass or steam reforming/partial oxidation of natural gas), with the ratios of H_2/CO determined by the water-shift reaction. If coal or biomass is used, a cleanup of residual products with heteroatoms and metal ions is necessary. The clean syngas is then passed into a reactor containing the required catalyst to the FTS of fuels. These reaction systems are summarized in Figure 32.8.14.

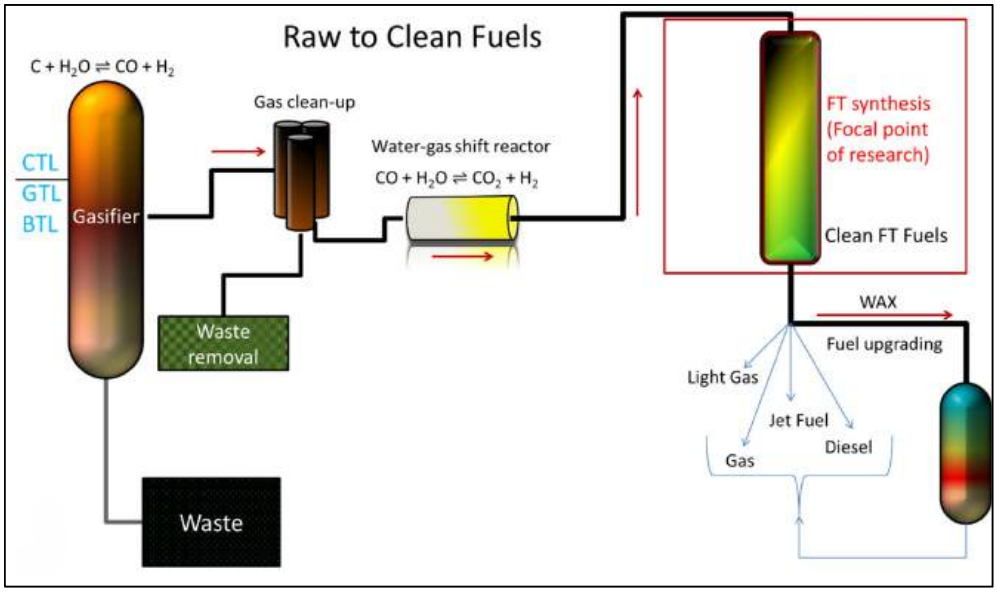


Figure 32.8.14: A simplified diagram of the Coal-to-Liquids (CTL), Gas-to-Liquids (GTL), and Biomass-to-Liquids (BTL) processes. Shafer, W.D.; Gnanamani, M.K.; Graham, U.M.; Yang, J.; Masuku, C.M.; Jacobs, G.; Davis, B.H. Fischer–Tropsch: Product

The FTS reaction is conducted at moderate temperatures and pressure to produce fuels (diesel and jet), lubricants, waxes, and chemical feedstocks. The main representative reactions are given by:

N-alkane production:

$$(2n+1)H_2(g) + nCO(g) \rightarrow C_nH_{2n+2} + nH_2O \quad (32.8.7)$$

Alkene production:

$$(2n)H_2(g) + nCO(g) \rightarrow C_nH_{2n} + nH_2O \quad (32.8.8)$$

The products can have from one carbon (CH_4) to over 70, depending on the catalyst, T, P, and H_2/CO ratio. The FTS is a polymerization reaction, which starts with adsorption of the gas to the catalytic surface, followed by multiple cycles of free radical initiation, propagation, termination, desorption, and reabsorption.

Since this book focuses on structure/function and reaction mechanisms, we would be remiss not to include at least a simplified mechanism for these complex reactions. Several mechanisms have been proposed. Since iron was/is used in the catalyst, and since iron can form iron carbides, Fisher proposed a carbide mechanism. A second enol mechanism has also been proposed. Both are shown in Figure 32.8.15 below.

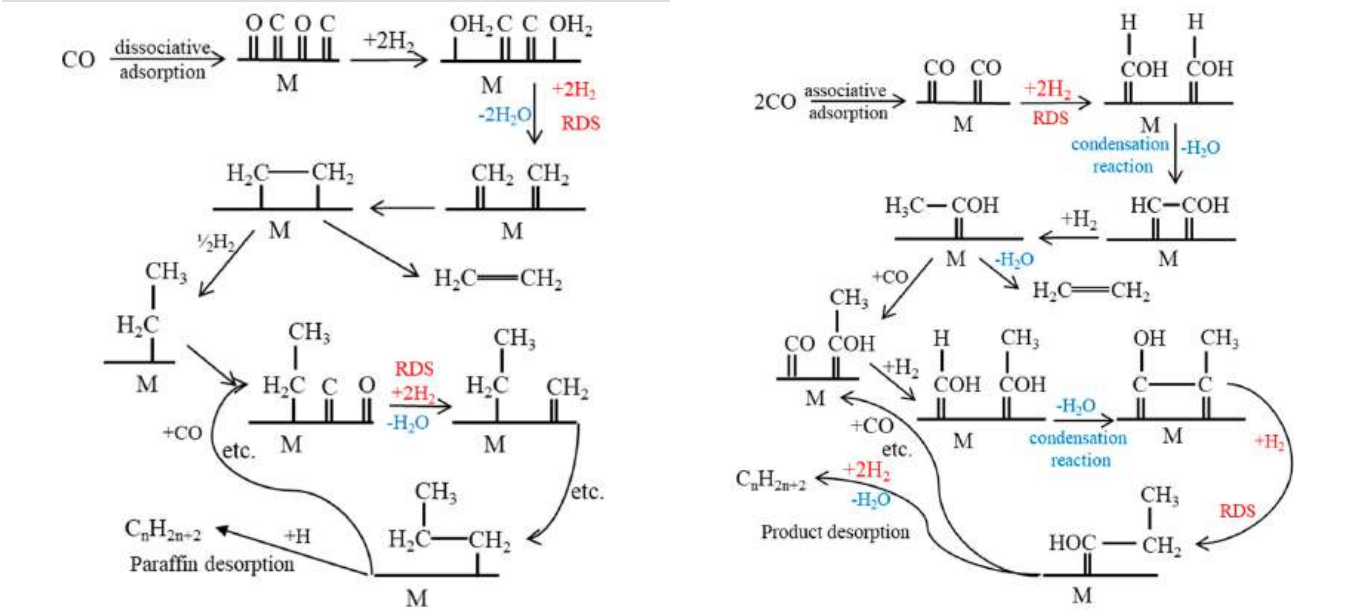


Figure 32.8.15: Proposed mechanisms for the Fischer-Tropsch reaction. Left: A proposed FTS route based on the carbide mechanism; Right: A proposed FTS route based on the enol mechanism. M is the metal surface.

In the carbide mechanism, CO adsorbs on the metal catalyst and dissociates in C and O atoms that cover the surface. These are hydrogenated to form H_2O and CH_2 (methylene), and H_2 adds as the reaction proceeds, as shown.

In the enol mechanism, CO adsorbs without dissociation into atoms. It reacts on the surface to surface-bound H atoms (that arise from the dissociation of adsorbed H_2) to form hydroxymethylene (M-CHOH). This enol grows on condensation with adjacent hydroxymethylenes. The rate-determining step is the hydrogenation of adsorbed CO.

Another proposed mechanism, CO insertion, is shown in Figure 32.8.16

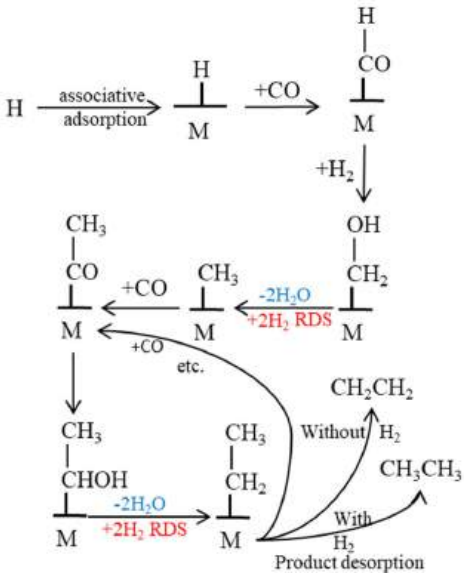


Figure 32.8.16: A proposed FTS route based on the CO insertion mechanism.

CO in this model inserts into a bond from a hydrogen atom to a metal on the catalyst. The rate-limiting step here is the hydrogenation of CO to the CH_2 methylene group. The assumed monomer for this mechanism is simply CO through its insertion into metal-carbon bonds.

Generally, the FTS reaction catalyst has either cobalt or iron ions. The metal catalyst can also be doped with potassium and copper ions and bind silica and alumina. Iron is abundant and cheap and better promotes the water-gas-shift reaction, so it is best for FTS synthesis of fuels from coal and biomass since syngas derived from them have a lower H_2/CO ratio.

CHEMICAL SYNTHESIS OF FOOD

If fuels can be synthesized from nonpetroleum sources and from agricultural waste, why not use the reactions described above (like syn gas production, the Fischer-Tropsch, and electrochemical reduction) to make food, not just fuel? The agricultural sector accounts for about 25% of total greenhouse gas emissions, so it is ripe (no pun intended) for novel ways to produce foodstuffs (think of them as feedstocks for human consumption).

It should be obvious from the chemistry, that the simplest type of food to create from these reactions is fats, and not the more synthetical and stereochemically complicated proteins and carbohydrates. Presently the world relies on palm oil derived from palm trees for fats found in crackers, cookies, breads. and other food products. Typical production of palm oil has been associated with large-scale deforestation of rainforests, with its associated climate and biosphere impacts. Americans consume about 8 kg (17.6 lbs) of palm oil products each year, including products like cosmetics, cleaning products, and waxes. Figure 32.8.17 below shows the scale of palm oil production in the world, which might come as a surprise since we more frequently encounter soybean, sunflower, rapeseed (source of canola oil), and olive oils in our stores.

Figure 32.8.17: Vegetable Oil Product in the World. Our World in Data. https://ourworldindata.org/palm-oil

The general chemical pathways used to chemically synthesize proteins, fat,s and carbohydrates are illustrated in Figure 32.8.18 below. The fat synthesis pathways relevant for this discussion are represented in the central part of the figure.

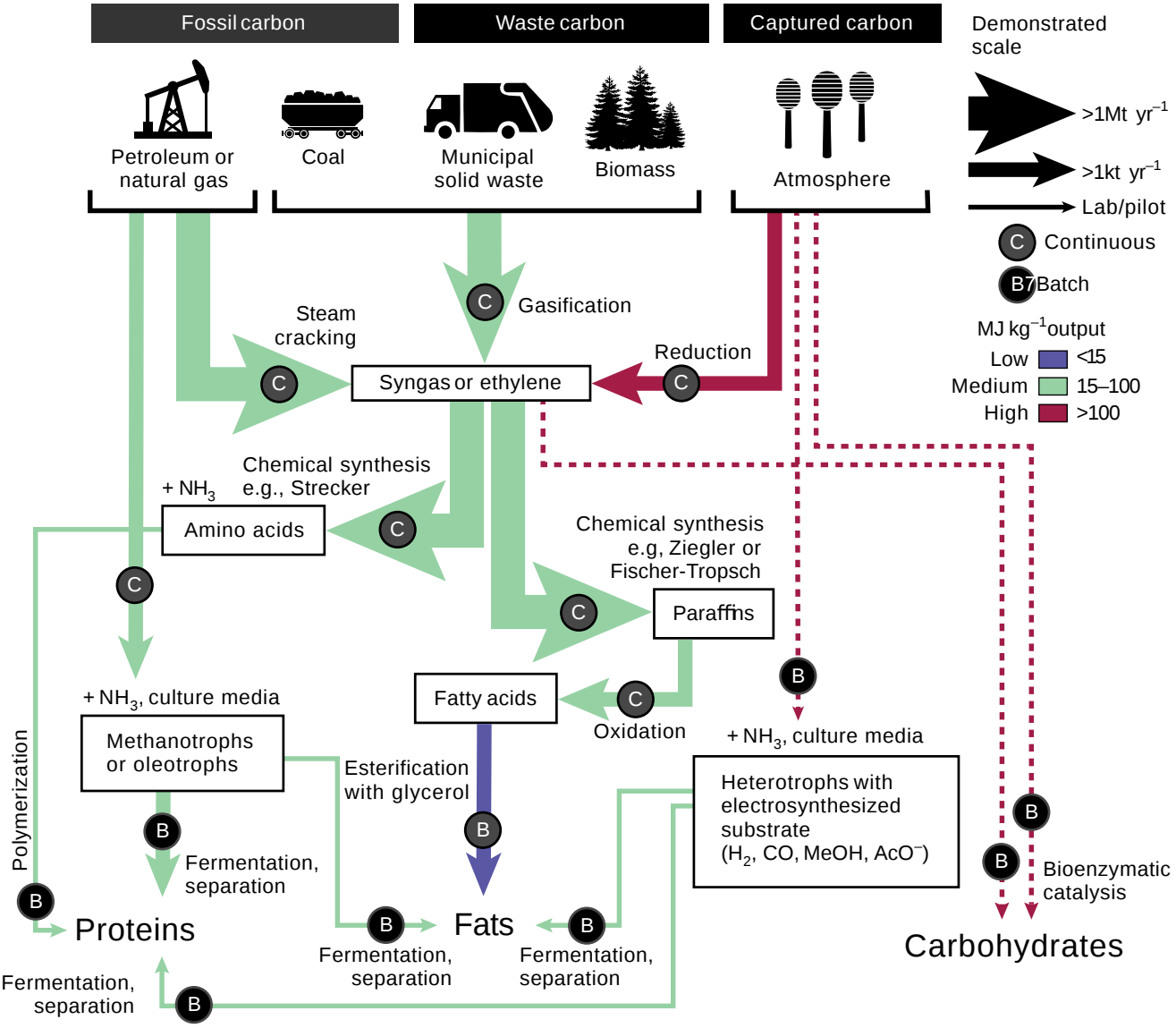


Figure 32.8.18: **Schematic of potential pathways to synthesize food without agriculture.** Davis, S.J., Alexander, K., Moreno-Cruz, J. *et al.* Food without agriculture. *Nat Sustain* (2023). https://doi.org/10.1038/s41893-023-01241-2. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Proteins, fats, and carbohydrates can be synthesized from a range of carbon feedstocks via multiple chemical and biological pathways (arrows). The weight and color of the arrows indicate the scale at which the different processes have been demonstrated and the energy required per mass unit output, respectively. Dashed lines indicate where energy requirements remain highly uncertain. Circular labels on each arrow further indicate whether the process is typically continuous (C) or batched (B). NH_3 is ammonia, H_2 is hydrogen gas, MeOH is methanol, and AcO^- is acetate (agriculturally produced carbon feedstocks are omitted).

It is estimated that the chemical of synthesis of fats could lead to the emission of <0.8 g $CO_{2\ equiv}$/kcal compared to >1.5 g CO_2 equiv/kcal for the present biological synthesis and processing of palm oil in Brazil or Indonesia. This advantage would increase if the CO_2 for the chemical synthesis could be captured from the atmosphere, yielding net 0 emissions.

Figure 32.8.19 shows both land use and energy emissions of $CO_{2\ equiv}$/kcal for agricultural-produced fats (panel a) versus equivalents for chemical-synthesized fats.

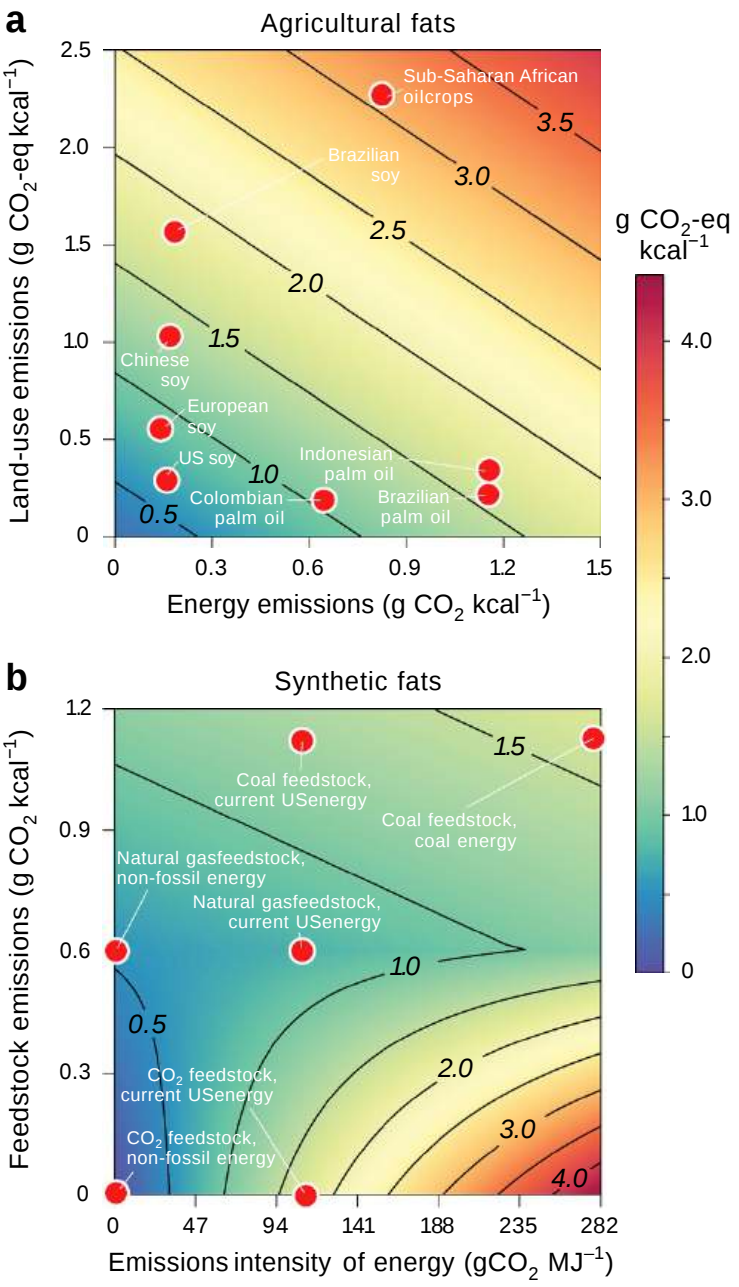


Figure 32.8.19: **Comparison of emissions per calorie of edible fats.** Davis, S.J et al., ibid.

Panels a and **b** show shading and contours show grams of CO_2-equivalent GHG emissions per kilocalorie of edible fat produced by conventional agriculture (Panel **a**) and chemical synthesis (Panel **b**). Agricultural emissions are shown as the sum of land-use emissions (*y*-axis) and energy-related emissions (*x*-axis), and emissions from synthesis are shown as a function of feedstock emissions intensity (*y* axis) and energy emissions intensity. Red circles denote specific estimates based on literature and assumed values. Feedstock emissions include process-related conversion of feedstock to CO_2—for example, during extraction of natural gas, gasification of coal, and the eventual human respiration of fossil feedstock.

BIOAVIATION FUEL

Now we can turn our attention to the production and analysis of bioaviation fuel, which is more like kerosene and diesel fuel than ethanol in its composition. As we saw for bioethanol productions, the feedstocks can be 1st, 2nd and 3rd generation, as shown in Table 32.8.3 below.

First-generation (1-G)	Second-generation (2-G)	Third-generation (3-G)	Fourth-generation (4-G)
• Oil-seed crops: camelina, oil palm, rapeseed, soybean, sunflower, salicornia • Sugar and starchy crops: corn, wheat, sugarcane, sugar beets	• Oil-seed energy crops: jatropha, castor bean • Grass energy crops: switchgrass, miscanthus, Napier grass • Wood energy crops: poplar, willow, eucalyptus • Agricultural and forestry residues: corn stover, sugarcane bagasse, wood harvesting/processing residues • Food and municipal waste: used cooking oil, animal fats, biogenic fraction of municipal solid waste	• Algae: microalgae	• Genetically modified organisms • Non-biological feedstocks: CO2, renewable electricity, water

Table 32.8.3: Feedstocks for bio-aviation fuel production. Doliente SS, Narayan A, Tapia JFD, Samsatli NJ, Zhao Y and Samsatli S (2020) Bio-aviation Fuel: A Comprehensive Review and Analysis of the Supply Chain Components. *Front. Energy Res.* 8:110. doi: 10.3389/fenrg.2020.00110. .

A typical jet aviation fuel (Jet A-1) contains n- and branched-alkanes (often called paraffins) and some alkenes (often called olefins) with 8-16 carbon atoms, cycloalkanes, and aromatics. A comparison of the components of Jet-A1 with a typical bioaviation fuel, Bio-Jet, is shown in Figure 32.8.20 below.

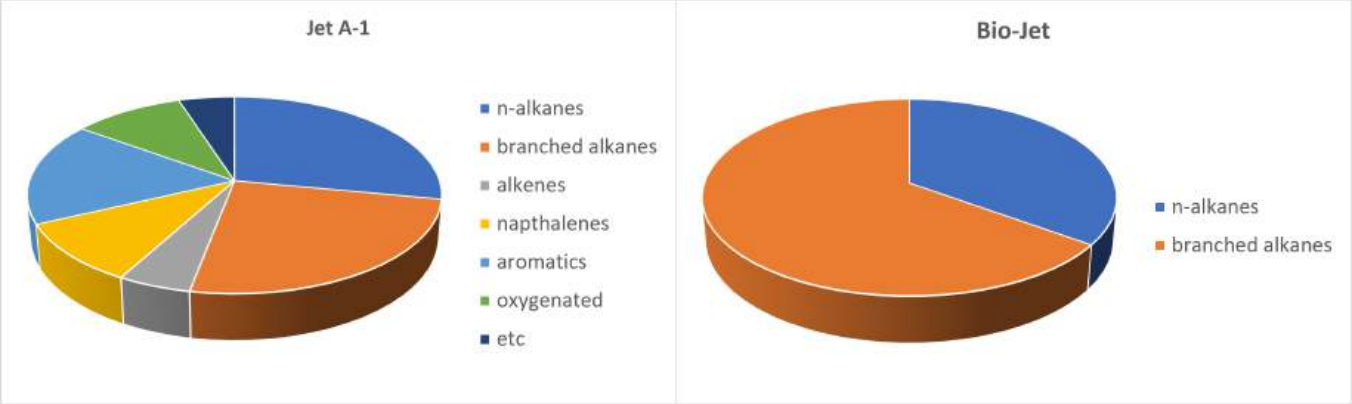


Figure 32.8.20: Molecular-class compositions of (a) Jet A-1 and (b) bio-jet identified by the relative signal area percentage analysis of GC–MS. After Cheon Hyeon Cho, Hee Sun Han, Chae Hoon Sohn, and Jeong Sik Han. *ACS Omega* **2021** *6* (40), 26646-26658. DOI: 10.1021/acsomega.1c04002.

Bioaviation fuel can be made from feedstocks containing triacylglycerol or, most readily using feedstocks shown in Table 2 above to create syngas for use in the Fischer-Tropsch reaction. We've already discussed those reactions above. Instead, let's focus on whether bioaviation fuel, better termed sustainable aviation fuels (SAF), is good for our climate. This presupposes that battery-powered planes and jets are not scalable to our current environmental needs.

LIFE CYCLE ANALYSIS - JET FUEL FROM GRASSES

As expected, the US is the largest user of aviation fuel and causes 25% of aviation CO_2 emissions, as much as all the greenhouse gases emitted through fuel use in Spain. It is estimated that the US will need around 30 billion gallons/per (BGY) of jet fuel in 2040. A biojet fuel industry based on cellulose as a feedstock could theoretically produce that amounts. If the industry costs are estimated at $123 billion, then the cost of the jet fuel would be $4.30/gal. About 60% of the costs would arise from the conversion of biomass to sustainable aviation fuel (SAF), which we described above. Those costs include the building of the biorefineries. Their cost would be distributed over their lifetimes of the plants. The fuel would be derived from syngas from the Fisher-Tropsch reaction, a reasonably mature technology. The costs would be much lower (closer to $1/gal) if the infrastructure costs were not included. These total costs are comparable to the price paid for regular jet fuel (around $2.2/gal in 2021). Consumer costs would not go up 2-fold since fuels are only part of the cost paid by passengers (15-25%)

Consumers' current price for fossil fuel is much less than its actual cost. The present price/gal does not include external costs associated with the use of fossil fuels. These include climate change effects on infrastructure, agriculture, industry, etc., and on human health (mostly from negative health consequences and diseases exacerbated by fossil pollution). We all ultimately pay for these hidden costs resulting from a failed market for pricing fossil fuels. On top of this, the fossil fuel industry has been massively subsidized for decades.

Different prices have been placed on carbon emissions from fossil fuels to resolve this error in the market. The carbon price is based on the number of tons of CO_2 equivalent emitted ($US/ton CO_2e). If a reasonable carbon price is added, biojet fuels made from lignocellulosic stocks through the Fisher-Tropsch reaction would be theoretically competitive with conventional jet fuels made from fossil fuels. The extra added cost for sustainable aviation fuel (SAF) compared to traditional aviation fuel at different prices placed on carbon placed on each are shown in Table 32.8.4 below, assuming an average cost of $2.20/gal for traditional aviation fuel.

Price on carbon ($US/t CO2e)	SAF - Traditional Aviation Fuel ($/gal)
0	+$1.90/gal
$50	+$0.60/gal
$175	$0

If the cost of traditional jet fuel went to $3/gal, as it did in the US in March 2022, SAF and conventional jet fuel would cost the same if a price on carbon of $100/t CO2e were included for both. The actual "social" cost of carbon has been calculated (9/22) to be $175/ton CO_2e.

Other factors other than the cost of carbon should be included in these analyzes. These include the issue of sustainable land use to grow feedstocks for the SAF. A recent analysis shows that it would be possible to produce 30 billion gall/yr of cellulosic SAF by planting 23.2 Mha (Million hectares, about the size of the state of Wyoming) of marginal agricultural lands (about one-third of croplands and 2/3 noncrop lands) in the Midwest with miscanthus, with a net cost of $4.1/gal and assuming a carbon price of a $50/ t CO2e. Miscanthus is a rapidly growing tall perennial grass with a high yield that grows in moderate climates. The life cycle analysis included interactions among atmospheric, land surface, ecosystem, and economic systems. Miscanthus gigantheus is shown below in Figure 32.8.21.

Figure 32.8.21:Miscanthus gigantheus. https://commons.wikimedia.org/wiki/F...us_Bestand.JPG. Creative Commons Attribution-Share Alike 3.0 Unported

Figure 32.8.22 shows some data from the study. Four different scenarios are offered, each represented by bar graphs.

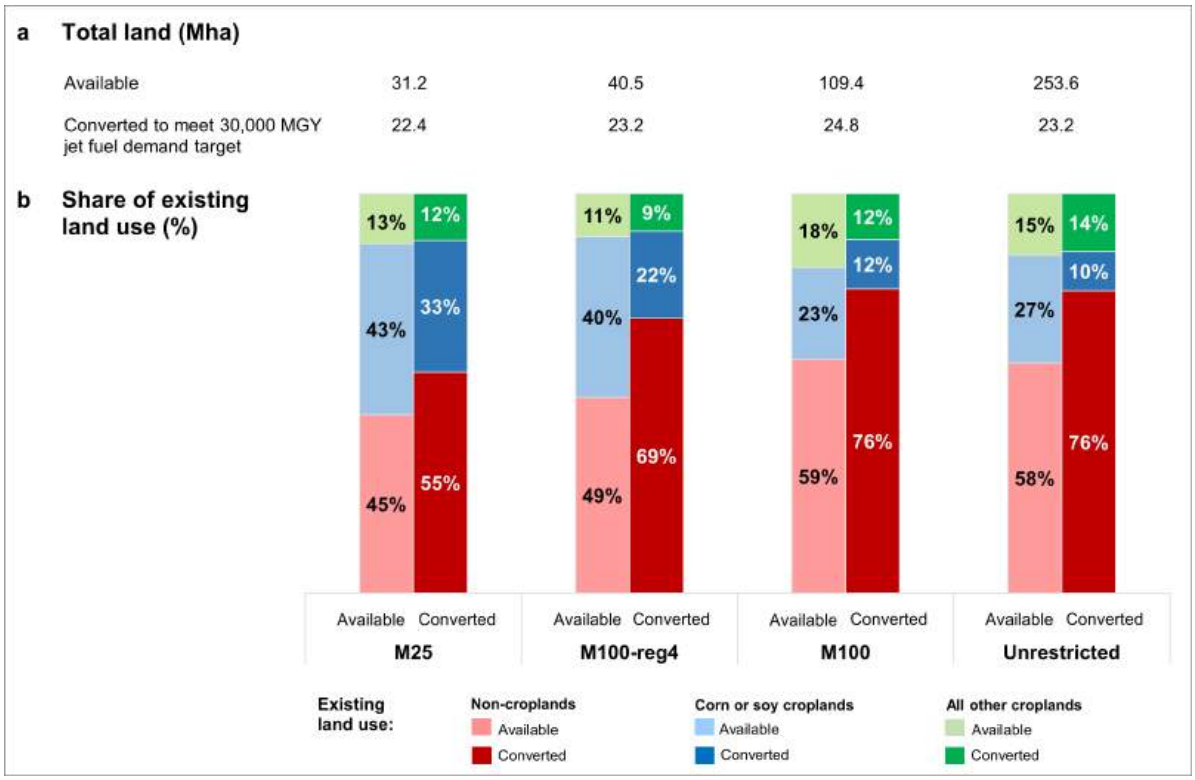


Figure 32.8.22: Land availability and conversion by existing use. Excel data and graph from https://dataverse.harvard.edu/datase...910/DVN/VBFLI2. CC0 1.0 Public Domain.

Four scenarios (left to right) were used to produce 30 MG/yr of SAF using a carbon price of $50/t CO2e.

- M25: only 25% of the marginally useful land was used
- M100-reg4: marginal land bases with the lowest hydrological and climatic risks
- M100: all of the marginally included land was made available
- Unrestricted

Panel (a) shows that in each case, about 23.2 Mha of land was converted to growing miscanthus out of the available land. Pane (b) shows stacked bars showing the percentages of each type of land available and converted for each scenario. The last scenario shows that demand can be made with the lowest % conversion of the marginal lands now used for corn/soybeans and other crops. It would appear that the marginal croplands converted to SAF production would be the same lands diverted to bioethanol production. Nevertheless, it would appear that up to 76% of projected aviation fuel needs could be met by planting marginal cropland and noncrop land for cellulosic SAF production. The study found that using available lands in the Plains was not feasible.

KEY POINTS - BETA VERSION FROM CHAT.OPENAI

1. Biodiesel, syngas and bioaviation fuel are alternative biofuels derived from biomass that can be used to reduce the dependence on fossil fuels.
2. Biodiesel is a liquid fuel that can be used in diesel engines, it is produced by chemically converting vegetable oils or animal fats into a fuel that can be used in place of diesel fuel.
3. Syngas (synthetic gas) is a mixture of hydrogen and carbon monoxide that can be produced from biomass through processes such as gasification. It can be used as a fuel for heat and power generation or further converted into chemicals and liquid fuels.
4. Bioaviation fuel is a form of biofuel that can be used in aviation, it is made from biomass such as algae or woody biomass, it can be blended with traditional jet fuel and can help reduce emissions from airplanes.
5. These biofuels have different advantages and limitations, and their production process is still in the research and development stage.
6. Biodiesel has the advantage of being able to be used in existing diesel engines with little or no modification, it reduces emissions and it is renewable.
7. Syngas has the advantage of being able to be converted into different chemicals and liquid fuels, it has a higher energy content than bioethanol and it can be used for power generation.

8. Bioaviation fuel has the advantage of reducing emissions from airplanes, it can reduce dependence on fossil fuels in the aviation industry and it can be renewable.

32.09: BIOHYDROGEN - AN INTRODUCTION

Search Fundamentals of Biochemistry

32.09.1: H_2 AS A FUEL

Hydrogen gas would be ideal if it could be produced at scale, easily transported and stored, or produced at local sites on demand. The reaction for the "burning" of hydrogen shows that the only greenhouse gas emitted is H_2O.

$$2H_2(\ g) + O_2(\ g) \rightarrow 2H_2O(g) \tag{32.9.1}$$

H_2O comes and goes in our atmosphere in short timescales and does not continually build up, as does CO_2 from burning fossil fuels. The standard heat of combustion (in kJ/g or kcal/kg) for H_2 is far higher than any other fuel, as shown in Table 32.9.1 below, making it an ideal fuel.

Name	Formula	State	$-\Delta H_c°$ *kJ/mol*	$-\Delta H_c°$ *kJ/g or MJ/kg*	$-\Delta H_c°$ *kcal/kg*
Ammonia	NH_3	gas	383	22.48	5369
Butane	C_4H_{10}	gas	2878	49.50	11823
Carbon (graphite)	C	cry	394	32.81	7836
Ethanol	C_2H_6O	liq	1367	29.67	7086
Hydrogen	**H_2**	**gas**	**286**	**141.58**	**33817**
Methane	CH_4	gas	891	55.51	13259
methyl stearate (biodiesel)	$(CH_3(CH_2)_{16}(CO)CH_3$	liq	1764	40	9560
Naphthalene	$C_{10}H_8$	cry	5157	40.23	9609
Octane	C_8H_{18}	liq	5470	47.87	11434
Propane	C_3H_8	gas	2220	50.33	12021
wood (red oak)	-	solid	-	14.8	3540
coal (lignite)	-	solid	-	15	3590
coal (anthracite)	-	solid		27	4060

Table 32.9.1: Energy values for various fuels. Data source: https://www.engineeringtoolbox.com/s...nt-d_1987.html

We won't discuss large-scale H_2 storage or transport, two fundamental engineering problems. Instead, we will focus on the production of "biohydrogen." Of course, the prefix "bio" can mean many things, including the production of H_2 in syngas using cellulose as a feedstock, the electrolysis of water powered by solar/wind energy, and its production by hydrogenases, enzymes found in some microbes.

The fuel industry uses different colors as descriptors of hydrogen based on how it is produced. They are shown in Table 32.9.2.

Color	Method of production
Green	electrolysis of H_2O using solar/wind to generate electricity (expensive at present)
Blue	steam reforming of natural gas (CH_4) with the other product, CO_2 captured and stored (CCS)
Grey	steam reforming of natural gas (CH_4) without CO_2 capture and storage
Black (coal/oil)	gasification to form syngas
Pink (purple/red)	electrolysis powered by nuclear energy, which does not emit CO_2; heat emitted produces steam for blue/gray H_2 production
Turquoise	methane pyrolysis (heat in the absence of O_2) to form H_2 and C
Yellow	electrolysis using solar power without conversion to electricity as the power source.
White	underground H_2 released through fracking

Table 32.9.2: Different "colors" of hydrogen based on production methods

Of course, H_2 in syngas can be produced from biomass, as described in Chapter 32.8, but it is unclear if a hydrogen color has been assigned to it.

At present, the important feedstocks for H_2 production around the world are natural gas (48%), oil (30%), coal (18%), and electrolysis (4%) - mostly all fossil fuels.

32.09.2: METHODS OF PRODUCTION

H_2 production is also classified based on the chemical processes used to produce it. These processes include

1. **Biological** (use of live bacteria and algae cells)
2. **Thermochemical**: (gas and liquid fuel reforming, coal and biomass gasification),

3. **Electrochemical (electrolytic)**: (photothermal, photoelectrolytic, and photobiological)

We will organize this chapter section using these three processes. We will start with Biological (1), followed by Electrochemical/Electrolytic (3), and end with Thermochemical (2). They are summarized in Figure 32.9.1.

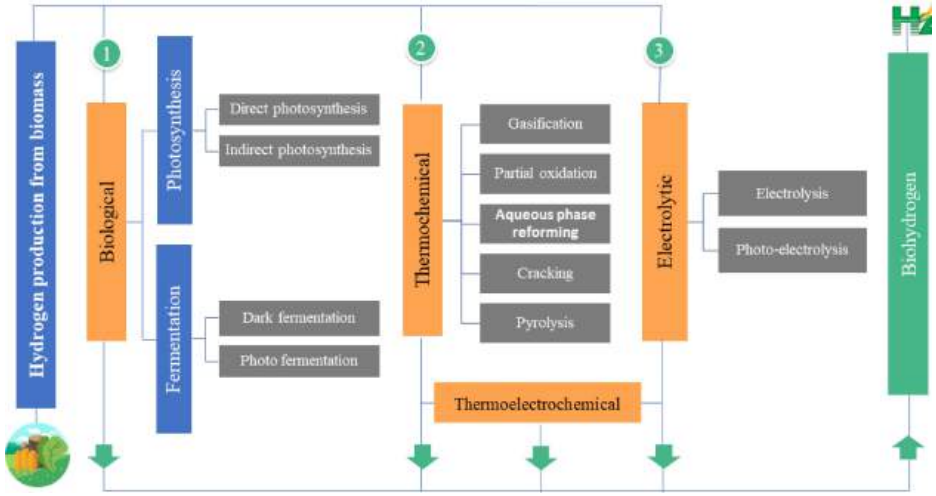


Figure 32.9.1: The main pathways for H_2 production based on biomass. M.G. Eloffy et al., Chemical Engineering Journal Advances, 12 (2022). https://doi.org/10.1016/j.ceja.2022.100410. Creative Commons license

Biomass can be used as the feedstock for all of these methods, so the resulting product can be called biohydrogen. Of course, nonbiological sources of feedstocks are the predominant ones used in thermochemical and electrochemical methods as well, as we discussed in the previous chapter section.

32.09.3: $2H^+ \leftrightarrow H_2$: AN OVERVIEW

We will mostly discuss the production of H_2 as a society energy source. For industry use, it can be used in fuel cells to power spacecraft and cars, as shown in the reaction below.

$$\begin{aligned} O_2 + 4H^+ + 4e^- &\longrightarrow 2H_2O \\ H_2 &\longrightarrow 4H^+ + 4e^- \end{aligned} \qquad (32.9.2)$$

In the next chapter section, we will discuss in great detail the **hydrogenases** that produce and use H_2 in microbes so that this chapter will treat them very generally. However, we need to review the topic.

Use of H_2 as a *source* of electrons for reduction reactions.

Each hydrogen in H_2 has an oxidation number of 0. Each hydrogen can be oxidized to H^+ (oxidation number +1) with the 2 electrons passed on to a substrate/cofactor or a sequential series of substrates with higher and higher standard reduction potentials (better oxidizing agents), leading to the formation of reduced products.

H_2 + $(substrate)_{OX}$ → $2H^+$ + $(product)_{RED}$

This general reaction is analogous to the mitochondrial electron transport chain, in which electrons are passed from a source (NADH) to oxidized forms of acceptors. The general reaction below shows each redox pair in the electron transport chain.

$NADH/NAD^+$ → $FAD/FADH_2$ → UQ/UQH2 → Cyto C_{OX}/Cyto C_{RED} → O_2/H_2O

Some organisms have evolved to produce energy by the oxidation of H_2. This reaction is analogous to those used by photosynthetic organisms to obtain energy through the oxidation of water. In photosystem II, oxygen in H_2O (oxidation number -2) gets oxidized by the oxygen-evolving complex to produce O_2 (oxidation number 0). Some redox pairs, starting with H_2O/O_2, are shown below for photosystem II.

H_2O/O_2 → P680/P680* → $(Plastoquione)_{OX}/(Plastoquione)_{RED}$

The first reaction is endergonic and requires an energy source photons.

Use of H^+ as a *sink* for electrons for oxidation reactions that produce H_2.

H^+ has an oxidation number of +1. Hence it can be reduced to H_2 (oxidation number of 0) as it gains electrons from substrates/cofactors, which get oxidized. This general reaction is shown below.

$2H^+$ + (substrate/cofactor)RED → H_2 + (substrate/cofactor)OX

Many microorganisms can produce H_2 through variants of photosynthesis or through fermentation, both of which provide the two electrons needed. E. Coi has four hydrogenases (Hyd 1, 2, 3, and 4). It forms H_2 through two reactions catalyzed by:

- formate (HCO_2^-) dehydrogenase (FDH): $2HCO_2^- \rightleftharpoons 2CO_2 + 2H^+ + 2e^-$
- hydrogenase (H_2ase): $2H^+ + 2e^- \rightarrow H_2$

The C in formate has an oxidation number of +2 and is oxidized to CO_2, in which the C has an oxidation number of +4.

NOTHING IS SIMPLE: H_2 IS AN INDIRECT GREENHOUSE GAS.

H_2 itself is not a greenhouse gas as it doesn't have any bond vibrations that produce transient dipoles and hence does not absorb in the infrared region of the spectrum. Yet by affecting atmospheric levels of methane, a very potent greenhouse gas, as well as levels of ozone, it can lead to warming. It's not emission from the combustion of H_2 but rather the leakage into the atmosphere of transported and stored H_2 gas that is problematic.

Most of the H_2 that finds its way into the atmosphere diffuses into the soil and is taken up by bacteria. The rest reacts with hydroxy radicals (·OH) in the atmosphere, as shown in the reaction below.

$$H_2 + \cdot OH \rightarrow H_2O + H\cdot \text{ (atomic hydrogen)}$$

The reaction of ·OH with H_2 decreases the hydroxy radical's availability to react with the very potent greenhouse gas methane, CH_4. That reaction is shown below.

$$CH_4 + \cdot OH \rightarrow \cdot CH_3 + H_2O$$

The methyl radical $\cdot CH_3$ reacts rapidly with oxygen to form the methylperoxy radical ($CH_3O_2\cdot$). This eventually forms formaldehyde, a water-soluble molecule that is removed from the atmosphere on precipitation. Hence the reaction of H_2 with ·OH increases the half-live of CH_4 in the atmosphere.

·OH is a key molecule in the troposphere and is considered a methane "sink" that leads to the drawdown of methane. We discussed the extreme reactivity of ·OH in Chapter section's 12.3 and 12.4. It's so reactive that its half-life is in the order of seconds. It is also at very low concentrations of less than 1 part per trillion.

·OH is produced from ozone, O_3, by the following reactions:

$$O_3 + h\nu \text{ (UV)} \rightarrow O_2 + \cdot O$$

$$\cdot O + H_2O \rightarrow 2\ \cdot OH$$

The first reaction is a photolysis, and experiments during a solar eclipse have shown the production of ·OH in the atmosphere shuts down!

Dr. Paul Crutzen, Nobel Prize winner in Chemistry, described ·OH as the "detergent of the atmosphere" since it can react with and oxidize many deleterious trace gases in the troposphere, making them more water-soluble, leading to their elimination from the atmosphere. A main reaction of ·OH is carbon monoxide (CO). It also reacts with volatile organic compounds (VOCs) and NOx (NO + NO_2), which are precursors of tropospheric ozone, a health hazard. Even though dioxygen, which comprises 20% of the atmosphere, is also an excellent oxidizing agent, it is kinetically slow to react.

Very few gases are not oxidized by ·OH. One set includes the refrigerant gases chlorofluorocarbons, which without oxidation by ·OH enter the stratosphere, where they react with stratospheric ozone and reduce its protective effect against dangerous UV light. It does react with hydrochlorofluorocarbons (HCFCs).

Figure 32.9.2 below summarizes the adverse climatic effects of the oxidation of H_2 in the atmosphere.

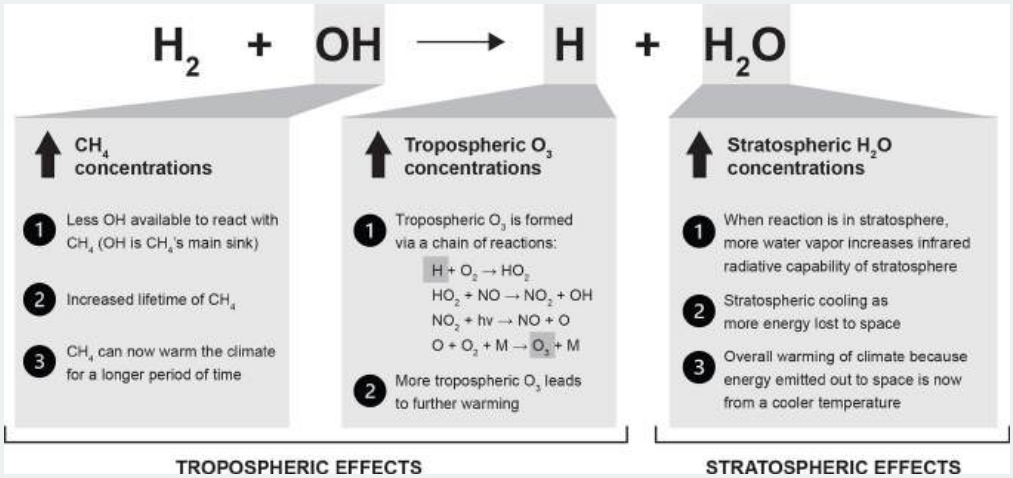


Figure 32.9.2: Effects of hydrogen oxidation on atmospheric greenhouse gas concentrations and warming. I. Ocko and Steven P. HamburgAtmos. Chem. Phys., 22, 9349–9368, 2022. https://doi.org/10.5194/acp-22-9349-2022. Creative Commons Attribution 4.0 License.

Note in the central panel that H· (atomic hydrogen) can start a free radical change reaction to produce tropospheric ozone, O_3, a pollutant that is not only a greenhouse gas but which also causes serious health consequences.

The message is this: Care has to be taken to minimize methane and H_2 leakage during their production and use as fuels.

32.09.4: BIOHYDROGEN FROM MICROALGAE

We will focus most of our attention on the Biological (1) and Electrolytic (3) processes for producing biohydrogen from microalgae. The Biological processes (1) require hydrogenases for H_2 production within cells. The Electrolytic (3) processes use microalgae as a feedstock to provide substrates that other microbes can ferment. These can be combined to increase production. Figure 32.9.3 below summarizes the Biological (1) and Electrolytic (3) metabolic processes that can be used for microalgae H_2 production.

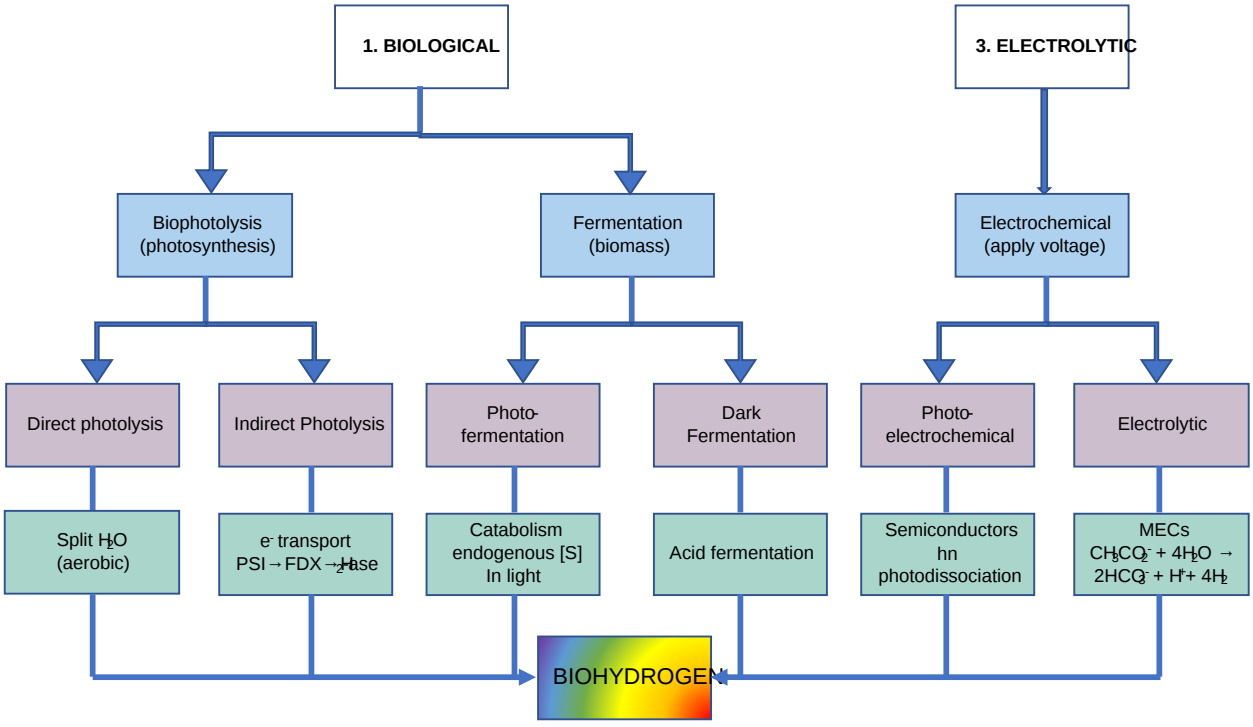


Figure 32.9.3: Metabolic pathways of biohydrogen production by micro-algal biomass. modified from Ahmed SF et al. *Front. Energy Res.* 9:753878. doi: 10.3389/fenrg.2021.753878. Creative Commons Attribution License (CC BY).

These are mainly classified into three categories: i) the photobiological process through which biohydrogen is produced *via* direct and indirect photolysis in the microalgae, ii) fermentation, and iii) the electrochemical process that comprises photoelectrochemical and electrolytic.

32.09.4.1: BIOLOGICAL (1) - BIOPHOTOLYSIS (PHOTOSYNTHESIS)

This consists of two processes, Direct and Indirect Photolysis (photosynthesis). Both use light to drive the ultimate reduction of $2H^+$ to H_2 using hydrogenase or nitrogenase. We will explore the details in the next chapter section. The biophotolysis process is divided into indirect (using electrons from substrates) and direct (using electrons from water). These processes are simplified in Figure 32.9.4.

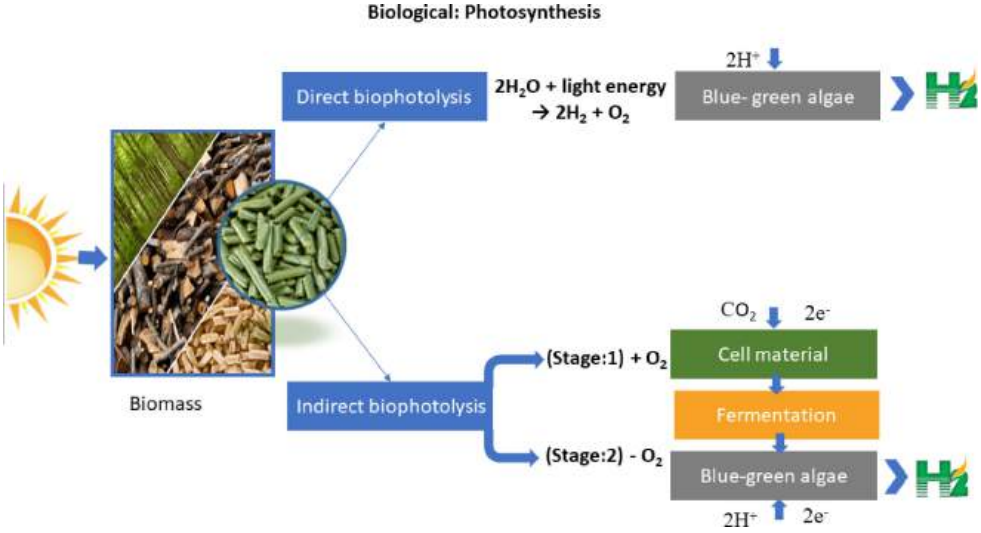


Figure 32.9.4: Schematic diagram for biological (biophotolysis) process. M.G. Eloffy et al.

Direct Biophotolysis (photosynthesis)

In direct biophotolysis (photosynthesis), water molecules are oxidized in Photosystem II, which contains the Oxygen Evolving Complex (OEC). This endergonic process is driven by light. The electrons lost from water are passed through Cytochrome b_6f and Photosystem I to ferredoxin then $NADP^+$, which gets reduced to NADPH (as discussed in Chapter 20). These reactions are illustrated in Figure 32.9.5.

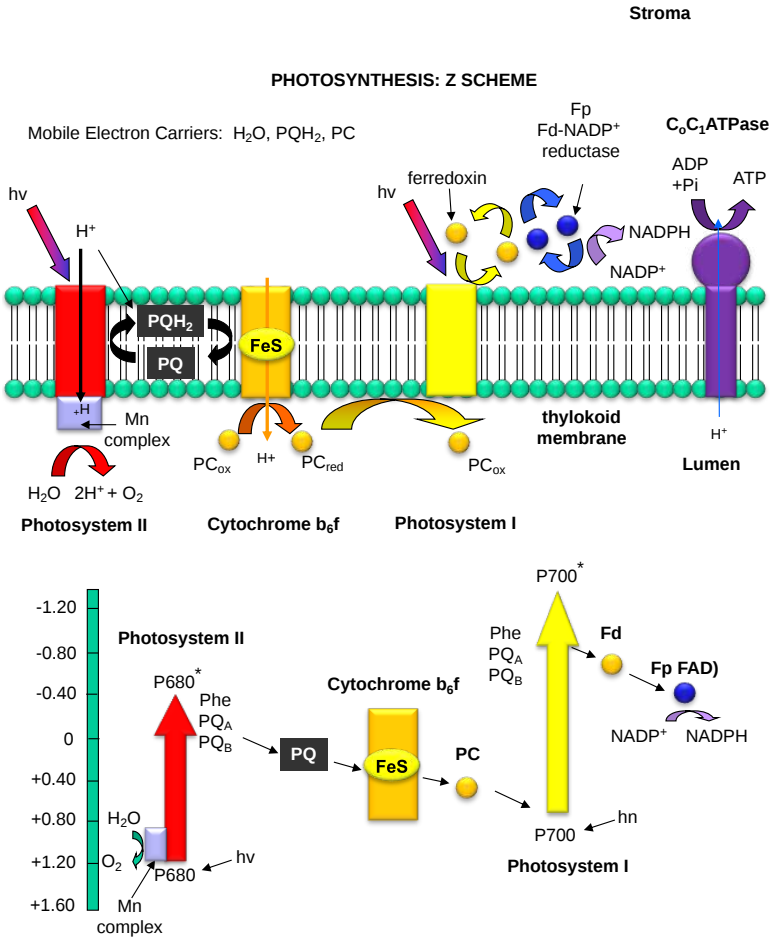


Figure 32.9.5: Light reaction of photosynthesis and associated standard reduction potentials

In direct photolysis, electrons are passed directly from reduced ferredoxin to $2H^+$ in a reaction catalyzed by a hydrogenase, as shown in Figure 32.9.6 below.

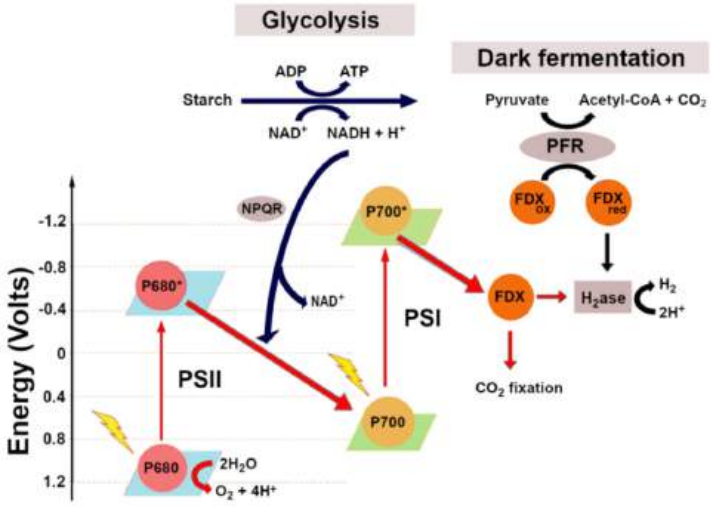


Figure 32.9.6: Metabolic hydrogen production pathways used by *Chlamydomonas reinhartii*.FDX: ferredoxin; H_2ase: hydrogenase; NPQR: NADPH−plastoquinone oxidoreductase; PFR: pyruvate:ferredoxin oxidoreductase; PSI: photosystem I; PSII: photosystem II. Touloupakis, E.; Faraloni, C.; Silva Benavides, A.M.; Torzillo, G. Recent Achievements in Microalgal Photobiological Hydrogen Production. *Energies* **2021**, *14*, 7170. https://doi.org/10.3390/en14217170.

The overall reaction is simplified in the equation below.

$$2H_2O + \text{Light} \rightarrow 2H_2 + O_2 \tag{32.9.3}$$

One problem with direct photolysis is that O2 can damage hydrogenases. Again we will discuss the biochemistry of hydrogenases in great detail in the next chapter section.

Indirect Biophotolysis

This process bypasses the damaging effects of O_2 on hydrogenase by being carried out in the absence of O_2 using fermentation to provide electrons for the hydrogenase reduction of $2H^+$ to H_2. Photosynthesis is required to make the carbohydrates necessary for fermentation. Glucose can then be oxidized anaerobically (in the dark to avoid O_2 formation from photosynthesis) to form pyruvate through the glycolytic pathway. Pyruvate can then be oxidatively decarboxylated through the **pyruvate:ferredoxin oxioreductase (PFR)** as ferredoxin gets reduced. It then passes its electrons on through hydrogenase to produce H_2. The pathway is illustrated in the top/right parts of the above figure and the reaction diagram in Figure 32.9.7 below.

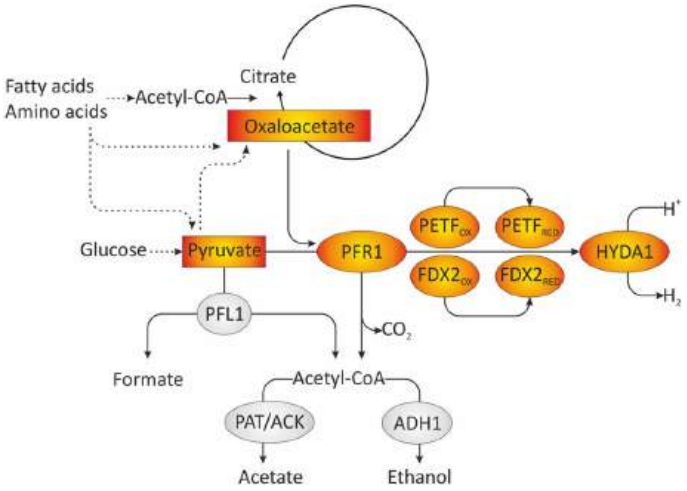


Figure 32.9.7: **Model of fermentative pathways involved in dark anaerobic H_2 production in *C. reinhardtii*.** Proteins are shown as ovals. Photosynthetic ferredoxin (PETF). Jens Noth et al., Journal of Biological Chemistry, 288 (2013). https://doi.org/10.1074/jbc.M112.429985. Creative Commons license.

Glucose and some amino acids can be converted into pyruvate, a substrate for PFR1 in the single-cell algae *C. reinhardtii*. PFR1 converts pyruvate to acetyl-CoA and CO_2 with the electrons used to reduce ferredoxin. The reduced FDX2 passes electrons through hydrogenase (HYDA1) to form H_2.

Another enzyme used to continue fermentation, **pyruvate:formate lyase (PFL1),** converts pyruvate to formate and acetyl-CoA, which can be metabolized further to acetate and ethanol. A shift to pyruvate oxidation to PFR1 occurs if PFL1 is mutated or long term anoxic conditions.

The key enzyme, pyruvate:ferredoxin oxioreductase (PFR), uses thiamine pyrophosphate (TPP) as a cofactor for the oxidative decarboxylation of the α-keto acid pyruvate, as expected. Figure 32.9.8 shows an interactive iCn3D model of the pyruvate ferredoxin oxidoreductase (PFOR) from Desulfocurvibacter africanus in anaerobic conditions (7PLM).

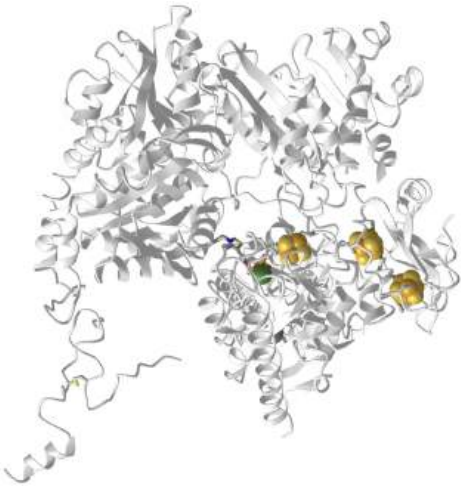


Figure 32.9.8: Pyruvate ferredoxin oxidoreductase (PFOR) from Desulfocurvibacter africanus in anaerobic conditions (7PLM). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...XRFMoVhaoRbW86

PFOR, abbreviated here, is a 267 kDa homodimer containing three [Fe_4S_4] clusters (spacefill) per monomer. Only one monomer is shown, and TPP is shown in sticks.

Again indirect photolysis occurs in the absence of O_2. Light illumination leads to only transient H_2 synthesis. If sulfur is limited in the growth media of the algae, more sustained H_2 production occurs, as the lack of sulfur reduces PSII activity. Hence H_2 production can be maximized by depleting sulfur and minimizing O_2 even in the presence of light. In the absence of O_2, hydrogenase gene expression increases. Nutrient depletion also leads to the production of formate and acetyl-CoA through the enzyme pyruvate:formate lyase (PFL1). This is predominant in *Chlamydomonas* cells in the dark.

The green microalgae *C. reinhardtii* makes most of its H_2 (approximately 90%) using direct photolysis. Commercially, the production of H_2 in indirect photolysis is carried out in a separate sealed bioreactor to avoid O_2. Indirect photolysis is shown in the above figures.

The reactions to this process are as follows :

$$\begin{aligned} 12H_2O + 6CO_2 + h\nu &\rightarrow C_6H_{12}O_6 + 6O_2 \\ C_6H_{12}O_6 + 12H_2O + h\nu &\rightarrow 12H_2 + 6CO_2 \end{aligned} \tag{32.9.4}$$

32.09.4.2: BIOLOGICAL (1) - FERMENTATION

We have just discussed fermentation processes **within living microalgae cells**. Now let's consider fermentation processes using **nonliving biomass feedstocks** supplied to microbes to produce H_2. This offers a significant way to make biohydrogen. A schematic diagram for Biological (1) fermentation is shown below in Figure 32.9.9:

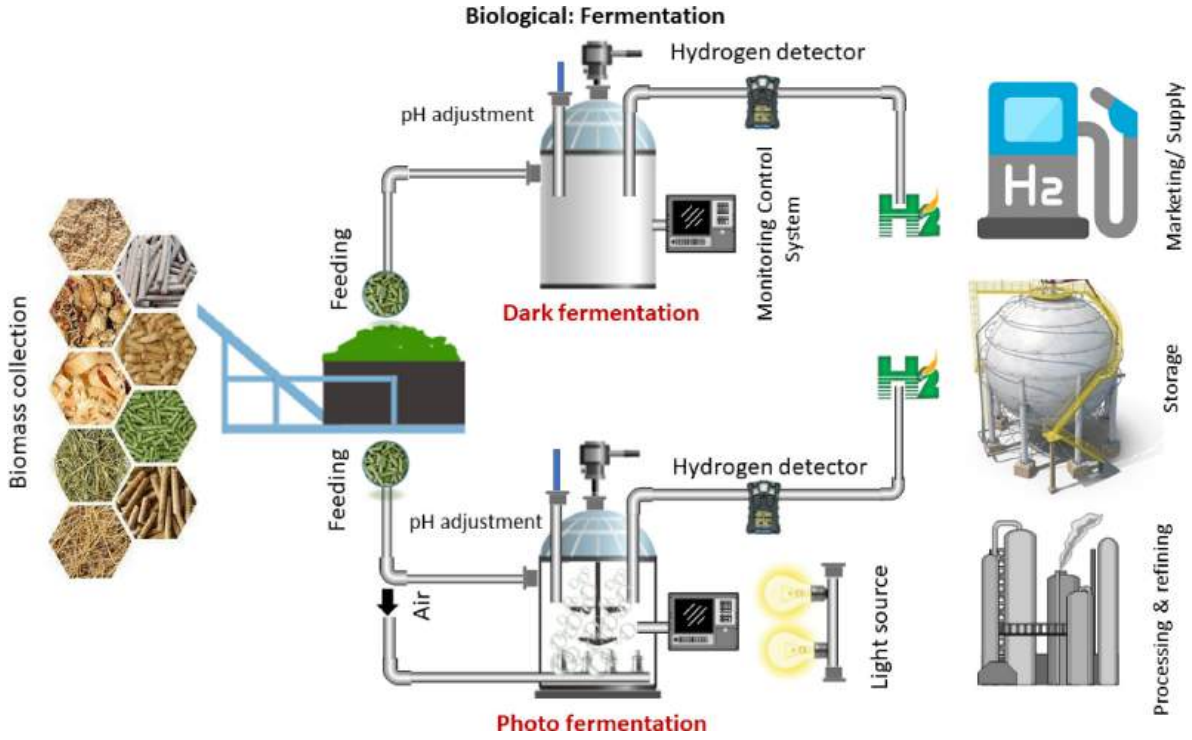


Figure 32.9.9: Schematic diagram for biological (fermentation) process. M.G. Eloffy et al.

Fermentation involves the decomposition of organic biomass to produce CO_2 and H_2. The fermentation process can be separated into photofermentation (light fermentation) and dark fermentation.

Photofermentation

Some photosynthetic bacteria and microalgae use Photofermentation to produce H_2 from organic acids like acetic, butyric, lactic, and succinic acids. Oxidation of the acids produces CO_2 as well as H^+s and e^- for H_2 production. Electrons are transferred through photosystem I and eventually, believe it or not, nitrogenase. It is a fermentation process as the process is anoxic.

Some photosynthetic bacteria, like the purple nonsulfur bacteria, a facultative anoxygenic phototroph, and some microalgae, can produce H_2 using a simplified system that has only one photosystem and uses the enzyme nitrogenase to produce H_2. The photosystem can not generate an oxidizing agent strong enough to oxidize H_2O, but under anaerobic conditions, they can oxidize organic acids and even H_2S to provide electrons from H_2 production. These reactions are shown below in Figure 32.9.10.

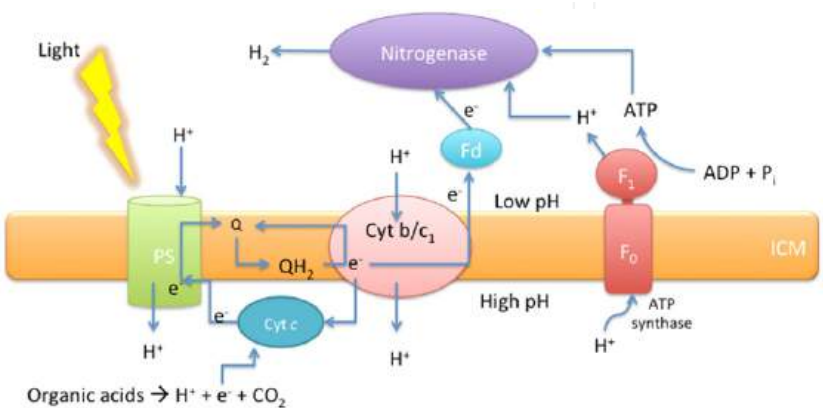


Figure 32.9.10: Photofermentative hydrogen production in PNSB.

Deo, D., Ozgur, E., Eroglu, I., Gunduz, U., & Yucel, M. (2012). Photofermentative Hydrogen Production in Outdoor Conditions. Hydrogen Energy - Challenges and Perspectives. doi: 10.5772/50390. Creative Commons Attribution 3.0 License,

We studied nitrogenase in Chapter x.xx. The net reaction for the fixation of nitrogen is shown below.

$$N_2 + 8H^+ + 8e^- + 16ATP \rightarrow 2NH_3 + H_2 + 16ADP + 16Pi \quad (32.9.5)$$

In this reaction, N_2 is reduced as the N atoms go from a 0 oxidation state to +3 in NH_3. The needed electrons are made from organic acids and fed into the system and eventually go to ferredoxin, which transfers them to protons. The ratio of N_2 to H_2 produced is 1:1, at the expense of 16ATPs per H_2 produced.

The ATP produced by the collapse of the produced proton gradient through FoF1ATPase powers the reaction.

In the absence of N_2, the net reaction becomes

$$2H^+ + 2e^- + 4ATP \rightarrow H_2 + 4ADP + 4Pi \quad (32.9.6)$$

The electrons are still fed into nitrogenase, but in the absence of the substrate N_2, they are used to reduce $2H^+$ to H_2. Note that only 4 ATPs are required per each H_2 produced, a significant energy gain.

ATP produced during photosynthesis would be used for anabolic biosynthesis contributing to biomass, so extra ATP is needed to support H_2 synthesis past that needed for growth. As anabolism is a reductive process (compared to oxidative catabolism), adequate sources of electrons for reduction are required. Multiple pathways need electrons, including CO_2 fixation, N_2 fixation (with associated H_2 production, and organic acids like polyhydroxbutyrate. The bacteria use photosynthesis and the Calvin cycle under photoautotrophic conditions to fix CO_2. When external energy supplies from organic acids are present, the bacteria can become photoheterotropic. Under these conditions, the Calvin cycle is used to maintain redox balance.

Dark Fermentation

We studied this indirectly above section in our discussion of hydrogenases in microalgae. Hydrogenases are induced in dark conditions, and this pathway involved heterotrophic fermentation (anaerobic) in some bacteria and microalgae. Many microbial species are used. Industrial wastewater enriched in organic material can be used as a feedstock.

Feedstock materials are hydrolyzed and subjected to fermentation, during which H_2 can be produced. For example, pyruvate produced by glycolytic fermentation can be oxidatively decarboxylated to acetyl-CoA and CO_2 by pyruvate:ferredoxin oxidoreductase with electrons passed on to ferredoxin and even through hydrogenase to form H_2 (as we described above). Addition H_2-produced steps after fermentation include acetogenesis and methanogenesis. These processes are illustrated in Figure 32.9.11 below.

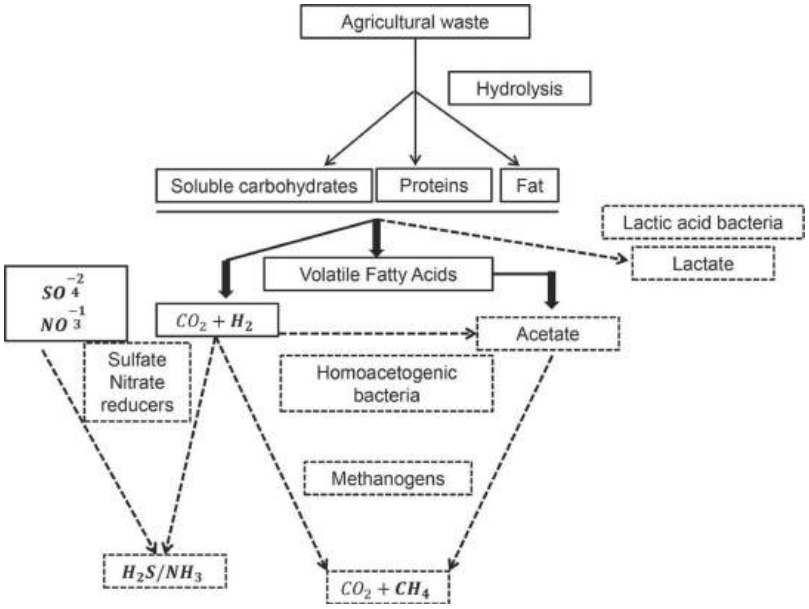


Figure 32.9.11: The steps involved in anaerobic digestion [9]. Rosa, P. R. F., & Silva, E. L. (2017). Review of Continuous Fermentative Hydrogen-Producing Bioreactors from Complex Wastewater. Frontiers in Bioenergy and Biofuels. doi: 10.5772/65548. Creative Commons Attribution 3.0 License

Examples of acidogenic (formation of short carboxylic/fatty acid), acetogenic (formation of acetic acid), and methanogenic (formation of methane) reactions that produce (and a few that consume) H_2 are shown in Table 32.9.3 below.

Acidogenic reactions
$C_6H_{12}O_6 + 2H_2O \rightarrow 2CH_3COOH + 2CO_2 + 4H_2$
$C_6H_{12}O_6 + 2H_2O \rightarrow CH_3CH_2CH_2COOH + 2CO_2 + 2H_2$
Acetogenic reactions
$CO_2 + 4H_2 \rightarrow CH_3COOH + 2H_2O$
$CH_3CHOHCOOH + H2O \rightarrow CH_3COOH + CO_2 + 2H_2$
$CH_3CH_2OH + H2O \rightarrow CH_3COOH + 2H_2$
$CH_3CH_2COOH + 2H_2O \rightarrow CH_3COOH + CO_2 + 3H_2$
$CH_3(CH_2)_2COOH + 2H_2O \rightarrow 2CH_3COOH + 2H_2$
Methanogenic reactions
$4H_2 + CO_2 \rightarrow CH_4 + 2H_2O$
$CH_3COOH \rightarrow CH_4 + CO_2$
$2CH_3(CH_2)_2COOH + 2H_2O + CO_2 \rightarrow 4CH_3COOH + CH_4$

Table 32.9.3: Example of acidogenic, acetogenic, and methanogenic reactions in dark fermentation. Adapted from Rosa, P. R. F., & Silva, E. L., ibid.

We have described a few of the enzymes involved in acidogenic reactions above. Figure 32.9.12 shows a summary of the steps in acidogenesis.

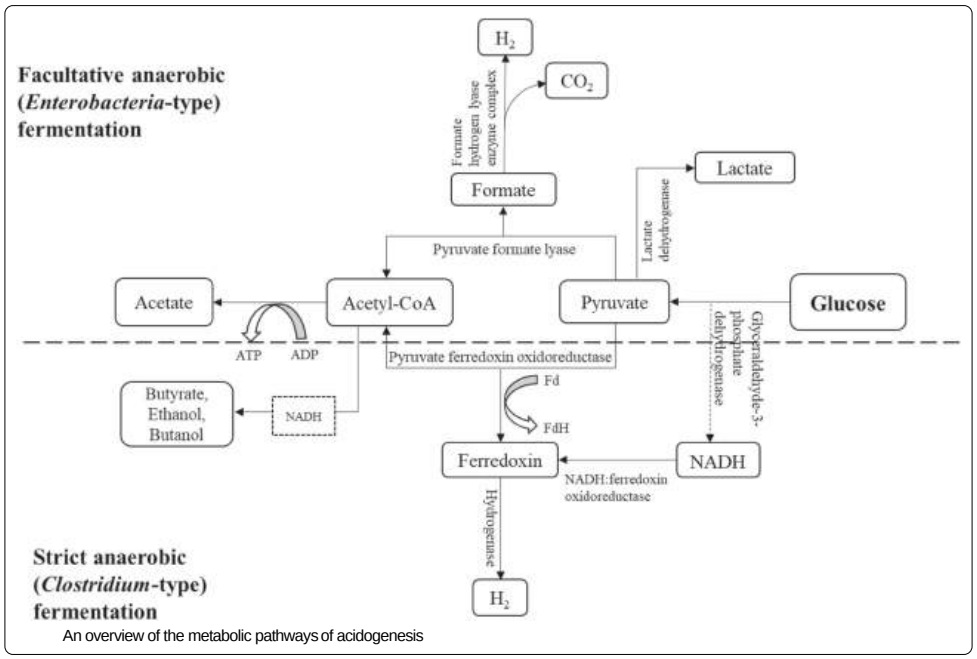


Figure 32.9.12: An overview of the metabolic pathways of acidogenesis. Dzulkarnain, E.L.N., Audu, J.O., Wan Dagang, W.R.Z. *et al.* Microbiomes of biohydrogen production from dark fermentation of industrial wastes: current trends, advanced tools and future outlook. *Bioresour. Bioprocess.* **9**, 16 (2022). https://doi.org/10.1186/s40643-022-00504-8. http://creativecommons.org/licenses/by/4.0/.

A more complex list and summary of dark fermentation reactions are shown in Figure 32.9.13 below.

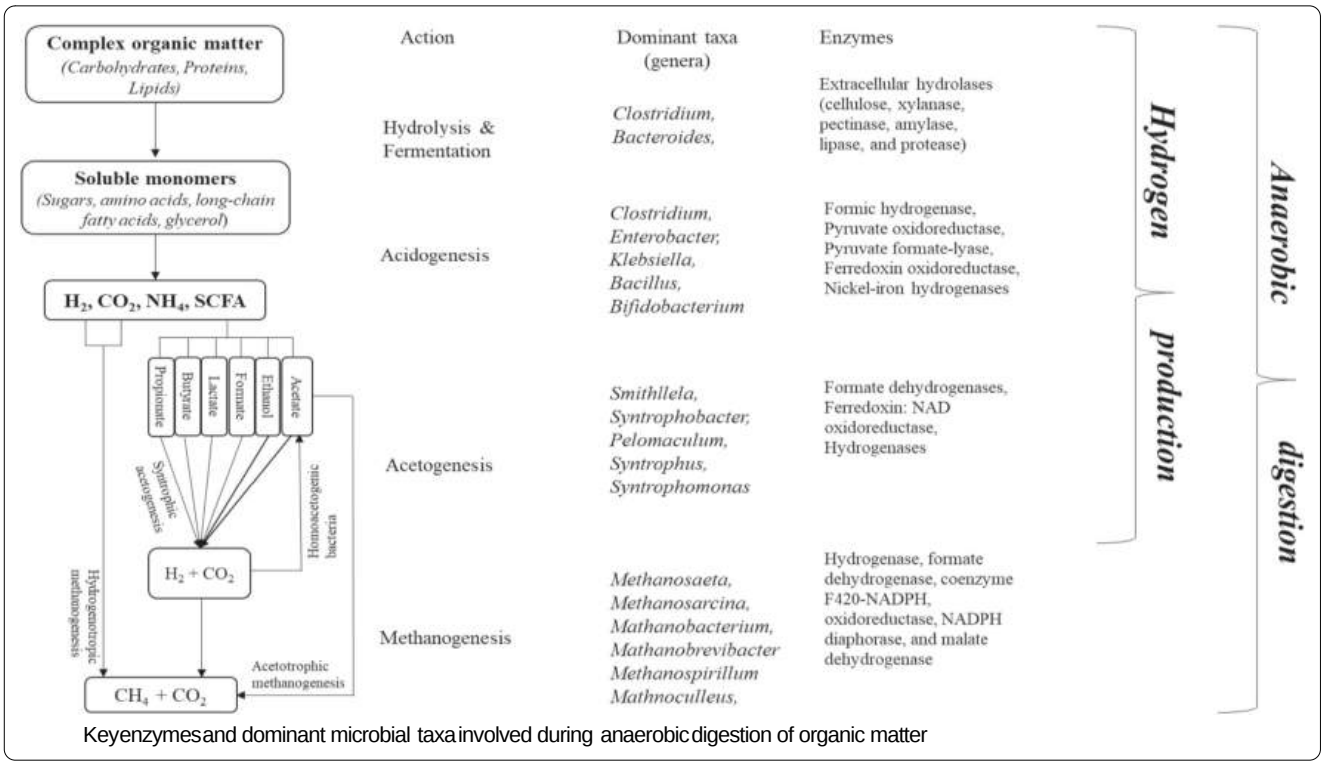
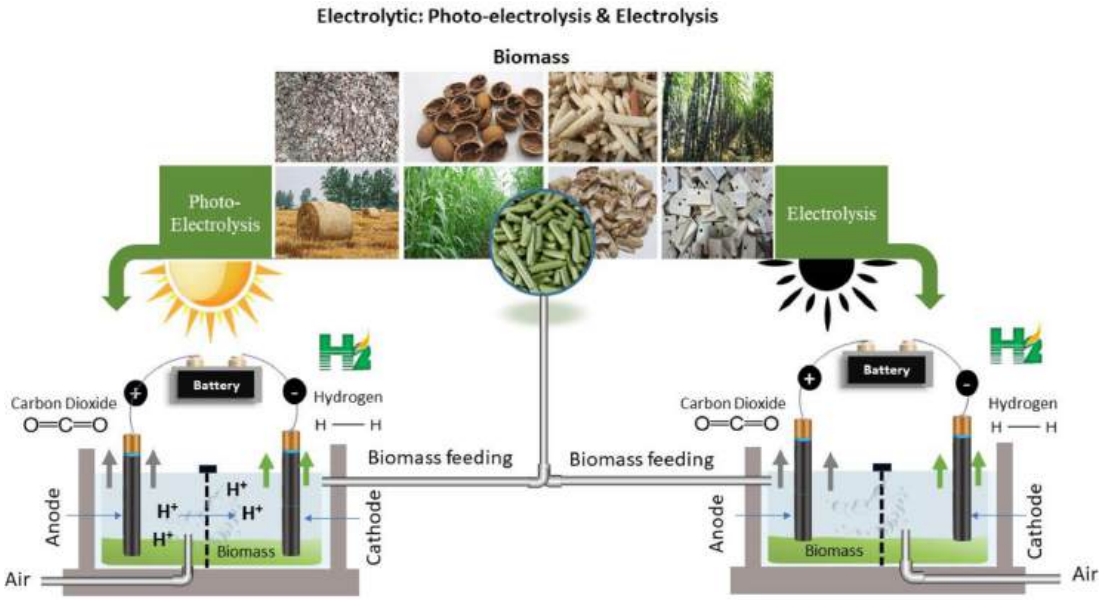


Figure 32.9.13: Key enzymes and dominant microbial taxa involved during anaerobic digestion of organic matter. Dzulkarnain, E.L.N.et al. Ibid.

32.09.4.3: ELECTROCHEMICAL/ELECTROLYTIC (3)

Two primary electrochemical/electrolytic methods for H_2 production are photoelectrochemical and electrolytic, as shown in Figure 32.9.14 below.



djfkjdk

Figure 32.9.14: Schematic diagram for the electrochemical process. M.G. Eloffy et al.

Electrolysis

In a microbial electrolytic cell (MEC), microalgae/cyanobacteria use industrially- and metabolically-processed feedstocks to oxidize organic substrates (for example, acetic acid) to CO_2. The released electrons move from the anode (where the oxidation occurs) to the cathode for H^+ reduction to H_2. An external voltage is applied to increase electron flow to the cathode to facilitate the process. This increases the production of H_2 over and above that of just fermentation by microbes in the electrolytic cell. Cyanobacteria and a mix of green microalgae are used, as well as bacteria, which can use dark fermentation (i.e., combining the processes described above).

Photoelectrochemical

Microbial photoelectrochemical cells (PEC) use light-sensitive semiconductor electrodes for water electrolysis. A membrane separates the two electrodes so the protons can be reduced to form H_2. for the separated by a membrane,

32.09.5: 2. THERMOCHEMICAL FROM BIOMASS

We have already explored thermochemical methods to produce syngas (H_2 and CO) and further use in the Fishcer-Tropsch reaction to make small and large molecules for chemical feedstocks and fuels. We also discussed electrochemical methods to produce syngas and other small organic molecules like formate and ethanol from CO_2. Figure 32.9.15 shows a schematic diagram for thermochemical (gasification) processes to produce H_2.

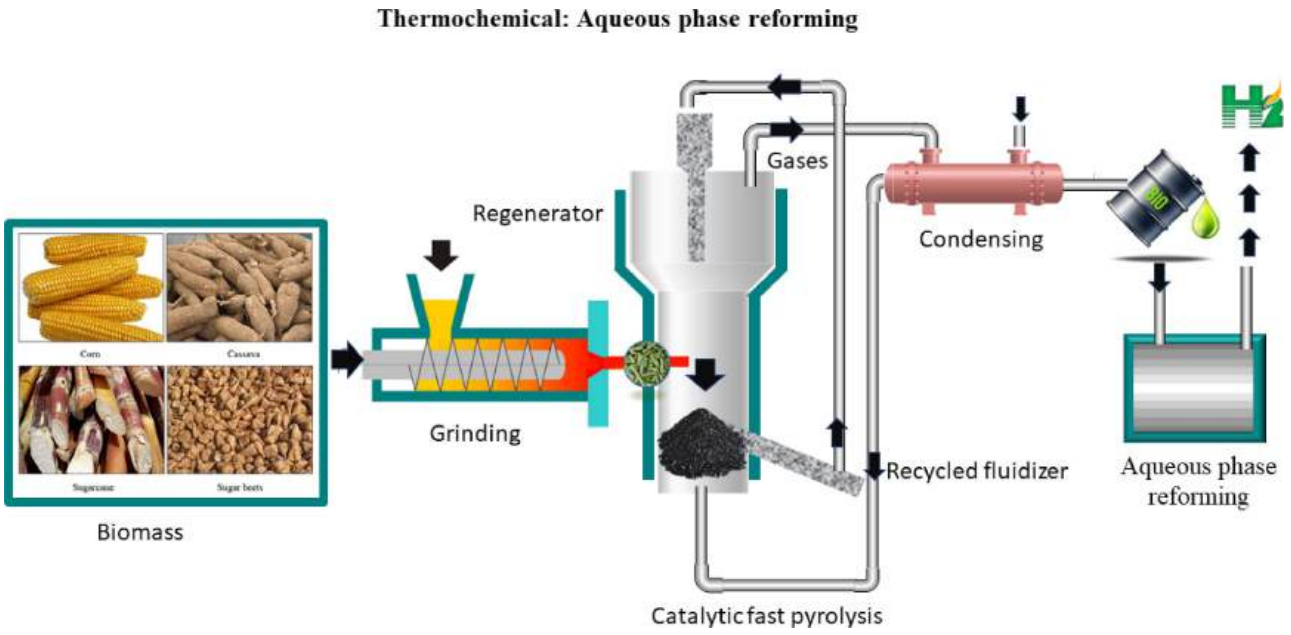


Figure 32.9.15: Schematic diagram for thermochemical (aqueous phase reforming) process. M.G. Eloffy et al.

KEY POINTS - BETA VERSION FROM CHAT.OPENAI

1. Biohydrogen is a form of biofuel that is produced from biomass through a process known as biological hydrogen production.
2. Biological hydrogen production involves the use of microorganisms, such as bacteria and algae, to convert organic matter into hydrogen gas.
3. Biohydrogen has the potential to be a clean, renewable, and sustainable source of energy, as it produces only water when burned and does not produce greenhouse gases.
4. Photosynthetic microorganisms such as algae and cyanobacteria are the most promising organisms for biohydrogen production, they can convert water and CO2 into hydrogen and oxygen through the process of photosynthesis.
5. Fermentative microorganisms such as bacteria and fungi can also be used to produce biohydrogen, they can convert organic materials such as sugars and starches into hydrogen through the process of fermentation.
6. Biohydrogen can be produced through different processes, including dark fermentation, light-driven fermentation, and photo-biological hydrogen production.
7. Dark fermentation is the process of using microorganisms to ferment organic matter in the absence of light to produce hydrogen gas.
8. Light-driven fermentation is the process of using microorganisms to ferment organic matter in the presence of light to produce hydrogen gas.
9. Photo-biological hydrogen production is the process of using algae or photosynthetic bacteria to produce hydrogen gas through photosynthesis.
10. Biohydrogen production is still in the early stages of development and research is ongoing to improve the efficiency and cost-effectiveness of the process.
11. Biohydrogen production from algae is considered more sustainable and environmentally friendly than traditional hydrogen production methods, which are often based on fossil fuels.
12. Biohydrogen has the potential to significantly reduce carbon emissions and to help decarbonize various sectors, including transportation and energy.

32.10: BIOHYDROGEN - HYDROGENASES

Search Fundamentals of Biochemistry

INTRODUCTION

In the last section, we described different ways to produce H_2 and the colors ascribed to them based on the environmental impacts. Many look to the production and use of H_2 to provide energy without releasing CO_2. H_2 can be used in fuel cells to power spacecraft and cars, as shown in the reaction below.

$$\begin{aligned} O_2 + 4H^+ + 4e^- &\longrightarrow 2H_2O \\ H_2 &\longrightarrow 4H^+ + 4e^- \end{aligned} \tag{32.10.1}$$

Cars are already available that run on H_2 considered a zero-emission fuel. These fully electric cars use fuel cells powered by the oxidation of H_2 to produce electrical energy.

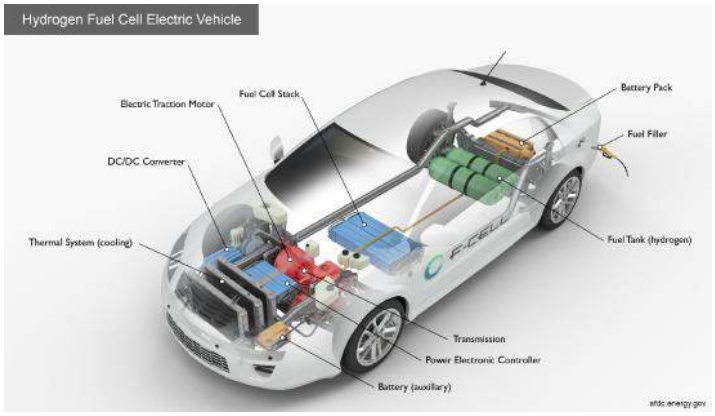


Figure 32.10.1: https://afdc.energy.gov/vehicles/how...tric-cars-work

Given the scale needed, most H_2 is presently derived from the steam reformation of natural gas and the electrolysis of water. From a biochemical perspective, cells have evolved to make H_2 and use H_2 as an energy source. It's unlikely that direct microbial production of H_2 would meet society's energy needs. The 2023 International Energy Agency (IEA) report, "Hydrogen patents for a clean energy future", doesn't mention direct microbial production. However, much can be learned by studying how **hydrogenases (H_2ases, Hyd)**, enzymes that make or use H_2, work. (Don't confuse hydrogenases with dehydrogenases that directly use NAD^+/NADH and FAD/$FADH_2$) for redox reactions. Transition metal active site mimetics can be made as potential catalysts for more industrial-level production of H_2.

Although the reversible formation of H_2 involves the most elemental particles in chemistry, H^+ and e^-, the biological reactions that produce and consume H_2 are complex. Before we proceed, let's see how these reactions are similar to other biochemical reactions and pathways we have already discussed.

Use of H_2 as a *source* of electrons for reduction reactions.

Each hydrogen in H_2 has an oxidation number of 0. Each hydrogen can be oxidized to H^+ (oxidation number +1) with the 2 electrons passed on to a substrate/cofactor or a sequential series of substrates with higher and higher standard reduction potentials (better oxidizing agents), leading to the formation of reduced products.

$$H_2 + (\text{substrate})_{OX} \rightarrow 2H^+ + (\text{product})_{RED}$$

This general reaction is analogous to the mitochondrial electron transport chain, in which electrons are passed from a source (NADH) to oxidized forms of acceptors. The general reaction below shows each redox pair in the electron transport chain.

$$NADH/NAD^+ \rightarrow FAD/FADH_2 \rightarrow UQ/UQH2 \rightarrow \text{Cyto } C_{OX}/\text{Cyto } C_{RED} \rightarrow O_2/H_2O$$

Some organisms have evolved to produce energy by the oxidation of H_2. This is analogous to photosynthetic organisms obtaining energy through the oxidation of water. In photosystem II, oxygen in H_2O (oxidation number -2) gets oxidized by the oxygen-evolving complex to produce O_2 (oxidation number 0). Some redox pairs, starting with H_2O/O_2, are shown below for photosystem II.

$$H_2O/O_2 \rightarrow P680/P680^* \rightarrow (\text{Plastoquione})_{OX}/(\text{Plastoquione})_{RED}$$

The first reaction is endergonic and requires as an energy source photons.

Use of H^+ as a *sink* for electrons for oxidation reactions that produce H_2.

H^+ has an oxidation number of +1. Hence it can be reduced to H_2 (oxidation number of 0) as it gains electrons from substrates/cofactors, which get oxidized. This general reaction is shown below.

$$2H^+ + (\text{substrate/cofactor})RED \rightarrow H_2 + (\text{substrate/cofactor})OX$$

Many microorganisms can produce H_2 through variants of photosynthesis or through fermentation, both of which provide the two electrons needed. E. Coi has four hydrogenases (Hyd 1, 2, 3, and 4). It forms H_2 through two reactions catalyzed by:

- formate (HCO_2^-) dehydrogenase (FDH): $2HCO_2^- \rightleftharpoons 2CO_2 + 2H^+ + 2e^-$
- hydrogenase (H_2ase): $2H^+ + 2e^- \rightarrow H_2$

The C in formate has an oxidation number of +2 and is oxidized to CO_2, in which the C has an oxidation number of +4.

The formate hydrogenlyase (FHL) complex contains both the formate dehydrogenase (FDH) and a hydrogenase (H_2ase) and reversibly interconverts HCO_2^- and H_2. The *E. coli* FHL-1 complex, which makes H_2 using fermentation, is shown below in Figure 32.10.2. The complex can be immobilized on a Macro-mesoporous inverse opal (IO) indium tin oxide (ITO) electrode (IO-IPO) or ITO nanoparticles (NP), which can relay electrons.

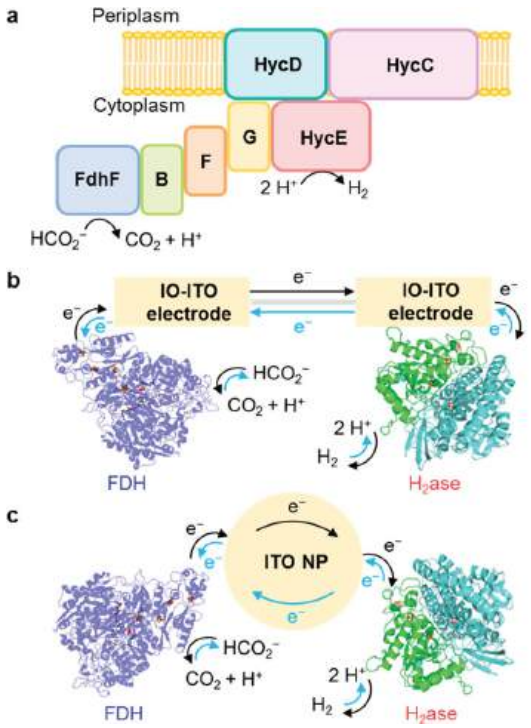


Figure 32.10.2: Katarzyna P. Sokol et al. J. *Am. Chem. Soc.* 2019, 141, 44, 17498–17502. https://doi.org/10.1021/jacs.9b09575. **CC-BY license**

Panel (a) shows the biological *E. coli* FHL-1 complex. FdhF, [Mo]-FDH; B/F/G, Fe–S cluster-containing proteins; HycE, [NiFe]-H_2ase; HycD/C, membrane proteins. (17)

Panel (b) shows a IO-ITO|FDH||IO-ITO|H_2ase cell: IO-ITO|FDH wired to IO-ITO|H_2ase electrode.

Panel (c) shows a FDH–ITO–H_2ase nanoparticle (NP) system with enzymes immobilized onto ITO NP in solution. Species size not drawn to scale.

All you need to synthesize H_2 are 2 protons and 2 electrons (potentially derived from photosynthesis). Let's take a deeper look at the hydrogenase that catalyzes H_2 production.

HYDROGENASES (H_2ASES)

Hydrogenases catalyze the reversible conversion of $2H^+ \rightarrow H_2$. A hydrogenase database, HydDB, a web tool for hydrogenase classification and analysis of sequence, shows their high diversity and metabolic roles. There are three classes of hydrogenases, the **Ni-Fe** (most abundant, primarily for H_2 conversion to $2H^+$), the **Fe-Fe** (highest k_{cat} for H_2 production), and the **single Fe** hydrogenases, as shown in Table 32.10.2 below. We won't discuss the single Fe hydrogenases.

CLASSES AND SUBCLASSES OF HYDROGENASES

[NiFe] Group 1: Respiratory H_2-uptake [NiFe]-hydrogenases		
1a	Periplasmic	Electron input for sulfate, metal and organohalide respiration. [NiFeSe] variants.
1b	Prototypical	Electron input for sulfate, fumarate, metal and nitrate respiration.
1c	Hyb-type	Electron input for fumarate, nitrate and sulfate respiration. Physiologically reversible.
1d	Oxygen-tolerant	Electron input for aerobic respiration and oxygen-tolerant anaerobic respiration.
1e	Isp-type	Electron input primarily for sulfur respiration. Physiologically reversible.
1f	Oxygen-protecting	Unresolved role. May liberate electrons to reduce reactive oxygen species.
1g	Crenarchaeota-type	Electron input primarily for sulfur respiration.
1h	Actinobacteria-type	Electron input for aerobic respiration. Scavenges electrons from atmospheric H_2.
1i	Coriobacteria-type (putative)	Undetermined role. May liberate electrons for anaerobic respiration.
1j	Archaeoglobin-type	Electron input for sulfate respiration[π].
1k	Methanophenazine-reducing	Electron input for methanogenic heterodisulfide respiration.
[NiFe] Group 2: Alternative and sensory uptake [NiFe]-hydrogenases		
2a	Cyanobacteria-type	Electron input for aerobic respiration. Recycles H_2 produced by other cellular processes.
2b	Histidine kinase-linked	H_2 sensing. Activates two-component system controlling hydrogenase expression.
2c	Diguanylate cyclase-linked (putative)	Undetermined role. May sense H_2 and regulate processes through cyclic di-GMP production.
2d	Aquificae-type	Unresolved role. May generate reductant for carbon fixation or have a regulatory role.
2e	Metallosphaera-type (putative)	Undetermined role. May liberate electrons primarily for aerobic respiration.
[NiFe] Group 3: Cofactor-coupled bidirectional [NiFe]-hydrogenases		
3a	F_{420}-coupled	Couples oxidation of H_2 to the reduction of F_{420} during methanogenesis. Physiologically reversible. [NiFeSe] variants.
3b	NADP-coupled	Couples oxidation of NADPH to the evolution of H_2. Physiologically reversible. May have sulfhydrogenase activity.

3c	Heterodisulfide reductase-linked	Bifurcates electrons from H_2 to heterodisulfide and Fd_{ox} in methanogens. [NiFeSe] variants.
3d	NAD-coupled	Interconverts electrons between H_2 and NAD depending on cellular redox state.
	[NiFe] Group 4: Respiratory H_2-evolving [NiFe]-hydrogenases	
4a	Formate hydrogenlyase	Couples formate oxidation to fermentative H_2 evolution. May be H^+-translocating.
4b	Formate-respiring	Respires formate or carbon monoxide using H^+ as electron acceptor. Na^+-translocating via Mrp.
4c	Carbon monoxide-respiring	Respires carbon monoxide using H^+ as electron acceptor. H^+-translocating.
4d	Ferredoxin-coupled, Mrp-linked	Couples Fd_{red} oxidation to H^+ reduction. Na^+-translocating via Mrp complex.
4e	Ferredoxin-coupled, Ech-type	Couples Fd_{red} oxidation to H^+ reduction. Physiologically reversible via H^+/Na^+ translocation.
4f	Formate-coupled (putative)	Undetermined role. May couple formate oxidation to H_2 evolution and H^+ translocation.
4g	Ferredoxin-coupled (putative)	Undetermined role. May couple Fd_{red} oxidation to proton reduction and H^+/Na^+ translocation.
4h	Ferredoxin-coupled, Eha-type	Couples Fd_{red} oxidation to H^+ reduction in anaplerotic processes. H^+/Na^+-translocating.
4i	Ferredoxin-coupled, Ehb-type	Couples Fd_{red} oxidation to H^+ reduction in anabolic processes. H^+/Na^+-translocating.
	[FeFe] Hydrogenases	
A1	Prototypical	Couples ferredoxin oxidation to fermentative or photobiological H_2 evolution.
A2	Glutamate synthase-linked (putative)	Undetermined role. May couple H_2 oxidation to NAD reduction, generating reductant for glutamate synthase.
A3	Bifurcating	Reversibly bifurcates electrons from H_2 to NAD and Fd_{ox} in anaerobic bacteria.
A4	Formate dehydrogenase-linked	Couples formate oxidation to H_2 evolution. Some bifurcate electrons from H_2 to ferredoxin and NADP.
B	Colonic-type (putative)	Undetermined role. May couple Fd_{red} oxidation to fermentative H_2 evolution.
C1	Histidine kinase-linked (putative)	Undetermined role. May sense H_2 and regulate processes via histidine kinases.
C2	Chemotactic (putative)	Undetermined role. May sense H_2 and regulate processes via methyl-accepting chemotaxis proteins.
C3	Phosphatase-linked (putative)	Undetermined role. May sense H_2 and regulate processes via serine/threonine phosphatases.
	[Fe] Hydrogenases	

All	Methenyl-H_4MPT dehydrogenase	Reversibly couples H_2 oxidation to 5,10-methenyltetrahydromethanopterin reduction.

The three main types have different main functions *in general*. The Ni-Fe, Fe-Fe, and Fe H2ases generally oxidize H_2, produce H_2 and promote H^- (hydride) transfer, respectively, as shown in Figure 32.10.3 below.

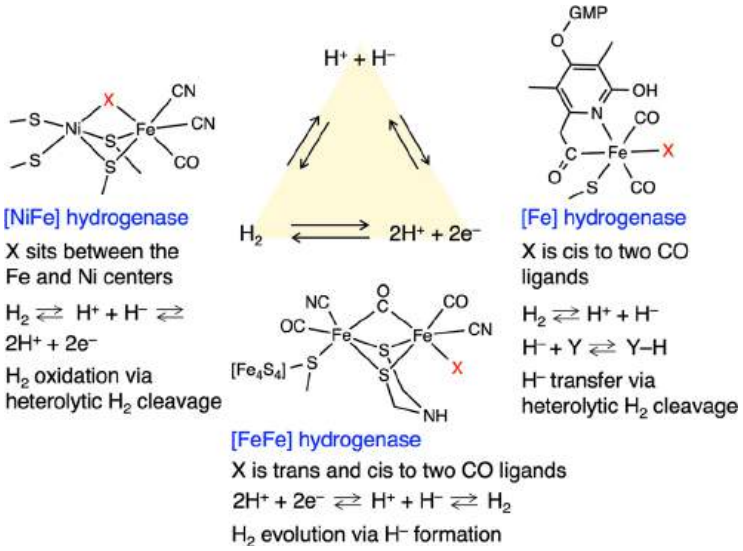


Figure 32.10.3: The active site structures of [NiFe] H_2ases that mainly catalyze H_2 oxidation reactions, [FeFe] H_2ases that mainly catalyze H_2 evolution reactions, and [Fe] H_2ases that catalyze H^- transfer to the substrate via heterolytic H2cleavage. X, possible H_2 active sites; Y, methenyltetrahydromethanopterin; GMP, guanosine monophosphate. Seiji Ogo et al., Science Advances. (2020). DOI: 10.1126/sciadv.aaz81. Creative Commons Attribution-NonCommercial License 4.0 (CC BY-NC).

Much effort has been devoted to making transition state analogs of the active site to act as catalysts for H_2 production for fuel cells. Transition metal catalysts that mimic the structures and activities of the three hydrogenases have been made. Three specific ones are shown below in Figure 32.10.4.

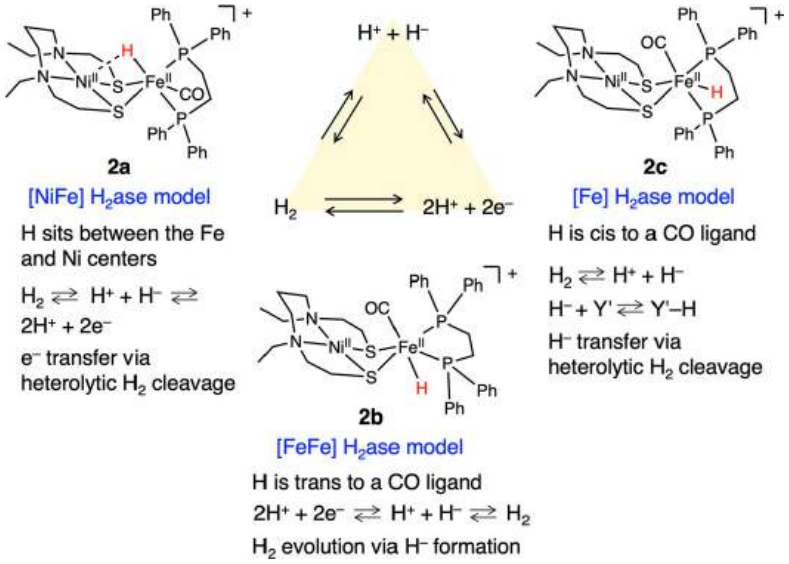


Figure 32.10.4: The differing reactivity of the three isomers. Y′, methylene blue $[MB]^+$.Seiji Ogo et al., ibid.

The ligand containing P and PH is bis(diphenylphosphino)ethane.

First, we will explore the Ni-Fe hydrogenases.

NI-FE H_2ASES (HYD):

We'll discuss two examples of Ni-Fe H_2ases

Group 1a periplasmic (membrane-bound) hydrogenases - MBH

These are used in fuel cells and H_2-producing devices since they can adhere to surfaces that can be useful heterogeneous (not in solution) catalysts. Some also are damaged by O_2. The enzyme consists of a large subunit) found in the periplasm and small subunit, which anchors the protein in the plasma membrane of bacteria. . This enzyme oxidizes H_2: $H_2 \rightarrow 2H^+ + 2e^-$. The electrons enter the bacterial respiratory chain through quinones. The transmembrane part of the small subunit binds cytochrome b, which is involved in electron transfer with the quinones, as we saw in Complex II of mitochondrial electron transport. Some soil bacteria (like *Ralstonia eutropha*,) can use H_2 as their sole energy source. The orientation of a NiFe MBH within a bacterial cell is shown in Figure 32.10.5 below.

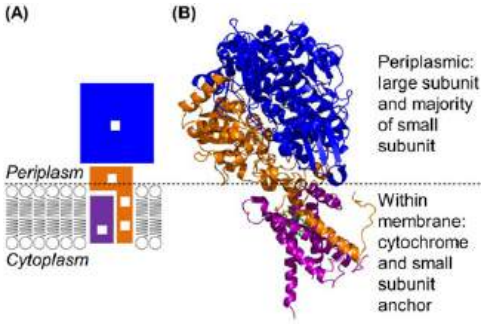


Figure 32.10.5: The orientation of a NiFe MBH within a bacterial cell. Lindsey A. Flanagan* and Alison Parkin. Biochem Soc Trans. 2016 Feb 15; 44(1): 315–328 (2016). doi: 10.1042/BST20150201 Creative Commons Attribution Licence 3.0.

Panel (**A**) shows a cartoon depiction of how a NiFe MBH is located within the cytoplasmic membrane, with white boxes representing the redox active metal centers and blue, orange and purple blocks indicating the large, small and cytochrome subunits, respectively.

Panel (**B**) shows how the *E. coli* hydrogenase-1 large (blue ribbon), small (orange ribbon) and cytochrome (purple ribbon) subunits can interact.

Figure 32.10.6 shows an interactive iCn3D model of the O_2-Tolerant Membrane-Bound Hydrogenase 1 from Escherichia coli in Complex with Its Cognate Cytochrome b (4GD3). The same color coding is used for the subunits as in the above figures.

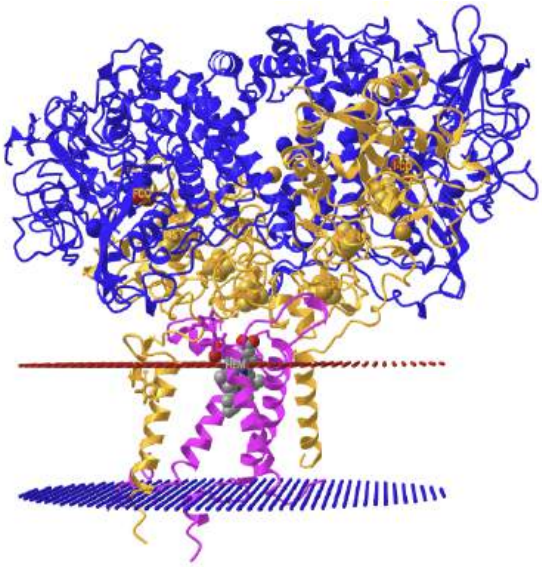


Figure 32.10.6: O_2-Tolerant Membrane-Bound Hydrogenase 1 from Escherichia coli in Complex with Its Cognate Cytochrome b (4GD3). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...9p4CSQRDPM8nL7

The large and small chains of hydrogenase are shown in blue and orange, respectively, while cytochrome b is sown in magenta. Cofactors and key metals are shown in spacefill. F4S is the Fe4-S3 cluster, SF4 is the iron-sulfur cluster, HEM is hemoglobin, FCO is carbonmonoxide-(dicyano) Fe, and Ni is nickel.

The biological (functional) unit consists of two heterodimers. Under low O_2 levels, one dimer can reactivate the other if exposed to O_2. The enzyme is found in the highest concentration during anaerobic fermentation. Remember that E. Coli is a facultative anaerobe and can shift its metabolic pathways to fit conditions. Perhaps its primary role is to reduce O_2 to water and protect enzymes sensitive to it. The function of Cytochrome b may be mostly to anchor the dimeric H_2ase in the membrane.

A bifurcating Ni-Fe H_2ase

These enzymes are more complicated. They oxidize H_2 and move the two electrons through a complicated path that bifurcates electron flow to different substrates/cofactors. They move an electron to a low-potential (i.e. not a great oxidizing agent), high-energy species, which gets reduced in an endergonic process. The other electron simultaneously moves to a high-potential (i.e., great oxidizing agent), low-energy species that also gets reduced in an exergonic process. The overall electron transfer is thermodynamically favorable. An example might prove helpful. NADH ($E^{0'}$ = -280 mV, higher potential) can reduce the protein ferredoxin ($E^{0'}$ = -500 mV, lower potential), which can then pass its electrons in other reactions, including the formation of H_2, CH_4, and NH_3. ATP is not required.

BIFURCATING ENZYMES - WE'VE SEEN ONE BEFORE!

Four classes of bifurcating enzymes that use $FAD/FADH_2$ or $FMN/FMNH_2$ are known. They are optimal since they can participate in either 1 or 2 electron transfers. We will see an example of a Fe-Fe H_2ase further below.

In electron transport, we encountered an electron bifurcating complex in the Q-cycle of Complex III. Reduced ubiquinone (UQH_2, or ubiquinol) is oxidized and the two lost electrons bifurcated to cytochrome C in a high potential pathway and to UQ to reform UQH2, as shown in Figure 32.10.7 below.

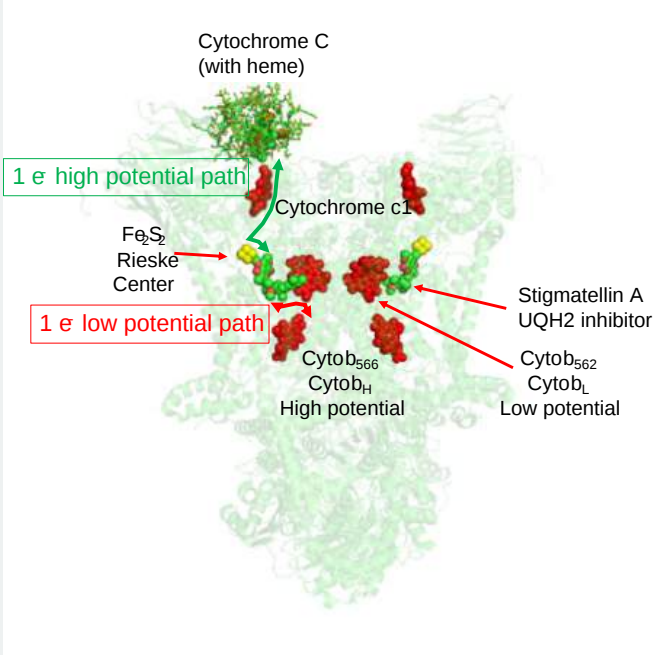


Figure 32.10.7: Electron bifurcation in Complex III

When ubiquinone, UQH "Electron bifurcation was first described in the Q-cycle of the respiratory complex III where the two electrons originating from ubiquinol oxidation are bifurcated via a *high*-potential pathway to cytochrome c, and via a *low*-potential path to reduce ubiquinone to ubiquinol.

One example of a bifurcating Ni-Fe H_2ase is the NiFe-HydABCSL protein from the bacteria *A. mobile*. The general structure of the pentameric form of the functional decamer is shown in Figure 32.10.8 below.

Figure 32.10.8: Structure of the *A. mobile* NiFe-HydABCSL pentamer. XIANG FENG et al. ***SCIENCE ADVANCES.*** 2022. DOI: 10.1126/sciadv.abm7546. Creative Commons Attribution License 4.0 (CC BY). https://creativecommons.org/licenses/by/4.0/

The five subunits are called HydA (Hyd = hydrogenase), HydB, HydC, HydL (large subunit), and HydS (small subunit). Pane (**A**) shows the domain structure of the five subunits. The NiFe-HydB NTD and CTDs are partially flexible, as indicated by dashed outlines. Panel (**B**) shows the subunit organization of the NiFe-Hyd complex and their associated cofactors.

NiFe-HydABCSL hydrogenase can reversibly oxidize H_2 with the two electrons reducing ferredoxin in an endergonic process and reducing NAD in an exergonic process. FMN is surrounded by an FeS cluster and appears to be the center of bifurcation. The reaction is as follows:

- The HydL oxidizes H_2 with two electrons passing through the FeS centers in HydA to HybB.
- The electrons are passed to FMN, where the bound NAD gets reduced.

Figure 32.10.9 jdkfjdkjfdkjff

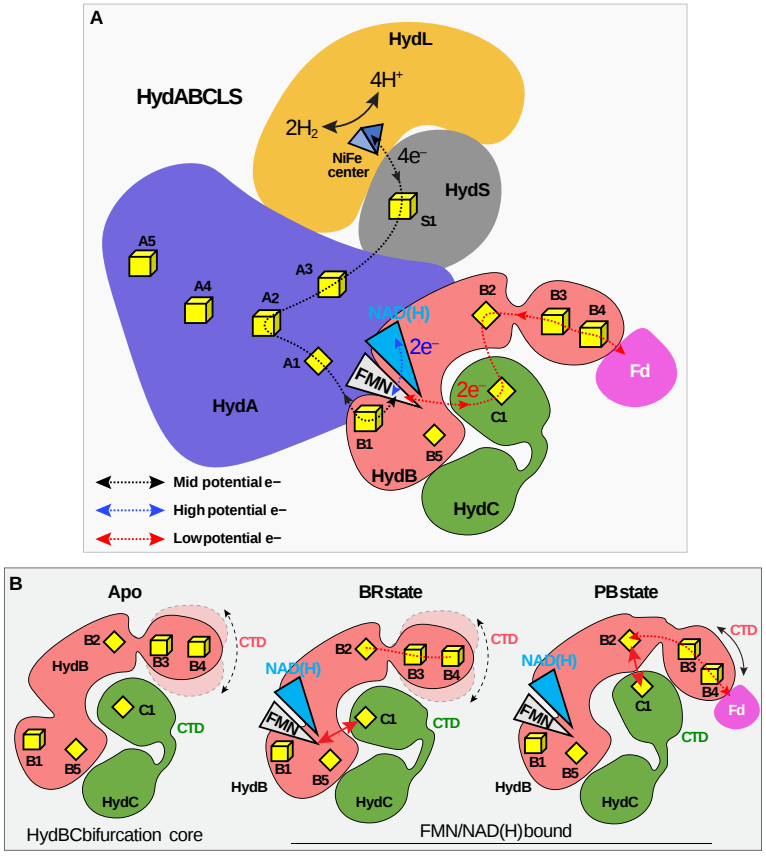


Figure 32.10.9: Proposed mechanism of electron bifurcation/confurcation in *A. mobile* NiFe-HydABCSL.

(**A**) Overall electron transfer pathway, highlighting the three branches of the electron transfer path. The mid-potential path is a black dashed line, the exergonic path is a blue dashed line, and the endergonic path is a red dashed line. (**B**) Conformational changes in the HydBC bifurcation core from the electron bifurcation state (BR state) to the electron transduction state (PB state).

Figure 32.10.10 shows an interactive iCn3D model of the electron bifurcating Ni-Fe hydrogenase complex HydABCSL in FMN/NAD(H) bound state **7T30**

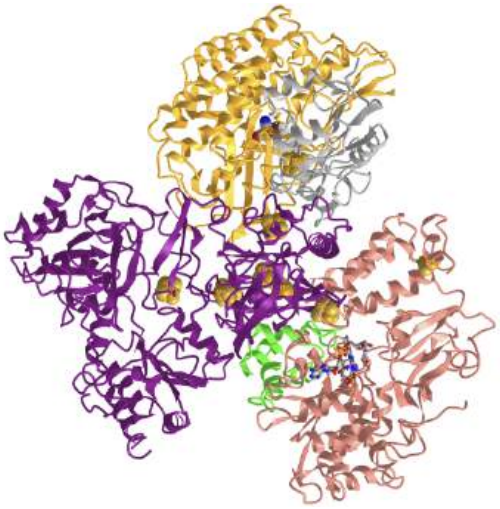


Figure 32.10.10 : Electron bifurcating Ni-Fe hydrogenase complex HydABCSL in FMN-NAD(H) bound state **(7T30)**. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...G9nKGwqvuuh6G7

Figure 32.10.11 shows an interactive iCn3D model of the cofactors in the electron bifurcating Ni-Fe hydrogenase complex HydABCSL in FMN/NAD(H) bound state **7T30**

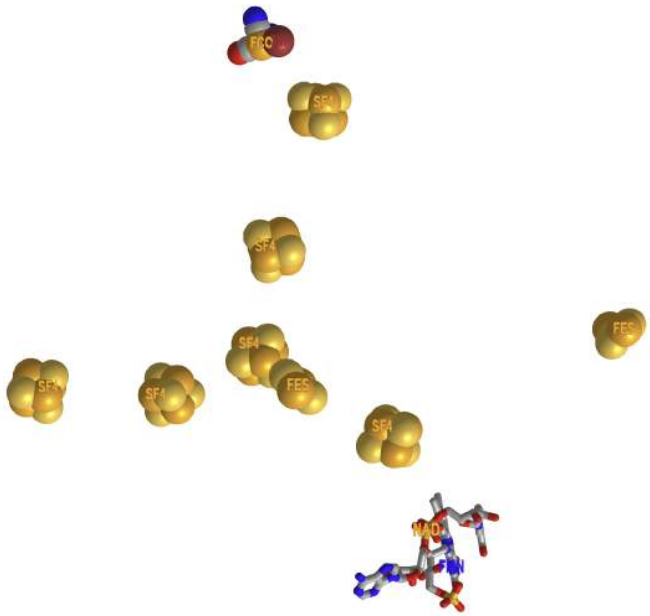


Figure 32.10.11 : Cofactors in the electron bifurcating Ni-Fe hydrogenase complex HydABCSL in FMN-NAD(H) bound state **(7T30)**. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...hHuzaWUDYCMnAA.

Zoom into the Ni-Fe center catalytic site. The ligands that form coordinate covalent bonds to the Fe are called FCC, or carbon monoxide-(dicyano)-Fe Figure 32.10.12 below. There are also bridging sulfurs between Fe and Ni.

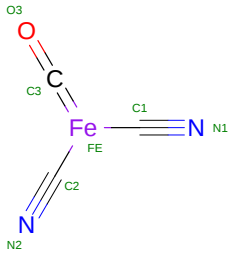


Figure 32.10.12 : Carbonmonoxo-dicyano-Fe and is shown in detail in

FE-FE HYDROGENASES

These enzymes catalyze a variety of reactions as illustrated in Figure 32.10.13 below.

Figure 32.10.13 : [FeFe]-hydrogenases phylogeny and known functions. Morra S. Front Microbiol. 2022 Mar 2;13:853626. doi: 10.3389/fmicb.2022.853626. PMID: 35308355; PMCID: PMC8924675. Creative Commons Attribution License (CC BY)

A phylogenetic tree shows the phylogeny of [FeFe]-hydrogenase sequences from public databases, as previously proposed. Enzymes that have been experimentally characterized are indicated on the tree to show their relative position. The proposed physiological function of each enzyme is also presented, where known. Hyd, hydrogenase subunit; FdhF, formate dehydrogenase subunit; $Fd_{rex/ox}$, reduced/oxidized ferredoxin; $NADH/NAD^+$, reduced/oxidized nicotinamide adenine dinucleotide. They are found in prokaryotic and eukaryotic microorganisms, but not in Archaea.

These are the most active for H_2 production with a k_{cat} around 10,000 s^{-1}. They contain a Fe_2S_2 cluster with CO and CN ligands forming bonds to the iron with the iron ions bridged by a $-SCH_2-NH-CH_2S-$ (aza-dithiolate). A cysteine links the Fe2S2 to a Fe_4S_4 cluster. These two are called the **H-cluster (or $[Fe]_H$**. Within this class are a soluble, monomeric cytoplasmic form, a heterodimeric periplasmic form, and a soluble, monomeric chloroplastic form. This one has a ferredoxin, connecting it to the electron transport chain in photosynthesis. Some in this group using both NADH and ferredoxin are called bifurcating types, as they send two electrons from a donor in two different directions. More on this later.

They contain multiple FeS clusters. The H-cluster consists of a Fe_2S_2 linked to a Fe_4S_4 cluster (cubane-like) by a cysteine. The Fe_2S_2 group has CO and CN ligands, and the two Fe ions of Fe_2S_2 unit are coordinated by an azadithiolato ligand, as shown below in Figure 32.10.14 .

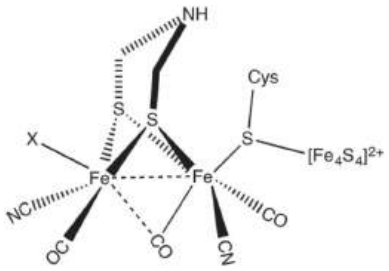


Figure 32.10.14 : Chemical structure of the H-cluster, which is the active site of the [FeFe] hydrogenase enzyme. Rakesh C. Puthenkalathil et al., Phys. Chem. Chem. Phys., 2020, 22, 10447. https://pubs.rsc.org/fa/content/arti.../cp/c9cp06770a. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence

The Fe in the [2Fe2S] cluster is linked to the cubane [4Fe-4S] and has six ligands, so it is saturated. The other Fe has an extra coordination site denoted by X, which can bind H^+ or H_2. The cluster is buried in a hydrophobic catalytic site which helps restrict O_2 access.

As we did for the Fe-Ni H_2ases, we will study two examples of Fe-Fe H_2ases.

Fe-Fe hydrogenase (CpI) from Clostridium pasteurianum

Figure 32.10.15 shows an interactive iCn3D model of the H-Cluster (HC1) of Fe-Fe hydrogenase (CpI) from Clostridium pasteurianum (1FEH).

NCBI iCn3D Figure 32.10.15 : H-Cluster (HC1) of Fe-Fe hydrogenase (CpI) from Clostridium pasteurianum (1FEH). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...TksW8FGn7crh39

In this model, the CN ligands are all displayed as CO. The sulfurs are shown in green. Hover over the atoms/ions to identify them. (In iCn3D, choose, **Select, Select on 3D, atom**). The [4Fe-4S] subcluster forms coordinate covalent bonds with four cysteines (300, 355, 499, and 503) with one cysteine (503) forming a bridge to the [2Fe] cluster. The Fe ions in that cluster have an octahedral arrangement of ligands surrounding them. One of the ligands is water (no connecting C atom).

Figure 32.10.16 shows an interactive iCn3D model of the Fe-Fe hydrogenase (CpI) from Clostridium pasteurianum (1FEH)

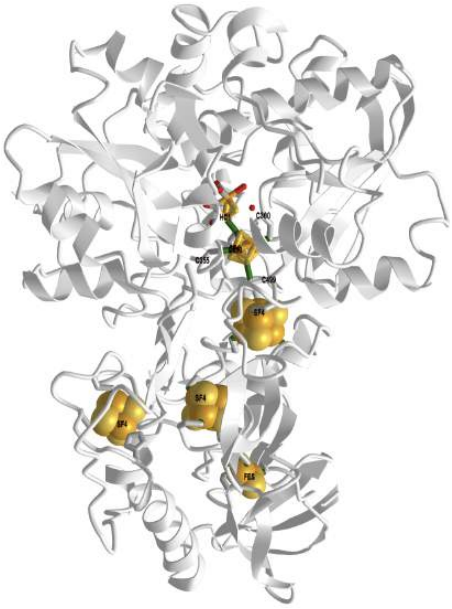


Figure 32.10.16: Fe-Fe hydrogenase (CpI) from Clostridium pasteurianum (1FEH). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...QJsoZ8zbvkpDc7.

As we mentioned above, the net reaction is:

$$2H^+ + (\text{substrate/cofactor})RED \rightarrow H_2 + (\text{substrate/cofactor})OX$$

Many microorganisms can produce H_2 through variants of photosynthesis or through fermentation, both of which provide the two electrons needed. E. Coi has four different hydrogenases, (Hyd 1, 2, 3 and 4). It forms H_2 through two reactions catalyzed by:

- formate (HCO_2^-) dehydrogenase: $2HCO_2^- \rightleftharpoons 2CO_2 + 2H^+ + 2e^-$
- hydrogenase 3: $2H^+ + 2e^- \rightarrow H_2$

Figure 32.10.17 below shows a reaction scheme for the production of H_2 linked to photosystem I in the chloroplast of microalgae under anaerobic conditions. It starts with absorption of a photon by P700, which in the excited state transfers an electron to a 4Fe4S cluster in ferredoxin, which then passes an electron to the HC-cluster and then onto H^+.

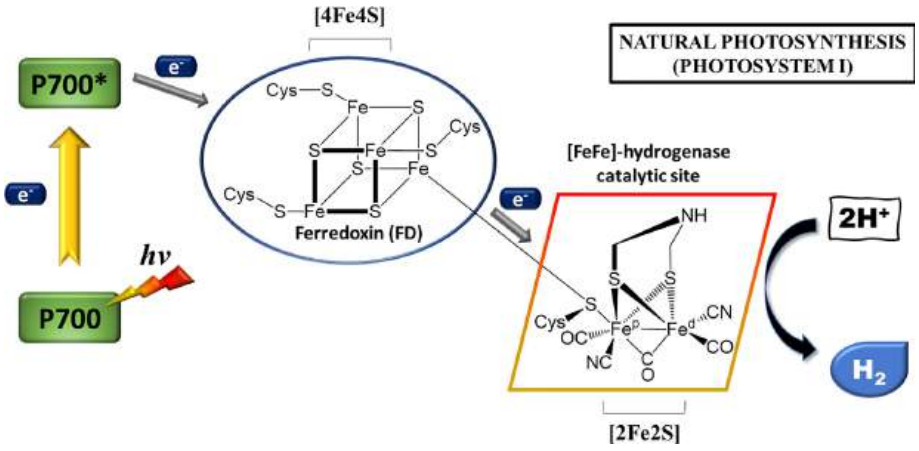


Figure 32.10.17: Schematic representation of electron flow from Photosystem I to an [FeFe]-hydrogenase via a ferredoxin redox mediator (Photosystem I). JuanAmaro-Gahete et al.,Coordination Chemistry Reviews. 448, December 2021. https://doi.org/10.1016/j.ccr.2021.214172. Creative Commons CC-BY

A possible mechanism for the formation of H_2 in the H clusteris shown below in Figure 32.10.18 below.

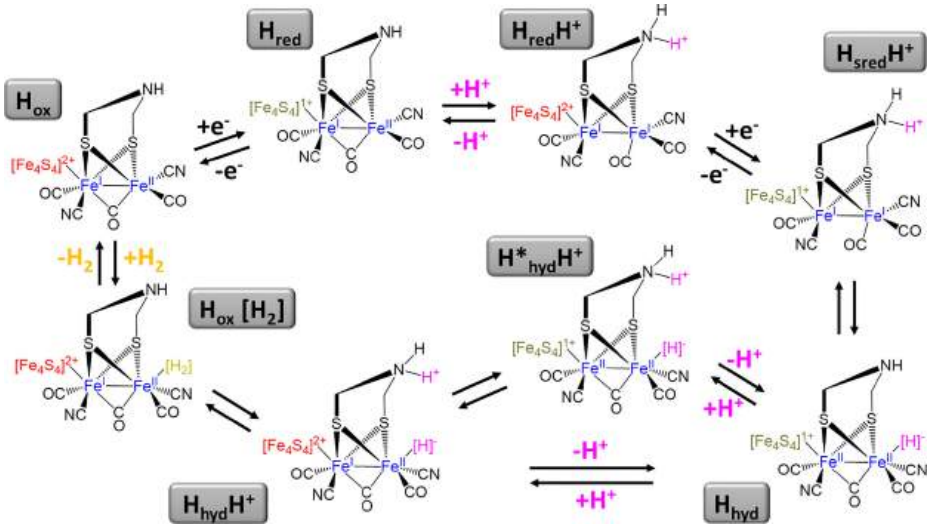


Figure 32.10.18: Proposed mechanistic cycle for hydrogen evolution in the H cluster by [FeFe]-hydrogenase adapted from Lubitz et al. JuanAmaro-Gahete et al., ibid

Start at the top left which shows the resting oxidized state. In the enzyme's most oxidized resting system (H_{ox}), the [4Fe4S] cubane is in a 2 + oxidation state while the catalytic subcluster [2Fe] is a mixed-valence $Fe^{I}Fe^{II}$ state. The first one-electron reduction results in the formation of the H_{red} state, where the [4Fe4S] subcluster is reduced to a 1 + oxidation state. Protonation of the N of aza-propane-1,3-dithiolate ligand (adt-*N*) triggers an intramolecular charge shift to form $H_{red}H^{+}$ in which the [4Fe4S] cubane is in the 2 + state and the [2Fe] subsite reduced to a homovalent $Fe^{I}Fe^{I}$ state. Subsequent one-electron reduction of the subcluster [4Fe4S] gives rise to the "super-reduced" state $H_{sred}H^{+}$. In the next step of the catalytic cycle, an intermediate hydride state [H_{hyd}] is formed by an intramolecular proton shift from the adt-*N* to the distal iron F_d. This process is coupled to an electron rearrangement in the [2Fe] subsite, leading to a formal $Fe^{II}Fe^{II}$ oxidation state. Addition of a second proton coupled to another charge shift from the reduced [4Fe4S] to the [2Fe] subsite either in one or two discrete steps gives rise to [$H_{hyd}H^{+}$] that is characterized by a formal $Fe^{I}Fe^{II}$ oxidation state. At this point, there is an equilibrium between the $H_{hyd}H^{+}$ and $H_{ox}[H_2]$ in which the hydride and the proton are combined into a hydrogen molecule at the distal iron of the system. The catalytic cycle is closed by H_2 release, returning to the initial H_{ox} configuration.

A bifurcating [Fe-Fe] hydrogenase from *Thermotoga maritima* (HydABC)

This enzyme, functionally a heterododecamer, uses NADH as a source of electrons, which passes electrons to FMN, the bifurcation site, with an electron going to oxidized ferredoxin (Fd_{ox}) and another to H^{+}s for reduction to FD_{RED} and H_2. The enzyme consists of a dimer of a trimer of subunits HydA, HydB, and HydC, with dimer $(HydABC)_2$ interacting with another $(HydABC)_2$ to form a heterododecamer, with both halves acting independently. The two trimers (HydABC) in the dimer $(HydABC)_2$ are connected by a [4Fe–4S] cluster. A flexible loop in the B and A chain has a "closed" and "open" bridge conformation with a nearby Zn^{2+} important in the loop conformation.

Figure 32.10.19 below shows the cryo-EM structure of the HydABC tetramer and the arrangement of the redox cofactors.

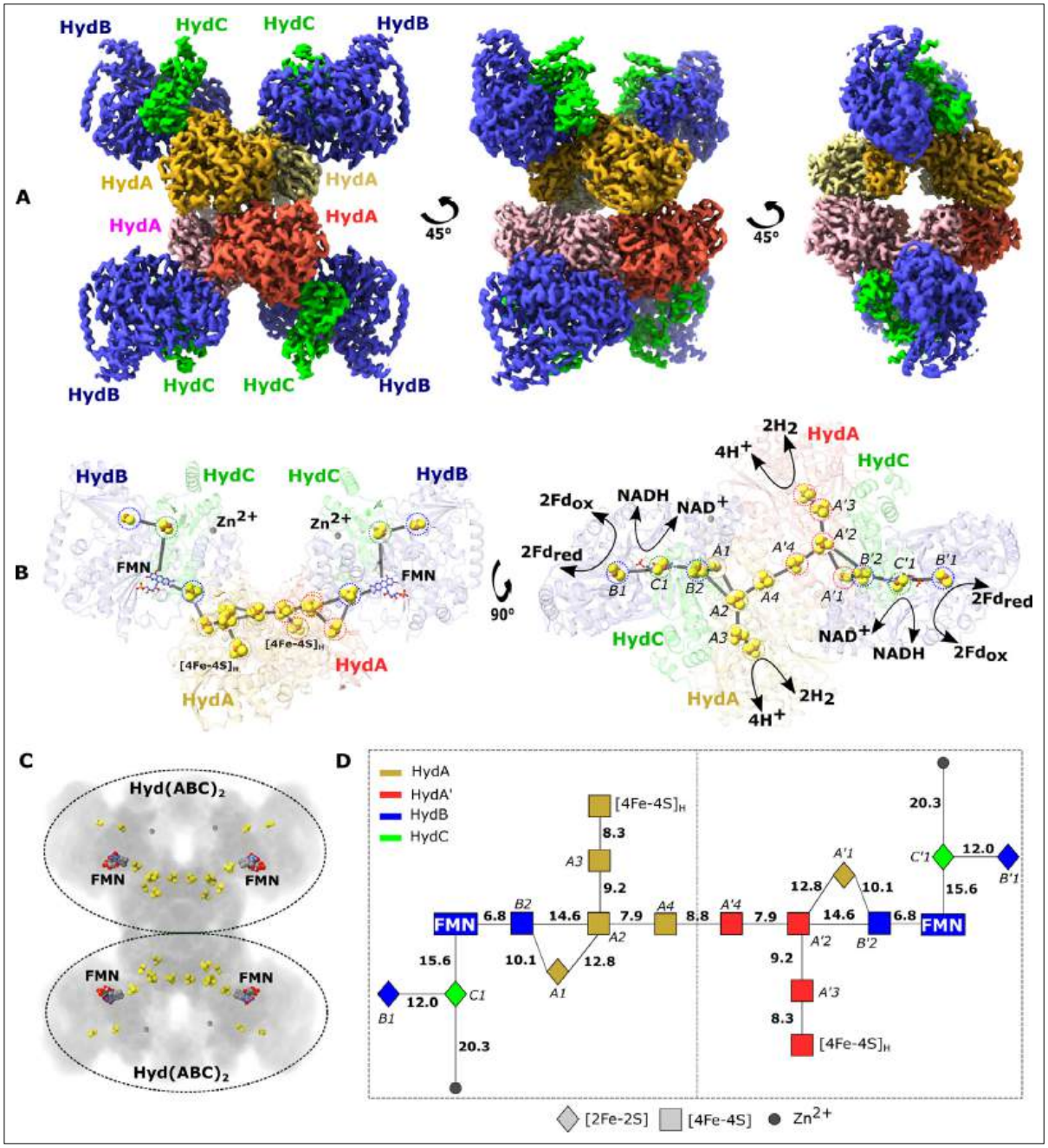


Figure 32.10.19 : Cryo-EM structure of the HydABC tetramer and arrangement of the redox cofactors. Chris Furlan et al. (2022) eLife 11:e79361. https://doi.org/10.7554/eLife.79361. Creative Commons Attribution License

Panel (**A**) shows the unsharpened 2.3 Å map of Hyd(ABC)$_4$ with D2 symmetry enforced, showing a tetramer of HydABC heterotrimers. All four copies of HydB and C are colored blue and green, respectively. The four HydA copies that make up the core of the complex are orange, yellow, pink, and red. The top and bottom halves of the complex are constituted by dimers of HydABC protomers (each HydABC unit is a protomer); the two protomers within the same dimer are strongly interacting, while a weaker interaction is present between the top and bottom dimers.

Panel (**B**) shows the HydABC dimer highlighting the iron–sulfur clusters and flavin mononucleotide (FMN) constituting the electron transfer network.

Panel (**C**) shows the arrangement of redox cofactors within the protein complex, showing two independent identical redox networks (dashed circles); each redox network is composed of iron–sulfur clusters belonging to a Hyd(ABC)$_2$ unit consisting of two strongly interacting HydABC protomers.

Panel (**D**) shows a schematic of the electron transfer network of one of the two identical $Hyd(ABC)_2$ units showing edge-to-edge distances (in Å) between the various cofactors. Note that our structure is of apo-HydABC and contains only the $[4Fe–4S]_H$ subcluster of the H-cluster. The $2H^+/H_2$ interconversion reaction in (**B**) illustrates the site at which this reaction occurs, but this will only occur in the fully assembled H-cluster, including $[2Fe]_H$.

Figure 32.10.20 shows an interactive iCn3D model of the electron-bifurcating [FeFe] hydrogenase from *Thermotoga maritima* (HydABC) (7P5H), using the same colors as the figure above.

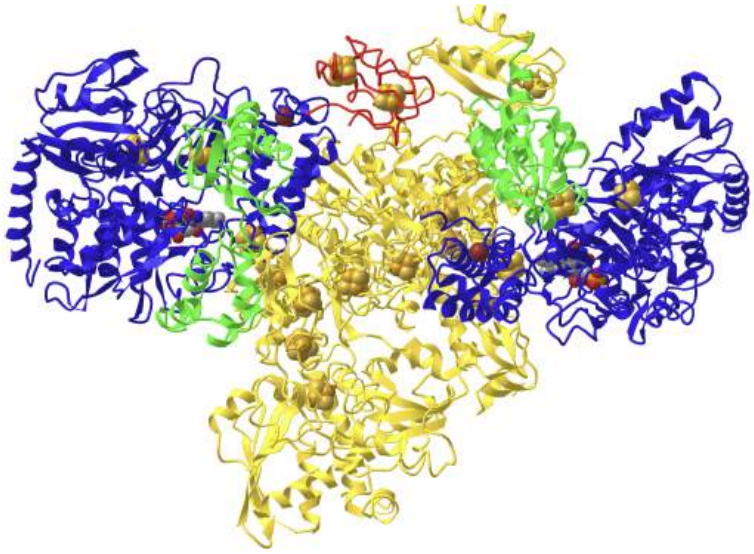


Figure 32.10.20 : electron-bifurcating [FeFe] hydrogenase from *Thermotoga maritima* (HydABC) (**7P5H)**. (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...2hwCUnJ9BgwJi8.

Only the dimer $(HydABC)_2$ is shown. The A (gold), B (blue), and C (green) chains are colored as in the previous figure. The conformationally flexible loop at the C-terminal of a B chain in the closed state is shown in red. The gate also includes the C-terminal part of the A subunit near it.

Figure 32.10.21 shows the closed-bridge and open-bridge conformations of HydABC (the closed loop was shown in the model above).

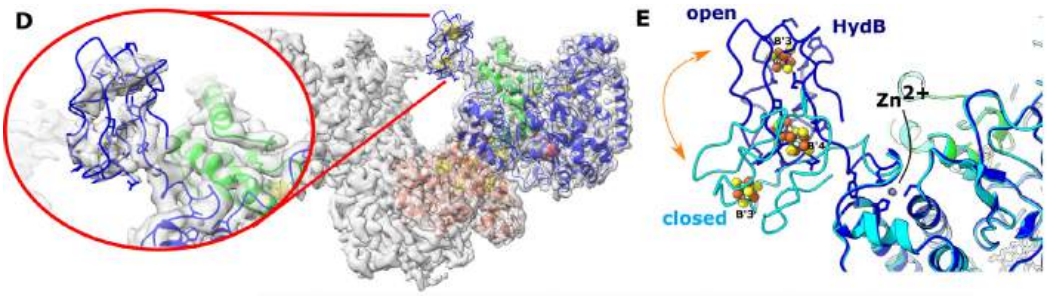


Figure 32.10.21 : closed-bridge and open-bridge conformations of HydABC from *Thermotoga maritime*.

Panel (**D**) shows the HydB bridge domain in the open position and its fitted model. Panel (**E**) shows a Zn^{2+} hinge region and the two possible conformations of the HydB bridge domain, open (blue) and closed (light blue).

The similarities in cofactor arrangement in the *Thermotoga maritime* Hyd A, B and C subunits compared to the Nqo1, Nqo2, and Nquo3 subunits in Complex I from *Thermus thermophilus* (discussed in Chapter 19.1) are shown in Figure 32.10.22 below.

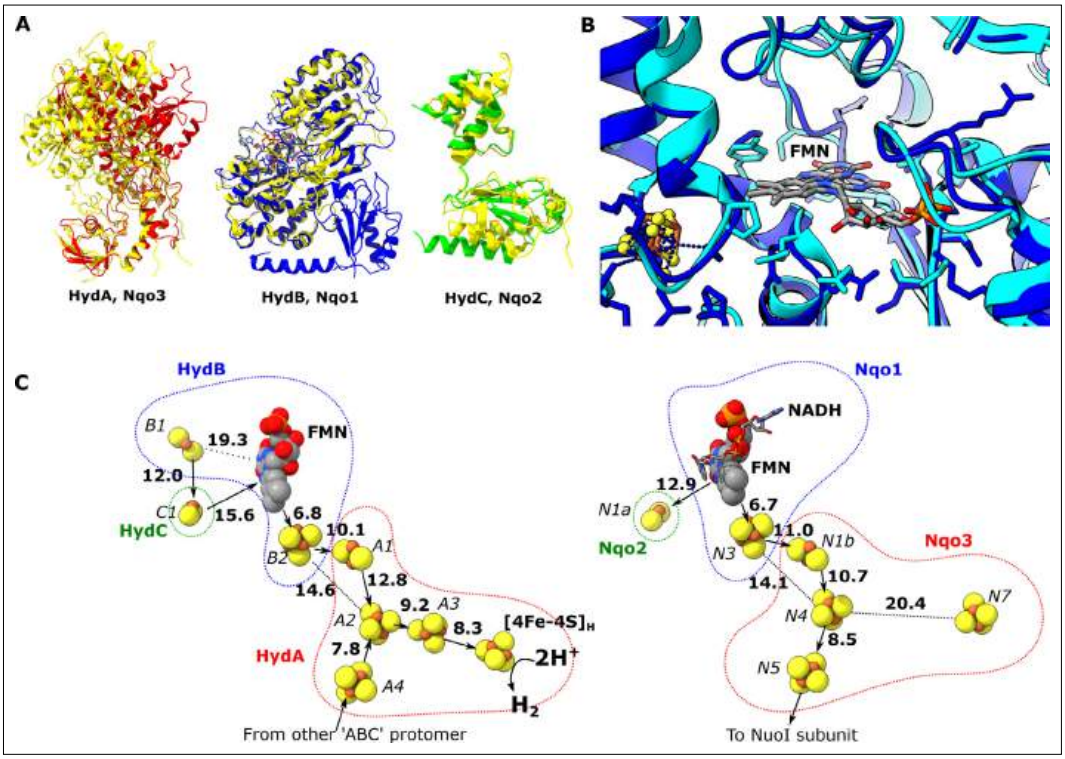


Figure 32.10.22 : Comparison of the HydA, B and C subunits of the electron bifurcating [FeFe] hydrogenase from *Thermotoga maritima* with the Nqo3, 1 and 2 subunits from respiratory complex I from *Thermus thermophilus*.

Panel (**A**) shows the subunits HydA (red), HydB (blues), and HydC (green) overlaid with, respectively, Nqo3, Nqo1, and Nqo2 (all yellow) of complex I from *T. thermophilus* (Gutiérrez-Fernández et al., 2020, PDB: 6ZIY).

Panel (**B**) shows a comparison of the NADH-binding site of the Nqo1 subunit of complex I from *T. thermophilus* (light blue) with the flavin mononucleotide (FMN) site in HydB; the high similarity suggests NADH binds in the proximity of FMN in HydABC similar to complex I.

Panel (**C**) shows an electron transfer network in HydABC compared to complex I from *T. thermophilus* with edge-to-edge distances indicated in bold. The red, blue, and green dotted lines indicate the cofactors present in the HydA (Nqo3), HydB (Nqo1), and HydC (Nqo2) subunits, respectively. Note that our structure is of the apo-HydABC and lacks the $[2Fe]_H$ subcluster of the H-cluster. The $2H^+/H_2$ interconversion reaction in (**C**) illustrates the site at which this reaction occurs, but this will only happen in the fully assembled H-cluster, including $[2Fe]_H$.

Here is a link to a video showing the conformational change observed between the 'Bridge closed forward' (7P8N) and 'Open bridge' (7PN2) classes.

In the video, the HydB C-terminal iron–sulfur cluster domain is colored blue, and the HydA C-terminal iron–sulfur cluster domain is colored orange. The zinc ion (gray sphere) and ligating residues (three cysteine ligands and one histidine) are also shown. The location of the HydA C-terminal domain when the bridge is open is unknown, so it is shown transparently in both states for reference.

The geometric separation of catalytic sites and the bifurcation mechanism **prevents** these thermodynamically favored reactions from happening

- H_2 production from ferredoxin oxidation (in the absence of NADH oxidation)
- NAD^+ reduction by H_2 (in the absence of ferredoxin reduction)
- ferredoxin oxidation by NAD^+

OXYGEN SENSITIVITY OF FE-FE H_2ASES

We have alluded to the fact that Fe-Fe H_2ases can be sensitive to O_2. A possible mechanism involves the interaction of one the Fe ions (Fe_d, the distal Fe) with oxygen, leading to the formation of damaging free radicals. As CO binds more strongly than O_2 to the iron in hemoglobin, its interaction with the H-center can help protect the H_2ases. Sulfides can also afford protection. These mechanisms are illustrated in Figure 32.10.23 :

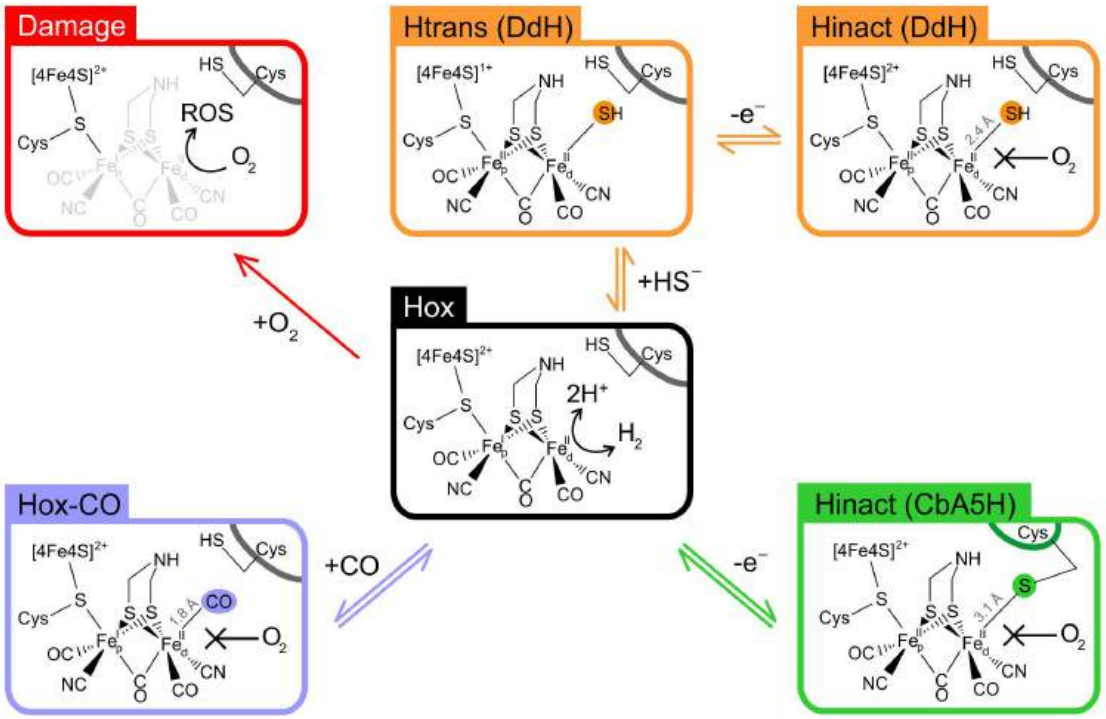


Figure 32.10.23 : Oxygen tolerance strategies in [FeFe]-hydrogenases. Morra S, ibid.

Schematic representation of the H-cluster in the oxidized active state Hox (centre). In the absence of any exogenous protectant, numerous [FeFe]-hydrogenases undergo irreversible inactivation due to H-cluster damage with loss of Fe atoms (red pathway); carbon monoxide acts as a protective agent due to its ability to form Hox-CO, by binding reversibly to the H-cluster at the same site as O_2 (purple pathway); in DdH, a similar mechanism occurs when sulphide binds to the H-cluster forming Hinact, *via* the Htrans intermediate (orange pathway); in CbA5H, a conformational change in the protein structure allows for a conserved cysteine to directly bind to the H-cluster, forming Hinact (green pathway). Fe_p, proximal iron atom; Fe_d, distal iron atom; Cys, cysteine residue.

KEY POINTS - BETA VERSION FROM CHAT.OPENAI

1. Hydrogenases are enzymes that catalyze the reversible conversion of hydrogen gas (H2) to protons and electrons.
2. There are two main types of hydrogenases: [NiFe]-hydrogenases and [FeFe]-hydrogenases.
3. [NiFe]-hydrogenases are found in a variety of microorganisms and mostly used to oxidize H_2.
4. [FeFe]-hydrogenases are mainly used to produce H_2.
5. Hydrogenases play a crucial role in the metabolism of microorganisms, allowing them to produce or consume hydrogen gas as needed.
6. The activity of hydrogenases is regulated by various factors, including the availability of hydrogen and the presence of inhibitors such as oxygen.
7. Research is ongoing to improve the efficiency and cost-effectiveness of hydrogenase-based biohydrogen production and to understand the mechanisms of hydrogenase enzymes to develop more efficient and sustainable ways of producing hydrogen.
8. Genetic engineering techniques can be used to improve hydrogenase activity in microorganisms, and also to increase the tolerance of microorganisms to the toxic effects of H2.

32.11: A WARMER WORLD: TEMPERATURE EFFECTS ON CHEMICAL REACTIONS

Search Fundamentals of Biochemistry

Inspiration for the chapter comes from Biochemical Adaptation by Hochachka and Somero.

Organisms adapt to their environment, with one of the main drivers being temperature. This has occurred over geological time (think of arctic camels 3.4 million years ago!) and space with temperature gradients in terrestrial and aquatic environments. This is evident in the different species that thrive at different mountain heights and ocean depths. Species that can move have advantages in selecting an environment best suited to their thermal needs. Historically, homo sapiens have engaged in seasonal migration, and aquatic species in vertical migrations.

Temperature effects are universal throughout life, and physiology and biochemistry adaptations are ubiquitous. Metabolically-active life can exist from around -15^{o}C to about 121^{o}C (thermal saline springs). Unless greenhouse gas emissions are significantly decreased from present levels, parts of the world will become increasingly uninhabitable due to high temperatures and sea level rise. Estimates for the number of climate refugees range up to 1 billion people by 2050.

Two similar questions arise. Can organisms adapt to increasing temperatures as the climate changes, and are organisms living close to their maximal survivable temperatures?

Before we study the effects of temperature on chemical/biochemical reactions, let's review the basics of thermoregulation. The following classification of organisms by types of thermoregulation is from BioLibre text.

TYPES OF THERMOREGULATION (ECTOTHERMY VS. ENDOTHERMY)

Thermoregulation in organisms runs along a spectrum from endothermy to ectothermy. Endotherms create most of their heat via metabolic processes, and are colloquially referred to as “warm-blooded.” Ectotherms use external sources of temperature to regulate their body temperatures. Ectotherms are colloquially referred to as “cold-blooded” even though their body temperatures often stay within the same temperature ranges as warm-blooded animals.

Ectotherm

An ectotherm, from the Greek (ektós) “outside” and (thermós) “hot,” is an organism in which internal physiological sources of heat are of relatively small or quite negligible importance in controlling body temperature. Since ectotherms rely on environmental heat sources, they can operate at economical metabolic rates. Ectotherms usually live in environments in which temperatures are constant, such as the tropics or ocean. Ectotherms have developed several behavioral thermoregulation mechanisms, such as basking in the sun to increase body temperature or seeking shade to decrease body temperature. The cCommon frog is an ecotherm and regulates its body based on the temperature of the external environment

Endotherms

In contrast to ectotherms, endotherms regulate their own body temperature through internal metabolic processes and usually maintain a narrow range of internal temperatures. Heat is usually generated from the animal’s normal metabolism, but under conditions of excessive cold or low activity, an endotherm generate additional heat by shivering. Many endotherms have a larger number of mitochondria per cell than ectotherms. These mitochondria enables them to generate heat by increasing the rate at which they metabolize fats and sugars. However, endothermic animals must sustain their higher metabolism by eating more food more often. For example, a mouse (endotherm) must consume food every day to sustain high its metabolism, while a snake (ectotherm) may only eat once a month because its metabolism is much lower.

Homeothermy vs. Poikilothermy

Two other descriptors are also used. A **poikilotherm** is an organism whose internal temperature varies considerably. It is the opposite of a **homeotherm**, an organism which maintains thermal homeostasis. Poikilotherm’s internal temperature usually varies with the ambient environmental temperature, and many terrestrial ectotherms are poikilothermic. Poikilothermic animals include many species of fish, amphibians, and reptiles, as well as birds and mammals that lower their metabolism and body temperature as part of hibernation or torpor. Some ectotherms can also be homeotherms. For example, some species of tropical fish inhabit coral reefs that have such stable ambient temperatures that their internal temperature remains constant. Figure 32.11.1 below shows the energy output vs temperature for a homeotherm (mouse) and poikilotherm (lizard).

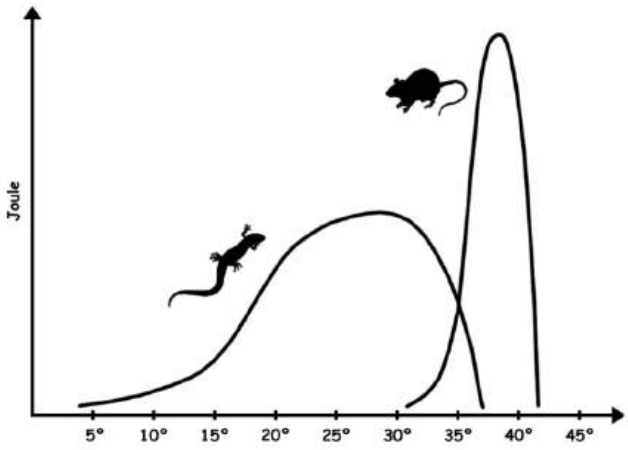


Figure \(\
PageIndex
{1}\):

Homeotherm vs. Poikilotherm: Sustained energy output of an endothermic animal (mammal) and an ectothermic animal (reptile) as a function of core temperature. In this scenario, the mammal is also a homeotherm because it maintains its internal body temperature in a very narrow range. The reptile is also a poikilotherm because it can withstand a large range of temperatures.

Another term is also employed, **heterothermy**, in which the temperature of a homeotherm can vary in different regions of the body (spatially) and also at different times (daily or seasonally as in hibernation). The core body of a homeotherm is usually warmer than the extremities that allow cooling when needed. In hibernation (or sustained torpor), both the body temperature and metabolic rates are decreased.

Means of Heat Transfer

Heat can be exchanged between an animal and its environment through four mechanisms: radiation, evaporation, convection, and conduction. Radiation is the emission of electromagnetic “heat” waves. Heat radiates from the sun and from dry skin the same manner. When a mammal sweats, evaporation removes heat from a surface with a liquid. Convection currents of air remove heat from the surface of dry skin as the air passes over it. Heat can be conducted from one surface to another during direct contact with the surfaces, such as an animal resting on a warm rock.

Key Points

- In response to varying body temperatures, processes such as enzyme production can be modified to acclimate to changes in temperature.
- Endotherms regulate their own internal body temperature, regardless of fluctuating external temperatures, while ectotherms rely on the external environment to regulate their internal body temperature.
- Homeotherms maintain their body temperature within a narrow range, while poikilotherms can tolerate a wide variation in internal body temperature, usually because of environmental variation.
- Heat can be exchanged between the environment and animals via radiation, evaporation, convection, or conduction processes.

Key Terms

- **ectotherm**: An animal that relies on the external environment to regulate its internal body temperature.
- **endotherm**: An animal that regulates its own internal body temperature through metabolic processes.
- **homeotherm**: An animal that maintains a constant internal body temperature, usually within a narrow range of temperatures.
- **poikilotherm**: An animal that varies its internal body temperature within a wide range of temperatures, usually as a result of variation in the environmental temperature.

These terms are diagramed in Figure 32.11.2 below.

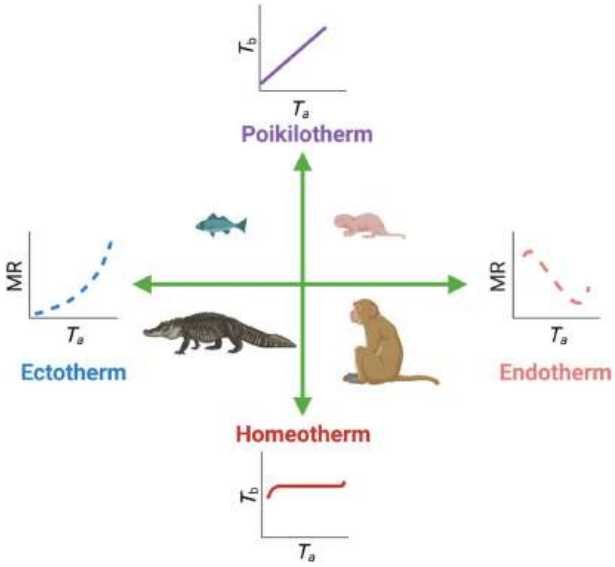


Figure 32.11.2: Thermoregulatory Term. Buffenstein et al., Biol. Rev. (2021), doi: 10.1111/brv.12791. Creative Commons Attribution License

We have discussed in previous chapter sections how temperature can affect macromolecules such as proteins (Chapter 4), nucleic acids (Chapter 9.1) as well as supramolecular assemblies such as membranes (Chapter 10.3). Temperature effects on small molecules and ions (such as salts in the Hofmeister series and glycerol, Chapter 4.9) in the environment that regulate the function/activity of these larger molecules and assemblies are also important. Hence we'll review and discuss the effects of temperature on these key molecular species in the next chapter section. First, we'll delve deeper into the general impact of temperature on chemical and biochemical reactions.

TEMPERATURE EFFECTS ON THE RATES OF CHEMICAL REACTIONS

To understand temperature effects on metabolic processes, let's first review temperature effects on ordinary chemical and biochemical reactions. You may remember the general rule that the rate of a chemical reaction approximately doubles when the temperature is increased 10^o C (10 K). How does that arise? This is generally true in a specific temperature range, as we will see below.

The rates of reactions, either endothermic or exothermic, depend on the activation energy (E_a). The activation energy is required to move from a reactant to the transition state, which then can go on to form product.

The activation energy can be obtained from the Arrhenius equation (that you learned in introductory chemistry), which shows how the rate of an individual chemical reaction depends on temperature.

$$k = Ae^{-E_a/RT} \tag{32.11.1}$$

where k is the rate constant, E_a is the activation energy, E_a/RT is the average kinetic energy, and A is a constant (the "preexponential" factor).

By taking the natural log (ln) of each side and rearranging the equation, you get a "linearized" equation that is easier for most.

$$\ln k = \ln A - \frac{E_a}{RT} \tag{32.11.2}$$

A plot of ln k vs 1/T has a slope = E_a/R, from which the activation energy can be calculated.

An alternative form can be derived:

$$\ln \frac{k_2}{k_1} = \frac{E_a}{R}\left(\frac{1}{T_1} - \frac{1}{T_2}\right) \tag{32.11.3}$$

? A DERIVATION

Here it is!

Derivation

From

$$\ln k_1 = \ln(A) - E_a/RT_1 \tag{32.11.4}$$

solve for lnA

$$\ln(A) = \ln(k_1) + E_a/RT_1 \tag{32.11.5}$$

Substitute into the equation for ln(k_2) gives

$$\ln(k_2) = \ln(k_1) + E_a/RT_1 - E_a/RT_2 \tag{32.11.6}$$

Rearrange to get

$$\ln(k_2) - \ln(k_1) = E_a/RT_1 - E_a/RT_2 \tag{32.11.7}$$

Simplify to get the final equation!

$$\ln\left(\frac{k_2}{k_1}\right) = \frac{E_a}{R}\left(\frac{1}{T_1} - \frac{1}{T_2}\right) \tag{32.11.8}$$

Solving for E_a gives

$$E_a = \frac{R\ln\frac{k_2}{k_1}}{\frac{1}{T_1} - \frac{1}{T_2}} \tag{32.11.9}$$

Let's use this equation to calculate an E_a that will give a doubling of the reaction rate ($k_2/k_1 = 2$) going from T_1 = 295 K (21.9^0 C, 71.3^o F) to T_2 = 305 K (21.9^0 C, 89.3^o F), a 10^{o}C temperature rise.

$$\begin{aligned} E_a &= \frac{(8.314)(\ln 2)}{\frac{1}{295} - \frac{1}{305}} \\ &= \frac{\left(8.314\ \text{J mol}^{-1}\ \text{K}^{-1}\right)(0.693)}{0.00339\ \text{K}^{-1} - 0.00328\ \text{K}^{-1}} \\ &= \frac{5.76\text{Jmol}^{-1}\ \text{K}^{-1}}{\left(0.00011\ \text{K}^{-1}\right)} \\ &= 52,400\text{Jmol}^{-1} = 52.4\ \text{kJ mol}^{-1} \end{aligned} \tag{32.11.10}$$

Hence if a reaction has an activation energy E_a of about 54 kJ/mol, increasing the temperature from 295 to 305^{o}C (i.e, by 10^{o}C) doubles the reaction rate.

Assuming that the activation energy is constant, the rate constants increase with temperature since a larger fraction of the molecules have the energy ($\geq E_a$) necessary to react. This is illustrated in Figure 32.11.3 below.

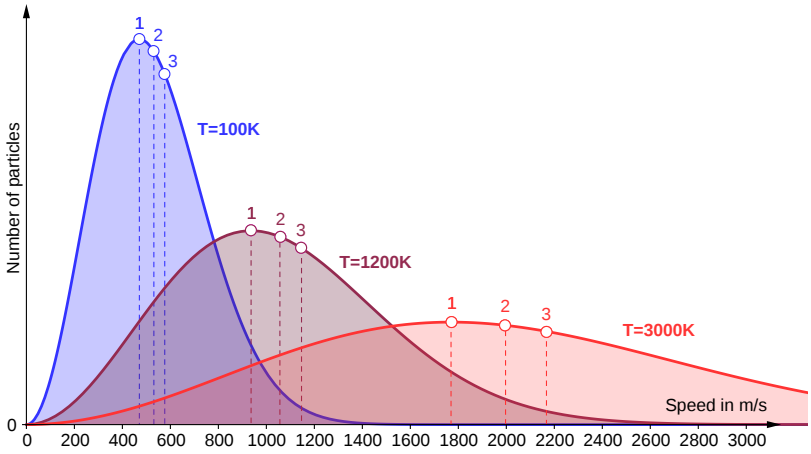


Figure 32.11.3: Plot of a Maxwell-Boltzmann distribution of speeds for different temperatures T=100K, T=1200K, T=5000K. Points along the curve show (1) most likely speed, (2) average speed, and (3) thermal speed (velocity that a particle in a system would have if its kinetic energy were equal to the *average* energy of all the particles of the system). https://commons.wikimedia.org/wiki/F...xis-labels.svg. Creative Commons Attribution-Share Alike 4.0 International license.

Let's look a the brown vertical line around 950 m/s. If we take that as the activation energy, very few molecules in the blue distribution have the required kinetic energy $\geq$ Eact. Ar progressively higher temperatures, great fractions (as measured by the area under the curve to

the right of the dotted line at 950 m/s) have the required energy, hence the rates increase with temperature.

When the **temperature change is 10^oC**, the ratio of the rate constants (or rates), k_2/k_1 is often called **Q_{10}**, the **temperature coefficient** (unitless). Q_{10} is not a constant, since it depends on the two temperatures that differ by 10^0 C (10 K). Hence the Q_{10} value for the 10^0 range from 273-283K is different than the Q_{10} value from 373-383K) . Q_{10} for many reaction is around 2 (doubling of the reaction rate) - 3 (tripling the reaction rate) at physiological temperature . Q_{10} =2 for a given E_a only at one set of temperatures that differ by 10^oC. The variation in Q_{10} values is illustrated in Table 32.11.1 below for a reaction in which E_a = 44.5 kJ/mol. Q_{10} **decreases** from 2 as the temperatures T_1 and T_2=T_1+10^oC increase.

T_1 in K (^oC)	T_2 (K) (^oC)	k_2/k_1 (Q_{10})
273 (- 0.15 ^oC)	283 (9.85 ^oC)^oC	2
373 (99.9 ^oC)	383 (110 ^oC)	1.45
473 (200 ^oC)	483 (210 ^oC)	1.26

Table 32.11.1: Q_{10} = k_2/k_1 values at different temperatures T_1 and T_2 that differ by 10^oC.

We will see how this is important in biological settings in a bit. If Q_{10} = 1, the reaction is independent of temperature, and a Q_{10} <1 shows a reaction that is not functioning. An example might be an enzyme-catalyzed reaction in which the threshold is reached at a higher temperature T_2 = T_1+10, at which the enzymes lose an active conformation and starts to unfold.

The same equation and the Q_{10} parameters apply to enzyme-catalyzed reactions. The activation energies (E_a) for four enzymes involved in the degradation of lignocellulose in the surface soil and subsoil are shown in Table 32.11.2 below. The enzymes include two hydrolases, β-glucosidase (BG) and cellobiohydrolase (CB), which cleave cellulose, and two oxidases, peroxidase (PER) and phenol oxidase (POX), which help degrade lignin. The overall average E_a for these enzymes is about 44.7 kJ/mol, similar to the example used in Table 1 above.

Soil	Type	Ea (kJ/mol)			
		BG	CB	PER	POX
Arctic	surface	35.4	39.4	12.7	81.8
Subarctic	surface	36.5	38.6	21.1	45.7
	subsoil	52.2	41.5	22.4	39.4
Temperate 1	surface	40.9	38	64.9	102
	subsoil	49.4	21.2	28	94.8
Temperate 2	surface	31	43.4	25.4	49.5
	subsoil	40.9	39.9	19.8	47.5
Temperate 3	surface	51.5	53.6	28.8	73.2
	subsoil	58.8	46.7	54.2	29
Tropical 1	surface	47.8	50.5	26.5	47.7
	subsoil	56.6	47	47.1	27.1
Tropical 2	surface	39.3	42.5	58.3	82.5
	subsoil	42.8	43.3	22.8	45.5
Avg		44.9	42.0	33.2	58.9

Table 32.11.2: Activation Energies (E_a, kJ mol^{-1}) for extracellular soil enzymes involved in the degradation of lignocellulose. Adapted from Steinweg JM et al. (2013) PLOS ONE 8(3): e59943. https://doi.org/10.1371/journal.pone.0059943. Creative Commons CC0 public

domain

Q_{10} temperature coefficients are also used to describe biological processes like respiration, speed of neural signal propagation, metabolic rates, etc. Many biological processes are affected by temperature, especially for ectotherms that adjust temperatures to outside environments, including daily and seasonal temperature shifts. Mammals and birds alter metabolic rates with temperature. This is true for hibernating animals.

The Q_{10} **temperature coefficient** can be considered to be the *factor* by which the reaction rates (k or R) increase (factor of 2, 3, 1.5, etc) for each **10-degree** K or C temperature increase. It is given by the following equation:

$$Q_{10} = \left(\frac{k_2}{k_1}\right)^{10^\circ\text{C}/(T_2 - T_1)} \tag{32.11.11}$$

It is also called the **van't Hoff's temperature coefficient**. To help understand Q_{10}, let's consider some examples.

- If T_2-T_1=10^o, then Q_{10} is simply k_2/k_1 for the specified temperature pairs separated by a 10^0 C range (T_1 and T_2=T_1+10). Remember that Q_{10} is not a constant but depends on the temperature pairs and that it goes down with increasing temperature.
- If the temperature range is > 10^0 C, the the measured ratio k_2/k_1 is a factor > 1 x Q_{10}
- If the temperature range is < 10^0 C, the the measured ratio k_2/k_1 is a fraction of Q_{10}

This equation can be converted to

$$k_2 = k_1 Q_{10}^{(T_2 - T_1)/10^\circ\text{C}} \tag{32.11.12}$$

where the rate constant k_2 is related to a "base" rate k_1 at a base temperature of T_1. An interactive graph of the above equation is shown in Figure 32.11.4 below.

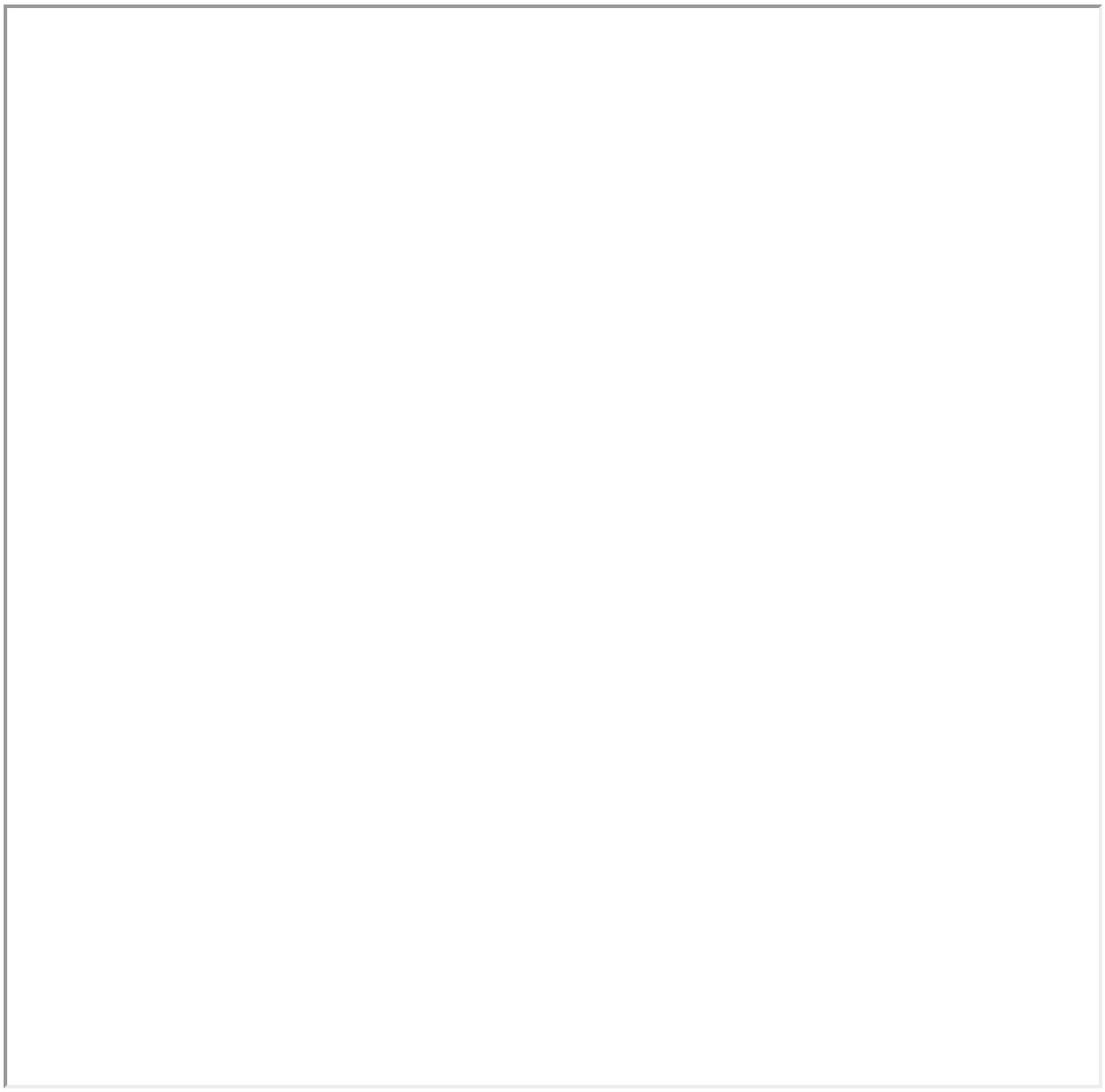


Figure 32.11.4: Interactive graph of k_2 (rate 2) vs. delta T at different base rates k_1.

Change the base rate constant, k_1, at a base temperature of T_1, and Q_{10} coefficient to see how they change k_2.

Note that if Q_{10} =1, k_2 at T_1+10 = k_1 at T_1, the rate is independent of the temperature.

For most biological systems, the Q_{10} value is ~ 2 to 3 under physiological relevant conditions. The ratios of the rates (R_2/R_1) for different Q_{10} values are shown in Figure 32.11.5 below.

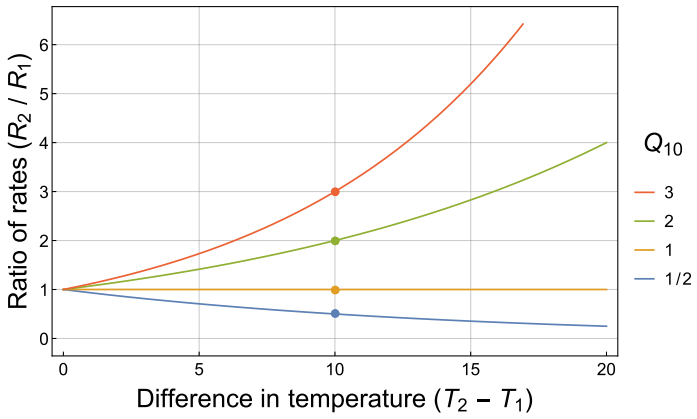


Figure 32.11.5: Idealized graphs showing the dependence on temperature of the rates of chemical reactions and various biological processes for several different Q10 temperature coefficients. The dots on the graph show how the rate change with a temperature difference of 10oC. Wikipedia. https://en.wikipedia.org/wiki/Q10_(t...e_coefficient). CC BY-SA 4.0

Again this hypothetical graph is meant to show the general meaning of Q_{10} values.

The "Q" model has been used to fit complex reaction systems, not just individual reactions. Figure 32.11.6 below shows the daily mean soil respiration rate as a function of soil temperature. In these graphs, the x-axis is Temperature, not ΔT.

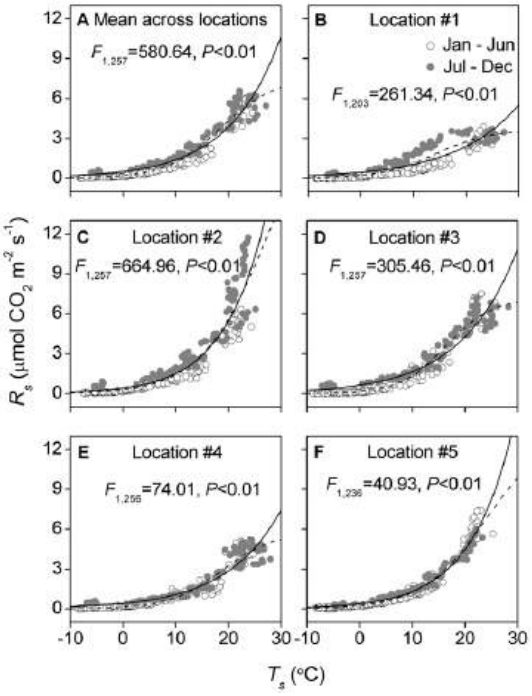


Figure 32.11.6: Relationships between daily mean soil respiration (R_s) and soil temperature (T_s). Jia X et al., PLoS ONE 8(2): e57858. https://doi.org/10.1371/journal.pone.0057858. Creative Commons Attribution License

The soil temperature, T_s, was measured at a 10-cm depth. Open circles are from January to June; closed circles are from July to December. The solid lines use a Q_{10} model, in which the observed R_s vs T_s data are fit with an equation that optimizes the Q_{10} parameter The dashed lines are fitted by a logistic model, which we used in Chapter 5.7 for fitting ELISAs data. R_s is significantly different between the first and second half of the year.

The soil respiration rate, R_s, at 10 cm depth was strongly affected by temperature, with an annual Q_{10} value of 2.76. Daily estimates of Q_{10} averaged 2.04 and decreased with increasing T_s. A study of seagrass showed the Q10 values are affected by plant tissue age and that Q10 varied significantly with the initial temperature and temperature ranges.

The use of Q_{10} values from the Arrhenius equation is based on the assumption that the chemical/biochemical processes are exponential functions of temperature. For complex processes like the decay of organic matter, it would be better to model the whole system by looking at the individual enzymes involved. One problem with using Q_{10} values for very complex systems is the choice of the base temperature value for rate comparisons. The anaerobic decomposition of organic matter is generally a linear function of temperature

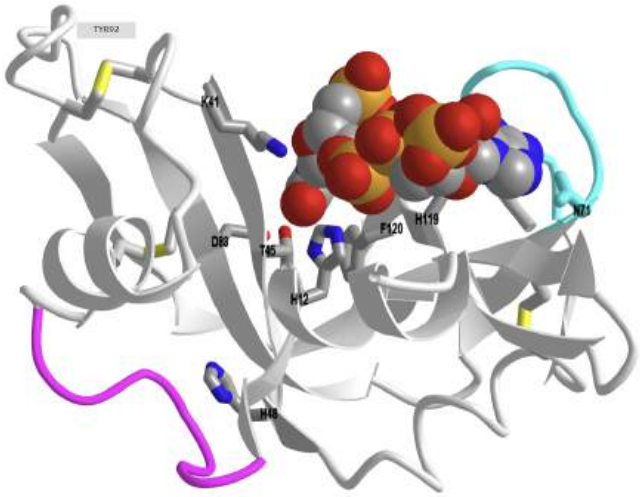
between 5°C and 30°C, which shows that a Q_{10} modeling system is not ideal. A more complex systems biology approach using programs like Vcell and COPASI would be better and less likely to cause errors in predicted CH_4 emissions from the decomposition process, for example.

GETTING BACK TO PROTEINS

In Chapter 6.1 we explored the mechanisms used by enzymes to catalyze chemical reactions. These included general acid/base catalysis, metal ion (electrostatic) catalysis, covalent (nucleophilic) catalysis, and transition state stabilization. Some physical processes included intramolecular catalysis and strain/distortion. The rate-limiting step in enzyme-catalyzed reactions can include actual bond breaking in the substrate, dissociation of product, and conformational change required to facilitate binding, catalysis, and dissociation. A rate-limiting conformational change may occur not in the active site pocket but in nearby loops that modulate the accessibility of reactant to and dissociation of product from the active site. These all may be influenced by temperature, with localized conformational flexibility especially important.

An interesting example of localized conformational changes affecting enzyme activity is RNase A. His 48, 18 Å from the enzyme active site, is involved in the rate-limiting enzymatic step involving product release.

Figure 32.11.7 shows an interactive iCn3D model of bovine pancreatic Ribonuclease A in complex with 3'-phosphothymidine (3'-5')-pyrophosphate adenosine 3'-phosphate (1U1B)



NCBI iCn3D Figure 32.11.7: Bovine pancreatic Ribonuclease A in complex with 3'-phosphothymidine (3'-5')-pyrophosphate adenosine 3'-phosphate (1U1B). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...SnrGYXSVXcLCk6

The substrate is shown in spacefill. The active site side chains and the distal His 48 are shown as sticks and labeled. Two flexible loops, Loop 1 (magenta) near His 48, and Loop 4 (cyan) near the active site are highlighted. On ligand binding, the loops move a few angstroms to make the active site more closed, inhibiting product release. Product release is associated with mobile regions including Loops 1 (20 Å from the active site) and 2. Loop 4, near the active site, is involved in the specificity for purines 5' to the substrate cleavage site. His 48 is conserved in pancreatic RNase A. If mutated to alanine, the k_{cat} decreases greater than 10X, indicating a change in the rate-determining conformational motion. The enzyme is still very active compared to the uncatalyzed reaction. His 48 appears to regulate coupled motions in the protein that are rate-limiting.

Figure 32.11.8 below shows the subtle shift in the conformation of apo-RNase A (magenta, no ligand, 1FS3) to the substrate-bound form (cyan, ligand in sticks), 1U1B). Note the small motion in His 48 shown in sticks at the bottom of the image.

Figure 32.11.8: Conformational changes apo-RNase A (magenta, no ligand, 1FS3) on conversion to the substrate-bound form (cyan, ligand in sticks, 1U1B).

We will explore temperature effects on protein structure and function more in the next chapter section.

32.12: A WARMER WORLD: TEMPERATURE EFFECTS ON PROTEINS

Inspiration for the chapter comes from Biochemical Adaptation by Hochachka and Somero.

INTRODUCTION

In the previous chapter section, we discussed the generalized effects of temperature on chemical/biochemical reactions. The rate for chemical reactions, including enzyme-catalyzed ones, typically increases 2-3 fold (Q_{10} values) with a 10^0 C temperature increase over an organism's typical temperature range. Q_{10} values decrease at higher temperature pairs differing by 10^0 C. At too high a temperature, a protein enzyme destabilizes, and Q_{10} values can fall to less than 1, a sign of potential trouble for an organism subjected to that higher temperature.

As we saw in Chapter 32.11 and will here, two competing processes affect protein enzymes as temperatures increase. They are increased rates for the catalyzed reaction and increased conformational dynamics, which leads to eventual denaturation at high enough temperatures. Hence evolution would presumably select for increased protein stability for organisms adapted to higher temperatures.

For cold-adapted organisms, the rates of catalysis are expected to decrease. Hence evolution might lead to higher k_{cat} values for cold-adapted organisms. However, we saw in Chapter 32.11 that the rate-limiting step for enzyme-catalyzed reactions often involved localized conformational changes, which would be disfavored at colder temperatures. Hence evolution would also select for enzymes that could maintain flexibility at low temperatures. In Chapter 4.9, we discussed low-temperature protein denaturation. Proteins can be destabilized at low temperatures. In this section, we will study how enzymes can adapt to higher temperatures. We won't discuss how proteins adapt to cold since this topic is less relevant for human-caused climate change.

We will follow the approach used throughout this book - that structure mediates function. We will use a lot of enzyme kinetics since kinetic parameters can tell us much about how bound substrate goes to product at high substrate concentrations (k_{cat} or V_M) or low substrate concentrations (k_{cat}/K_M) at different temperatures and for different orthologs of enzymes from species that have adapted to grow at low (psychrophile), medium (mesophile) or high temperatures (thermophiles). We'll next look at the structure of enzymes and which features allow them to adapt to their optimal temperature for growth. Finally, we will look at entire pathways to discern clues as to how they adapt to increased temperatures.

AN OVERVIEW - SOIL ENZYMES

Soil microbes play a key part in the carbon cycle as they can both store and release carbon. Soil temperatures influence this balance between uptake and release of CO_2 into the atmosphere.

SOIL ORGANIC CARBON - SOC

The soil is a sink for carbon and stores about 1500 gigatons [Gt] = 1.5 Pt = 1500 Pg), more than the atmosphere and vegetation combined. SOC derives ultimately from photosynthetic organisms. When they die, their carbon is used by heterotrophs for energy and biosynthesis. Carbon can also be returned to the atmosphere by aerobic oxidation by microorganisms, but this requires O_2, which diminishes rapidly with soil dept. Oxygen levels depend on soil porosity, relative amounts of sand and clay, and hydration. Tilling of soil increases O_2 exposure and hence oxidative respiration of SOC, increasing atmospheric CO_2. No-till farming hence can decrease CO_2 release into the atmosphere. Inorganic carbon from CO_2 (HCO_3^-, and CO_3^{-2}) bind with cations in the soil (mineralization) or is released into the atmosphere as CO_2.

Carbon input into the soil is mostly determined by photosynthesis, which correlates with root mass, and decay, while export is determined by soil microbial (bacteria, fungi, protists, animals) respiration. Microorganisms play a key role in the balance of carbon input and release in the soil and hence are prime determinants of SOC.

SOC is high in northern latitudes since colder temperatures promote lower respiration rates and accumulation of SOC over time. SOC is low in the lush tropics (even given the high photosynthetic rates) because of a high microbial respiration rate at higher temperatures. Deforestation of the lush Amazon Rain Forest will leave soil poor SOC with little to balance CO_2 release from the decomposition of what's left by the abundant soil microorganisms.

About 21 Gt of the 1500 Gt of SOC consists of microbial mass (12 Gt fungi, 7 Gt bacteria and 2 Gt from animals). Fungi hence are key players in soil metabolism. They are involved in the slow decomposition of decaying organic matter and promote the growth of slow-growing organisms like trees. In contrast, bacteria are fast metabolizers, and are found in abundance in grasslands. Northern attitudes have a higher soil microbial mass than in the tropics, but they are less active, allowing great SOC stores.

We often think of enzymes working in an aqueous environment in a test tube or a cell (which is very crowded with other molecules). Figure 32.12.1 below represents the microenvironment of soil enzymes involved in the decomposition of SOC, like cellulose.

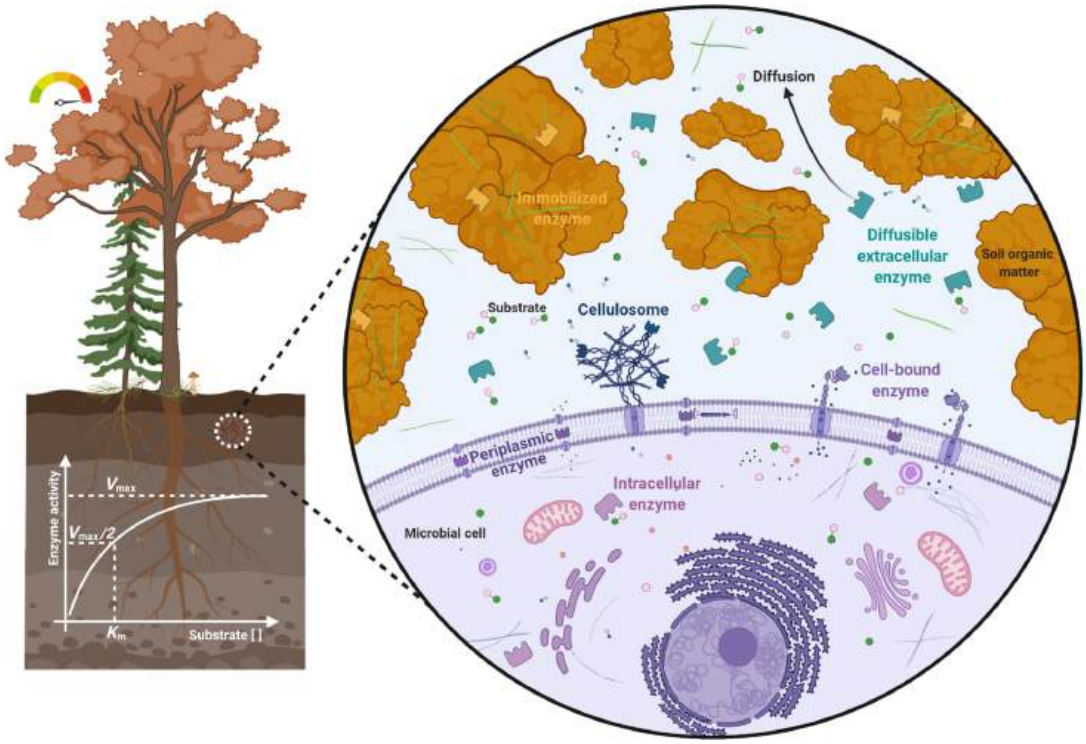


Figure 32.12.1: Location of enzymes in soils and their importance for carbon and nutrient cycling. Fanin, N. et al. (2022). Soil enzymes in response to climate warming: Mechanisms and feedbacks. *Functional Ecology*, 36, 1378– 1395. https://doi.org/10.1111/1365-2435.14027. Permission from John Wiley and Sons and Copyright Clearance Center.

Figure 32.12.2 below give a review and an overview of the effects of increasing temperature on soil enzymes.

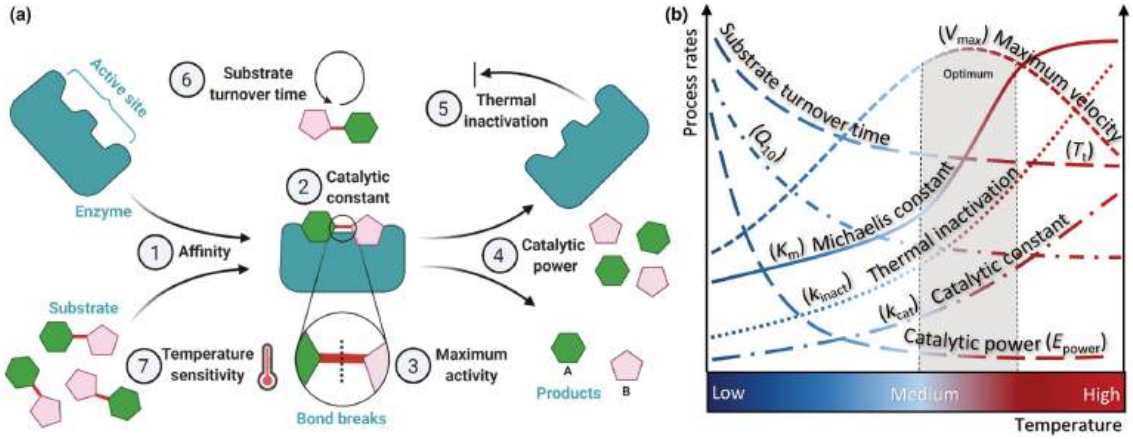


Figure 32.12.2: Effects of temperature at the enzyme scale. Fanin, N. et al., ibid. Permission from John Wiley and Sons and Copyright Clearance Center.

Panel (a) shows the many steps involved in enzyme catalysis that can be affected by temperature changes. Step 1 shows the binding of substrates. The K_M (units M) for the enzyme gives a "measure" of the strength of the interaction (but remember that $K_M = K_D$ - the dissociation constant - only under rapid equilibrium conditions). Step 2 reflects k_{cat}, the "net" rate constant for converting bound substrate to product.

Panel (b) shows how key constants change with increasing temperature. The figure shows that k_{cat} increases with temperature, consistent with the Arrhenius equation, as the temperature coefficient Q_{10} decreases (as discussed in Chapter 12.11). (Remember that Q10 is the *factor* by which the reaction rate, k or R, changes for each **10-degree** K or C temperature increase))The enzyme's thermal inactivation

rate, k_{inact}, also increases with temperature, leading to the bell-shaped V_M curve. K_m, a measure of the apparent K_D of the substrate for the enzymes, increases, reflecting weaker binding. The catalytic power, $E_{power} = k_{cat} / k_{inact}$, also decreases with increasing temperature as the slope of k_{inact} is generally greater than that of k_{cat}. The values for the temperature axis would be different for microbes that grow best at cold temperatures (psychrophilic), moderate temperatures (mesophilic), and high temperatures (thermophilic).

The graphs above represent temperature effects at the enzyme level. The gray rectangle represents the optimal growing conditions, which show that enzymes are poised near V_M for substrate conversion (assuming abundant substrate) but with low catalytic power. Increasing temperatures also have an effect at the microbial community level. These can affect SOC. For example:

- After an increase in the decomposition of SOC at higher temperatures, subsequent decreases in SOC can occur due to the depletion of available substrates (as enzymes are running at V_M) and changes in carbon use among the microbial communities;
- Additional decreases in SOC due to increased oxidation and shifts in the composition of the microbial community occur;
- The levels and types of substrates for enzymes likely change;
- Increased temperatures can lead to increases in microbial community mass, which requires more substrate, but in the long-term metabolic shifts might lead to a decrease of extracellular enzymes and microbial biomass;
- Soil conditions also change with increasing temperatures, which affects biomass by changing substrate availability;
- Increased temperature lead to short-term increases in CO_2 emission due to higher microorganism metabolic rates explained by the Arrhenius equation, but additional effects caused by accompanying changes in the microbial community occur.

Complex mathematical modeling (as we saw using Vcell with metabolic and signal transduction pathways) would be needed to understand the effects of warming on SOC stores and their return to the atmosphere as CO_2.

ENZYME PROPERTIES WITH ALTITUDE - MOUNT KILIMANJARO

It is possible that the loss of SOC with climate change may be mitigated to some extent as the soil microbial community thermally adapts to a lower respiration rate/microbial biomass. As Figure 1 shows, both extracellular and intracellular enzymes must be considered. Extracellular enzymes break down polymers like cellulose into monomers, which are transported into the cell for intracellular respiration and the formation of CO_2 by intracellular enzymes. The extracellular (lytic) and intracellular (oxidative) enzymes might respond differently to higher temperatures. Polymers that are hard to degrade have high activation energies, making soils with higher concentrations of these polymers more sensitive to climate warming (based on the Arrhenius equation).

Changes in the microbial community might include shifts in the fungal/bacterial ratio, causing changes in degradation pathways and the rates of enzyme-catalyzed reactions. At higher temperatures, such changes increase conformational flexibility, which could increase k_{cat} but also decrease the apparent affinity of the enzyme for the substrate, as reflected in increased K_M values. These compensatory effects might leave catalysis unaffected by increasing temperatures.

Studies have been conducted on individual degradative enzymes in soil samples from Mount Kilimanjaro. Enzyme kinetic analyses were done at two different temperatures differing by 10^0 C (10^0 C and 20^0 C), so Q10 values could be evaluated. The soil samples were obtained from different heights on the mountain to allow for the comparison of the kinetic parameters of enzymes from microbes adapted to different heights. All organisms in soil from different heights would experience 10 °C, while those at the highest altitudes (3000 m) would encounter 20^0 C only in the summer. The microbes presumably would have slightly different optimal temperatures for growth, and would likely use different adaptive mechanisms at low and high altitudes. Keep in mind in interpreting the results below that the Q10 value determines the sensitivity of a parameter (v_0, K_M, V_M, etc.) to an increase in the temperature of 10^0 C.

Extracellular enzyme activities in soil from one altitude

The activity of three extracellular enzymes in soil samples were studied: β-1,4-glucosidase (degrades cellulose), N-acetylglucosaminidase (degrades chitin from fungi and peptidoglycans from bacteria), and acid phosphatase. The first two enzymes catalyze "recalcitrant" reactions with higher activation energies. Michaelis-Menten plots for the three enzymes in soil samples taken at one height, 2010 meters (m), are shown in Figure 32.12.3 below. In addition, a plot of glucose conversion to CO_2, an intracellular process, which the authors termed "glucose" mineralization" (probably because they trapped CO_2 using OH^- to form HCO_3^-), is also shown.

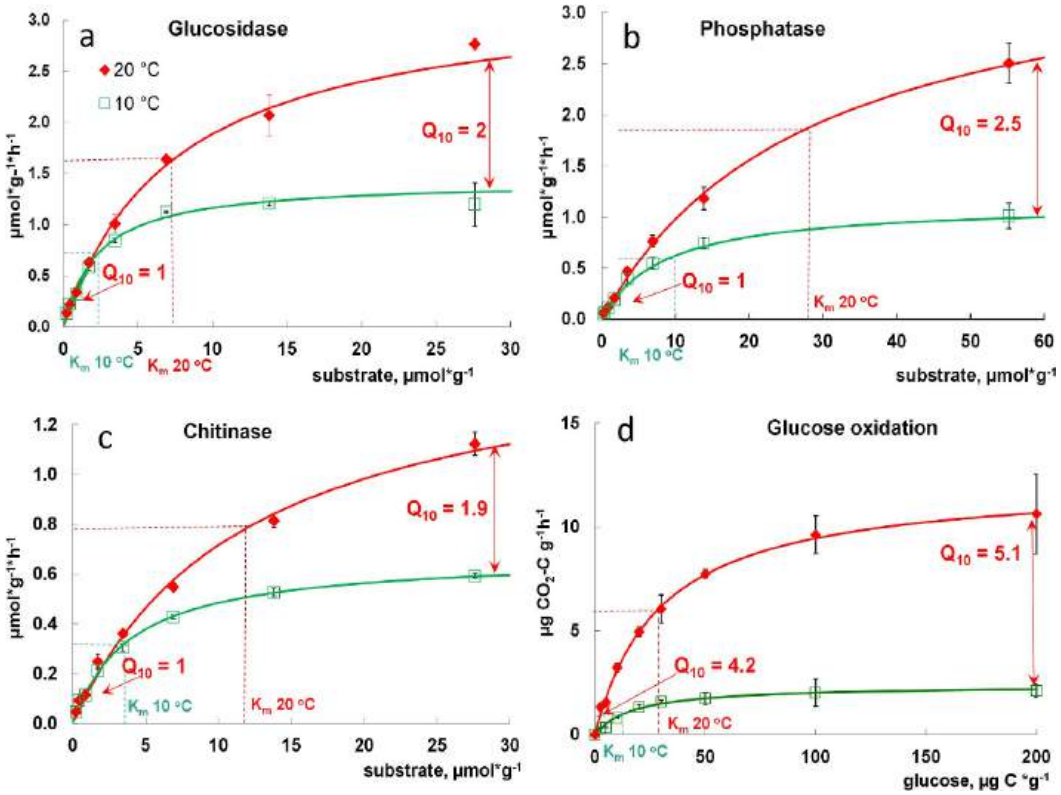


Figure 32.12.3: Rates of reactions mediated by hydrolytic enzymes (**a–c**) and rates of glucose oxidation to CO_2 (**d**) as dependent on substrate concentration at 10 and 20 °C for the site located at 2010 m a.s.l. Symbols – experimental data, lines – approximation by Michaelis–Menten kinetics. The **red** lines indicate assays performed at 20^0 C, while the **green** lines are for assays run at 10^0 C. Blagodatskaya, E. *et al*. Temperature sensitivity and enzymatic mechanisms of soil organic matter decomposition along an altitudinal gradient on Mount Kilimanjaro. *Sci Rep* **6**, 22240 (2016). https://doi.org/10.1038/srep22240. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/

The mineralization rate was determined using trace amounts of ^{14}C-labeled glucose, which if fully oxidized, is converted to $^{14}CO_2$. Given the conditions of the reactions, the added glucose did not cause microbial proliferation. The authors used a fluorophore (4-methylumbelliferone or MUF)-labeled small substrate analogs for cellulose (MUF-β-D-glucopyranoside), chitin (MUF-N-acetyl-β-D-glucosaminide dehydrate) and for acid phosphatase (4-MUF-phosphate). Reactions were carried out in soil samples and valid initial velocities for the reaction were determined. The Km values for the glucose oxidation (mineralization) is **not** a valid K_M value since CO_2 would be produced from the combined actions of the enzymes in glycolysis, pyruvate dehydrogenase and the citric acid cycle. It could be better called an "operational K_M".

For the three enzymes, the K_M at 20 °C was 25–42% larger than the $V_M = k_{cat}E_{T\ at}$ 20 °C, causing $Q_{10}^{KM} > Q_{10}^{VM}$. These compensatory changes canceled any increases in enzyme activity at low substrate concentrations, but not at high ones when the enzyme was saturated. Hence Q10 for the catabolic depolymerization reactions increased with substrate concentration. Note, however, that the rate of intracellular glucose oxidation (mineralization) increased at all substrate concentrations going from 10 to 20 °C and the canceling effect was not detected, even at low substrate levels. The temperature response of monomer oxidation showed a strongly accelerated reaction rate instead of a canceling effect

Extracellular Enzyme activities in soils from different altitudes

Next, kinetic analyses were performed on soil samples from 2010 m (warm-adapted microorganisms) and 3020 m (cold-adapted) on the mountain. Plots of Q10 (**not** v_0 as in Michaelis-Menten plots) vs. substrate for these studies are presented in Figure 32.12.4 below to show the temperature adaption capacities of the enzymes.

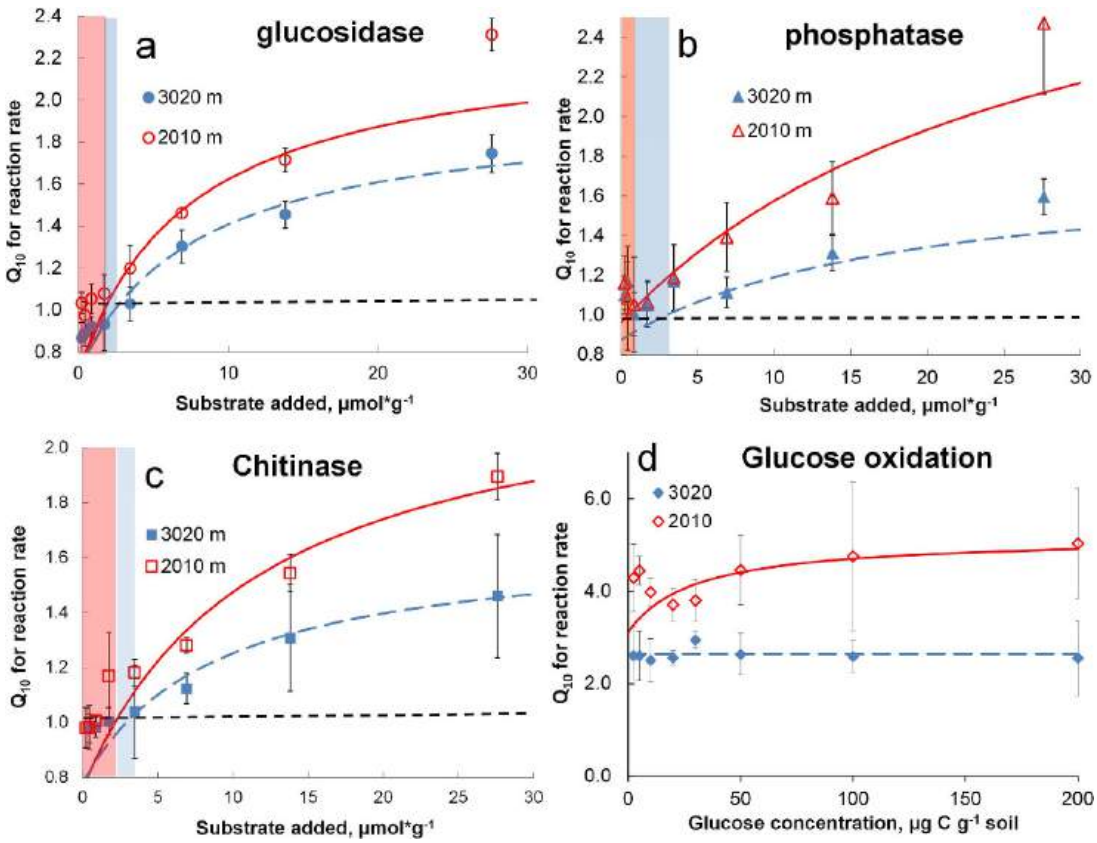


Figure 32.12.4: The Q_{10} values for enzymatic activities (**a–c**) and glucose oxidation to CO_2 (**d**) as dependent on substrate concentration at two altitudes. The blue and red rectangles show the concentration range at which no temperature effects occur (i.e $\leq S_{crit}$) with shading colors corresponding to different altitudes. The Q_{10} values derived from experimental data are shown as symbols. The model simulations based on experimentally obtained parameters of Michaelis–Menten kinetics are shown as curves (**a–c**). For glucose oxidation (**d**) at 3020 m elevation, a non-linear trend was very weakly expressed. Bars show standard deviations of the means (n = 3). Blagodatskaya, E. *et al. ibid.*

The graphs show that Q10 for polymer degradation increased with increasing substrate concentration. The authors defined a substrate concentration threshold (S_{crit}) below which K_M and V_M values canceled, so no increase in rate was seen with increasing temperature based. The width of the rectangles is based on the best-fit dashed **blue** and **red** lines, not the data points. The S_{crit} values were 35–42% larger for the 3020 m (**blue** rectangle) soil samples than at 2010 m (**red** rectangle), even though SOC was lower at that elevation. Q10 values were always lower at $S > S_{crit}$ at higher altitudes. This demonstrates that the enzymes responded less to higher temperatures at higher altitudes (**blue** dots and dashed lower curves), implying that the enzymes at higher latitudes demonstrated significant compensatory changes useful for microorganisms that experience a greater temperature range with larger shifts at these higher altitudes.

Figure 32.12.5 below shows that the Q values generally decrease over a larger range of altitudes for both enzyme activity (panel A) and glucose oxidation (Panel B).

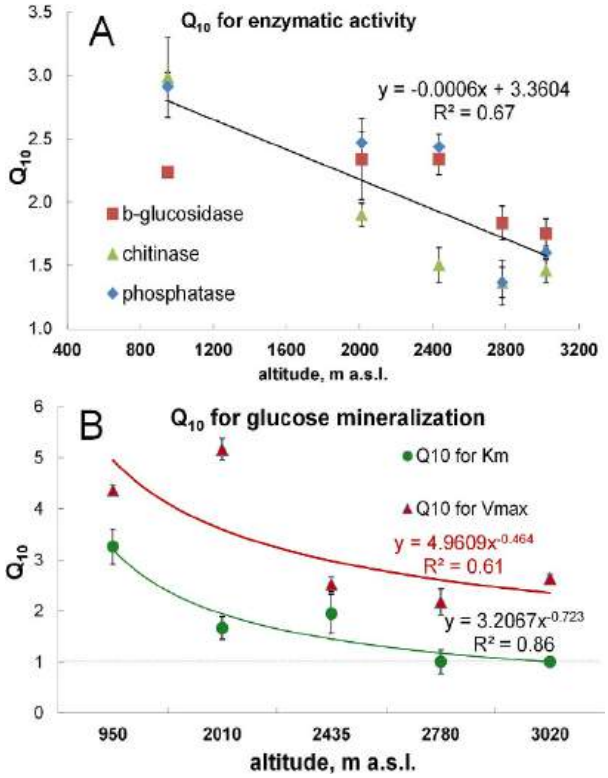


Figure 32.12.5: The $\mathbf{Q_{10}^{total}}$ values for hydrolytic enzyme activity at saturating substrate concentrations (**A**) and the increase in V_{max} and K_m induced by a temperature increase from 10 to 20 °C for ^{14}C-glucose oxidation (**B**) depending on altitude. Symbols – experimentally derived values for $\mathbf{Q_{10}^{total}}$ (**B**), $\mathbf{Q_{10}^{Vmax}}$ and $\mathbf{Q_{10}^{Km}}$ (**A**). Lines are the trend lines obtained by the best fitting of power (**A**) and linear functions (**B**) at P values < 0.05, bars show standard deviations of the means (n = 3). Blagodatskaya, E. *et al. ibid.*

Panel B shows how Q10 values for K_M apparent (**green**) and V_M (**red**) for intracellular glucose oxidation/mineralization both decreased with increasing altitude. Again consider these K_M and V_M values to be operationally defined and apply not to an individual enzyme but in less rigorous way to all the enzymes involved in the intracellular oxidation of glucose to CO_2. Investigators could determine these values only by fitting the kinetic data for CO_2 vs [glucose] to the Michaelis-Menten equation. At altitudes < 2435 m, all showed Q10 K_M values > 1.9, showing that the apparent collective K_M values were very sensitive to temperature increase. This implies the "collective" set of enzymes responsible for intracellular glucose oxidation was more conformational flexible, and higher temperatures caused significant increases in apparent K_M values. However, at high altitudes, the Q10 values for the apparent K_M were about 1, suggesting no temperature effects on the generic structure and apparent K_M values for the enzymes. The high-altitude enzymes were effectively temperature-stable with respect to K_M values. V_M values were more sensitive to increasing temperatures at all altitudes, but little change was seen going from 2435 to 3020 m. This again shows that the enzymes for microorganisms from high altitudes were more strongly adapted to temperature changes, especially at lower substrate concentrations.

The next part can be a bit confusing. The K_M for a given enzyme increases with increasing temperature, as shown in Figure 2. This suggests that the apparent affinity of the given enzyme for substrates decreases, which makes intuitive sense. Figure 32.12.6 shows that the Q10 for K_M (i.e. the sensitivity of K_M to a temperature increase of 10^0 C) decreases with increasing elevation for each enzyme studied.

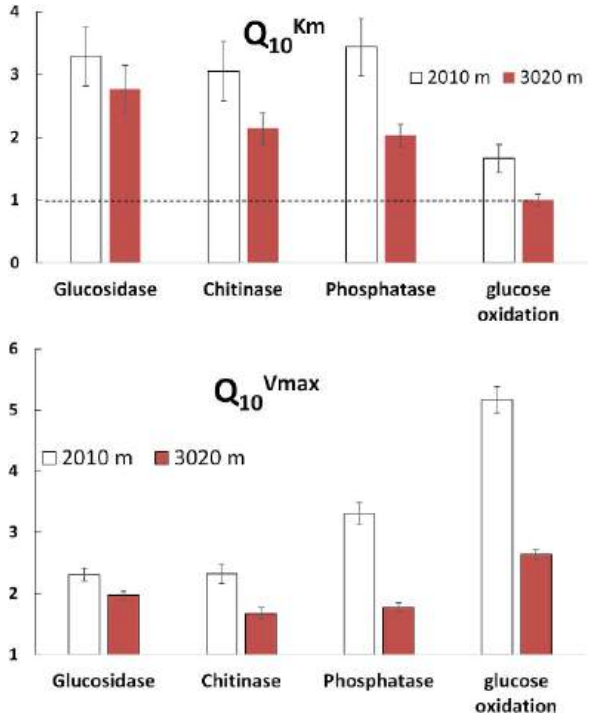


Figure 32.12.6: The values of $\mathbf{Q_{10}^{Km}}$ (**a**) and $\mathbf{Q_{10}^{Vmax}}$ (**b**) for hydrolytic reactions and for reactions of glucose oxidation at low and high altitudes. Bars show standard deviations of the means (n = 3).

The figure shows that both Q_{10}^{Km} and Q_{10}^{Vmax} were lower at high altitudes. Hence the enzymes in organisms from higher altitudes did not respond as strongly to temperature changes. This again suggests that larger compensatory changes are found in enzymes at high latitudes, allowing them to better adapt to the greater temperature range they would experience.

These data suggest that thermal adaptions in the intracellular enzyme are driven more by a large range of temperatures experienced by the organisms and not the mean temperature. Compensatory and canceling changes in K_M and V_M at low substrate concentration led to a higher S_{crit} in cold-adapted organisms.

STRUCTURAL MECHANISMS FOR ENZYME TEMPERATURE ADAPTATION

Structure determines function. A detailed understanding of how proteins, and more directly enzymes, adapt to temperature changes must come from detailed structural analyses that can be correlated to functions such as enzyme catalytic activity. Two approaches have been used to study the structural bases of enzyme temperature adaptation. One involves structural analysis of a single enzyme in organisms adapted to different temperatures. The other involves large computational analyses of databases of protein structure. We'll discuss both. First, let's explore orthologs (in this case, a protein from the same gene in different species) of a single enzyme, **ketosteroid isomerase (KSI),** from mesophilic (grown in moderate temperatures) and thermophilic (grown in warm temperatures) bacteria.

Figure 32.12.7 below shows structural and functional features that document KSIs temperature adaptation through changes in activity and stability. Let's step through each of the panels in order.

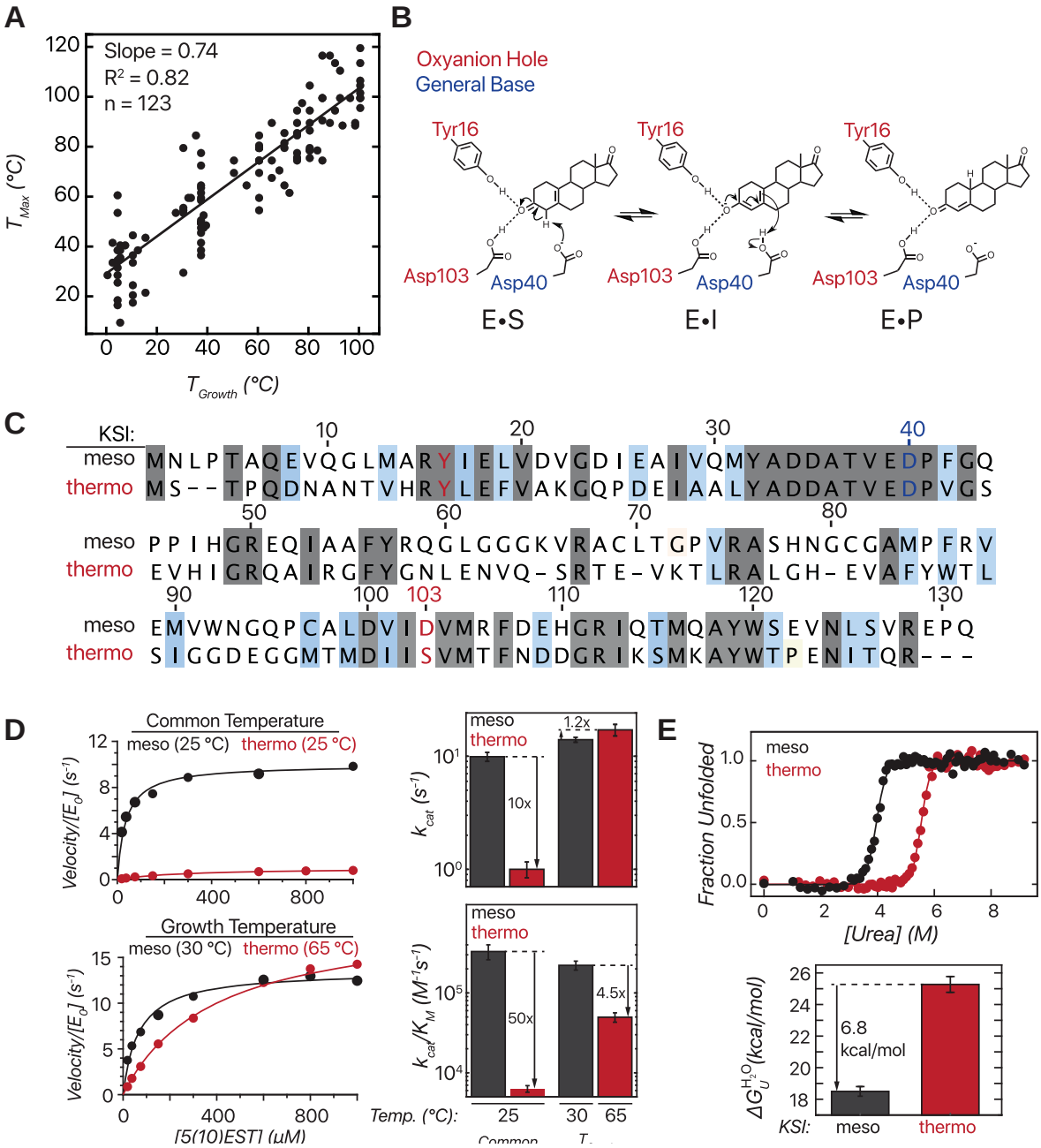


Figure 32.12.7: Enzyme temperature adaptation through changes in activity and stability. Pinney et al., Science, 371. (2021). DOI: 10.1126/science.aay2784. Reprinted with permission of the American Association for the Advancement of Science and Copyright Clearance Center.

Panel (**A**) shows that as the optimal growth temperature for an organism increases, the optimal temperature for selected enzyme activity also increases. Each dot represents a different species with the enzymes broadly chosen across all enzyme commission classes (i.e. the dots do not represent just KSI). There is a strong linear correlation.

Starting with Panel (**B**), we look at KSIs. Panel B shows the mechanism of isomerization of the steroid substrate, 5(10)-estren-3,17-dione [5(10)EST] by KSI. KSI has one of the highest k_{cat} values of enzymes. The reaction changes the position of the C=C double bonds and proceeds through an enolate/oxyanion intermediate (EI) formed after the abstraction of a proton by Asp 40. The transition state, which would have a developing negative charge on the O atom, is stabilized by proximal Tyr 16 and Asp 103 in a developing oxyanion hole.

Panel (**C**) shows the KSI sequences from *P. putida* (a mesoKSI) and *M. hassiacum* (thermoKSI). The sequences are 33% identical, but some key resides (gray) are identical. Similar ones are shown in blue. The thermophilic KSI (thermKSI) has Ser 103 instead of the often conserved Asp 103 in mesophilic organisms (mesoKSI). D103 and S103 are shown in red.

Now let's look at **Panel (D),** which shows the activity (v_0/E_0) vs substrate [5(10)EST], for the meso- and thermophilic enzymes. At a nominal temperature (25^0 C, top left panel), the thermoKSI shows little activity. At their optimal growth temperatures (30^0 C for the mesoKSI and 65^0 C for thermoKSI, bottom left), the thermoKSI has both a higher $V_M=k_{cat}E_0$, and K_M at 65^0 C. The higher K_M is consistent with the idea that K_m values are usually higher at higher temperatures. The derived k_{cat} and k_{cat}/K_M values are shown in the adjacent histogram. Remember that k_{cat} is a measure of how many bound substrate molecules are converted to product per sec (at saturating substrate concentrations). The parameter k_{cat}/K_M is a measure of the effective biomolecular rate constant for product formation at low substrate concentrations ([S] $<<K_M$).

From an evolutionary perspective, early ancestral enzymes probably arose in warmer environments. When the earth cooled, enzymes had to "solve" the problem easily apparent from the Arrhenius equation (the rate of reactions decreases with decreasing temperatures). Hence evolution would lead to structural changes that would facilitate either higher k_{cat} values, lower K_M values, or both, at lower temperatures. We've seen in the examples above that K_M decreases with decreasing temperatures, so evolutionary pressures would more likely lead to increased k_{cat} values. This is consistent with localized changes in protein dynamics as modulators of k_{cat}. The evolutionary pressure to maintain stability would be low since proteins become more stable at lower temperatures.

Now the question arises, how thermally stable are the meso- and thermKSI? Urea denaturation experiments were used to determine the $\Delta G^{0\ H2O}{}_U$ for unfolding of the enzyme, as described in Chapter 4.12. Panel (**E**) shows urea denaturation curves for mesoKSI (black) and thermoKSI (red) monitored by changes in internal tryptophan fluorescence (top) and stabilities extrapolated to 0 M urea ($\Delta G^{H2O}{}_U$) (bottom). The denaturation curve for thermoKSI is significantly shifted to the right. The calculated values for $\Delta G^{0\ H2O}$ unfolding are +18.5 kcal/mol (mesoKSI)and about +25.3 (thermoKSI), a +6.8 kcal/mol difference. These values would be - 18.5 kcal/mol (mesoKSI) and about -25.3 (thermoKSI) for the reverse folding reaction (unfolded ↔ folded transition).

Figure 32.12.8 shows an interactive iCn3D models of the bacterial ketosteroid isomerase from the mesophilic bacteria Pseudomonas Putida (left) and from the thermophilic bacteria Mycobacterium hassiacum (right).

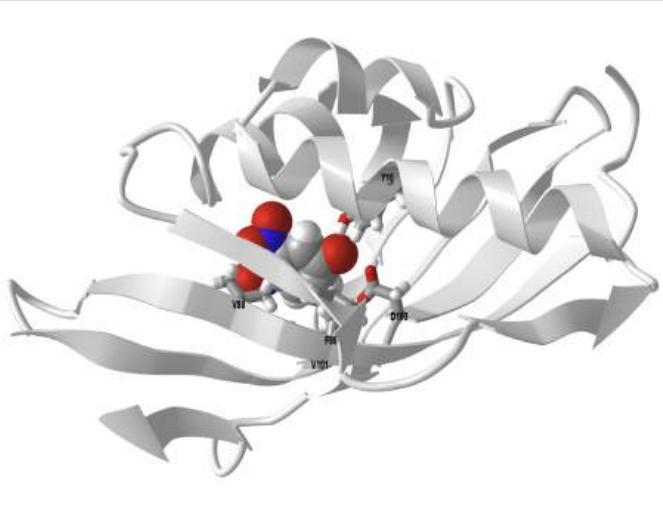


Figure 32.12.8*a*: Mesophilic Ketosteroid Isomerase D40N mutant (monomer) from Pseudomonas Putida (pKSI) bound to 3,4-dinitrophenol (**6C17**). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...kH877KZKPxWsGA

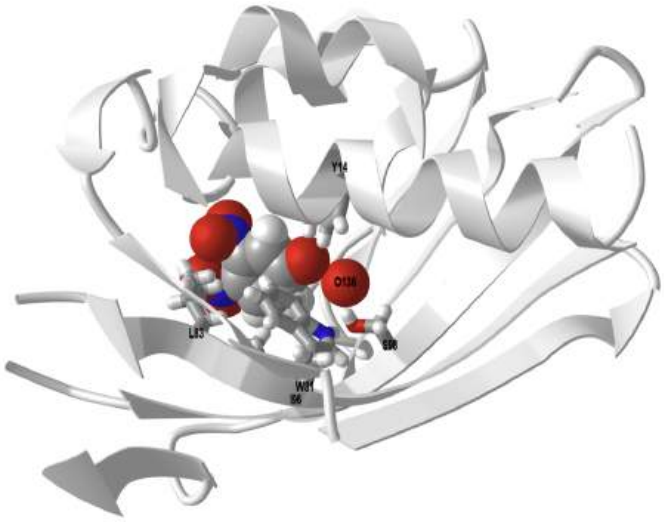


Figure 32.12.8*b*: Thermophilic Ketosteroid Isomerase D38N mutant (monomer) from Mycobacterium hassiacum (mhKSI) bound to 3,4-dinitrophenol (**6P44**). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...YbGw5zr2dRpnK8

Both structures have a bound 3,4-dinitrophenolate, a stable oxyanion and transition state analog. The corresponding active site side chains are shown in color sticks. Just one monomer of the dimer for each protein is shown for clarity. The key active-site residues F86, V88, V101, and **D103** in mesoKSI are replaced by W86, L88, I101, and **S103** in thermoKSI (all very conservative except for D103S). A water molecule forms a bridging hydrogen bond from S103 to the oxyanion in the phenolate.

Figure 32.12.9 below shows the key change in Asp 103 (see the mechanism in Fig. 7B) in mesoKSI to Ser 103 + H_2O in thermoKSI

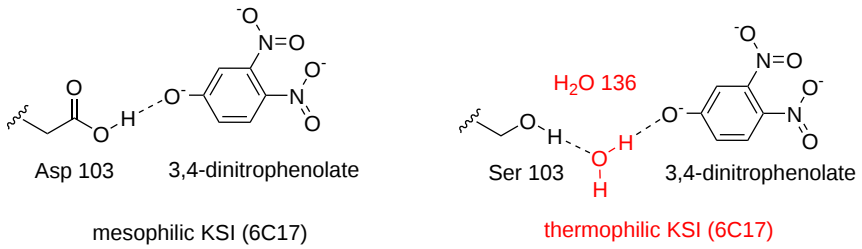


Figure 32.12.9: Change from Asp 103 in mesoKSI to Ser 103 + H_2O in thermoKSI

Asp 103 is initially protonated with a pKa is >9, much higher than in an aqueous solution (3.7). The hydrogen bond from the mesoKSI D103 to the phenolate is stronger than the bridging one from Ser 103 in the thermoKSI. This arises from the increased polarity of the OH on a carboxylic acid (Asp) compared to an alcohol (Ser). In addition, the hydrogen bond distance from the bound water to the phenolate is longer than from Asp 103. Hence the Asp 103 in the thermoKSI improves enzymatic stability. In contrast, at lower temperatures, thermoKSI is less active, but at higher temperature it is more stable. The role of Ser 103 in stabilizing the folded state is shown in Figure 32.12.10 below

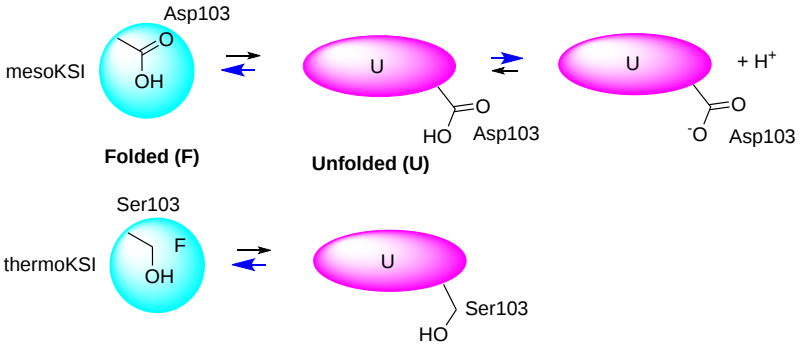


Figure 32.12.10: Roles of Asp 103 and Ser 103 in the folding to unfolding transition of KSI

When mesoKSI unfolds, Asp 103, which is protonated and not accessible to solvent in the native state, becomes solvent exposed on denaturation. Its pka drops, which leads to its deprotonation. This extra deprotonation step stabilizes the unfolded state (pulling the reaction to the right), making thermoKSI less thermally stable. In contrast, Ser 103, on solvent exposure, does not deprotonate, so the unfolded state is not additionally stabilized. Hence Ser 103 leads to greater stability of the folded thermoKSI. The D103/S103 change is found in many KSI from many bacteria.

Here are some additional findings:

- The structures of mesoKSI and thermoKSI are highly similar even though they have only 33% sequence identity. Figure 32.12.11 shows the conformational changes in the monomeric ketosteroid isomerase (KSI) going from the mesophilic enzyme (6C17, cyan) to the thermophilic one (6P44, magenta) enzymes

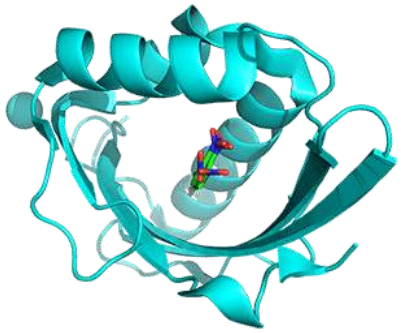


Figure 32.12.11: Conformational changes in the monomeric KSI going from the mesophilic enzyme (6C17, cyan) to the thermophilic one (6P44, magenta) enzyme

- The F86W change in the thermoKSI stabilizes the conformation of S103 to maximize its stabilization of the oxyanion in the oxyanion hole and allows high-temperature activity.
- Other KSI orthologs and mutants show higher enzyme activity if they have D103.

- Changing key residues at 86, 88, and 101 in mesoKSI to those found in the thermoKSI additionally increased the stability of the mesoKSI
- Analysis of 1140 KSIs showed that the fraction containing D103 decreases with increasing growth temperature, and the fraction containing S103 increases.

It appears that the thermal adaption of KSI occurs mostly through one amino acid change (D103S). The change reduces k_{cat} for the mesoKSI 10x at low temperatures but greatly increases stability at high temperatures.

This switch and compensatory changes in activity and stability suggest that protonated Glu and Asp side chains involved in activity might be changed to other amino acids that confer more stability but reduce activity. A conserved and protonated active site Glu is found in glycosidases from high-temperature orthologs. Likewise, a protonated Asp side chain is found in thioredoxin. As a control, a change from a protonated Glu in triosphosphate isomerase distal to the active site to a Gln shows no effect on catalysis and was not found in thermophiles. It should be noted that not all stabilizing mutations decrease activity since many examples are known that don't. 67 protonated Asp and Glu side chains were identified in the PDB, 14 of which were replaced in high-temperature orthologs.

Conservation of chaining pairs of amino acids between mesophilic and thermophilic organisms.

Are there other broadly found changes in amino acids (in addition to Asp/Glu) at each position in a mesophilic protein and its thermophilic ortholog? Computational analysis in 2194 enzyme families in 5582 bacterial species (for a total of 17 million amino acid pairs) were performed to explore this question. Half of all families had an amino acid at a given position which correlated with growth temperature, resulting in almost 160,000 key positions. The results of this study are broadly outlined in Figure 32.12.12 below.

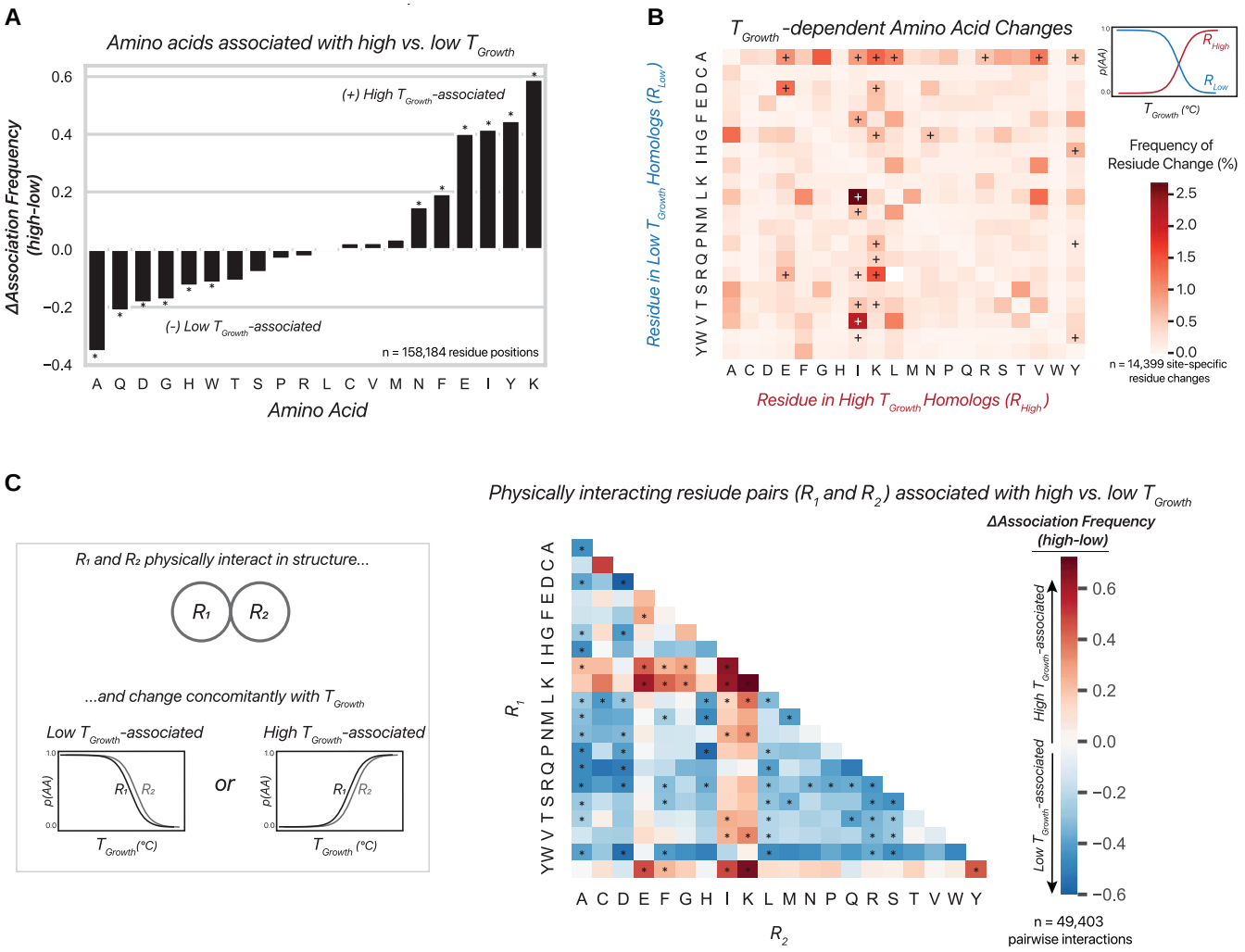


Figure 32.12.12 : Examination of temperature-associated residues and their interactions. Pinney et al., ibid. Reprinted with permission of the American Association for the Advancement of Science and Copyright Clearance Center.

Let's explain each part of this complex figure.

Panel (**A**) - Preferences in the types of amino acids associated with high or low-temperature growth: Panel (**A**) shows the difference in frequency of association of given amino acids with high to low temperatures. Some amino acids were associated with low-temperature growth. These include alanine (A), glutamine (Q), and aspartate (D), as in KSI). Others were associated with high-temperature growth (E, I, Y and K). F

Panel (**B**) - Identities and frequencies of site-specific residue changes across temperature growth conditions: Now let's see how a specific acid changes in going from mesophilic (low T) to thermophilic (high T), which show key temperature-dependent amino acid changes. Panel (**B**) shows the frequency of observed site-specific changes in temperature-associated residues. The "+" indicates that the frequency of a change (such as Leu to Ile) is significant in comparison to the opposite change (Ile to Leu). Indeed the darkest square is from a Leu (at low temperatures) to Ile (at high temperatures change. Other darker squares show these changes occur going from low to high temperature-adapted bacteria: V → I, R → L, and D → E. The specific one found for KSI, D → S, does not stand out, most likely because of the diverse types of enzymes included in the analysis.

Panel (**C**) - Identities and frequencies of physically interacting temperature-associated residue pairs (in a single protein): If one amino acid changes in going from a low to a high-temperature ortholog, it is likely that the original and changed amino acids physically contact different nearby amino acids in their respective protein. Panel (**C**) shows the difference in frequency of association with high vs. low-temperature growth bacterial enzymes for "all possible physically interacting pairs of residues (made up of residues R_1 and R_2) that change concomitantly with *the* growth temperature". The darker blue squares show interacting residue pairs found more often in low-growth temperature enzymes, while darker maroon squares show contacting pairs in high-growth temperature proteins. The asterisk shows pairs are statistically significant.

One of the darkest maroon squares (high-temperature enzymes) shows an Ile-Ile interaction pair. The vertical column above Ile (high growth) in Panel B shows that changes to Ile in high-temperature enzymes occur frequently from 6 different amino acids found in low-temperature enzymes. Ile might be favored over Val to maximize buried hydrophobic surface area. Also, compared to Leu, it has great conformation flexibility and could better pack empty spaces in protein interiors.

A Lys to Glu (K to E) interaction is strongly associated with high-temperature enzymes, while an Arg to Asp (R to D) interaction is strongly associated with the low-temperature enzyme. This implies that simply increasing the number of salt bridges (ion-ion interactions) does not make a protein more thermally adaptable. Lys salt bridges would have greater conformational flexibility than those using Arg. The same applies to Glu compared to Asp. Lys also has the largest hydrophobic surface area, which could enhance hydrophobic packing. Since Arg has hydrogen bond donors requiring more adjacent hydrogen bond acceptors, Arg use might depend more on adjacent amino acids and not just a binary pair.

In summary, these results are more nuanced than previous explanations for high-temperature stability:

- increasing branched chain residues like Ile, Leu, and Val in the packed hydrophobic core. Indeed, as seen in Panel B, the most frequent amino acid change observed are from Leu/Val in low-temperature growth orthologs to Ile in high-temperature growth ones. These are all branched-chain amino acids. They occur 2-3x more frequently than the reverse, Ile in low-temperature growth to Leu/Val in high-temperature growth enzymes. In fact, Panel C shows that Ile preferentially interacts with another Ile in high temperature adapted enzyme. Hence stability is not just improved by substitution with any hydrophobic side chain.
- increasing number of salt bridges (ion-ion interactions) and hydrogen-bonding interaction charged side chains. There are more charge side chains and salt bridged in thermophilic proteins. However, the above data shows a clear preference for Lys in thermophilic proteins as changes from Arg to Lys are common in that group as shown in Panel B. Likewise, Asp to Glu changes from low to high temperature-adapted proteins are 3 times more probably than the reverse. In addition, Panel C shows that interactions between Lys and Glu are most strongly associated with high temperature-adapted proteins, while Arg and asp interactions are most often in low temperature-adapted proteins.

ALLOSTERISM AND THERMAL CHANGES

As we mentioned before, increasing temperature can cause local effects in a protein, instead of large global changes that obviously lead to denaturation at high temperatures. Changes occurring in the active site can clearly affect k_{cat}, K_M, and k_{cat}/K_M. Allosteric effects also occur. The protein imidazole glycerol phosphate synthase (HisFH) is a heterodimer with two active sites, so it is considered a bienzyme. The H subunit is a glutaminase, which cleaves glutamine into glutamine acids and ammonia, which can diffuse through a channel to the active site of the F subunit, the cyclase. The active form of the enzyme occurs only when both substrates are bound resulting on long-range allosteric activations. In particular, the oxyanion hole in the H subunit is formed in the activated form which allows the stabilization of the tetrahedral transition intermediate in the hydrolysis of glutamine.

Figure 32.12.13 shows an interactive iCn3D models of the heterodimeric imidazole glycerol phosphate synthase complex (7AC8)

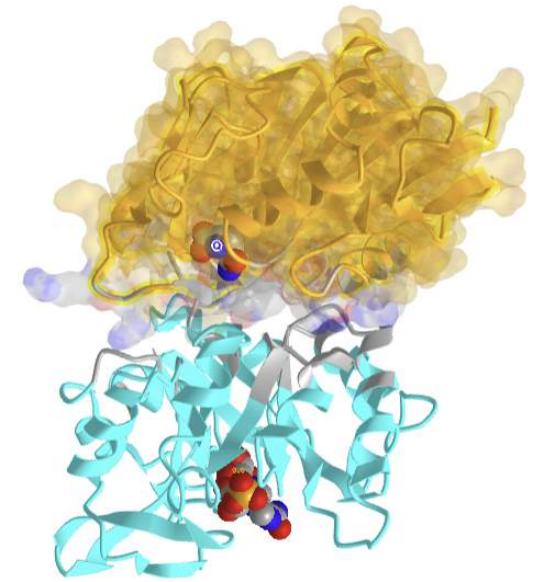


Figure 32.12.13: (Copyright; author via source). Heterodimeric imidazole glycerol phosphate synthase complex (7AC8). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...S5MbiiA9bDvDQ8.

The HisH (glutaminase) subunit is shown in orange with a molecular surface and the bound substrate (spacefill), glutamine. The HisF subunit is shown in cyan with a bound substrate.

An allosteric effector molecule can bind in the active site of His F and induce long-range conformational changes in this HisH active site which increases its activity 5000x. The structures involved are shown in Figure 32.12.14 below.

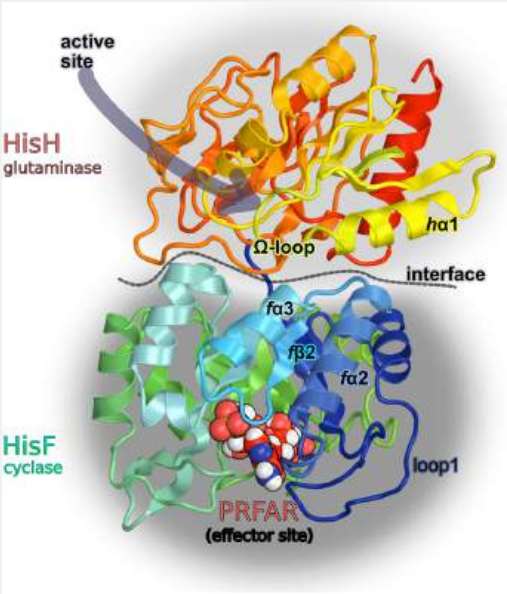


Figure 32.12.14: Imidazole glycerol phosphate synthase (IGPS) from the thermophile *Thermotoga maritima* (*T. maritime*). Maschietto, F., Morzan, U.N., Tofoleanu, F. *et al.* Turning up the heat mimics allosteric signaling in imidazole-glycerol phosphate synthase. *Nat Commun* **14**, 2239 (2023). https://doi.org/10.1038/s41467-023-37956-1. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

IThe HisF and HisH subunits are colored respectively in green-to-blue and red-to-yellow gradients, respectively, and separated by a dotted line which marks the interface between HisF and HisH. The labels (*f*α2, *f*α3, *f*β2, loop1, *h*α1, Ω-loop) indicate secondary structure elements that are directly involved in the allosteric regulation.

Molecular dynamics and NMR studies have shown that increases in the temperature lead to conformational changes resembling those that occur on binding of the allosteric effector at room temperature. As the temperature of the apoenzyme (no substrates or effects bound) increases from 30 °C to 50 °C, the dynamics and structure increasingly resemble the state induced by the effector. Increasing

temperatures from climate change are likely to cause subtle conformational and dynamic changes in all proteins with some having negative consequences.

THERMAL DETERMINANTS OF YEAST METABOLISM

With this basic background, we can attempt to understand the thermal determinants for entire metabolic pathways. This has been attempted in the yeast *Saccharomyces cerevisiae*, a eukaryotic organism with optimal growth around ~30 °C, extremely limited growth at 40 °C, and no growth/death at 42 °C.

Mathematical modeling all metabolic pathways in a cell is a daunting task. Accurate concentrations, rate constants, and dissociation constants are needed for all reactions. **Genome-scale metabolic models** (GEM) use a multitude of constants that are experimentally or computationally determined. There are usually significant uncertainties in the parameters used in the model. Bayesian statistics has been used to decrease these uncertainties. In Bayesian statistics, parameters and models are updated with the known values and information. It is similar to machine learning models, which uses data to train the model and refine it.

A Bayesian model for S. cerevisiae was used as it is the most abundantly used organism in industry and has many GEMs. The GEM used was the **enzyme-constrained GEM** (**ecGEM**). It was then further developed into the **enzyme and temperature-constrained GEM** (**etcGEM),** which, in addition, incorporates the temperature dependence of both the concentration of the native enzyme (E_N) and kcat for the enzyme. For each enzyme, the melting point (T_M), the change in heat capacity (ΔCp) for the transition state, and the optimal temperature (T_{opt}) were included. We discussed both (T_M) and the change in heat capacity (ΔCp) for proteins in Chapter 4.9. In addition, another term for **non-growth associated maintenance (NGAM)** of the cells, which is also temperature-dependent, was included. Examples of non-growth associated maintenance include maintaining membrane potential, turgor pressure, normal protein refolding and DNA repair.

The Bayesian models reproduced the datasets well. Using the models, key enzymes that control the flux through metabolic pathways were determined at each temperature. The most rate-limiting enzyme at superoptimal temperature in yeast was squalene epoxidase (ERG1), found in sterol metabolism pathways. Replacing the S. cerevisiae enzyme with one from a thermotolerant yeast strain led to better growth than the wild-type cells.

Figure 32.12.15 below the complexity and extent of the metabolic proteins and pathways of the enzyme-constrained GEM (ecGEM).

Figure \(\PageIndex{15}\): Metabolic proteins and pathways included in the enzyme-constrained GEM. Benjamín J Sánchez et al., Molecular Systems Biology (2017)13:935. https://doi.org/10.15252/msb.20167411. Creative Commons CC BY

(add soon: ecYeast7 model (both constrained and unconstrained; each as.mat,.sbml and.txt files): GitHub (https://github.com/SysBioChalmers/GECKO/tree/v1.0/Models)

Now, let's explore some of the results of the study. Figure 32.12.16 below shows how the temperature dependencies of proteins were incorporated into the enzyme-constrained GEMs.

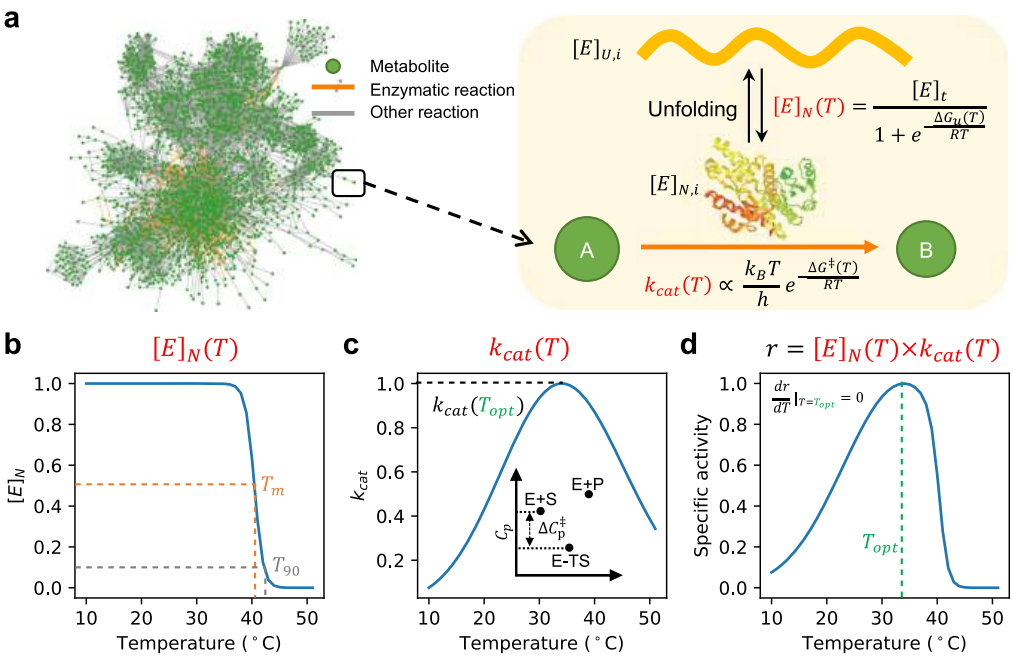


Figure 32.12.16 : **Using Bayesian statistical learning to integrate temperature dependence in enzyme-constrained GEMs.** Li, G et al., *Nat Commun* **12**, 190 (2021). https://doi.org/10.1038/s41467-020-20338-2. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Panel a shows the complexity of the metabolic network used to produce the ecYeast7.6, shown in greater detail in Fig. 13 above, and how E_N and k_{cat} depend on temperature.

Panel b shows how a two-state denaturation model was used. $[E]_N$ is the enzyme concentration in the native state; T_{opt} is the optimal temperature at which the specific activity is maximized; T_m and T_{90} are temperatures at which there is a 50 and 90% probability that an enzyme is in the denatured state, respectively. In Chapter 4.4 we indicated that for a reversible two-state transition, T_M is the temperature at which half of the enzyme is native, and half denatured, giving a K_{eq} for the N ↔ D equilibrium of 1.

Panel c shows how k_{cat} depends on temperature. The insert show how the heat capacity change from E+S to the E-transition state. We have previously seen that the $\Delta C_P > 0$ for protein unfolding, and this $+\Delta C_P$ value is a signature of the hydrophobic and occurs when nonpolar groups become more solvent exposed. We have shown previously that enzymes bind the transition state more tightly than the substrate. This $-\Delta C_P$ is more in line with the latter.

Panel d shows the temperature dependence of enzyme's **specific activity, r**, as a function of temperature, which is determined by E_N (panel B) and k_{cat} (panel C).

Now we can explore the outputs of the ecGEM run with the separate added effects of temperature on NGAM, k_{cat}, and E_N (denaturation). Finally, the combined etcGEM was run. The combined etcGEM was able to produce the observed outcomes in yeast growth. Given that the contributions of each of the three factors, NGAM, k_{cat}, and protein denaturation, to whole-cell growth could be modeled. The outcomes from these models as a function of temperature are shown in Figure 32.12.17. They support the notion that the growth rate of yeast is explained by temperature effects on its enzymes.

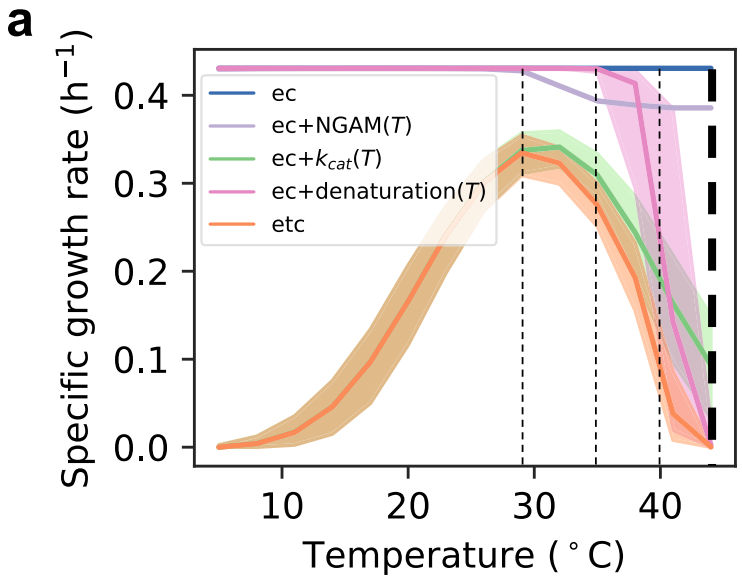


Figure 32.12.17 : **Fig. 3: Yeast growth rate is explained by temperature effects on its enzymes.** Li, G et al., ibid

This figure shows how the temperature dependence of different processes combines to affect the growth rate. EC is the prediction from the enzyme contained GEM, —predictions with the enzyme-constrained model; ec+NGAM includes temperature effects on nongrowth associated maintenance; ec+$k_{cat}(T)$ incorporates the temperature effects on enzyme k_{cat} values; ec+denaturation(T) incorporates the temperature effects on enzyme denaturation. Finally, in the etc model, the enzyme and temperature-constrained model incorporates temperature effects on all three (NGAM, k_{cat} and enzyme denaturation) into ec model. The solid lines indicate median values and shaded areas indicate regions between the 5th and 95th percentiles (n = 100).

Here are some summarized results:

- <29 °C, only temperature-dependent k_{cat} affected the cell growth rate (green line under the orange etc line <29 °C;
- at 29 °C <T < 35 °C, both k_{cat} and NGAM (gray line) determined the growth rate;
- at T > 35 °C, enzyme denaturation at major effects and by 40 °C was dominant,

Figure $32.12.18 d-e$ shows using images how the temperature dependencies of the factors E_N, k_{cat} and r (specific activity) on yeast growth. A phenomenal amount of data is displayed in these images.

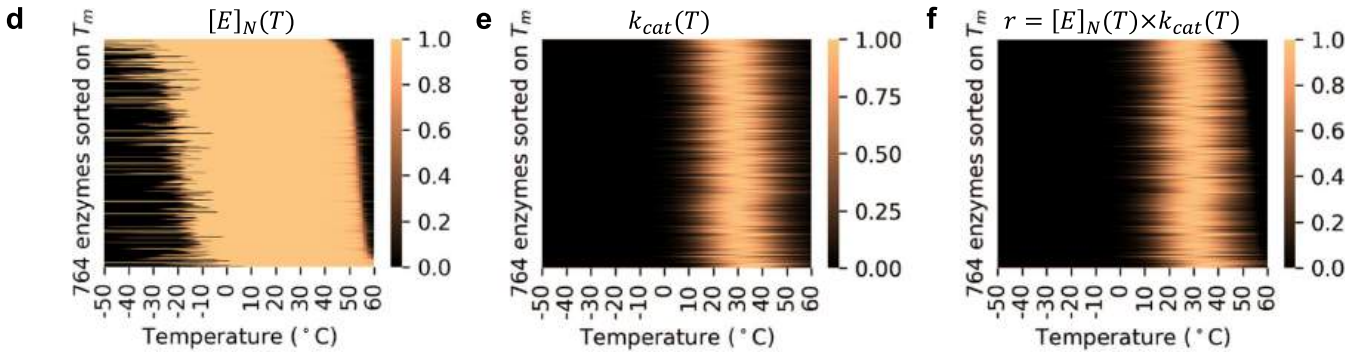


Figure $32.12.18 d-e$: **Fig. 3: Yeast growth rate is explained by temperature effects on its enzymes.** Li, G et al., ibid

Panel d shows the probability that a given enzyme is in the native state. Think of the y-axis as increasing 1 pixel at a time from bottom to top, with each new added pixel representing a different enzyme for a total of 764 enzymes along the y-axis. The x-axis shows with each pixel at a given y-axis if the enzyme is native that temperature

The interface between bright (native) and black (denatured) on the right side of the image shows that some enzymes (top) become unfolded at 40^0C, while some don't unfold until close to 60^0C

Panel e shows normalized k_{cat} values of 764 enzymes at different temperatures. The brightest white pixels show the highest k_{cat} values. The image clearly shows the brightest vertical band at around 30^0C, the optimal growth temperature of yeast.

Panel f shows the normalized specific activities (r) of 764 enzymes at different temperatures. Again the highest specific activity is centered around 30^0C.

Note in Fig. 15d that most enzymes also denature at temperatures < -10 ^{0}C, but cells were not viable under those conditions.

The etcGEM was able to replicate a finding that above 37 °C, yeast cells switch from respiration to partial fermentation accompanied by a larger flux through glycolysis. This occurred because of a decrease in specific activities of enzymes with increasing temperature, which constrains metabolism. In addition, the total protein concentration reaches a limit and can't increase further, which could have increased enzyme activity. Along with an increase in glycolysis, mitochondrial ATP production decreases. Respiration produces more ATP per mole of glucose, but glycolysis/fermentation produces more ATP/protein mass so when protein concentration reaches a maximum, glycolysis is more efficient.

Lastly, the etcGEM was used to find the enzymes whose flux changed most at superoptimal temperatures. (We introduced the flux control coefficient in Chapter 14.3). The results of this modeling are shown in Figure 32.12.19 below.

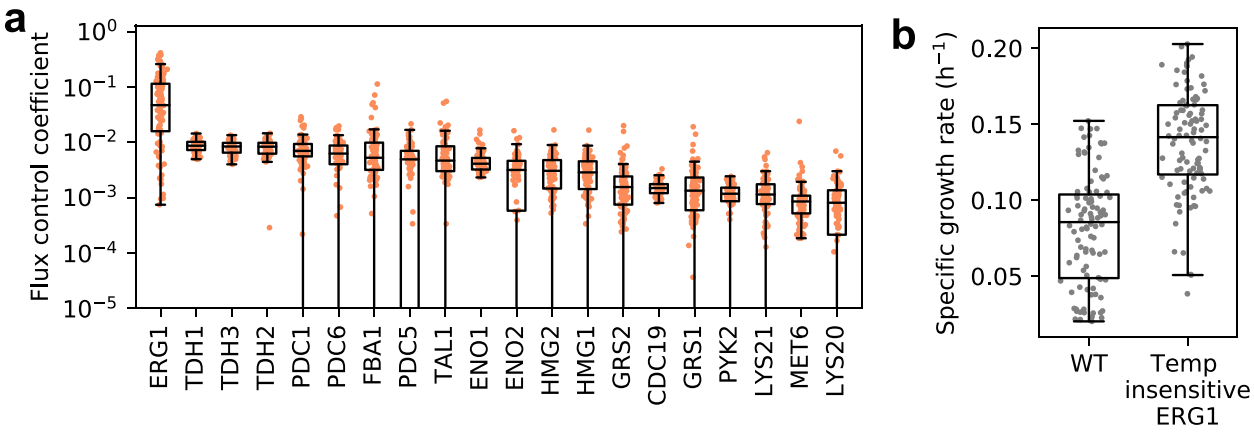


Figure 32.12.19 : Flux control coefficients at superoptimal model for yeast.

One enzyme stands out in Panel a, ERG1, squalene epoxidase, an enzyme involved in sterol oxidation. If the wildt-type enzyme was replaced with a temperature-insensitive ERG1, the specific growth rate increased significantly (over 55%).

WHAT DOES THIS HAVE TO DO WITH CLIMATE CHANGE? LIFE'S THERMAL TOLERANCE AND LIMITS

We've just explored the molecular and metabolic adaptations that allow organisms to thrive at higher temperatures. The question is how quickly present life can adapt to global warming caused by increasing greenhouse gases. Life on the planet will adapt, but the diversity of life forms in the biosphere will change. Five mass extinction have occurred over the last 450 million years up until the present, as shown in Figure 32.12.20 below.

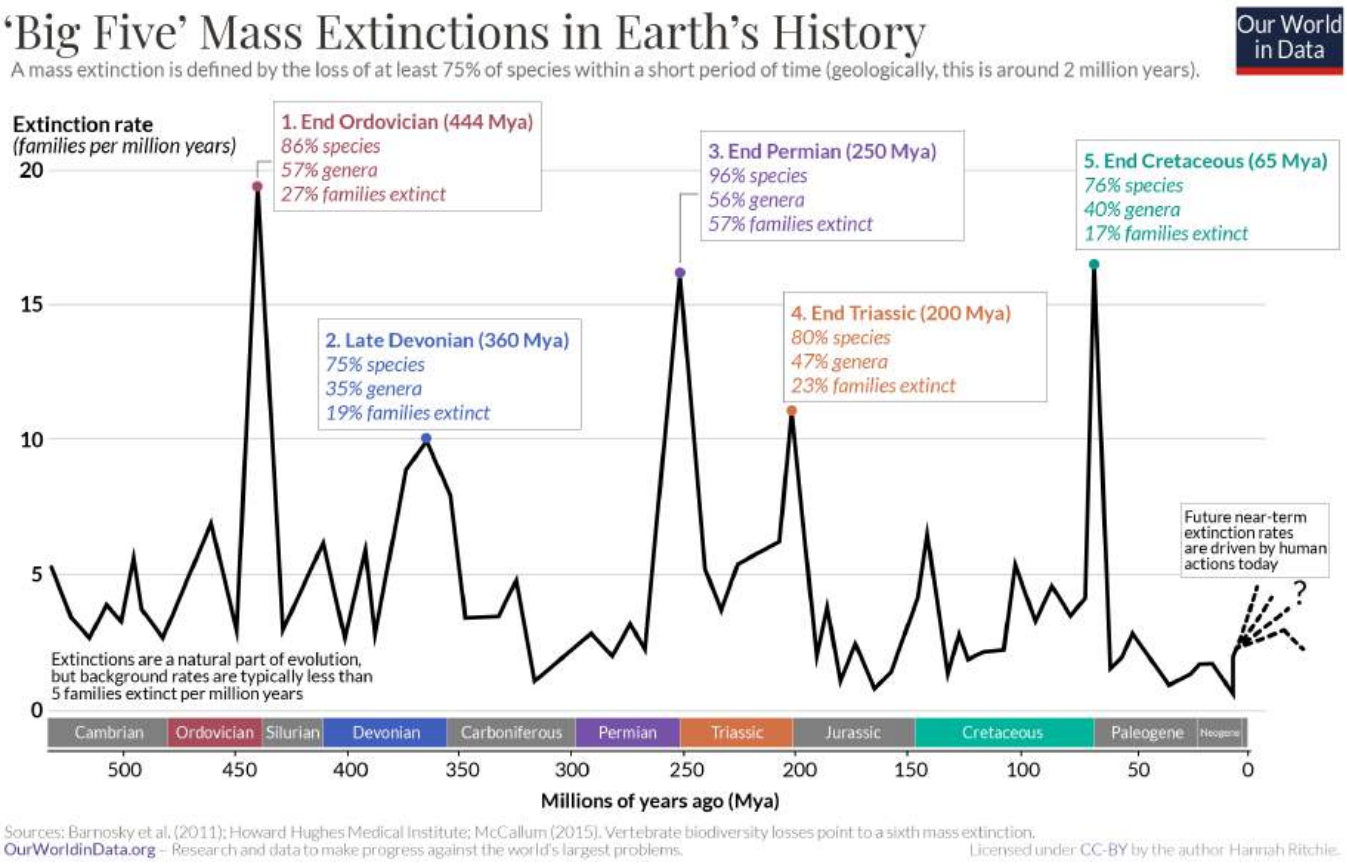


Figure 32.12.20 : Five mass extinction over the last 450 million years. Hannah Ritchie. Our World in Data. https://ourworldindata.org/mass-extinctions. Creative Commons BY license.

Many factors, often interrelated, can contribute to mass extinction. These include volcanism, asteroid impact, climate change, ocean anoxia, and the release of methane from ocean hydrates and permafrost. All are correlated with increased temperatures. as shown in Figure 32.12.21 below. It documents the correlation between temperature changes and the extinction rate for marine animals over the last 450 years. The gray bars highlight the extinction cycle.

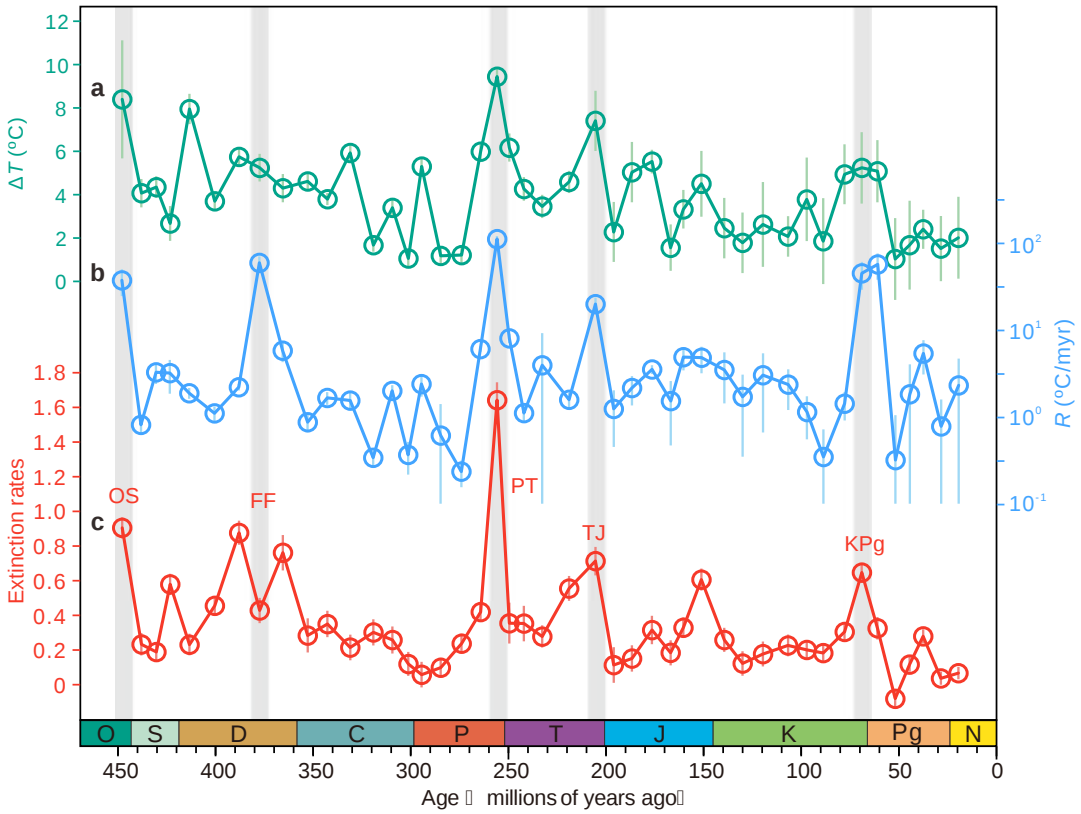


Figure 32.12.21 : **Temperature change and extinction rate over the past 450 million years.** Song, H., Kemp, D.B., Tian, L. *et al. Nat Commun* **12**, 4694 (2021). https://doi.org/10.1038/s41467-021-25019-2. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Pane a shows the largest magnitude of temperature change (Δ*T*, absolute value) in each time interval (bin)

Panel b shows the highest rate (*R*, absolute value) of temperature change in each time bin, defined at the million-year (Myr) scale.

Panel c shows the generic extinction rates of marine animals calculated using gap-filler methods using data from the Paleobiology Database. The Big Five extinctions occurred in the end-Ordovician (OS), Frasnian-Famennian transition (FF), Permian-Triassic transition (PT), Triassic-Jurassic transition (TJ), and Cretaceous-Paleogene transition (KPg). Vertical bars show mean ± 1 x standard deviation (see Methods). O Ordovician, S Silurian, D Devonian, C Carboniferous, P Permian, T Triassic, J Jurassic, K Cretaceous, Pg Paleogene, N Neogene. Dark cyan, blue, and red dots represent Δ*T*, *R*, and extinction rate, respectively.

The authors suggest that a temperature change of >5.2 °C and a rate of >10 °C/Myr would lead to a new extinction rate comparable to the "Big Five". A rise of 5.2 °C is in the upper range (but likely only with human inaction to prevent it) of IPCC projection. The present rate of temperature increase (almost 2 °C/200 years) is unprecedented in at least the last 3 million years. So it appears that we are headed to a 6th mass extinction, caused by a combination of environmental degradation, pollution and climate change created by humans.

Rothamn has suggested that we can tip over into an "official" mass extinction depending on the magnitude and the rate of change in the carbon cycle. If long-scale changes occur too quickly and organisms can't adapt, extinction follows. If the changes occur on a short time scale (as it happening now), the size of the change is a key factor. His analysis suggests that a key factor is anthroprogenic-related increases in ocean CO_2. A threshold of 310 Pg (Gt), which we could reach by 2100, could officially trigger the sixth mass extinction. However, many would argue that we are already in the 6th extinction cycle. For example, of about 30,000 terrestrial vertebrate species, about 1.7% (515 species) are on the verge of extinction (having <1000 individuals). 77 mammal and bird species have lost most of their populations.

When it comes to how organisms will adapt to our present climate change, Hochahka and Somero ask 6 relevant questions:

1. What is the thermal optima and how much change leads to suboptimal or lethal conditions?
2. What mechanisms set upper limits?
3. How close do organisms live to the limits of their thermal tolerance?
4. To what extent can organisms acclimatize to temperature increase?
5. When are genetic changes necessary for survival?
6. Does the acquisition of heat tolerance reduce cold tolerance?

More simply, we can ask if organisms can survive and thrive and at what cost. Humans most assuredly can survive a warming world, but if temperatures become too high, they will not thrive and parts of the world we become inhabitable to them without great economic, social, political, and cultural costs. Some organisms will become extinct. Mass migrations of all species will occur as they seek more habitable environments.

In an evolutionary sense, species occupy environments in which they can survive and adapt. Our climate has been fairly constant since the end of the last ice age, about 12,000 years ago. Climatic changes that make an environment suboptimal depend on how sensitive an organism is and how close it lives to its thermal limits. We have already documented the thermal limits of individual enzymes as well as whole metabolic systems

As you learned in introductory chemistry, multiple linked reactions can only go as fast as their slowest step. An analogous insight is that a chain is only as strong as its weakest link. It may be the membranes, not proteins, are the weakest link with respect to temperature adaptation as higher temperatures alter phases and subphases like rafts with lipid bilayers which in turn can disrupt the activity of membrane proteins. Membrane functions (permeability, endo- and exocytosis, and maintenance of transmembrane potentials) are also very important.

Temperature effects on neural communication in synapses might be key. An interesting example is a particular Antarctic fish, which after long exposures to 4 ^{0}C, dies if temperatures are raised to around 9^{0}C. Synaptic transmission through the acetylcholine receptor is altered as the amount of acetylcholine in the synapse increases and the rate at which it is degraded by acetylcholinesterase decreases.

The lethal temperature for an organism can depend on many factors, including the rate of temperature increase, the length of exposure, previous long-term acclimatization, and for humans, humidity. Humans can't dissipate body heat without sweating if the external temperature is >37^{0}C (average body temperature). If the humidity is high enough, sweating is ineffective in dissipating heat. Since metabolism produces heat, the body temperature can increase past 37^{0}C even if the external temperature is lower. A lethal limit for humans can occur as low as 29^{0}C (84^{0}C) at a relative humidity of 85%. Even now, about 30% of the human population is exposed to lethal thresholds at least 20 days/yr, and this will only grow as temperatures and humidity increase.

Comparisons of similar species (congeneric) that live in intertidal regions (region above the water level at low tide and underwater at high tide) vs. subtidal regions (close to a shore but always submerged) show significant differences. Organisms living in intertidal regions will experience a higher range of temperatures. 19 species of congeneric porcelain crabs exist that inhabit different latitudes and vertical depths in intertidal and subtidal regions. Intertidal crabs that experience a highly variable temperature region are exposed to temperatures much closer to their lethal temperature limit, showing the T_{lethal} - $T_{max\ habitat}$ get smaller. Heart function and neural activity at higher temperatures appear to be weak links in survival. Heat shock protein function is also different in congeners of marine snails. Lethal temperatures appear to be those at which protein synthesis dramatically decreases. Repair to damaged protein is decreased, which limits vertical migration to lower temperatures and can make organisms more susceptible to predation. These effects can occur over a small increase in temperature (a few degrees Celsius). It becomes more important to heed the warnings by the IPCC to limit global average temperature increases to 1.5 to the more likely 2.5^{0}C range.

Laboratory experiments in directed evolution show that an organism's thermal tolerance can be increased, but there are limits. E. Coli can live at a variety of temperatures and experiments to evolve them at 37 ^{0}C (human gut temperature), at 42 ^{0}C (close to their lethal limit), and in an environment with fluctuating temperatures (between 32 ^{0}C - 42 ^{0}C) have been conducted. Over 1000s of generation, the rate of evolution was highest in the 42 ^{0}C group. The group that evolved at fluctuating temperatures was better adapted to both higher and lower temperatures. Under no circumstance did E. Coli evolve into a thermophilic species, which would require novel gene function and not just a small set of mutations that allow for more optimal values of k_{cat} and K_M for enzyme catalysis.

32.13: BIOCHEMISTRY, CLIMATE CHANGE AND HUMAN HEALTH

Search Fundamentals of Biochemistry

32.13.1: INTRODUCTION

Climate change affects human health and of course the health of the biosphere to which we are inextricably linked. Many, including the US CDC, describe the concept of One Health as "a collaborative, multisectoral, and transdisciplinary approach — working at the local, regional, national, and global levels — with the goal of achieving optimal health outcomes recognizing the interconnection between people, animals, plants, and their shared environment."

We all know that pollution from the use of fossil fuels also has severe health consequences independent of effects mediated more directly by climate change. A solution to both is to dramatically decrease the use of fossil fuels and mitigate pollution from their use. People most likely do not understand the extent which climate change and fossil fuel use are linked to human health. If they did, perhaps they would become advocates for climate change action. This section will cover how climate change and fossil fuel use affect aspects of human health and diseases. In this section, we will focus on heat-related illnesses and pulmonary/cardiovascular diseases. We'll tackle climate change, emerging diseases, and pandemics in the next chapter section.

Fossil fuel use and climate change obviously affect other diseases as well, some by indirect means For instance, increases in cancer deaths will occur due to lower availability of health care arising from extreme weather disasters that impact health facilities and peoples' access to them. An increase in cancer deaths occurred during the Covid pandemic since people deferred preventive healthcare treatments as well as cancer surgeries during the pandemic. Allergic illness will increase as growing seasons lengthen and species that cause allergic reactions shift to new growth regions.

32.13.2: HEAT-RELATED ILLNESS

Heat illnesses include heat cramps, heat exhaustion, heat syncope (fainting), and heat stroke, the latter of which can quickly become fatal. Heat affects normal physiology and health. It's estimated that about 1% of all cardiovascular deaths are linked to extreme temperatures. We know that the number of warm days and the number of heat waves has increased with climate change.

Figure 32.13.1 below shows how heat waves have changed in the US in the decades from 960 to 2022 using data from 50 large metropolitan areas. The 2020s is not even a third over but already shows increases in heat waves over the previous decade.

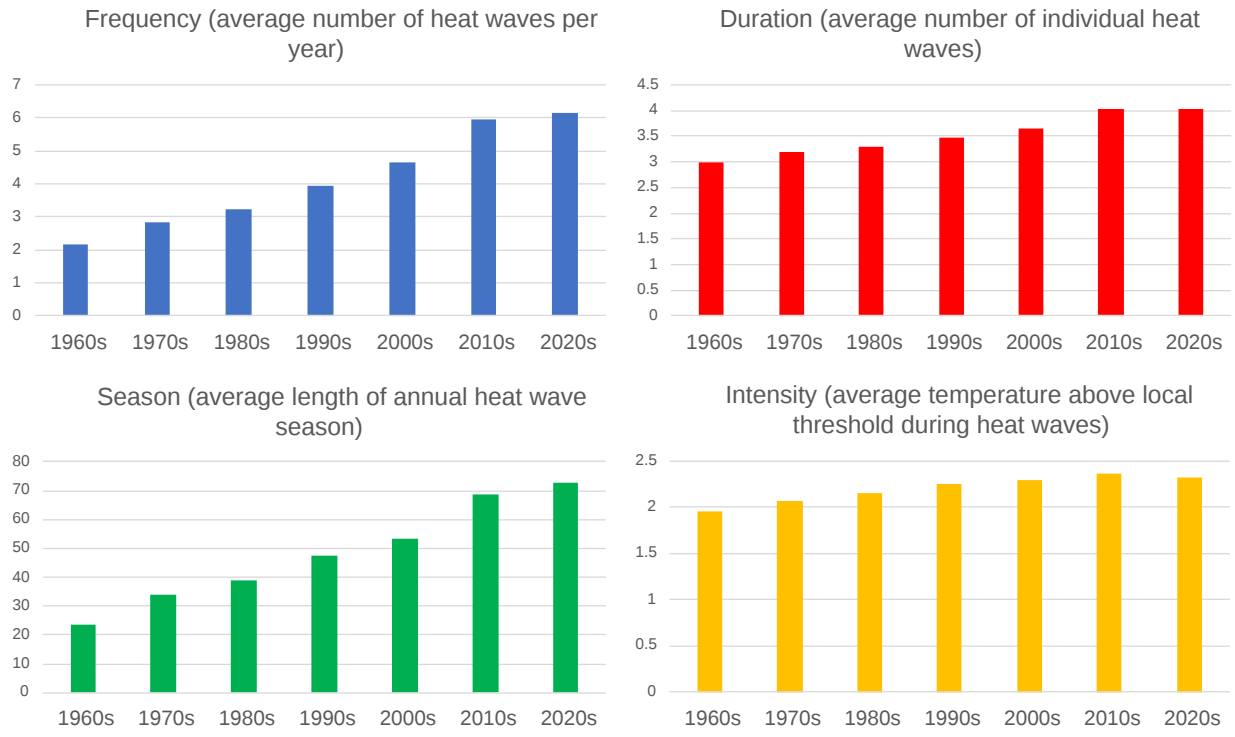


Figure 32.13.1: Climate Change Indicators: Heat Waves. EPA. https://www.epa.gov/climate-indicato...ors-heat-waves. Data source: NOAA (2022)

Many historical heat waves have occurred in recent times. A heat wave in Europe in 2003 caused 15,000 heat-exposure-related deaths in France and 70,000 throughout Europe. On June 28, 2019, France recorded a temperature of 45.9 °C (115 °F). In 2022 China experienced a two-month heat wave and drought, with a record temperature of 45 °C (113 °F) set in Chongqing. The Chicago Heat Wave of 1995, with a maximal temperature of 106 °F, caused by high temperatures and humidity, killed over 500 people. The worst might be the Russian Heat Wave of 2010 when temperatures were 5 °C (9 °F) higher than normal and reached 40 °C (104 °F). Around 55,000 people died.

Figure 32.13.2 shows temperature and excess mortality from the 2022 heat wave in Europe during which 60,000 people died from heat-related causes.

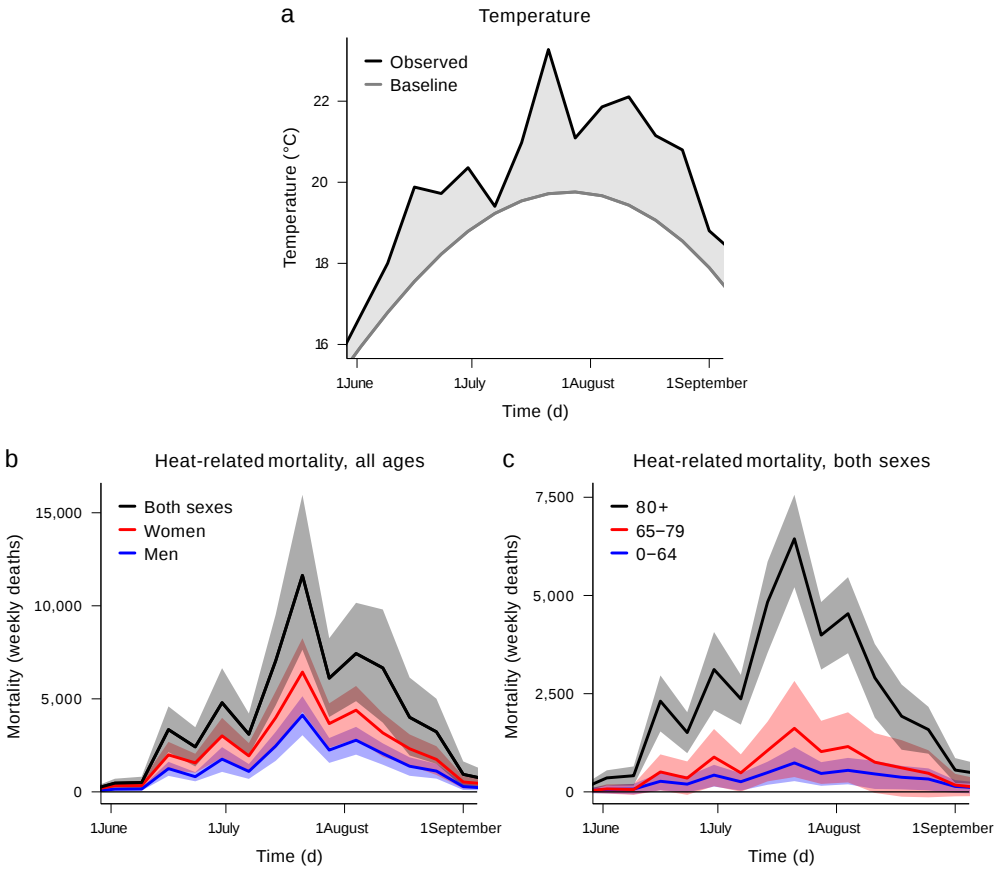


Figure 32.13.2: **Weekly temperature and heat-related mortality numbers in Europe during the summer of 2022.** Ballester, J., Quijal-Zamorano, M., Méndez Turrubiates, R.F. *et al.* Heat-related mortality in Europe during the summer of 2022. *Nat Med* (2023). https://doi.org/10.1038/s41591-023-02419-z. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/

Panel a shows the weekly baseline (gray line) and observed (black line) temperature (°C) averaged over Europe. Temperature anomalies are defined as the difference between observed and baseline temperatures (gray shading). Baseline temperatures were computed as the mean annual cycle of observed temperatures in the reference period 1991–2020.

Panels b and c show weekly heat-related mortality (weekly deaths) aggregated over Europe for the overall population (black), women (red) and men (blue) (**b**) and people aged 0–64 (blue), 65–79 (red) and 80+ (black) years (**c**), together with their 95% CIs (shadings). The numbers for women and men in **b** do not include the United Kingdom; values for the age groups in **c** do not include Germany, Ireland, and the United Kingdom.

Figure 32.13.3 below shows the steady increase in deaths as the average summer temperature in Europe has increased. The outlier in 2022 is shown as a red dot.

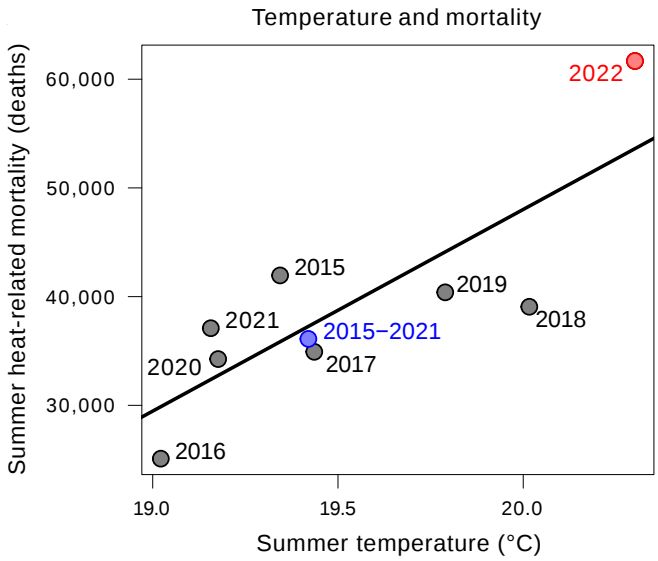


Figure 32.13.3: **The summer of 2022 within the context of rising temperatures in Europe.** Relationship between summer mean temperature (°C) and summer heat-related mortality (summer deaths) in the analyzed European countries. The straight line shows the linear fitting for the 2015–2022 period. Ballester et al., ibid.

In exercise studies under controlled conditions, the heart rate increases and plateaus after a temperature increase, but with further temperature increases, the heart rate increases with plateauing, which is a sign of cardiovascular strain. In humid conditions, cardiovascular strain develops even on slow walking at 34 °C (93.2 °F). Under dry conditions, the strain developed at around 41 °C (106 °F). The strain (as indicated by an increasing heart rate) proceeds by about 20 minutes a rise in core temperature.

With each temperature increase over pre-Industrial Revolution values, the annual probably of heat waves with apparent temperatures of 40 °C/104 °F (dangerous with a high incidence of heat cramps, heat exhaustion, and heat strokes) and 55 °C./131 °F (very dangerous with heat stoke very likely) increases, as shown in Figure 32.13.4 below.

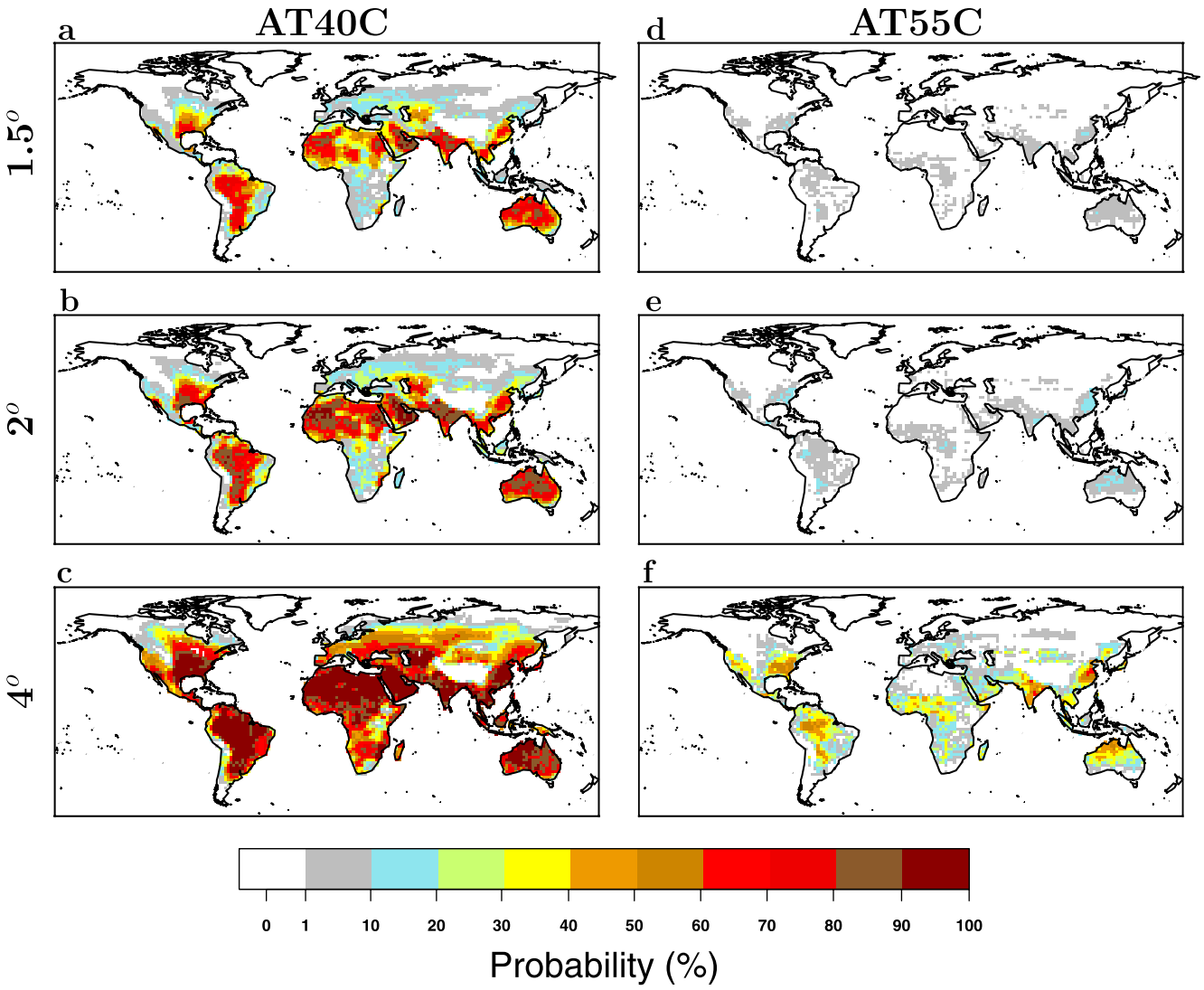


Figure 32.13.4: Annual probability of occurrence of heat waves with apparent temperature (with contributions from humidity) peaks greater than 40 °C and 55 °C. Russo, S., Sillmann, J. & Sterl, A. Humid heat waves at different warming levels. *Sci Rep* **7**, 7477 (2017). https://doi.org/10.1038/s41598-017-07536-7. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Panels (**a**–**c**) show the probability of occurrence of heat waves with $AT_{peak} \geq 40$ (AT40C) calculated at each grid point for all model years with global mean temperature anomaly relative to 1861–1880 at 1.5, 2, and 4 degrees warming (see Fig. 2), respectively.

Panels (**d**–**f**) are similar to panels (**a**–**c**) but show the occurrence of heat waves with $AT_{peak} \geq 55$ (AT55C).

32.13.3: HEAT EXPOSURE AND KIDNEY DISEASE

Cumulative exposure to high effective temperatures caused by sublethal combinations of heat and humidity leads to chronic kidney disease. This is happening with increasing frequency to workers in poor agricultural areas and others who work in such hot conditions in industry and outdoors. This type of kidney disease is not caused by diabetes, hypertension, and other known diseases of the glomeruli. Chronic kidney disease in such workers was noted in El Salvador and elsewhere in Central America. The disease has a high mortality rate. Just in El Salvador, the death from kidney disease is ten times higher than in the US. Initially, the disease was called Mesoamerican nephropathy.

Biochemical correlates of the diseases are yet unclear. Serum creatinine levels are increased, which might affect renal perfusion and lead to kidney damage. The effects might be generally cumulative or from repetitive episodes of exposure. Sugarcane field workers who report nausea, vomiting, headaches, muscle weakness, back pain, and fevers have high levels of creatinine. Kidney biopsies show inflammation and kidney fibrosis.

Similar diseases in other parts of the world that seem to have the same presentation include Sri Lankan nephropathy and Uddanam (in the Indian state of Andhra Pradesh) nephropathy. Some have categorized it as CKDu or Chronic Kidney Disease of Unknown

etiology/Uncertain cause or as chronic kidney disease of non-traditional origin (CKDnt).

Figure 32.13.5 below shows the distribution of kidney disease in the Western Hemisphere in 2019.

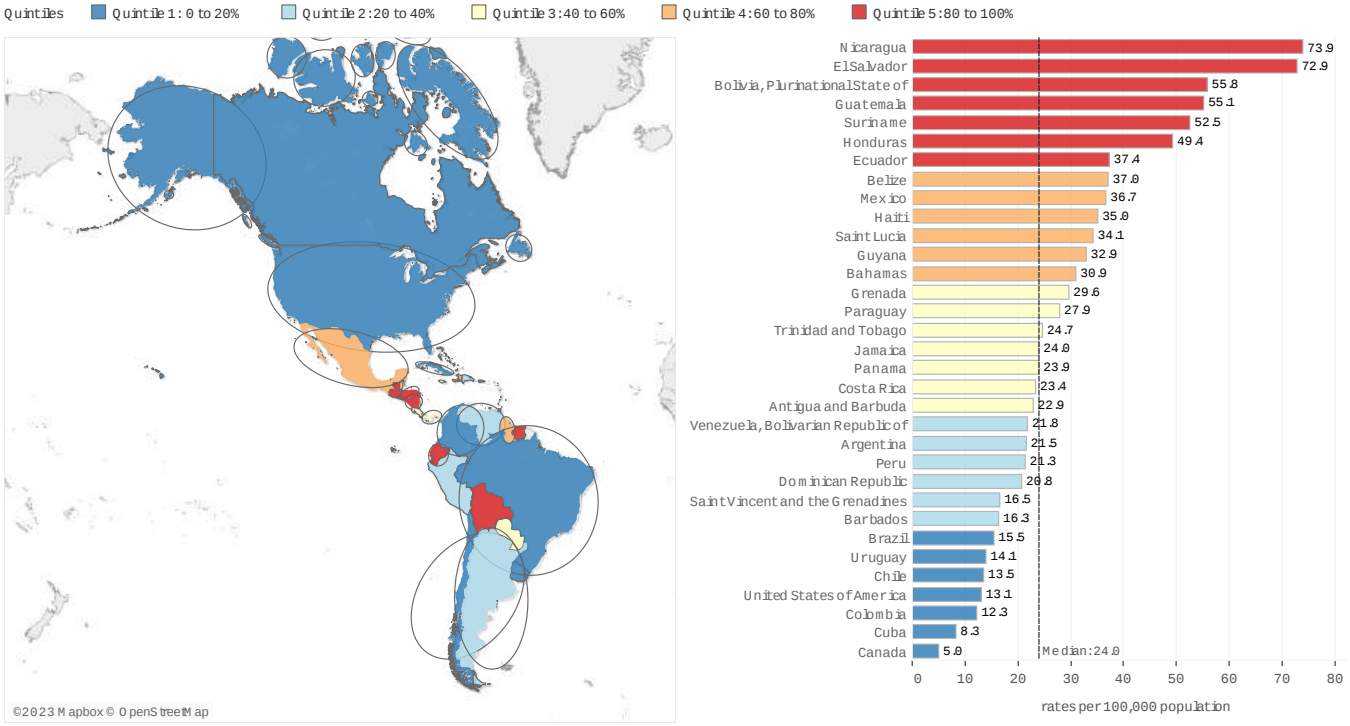


Figure 32.13.5: Burden of Kidney Disease. 2019. Men and Women. Pan American Health Organization/WHO.

Other factors such as increased exposure to herbicides, heavy metals, and microbial agents might also cause or contribute to the disease. However, the disease is most prevalent in the hotter regions of affected countries, as the incidence is lower when workers work at high altitudes. Increased incidence of chronic kidney disease also appears to be occurring in workers in Florida and California.

32.13.4: BIOCHEMICAL MECHANISM FOR HEAT STROKE

The actual biochemical mechanisms of heat stroke effects (circulatory failure, organ injury, uncontrolled clotting, death) are not fully understood. Certainly, cell death plays a major role, but not through the classical apoptotic pathway which depends on the activation of **caspases** (see Chapter 28.14). Rather, cell death occurs through **necroptosis**, a caspase-independent pathway. In necroptosis, an upstream protein kinase **RIPK3** (receptor-interacting serine/threonine protein kinase 1) activates through phosphorylation the effector protein **MLKL** (mixed lineage kinase domain-like protein). Phosphorylated-MLKL then translocates to the cell membrane where it leads to calcium influx and plasma membrane damage in the final "execution" phase of cell necrosis.

The activation of RIPK3 and MLKL through other receptors, including the toll-like receptors (TLR3 and TLR4), and tumor necrosis factor receptor 1 (TNF-R1), is shown in Figure 32.13.6 below.

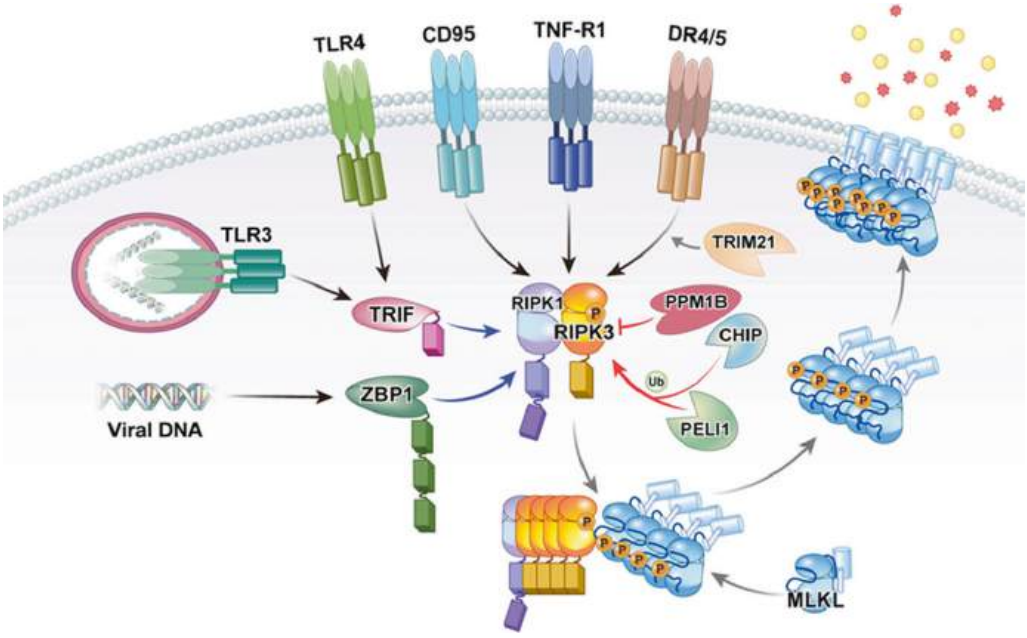


Figure 32.13.6: **Activation of RIPK3 by multiple stimuli.** Morgan MJ, Kim YS. Roles of RIPK3 in necroptosis, cell signaling, and disease. Exp Mol Med. 2022 Oct;54(10):1695-1704. doi: 10.1038/s12276-022-00868-z. Epub 2022 Oct 12. PMID: 36224345; PMCID: PMC9636380. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/

RIPK3 can be activated via various receptors when bound by their respective ligands. These are TNF receptor 1 (TNF-R1), CD95, death receptors (DR4/5), Toll-like receptors (TLR3/4), and Z-DNA-binding protein-1 (ZBP1)/DAI. In the first three of these pathways (but not TLR3/4 or ZBP1), RIPK1 is required and binds to RIPK3 through its receptor-interacting protein homotypic interaction motif (RHIM). In the case of ZBP1, RIPK3 is recruited directly via the ZBP1 RHIM domain, while in the case of TLR3/4, RIPK3 is recruited indirectly via the RHIM domain of TRIF. Once activated, RIPK3 autophosphorylates and then phosphorylates and activates MLKL to induce a conformational change and translocation to the membrane, where membrane permeabilization follows. During this process, post-translational modifications positively and negatively regulate the necroptosis pathway. Two E3 ligases, Pellino-1 (PELI1) and carboxy terminus of HSC70-interacting protein (CHIP), may control the basal threshold of necroptosis. Another E3 ubiquitin ligase, TRIM21, is proposed to be a regulator of necroptotic cell death in response to TRAIL. PPM1B suppresses necroptosis by dephosphorylating RIPK3.

The domain structure and phosphorylation sites on human RIPK3 are shown in Figure 32.13.7 below.

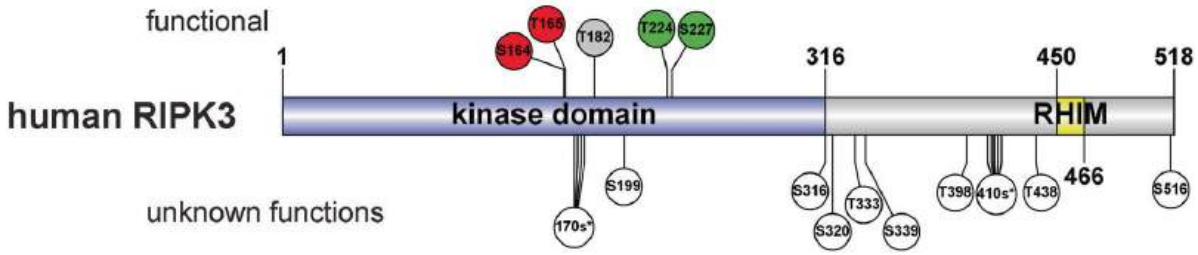


Figure 32.13.7: Meng, Y., Horne, C.R., Samson, A.L. *et al.* Human RIPK3 C-lobe phosphorylation is essential for necroptotic signaling. *Cell Death Dis* **13**, 565 (2022). https://doi.org/10.1038/s41419-022-05009-y. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/

"Schematic of human RIPK3 domain architecture and the phosphorylation sites identified. Phosphorylation sites with proposed functions are shown at the top. pT224 and pS227 positively regulate necroptosis (green) by recruiting MLKL. pS164 and pT165 negatively regulate necroptosis by inhibiting RIPK3 kinase activity (red) [38]. Phosphorylation of T182 (grey) was proposed to promote RIPK3 kinase activity and to recruit PELI1 to mediate proteasomal degradation of RIPK3 [54]. Phosphorylation sites with unknown functions are shown on the bottom (white). Asterisks (*) denotes multiple serine/threonine on the same peptide, as such the exact site of phosphorylation could not be unambiguously identified."

Figure x shows a complex, the **necrosome,** containing multiple activated RIPK3s along with RIPK1. Aggregation of RIPK3 occurs through the **RHIM** (RIP homotypic interaction motifs) domain through the formation of amyloid fibers. The necrosome then phosphorylates

MLKL, which forms oligomers and traffics to the membrane.

Figure x above also shows that an internal sensor protein for viral DNA can also activate RIPK3. That protein is **ZBP1**, or Z-DNA-Binding Protein 1, which also binds Z-RNA. Nuclear Z-RNA can derive from viruses like influenza A, leading to the activation of the same pathway. Cytokine expression then produces a systemic inflammatory response.

In addition to apoptosis and necroptosis, another type of programmed cell death caused by inflammation is called **pyroptosis**. Usually occurring in bacterial-infected macrophages, pyroptosis leads to the activation of intracellular inflammasomes, which then activate inflammatory cytokines through selective proteolysis by caspases. In pyroptosis, proteins called gasdermins are cleaved by caspases and their N-terminals self-associate in the cell membrane to form pores, from which the inflammatory cytokines IL-1β, and IL-18 are released.

A final programmed cell death pathway for virally-infected cells is called PANoptosis, which uses the PANoptosome complex with downstream results not explained by the other three programmed cell death pathways (pyroptosis, apoptosis, and necroptosis) ZBP-1 leads to the activation of RIPK3, caspase-8 (key in the apoptosis pathway) and the NLRP3 inflammasome.

ZBP-1 seems to play a key role in heat stroke. Its concentration increases with heat stress mediated by the heat shock transcription factor 1 (HSF1., which itself is induced by cellular stress. HSF1 induces a heat shock response which causes increased transcription of chaperones and heat shock proteins (HSPs) such as ZBP-1. Deletion/inactivation of ZBP-1, RIPK3, or MLKL and caspase 8 decreases heat stroke. The main role of ZBP-1 in cell death from heat stroke arises from the RIPK3/MLKL pathway and to less extent through cross-talk with the classical apoptosis pathway through caspase 8.

How does ZBP-1 activate cell death during heat stroke without binding to and activating dsDNA or RNA derived from a viral infection? Does ZBP-1 have an endogenous ligand other than viral Z-RNA or Z-DNA? Let's first explore the domain structures of some key proteins in the RIPK3 activation pathway. Figure 1 above shows three key proteins, RIPK1, TRIFF, and ZBP1 that interact with RIPK3. Each of these proteins and RIPK3 have a RHIM domain for protein-protein interactions. Figure 32.13.8 shows the domain structure of our key protein, ZBP-1, the cytosolic Z-DNA/Z-RNA sensor.

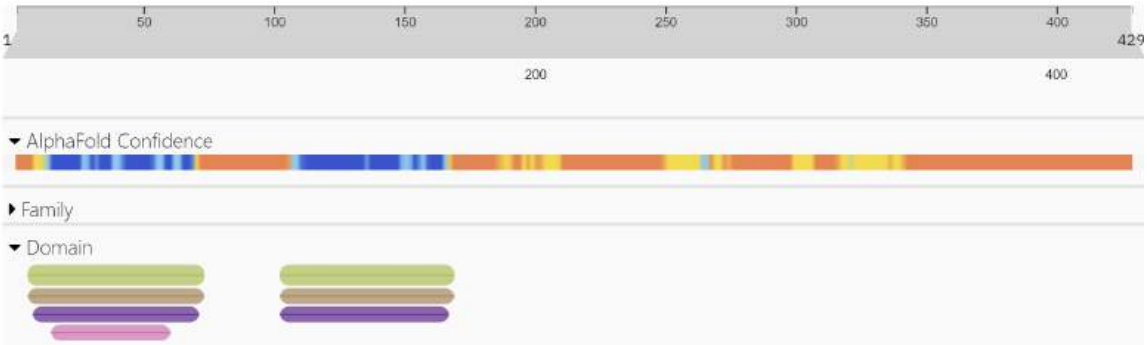


Figure 32.13.8: Domain structure of ZBP-1 (http://www.ebi.ac.uk/interpro/protein/UniProt/Q9H171/)

The green bars in the N-terminal part of the protein are the Z-DNA binding domain. These are also called Zα domains. These regions are the most ordered in the protein, as indicated by the blue in the AlphaFold confidence bar.

Figure 32.13.9 shows an interactive iCn3D models of the AlphaFold-predicted model of human Z-DNA-Binding Protein 1 (ZBP1), (Q9H171)

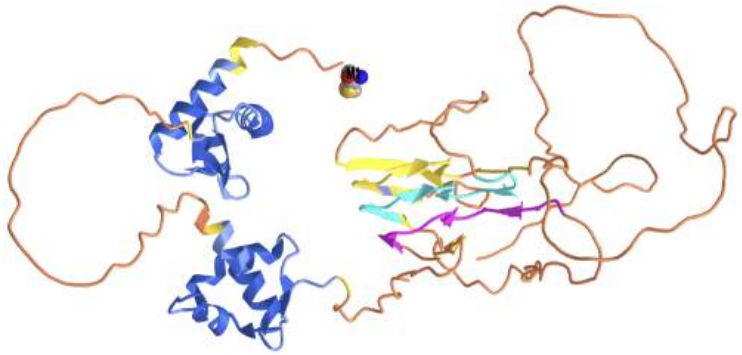


Figure 32.13.9: AlphaFold-predicted model of human Z-DNA-Binding Protein 1 (ZBP1), (Q9H171). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...FMCRhJwHFVb1X7

The spacefill atoms labeled M1 represent the N-terminal methionine of the protein. The two Z-DNA binding domains follow, are well-ordered, and are shown as blue cartoons. Much of the protein can't be predicted as it is most likely intrinsically disordered. Two fairly well-structured motifs, shown in magenta and cyan are the RHIM1 and RHIM2 protein interaction motifs, which can be shown self-

associated through their amyloid-like structures. These motifs allow ZBP1 to bind to other proteins with RHIM motifs and on to cell death through necrosis. The C-terminal domain appears to be involved in signal transduction type I interferon-mediated by DNA.

Figure 32.13.10 shows an interactive iCn3D model of the second Z-DNA binding domain of human DAI (ZBP1) in complex with Z-DNA (3EYI)

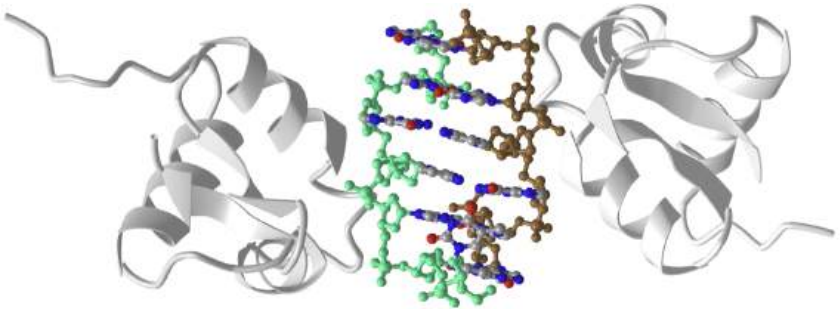


Figure 32.13.10: Second Z-DNA binding domain of human DAI (ZBP1) in complex with Z-DNA (3EYI). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...SLBmo2xqV1WpS6

In the absence of viral DNA or RNA, the Zα domain can bind endogenous ligands. Moreover, it appears that a deficiency in RIPK1 or of the RHIM in RIPK1 also triggers ZBP1 to induce necroptosis and inflammation. and that its Zα domain is required. If nuclear export was stopped, ZBP1 activates nuclear RIPK3 and then necroptosis. This suggests that nuclear ZBP1 interacts with endogenous nuclear Z-nucleic acids, probably Z-RNA from retroelements to activate RIPK3-dependent necroptosis and could lead to some forms of chronic inflammation.

Here are a series of finding on RIPK3-dependent cell death on heat stress in mouse fibroblasts that show that Z-nucleic acid binding to ZBP1 is not required for heat stress effects:

- Heat (43°C for 2 hr) induces phosphorylation of RIPK3 and MLKL within 2 hours, and cleavage of pro-caspases and GSDME in 6 hours but none occurred if RIPK3 was deleted.
- Deletion of ZBP1 but not RIPK1, TRIF, affect heat induce death so so heat stress acts through ZBP1 and RIPK3.
- In mice without ZBP1, the effects of heat stress (clotting, inflammation, organ injury, and death) were prevented.
- Mutations in the RHIM domain, but **not** the Zα domains (made to prevent Z-nucleic acid binding) or in the C-terminal signaling region (to stop signaling) prevented death from heat stress . Hence Z-nucleic acid binding is **not** required but may contribute to cell death from heat stress.
- Heat stress caused the aggregation of a ZBP1-GFP (green fluorescent protein) fusion protein through the RHIM domains of ZBP-1

Hence ZBP1 is an innate pathogen sensor and also an initiator of heat-related death in the absence of pathogens.

32.13.5: HEAT STROKE-INDUCED EPIGENETIC CHANGES

Short of death, heat stroke can also cause long-term health issues. Increasing global temperatures are forcing people to work at more dangerous temperatures and at night to reduce heat exposure. Data suggests that people who have had a heat-related illness are more susceptible to additional heat exposure health consequences. This has been noted in exertional heat illnesses. (such as in athletes). Additional long-term effects on immune regulation have been observed. Epigenetics may play a role in long-term effects such as greater vulnerability to additional heat challenges. Studies show that a single episode of exceptional heat stroke changes DNA methylation patterns in bone marrow-derived monocytes from mice. The monocytes become immunosuppressed allowing for increased microbial disease and reduced heat shock responses. The epigenetic changes are passed onto progeny monocytes which also shows compromised function. The epigenetic changes persist for 30 days or more and we clearly noted in inflammatory cell signaling pathways. This suggests a mechanism for the reduced tolerance to those with previous heat-related illnesses.

32.13.6: CARDIOVASCULAR AND PULMONARY DISEASES

Many factors can cause cancer and lead to mortality. For example, mortalities from cancer can increase due to lower availability of health care arising from extreme weather disasters that impact health facilities and access to them. Early detection of many cancers is key to survival. Instead of discussing climate change links to cancer we will focus on pollution and in particular small particles.

32.13.6.1: PM2.5 PARTICULATE POLLUTION EFFECTS ON CARDIOVASCULAR AND PULMONARY HEALTH

Pollution from the combustion of fossil fuels contributes to many chronic diseases. Here we will focus on one type of pollutant, particulate matter which can be inhaled. These particles can be liquids, solids, or combinations of both. They are classified according to size with common categories including:

- PM10: diameters $\leq$ 10 uM = 10,000 nm;

- PM2.5: fine particles with diameters $\leq$ 2.5 um = 2500 nm
- PM0.1: ultra-fine particles with diameters $\leq$ 0.1 um = 100 nm (also called nanoscopic particulate matter or NPM)

Figure 32.13.11 below shows the relative sizes of PM10 and PM2.5 particles compared to other biological structures.

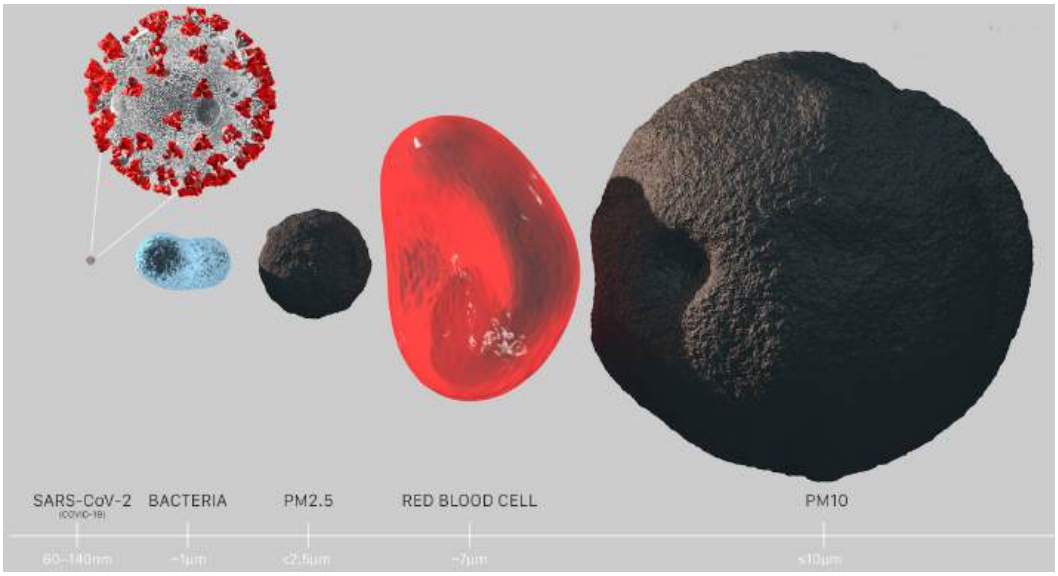


Figure 32.13.11 : Relative sizes of PM particles compared to biological structures. Sotirios Papathanasiou. Particulate Matter (PM2.5) Mega Guide. With Permission. https://seetheair.org/2022/05/16/par...-5-mega-guide/

Figure 32.13.12 below shows the relative sizes of PM0.1 particles compared to a PM2.5 particle.

Figure 32.13.12 : Relative sizes of PM01 particles compared to a PM2.5 particle.Sotirios Papathanasiou, ibid.

32.13.7: COMPOSITION OF PM2.5 PARTICLES

PM2.5 particles obtained by collection from polluted city air can actually be purchased from the National Institute of Standards and Technology (NIST, SRM1648a) and used for experimental studies on living cells. It is typically added to water and a suspension produced through sonication. PM2.5 particles are derived from human sources such as emissions from vehicles and industry, and both human-caused and natural processes such as the burning of biomass, and the release of dust from land. They can also include salts from land and ocean sources.

They arise from the burning of fossil fuels and wear and tear of products such as automobiles (including tires). PM2.5 particles contain mainly black carbon, polycyclic aromatic hydrocarbons (PAH), aryl hydrocarbon, volatile organic compounds (VOCs) as well as minerals,

ions (sulfate, nitrate, ammonium), and general biological materials. The metal composition includes Group 1A (K, Na, Fr), Group 2A (Ca, Mg), Group 3A (Al), transition metals (Al, As, Cr, Fr, Mn, Pb, Ti, Zn), and counter ions Br and Cl. They also contain silicon and silicates. Of course, particles in the air, including dust, also derive from non-anthroprogenic sources. Atmospheric dust is also produced from land by winds and also by volcanic eruptions. In homes, dust has an abundance of dead skin cells. along with pollens, hair, fur, and fibers from clothes and paper. Humans have evolved with particulates in the air, but large increases in their abundance caused by human activities in many parts of the world pose serious health consequences.

In addition, the reaction of pollutants in the atmosphere produces "secondary" pollutants. One, the tropospheric gas ozone (O_3) produced from hydrocarbons and nitrogen oxides, is a known health risk, and its levels are increased in cities on sunny, hot, and humid days. Secondary organic carbon (SOC) is also generated from primary organic carbon, typically volatile organic compounds (VOC) through oxidative photochemical reactions. These VOCs (like m-xylene and 1,2,4-trimethylbenzene) can produce aerosols (larger particles, called secondary organic aerosols by reacting with each other to produce larger structures. Terpenes containing isoprene units like α-pinene and limonene (a monoterpene found in large abundance in fruit peels) are reactants for the products of much larger structures. (See Chapter 10.1 and Chapter 21.6 for a review of isoprene and terpenes). Figure 32.13.13 below shows generalized pathways for the formation of particulate SOCs from smaller terpenes.

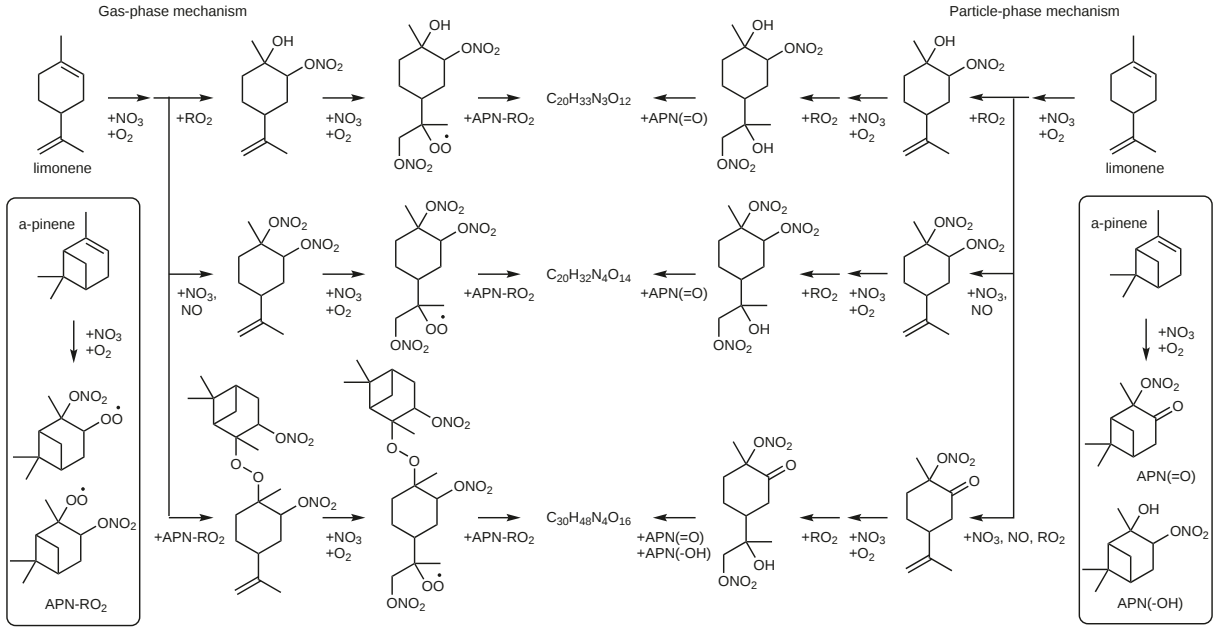


Figure 32.13.13 : **Proposed mechanisms for the formation of $C_{20}H_{33}N_3O_{12}$, $C_{20}H_{32}N_4O_{14}$, and $C_{30}H_{48}N_4O_{16}$ in the simultaneous oxidation (MIX) experiment.** Takeuchi, M., Berkemeier, T., Eris, G. *et al.* Non-linear effects of secondary organic aerosol formation and properties in multi-precursor systems. *Nat Commun* **13**, 7883 (2022). https://doi.org/10.1038/s41467-022-35546-1. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Gas-phase mechanism via cross-reactions of α-pinene and limonene peroxy radicals (RO_2·), and particle-phase mechanism via hemiacetal formation involving α-pinene and limonene oxidation products. APN-RO_2, APN(=O), and APN(-OH) represent α-pinene oxidation intermediate (i.e., RO_2·), α-pinene oxidation products with carbonyls, and α-pinene oxidation products with hydroxyl functional groups, respectively.

Figure 32.13.14 below shows the morphology of PM2.5 particles.

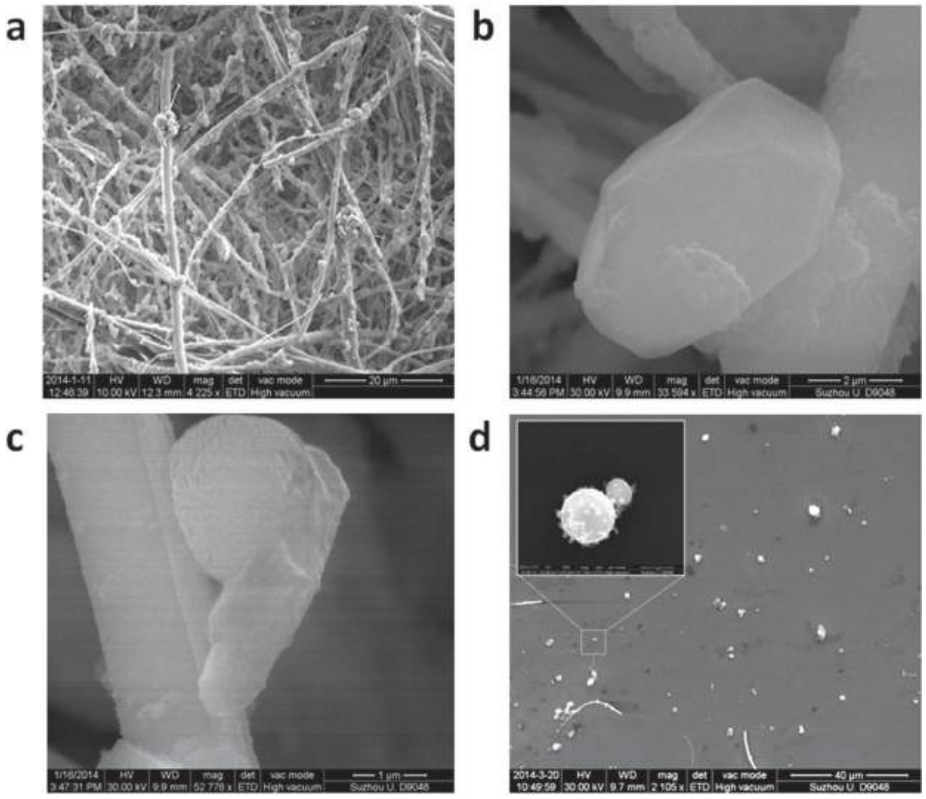


Figure 32.13.14 : Morphology of PM2.5 particles. The scale bars are 20 µm for image (a) 2 µm for image (b) 1 µm for image (c) 40 µm for image (d). Shi, Y., Ji, Y., Sun, H. *et al.* Nanoscale characterization of $PM_{2.5}$ airborne pollutants reveals high adhesiveness and aggregation capability of soot particles. *Sci Rep* **5**, 11232 (2015). https://doi.org/10.1038/srep11232. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Panel (a) shows a large area image of as-collected PM2.5 on the filamentary filter.

Panel (b) and (c) show SEM images of a particle with flat and rough top surface, respectively.

Panel (d) shows a SEM image of the PM2.5 transferred on a Silicon substrate. Inset, zoom-in SEM image of an Iron-rich particle.

Many of the PM2.5 particles have rough surfaces which can deform more easily interact with (i.e. stick to) other particles through noncovalent interactions to produce even larger particles.

Figure 32.13.15 below shows the elemental composition of rough, semi-rough, and flat PM2.5 particles.

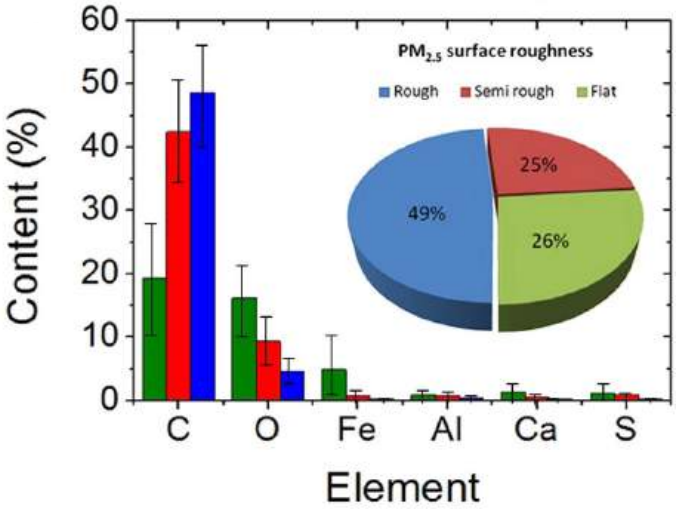


Figure 32.13.15 : EDAX Chemical composition histogram of the particles collected with SEM/EDAX classified by surface roughness; a larger surface roughness (and therefore, stickiness and deformation) are linked to a larger content of Carbon, while particles with a flat

surface (low stickiness and viscosity) are richer in Oxygen and metals. Shi, Y. et al., ibid.

Figure 32.13.16 below shows electron micrographs of actual airborne particles. Most are PM2.5 particles with diameters $\leq$ 2.5 uM = 2500 nm.

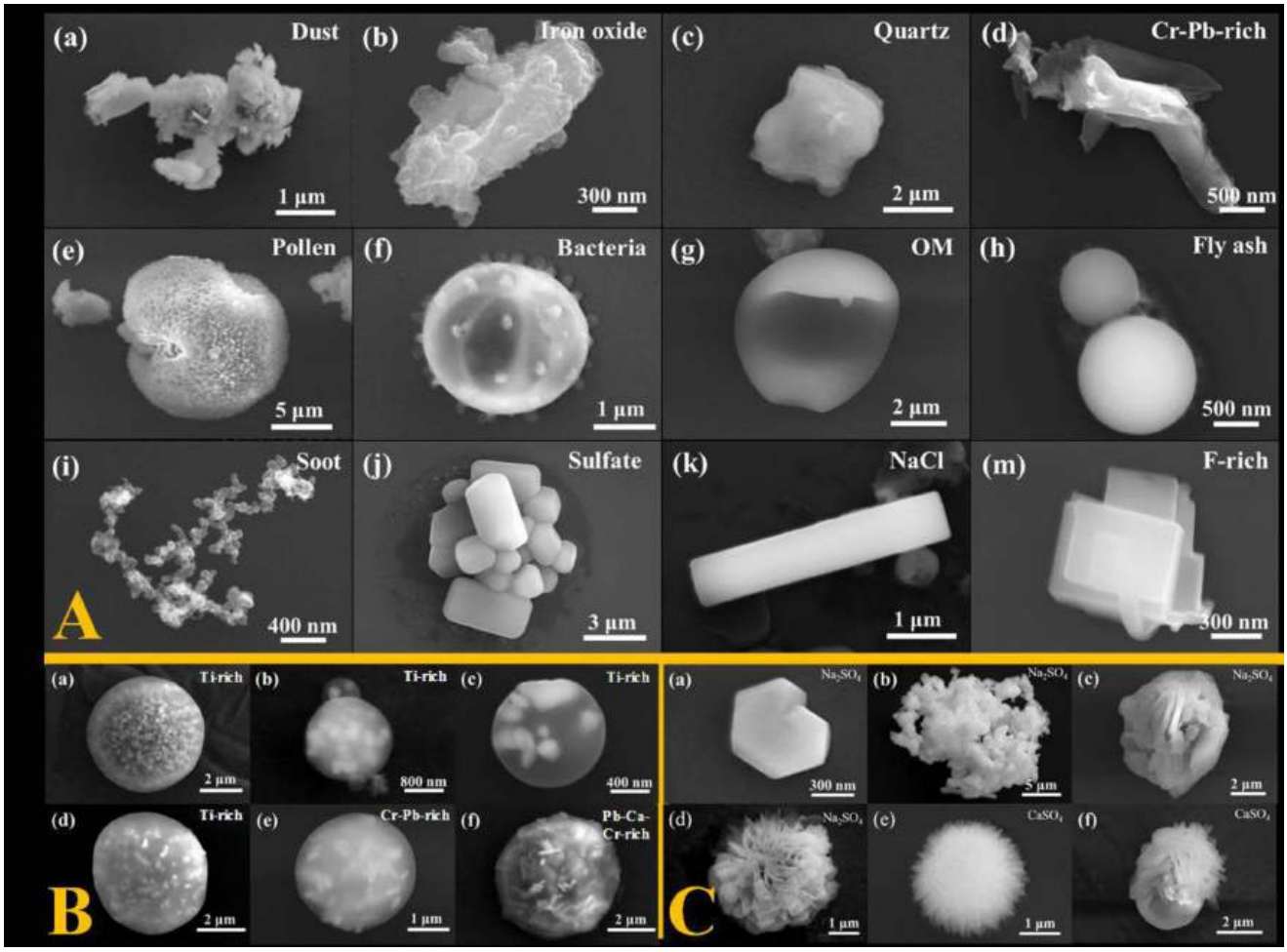


Figure 32.13.16 : A collage of SEM images for airborne particulates. (A) General classification of airborne particulates; (B) particulates with seeds-coating composite morphology; (C) sulfate particulates with different morphologies. Clara Yuan Li et al, Journal of Environmental Protection, Vol.7 No.10, 2016. https://www.scirp.org/journal/paperi...?paperid=71021. Creative Commons Attribution 4.0 International License.

Given their composition and their structures, it doesn't take much thought to realize that the particles must cause significant health effects. Would you want to breath in these particles routinely?

32.13.8: HEALTH EFFECTS OF PM2.5 PARTICLES

PM2.5 particles are associated with just about every type of illness, including cardiovascular and pulmonary diseases, including asthma, as well as cancer. Since they are small, they can easily be inhaled and deposited in lung alveoli from where they can actually enter the bloodstream and be deposited in tissue. Particles up to 240 nm (0.25 uM) can cross the placenta and black carbon particles have been found to cross the placenta. PM2.5 particles can cause inflammation, DNA damage, organelle dysfunction and can also generate free radicals which are most likely involved in these toxic health effects.

THE GREAT LONDON FOG OF 1952

Aerosol particles can even have acute and lethal effects. During The Great London Smog in London in December 1952, around 12,000 people died in two weeks from its effects. Figure 32.13.1 people in the thick smog from that event. The smog consisted of acidified water droplets arising from SO_2 and NO_2 released on burning coal that contains sulfur. This gas can be oxidized to sulfate in

gas-phase reactions probably through the ·OH free radical or in aqueous phase reactions using O_3, peroxides, and NO_2 as reactants/catalysts.

The reaction of SO_2 with NO_2 in aqueous droplets is shown below:

$$SO_2(g) + 2NO_2(g) + 2H_2O(aq) \rightarrow 2H^+(aq) + SO_4^{2-}(aq) + 2HONO(g) \quad (32.13.1)$$

Given the stoichiometry of the reaction with NO_2, the reaction proceeds significantly only in the presence of high NO_2. The reaction is favored under high relative humidity. In large water drops in clouds, the droplets are not very acidic but as water evaporates from the drops, the sulfate concentration and acidity dramatically increase. However, as the acidity increases the rate of oxidation and the solubility decreases.

Major cities in China, including Beijing and Xian, have experienced high levels of haze and particulate matter in the atmosphere until recently. Yet these smogs were not lethal (very acidic) as in London due to the addition of NH_3 in the droplets, which neutralizes the acidic particles. Atmospheric ammonia is derived from large amounts of agricultural fertilizers that get aerosolized and also from vehicles, which produce NH_3 in catalytic converters and from urea used in the catalytic reduction in diesel engines. The relevant production of sulfates in the presence of NH_3 is shown in the equation below.

$$2NH_3(g) + SO_2(g) + 2NO_2(g) + 2H_2O(aq) \rightarrow 2NH_4^+(aq) + SO_4^{2-}(aq) + 2HONO(g) \quad (32.13.2)$$

Figure 32.13.17 below shows a ghostly image of pedestrians in London during the Great Fog.

Figure 32.13.17 : Ghost-like pedestrians making their way through the smog. https://heritagecalling.com/2022/12/...f-london-1952/. Public Domain

The prevailing weather conditions (cold temperatures) increased emissions from coal use. A stalled high pressure system and resulting low winds caused the buildup of stagnant air with increasingly acidic PMs. The appalling death toll led politicians to pass the Clean Air Act in 1954 which over many years led to huge improvements in air quality in London and dramatically reduced negative health effects and deaths. This was a prelude to the Clean Air Act in the US which dramatically improved air quality as well.

High levels of PM2.5s still are prevalent in much of the world, although there have been dramatic decreases in the US. The notable exceptions occur during forest fires that are exacerbated by climate change. In the US, the Air Quaility Index (AQI) is used as an indicator of health risk. It measures the value of 5 pollutants, fine particles (PM2.5 and PM10), ground-level ozone, SO_2, NO_2 and CO. The value of AQI at a given time is determined by which pollutant is highest. In haze produced by smoke, the reported AQI represents PM2.5 particles. AQI values < 50 or below represent good air quality, while an AQI value over 300 represents hazardous air quality. Western forest fires in Oregon in September 2020 lead to an AQI of 611 in Madras, Oregon. Forest fires in Eastern Canada, along with slow-moving weather system, led to PM2.5 levels over 800 (mg/m^3) in New York City on June 7, 2023. No place is immune from PM2.5 particles from wild fires and human-caused pollution.

- AirNow - gives present pollution data, primarily PM2.5 levels, based on US zip codes

The following interactive graphs show the changes in PM2.5 particles over time (from Hannah Ritchie and Max Roser (2019), OurWorldInData.org/outdoor-air-pollution • CC BY. Source: Brauer et al. (2017) via World Bank).

Figure 32.13.18 shows the share of the population exposed to PM2.5 levels higher than those suggested by the World Health Organization.

Figure 32.13.18: Share of the population exposed to PM2.5 levels higher than those suggested by the World Health Organization.OurWorldInData.org/outdoor-air-pollution • CC BY. Source: Brauer et al. (2017) via World Bank

Figure 32.13.19 below shows the share of the population in the US and in India exposed to PM2.5 levels higher than those suggested by the World Health Organization.

Figure \(\PageIndex{19}\): PM2.5 levels in the US and India.Our World in Data. ibid

Finally, Figure 32.13.20 below shows the death rate from PM2.5/100,000 people in 2017 in countries around the world.

Figure 32.13.20 : Death rate from PM2.5/100,000 people in 2017 in countries around the world.

32.13.9: MECHANISMS FOR PM2.5 HEALTH EFFECTS

In Chapter 5.4, we discussed how solids such as silica, cholesterol crystals, uric acid crystals, and even aggregated proteins such as prions can be engulfed by monocytes/macrophages (much as they engulf bacteria as part of their immune function) in a process called phagocytosis. The particles are enveloped in plasma bilayer-derived membrane which buds off into the cell. This vesicle merges with a lysosome which gets damaged in the process. They then release ATP into the cytoplasm which acts as a damage signal to activate inflammation.

PM2.5 particles can also be taken up by phagocytosis to produce intracelluar phagosomes. They can also be taken up by pinocytosis, and caveolin and clathrin-mediated endocytosis to form endosomes. These can fuse with lysosomes and mitochondria and induce damage. The heavy metals from PM2.5 particles that are released into the cell also contribute to damage.

Smaller PM2.5 particles of diameter less than 0.1 uM (100 nM), sometimes called **nanoscale particulate matter (NPM)**, offer large surface areas to which proteins can adhere. The adsorbed proteins form a "crown" called a protein **corona**. The corona is actually larger than the NPM. Proteins bounds to the particles include hemoglobin, albumin and fibrinogen. The corona also mediates cellular interactions and participates in the mechanisms that lead to inflammation and cellular dysfunction. As the extracellular matrix includes protein components such as collagen and fibrin, the interaction of PM2.5 with fibrin, has been studied as a model for how the particles might interact with cells. In particular, lung fibroblasts embedded in a 3D matrix of fibrin (i.e. a 3D organotypic culture) were exposed to

NPM with protein coronas, and the effects of the NPM particles on cell proliferation, oxidative stress, etc. were monitored. Figure \(\PageIndex{21}\) below characterizes the interaction of the NPMs with the fibroblast in the 3D culture.

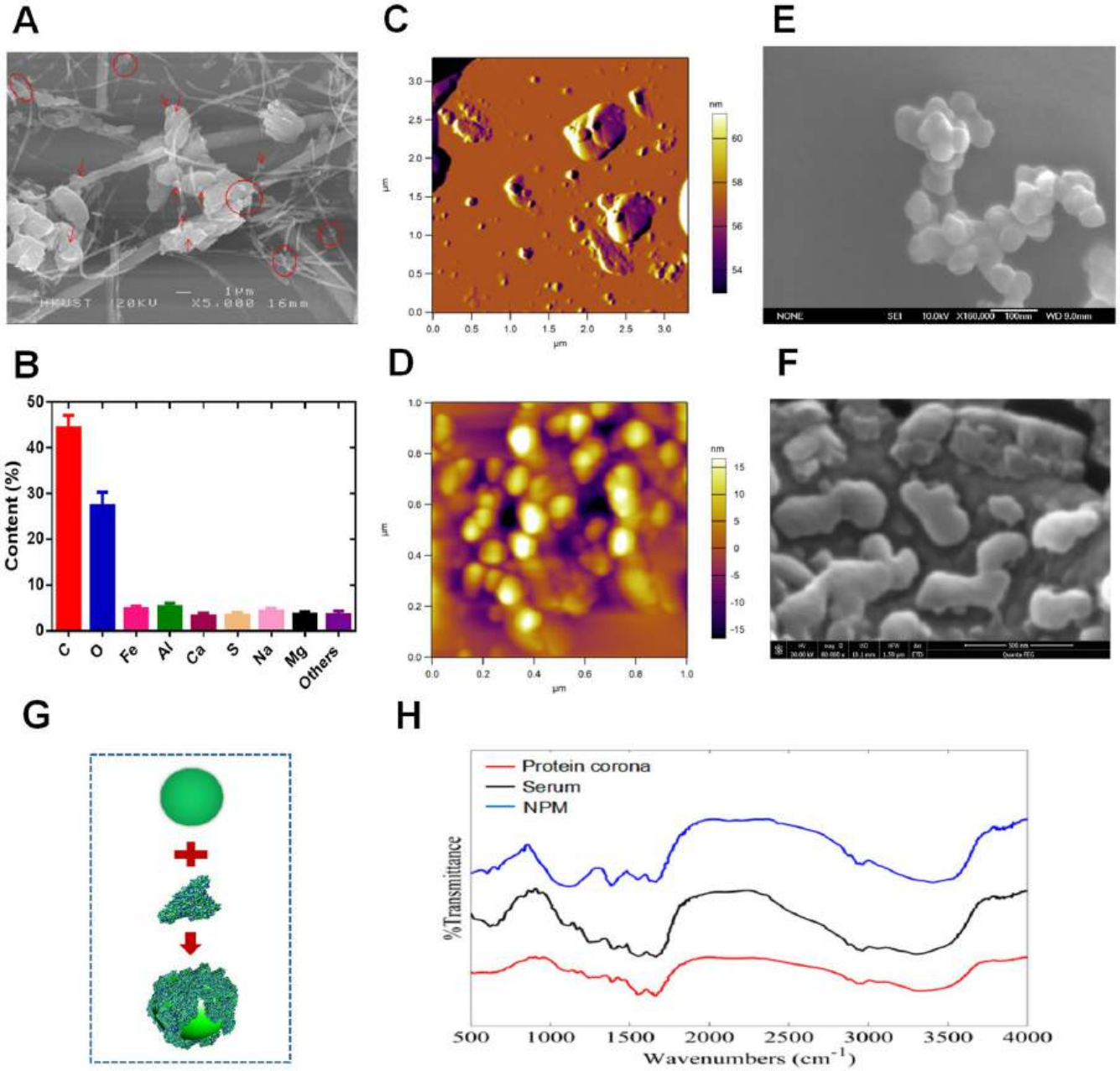


Figure 32.13.21 : Physicochemical characterization of airborne particulate matter. Li, Y., Wang, P., Hu, C. *et al*. Protein corona of airborne nanoscale PM2.5 induces aberrant proliferation of human lung fibroblasts based on a 3D organotypic culture. *Sci Rep* **8**, 1939 (2018). https://doi.org/10.1038/s41598-018-20445-7. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Panel (**A**) shows SEM images of airborne PM2.5 (nanoscale PM2.5 is marked in red).

Panel (**B**) shows a chemical element analysis of PM2.5 by EDX analysis.

Panel (**C**) shows an atomic force microscopy (AFM) image of nanoscale PM2.5 (NPM) from air pollutant samples (scale 3.5 μm).

Panel (**D**) AFM image of airborne NPM from air pollutant samples (scale 1 μm).

Panel (**E**) shows an SEM image of airborne NPM from air pollutant samples.

Panel (**F**) shows an SEM image of an NPM-protein corona.

Panel (**G**) shows a schematic diagram of the biological interaction between NPM and protein.

Panel (**H**) shows a FTIR spectra of NPM, serum, and NPM-protein corona.

The NPM-protein corona particle leads to the proliferation of the 3D-cultured human lung fibroblast cells over and above stimulation of the cells with just NPMs or serum alone. The bigger the size of the corona, the greater the proliferative effect. This is consistent with the extensive fibrosis of the lungs seen after chronic exposure to PM2.5 particles. Reactive oxygen species also increased in the presence of NPM-corona particles. These data suggest that NPM-protein corona are important in PM2.5-induced lung fibrosis and pulmonary disease.

32.13.10: NEURAL EFFECTS OF PMS

PM particles, particularly the ultrafine PM0.1 particles (UFP), can enter the brain, and affect neural function. Exposure to PM2.5 over long periods of time is associated with increased incidence of dementia and Alzheimer's Disease (AD). Four particular components of PMs (SO_4^{2-}, NH_4^+, black carbon, and organic matter) were most associated with a higher risk of dementia and AD. In the US, the first admission to a hospital for Parkinson, AD, and other dementia is "significantly" associated with the average annual mean PM2.5 exposure.

Even low levels of exposure pose risks. Transgenic mice containing mutant forms of human presenilin 1, amyloid precursor protein, and the tau protein, were exposed to subchronic, "real-world" levels of PM2.5 through inhalation for 3 months. Neuronal loss was observed in the cortex but no motor or cognitive impairment was noted. Increased levels of phosphorylated forms of Tau were observed and free radical formation, as evidenced by levels of malondialdehyde (a marker of oxidative stress), was seen in the hippocampus and olfactory centers, consistent with inhalation of PM2.5 through the noise. No abnormal amyloid plaques were observed in this short exposure time.

PMs damage to neurons occur through the generation of reactive oxygen species increased inflammatory responses and organelle damage (all of which are interrelated). Even in skin cells (keratinocytes), exposure to PM2.5 particles led to increased levels of ROS and malondialdehyde, decreased levels of superoxide dismutase, and increased DNA damage. Inflammatory Caspase levels also increased.

Cells have mechanisms to detect and eliminate aberrant species before they led to worse biological effects. One process is **autophagy** which degrades misfolded protein and damaged organelles, processes important for normal neural function. Aberrant autophagy is a key player in the pathogenesis of dementia. A second pathway is **ferroptosis**, a kind of apoptotic pathway, which kills cells that have accumulated large amounts of iron ions, which of course are free radicals themselves. Through the Fenton reaction and others, iron ions can lead to the generation of damaging reactive oxygen species and oxidation of lipids, proteins, and nucleic acids (see Chapter 12.3).

Several key proteins are involved in antioxidant defense and autophagy:

- NRF2 (Nuclear factor erythroid 2-related factor 2): This transcription factor binds to antioxidant response elements (ARE) in front of protective genes.
- Keap1 (Kelch-like ECH-associated protein 1): An adapter program that targets NRF2 for ubiquitination and as such is a sensor for oxidative stress. Under those conditions, electrophilic metabolites lead to post-translational modification of reactive Cys side chains which leads to the inactivation of ubiquitin ligase activity. This increases NRF2 and subsequent transcription of antioxidant genes.
- SQSTM1 aka p62 (Sequestosome-1): This protein bridges autophagosomes and polyubiquitinated proteins (cargo for degradation).

Hence under low oxidative stress, NRF2 are low through its degradation mediated by KEAP1. If SQSTM1 (p62), increases, and reduced autophagy, p62 biding to the NRF2 sites on KEAP1, leading to the release of NRF2, its transfer to the nucleus, and activation of gene transcription of protective genes. These three proteins also protect cells from ferroptosis. In the long term, neural cell death occurs during dementia. Hence apoptosis, necroptosis, and pyroptosis (from activation of the inflammatory response mediators capspase, Gasdermin, and key cytokines like IL-1β and IL-18) do "win out" to eventually kill cells.

Figure 32.13.22 below shows the domain structure of SQSTM1 (panel A) and the signaling process described above.

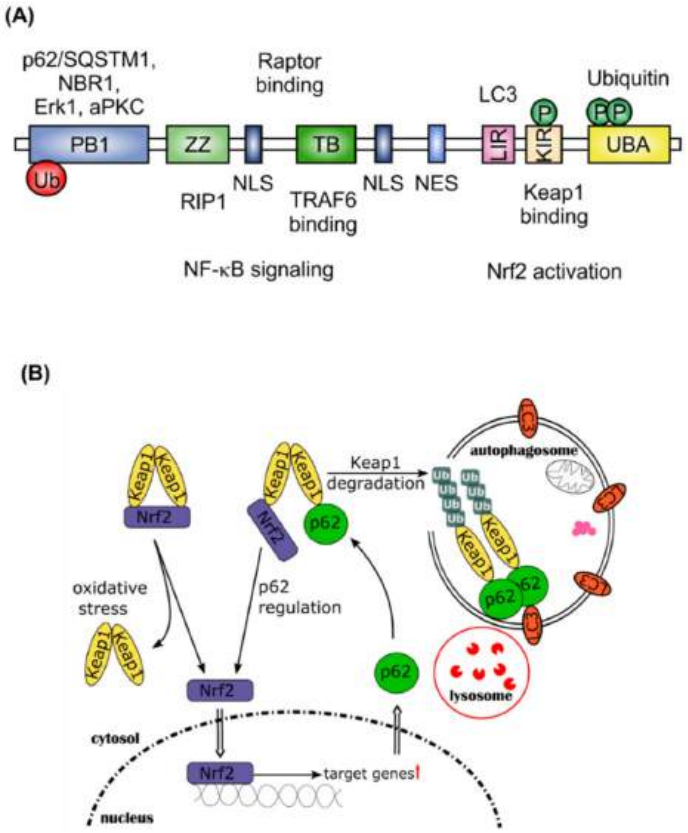


Figure 32.13.22: Positive feedback-loop of Nrf2 activation by *p62/SQSTM1*. Vomund S, Schäfer A, Parnham MJ, Brüne B, von Knethen A. Nrf2, the Master Regulator of Anti-Oxidative Responses. Int J Mol Sci. 2017 Dec 20;18(12):2772. doi: 10.3390/ijms18122772. PMID: 29261130; PMCID: PMC5751370. Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

Panel (**A**) shows the domain structure of p62/SQSTM1;

Panel (**B**) shows p62/SQSTM1 as an important protein for selective autophagy, as binds to Keap1 and other long-lived proteins and forms polyubiquitinated protein aggregates. Furthermore, it binds to the autophagy marker LC3 within the autophagosome, thereby leading the aggregated proteins into the autophagosome. After fusion with a lysosome, proteins and organelles, such as mitochondria, are degraded within the autophagosome. By binding to Keap1, p62/SQSTM1 stabilizes Nrf2 and enhances its translocation into the nucleus, where Nrf2 activates its target genes (↑ = upregulation of Nrf2 target genes). One of these genes is *p62/SQSTM1*

PM2.5 particles intefer with the protective autophagy and ferroptosis pathways, leading to increased NRF2 activity and expression of antioxidant genes which are beneficial processes to rid cells of aberrant particles and kill damaged cells in normal conditions but not on exposure of neuronal cells to PM2.5 particles.

Lysosomal membrane permeabilization (LMP) as well as mitochondrial and ER damage might be a likely mechanism to initiate ultimate neuronal death on long-term exposure to PM2.5 particles. PM2.5 particles inhibit lysosomal activity and increase their permeability and release to degradative enzymes into the cytoplasm. Ultimately increased or decreased activation of Nrf2 lead to disease states.

32.13.11: PMS AND CANCER

Long-term exposure can also cause lung cancers. This is associated with a conversion of lung cells from normal epithelial to mesenchymal cells (called the EMT transition). Phenotypically, mesenchymal cells can migrate, invade other tissues and cause enhanced production of the extracellular matrix, all hallmarks of tumor cells. This EMT transition is associated with significant changes in cell signaling and the production of transcription factors. These changes are documented in Figure 32.13.23 below and its caption.

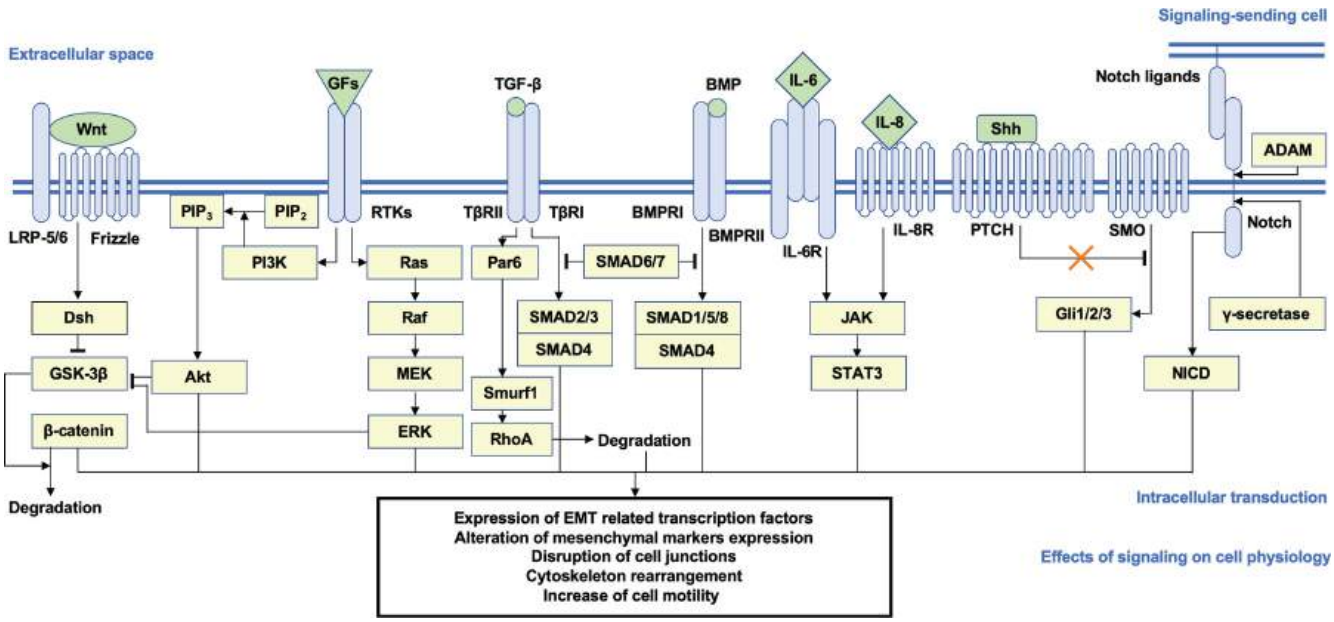


Figure 32.13.23 : Brief schema of the putative signaling transduction mechanisms underlying EMT. Xu et al., **Front. Physiol., 29 November 2019. Sec. Renal Physiology and Pathophysiology.** Volume 10 - 2019 | https://doi.org/10.3389/fphys.2019.01404. **Creative Commons Attribution License (CC BY)**

Activation of the Wnt/β-catenin, PI3K/Akt, Ras/ERK, TGF-β/SMAD2/3, BMP/SMAD1/5/8, JAK/STAT3, Shh, and Notch pathways is highly correlated with EMT. After ligand-receptor binding, intracellular secondary messengers are activated and initiated downstream transduction, which generally induce the nuclear translocation of signaling-specific TFs and the transcriptional regulation of EMT-related genes, such as CDH1 and CDH2, EMT TFs, and mesenchymal markers, accompanied by a series of alterations on cellular physiological or pathological activities (e.g., dysjunction of adherin junctions, cytoskeleton remodeling, and increase of cellular motility). Arrows represent the molecular interactions in which downstream messengers are activated; T shape arrows represent inhibitive molecular interactions. EMT, epithelial-mesenchymal transition; PI3K, phosphoinositide 3-kinase; ERK, extracellular signal-regulated protein kinase; TGF-β, transforming growth factor β; JAK, Janus kinase; Shh, sonic hedgehog; TFs, transcription factors.

PM2.5 particles and their deleterious contents (heavy metal ions, PAHs, etc), as well as ROS generated by them are associated with the changes seen in the EMT transition. These include activation of transforming growth factor β (TGF-β)/SMADs, NF-κB, growth factor (GF)/extracellular signal-regulated protein kinase (ERK), GF/phosphatidylinositol 3-kinase (PI3K)/Akt, wingless/integrated (Wnt)/β-catenin, Notch, Hedgehog, high mobility group box B1 (HMGB1)-receptor for advanced glycation end-products (RAGE), and aryl hydrocarbon receptor (AHR) signaling cascades and to cytoskeleton rearrangement.

32.14: CLIMATE CHANGE, INFECTIOUS DISEASE AND PANDEMICS

Written by Henry Jakubowski

32.14.1: INTRODUCTION

Microorganisms can cause both chronic and acute diseases, both of which if left untreated can lead to death. Infections with the Human Papillomavirus (HPV) can cause cancer of the cervix, vagina, vulva, penis, anus, and throat. Modern vaccines against HPV can prevent over 90% of these cancers. The bacteria H. pylori can, in some people, cause stomach illness (such as severe chronic gastritis and ulcer) which can lead to stomach cancer. The Coxsackie virus, through binding to receptors on cardiac myocytes, can cause heart disease (acute myocarditis and cardiomyopathy), and ultimately death.

Acute microbial diseases that occur immediately after infection can cause epidemics and pandemics (worldwide epidemics). Everyone has experienced the COVID-19 pandemic caused by the SARS-CoV-2 coronavirus. Johns Hopkins estimates as of 3/10/23 (end of their data collection), that there were about 677 million reported cases of Covid-19 and about 6.9 million deaths. The WHO estimates that just for the first two years of the pandemic (2020 and 2021), there were 14.83 million excess deaths globally, 2.74 times more than the number or reported death (5.42 million) from the virus. Machine learning models suggest that there have been closer to 20 million excess death through the end of March 2023, as shown in Figure 32.14.1 below.

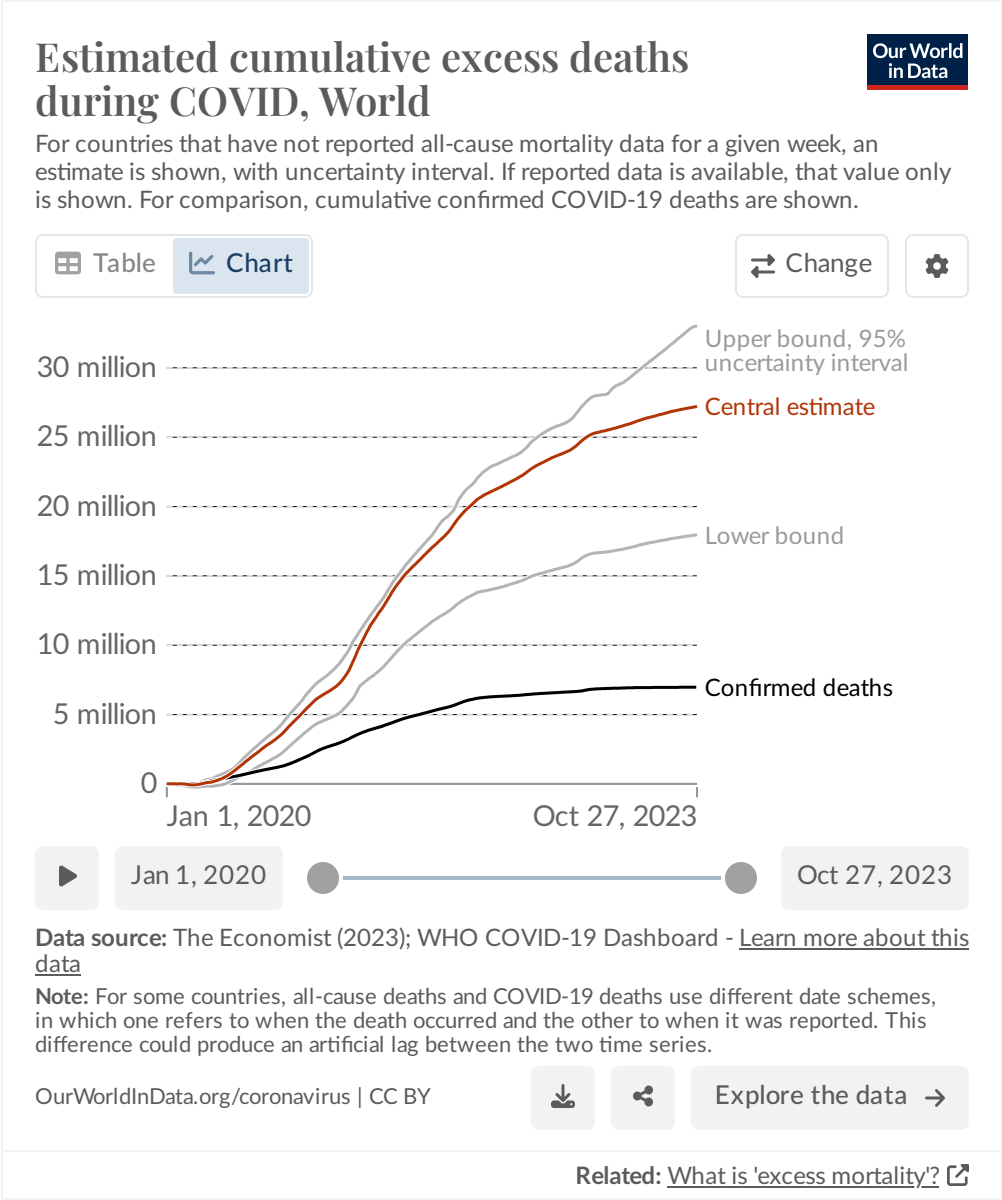


Figure 32.14.1: https://ourworldindata.org/excess-mortality-covid

The data from Johns Hopkins suggest that the average mortality rate was about 1% (deaths/infections). If there have been 20 million cases (based on excess deaths) out of a world population of 8 billion, the mortality rate was close to about 0.25% for the entire world.

Another indicator of the severity of pandemics is a decrease in life expectancy. Figure 32.14.2 below offers an interactive graph that shows the general rise in life expectancies since 1750 punctuated by steep drops.

Figure 32.14.2: Life expectancies since 1750

Note the small drop in 2020 in the United State was caused by the Covid-19 pandemic with some contribution from opioid-associated deaths. The graph is dominated by a stunning decline in 1918 due to the 1918 Flu Pandemic (also historically and inaccurately named the Spanish flu). The large drop in life expectancy in Sweden in 1772-1773 was probably attributed to the Russian plague epidemic of 1770–1772, also known as the Plague of 1771. Figure 32.14.3 below shows a history of pandemics back to the Antonine Plague of 165-180 CE. In viewing the figure, remember that the numbers of deaths are estimates at best, especially for the historically early pandemics.

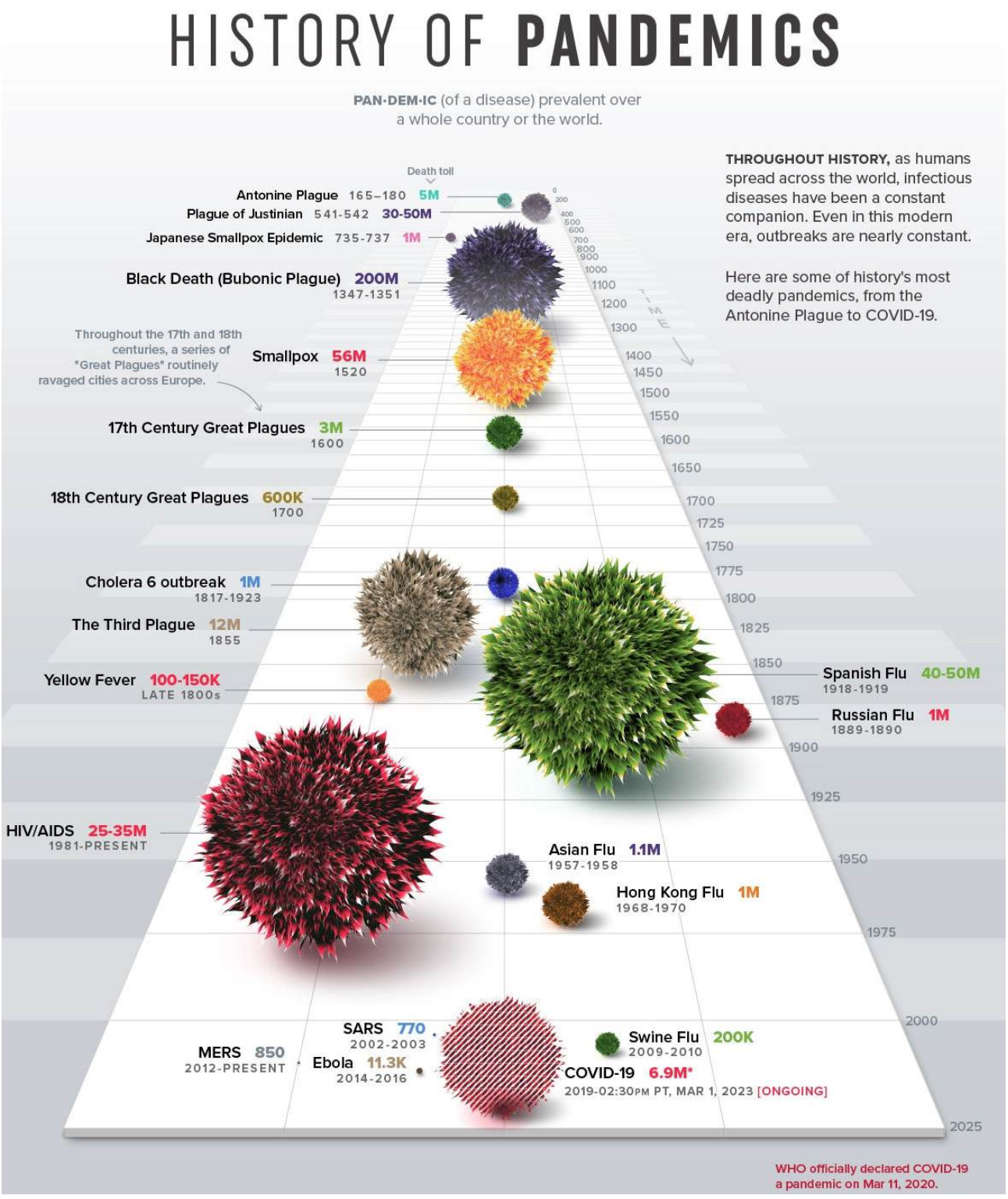


Figure 32.14.3: **Visualizing the History of Pandemics.** Attribution Visual Capitalist. https://www.visualcapitalist.com/his...ics-deadliest/

The graphic misses another key plaque in world history, the Plaque of Athens, which hit the city from around 430 BCE - 427 BCE, during which up to 25% of the city's population died. Smallpox has emerged as a possible candidate for that outbreak.

Figure 32.14.4:

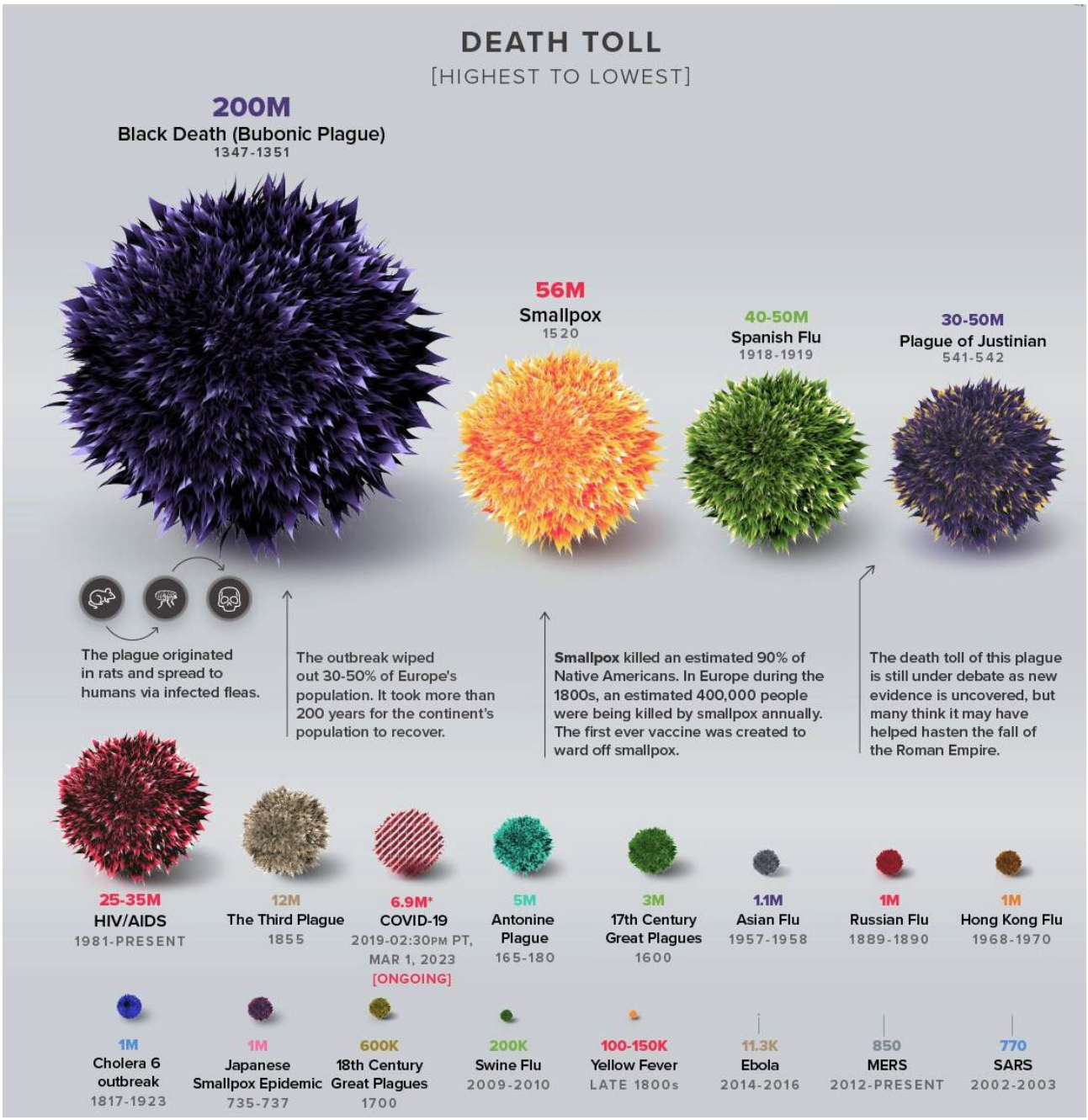


Figure 32.14.4: **Visualizing the History of Pandemics.** Attribution Visual Capitalist. https://www.visualcapitalist.com/his...ics-deadliest/

The **B**lack Death (also called the **B**ubonic Plague) was caused by the **b**acterium, *Yersinia pestis*. Humans usually get the bubonic plague after being **b**itten by a rodent flea that is carrying the plague bacterium or by handling an animal infected with the plague (notice the bold letter **B** to help you remember **B**lack, **B**ubonic, **B**acterium, **B**ite). The Black Death/Bubonic Plagues derives its name from the fact that many had **b**lack tissue from gangrene. Large buboes, and inflammatory swellings of lymphatic glands, especially in the groin or armpit, were common. Another variant caused by Y. pestis is the **Pneumonic Plague**, caused by breathing particles containing Y. pestis into the lungs, which leads to death from pneumonia and its complications. That was more infectious since it could be spread from person to person. Modern antibiotics are used to treat the plague, which still occurs.

Measles: A disease not shown in the figure is measles, which probably has killed upwards of 200 million people throughout time. It emerged from a viral infection, rinderpest, which infects cattle, deer, and buffalo. In 2021 there were 128,000 deaths out of 9 million cases worldwide, even though there is a highly effective vaccine. Vaccinations have decreased since the Covid-19 pandemic. Since it is one of the most infectious viruses known, and one contract leads to life-long immunity, a large population (250,00-500,000) is needed for it to self-sustain. The most recent analysis of historical sequences suggests that it emerge (jumped to humans) around the 6th century BCE, around

the time when cities of high enough population formed to allow its emergence. Measles is caused by an RNA virus, and since RNA is much more labile than DNA, few historical traces of the measles virus are available. The oldest one is from 1911 and it was from this and newer viruses that a phylogenetic RNA tree using a molecular clock model was constructed that led to the 6th century BCE time of emergence. Cities with a critical number of people for sustaining an emergence existed about 300 BCE in North Africa, India, China, Europe, and the Near East. A disease similar to measles was mentioned by Rhazes (Persia, 10th century CE). Past pandemics of unclear etiology could be attributed to measles, but it's difficult to know for sure given the difficulty in differentiating measles from other diseases.

Influenza: The genome of the influenza virus consists of 8 separate segments of ssRNA, much like the human genome resides on 23 different "segmented" chromosomes. Because its genome consists of RNA, past traces of it that point to its origin is lacking. The human influenza virus arose from swine (causing swine flu) and birds (avian flu). Hippocrates wrote of a disease with similar symptoms in 412 BCE. In 1357 an epidemic called “influenza di freddo,” or cold influence, swept Florence, Italy. The influenza RNA genome and transcribed proteins are shown in Figure 32.14.5 below.

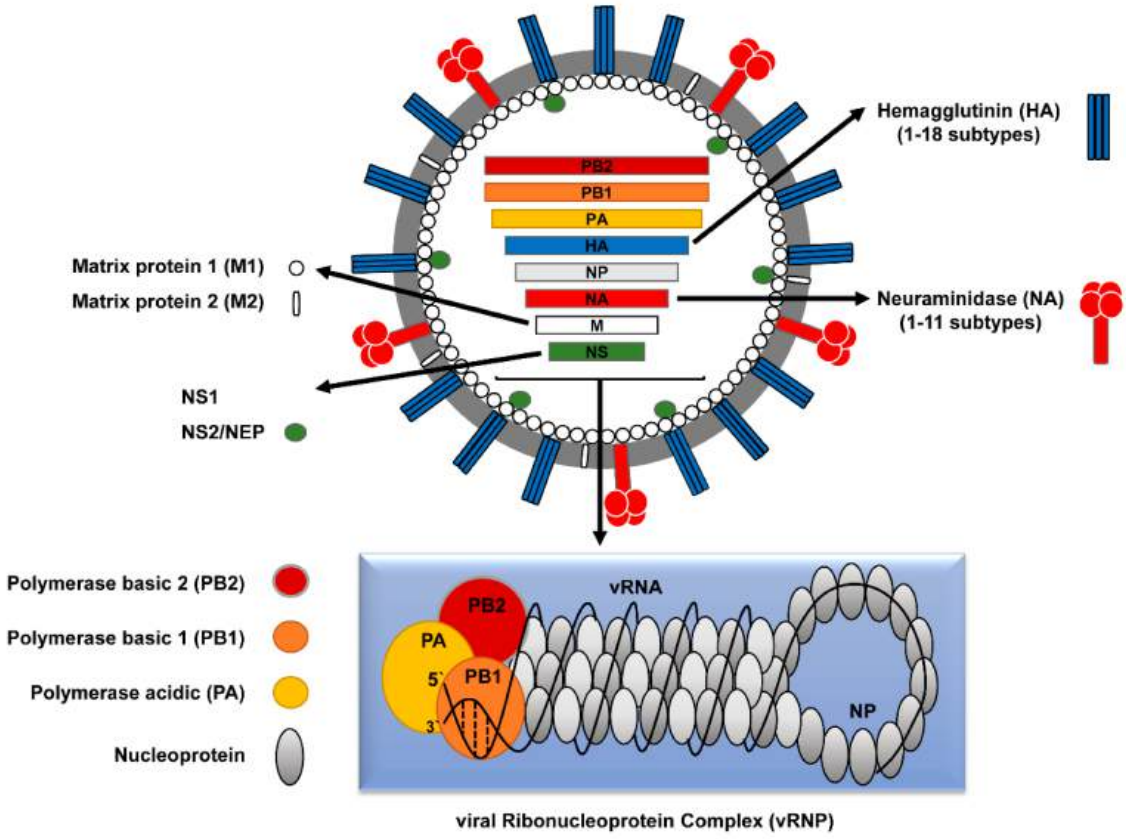


Figure 32.14.5: Influenza RNA genes and their protein products. Ahmed Mostafa, Elsayed M. Abdelwhab, Thomas C. Mettenleiter, and Stephan Pleschka - mdpi.com/1999-4915/10/9/497/htm, CC BY 4.0, https://commons.wikimedia.org/w/inde...curid=92987475

The **hemagglutinin** (H) membrane protein, responsible for viral binding to host cells, and **neuraminidase** (N), required for the exit of newly replicated viruses and hence viral propagation, are especially key in understanding past and future pandemics. There are 18 different subtypes of H and 11 subtypes of N comprising 4 different types of viruses (A-D) with A and B being the most common. The main types in circulation in 2022 were types A (H3N2) and B (H1N1).

Since the genome consists of RNA replicated by a RNA polymerase, which does not have proofreading functions, mutations occur on viral replication. This leads to slow changes in viral protein sequence and structure, called **antigen drift**, and hence to viruses less recognized by the host immune system. This is why new influenza vaccines are formulated each year (through a process that requires growing the virus in eggs).

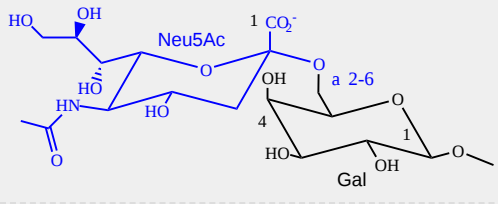
Large-scale pandemics occur through **antigen shifts**. This occurs when an animal such as a pig gets infected with an avian virus, a not unlikely occurrence given the co-farming of these animals in many places in the world. Newly replicated pig viruses could then contain some avian viral segments, which when transmitted to humans could produce lethal disease since they have no immunological memory in the host to produce an immediate immune response. Analyzes show that the horrific 1918 flu pandemic was caused by an avian influenza virus. An ancestral virus from the late 1880s is related to the horse (equine) H7N7 and equine H3N8 as well as to birds, humans, and swine viruses, and was the likely precursor of the 1918 flu virus. This ancestral virus led to a global change in the avian influenza virus which

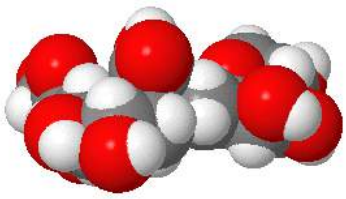
contributed most of the RNA segments to the 1918 pandemic. Smaller pandemics in the last half of the 20th century were likely caused by quick replacements of H3N2 and H1N1 genes leading to evolutionary fitness and ease of transmission. We should be on guard as there is an ongoing, worldwide highly virulent avian flu (H5N1) pandemic in wild birds and domestic poultry that has jumped to some animals, including humans who handle infected birds.

The hemagglutinin protein, homotrimer (3 identical protein subunits), MW 220,000, is the most abundant protein on the viral surface. Only three have adapted to humans in the 20th century, giving pandemic strains H1 (1918), H2 (1957), and H3 (1968). Three recent avian variants (H5, H7, and H9) can jump directly to humans but have low human-to-human transmissibility.

The viral hemagglutinin binds to glycoprotein receptors on human and other animal cells. The receptor binding site on host cells contains a terminal sialic acid (Sia) covalently attached to a galactose. The sialic acid is usually connected through an α(2,3) or α(2,6) link to galactose on N-linked glycoproteins. The viral subtypes found in avian (and equine) influenza bind preferentially to host Sia α(2,3) Gal which predominates in the avian GI tract where viruses replicate. Human influenza binds preferentially to Sia α(2,6) Gal links on human cells. The swine influenza HA binds to both Sia α(2,6) Gal and Sia α(2,3) Gal. The structures of the Sia-Gal disaccharide are shown in Figure 32.14.6 below.

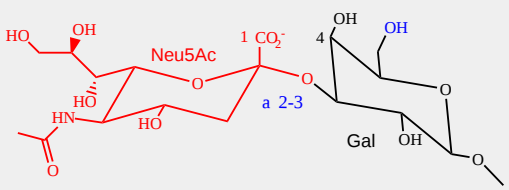
Sia α(2,6) Gal (Human and swine)

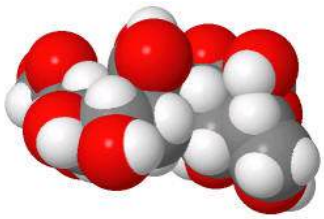




(made with Sweet, with an OH, not AcNH on sialic acid on C5)

Sia α(2,3) Gal (Avian and Swine)





(made with Sweet, with an OH, not AcNH on sialic acid on C5)

Figure 32.14.6: Structures of Sia α(2,6) Gal (human) and Sia α(2,3) Gal Gal (avian/swine)

The H5N1 avian virus is deadly but lacks human-to-human transmissibility. Why? One reason is that it appears to bind deep in the lungs and is not released easily on coughing or sneezing. It appears that cell surface glycoproteins deeper in the respiratory tract have Sia (α2,3) Gal linkages which account for this pathology.

Small changes in the amino acids of the viral hemagglutinin (HA) could change the preference for binding between Sia α(2,6) Gal (predominant human form) and Sia α(2,3) Gal (predominate form in birds) on host cells, and could dramatically affect both human lethality and transmission. Even though it was mostly of avian origin, the predominant 1918 hemagglutinin bound to the human Sia α(2,6) Gal.

Figure 32.14.7 shows an interactive iCn3D model of a H1 1918 hemagglutinin with a human receptor (2WRG), in this case, just the Sia α(2,6) Gal disaccharide from a target N-linked glycoprotein.

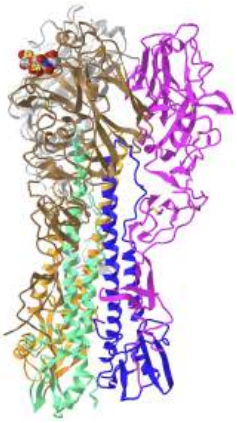


Figure 32.14.7: H1 1918 hemagglutinin with the human receptor - the Sia α(2,6) Gal dissachharide (2WRG). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...svY14r7wmnM1r5

The HAs in each of the 20th-century influenza pandemics, 1918 (H1N1), 1957 (H2N2), and 1968 (H3N2), preferentially bound to the Sia α(2,6) Gal even though the 1918 viruses and presumably the other, arose from avian viruses with a Sia α(2,3) Gal preference.

Figure 32.14.8 shows an interactive iCn3D model of α-2,6-linked sialyl-galactosyl ligand binding to the H1 1918 hemagglutinin (2WRG).

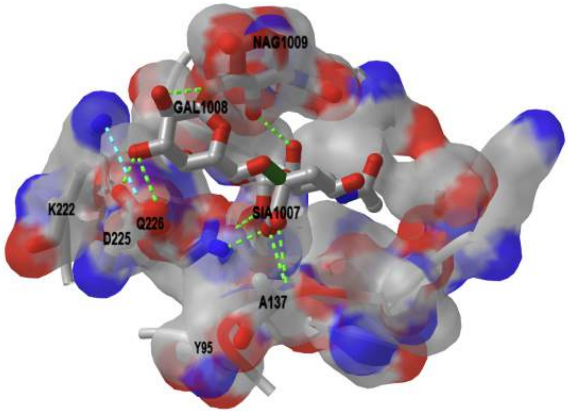


Figure 32.14.8: α-2,6-linked sialyl-galactosyl ligand binding to H1 1918 hemagglutinin (2WRG). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...yNJHQaQQebSuT9

A variant of the 1918 virus, A/South Carolina/1/18 (18H1), also circulated at the time. It contained a single amino acid mutation, D225G. in the HA protein. That variant switched the HA binding specificity on its target from Sia α(2,6) Gal to both Sia α(2,6) Gal and Sia α(2,3) Gal. This change eliminated a salt bridge (ion-ion interaction) between K222 and D225 in the main variant (see the blue-dotted line in the above model). This in turn allowed another key residue, Q226, to bind to the host receptor.

The viral HA in the 2009 human influenza pandemic had K222 and D225, giving it specificity for Sia α(2,6) Gal. Late in that pandemic (as occurred in the 1918 pandemic), a mutated version, D225G, that produced more severe symptoms was isolated. It had also gained dual specificity. Another mutant D225E did not as the salt bridge was maintained and the binding to Sia α(2,6) Gal was actually strengthened. Binding studies showed that the D225G mutants in the HA of both 18H1 and 09H1 viruses bound with higher affinity than the wild-type HAs which likely allowed binding to host glycoproteins deeper in the lung.,

The dissociation constant K_D and the on rate, k_{on}, and off rate, k_{off}, for the 09H1 and 18H1 hemagglutinins and relevant mutants, were determined by surface plasmon resonance spectroscopy (see Chapter 5.2 for a review of SPR). Table 32.14.1 below shows their values.

Hemagluttinin	Sia-Gal link	K_D (μM)	k_{on} (s^{-1})	k_{off} ($M^{-1}s^{-1}$)
09H1	α(2,6)	3.74	319	0.00119
09H1	α(2,3)	nd	nd	nd
09H1 D225G	α(2,6)	0.475	3650	0.00173
09H1 D225G	α(2,3)	2.24	1460	0.00327
18H1	α(2,6)	13.7	125	0.0017
18H1	α(2,3)	nd	nd	nd
18H1 D225G	α(2,6)	8.35	531	0.00444
18H1 D225G	α(2,3)	4.73	984	0.00466

Table 32.14.1: Dissociation and are constants for the interaction of hemagglutinins (H) from the 2009 and 1918 pandemics with Sia-Gal ligands. Adapted from Zhang et al., J Virol. 2013 May;87(10):5949-58. doi: 10.1128/JVI.00545-13. Epub 2013 Mar 20. PMID: 23514882; PMCID: PMC3648181.

It's remarkable how one amino change that can lead from no to strong binding can alter the specificity of a protein for its ligand and human history as well.

32.14.2: SLOWER-ACTING BUT VERY LETHAL MICROBIAL DISEASES HAVE TAKEN A VAST NUMBER OF LIVES

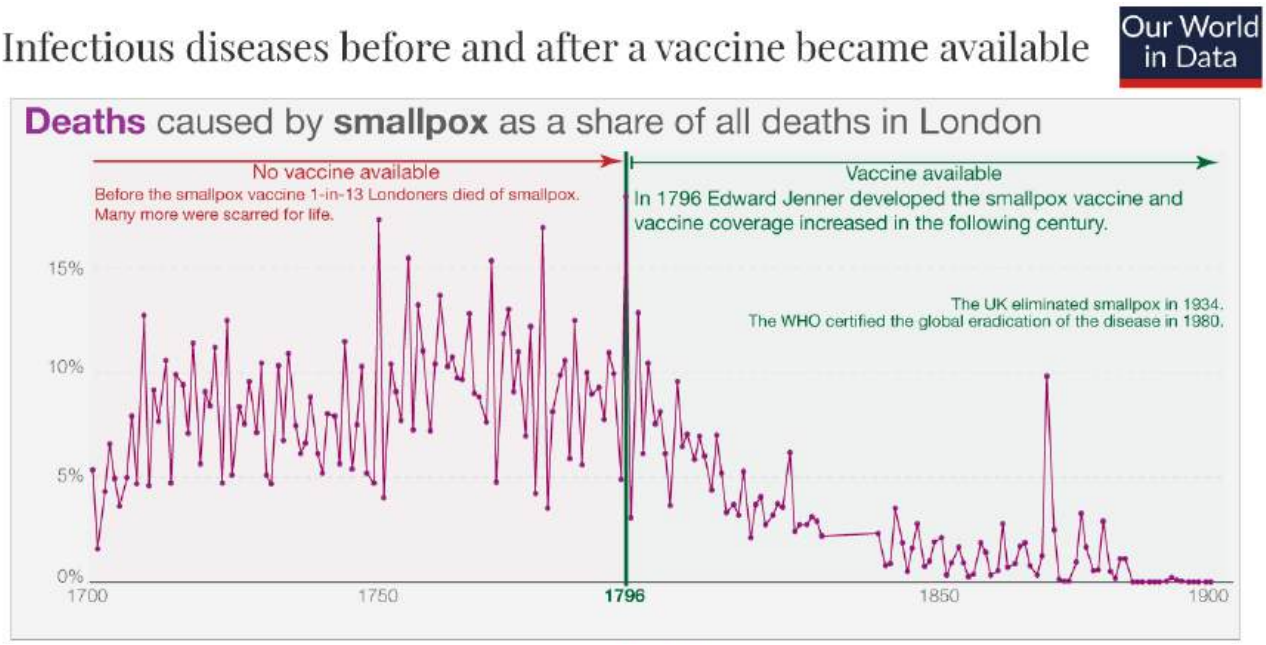
Malaria: Each year there are an estimated 300-500 million cases that result in about 2.7 million death. Most deaths are children under 5 in sub-Saharan Africa. The disease is caused mainly by the female Anopheles mosquito which transmits Plasmodium falciparum and the less lethal *Plasmodium vivax*. In the 100 years of the 20th century between 150-300 million deaths have been attributed to malaria (2-5% of all deaths). It was brought to the new world from Africa by the slave trade of over 7 million Africans, and from Portugal and Spain (the main colonial powers where malaria was endemic. The bacteria probably moved from gorillas to humans long ago in Africa. No effective vaccine has yet been developed to prevent this disease.

Proteins associated with the virus have been found in Egyptian samples from 3200 BCE and there were descriptions of the cyclic fevers associated with malaria in China in 270 BCE. It was also described by Homer (750 BCE), Plato, and Hippocrates in ancient Greece. It probably was first found in Rome around 0-100 ACE. The virus persisted in Europe for 2000 years.

Tuberculosis: This disease is caused by the *Mycobacterium tuberculosis* bacteria and is spread through the breath. Estimates are that up to 1 billion people have died of TB over history. The BCG vaccine is somewhat effective against TB but not often administered given its low prevalence and the availability of antibiotics. Tuberculosis (TB) was called “phthisis” in ancient Greece, and “tabes” in ancient Rome. The modern common ancestor of this bacteria arouse around 6000 years ago and is associated with disease in both the Old and New World.. Older strains were likely found in seals and sea lions. Genetic analysis showed that the modern strain was found in Peru before the arrival of Europeans to the New World. The disease in the Western Hemisphere probably derived from sea mammals which crossed the ocean.

Vaccines against some of our worst infectious disease agents have saved millions of lives. Here are some examples.

Figure 32.14.9:



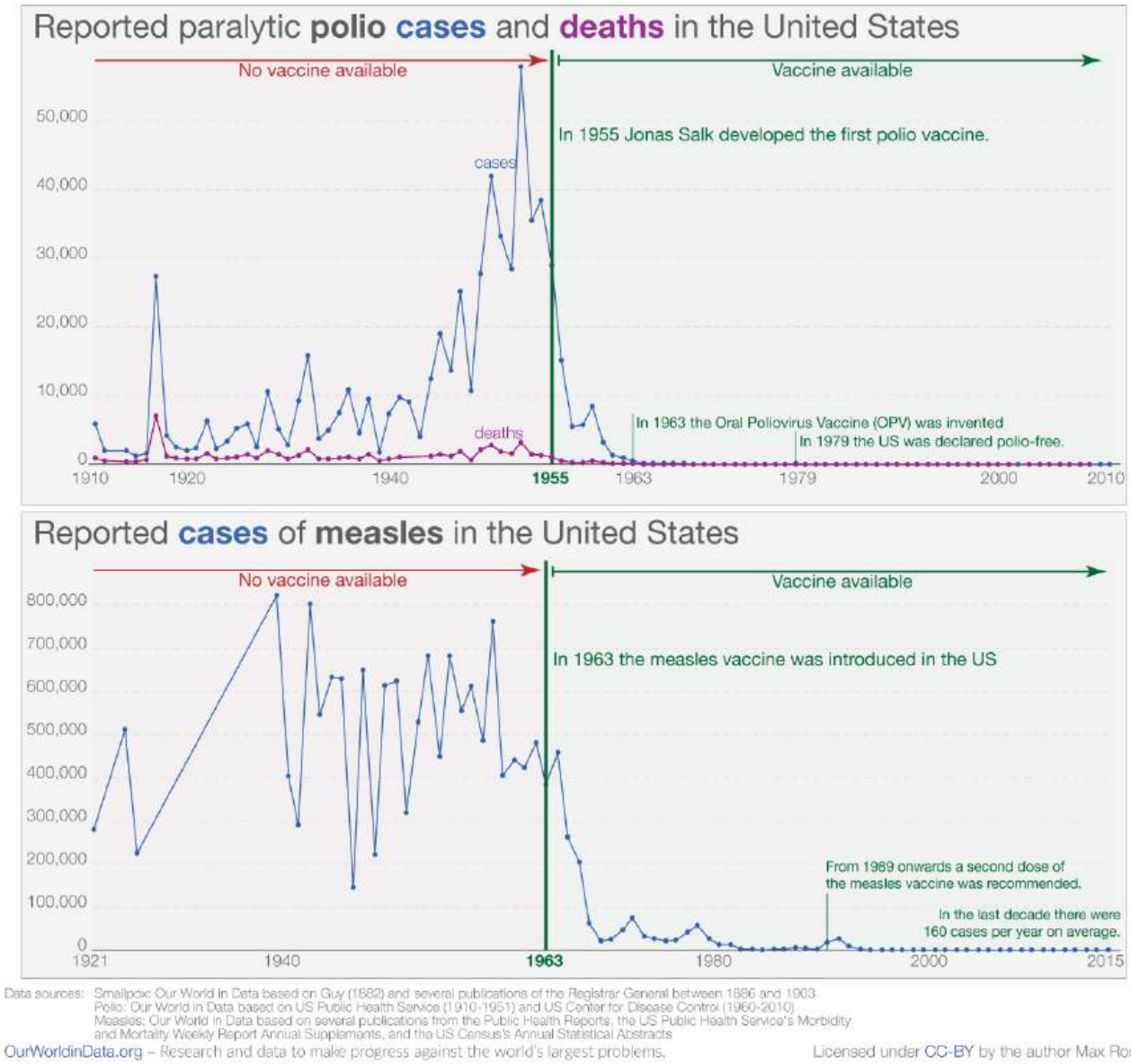


Figure 32.14.9: : https://ourworldindata.org/microbes-...ience-vaccines

Mathematical models show that from 12/20 through 12/22, Covid vaccines prevented over 120 million infections, 18.5 million hospitalizations, and 3.2 million deaths just in the United States. In the first year of the pandemic (12/20-12/21), models show that 14.4 million deaths (and 19.4 million excess deaths) were prevented in the whole world.

EPIDEMICS THAT DECIMATED INDIGENOUS PEOPLES IN THE NEW WORLD

Before Columbus came to the New World, there was no typhoid, flu, smallpox, or measles there. These diseases were present in Eurasia where people lived in increasingly populated areas in close quarters with domesticated animals. They would have developed some immunity over time. Their microbes likely derived from domestic animals before jumping species to humans, much as modern flu can be passed from swine to humans and less regularly but more lethally from birds to humans. Even with the buildup of some immunity, new pandemics were utterly devastating.

Indigenous peoples in the new world were never exposed to these pathogens before the arrival of people from the Ols World. They only utilized llamas for work and not generally for food and milk. Deaths were staggering. It's estimated that 90% of indigenous people died, a far higher proportion than seen even with the Black Plague in Europe. Imagine the loss of culture and civilization that would accompany a decline in the population of central Mexico from 15 million to 1.5 million in the 100 years after 1519.

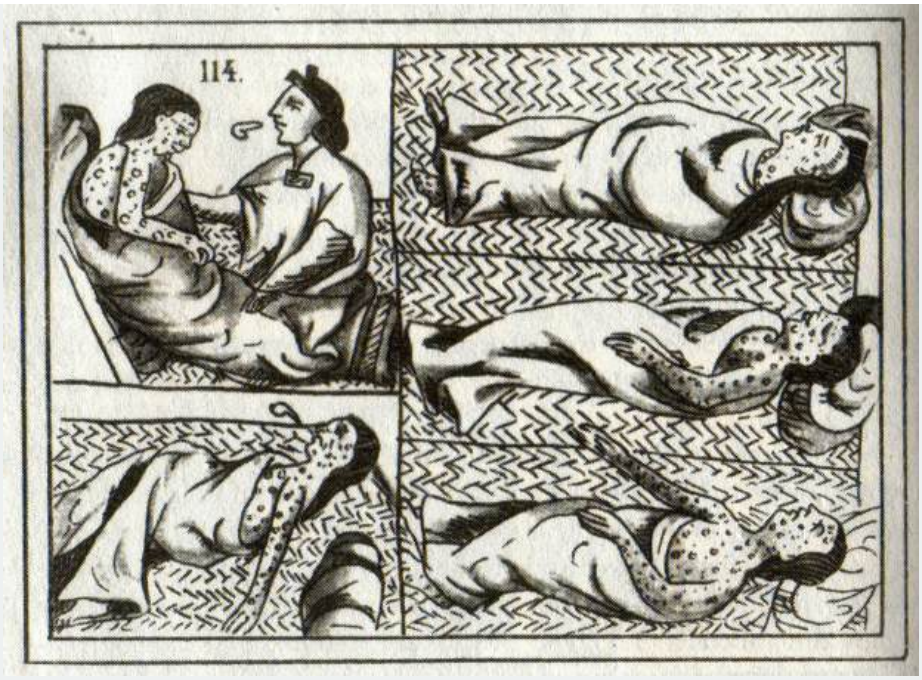


Figure 32.14.10: Sixteenth-century Aztec drawings of victims of smallpox. https://en.wikipedia.org/wiki/Native..._and_epidemics

Social conditions after the initial collapse of the indigenous people in the Americas led to their continued decline, even though they would have gained some immunity. An example is offered by Ostler who describes the health consequences of the Indian Removal Act of 1830, which led to the forced relocation of Native people east of the Mississippi River into "Indian Territory" (Oklahoma and Kansas). As an example, 16,000 Cherokee were expelled and forced to live in camps with few resources, where up to 2000 died of measles, malaria, dysentery, and whooping cough. 1500 more died as they moved west. More died in Oklahoma, leading to a death toll of 25% of the original group.

Cumulative death rates in the COVID-19 pandemic show that Indigenous peoples in the United State still have barriers to optimal health care.

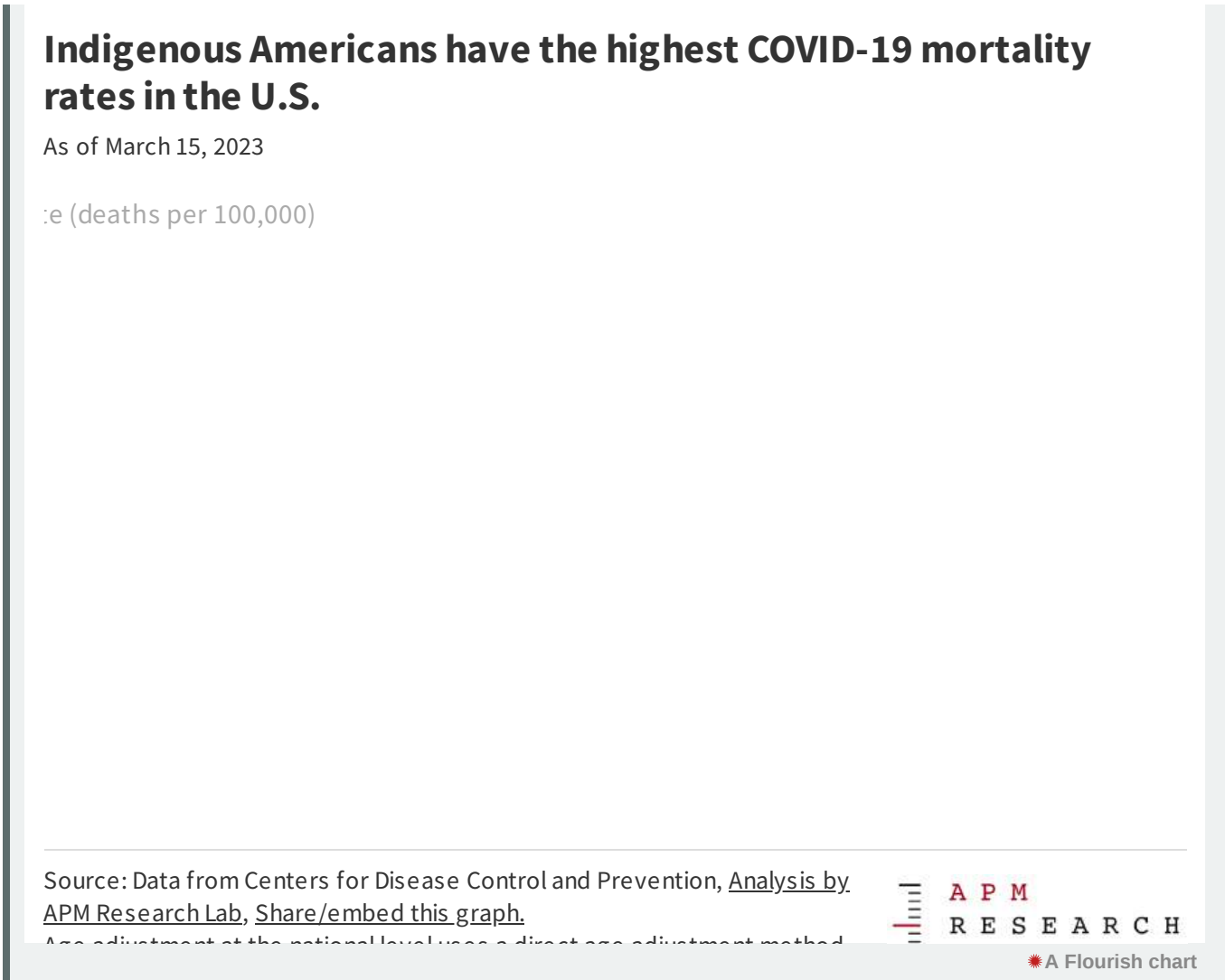


Figure 32.14.11 : Cumulative Deaths and Age-Adjusted Rates per 100,000 in the United States.

32.14.3: INFECTIOUS, EMERGING AND PANDEMIC DISEASES - LINKS TO CLIMATE CHANGE

Our understanding of infectious diseases clearly shows that the great epidemics and pandemics of the world have arisen when microbial pathogens make the jump from animals to humans who have not experienced them before. For example, HIV/AIDS arose when simian immunodeficiency viruses, to which non-human primates were adapted, jumped to humans in central Africa. The best available data suggest that the SARS-CoV-2 virus jumped from bats to animals (raccoon dogs or other animals from Wuhan China live animal markets) and then to humans, although some data suggest the possibility of a lab leak.

A **Zoonotic** disease (**zoonosis,** plural **zoonoses**) is a microbial infectious disease transmitted reversibly between animals and humans. The major types of zoonoses are viral, bacterial, parasitic, mycotic/fungal, rickettsial (obligate intracellular Gram-negative bacteria found in ticks, lice, fleas, mites, chiggers, and mammals), Chlamydial (bacteria that cause STDs), Protozoal or unconventional (such as prions). The ones most prevalent in the US are influenza, Salmonellosis, West Nile virus, Plague, coronaviruses, rabies, Brucellosi, and Lyme disease. **Vector-borne** diseases are caused by bacteria, viruses, and parasites transmitted through bites of **vectors** such as infected arthropods like mosquitoes, ticks, sandflies, and blackflies. The range of arthropod vectors expands with global warming as they are cold-blooded.

Several anthropogenic (human-caused) factors, including climate change, increase the chances of such jumps. These factors, many of which are interrelated include:

- movement of humans into environments where contact with disease-carrying organisms would increase transmission

- biodiversity loss which allows species and their microbes to move into new areas
- land use change (deforestation, farming, etc) that allows the expansion of species and microbes into new areas
- global warming, which encourages the movement of species and their microbes to new areas where human exposure is more likely
- climate change-induced changes in plant life that allow altered distributions of animals and microbes
- climate change-derived changes in precipitation patterns that affect the adaptation of species and their microbes.

Humans affect all of these factors by causing climate change, and land use changes including the expansion of agriculture, urbanization, and the rapid global movement of people, commodities, and other animals. Studies have shown that 58% of human infectious diseases have already worsened with climate change. Another study used databases of mammalian viruses and their host to see which ones might share viruses, an occurrence made much more likely when the species live in the same geographic area. Machine learning was used to model how mammals might share viruses and change their living range in a warming world through 2070. The study found over 4000 viruses could move among 3000 species, greatly enhancing the changes for the exchange of single and multiple viruses among species. Bats (see below) are especially worrisome as they harbor many viruses capable of infecting humans. As bats move habitats due to climate warming, their chances of infecting new species that could then infect humans are greatly increased.

25 years of land use changes in Australia led to altered bat (flying foxes) behavior and to their more permanent presence in agricultural land. This has resulted in viral"spillover" (transmission of a pathogen from a non-human vertebrate to a human) that is driven by periodic food shortages, especially in winters following El Nino weather patterns (characterized by less rain, warmer temperatures, and greater temperature extremes). These changes in bat behavior led to the emergence of the Hendra virus, which infected domestic horses (an intermediate vector), and could pass the virus to humans. The virus does not cause disease in bats but leads to a high mortality rate in horses (75%) and humans (57% based on just four deaths). With climate and land use change, bats persistently spent winters in agricultural lands close to horses. Spillovers occurred more frequently during low food conditions following an El Nino summer.

The Black Death (Second Great Pandemic, 1347-1351), occurred during the Little Ice Age in Europe (1300-1850), which also led to a great famine from 1315-1322. There is a link between the pandemic and climate change, but it's difficult to ascertain the strength of the association. An association exists between periodic warm springs and wet summers in Central Asia (using tree-ring data) and outbreaks of the Plague in Europe about 15 years later. This suggests a continual re-importation of *Yersinia pestis* in Asian rodents into Europe and could explain how long the Plague lasted in Europe.

The presence of the plague in gerbils in Kazakhstan would increase with warm spring and wet summer which increase gerbil and flee populations. This Moran effect (time correlation of two populations of a species with change in environment) is well known in population ecology. When gerbil populations collapse, the flea density of the remaining gerbils increases which also leads them to seek different hosts. The spread across geographic distance would take time. In the case of periodic import of flees to Europe, it has been proposed that the 4000 km from west central Asia to the Black Sea took 10-12 years (around 350km/y) .

PATHOGENS FROM THE NORTH

Most attention has been given to pathogens moving northward from the south as warmer temperatures allow them to thrive in traditionally colder climates. There is growing parallel concern about pathogens moving south from the Arctic as it warms. In fact, the high northern latitudes have experienced the greatest increase in temperature as the planet warms. The Arctic is predicted to be soon ice-free in the summers.

One major concern is that thawing of the permafrost (comprising almost 1/4 of the northern hemisphere that is "permanently" frozen) will allow the release of CO_2 (a metabolic product of microbes in the presence of oxygen) and CH_4 (a metabolic product of Archeal microbes in an anoxic environment) from organic molecules previously found in frozen soils. (A similar concern is the release of "frozen" methane clathrates from a warmer ocean).

An emerging concern is the "activation" of microbes from the permafrost that have been sequestered and dormant for 500,000 years or more. In fact, species like tardigrades, rotifers, and nematodes, that can go dormant and enter a state called **cryptobiosis** in harsh conditions such as freezing and dehydration, can reactivate. Two species of nematodes (roundworms), dating back to 46,000 years ago (based on radioactive dating), the last years of the Pleistocence, were recovered and revived. Periodic reoccurences of anthrax from the release of the Gram-positive bacterium *Bacillus anthracis* spores from permafrost thawed in summers have been reported. In 2016, when average temperatures were significantly elevated (see Figure 32.14.12 below), over 2000 reindeer died (close to a 90% mortality rate) and close to 100 people were hospitalized from an anthrax outbreak in Siberoia. On June 20, 2020, a record temperature of 38 °C (100 °F) in the Russian town of Verkhoyansk was recorded!

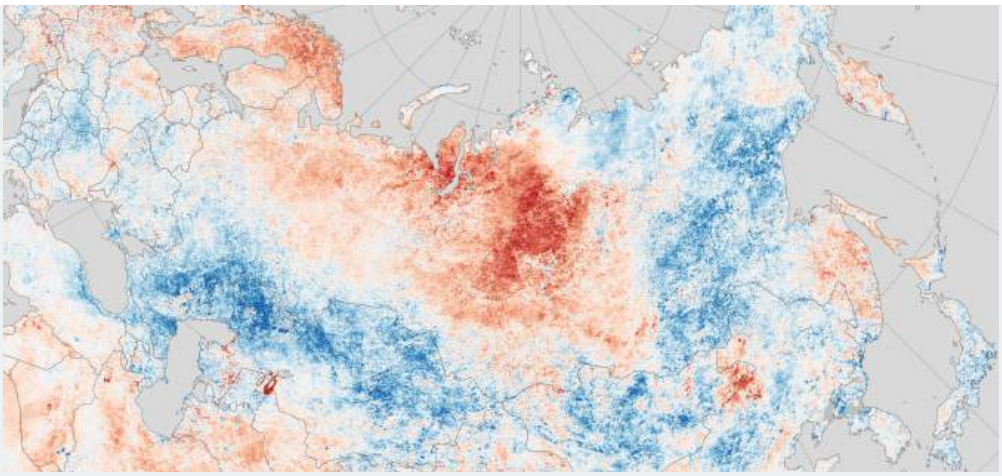


Figure 32.14.12 : Color map showing land surface temperature anomalies from -12 °C (-21.6 °F) (darkest red) to + 12 °C (+21.6 °F) (darkest red) during the week of July 20-27, 1916. https://earthobservatory.nasa.gov/im...n-extreme-year

We also have to worry about emerging viruses that are released from the thawing of the tundra. Just as for the SARS-Covid-2 virus, we would have no immunity to these viruses. Figure 32.14.13 below shows EM pictures of new infectious viruses isolated from seven different ancient Siberian permafrost samples.

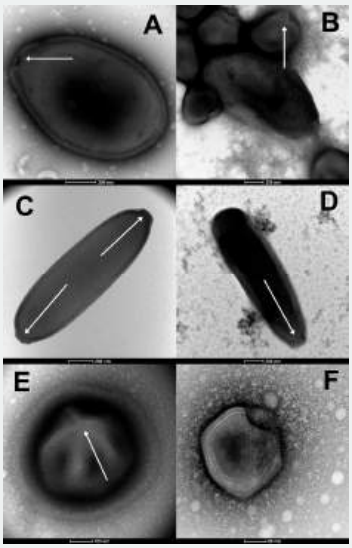


Figure 32.14.13 : Morphological features guiding the preliminary identification of newly isolated viruses (negative staining, TEM). Alempic, J.-M.; Lartigue, A.; Goncharov, A.E.; Grosse, G.; Strauss, J.; Tikhonov, A.N.; Fedorov, A.N.; Poirot, O.; Legendre, M.; Santini, S.; et al. An Update on Eukaryotic Viruses Revived from Ancient Permafrost. *Viruses* **2023**, *15*, 564. https://doi.org/10.3390/v15020564. Creative Commons Attribution (CC BY) license (**https://creativecommons.org/licenses/by/4.0/**).

Panel (**A**) shows large ovoid particle (1000 nm in length) of *Pandoravirus yedoma* (strain Y2) (showing the apex ostiole (white arrowhead) and the thick tegument characteristic of the *Pandoraviridae* family.

Panel (**B**) shows a mixture of *Pandoravirus mammoth* (strain Yana14) oblate particles and of *Megavirus mammoth* (strain Yana14) icosahedral particles exhibiting a “stargate” (white starfish-like structure crowning a vertex, white arrowhead)

Panel (**C**) shows the elongated particle of *Cedratvirus lena* (strain DY0) (1500 nm in length) exhibits two apex cork-like structures (white arrowheads)

Panel (**D**) shows the elongated particle of *Pithovirus mammoth* (1800 nm in length) exhibiting a single apex cork-like structure (white arrowhead).

Panel (**E**) shows the large (770 nm in diameter) “hairy” icosahedral particle of *Megavirus mammoth* (strain Yana14), showing the “stargate” (white arrowhead) characteristic of the *Megavirinae* subfamily

Panel (**F**) shows the smaller icosahedral particle (200 nm in diameter) of *Pacmanvirus lupus* (strain Tums2) typical of asfarviruses/pacmanviruses.

32.14.4: FUNGAL DISEASES

We have concentrated solely on bacteria and viral epidemics/pandemics. We also have to consider fungal outbreaks that affect human health but also the foods that sustain us. We have few medicines that treat fungal infections and no vaccines, so any outbreaks could be quite serious. Most grow optimally between 12-30 ^{0}C (54-86 ^{0}F) so they don't adapt easily to warm blooded animals like humans. Nevertheless, global warming will increase their range and lead to increased thermal tolerance.

Candida auris: This was first found in 2009 in Japan as a cause of an ear infection. It is now found around the world.

***Batrachochytrium dendrobatidis* (*Bd*)**: This affects amphibians and has caused a high loss in amphibian diversity on all continents.

Cryptococcus deuterogattii:, This was typically found in more tropical/subtropical climates but now is also found in western Canada and the Pacific Northwest. It causes infections in people and animals.

Puccinia striiformis: This causes wheat rust which devastates crops and is now moving into warmer areas.

Fusarium graminearum: This causes diseases in wheat and other food crops, especially in warm and wet conditions.

Coccidioides immitis: This fungus, which grows in desert soil, can also spread through severe dust storms that cause fungal spores to be blown over wide regions. An example is the dispersal of *Coccidioides immitis* from Bakersfield, where it was endemic, to Sacramento County, where it wasn't, in 1977. Another example is *Apophysomyces trapeziformis*, which caused disease in 2011 in Joplin, Missouri after a tornado. The fungus actively thrives in wet soils but forms spores in dry conditions These spores can last for decades and cause disease when blown into the air and inhaled. They are associated with a condition called Valley Fever which affects farm workers who spend much time outside. Latino, Asian, and Native American people get Valley Fever at 2-4 times the rate of others.

32.14.5: BATS, VIRUSES, AND CLIMATE CHANGE

We have seen that new infectious diseases arise from pathogen jumps to humans from other species. The more distant the species, the more unlikely humans have encountered the disease and the more likely it could cause severe illness and pandemics. A clear example is the avian flu that led to the 1918 flu pandemic. Yet we also have to worry about zoonoses from pathogen transfer from mammals, including rodents, bats, moles, shrews, monkeys, pigs, camels (a host of the deadly MERS virus), whales, cats, dogs, and seals (a likely source of the original TB virus).

Bats are a key source of zoonotic disease, including Middle East respiratory syndrome (MERS), which has a death rate of around 35%. Bats are the source of the Covid virus MERS-CoV which causes MERS. The virus spreads to people from camels. Severe acute respiratory syndrome (SARS) is another coronaviral disease, caused by the called SARS-associated coronavirus (SARS-CoV) which emerged in China in 2003. It had a death rate of around 12% but it was much higher in older people. Neither of these became lengthy full-blown pandemic, as with the SARS-CoV2 virus, the cause of the COVID-19 pandemic. In addition, viruses from bats include rabies, Ebola, and Marburg viruses, as well as the Nipah and Hendra viruses. Bats are more likely to be infected with zoonotic viruses than rodents.

Why are bat viruses so key in our worst zoonotic diseases? Two features are important. The same viruses that are so virulent to humans do not kill bats. A clue as to the special nature of bats is that they are the only flying mammals. What might protect bats from their own viruses is that their core temperatures are quite elevated during flight, which requires a high metabolic rate. These high temperatures likely prevent these viruses from harming bats but also make the viruses immune to the high-temperature fevers accompanying infection in humans. The average core temperature of flying bats derived from a variety of species was 39.6 ^{0}C or 103.3 ^{0}F. Many pathogens replicate optimally at temperatures less than normal body temperature.

Fevers in humans are regulated by the hypothalamus, mainly through prostaglandin E2 (PGE2). This response is part of the innate immune response and is elicited by most pathogens. PGE_2 binds to the E-prostanoid-3 receptor (EP3), a G protein-coupled receptor in the hypothalmus, which determines the "set point" for body temperature. Hypothalmic PGE_2 is produced from the endocannabinoid 2-arachidonylglycerol by the action of monoacylglycerol lipase, at least in mice stimulated with a bacterial cell wall component (LPS), that stimulates fever production.

Figure 32.14.14 shows an interactive iCn3D model of the human prostaglandin E receptor EP3 bound to prostaglandin E2 (6AK3).

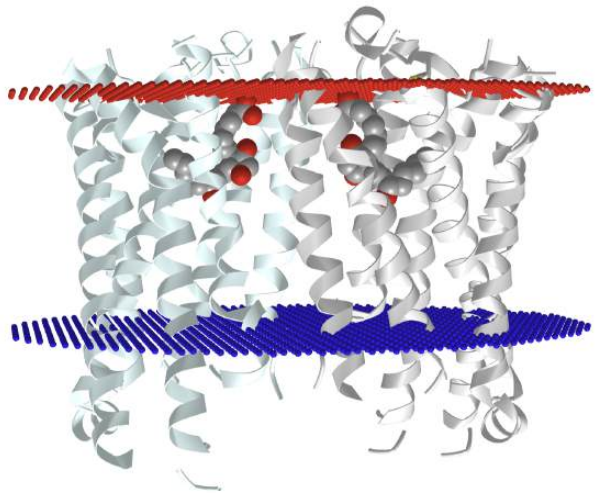


Figure 32.14.14 : Human prostaglandin E receptor EP3 bound to prostaglandin E2 (6AK3).. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...cCgTGEKu3KoGx9

The model shows a dimer of two, GPCRs each bound to 1 PGE2.

Research suggests that bats have also evolved to have a lower inflammatory response, mediated by the inflammasome (discussed in detail in Chapter 5.4). Here is a short review of the inflammasome modified from that chapter section. It's needed to give readers a more biochemical explanation for immunosuppression in bat cells, a topic critical to understand the role of bats in present and future pandemics.

The inflammasome, part of the innate immune system, is activated by a plethora of pathogens or damaged host molecules. Our innate system immune cells (dendritic cells, macrophages, eosinophils, etc) have receptors that recognize common **pathogen-associated molecular patterns (PAMPs)** such as lipopolysaccharides (LPS) on the surface of bacteria, mannose on bacteria, and yeast, flagellin from bacterial flagella, dsRNA (from viruses) and nonmethylated CpG motifs in bacterial DNA. These antigens are recognized by pattern recognition receptors (PRRs) - specifically the **Toll-like Receptors (TLRs)** 1-10. These include plasma membrane TLRs (TL4 for LPS, TL5 for flagellin, TLR 1, 2 and 6 for membrane and wall components of fungi and bacteria) and intracellular endosomal TLRs (TLR3 for dsRNA, TLR 7 and 8 for ssRNA and TLR9 for dsDNA).

Figure 32.14.15 shows the TLR family, their binding signals, and intracellular adapter proteins used to transmit signals into the cell.

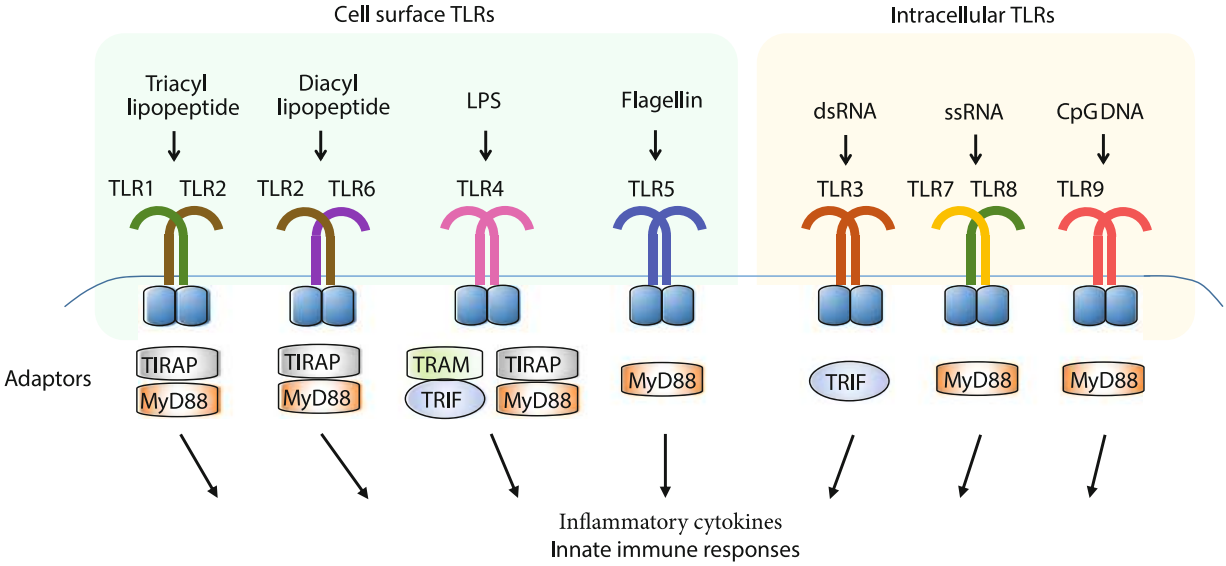


Figure 32.14.15: TLR family, their binding signals, and intracellular adapter proteins Ji-Yoon Noh, Suk Ran Yoon, Tae-Don Kim, Inpyo Choi, Haiyoung Jung, "Toll-Like Receptors in Natural Killer Cells and Their Application for Immunotherapy", *Journal of Immunology Research*, vol. 2020, Article ID 2045860, 9 pages, 2020. https://doi.org/10.1155/2020/2045860. This is an open access article distributed under the Creative Commons Attribution License

Inflammasomes are also activated by **Damage-associated molecular patterns (DAMPs).** These are typically found on molecules released from the cell or intracellular compartments on cellular damage (hence the name DAMP). Many are nuclear or cytoplasmic proteins released

from the cells. These would now find themselves in a more oxidizing environment which would further change their properties. Common DAMP proteins include heat shock proteins, histones and high mobility group proteins (both nuclear), and cytoskeletal proteins. Here are some other common non-protein DAMPS that can be released on cellular damage: ATP, uric acid, heparin sulfate, DNA, and cholesterol crystals. In the wrong location, these can be considered danger signals. They are sometimes referred to as "sterile" signals.

If TLRs recognize PAMPs, what recognizes DAMPs? They are recognized by another type of intracellular pattern recognition receptor (PRR) called **NOD** (Nucleotide-binding **O**ligomerization **D**omain (**NOD**)- **L**ike **R**eceptors or **NLRs**. NLRs also recognize PAMPs. The abbreviation **NLR** also comes from the **N**ucleotide-binding domain (NBD) and **L**eucine-**R**ich repeat (LRR)–containing proteins (**NLR**)s. This family of proteins participates in the formation of a large protein structure called the **inflammasome**. (Sorry about the multiple abbreviations and naming systems!)

As both PAMPs and DAMPs pose dangers, it would make sense that once they recognize their cognate PRRs (TLRs and NLRs, respectively), pathways leading from the occupied receptors might converge in a common effector system for the release of inflammatory cytokines from immune cells. Given that uncontrolled immune effector release from cells in an inflammatory response might be dangerous, it would be sometimes helpful to require two signals to trigger cytokine release from the cell.

Two such inflammatory cytokines are Interleukin 1-β (IL 1-β) and IL-18. Activation of TLRs by a PAMP leads to activation of a potent immune cell transcription factor, NF-kbeta, which leads to transcription of the gene for the precursor of the cytokine, pro-interleukin 1-β. Without a specific proteolytic cleavage, the active cytokine will not be released from the cell.

The protease required for this cleavage is activated by a signal arising when a DAMP activates a NLR, which then through a sequence of interactions leads to the proteolytic activation of another inactive protease, **procaspase 1**, on the **inflammasome**. The activated inflammasome activates procaspase to produce the active protein **caspase** (a cysteine-aspartic protease).

The convergence of the signals from the PAMP activation of a TLR and DAMP activation of a NLD at the inflammasome is shown in Figure 32.14.16 .

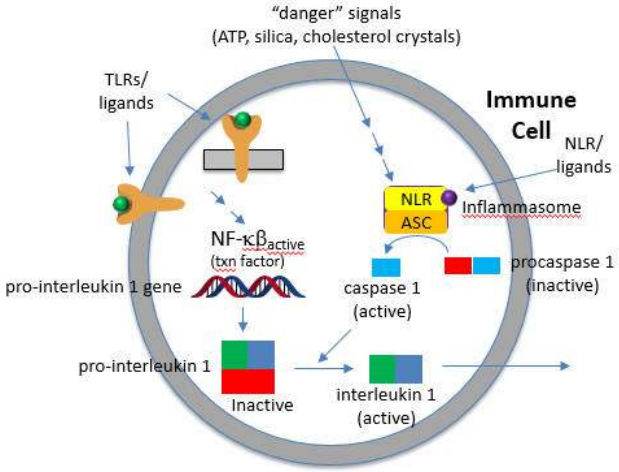


Figure 32.14.16 Signals from PAMP activation of a TLR and DAMP activation of a NLD at the inflammasome

The active cytokine interleukin 1-β helps recruit innate immune cells to the site of infection. It also affects the activity of immune cells in the adaptive immune response (T and B cels). Active IL-18 leads to the increase of another cytokine, interferon-gamma and it also increases the activity of T cells that kill other cells.

Figure 32.14.17 shows an interactive iCn3D model of the NLRP3 double-ring cage, 6-fold (12-mer) (7LFH)

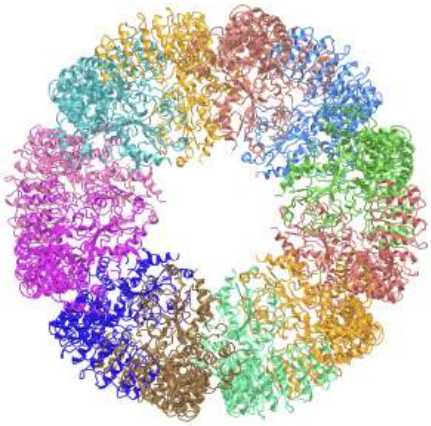


Figure 32.14.17 : NLRP3 double-ring cage, 6-fold (12-mer) (7LFH). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?DEbdkUoBtqRQ9bu59

The full-length mouse NLRP3 consists of 12- to 16-mer organized in a double-ring cage. It is held together by interactions between the leucine-rich repeats (LRR) domains. The pyrin domains are shielded by the structure, so they will not be activated without appropriate signals. The complex is also localized to the membrane. NLRP3 inflammasomes seem to be activated by cellular distress as well as cell exposure to pathogens. It is one of the main responders to a variety of microbial infections.

In summary, two signals are again needed:

Signal 1

The first signals are the bacterial and viral (influenza virus, poliovirus, enterovirus, rhinovirus, human respiratory syncytial virus, etc) PAMPs, which bind to TLRs and lead to the activation of the NFkb transcription factor. This activates not only the transcription of pro-interleukin 1-β and interleukin 18, but also the transcription of the NLRP3 sensor itself.

Signal 2

Signal 2 is delivered by PAMPs and DAMPs indirectly to the sensor NLRP3. This leads to the assembly of the inflammasome. These DAMPs appear to prime the activation of NLRP3 protein and subsequent formation of the active NLRP3 inflammasome. But what activates NLR3P3? After many studies, it became clear that the typical bacterial ligands that would activate TLRs and perhaps NLRs only prime NLRP3 for activation. They don't bind to it directly.

Extracellular ATP is a major activator of NLRP3. Nanoparticles are known to release ATP as well. Most studies show that K^+ efflux from the cell is an early signal.

32.14.6: BACK TO BATS

To survive the viruses they harbor, bats decrease their inflammatory response upstream at the level of PAMP and DAMP recognition, as well as downstream at the level of caspase-1 inhibition. In addition, additional cleavage sites in IL-1β cause its loss through proteolysis. These events decreased inflammasome signaling. That multiple steps are inhibited suggests that they have been selected through evolutionary pressures.

The activation of the bat NLRP3 inflammasome by a "sterile" agent (ATP), as well as 3 RNA viruses is significantly decreased in bat cells compared to responses to these signals in humans and mice cells. The viruses include:

- H1N1 influenza A virus (a negative-sense single-stranded RNA virus known to activate the NLRP3 inflammasome);
- the Melaka virus (a bat-borne zoonotic double-stranded RNA virus);
- MERS-CoV (a positive-sense (+) ssRNA zoonotic virus).

Even though the secretion of interleukin-1β is inhibited in bat cells, viral loads remained high in virally-infected bat cells. The altered NLRP3 activation in bat cells occurs in part through decreasing RNA splicing and an altered LRR domain in NLRP3.

It should be clear that if bat population increases or if they move to new habitats, both events which could be promoted by climate change, humans are at greater risks for bat-derived pandemics. Coronaviruses, with over 3000 species, make up about 1/3 of the bat's viral load. The region comprising southern Yunan (in China, Myanmar and Laos have very high populations of bats and it is from these areas that new bat zoonoses are likely to derive. Models of past climate (in the 1900s) and bat species richness compared to present climatic conditions, along with the knowledge of specific climate conditions necessary to support bat population and diversity, show that the greatest increase in bat species population in the 20th century occurred in SE Asia. Less important sites included regions in Central Africa and some pockets of Central and South America. The region in China, Myanmar, and Laos is likely the location for the origin of the SAR-CoV1 and SARS-CoV2 viruses.

32.15: Pandemic Diseases and Drug Discovery - Under Construction

Under construction - 9/26/23

Written by Valerie Doze (change title to what you want) by clicking pencil icon when hovering over the title.

Headings: Example

On the menu bar, select Normal, Heading 2

You can use Heading 3 for subheadings but adding sub-subheadings gets to look cluttered. For a sub-sub heading just use bold font for the title.

Numbering figures: Example

Figure 32.15.x below shows the reaction of one key carboxylase. x is the figure number you fill in to make them sequential in the document.

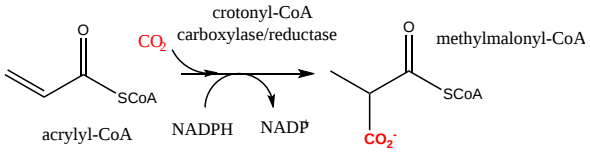


Figure 32.15.x Carboxylase used in the CETCH pathways to fix CO_2

---

iCn3D model: Example (to replace)

Figure 32.15.x shows an interactive iCn3D model of ribulose 1,5-bisphosphate carboxylase/oxygenase from Synechococcus PCC6301 (1RBL). (image is a png screen capture)

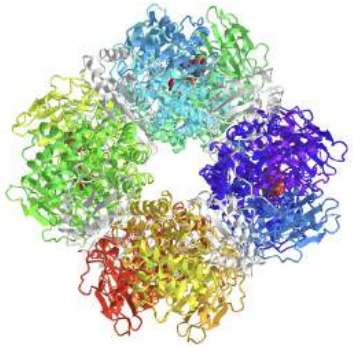


Figure 32.15.x: Ribulose 1,5-bisphosphate carboxylase/oxygenase from Synechococcus PCC6301 (1RBL) (Copyright; author via source). Click the image for a popup or use this external link:

I am sure you will have several to many iCn3D models. I can make them or teach you how to do it. Short but real learning curve. I sometimes use Pymol for some images when I feel that iCn3D doesn't show well.

Figure Types:

- png, jpeg, animated gifs (I can make them for you). Best file type is svg which can be made from PPTX images, ChemDraw, or exporting from other image file types using software like Adobe Illustrator or Affinity Designer
- for svg files, must drag them from a folder on your computer to desired spot in the new section. For other can drag them or use picture icon on menu bar to select and import them from your file.

References:

Incorporate as you like but essential for images et al directly used form a Creative Common CC BY document, for example.

32.16: FIXING CARBON FIXATION

Search Fundamentals of Biochemistry

CARBON FIXATION

Nature has produced many enzymes that can fix atmospheric CO_2, and, of course, everyone knows that photosynthesis is the source of most fixed CO_2 in the biosphere. Plankton, cyanobacteria, algae, and plants are key in removing atmospheric CO_2 through photosynthesis. Yet we can't plant enough trees to reduce CO_2 in the atmosphere in the next decade to avoid some of the worst consequences of anthropogenic climate change. Old-growth trees (mostly gone or under significant stress) are best at removing CO_2. New trees would take decades of growth before their effect on carbon drawdown would be consequential. We also need to fix carbon dioxide not only to decrease atmospheric CO_2 but also to make more food!

A lot of carbon (about 100 petagrams) is sequestered each year in net primary production (fixation of carbon into biomolecules). This is split almost equally between land and ocean organisms. The key enzyme in this process is Rubisco in the C3 (Chapter 20.4) and C4/CAM (Chapter 20.5) pathways. The enzyme is big containing many large subunits and an equivalent number of small ones. A chaperonin is required for folding. It is also a very slow enzyme with a k_{cat} of around 2-10 CO_2/sec. In addition, it can bind another substrate, O_2, and engage in a competing reaction of photorespiration as described in Chapter 20.4.

Figure 32.16.1 shows an interactive iCn3D model of ribulose 1,5-bisphosphate carboxylase/oxygenase from Synechococcus PCC6301 (1RBL)

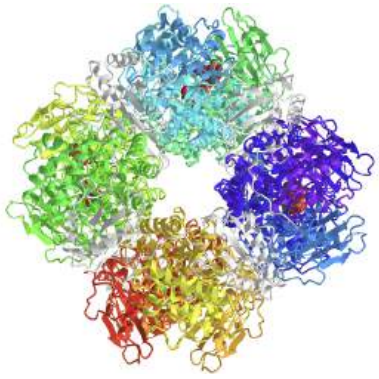


Figure 32.16.1: Ribulose 1,5-bisphosphate carboxylase/oxygenase from Synechococcus PCC6301 (1RBL) (Copyright; author via source). Click the image for a popup or use this external link (long load time): https://structure.ncbi.nlm.nih.gov/i...pxLGgEKR5gCYK8

This structure is a hetero 16-mer of 8 small chains and 8 large chains. The small chains are in gray and the large ones in differing colors. Each large subunit contains a bound reaction intermediate analog 2-carboxyarabinitol 1,5-bisphosphate.

This chapter section will focus on improving and designing new ways to capture carbon dioxide as one way to reduce its concentration in the atmosphere. Don't lose sight of the fact that the best way to deal with anthropogenic climate change is to stop putting CO_2 in the atmosphere from burning fossil fuels.

NATURALLY OCCURRING PATHWAYS TO FIX CARBON

(Much of this immediate section derives from the following reference: Natural carbon fixation. Sulamita Santos et al., Natural carbon fixation and advances in synthetic engineering for redesigning and creating new fixation pathways, Journal of Advanced Research, Volume 47, 2023, Pages 75-92, ISSN 2090-1232, https://doi.org/10.1016/j.jare. Creative Commons license

There are six naturally occurring pathways that fix carbon. These are illustrated in Figure 32.16.2 below.

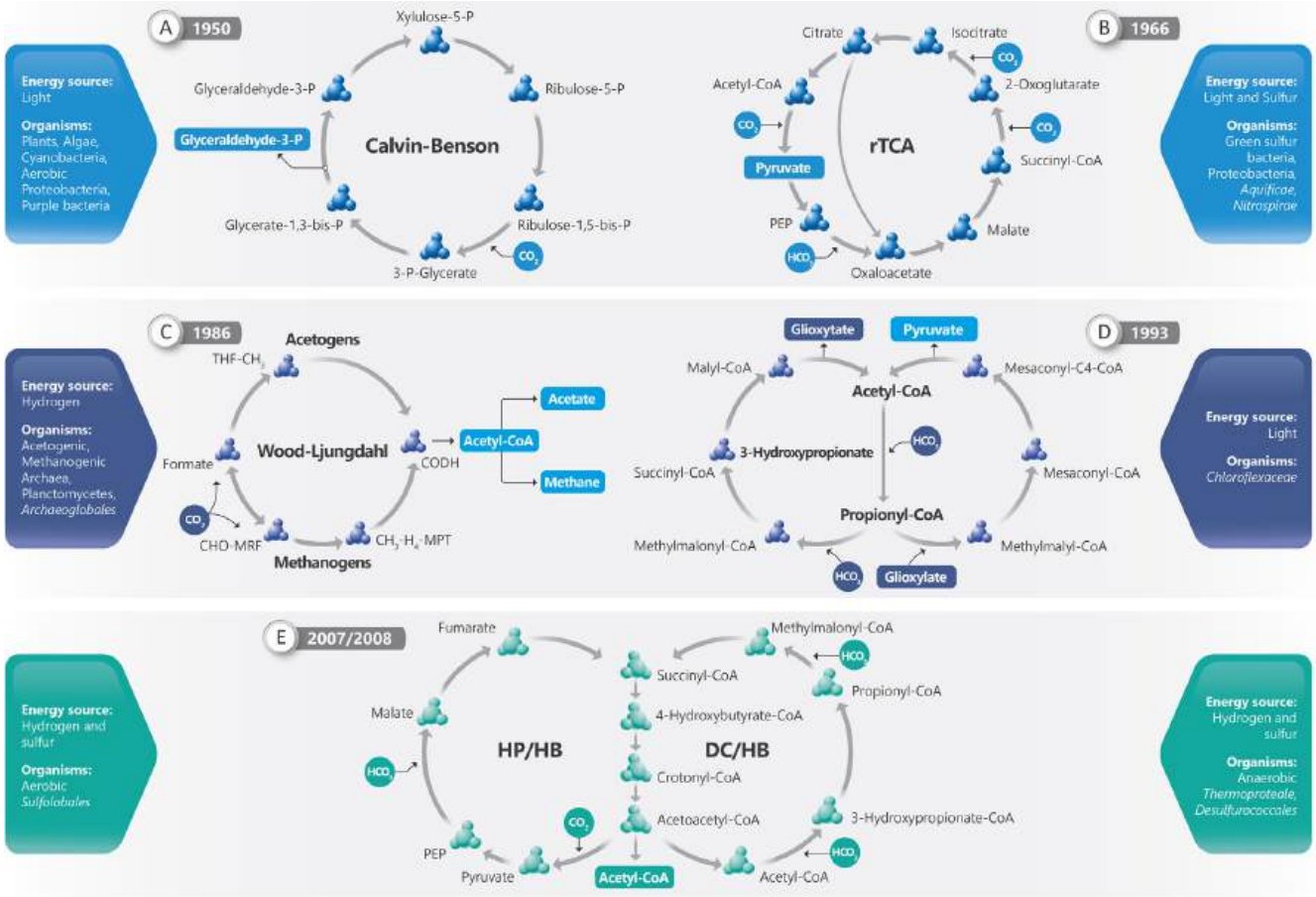


Figure 32.16.2: Natural carbon fixation. Sulamita Santos et al.,ibid.

Panel (A) shows the CBB cycle which we discussed in great detail in Chapter 20.4. Here are the enzymes: ribulose-1,5-bisphosphate carboxylase/oxygenase, 3-phosphoglycerate kinase, glyceraldehyde-3-phosphate dehydrogenase, ribulose-phosphate epimerase.

Panel (B) shows the reverse (reductive)-TCA cycle which we discussed in Chapter 16.4 in the section on the α-ketoacid pathway - A primordial, prebiotic anabolic "TCA-like" pathway. The enzymes include ATP-citrate lyase, malate dehydrogenase, succinyl-CoA synthetase, ferredoxin (Fd)-dependent-2-oxoglutarate synthase, isocitrate dehydrogenase, PEP carboxylase.

Panel (C) shows the Wood–Ljungdahl (or reductive Acetyl-CoA) cycle which we discussed in Chapter 30.1. At the top of the pathway are the acetogens Archaea and at the bottoms are the methanogens Archaea. The enzymes include MPT-methylene tetrahydromethopterin, MFR-methanofuran, THF, tetrahydrofolate.

Panel (D) shows the 3-hydroxypropionate (3HP) cycle. The enzymes include acetyl-CoA carboxylase, propionyl-CoA carboxylase, methylmalonyl-CoA epimerase, succinyl-CoA:(*S*)-malate-CoA transferase, trifunctional (*S*)-malyl-CoA, -methylmalyl-CoA, mesaconyl-CoA transferase, mesaconyl-C4-CoA hydratase.

Panel (E) shows the hydroxypropionate/4-hydroxybutyrate (HP/HB) and the dicarboxylate/4-hydroxybutyrate (DC/HB) cycles. The enzymes include pyruvate synthase, PEP-carboxylase, malate dehydrogenase, fumarate hydratase/reductase, acetyl-CoA/propionyl-CoA carboxylase, 3-hydroxypropionate-CoA ligase/dehydratase, methylmalonyl-CoA mutase, succinyl-CoA reductase, 4-hydroxybutyrate-CoA ligase, crotonyl-CoA hydratase, acetoacetyl-CoA-ketothiolase.

Table 32.16.1 below shows a comparative description of the natural and synthetic carbon fixation pathways.

Pathway	Organisms	Energy Source	Input	Output	Reductants	Key Enzyme
Calvin-Benson (N)	Plants, Algae, Cyanobacteria, Aerobic Proteobacteria, Purple bacteria	Light	3 CO_2, 9 ATP,6 NAD(P)H	Glyceraldehyde-3-phosphate	NAD(P)H	RuBisCO
rTCA (N) *	Green sulfur bacteria, Proteobacteria, Aquificae, Nitrospirae	Light and Sulfur	2 CO_2, 2 ATP,4 NAD(P)H	Pyruvate	NAD(P)H & ferredoxin	2-Oxoglutarate synthase, Isocitrate dehydrogenase
Wood–Ljungdahl (N) *	Acetogenic, Methanogenic Archaea, Planctomycetes, Sulfate. Archaeoglobales,	Hydrogen	2 CO_2, 1 ATP, 4 NAD(P)H	Acetyl-CoA	Ferredoxin	NAD-independent formate dehydrogenase, Acetyl-CoA synthase-CO dehydrogenase
3-HP (N)	Chloroflexaceae	Light	3 HCO , 5 ATP, 5 NAD(P)H	Pyruvate	NAD(P)H	Acetyl-CoA carboxylase, Propionyl-CoA carboxylase
HP/HB (N)	Aerobic Sulfolobales	Hydrogen/sulfur	2 HCO , 4 ATP, 4NAD(P)H	Acetyl-CoA	NAD(P)H	Acetyl-CoA-Propionyl-CoA carboxylase
DC/HB (N) *	Anaerobic Thermoproteale,Desulfurococcales	Hydrogen/sulfur	1 CO_2, 1 HCO , 3ATP, 4 NAD(P)H	Acetyl-CoA	NAD(PH &Ferredoxi	Pyruvate synthase, PEP carboxylase
RHP (CN)	Methanospirillum hungatei	Hydrogen	CO_2, 3 ATP, 2 NAD(P)H	Gluconeogenesis and glycolysis	NAD(P)H	RuBisCO
Natural Reductive Glycine (CD)	Candidatus phosphitivorax, anaerolimiDesulfovibrio desulfuricans	Phosphite	CO_2, ATP, NAD(P), H	Formate/ Pyruvate	NAD(P)H &Ferredoxi	CO_2-reducing formate dehydrogenase(fdhAB)
Reverse Otca (CD)	Desulfurella acetivorans	Hydrogen	CO_2, ATP, NAD(P) H	Acetyl-CoA	Ferredoxin	Citrate synthase
CETCH (S)	Theoretical	–	2 CO_2, 2 ATP, 3 NAD(P)H	Glyoxylate	NAD(P)H	CoA- dependent carboxylase
Reductive Glycine (S)	Demonstrated in E. coli as host	–	CO_2, NADH	Pyruvate	Ferredoxin	Glycine cleavage system
Synthetic malyl-CoA-glycerate (S)	Demonstrated in E. coli and Synechococcus elongatus PCC7942 host	–	CO_2, 3 ATP, 3 NADH	Acetyl-CoA	NAD(P)H	PEP-carboxylase, RuBisCO
SACA Pathway (S)	Demonstrated in E. coli as host	–	CO_2	Acetyl-CoA	–	NAD-independent formate dehydrogenase
Formolase pathway (S)	Theoretical	–	CO_2, NADH, ATP	Dihydroxyacetone-phosphate	NADH	NAD-independent formate dehydrogenase

Sulamita Santos et al., Natural carbon fixation and advances in synthetic engineering for redesigning and creating new fixation pathways, Journal of Advanced Research,
Volume 47, 2023, Pages 75-92, ISSN 2090-1232, https://doi.org/10.1016/j.jare.

Plants and microorganisms that are photoautotrophic fix CO_2 by the Calvin cycle using NADPH and ATP and produce O_2. Some anerobic photosynthetic bacteria don't produce O_2. In **chemolithoautotrophic** microorganisms, energy sources (i.e. electron donors like H_2, H_2S, sulfur, Fe^{2+}, nitrite and NH_3 found in bedrock - the lithosphere) other than NADPH, ATP and light can be used to drive CO_2 uptake. For **chemoorganotropic** microbes, reduced organic molecules such as sugars and amino acids serve as electron (donor) sources.

Genetic engineering and synthetic biology are now employed to improve preexisting enzymes and to create whole new pathways for carbon capture.

As a simple example, people are trying to engineer carbonic anhydrase, used to convert CO_2 to HCO_3^- for transport to the lung where it is converted back to CO_2 and released. It is a diffusion-controlled enzyme with a k_{cat}/K_M reported as high as 8×10^7 $M^{-1}s^{-1}$, so how can it be made better? One way is to engineer thermal stability into the enzyme. A critical problem for the forward reaction is that the enzyme is readily reversible so bicarbonate, HCO_3^-, is a competitive inhibitor of the forward reaction. In addition, the enzyme can be engineered to be more stable at higher pH to allow product (HCO_3^-) removal by the addition of OH^- in a process of mineralization, as shown in the equation below,

$$HCO_3^- + OH^- \rightarrow CO_3^{2-} + H_2O$$

where the carbonate anion can precipitate in the presence of divalent cations like Ca^{2+}, Mg^{2+} and Fe^{2+}. Here is a link to a Literature-based Guided Assessment on thermoengineering of carbonic anhydrase.

Now let's explore the use of new pathways created by synthetic biology to capture carbon. Some of these pathways are engineered to produce reactants (feedstocks) for biofuels, which we discussed in-depth in previous Chapter 32 sections, and chemical synthesis. We'll

focus on three: the CETCH pathway, the reductive glyoxylate and pyruvate synthesis (rGPS) cycle, and the malyl-CoA-glycerate (MCG) pathway.

CETCH (Crotonyl-CoA-EThylmalonyl-CoA-4Hydorxybutyl-CoA) pathway

To engineer a new pathway, Swander et al identified efficient carboxylases from known ones (acetyl-CoA carboxylase, Rubisco, propionyl-CoA carboxylase, PEP carboxykinase, 2-oxoglutarate carboxylase, and pyruvate carboxylase), all of which we have discussed in previous chapters. They created new pathways, calculated free energy and ATP/NADPH requirements, and then optimized the pathways. They chose CoASH–dependent carboxylases and enoyl-CoA carboxylases/reductases. Figure 32.16.3 below shows the reaction of one key carboxylase.

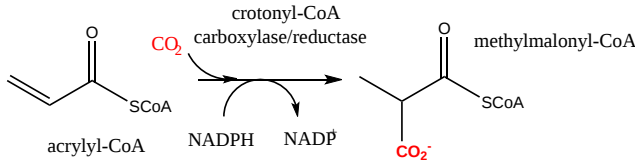


Figure 32.16.3: Carboxylase used in the CETCH pathways to fix CO_2

The pathway was named CETCH (**C**rotonyl-CoA-**ET**hylmalonyl-**C**oA-4**H**ydorxybutyl-CoA) which catalyzes this next reaction in cell lysates (in vitro):

$2CO_2 + 3NAD(P)H + 2ATP + FAD \rightarrow$ glycolate $+ 3NAD(P) + 2ADP + 2P_i + FADH_2$

The rate of CO_2 fixation by the CETCH pathways was similar to the rate of the Calvin cycle rates in cell lysates.

In a more expansive approach, Gleizer used synthetic biology to change E. Coli from a heterotroph to an autotroph in which its biomass (carbon reservoir) came from CO_2. Formate (HCO_2^-) was used as a source of reducing power (electrons) as it was oxidized by an added formate dehydrogenase to produce NADH for the autotropic fixation of CO_2 through the addition of Calvin cycle enzymes. Using isotopically labeled $^{13}CO_2$ to follow carbon flow, after 10 generations and evolution, the cells were completely autotrophic through fixation of CO_2. To accomplish this, they knocked out genes for phosphofructokinase (glycolysis) and glucose-6-phosphate-dehydrogenase (pentose-phosphate pathway) to impair these main metabolic pathways, and added carbonic anhydrase (to interconvert CO_2 and HCO_3^-) as well as Rubisco and phosphoribulokinase. As formate was ultimately converted to CO_2, the net effect was not exactlycarbon neutral but could be if atmospheric CO_2 was used to make formate (for a feedstock) by electrochemical reduction.

Figure 32.16.4 below shows the next reactions in the synthetic E. Coli autotrophs.

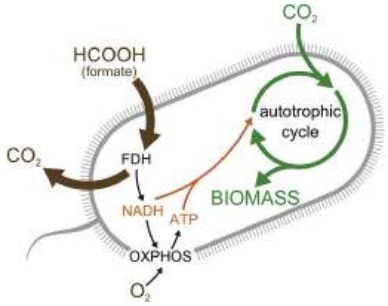


Figure 32.16.4: Schematic Representation of the Engineered Synthetic Chemo-autotrophic *E. coli.* Shmuel Gleizer et al., Conversion of Escherichia coli to Generate All Biomass Carbon from CO2, Cell, 179 (2019). https://doi.org/10.1016/j.cell.2019.11.009. Creative Commons license.

CO_2 (green) is the only carbon source for all the generated biomass. The fixation of CO_2 occurs via an autotrophic carbon assimilation cycle. Formate is oxidized by a recombinant formate dehydrogenase (FDH) to produce CO_2 (brown) and NADH. NADH provides the reducing power to drive carbon fixation and serves as the substrate for ATP generation via oxidative phosphorylation (OXPHOS in black). The formate oxidation arrow is thicker than the CO_2 fixation arrow, indicating a net CO_2 emission even under autotrophic conditions.

Figure 32.16.5 shows that almost 100% of carbon atoms after many generations are labeled with ^{13}C (detected by mass spec analysis) derived from $^{13}CO_2$.

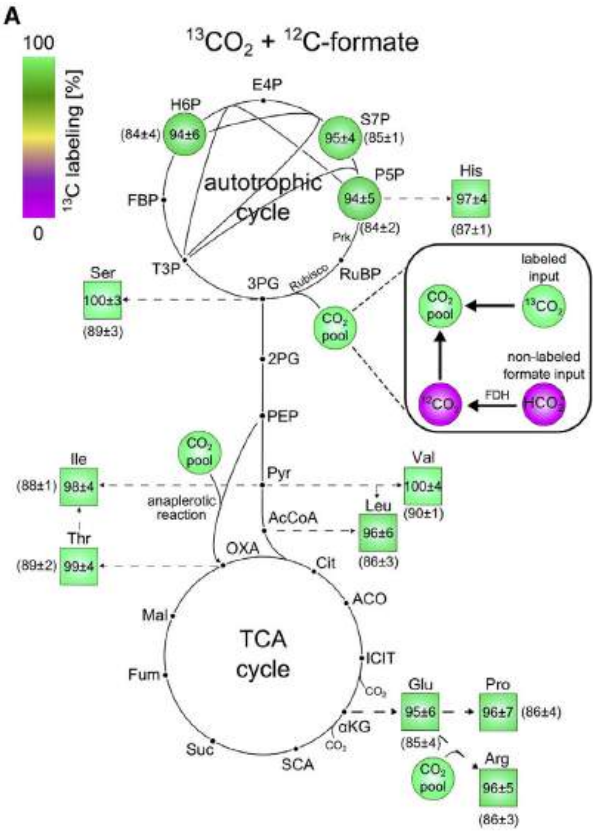


Figure 32.16.5: Isotopic Labeling Experiments Using ^{13}C Show that All Biomass Components Are Generated from CO_2 as the Sole Carbon Source. Gleizer et al., ibid.

(A) Values are based on LC-MS analysis of stable amino acids and sugar-phosphates. The fractional contribution of $^{13}CO_2$ to various protein-bound amino acids and sugar-phosphates of evolved cells grown on $^{13}CO_2$ and naturally labeled formate showed almost full ^{13}C labeling of the biosynthesized amino acids. The numbers reported are the ^{13}C fraction of each metabolite, taking into account the effective $^{13}CO_2$ fraction out of the total inorganic carbon (which decreases due to unlabeled formate oxidation to CO_2). The numbers in parentheses are the uncorrected measured values of the ^{13}C fraction of the metabolites.

Synthetic reductive glyoxylate and pyruvate synthesis (rGPS) cycle and the malyl-CoA-glycerate (MCG) pathways

These pathways were created to synthesize acetyl-CoA, pyruvate, and malate from CO_2 in cell-free systems to free the system from cell growth and regulation requirements, and to make it insensitive to oxygen. These molecules are also intermediates in the created cycle, which could operate continuously for hours at the same or greater rates of CO_2 fixation compared to photosynthesis. The cycle is shown in Figure 32.16.6 below.

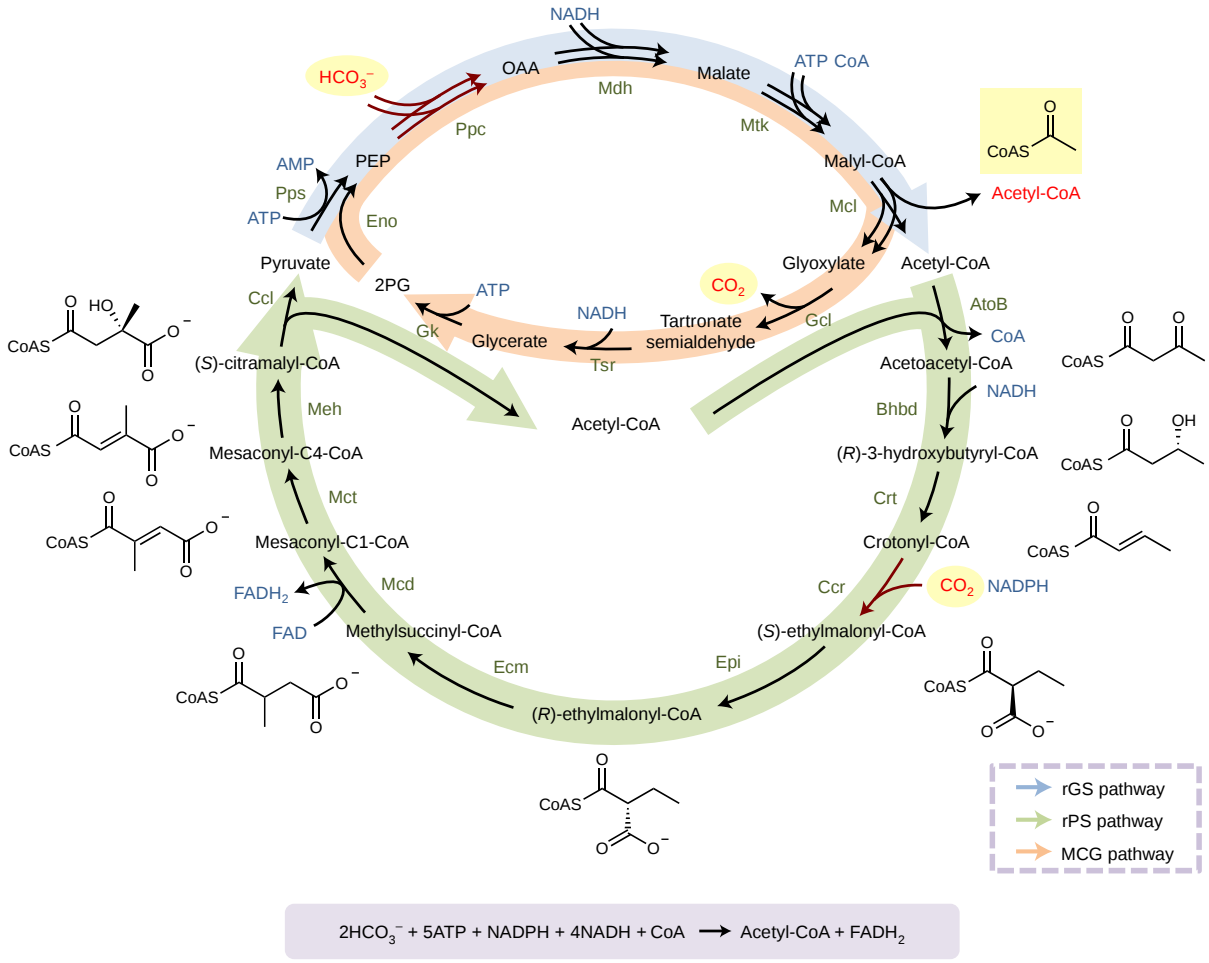


Figure 32.16.6: The rGPS–MCG cycle with acetyl-CoA as the end product. Luo, S., Lin, P.P., Nieh, LY. *et al*. A cell-free self-replenishing CO_2-fixing system. *Nat Catal* **5**, 154–162 (2022). https://doi.org/10.1038/s41929-022-00746-x . Creative Commons Attribution 4.0 International License, http://creativecommons.org/licenses/by/4.0/

The rGPS cycle consists of the reductive glyoxylate sythesis (rGS) pathway (blue) and the reductive pyruvate synthesis (rPS) pathway (green). The malyl-CoA-glycerate (MCG) pathway (orange) consists of the rGS pathway and the glycerate pathway. The red arrow indicates the carboxylation reaction. Gcl, glyoxylate carboligase; Tsr, tartronate semialdehyde reductase; Gk, glycerate 2-kinase; Eno, enolase and 2PG, 2-phospho-d-glycerate

Microbial electrosynthesis from CO_2

We mentioned above that if the formate used in the CETCH pathway could be synthesized through electrochemistry from CO_2, then the pathway would be truly carbon neutral. In fact, new microbial electrochemical methods are being designed to synthesize a variety of small molecules that could serve as feedstocks for chemical synthesis in industry. The carbon in CO_2 has an oxidation number of +4 while the carbon in formic acid has an oxidation number of +2. Hence two electrons must be added by electrochemically to make the CETCH pathway truly carbon neutral. Bigger electrochemical reductants of CO_2 require more electrons. Figure 32.16.7 below shows how key feedstocks could be made electrochemically from CO_2 and how they are usually made in industry.

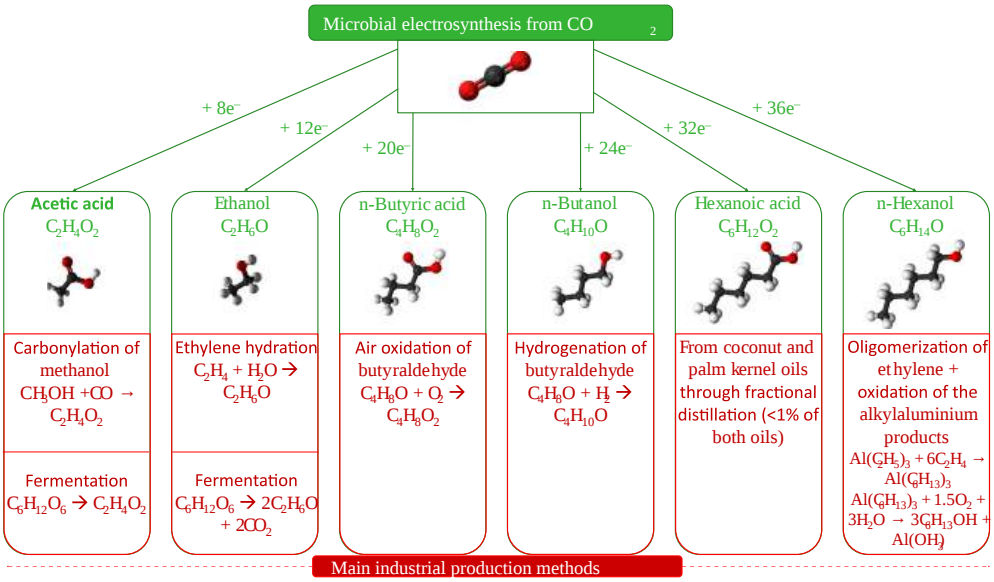


Figure 32.16.7 below. : **Overview of the Main Products Formed from Microbial Electrosynthesis (MES) From CO_2, Along With the Main Industrial Methods to Manufacture These Products.** Jourdin et al., Trends in Biotechnology, April 2021, Vol. 39, No. 4 https://doi.org/10.1016/j.tibtech.2020.10.014. Creative Commons Attribution (CC BY 4.0)

These feedstocks could be made by reductive electrosynthesis using electrons from the oxidation of water though electrolysis. Released electrons (oxidation number of O in water is -2 and 0 in O_2) move to a biocathode to reduce CO_2, as illustrated in Figure 32.16.8 below.

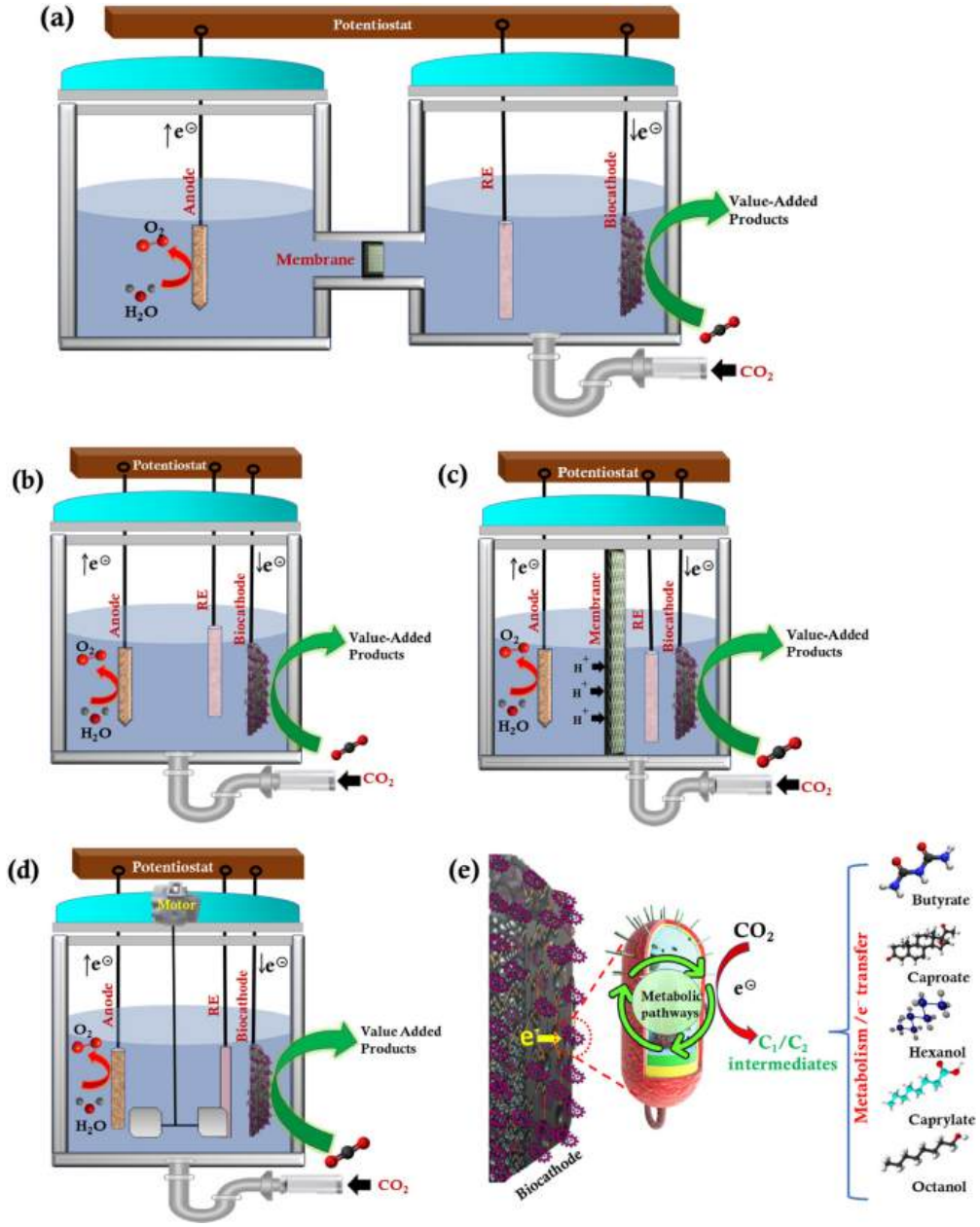


Figure 32.16.8: Reactor configurations for MES-based CO_2 conversion: (a) H-type, (b) single chamber, (c) dual chamber, (d) continuous stirred tank, and (e) schematic of electron transfer mechanism. **G. S. Lekshmi et al., Microbial electrosynthesis: carbonaceous electrode materials for CO_2 conversion.** *Mater. Horiz.*, 2023, **10**, 292-312. DOI: 10.1039/D2MH01178F. Creative Commons Attribution-Non Commercial 3.0 Unported Licence

The biocathode consists of a biofilm of cells printed onto a graphene, graphite, or carbon nanotube-laden support (all carbonaceous).

32.17: FIXING NITROGEN FIXATION

Search Fundamentals of Biochemistry

NITROGEN FIXATION

We spent most of Chapter 22.1 discussing the biochemistry of nitrogenase which fixes the stable molecule N_2 to form NH_3/NH_4^+. It's a very complicated reaction conducted by symbiotic microbes (prokaryotes) that fix N_2 for plants. The world uses the Haber-Bosch process to produce over 100 million metric tons of nitrogen fertilizer that supports half of the world's population's food supply. Only about half of the ammonium added to soil is taken up by plants. The rest is released into waterways or used by microbes, which can produce from it the potent greenhouse gas nitrous oxide (N_2O). It has a 300x greater effect than CO_2 based on weight. The oxidation number of N in NH_3 is -3 and in N_2O it is +1 showing that the NH_3 can be oxidatively metabolized for energy production by the microbes. Also, excess NH_4^+ goes into waterways and leads to eutrophication, the overproduction of algae and plankton, which depletes O_2 from the waters and kills other organisms.

N_2O emissions have increased dramatically since 1850, as shown in the interactive graph from Our World in Data in Figure 32.17.9 below.

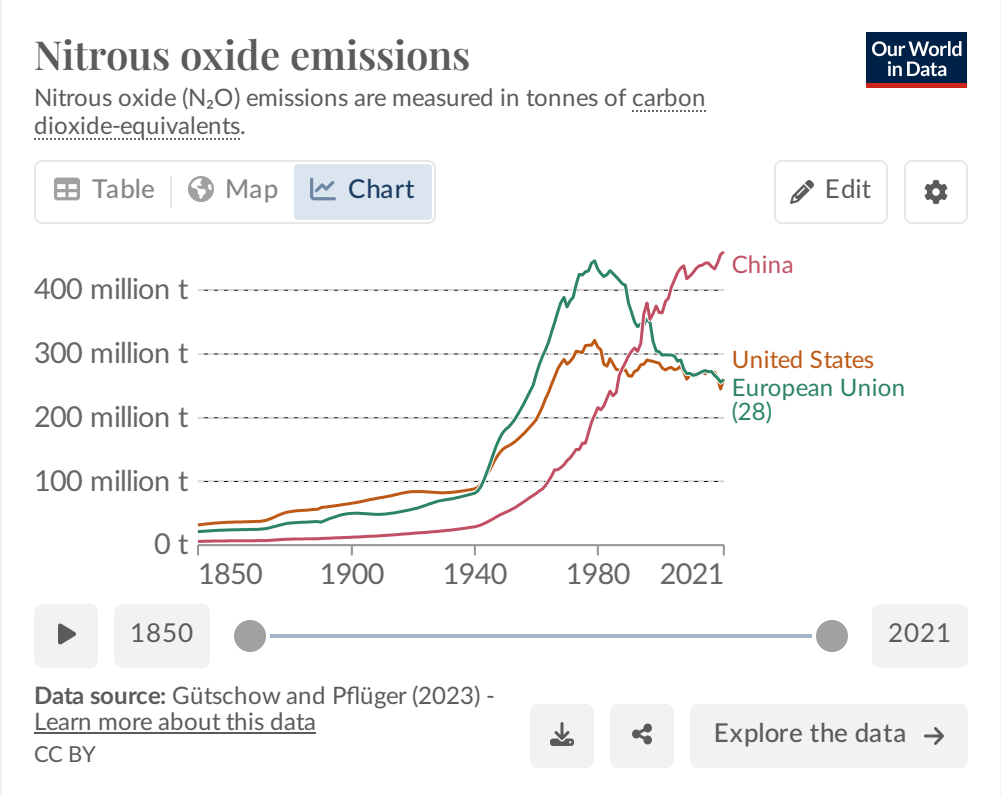


Figure 32.17.9: Our World in Data. https://ourworldindata.org/grapher/n...ions?tab=chart

So what can we do to "fix" nitrogen fixation to reduce our reliance on the Haber process and its collateral climate effects? Perhaps we could use mutagenesis to make nitrogenase a better and more efficient enzyme. That would prove very difficult given the complexity of both the enzyme and the mechanism of N_2 conversion to NH_3 that requires many metal ion cofactors. A better alternative would be to express nitrogenase in plants so they could synthesize their own nitrogen fertilizer!

Genes of Nitrogen Fixation

The catalytic nitrogenase enzyme complex is encoded by three genes:

- nifH gene for the subunit **Nitrogenase iron protein 1** (also called Nitrogenase component II, and Nitrogenase reductase). It is a homodimer that binds one [4Fe-4S] cluster per dimer.
- nifD gene for the subunit **Nitrogenase molybdenum-iron protein alpha chain** (also called Dinitrogenase and Nitrogenase component I). It catalyzes the key enzymatic reactions along with its partner nifK as part of a heterodimer. It binds one [8Fe-7S] cluster per heterodimer with nifK and binds 1 [7Fe-Mo-9S-C-homocitryl] cluster per subunit.
- nifK gene for the subunit **Nitrogenase molybdenum-iron protein beta chain** (also called Dinitrogenase and Nitrogenase component I). With its partner nifD, it catalyzes the key enzymatic reactions as part of a heterodimer with nifD. It binds one [8Fe-7S] cluster per

heterodimer with nifD and binds 1 [7Fe-Mo-9S-C-homocitryl] cluster per subunit.

The nitrogenase enzyme complex has regulatory proteins as well:

- nifA - **Nif-specific regulatory protein** required for activation of most nif operons. It senses N_2. If there is insufficient quantities of N_2, the protein NtrC activates NifA expression which activates the rest of the genes.
- nifB - **FeMo cofactor biosynthesis protein NifB** (also called FeMo-cofactor maturase NifB, Nitrogenase cofactor maturase NifB, and radical SAM assemblase NifB).
- nifL - **Nitrogen fixation regulatory protein** required for the inhibition of NifA activity (i.e. nitrogenase formation) in response to oxygen and low level of fixed nitrogen.
- nifE - **Periplasmic [NiFe] hydrogenase small subunit**.
- NifM - a possible peptidyl prolyl *cis-trans* isomerase (i.e. a protein chaperone) which helps in the folding of NifH.

Figure 32.17.10 below shows a summary of these gene products.

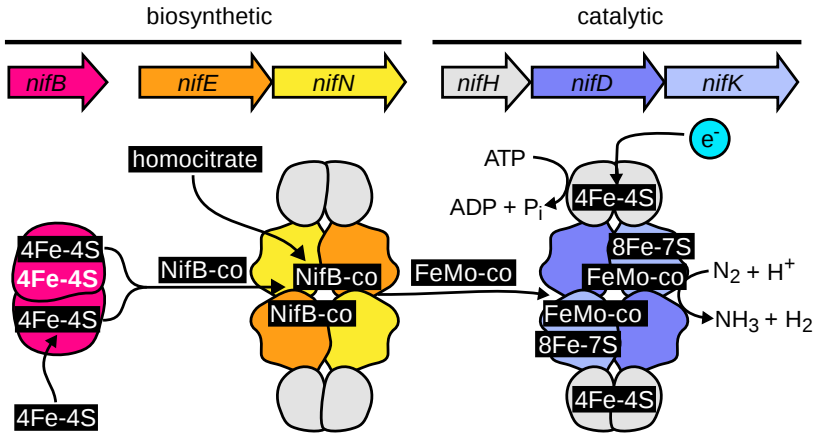


Figure 32.17.10 : Minimum set of *nif* genes essential for nitrogen fixation with molybdenum-iron nitrogenase. EMILY M. BENNETT et al., ***BIODESIGN RESEARCH.*** 10 Jan 2023, Vol 5, https://spj.science.org/doi/10.34133/bdr.0005 . Creative Commons Attribution License 4.0 (CC BY 4.0).https://doi.org/10.34133/bdr.0005

The stoichiometry depicted has not been adjusted. NifB contains one catalytic cluster (shown in white) and 2 substrate [4Fe-4S] clusters that react to produce the NifB cofactor. NifEN matures the NifB cofactor producing the FeMo cofactor. The molybdenum-iron (MoFe) nitrogenase (NifHDK) contains the FeMo cofactor at its active site. Electron donors transfer single electrons to the [4Fe-4S] cluster at the interface of the NifH homodimer. Electrons are moved from the [4Fe-4S] cluster into the active site of nitrogenase using energy produced by ATP hydrolysis by NifH. A minimum of 8 electrons are used to reduce each molecule of N_2.

It would be especially important to express nitrogenase in the main cereal food crops (rice, corn, and wheat) which get their nitrogen from soil microbes (in contrast to legumes, which contain nitrogen-fixing bacteria in nodules in their roots). It's a daunting task given the complexity of the protein complex, its metal cofactors, their inhibition by O_2, and the multiple genes required for nitrogenase regulation. Ideally, the relevant gene clusters could be moved into a chlorplast which is evolutionarily derived from bacteria so the gene regulation system might be more suitable. It also has low O_2 levels at night. However, O_2 is produced in chloroplasts, which is problematic given the sensitivity of nitrogenase to O2.

Using synthetic biology, *Saccharomyces cerevisiae* has been engineered to express the NifDK nitrogenase tetrameric protein in their mitochondria (after post-translational import). Yeast is a model organism and tools have been developed for synthetic biology experiments using yeast, so much can be learned that could apply to other eukaryotic organisms like plants. Mitochondria have high O_2 consumption (as opposed to production as in the chloroplast) and the ability to synthesize bacterial-type iron–sulfur clusters. The Nif gene clusters were engineered into the XV chromosome as shown in Figure 32.17.11 below.

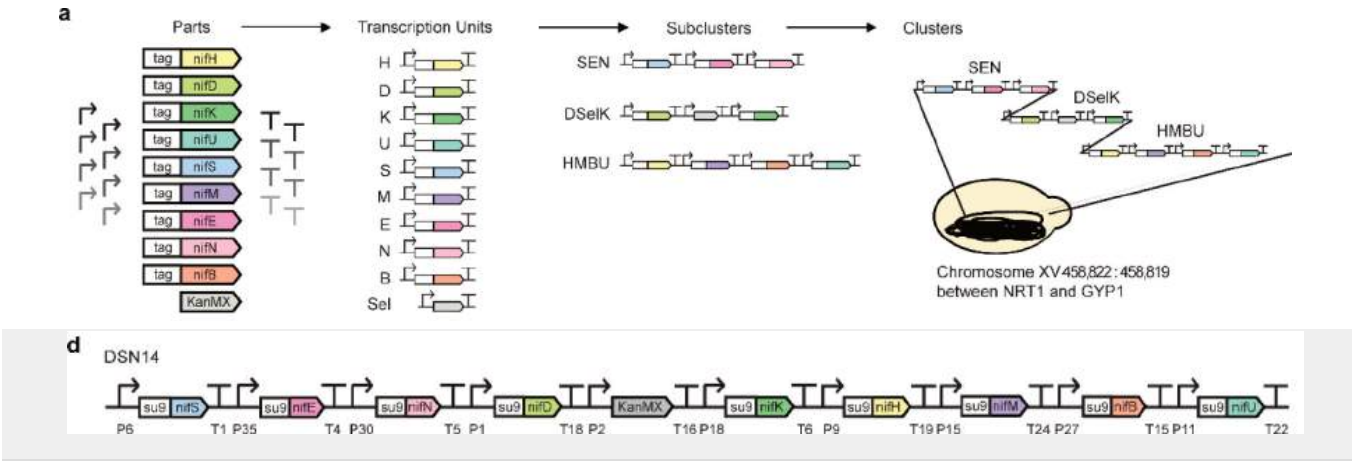


Figure 32.17.11 : nif gene assembly in yeast. *Buren et al., ACS Synth. Biol.* 2017, 6, 6, 1043–1055. https://doi.org/10.1021/acssynbio.6b00371. **CC-BY license**

Panel (a) shows the assembly strategy for transcription units, subclusters, and full clusters inserted by homologous recombination in the genome of *S. cerevisiae*. Panel (d) shows a disgram of *nif* gene organization in DSN14, a strain of *S. cerevisiae.*

Nitrogenase activity has also been functionally expressed in transgenic rice containing the NifH with a [4Fe-4S] cluster from *Hydrogenobacter thermophilus* and NifM (a peptidyl prolyl *cis-trans* isomerase from *Azotobacter vinelandii)* which helps NifH fold. They were correctly targeted to the mitochondria which again minimizes O_2 oxidative damage to the metal ion cofactors. The purified protein was able to transfer electrons to the MoFe protein (NfiDK dimer) but did so poorly. It also assisted in the assembly of the FeMo cofactor. However, the [4Fe-4S] cluster occupancy in the protein was poor. However purified protein was also able to reduce acetylene, HC≡CH (an alternative substrate similar to N≡N) after the addition of purified NifDK.

Many steps have to be optimized to create a functional nitrogenase in plants like rice, corn, and wheat. For example, mitochondrial-expressed NifD is readily cleaved by a mitochondria endoprotease. Some NifD subunits are more resistant to proteolysis and a single amino acid change (Y100Q) leads to enhanced stability on the protein. AI will likely be extremely useful in maximizing the expression of nitrogen in crop plants.

SYNTHETIC BIOLOGY TO EXPRESS NITROGENASE IN BACTERIA

Bacteria can express nitrogenase that can fix atmospheric N_2 but they won't work with the critical cereal crops unless the bacteria can interact with roots in the "rhizosphere", the layers of dirt intimately in contact with roots. This life in this area is often called the **holobiont**, which consists of the plant host and all species interacting with it in a symbiotic relationship. The metabolism in the holobiont is complex. For example, plant carbohydrates are used by other organisms in the holobiont. It's similar in a way to the gut biome, which consists of an ecosystem of human and microbial cells.

Bacteria have now been engineered that express nitrogenase AND interact with corn roots to fix N_2. The cells are derivatives of γ-proteobacterium (KV137), found on corns roots and which can fix N_2. They have been engineered to turn nitrogenase genes on when N_2 fixation is needed. The engineered bacteria is added to liquid fertilizer, reducing the need for chemical fertilizer by 25 lb/acre) and at the same time increasing yields. This bacterial-based fertilizer does not wash into waterways with its concomitant negative environmental effects. Likewise, no N_2O is produced on microbial metabolism of excess fertilizers, which decreases the release of this potent greenhouse gas. In 2021 it was used on 3 million acres of corn.

One problem with the use of biological N_2 microbes in agricultural settings is that high levels of chemical fertilizers can effectively inhibit microbial N_2 fixation, which is regulated (as mentioned above) to shut down if nitrogen is bioavailable. Still, bacteria that fix N_2 can produce up to 10% of the nitrogen requirement. Genetic engineering is needed to overcome this inhibition. Such bacteria are **diazotrophs** as they can fix N_2 and grow without exogenous sources of N_2. Rhizobia are one example. which can fix N_2 in the nodules of legumes. The diazotroph isolated from corn roots and mentioned above, Kv137, was gene-edited to produce a modified strain (Kv137-1036) that fixes N_2 without inhibition by applied nitrogen fertilizers.

The Kv137 strain has nifA and nifL genes and proteins which, as described above, regulate the expression of nitrogenase based on the need for nitrogen. These two genes are on one operon under the control of a single promoter. nifL was replaced with another promoter which removes the down-regulation of nifA since no nifL was present. This allowed the expression and activity of nitrogen even in the presence of exogenous fertilizer.

Figure 32.17.12 below shows that Kv137-1036 strain (red dots) does colonize corn roots.

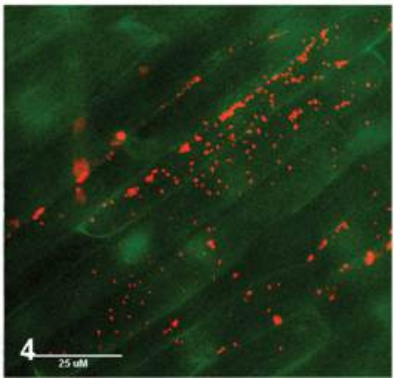


Figure 32.17.12 : Commercial efficacy of strain Kv137-1036 - Colonization of corn roots by microbes (red) after germination. Wen et al., Enabling Biological Nitrogen Fixation for Cereal Crops in Fertilized Fields. *ACS Synth. Biol.* 2021, 10, 12, 3264–3277. December 2021. https://doi.org/10.1021/acssynbio.1c00049. **Attribution 4.0 International (CC BY 4.0)**

32.18: TURNING TREES INTO PLEXIGLASS - SYNTHETIC BIOLOGY FOR PRODUCTION OF GREEN FOODS AND PRODUCTS

INTRODUCTION

Manufacturing of any kind is energy-consuming and environmentally damaging and contributes to climate change through the release of CO_2 and other pollutants. A manufactured item has a lifetime after which it must be disposed of in a fashion that often involves little recycling. A circular economy in which a used product is always recycled for further use if done well, would be highly beneficial for the environment.

Synthetic biology as a field seeks to genetically alter and redesign organisms to produce traditional or novel products in more sustainable ways with less energy input and polluting output. Although it is a nascent field, well-known products are binding produced through its use. We will explore several products made through synthetic biology as well as several in which novel cells themselves are the products.

PRODUCTS FROM CELLS

Burgers by Impossible Foods

Agriculture has literally transformed the planet. About 50% of all land (other than deserts and ice shields and comprising an area equal to the Americas) is used for agriculture and most of that is for animal production. In the US, about 10% of all greenhouse gases derive from the agricultural sector. Similar estimates are made for global greenhouse gas emissions from the production of food, which include agriculture, land use change, and supply chain emissions (transport, packaging, food processing, retail, cooking, and waste).

Estimates show that wild mammals comprise only about 6% of the total mass of all mammals (including people, livestock, and pets) on the planet). In a parallel finding, the mass of "stuff" (plastics, metals, asphalt, concrete, etc) created by humans now exceeds the entire biomass of the planet!

Most of agricultural land is used to produce meat and milk for human consumption. Collectively, cattle by far require the most land use, as shown in the interactive graph of Our World In Data in Figure 32.18.1 below, which shows how many square meters are required to produce 1000 kcals (1000 cal in the dietary sense) of food from each food type listed. The graph is similar when the measure is land use per 100 grams of protein produced.

Figure 32.18.1: Land use of foods/1000 kcals. https://ourworldindata.org/land-use-diets

The number of animals slaughtered **each day** in the world is unbelievably high, as shown in Table 32.18.1 below.

cows	goats	sheep	pigs	ducks	chickens	fish
900,00	1.4 million	1.7 million	3.8 million	11.8 million	202 million	100's of millions

Table 32.18.1: Animals slaughtered each day for food. Max Roser (2023) - "How many animals get slaughtered every day?" Published online at OurWorldInData.org. Retrieved from: 'https://ourworldindata.org/how-many-...ered-every-day' [Online Resource]

If you consider greenhouse gas emissions all aspects of food products through the entire food chain, the number approached 25%. Other factors contributing to the emission of greenhouse gases and the environmental impact of food production are shown in Figure 32.18.2 below.

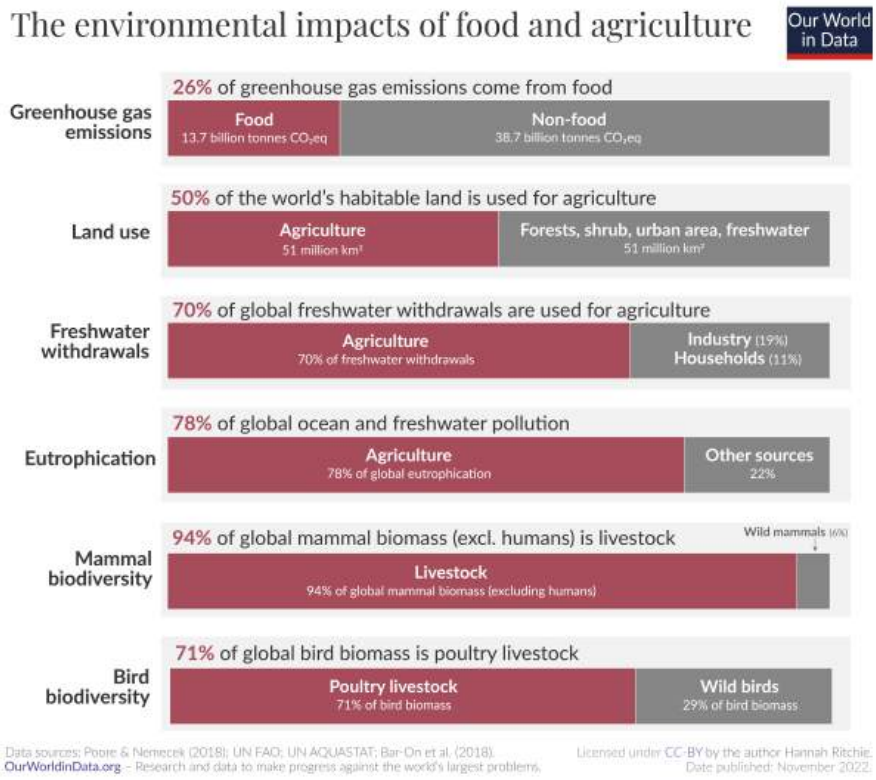


Figure 32.18.2: Environmental impact of food and agriculture

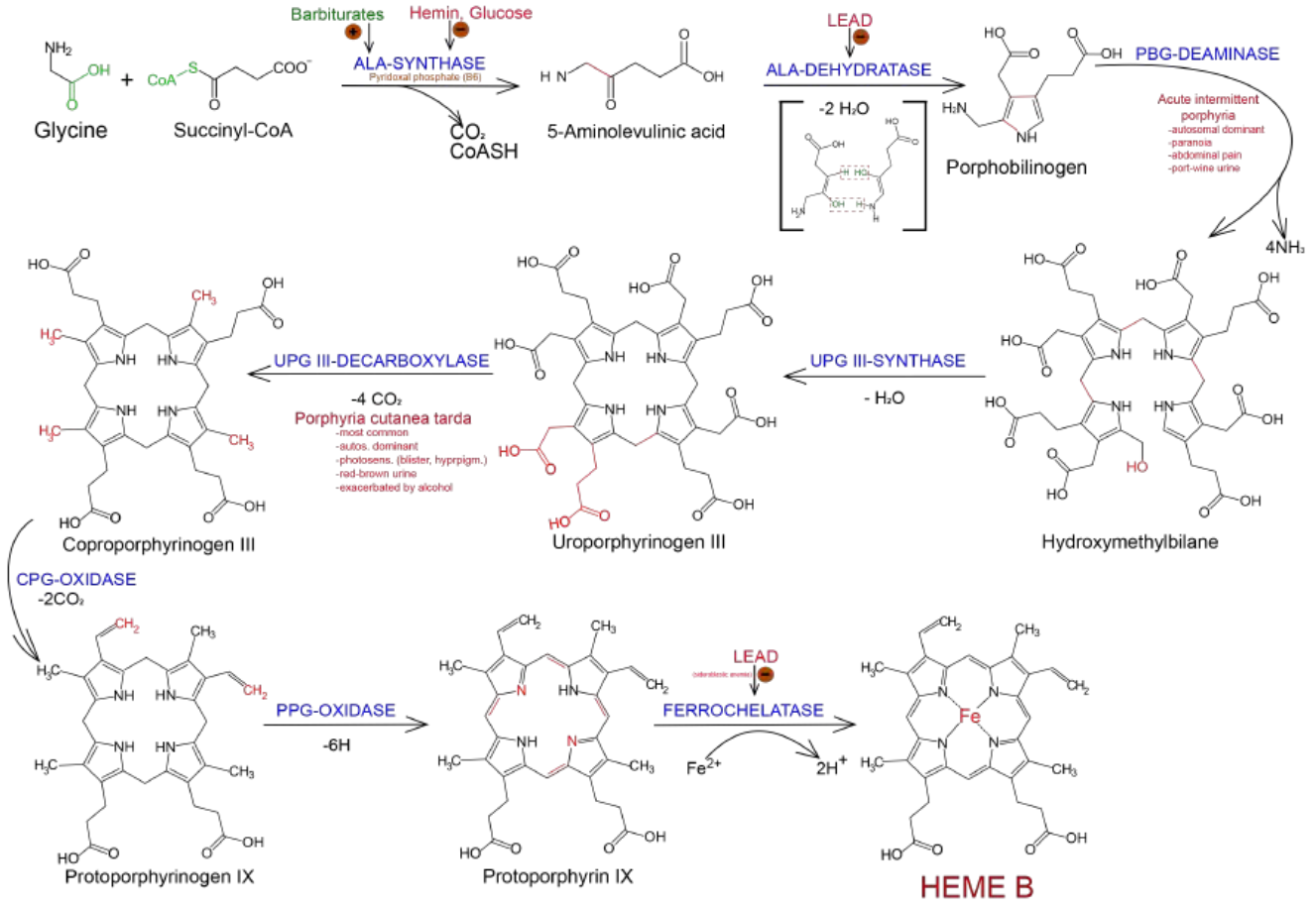
Making plant-based foods that taste more like meat, if people would eat them, could have a large effect on greenhouse gas emissions and climate change. One example is the Impossible Burgers and other similar meats from Impossible Foods. They have soy **leghemoglobin**, a monomeric heme-binding protein found in root nodules in legumes, to give the appearance and taste of blood in meat. As a single-chain heme-binding protein, it has a high affinity of O_2, similar to animal myoglobin. The high affinity derives from very high on-rates for binding O_2 (almost diffusion-controlled at around $2x10^8$ s^{-1}, and an off rate of around 20 s^{-1}. This high affinity keeps O_2 bound which would otherwise inhibit nitrogenase and nitrogen fixation by root-associated microbes. The heme is important for positive tastes when we eat red meat. Plant-based burgers containing leghemoglobin require much less land and lead to far lower greenhouse gas emissions.

Figure 32.18.3 shows an interactive iCn3D model of the alignment of sperm whale myoglobin (1MBO) and soy leghemoglobin (1BIN).

NCBI iCn3D Figure 32.18.3: Alignment of sperm whale myoglobin (1MBO, cyan) and soy leghemoglobin (1BIN, magenta). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...4pUi1S6qFtGn4A

The leghemoglobin in the Impossible burgers is produced in yeast so it can be scaled up easily. To produce leghemoglobin in yeast, large amounts of heme are required, which is also produced in the engineered cells on the introduction of the appropriate genes. The heme synthesis pathways (described in Chapter 22.3) for C4 (humans, animals, fungi, and purple non-sulfur phototrophic bacteria top) and C5

(archaea, plants, and other bacteria) for heme synthesis are shown in Figure 32.18.4 below. The succinyl-CoA is derived from the citric acid cycle.



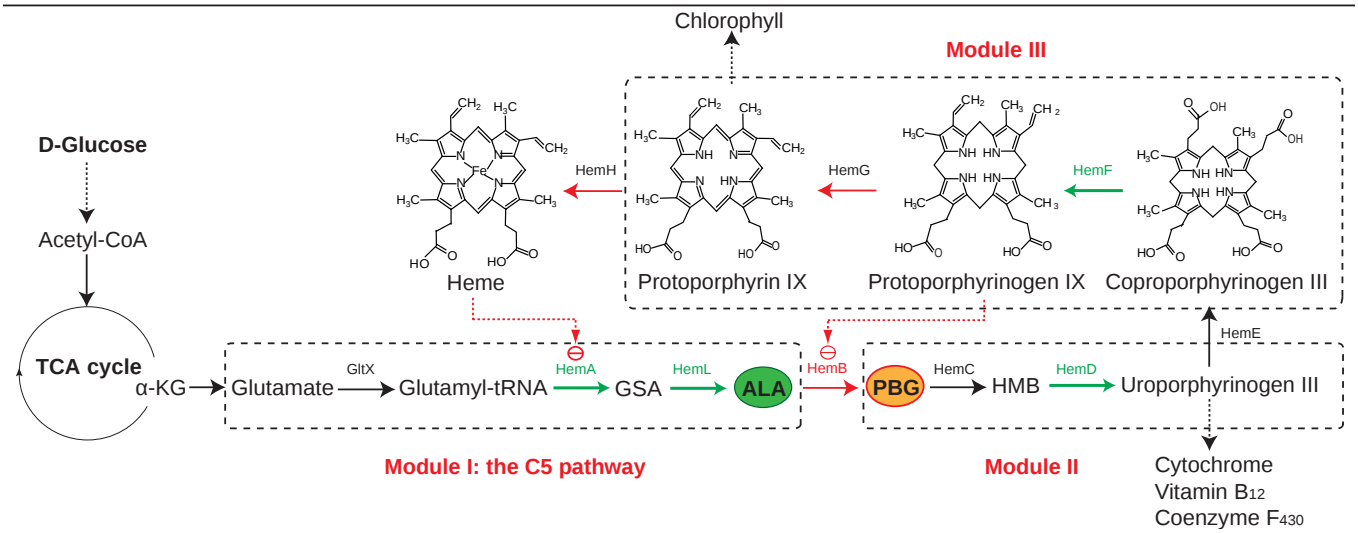


Figure 32.18.4: Top - heme synthesis pathway for C4 (humans, animals, fungi, and purple non-sulfur phototrophic bacteria top). Bottom - heme synthesis pathway for C5 (archaea, plants, and other bacteria). Heme biosynthetic pathway. Wikimedia Commonsile: Heme-Synthesis-Chemical-Details-Mirror (top) and Heme pathway in *E. coli*. Zhang, J., Kang, Z., Chen, J. *et al.* Optimization of the heme biosynthesis pathway for the production of 5-aminolevulinic acid in *Escherichia coli*. *Sci Rep* **5,** 8584 (2015). https://doi.org/10.1038/srep08584. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/ (bottom).

Note that both succinyl-CoA (C4 pathway) and α-ketoglutarate (C5 pathway) are derived from the citric acid cycle. The key precursor 5-aminolevulinate, ALA needs to be elevated, through engineering of the C4 or C5 pathways with the C5 pathways generally producing more ALA in engineered E. Coli.

Leghemoglobin from soy (species name Glycine max) can also be synthesized in the methylotrophic (uses methanol as a sole carbon source) yeast Pichia pastoris which is often used for recombinant protein expression. Three groups of enzymes are needed"

- (group 1: porphobilinogen synthase (PBGS)
- group 2: uroporphyrinogen III synthase (UROS), uroporphyrinogen III decarboxylase (UROD), coproporphyrinogen III oxidase (CPO)
- group 3: Ala synthase (ALAS), protoporphyrinogen oxidase (PPO), and ferrochelatase (FECH)

Transcription of these genes in P. pastoris can be controlled by the use of the methanol-induced alcohol oxidase (AOX1) promoter, which is often used to achieve high expression of recombinant proteins. Hence when the cells also contain two copies of leghemoglobin along with the rest of the genes, high levels of the protein were made.

A more detailed representation of the heme synthesis pathway is shown in Figure 32.18.5 below.

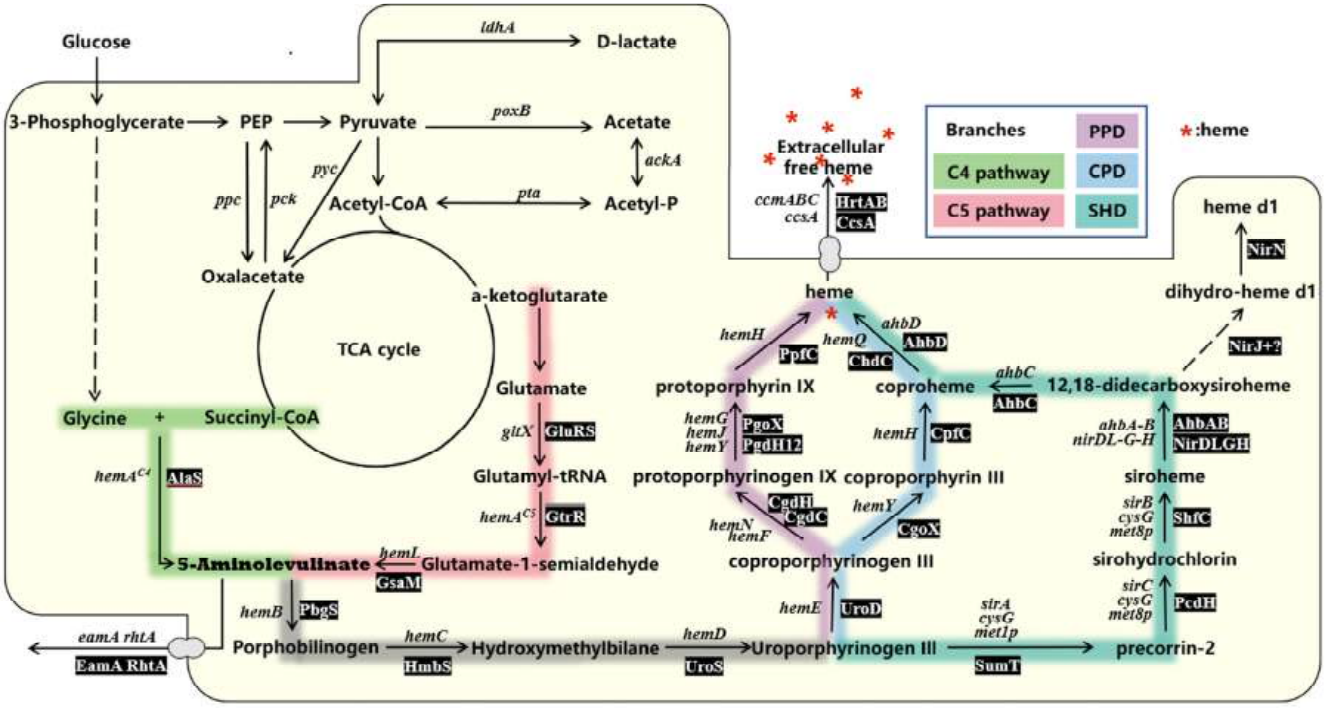


Figure 32.18.5: The biosynthetic pathway of heme. Su, H.; Chen, X.; Chen, S.; Guo, M.; Liu, H. Applications of the Whole-Cell System in the Efficient Biosynthesis of Heme. *Int. J. Mol. Sci.* **2023**, *24*, 8384. https://doi.org/10.3390/ijms24098384. Creative Commons Attribution (CC BY) license (**https://creativecommons.org/licenses/by/4.0/**)

The two biosynthetic pathways of 5-aminolevulinic acid (bolded to indicate its importance) are shown in green (C4 pathway) and red (C5 pathway). The three downstream synthetic pathways of heme are marked with blue (CPD), indigo (SHD) and purple (PPD). Solid lines indicate single reactions, and dashed lines indicate more than two reactions. The names of genes encoding the individual enzymes are in italics and some reactions have alternative genes. The abbreviations of the corresponding enzymes are shown in the grey rectangle. See **Table 1** for a list of names and abbreviations for heme synthesis enzymes

Figure 32.18.6 below complements this figure and shows the synthetic biology strategies to enhance heme production.

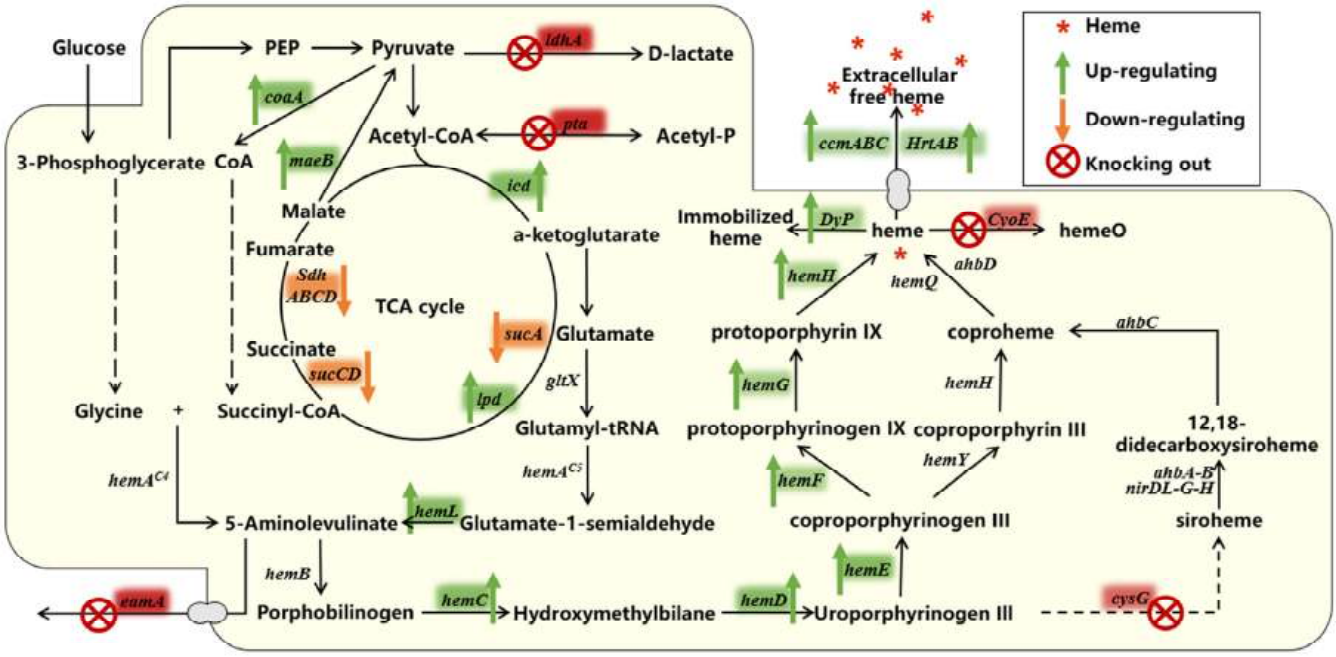


Figure 32.18.6: Synthetic biology strategies to enhance heme production. Green, orange, and red color blocks indicate genes that need to be up-regulated, down-regulated, and knocked out, respectively. Su, H. et al., ibid.

Computational tools such as AI can help the design of new pathways and novel enzymes to enhance production. It is becoming easier to transfer large pathways into yeast as well.

OTHER FOOD PRODUCTS FROM MICROBES AND SUSTAINABLE PLANTS

Significant effort is being devoted to growing meat in cell culture in the lab. This is a nascent field and has to overcome many problems, including consumer resistance to eating lab-created meat. At present meat grown in tissue culture is very expensive. Three key steps in growing meat are finding the best cells to grow, finding the nutrient conditions to maximize their growth, and adjusting conditions to make the lab meat taste like meat.

Muscle stem cells have been used as they can multiply many times, but these have growth limits. Alternatively, immortal cells, such as those derived from chicken fibroblasts, could be used. They can also be converted to fat cells. Yet they could accrue mutations with possible, but unlikely health consequences. Animal cells grown in culture often use fetal cow serum for its rich composition of growth factors and nutrients. However, it is very expensive and has ethical concerns as well since it's derived from animals. Synthetic growth medium can be used but it is also expensive. Whether lab-grown meat can overcome high costs and consumer resistance will determine its potential as a meat substitute.

More simply, people can use more peas, soy, grains, and nuts in their diet (i.e. being a vegan or vegetarian is the best approach to reducing your carbon footprint). Soy products have an extensive history of use as a source of protein but contain potential allergens (especially important in babies who use soy formulas) and isoflavones, which mimic human estrogen derivatives. Pea-based protein infant formulas are an increasing-used substitute.

Expressed recombinant proteins made in genetically modified bacteria and yeast are also becoming more popular. Examples (other than leghemogloblin) include the production in the fungus Trichoderma reesei of β-lactoglobulin, a cow whey protein, for dairy and animal-free milk products. The genetically-modified yeast Pichia pastoris has also been engineered to make milk casein proteins, egg-white proteins, muscle myoglobin, and human breast milk proteins. Enzymes used in the manufacture of cheese (derived from calves' stomachs) can be replaced by chymosin made in yeast. Production is linked to fermentation for many of these proteins. Filamentous proteins that have a texture similar to chicken fiber can be made through fermentation in the filamentous fungi Fusarium venenatumin. Macroalgae like seaweeds can provide high-protein food and have long been used in many cultures. Kelp farming can help not only provide protein but also capture carbon. Finally, insects, long eaten in many cultures, could become more climate-friendly source of protein.

If humans are in search of nonanimal sources of protein to fight climate change, why not produce and eat the most abundant protein in the biosphere, Rubisco? New products derived from the duckweed plant (genus Lemna) are coming to market. Figure 32.18.7:

Figure 32.18.7: Duckweed (and a frog). https://commons.wikimedia.org/wiki/F...7678481%29.jpg

Duckweed is high in nutrients, fast-growing, and a great source of Rubisco. It can be grown in aquaculture and does not require farmable land. It contains up to 50% protein. After harvesting, the plants are filtered, milled, and dried, which are all very simple technologies. Proteins, the most abundant being Rubisco, are then extracted. It can be used in baked goods and as meat and dairy substitutes. It is equal to eggs and meat in supplying all the essential amino acids required by humans.

GENETIC MANUFACTURING OF INDUSTRIAL FEEDSTOCKS

Let's look at one example in which synthetic biology and computational techniques are used to create products such as plexiglass from a biological source of acrylates. Acrylates are esters of acrylic acid (typically made from propylene) synthesized by reacting it with alcohols like methanol. Life cycle analyses show that almost 4000 kg CO_2 are produced per metric ton of acrylic acid made. To reduce the climate effect, biological feedstocks like glycerol and 3-hydroxypropanoic acid can be used, but large-scale supplies are needed. Figure 32.18.8 shows an overview of acrylate production fossil and biological feedstocks.

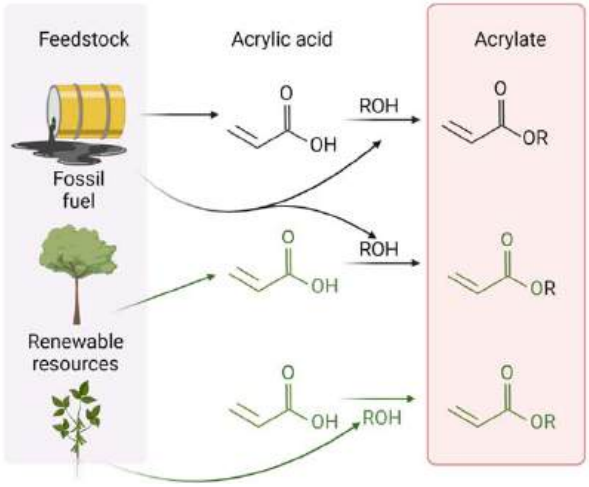


Figure 32.18.8: Production pathway of acrylates using fossil fuel and renewable resources. Souza, L.R.d.; Whitfield, B.; Al-Tabbaa, A. Biobased Acrylate Shells for Microcapsules Used in Self-Healing of Cementitious Materials. *Sustainability* **2022**, *14*, 13556. https://doi.org/10.3390/su142013556. Creative Commons Attribution (CC BY) license (**https://creativecommons.org/licenses/by/4.0/**)

When the alcohol is methanol, the final product is methylacrylate (MA). The structure of the cyclic acrylate monomer feedstock used to polymerize plexiglass (lucite) is methylene-butyrolactone, (MBL)whose structure is shown in Figure 32.18.9 below.

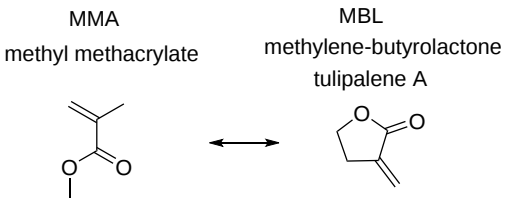


Figure 32.18.9: Structure of methyl methacrylate and its lactone (a cyclic ester)

Methyl methacrylate can undergo a free radical polymerization in the presence of an initiator (In·), as shown below in Figure 32.18.10.

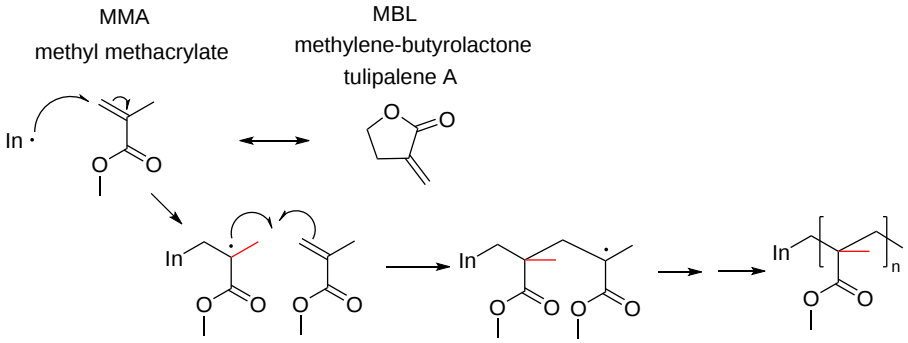


Figure 32.18.10: Mechanism of free radical polymerization of MMA

This reaction can form large polymers like plexiglass. The market for acrylic acid, the feedstock for plexiglass, is estimated to reach 12 million metric tons by 2030.

MBL, the lactone of MMA, is made in tulips from pathways that are not completely elucidated. It can also be used as a feedstock for the polymerization of plexiglass. Figure 32.18.11 shows the polymerization products from MMA and MBL.

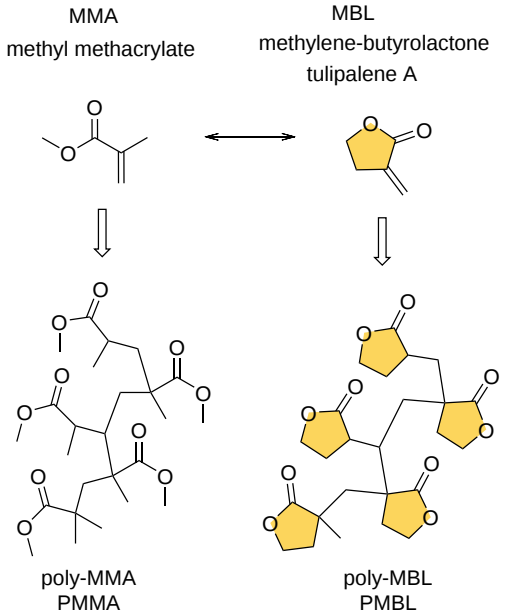


Figure 32.18.11: Structure of poly-MMA and poly-MBL

Using synthetic biology and advanced computational methods, plexiglass can now be made from biological sources instead of fossil fuels. To accomplish this, Azerda has designed synthetic pathways from millions of potential metabolic pathways (using a software package called Scylax™), and intelligently redesigned key enzymes to maximize their catalytic potential for the synthesis of MBL (using the software Archytas™). They used high-throughput DNA and protein analyses to maximize expression. Finally, they engineered expression strains and downstream purification processes to maximize the final output of MBL. In summary, the key steps in the process were:

- identifying a pathway from millions of reactions in databases of pathways that could produce MBL from simple sugars through a fermentation process;
- engineering pathway enzymes to greatly increase catalytic efficiency and decrease inhibition;

- producing test quantities of the products in cell strains;
- scaling up production to levels needed for purification and reactions of the MBL
- purifying sufficient amounts of MBL from large fermentation broths
- making the desired product (plexiglass, for example) from the feedstock.

Strains of bacteria, yeast, and filamentous fungi were modified to meet the above criteria. The ultimate substrate for the process was a lignocellulosic hydrolysate, so in the end the process converts trees to plexiglass (incredible to think about)! Of course, it is also amazing that CO_2 from the air, water, and minerals/ions from the soil can become a tree!

Starting with just a detectable level of product, the process was continually improved and scaled to eventually yield 5 g/L of broth, which is getting close to the 20 g/L required for commercial viability. Figure 32.18.12 below shows plexiglass created from the lignocellulosic stock!

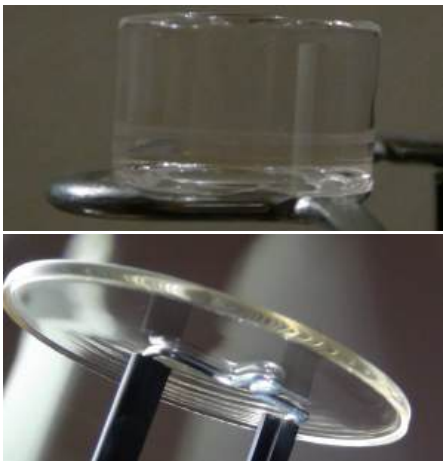


Figure 32.18.12: Plexiglass made from biosourced MBL.

Table 32.18.2 below compares the key physical properties of the polymers from Arzeda's PMBL compared to literature values for fossil-fuel-based PMBL and for PMMA.

Property	Measure	Lit PMBL	Arzeda PMBL	PMMA
Thermal	Glass transition pt T_g (^oC)	194-195	195	105
Mechanical	Elasticity (mPa)	1999/3439	5972	2855
	Tensile strength (mPa)	36.7/62.7	72.7	70
	Elongation at break	1.3%.6.5%	1.3%	2.5
Optical	Light transmission	NA	>88%	92%
Solvent resistance	toluene, 30 days, 20^oC	NA	Pass	Fail

Table 32.18.2: https://www.energy.gov/sites/default...-korkegian.pdf

CHAPTER OVERVIEW

Chapter 33: Your Contribution - Sandbox

Arthur Sikora Lipid rafts
BioMolViz Framework
Chetna IMF Test
Edmund - Structural Basis of Allostery, The Kinase Model
Emily Schmitt Sepiapterin Reductase-Beery Twins Story
Helena-Test
Inserting an iCn3D model into a FOB Chapter section
Inserting an interactive mathematic graph into FOB
Inserting a mathematical simulation (SBML) into a FOB
Insert an Image into FOB
Insert a question with hidden answer and floating hint
KP Procko Test
Make your own Chapter 33.x Section
Make Your Own Customized FOB: Assemble and Remix a Custom Book (Short Version)
Pallavi- Test
Pam Mertz Chapter 33 Test
Rebecca Roberts Chapter 33 testing
Rico Acevedo Testy
Samantha Wilner Ch 33 Test
Subhasish-test
Subhasish-TEST2

Arthur Sikora Lipid rafts

Exercise 1

What kinds of lipids are found in a eukaryotic cell membrane under normal conditions?

Here is a Hint if you need one

Answer

Glycerophospholipids, sphingolipids, sterols and fatty acids.

Authored by Arthur Sikora. Last update: 6/5/2023

Date of origin

1: Introduction

Lipid rafts are formed and dissipated over time. They are regions where specific lipids are enriched, frequently resulting in protein aggregation as well

New Heading

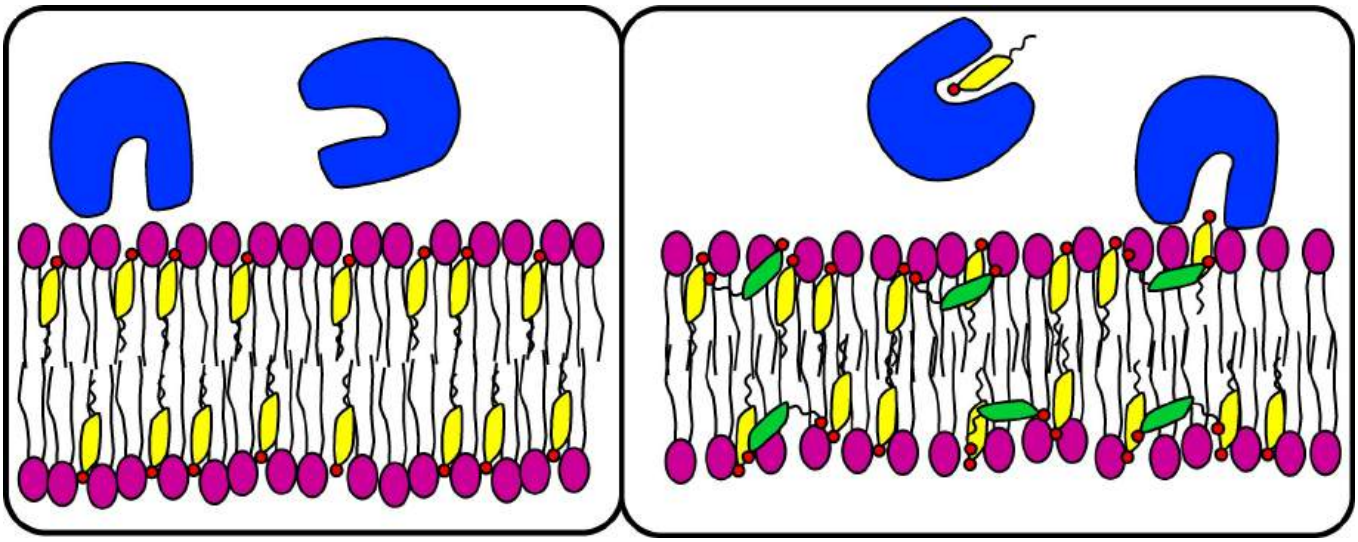
[ADD TEXT]

[ADD IMAGE] (saved to your computer and uploaded with picture icon from top menu bar or drag image file to location (required from svg image)

*Use the following under your picture:

Figure x: [Add caption]

*Center Picture and Caption together using top menu bar



Lipids being removed and added to a membrane

[ADD iCn3D Model]

Figure x is an interactive iCn3D model of [INSERT DESCRIPTIVE TEXT - often a modified version of the title from the actual PDB page 4Q7Q

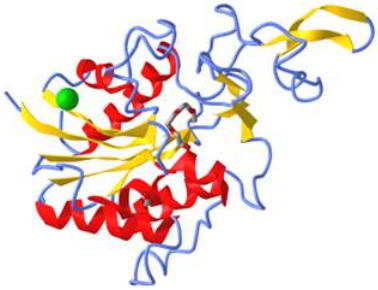
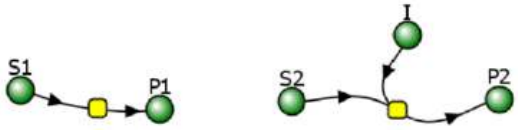


Figure 6: [INSERT THE PDB TITLE as above (INSERT PDB CODE)]. Click the image for a popup or use this external link: [https://structure.ncbi.nlm.nih.gov/i...PGDeccDmTH3Nq8]. (Copyright; author via source). iCn3D model made by **Arthur Sikora**

VCell MODEL

No inhibition (left) and Uncompetitive Inhibition (right)



Note that the Vcell reaction diagram is the same as for competitive and uncompetive inhibition. It doesn't explicitly show that the mixed inhibitor binds to both free and substrate-bound enzymes. Those interactions are addressed in the mathematical equations for mixed inhibition.

Initial values No Inhibition

Name	Description	Global	
J	reaction rate	☐	$\frac{Vmax \cdot S1}{(Km + S1)}$
Km	Km (1/2 max)	☐	5.0
Vmax	max reaction rate	☐	10.0
S1	Species Concentration	☑	Variable

Initial values With Uncompetitive Inhibitor

Name	Description	Global	
J	reaction rate	☐	$\frac{Vmax \cdot S2}{(Km \cdot (1.0 + \frac{I}{Kc}) + S2 \cdot (1.0 + \frac{I}{Ki}))}$
Km	user defined	☐	5.0
Vmax	user defined	☐	10.0
Kc	user defined	☐	1.0
Ki	user defined	☐	1.0
S2	Species Concentration	☑	Variable
I	Species Concentration	☑	Variable

I is fixed for each simulation (as it is not converted to a product) but can be changed in the simulation below.

Select Load [model name] below

Load NoInhib_MIXEDInhib_MM

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

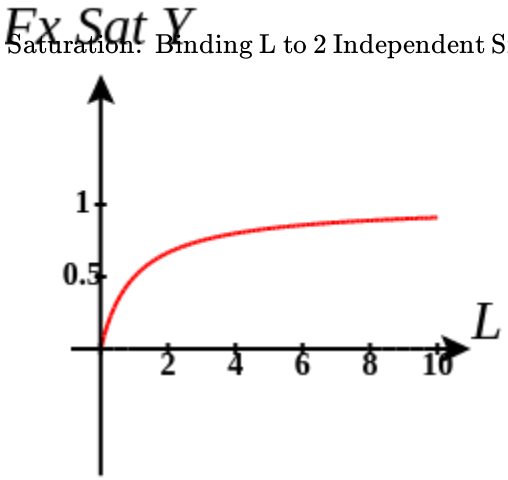
Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

---

sample modle

My modle



*your text and graph

Figure x: Add caption

[ADD VCELL/SBML SIMULATION - REUSE]

*your text before and after insert as needed

2: New Heading

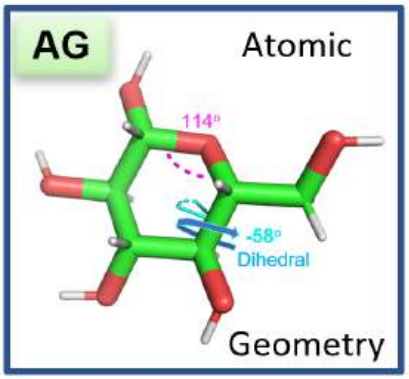
Add what you want

BIOMOLVIZ

Promoting Molecular Visualization Literacy

The BMV framework is used with permission from BioMolViz.Org

Copy the appropriate row when assigning a theme, goal, and objective to a designated iCn3D or other biomolecular visualization assessment

Atomic Geometry (AG)		Three-atom and four-atom (dihedral) angles, metal size and metal-ligand geometries, steric clashes

AG1. Students can describe the ideal geometry for a given atom within a molecule and deviations from the ideal geometry due to neighboring interactions.

AG1.01 Students can identify atomic geometry/hybridization for a given atom. (Novice)

AG1.02 Students can measure bond angles for a given atom. (Novice)

AG1.03 Students can identify deviations from the ideal bond angles. (Amateur)

AG1.04 Students can explain deviations from the ideal bond angles due to local effects. (Amateur, Expert)

AG1.05 Students can predict the effect of deviations from ideal bond angles on the structure and function of a macromolecule. (Expert)

AG1.06 Students can identify the geometric features of bonds (e.g., peptide bond, glycosidic, phosphoester).

AG2. Students can compare and contrast different structural conformations with regard to energy, the addition of substituents, and the impact on the structure/function of a macromolecule.

AG2.01 Students can describe different conformations that a structure can adopt using visualization tools. (Amateur)

AG2.02 Students can describe different conformations of atoms about a bond using visualization tools. (Novice)

AG2.03 Students can distinguish energetically favorable and unfavorable conformations that a structure can adopt. (Amateur)

AG2.04 Students can predict the effect of a given substituent on the structure and function of a macromolecule (e.g., substituent on a carbohydrate/ligand, R groups/rotamers, phosphorylation, methylation of nucleic acids, post-translational modifications). (Expert)

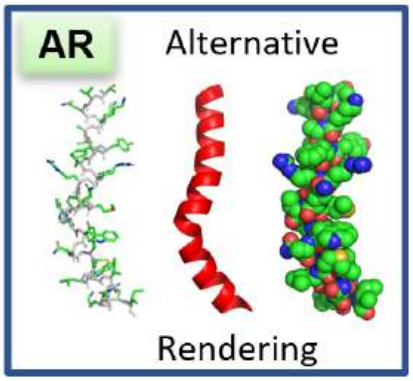
AG3. Students can describe dihedral/torsion angles in biomolecules.

AG3.01 Students can identify a dihedral/torsion angle in a three-dimensional representation of a molecule. (Novice)

AG3.02 Students can identify the planes between which a dihedral/torsion angle exists within a three-dimensional representation of a macromolecule. (Novice)

AG3.03 Students can identify phi, psi, and omega torsion/dihedral angles in a three-dimensional representation of a protein. (Amateur)

Alternate Renderings (AR)



Rendering of a macromolecular structure such as a protein or nucleic acid structure in various ways from the simplest possible way (connections between alpha carbons) to illustration of secondary structure (ribbons) to surface rendering and space filling.

AR1. Students can interpret or create molecular images that convey features such as secondary structure, CPK coloring, and active sites.

AR1.01 Students can manipulate rendered structures to illustrate molecular properties. (Novice)

AR1.02 REMOVED (integrated with SF2.02)

AR1.03 Students can describe or label structural differences among multiple structures. (Amateur, Expert)

AR1.04 Students can infer information from rendering a structure in different ways. (Novice, Amateur, Expert)

AR1.05 Students can create renderings that distinguish secondary structural features. (Novice)

AR1.06 Students can create an information rich rendering of a structure that depicts structural features found in the literature. (Amateur)

AR1.07 Students can create an information rich rendering of a structure containing ligands, covalent modifications, and noncanonical amino acids or nucleotides. (Amateur, Expert)

AR1.08 Students can use molecular visualization to tell a story about a macromolecular structure. (Expert)

AR1.09 REMOVED (integrated with MI1.02)

AR1.10 Students can convert textbook images of small molecules into 3D representations in a molecular visualization program. (Amateur)

AR2. Students can choose the best rendering of a macromolecule to use in a given situation.

AR2.01 Students can recognize that a cartoon rendering is a summary of the detail in a line rendering. (Novice, Amateur)

AR2.02 Students can describe the atoms and their representations in different renderings (e.g., coloring, showing hydrogens/double bonds). (Novice)

AR2.03 Students can identify or create a suitable rendering, or combination of renderings, for a specific purpose (e.g., a surface rendering overlaid with a cartoon to highlight the van der Waals surface alongside secondary structure, or active site sticks shown over a cartoon). (Novice, Amateur)

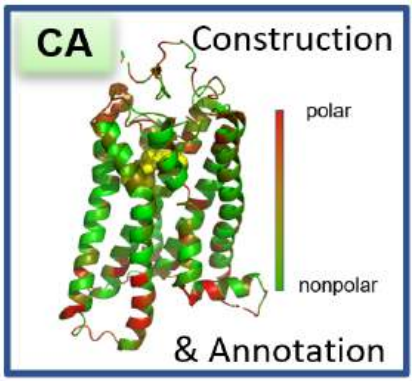
AR2.04 Students can identify the limitations in various renderings of molecular structures. (Amateur)

AR2.05 Students can understand the level of detail of different molecular representations. (Novice, Amateur, Expert)

AR2.06 Students can transition comfortably between equivalent 2D and 3D renderings of biomolecules. (Novice, Amateur, Expert)

AR2.07 Students can use and interpret color in the context of macromolecules to clarify and/or highlight features (e.g., coloring amino acids differently by property, different molecules uniquely in a complex, protein chains, secondary structure).

(Novice)

Construction and Annotation (CA)		Ability to build macromolecular models, either physical or computerized, and, where possible, add commentary, either written or verbal, to tell a molecular story.

CA1. Students can compose information-rich renderings of macromolecule-ligand interactions.

CA1.01 Students can construct and annotate a model of a macromolecule bound to a ligand. (Amateur)

CA1.02 Students can construct a model of a macromolecule bound to a ligand and identify the types of molecular interactions. (Amateur)

CA1.03 Students can construct a model of a macromolecule bound to a ligand and assess the importance of molecular interactions. (Expert)

CA1.04 Students can produce a model of a macromolecule based on a known structure of a related macromolecule. (Amateur, Expert)

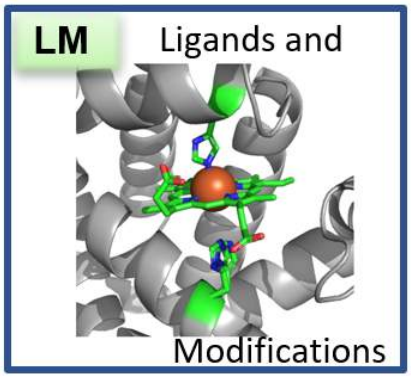
CA2. Students can compose a rendering to predict the cellular location of a protein (e.g., extracellular, membrane associated, or cytoplasmic) based on the properties and orientations of functional groups.

CA2.01 Students can design a rendering that conveys properties such as polarity, charge, secondary structure, etc. to suggest the cellular location of a macromolecule. (Amateur)

CA2.02 Students can create protein images with colored polar/nonpolar residues to determine whether they fold with a hydrophobic core. (Amateur)

CA2.03 Students can create images to display polar/nonpolar residues and propose a role for the protein and/or how it interacts with its environment - and that the predictions would be plausible based on the protein. (Amateur)

CA2.04 Students can make accurate predictions of the location/function of the protein that incorporates additional protein features, such as transmembrane helices, apparent docking surfaces, etc. (Expert)

Ligands and Modifications (LM)		Metals and metal clusters, additions such as glycosylation, phosphorylation, lipid attachment, methylation etc.

LM1. Students can identify ligands and modified building blocks (e.g., hydroxyproline, aminosaccharides, modified nucleobase) within a rendered structure.

LM1.01 Students can use the annotation associated with a pdb file to identify and locate ligands and modified building blocks in a given biomolecule. (Amateur)

LM1.02 Students can visually identify non-protein chemical components in a given rendered structure. (Amateur)

LM1.03 Students can distinguish between nucleic acid and ligands (e.g., metal ions) in a given nucleic acid superstructure. (Amateur)

LM1.04 Students can explain how a ligand in a given rendered structure associates with the biomolecule (e.g., covalent interaction with residue X). (Amateur)

LM1.05 Students can locate/identify ligands and modified building blocks in unannotated structures and describe their role. (Expert)

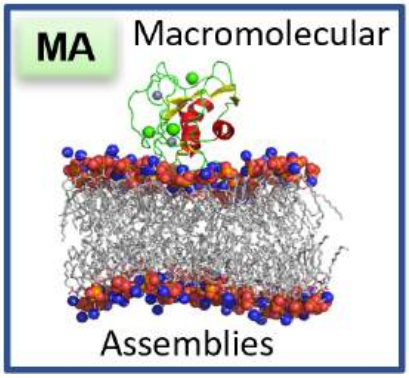
LM2. Students can describe the impact of a ligand or modified building block on the structure/function of a macromolecule.

LM2.01 Students can look at a given rendered structure and describe how the presence of a specific ligand or modified building block alters the structure of that biomolecule. (Amateur)

LM2.02 Students can explain how the removal of a particular ligand or modified building block would alter the structure of a given biomolecule. (Expert)

LM2.03 Students can use molecular visualization tools to predict how a specified ligand or modified building block contributes to the function of a given protein. (Amateur, Expert)

LM2.04 Students can predict how a ligand or modified building block contributes to the function of a protein for which the structure has been newly solved. (Expert)

Macromolecular Assemblies (MA)		Polypeptides, oligosaccharides, and nucleic acid and lipid superstructures (e.g. protein–nucleic acid complexes, lipid membrane-associated proteins)

MA1. Students can describe various macromolecular assemblies.

MA1.01 Students can identify individual biomolecules in a macromolecular assembly. (Novice, Amateur, Expert)

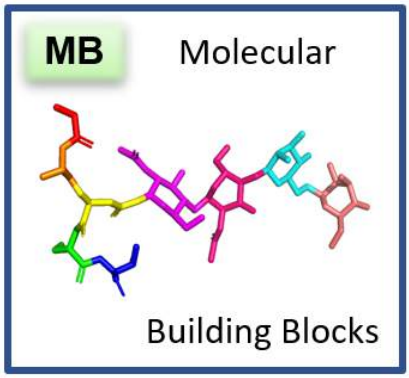
MA1.02 Students can describe functions of individual biomolecules within a macromolecular assembly. (Novice, Amateur, Expert)

MA1.03 Students recognize the various lipid ultrastructures (e.g., micelles, bicelles, vesicles, and lipid bilayers) in a 3D structure. (Novice)

MA2. Students can compose information-rich renderings of macromolecular assemblies.

MA2.01 Students can render a macromolecular assembly to highlight individual structures. (Amateur)

MA2.02 Students can render a macromolecular assembly to illustrate structural features (e.g., binding interfaces, symmetry, tertiary structure, etc.). (Novice, Amateur, Expert)

Macromolecular Building Blocks (MB)		Recognition of native amino acids, nucleotides, sugars, and other biomonomer units/building blocks. Understanding of their physical and chemical properties, particularly regarding functional groups.

MB1. Students can identify individual building blocks of biological polymers.

MB1.01 Given a rendered structure of a biological polymer, students can identify the ends of a biological polymer. (Novice, Amateur, Expert)

MB1.02 Given a rendered structure, students can divide the polymer into its individual building blocks. (Novice)

MB1.03 Given a rendered structure, students can identify the individual building blocks. (Novice)

MB2. Students can describe the contributions different individual building blocks make in determining the 3-D shape of the polymer.

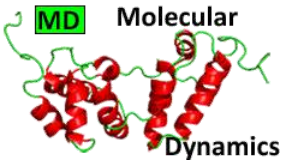
MB2.01 Students can describe the physical/chemical properties of an individual building block/functional group in a rendered structure of a polymer. (Amateur)

MB2.02 Students can describe the significance of the location of individual building blocks within the 3D structure of a polymer (protein, carbohydrate, or nucleic acid). (Novice, Amateur, Expert)

MB2.03 Students can identify physical/chemical properties of individual building blocks/functional groups in different local environments. (Amateur)

MB2.04 Using a visualized structure, students can identify stereochemistry (e.g., in carbohydrate, lipid, and protein structures). (Amateur)

MB2.05 Students can modify/mutate a building block to change the 3D structure of a polymer (protein, carbohydrate, or nucleic acid). (Amateur, Expert)

Molecular Dynamics (MD)		Animated motion simulating conformational changes involved in ligand binding or catalysis, or other molecular motion/dynamics.

MD1. Students can describe the impact of the dynamic motion of a biomolecule on its function.

MD1.01 Students can recognize that biological molecules have different conformations. (Novice, Amateur)

MD1.02 Students can correlate molecular movement with function. (Novice, Amateur, Expert)

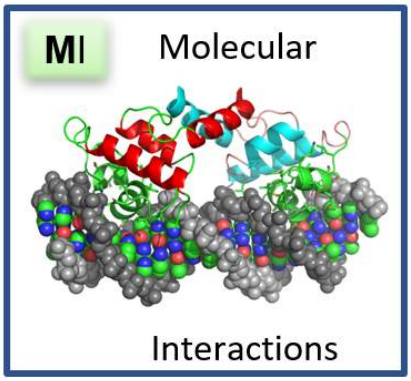
MD2. Students can predict limits to macromolecular movement.

MD2.01 Students can locate potential regions of flexibility and inflexibility in the structure of a biomolecule. (Novice, Amateur)

MD2.02 Students can recognize acceptable/unacceptable movement within a macromolecule by determining whether the movement is within allowable bond angles. (Expert)

MD2.03 Students can recognize acceptable/unacceptable movement within a macromolecule by determining whether the movement results in steric hindrance. (Amateur)

MD2.04 Students can recognize acceptable/unacceptable movement within a macromolecule by considering the atomic packing constraints. (Expert)

Molecular Interactions (MI)		Covalent and noncovalent bonding governing ligand binding and subunit-subunit interactions.

MI1. Students can predict the existence of an interaction using structural and environmental information (e.g. bond lengths, charges, pH, dielectric constant).

MI1.01 Students can distinguish between covalent and noncovalent interactions. (Novice)

MI1.02 Students can identify different noncovalent interactions (e.g., hydrogen bonds, ionic interactions, van der Waals contacts, induced dipole) given a 3D structure. (Amateur)

MI1.03 Students can predict whether a functional group (region) would be a hydrogen bond donor or acceptor. (Amateur)

MI1.04 Students can render the 3D structure of a biomolecule so as to demonstrate the ionic interactions and/or charge distribution of the different non-covalent interactions. (Amateur)

MI1.05 As it relates to a particular rendered structure, students can rank the relative strengths of covalent and noncovalent interactions. (Amateur)

MI2. Students can evaluate the effect of the local environment on various molecular interactions.

MI2.01 Students can identify regions of a biomolecule that are exposed to or shielded from solvent. (Novice)

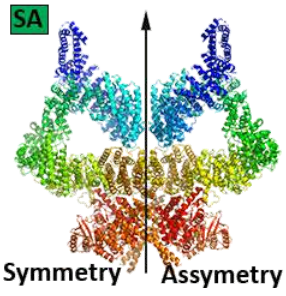
MI2.02 Students can identify other molecules in the local environment (e.g., solvent, salt ions, metals, detergents, other small molecules) that impact a molecular interaction of interest. (Novice)

MI2.03 Students can predict the impact of other molecules in the local environment (e.g., solvent, salt ions, metals, detergents, other small molecules) on a molecular interaction of interest. (Amateur)

MI2.04 Students can predict the pKa of an ionizable group based on the influence of its local three-dimensional environment. (Amateur)

MI2.05 Students can propose a change to the local environment that would yield a desired change in a molecular interaction. (Expert)

MI2.06 Using molecular visualization tools, students can determine which intermolecular force is most critical to stabilizing a given interaction. (Expert)

Symmetry/ Asymmetry Recognition (SA)		Recognition of symmetry elements within both single chain and multi-chain macromolecules.

SA1. Students can identify symmetric or asymmetric features in rendered molecules.

SA1.01 Students can identify symmetric features in a rendered molecule (shown in fixed orientation). (Novice)

SA1.02 Students can rotate a single macromolecule, multi-chain macromolecules (e.g., homo- or heteromers), complexes of macromolecules, and supramolecular assemblies to identify axes of symmetry. (Amateur)

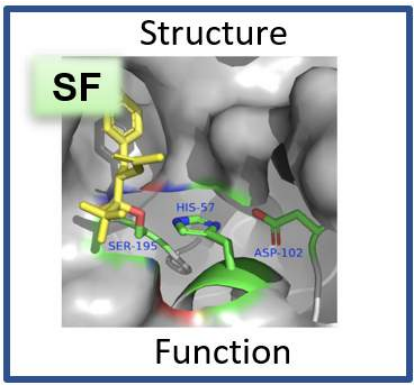
SA1.03 Students can identify symmetric and asymmetric features in rendered molecules after coloring a given rendered molecule to reveal structural features (charge, hydrophobicity, etc.). (Amateur)

SA2. Students can hypothesize the functional significance of symmetry or asymmetry in rendered molecules.

SA2.01 Students can explain the functional significance of rotational axes of symmetry (or asymmetry) in a given rendered molecule. (Novice, Amateur, Expert)

SA2.02 Students can predict functional significance of symmetry (or asymmetry) in a given rendered molecule. (Amateur, Expert)

Structure-Function Relationship (SF)



Active/binding sites, microenvironments, nucleophiles, redox centers, etc. (please also see LM2.03)

SF1. Students can evaluate biomolecular interaction sites using molecular visualization tools.

SF1.01 Students can identify functionally relevant cofactors, ligands or substrates associated with a macromolecule and describe their role (e.g., an active site magnesium ion). (Amateur, Expert)

SF1.02 Students recognize that the size and shape of the ligand must match the size and shape of the binding site. (Novice, Amateur)

SF1.03 Students recognize that the polarity or electrostatic potential of a surface complements that of the ligand or substrate. (Novice, Amateur)

SF1.04 Students recognize that the hydrophobicity of a surface complements that of the ligand or substrate. (Novice, Amateur)

SF1.05 REMOVED (integrated with SF1.03)

SF1.06 Students can use docking software to predict how the surface properties of a macromolecule guide and allow the binding of a ligand or substrate. (Amateur)

SF2. Using molecular visualization, students can predict the function of biomolecules.

SF2.01 Students can recognize structurally related molecules. (Novice)

SF2.02 Students can superimpose structurally related molecules. (Novice, Amateur)

SF2.03 Students can identify functionally relevant features of a macromolecule (e.g., an active site cysteine, a functional loop). (Amateur)

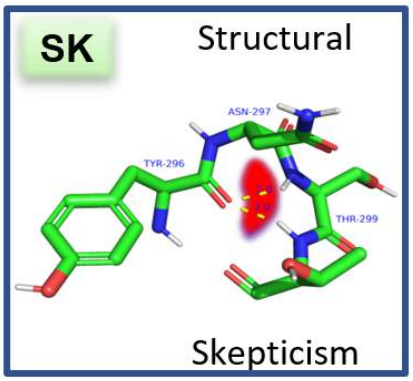
SF2.04 Students can predict molecular function given a binding site. (Amateur, Expert)

SF3. Using molecular visualization, students can predict the function of an altered macromolecule.

SF3.01 Students can structurally alter a macromolecule. (Novice)

SF3.02 Students can propose structural alterations to test interactions in a macromolecule. (Amateur)

SF3.03 Students can predict the impact of a structural alteration on the function of a macromolecule. (Amateur, Expert)

Structural Model Skepticism (SK)		Recognition of the limitations of models to describe the structure of macromolecules.

SK1. Students can critique the limitations of a structural model of a macromolecule.

SK1.01 Students can explain that the pdb file is a model based on data and that, as a model, it has limitations. (Novice, Amateur)

SK1.02 Students associate resolution with reliability of atom positions. (Amateur)

SK1.03 Students can identify building blocks (for example, amino acid side chains) whose orientation in a biopolymer is uncertain. (Expert)

SK1.04 Students can evaluate the flexibility/disorder of various regions of a macromolecular structure. (Novice, Amateur, Expert)

SK1.05 Students can reconcile inconsistent numbering of individual building blocks among species and structure files. (Novice)

SK1.06 Students can utilize a Ramachandran plot/steric clashes to interpret the validity of a structure. (Amateur, Expert)

SK1.07 Students can describe the limitations of a macromolecule-ligand docking simulation. (Amateur, Expert)

SK2. Students can evaluate the quality of 3D models including features that are open to alternate interpretations based on molecular visualization and PDB flat files.

SK2.01 Students can evaluate a crystal structure for crystal packing effects. (Novice, Amateur, Expert)

SK2.02 Students can resolve differences between the asymmetric unit and the functional biological assembly. (Expert)

SK2.03 Students can differentiate functional ligands (with biological/biochemistry role) from nonfunctional ligands (most solvents, salts, ions, and crystallization agents). (Novice, Amateur, Expert)

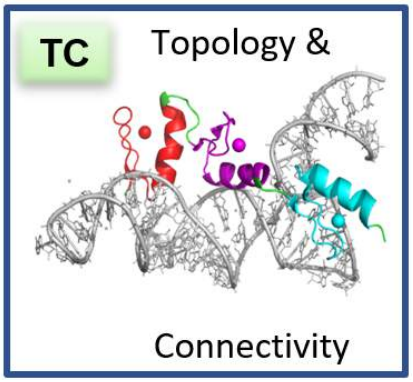
SK3. Students can discuss the value of experimentally altering a biomolecule to facilitate structure determination.

SK3.01 Students can identify non-native structural features. (Amateur)

SK3.02 Students can propose molecular modifications to facilitate structure determination. (Amateur, Expert)

SK3.03 Students can propose a purpose for the introduction of non-native structural features to facilitate structure determination. (Amateur, Expert)

Topology and Connectivity (TC)



Following the chain direction through the molecule, translating between 2D topology mapping and 3D rendering.

TC1. Students can describe or illustrate the linkages between building blocks within a macromolecule.

TC1.01 Students can trace the backbone of a macromolecule in three dimensions. (Novice, Amateur)

TC1.02 Students can use appropriate terms to describe the linkages/bonds/interactions that join individual building blocks together in a macromolecule or macromolecular assembly. (Novice, Amateur)

TC1.03 Given a virtual model of individual building blocks, students can predict the types of linkages/bonds/interactions that are possible or favorable. (Amateur)

TC1.04 Given individual building blocks, students can appropriately connect them to create a biological polymer (e.g., drawing carbohydrate linkages, a small peptide). (Amateur, Expert)

TC2. Students can describe the overall shape and common motifs within a 3D macromolecular structure.

TC2.01 Using molecular visualization software, students can describe the three-dimensional structure of a macromolecule, including overall shape and common structural motifs. (Novice, Amateur, Expert)

TC2.02 Students can identify common domains/motifs within a macromolecule. (Amateur, Expert)

TC2.03 Students can identify connectivity features between domains or subunits in a macromolecular structure. (Amateur)

TC2.04 Students can identify interactions between domains or subunits in a macromolecular structure. (Amateur, Expert)

TC2.05 Students can describe how domains/motifs in a macromolecule work together to achieve a concerted function in the cell. (Amateur, Expert)

TC2.06 Students can identify the levels of protein structure (e.g., parse a tertiary/quaternary structure into a series of secondary structures/motifs) and the ways in which they are connected from a three-dimensional structure. (Novice, Amateur, Expert)

TC3. Students can explain how any given biomolecular interaction site can be made by a variety of topologies.

TC3.01 Students can recognize that the groups that comprise a functional site only require proper arrangement in three-dimensional space rather than a particular order or position in the linear sequence. (Amateur)

TC3.02 Students can recognize similarities and differences in two similar - but not identical - three dimensional structures. (Amateur)

TC3.03 Students can describe dissimilar portions of homologous proteins as arising from genetic insertions/deletions/rearrangements. (Amateur)

Chetna IMF Test

Authored by [YOUR NAME]. Last update: [FILL IN DATE]

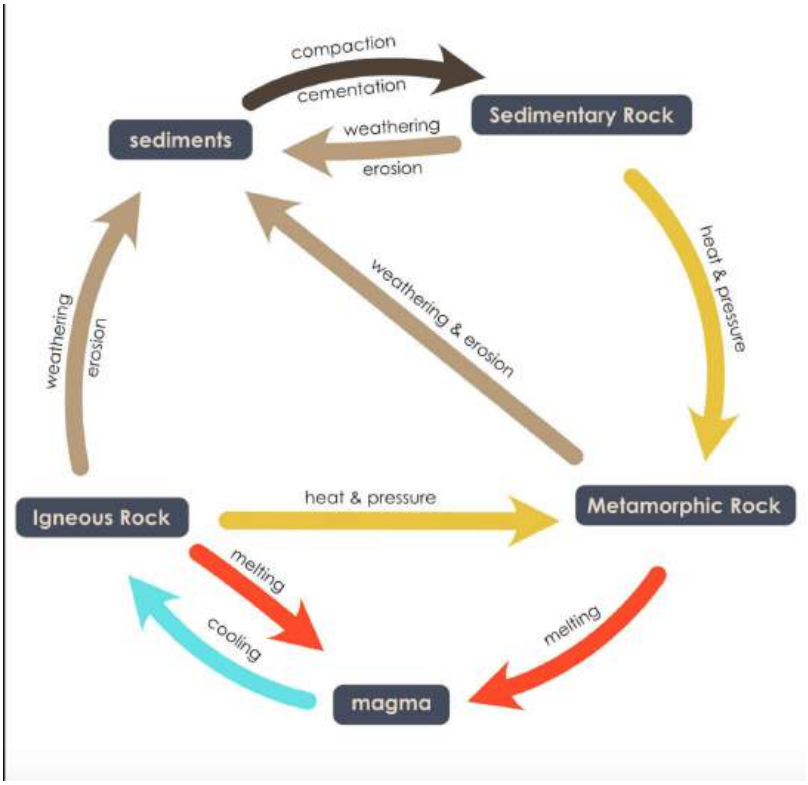
Date of origin

1: Introduction

[ADD CONTENT]

2: New Heading

[ADD TEXT]



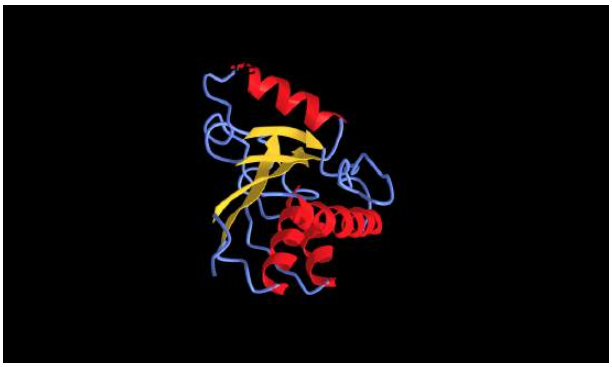
[ADD IMAGE] (saved to your computer and uploaded with picture icon from top menu bar or drag image file to location (required from svg image)

*Use the following under your picture:

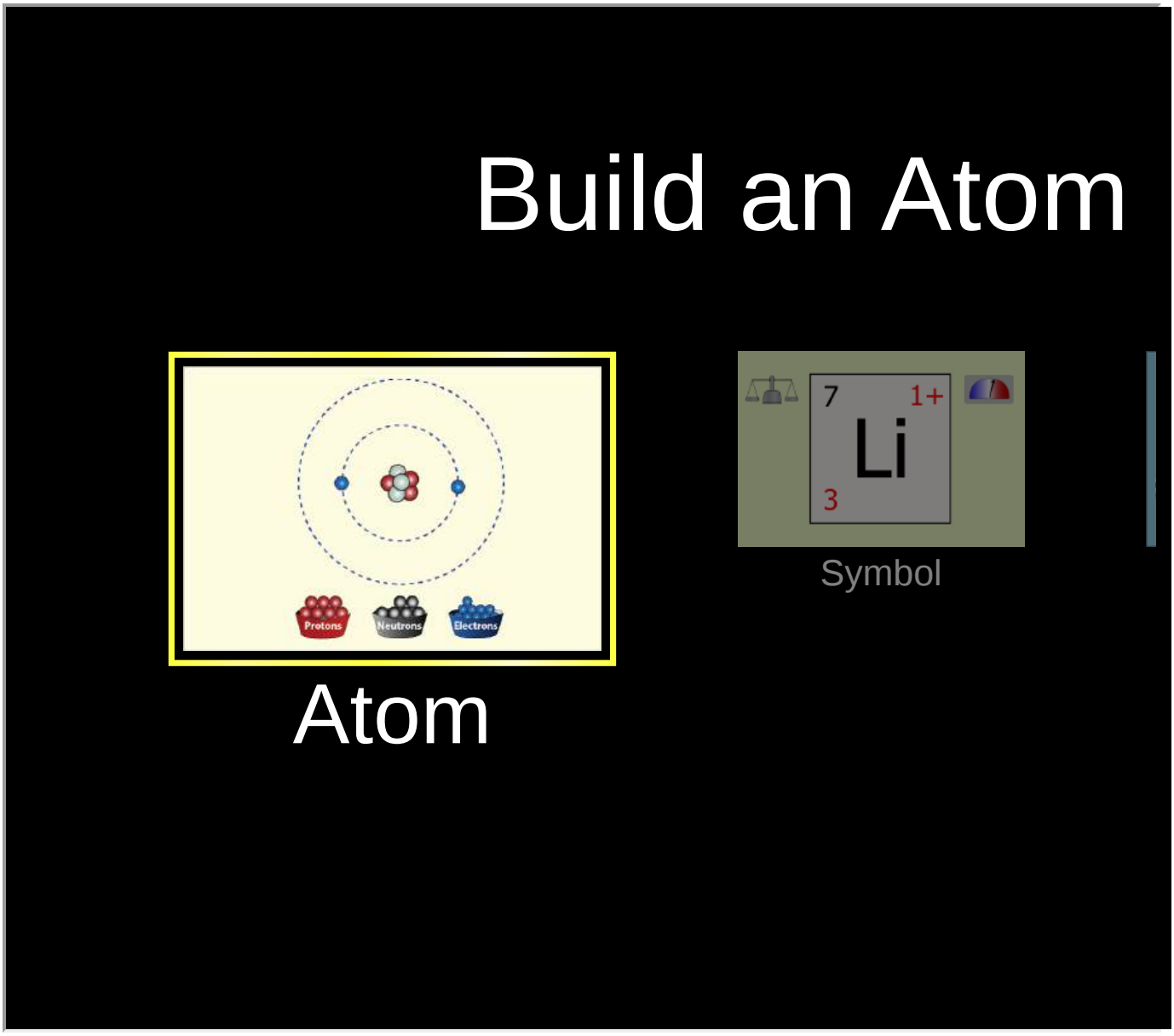
Figure x: [Add caption]

*Center Picture and Caption together using top menu bar

[ADD iCn3D Model]



[ADD MATHEMATIC GRAPH - REUSE]



d graph

Figure x: Add caption

[ADD VCELL/SBML SIMULATION - REUSE]

*your text before and after insert as needed

3: New Heading

Add what you want

Edmund - Structural Basis of Allostery, The Kinase Model

Authored by [YOUR NAME]. Last update: [FILL IN DATE]

Date of origin

1: Introduction

[ADD CONTENT]

2: New Heading

[ADD TEXT]

[ADD IMAGE] (saved to your computer and uploaded with picture icon from top menu bar or drag image file to location (required from svg image)

*Use the following under your picture:

Figure x: [Add caption]

*Center Picture and Caption together using top menu bar

[ADD iCn3D Model]

Figure x is an interactive iCn3D model of **[**INSERT DESCRIPTIVE TEXT - often a modified version of the title from the actual PDB page (INSERT PDB CODE)**]**

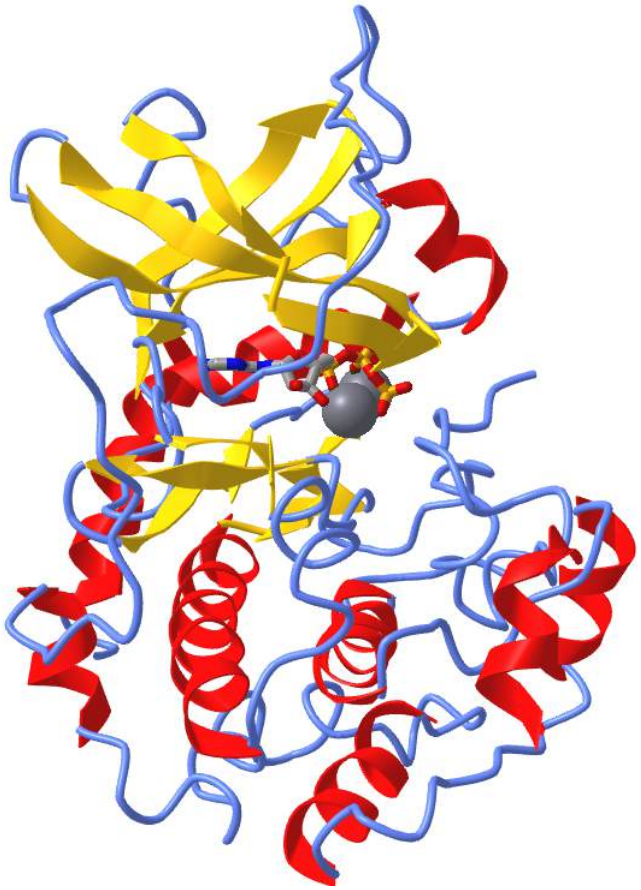
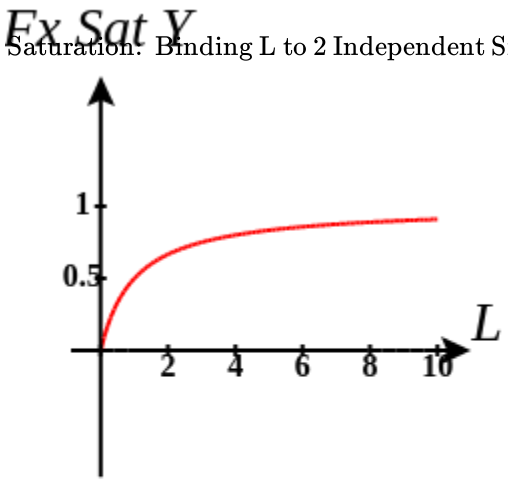


Figure 66: **[**INSERT THE PDB TITLE as above (INSERT PDB CODE)**]**. Click the image for a popup or use this external link: **https://structure.ncbi.nlm.nih.gov/i...KbCWx2etVfC2WA**. (Copyright; author via source). iCn3D model made by **[**YOUR NAME**]**

[ADD MATHEMATIC GRAPH - REUSE]



*your text and graph

Figure x: Add caption

[ADD VCELL/SBML SIMULATION - REUSE]

*your text before and after insert as needed

3: New Heading

? Exercise 1

What mutations to specific amino acid residues could be introduced to strengthen contacts between the upper and lower C-spine?

Answer

Mutations to create a greater extent of interaction or a stronger class of interaction are necessary to strengthen intramolecular contacts.

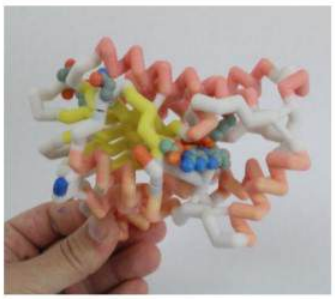
Example 1: Mutation of a small hydrophobic amino acid to a bulky hydrophobic or aromatic amino acid can be used to bridge the gap.

Example 2: Mutation to introduce a new hydrogen bond, salt bridge, or disulfide bond can be used to create a new link between these sections of the hydrophobic core.

Emily Schmitt Sepiapterin Reductase-Beery Twins Story

? Exercise 1

Which mutation was contributed by the mother that affected the Sepiapterian reductase (SPR) in the Beery Twins situation?



Answer

a) early stop of the protein (Lys to termination) Lys251X

b) Arg 150 Gly

c) there was no mutation from the mother that caused the disease

d) Lys 250 Glu

Here is a hint if you need one

Authored by [YOUR NAME]. Last update: [FILL IN DATE]

Date of origin

1: Introduction - Meet the Beery Twins

2: New Heading

[ADD TEXT]

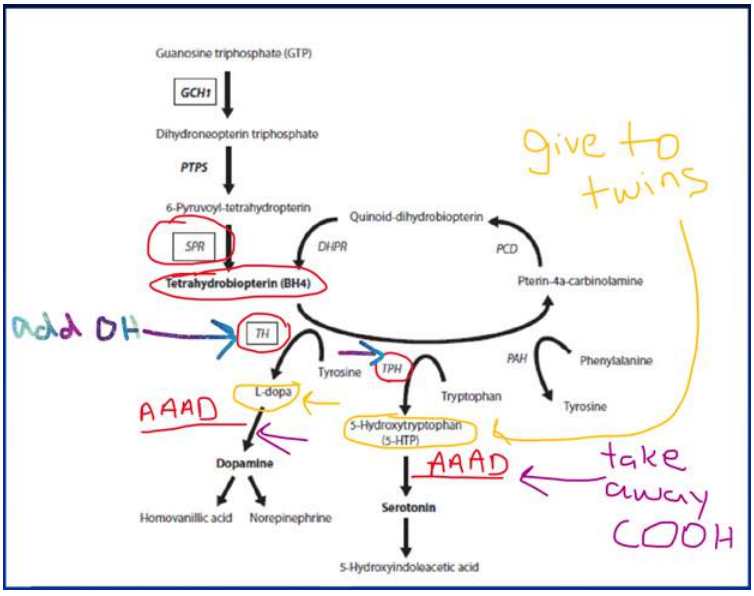


Figure 1: The pathway above shows where the drugs given to the twins exist in the pathway that involves Sepiapterin Reductase (SPR). Note that the twins were prescribed 5-HTP and L-Dopa which are downstream of the "broken" SPR. In this way dopamine and serotonin can still be produced in the twins.

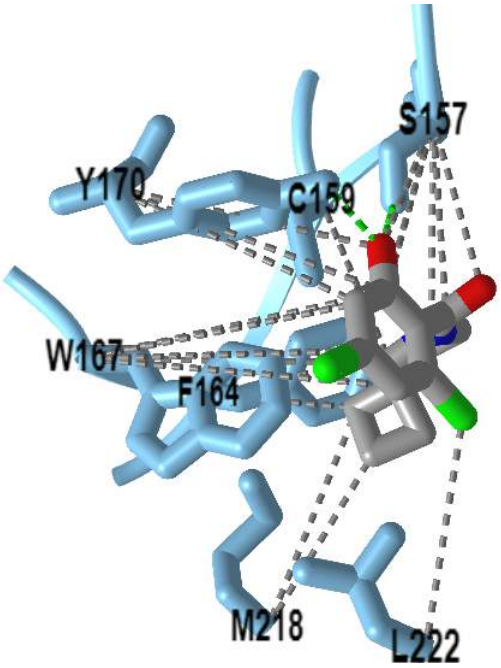


Figure 2: iCN3D recreation of Alen et al., 2019 Figure 5- Crystal structure of inhibitor 3 bound to SPR (PDB code 6I6P)

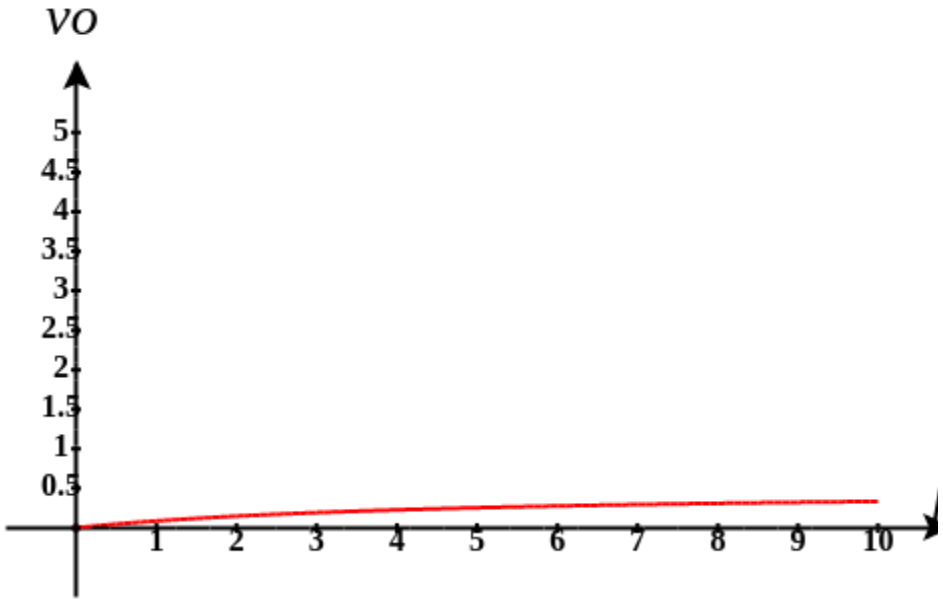
Figure x is an interactive iCn3D model of [INSERT DESCRIPTIVE TEXT - often a modified version of the title from the actual PDB page (INSERT PDB CODE)]

INSERT PNG (just a screen snip) OF YOUR iCn3D MODEL

Figure 6: [INSERT THE PDB TITLE as above (INSERT PDB CODE)]. Click the image for a popup or use this external link: [INSERT Lifelong short URL from File, Share Link in iCn3D)]. (Copyright; author via source). iCn3D model made by [YOUR NAME]

[ADD MATHEMATIC GRAPH - REU

— Content Reuse —



Page path to reuse:

SE]

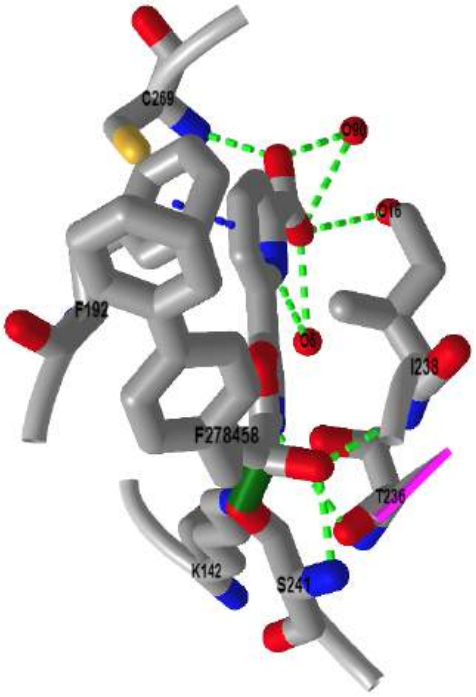
*your text and graph

Figure x: Add caption

[ADD VCELL/SBML SIMULATION - REUSE]

*your text before and after insert as needed

3: 3k83 Interactive iCn3D model



Helena-Test

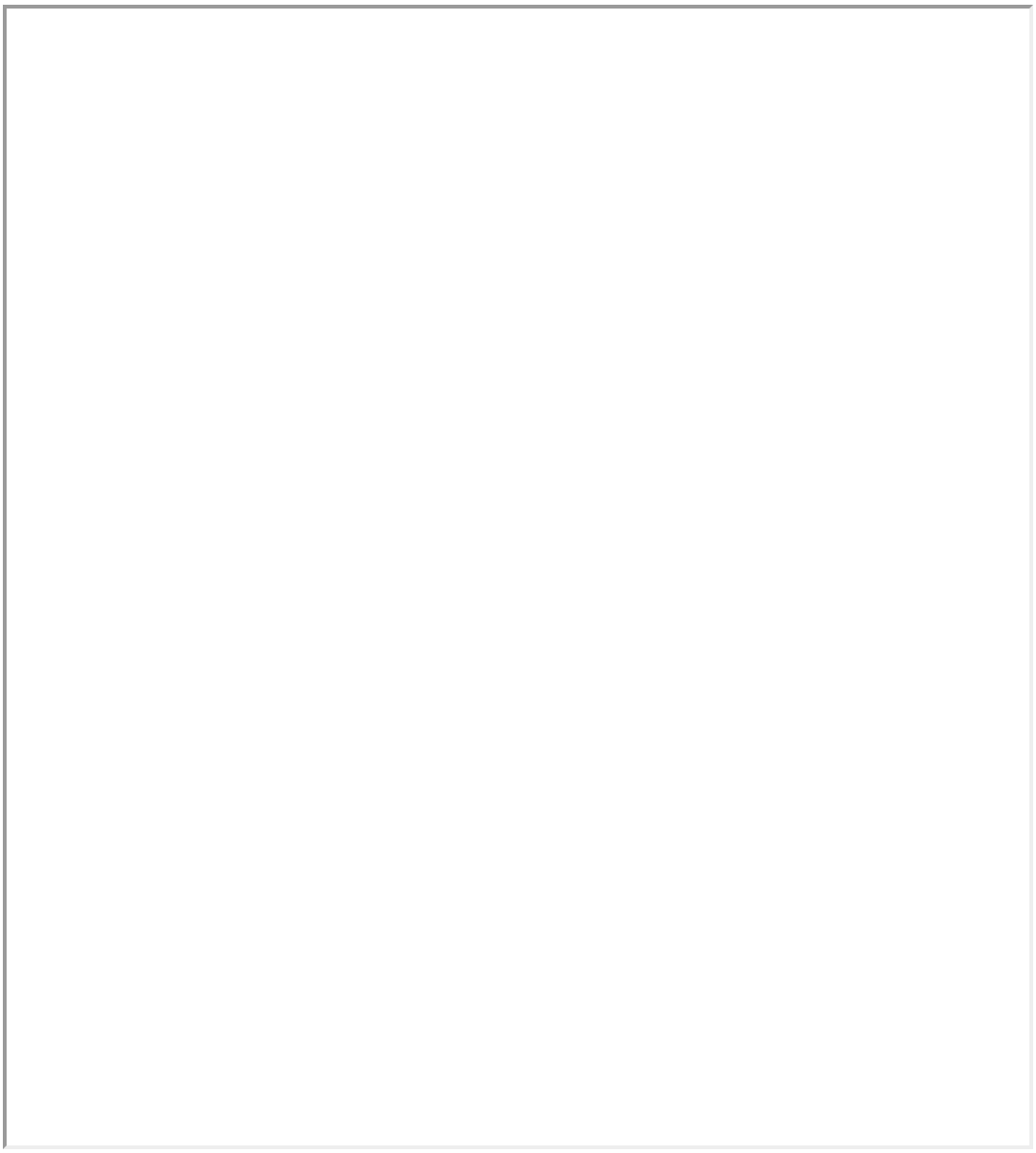
Exercise 1

Degradation of amino acids yields compounds that are common intermediates in the major metabolic pathways. Explain the distinction between glucogenic and ketogenic amino acids in terms of their metabolic fates.

Answer

Glucogenic amino acids are those which can be catabolized into pyruvate, oxaloacetate, a-ketoglutarate , fumarate, or succinyl-CoA, and thus can serve as glucose precursors.

Ketogenic amino acids are catabolized to Acetyl-CoA or acetoacetate, and thus can serve as precursors for fatty acids or ketone bodies.



Here is a hint if you need it.

Authored by Helena Prieto. Last update: 06.05.23

Date of origin 06.05.23

1: Introduction

[ADD CONTENT]

2: New Heading

[ADD TEXT]

[ADD IMAGE] (saved to your computer and uploaded with picture icon from top menu bar or drag image file to location (required from svg image)

*Use the following under your picture:

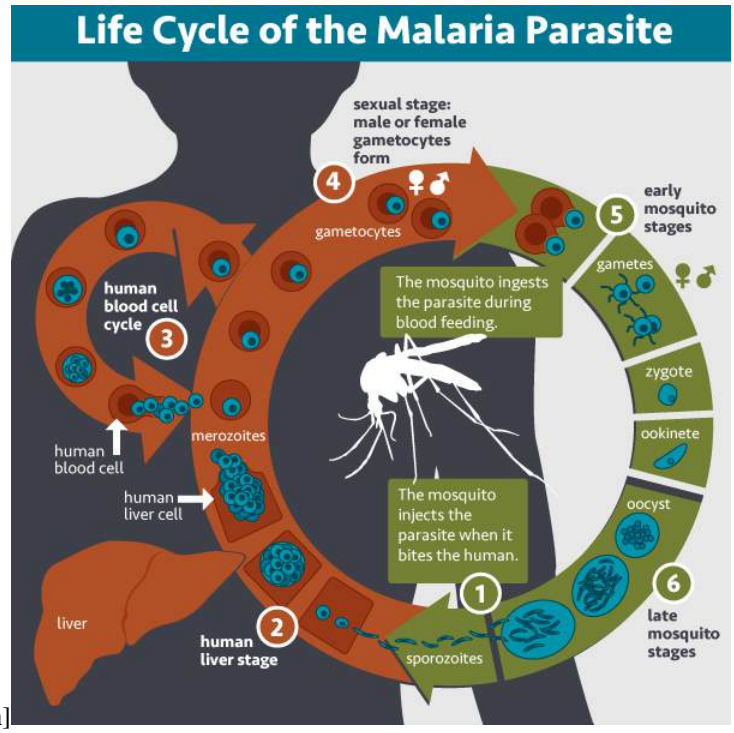


Figure x: [Add caption]

*Center Picture and Caption together using top menu bar

[ADD iCn3D Model]

[ADD MATHEMATIC GRAPH - REUSE]

*your text and graph

Reversible Competitive inhibition occurs when substrate (S) and inhibitor (I) both bind to the same site on the enzyme. In effect, they compete for the active site and bind in a mutually exclusive fashion. This is illustrated in the chemical equations and molecular cartoons shown in Figure 1.

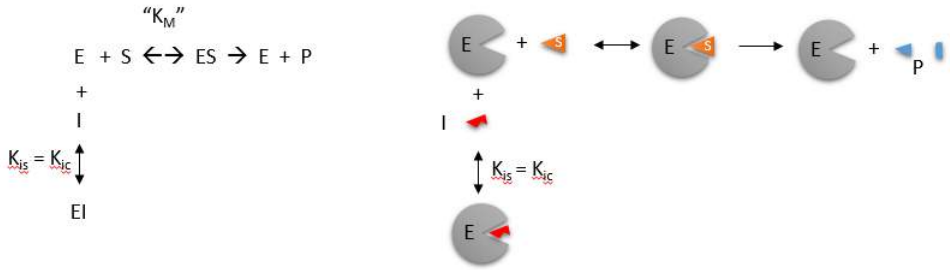


Figure 1: Competitive Inhibition

$$v_0 = \frac{V_M S}{K_M\left(1+\frac{I}{Kis}\right)+S} \tag{1}$$

There is another type of inhibition that would give the same kinetic data. If S and I bound to different sites, and S bound to E and produced a conformational change in E such that I could not bind (and vice versa), then the binding of S and I would be mutually exclusive. This is called allosteric competitive inhibition. Inhibition studies are usually done at several fixed and non-saturating concentrations of I and varying S concentrations.

The key kinetic parameters to understand are V_M and K_M. Let us assume for ease of equation derivation that I binds reversibly, and with rapid equilibrium to E, with a dissociation constant K_{IS}. The "s" in the subscript "is" indicates that the slope of the 1/v vs 1/S Lineweaver-Burk plot changes while the y-intercept stays constant. K_{IS} is also named K_{IC} where the subscript "c" stands for competitive inhibition constant.

A look at the top mechanism shows that even in the presence of I, as S increases to infinity, all E is converted to ES. That is, there is no free E to which I could bind. Now, remember that $V_M = k_{cat}E_0$. Under these conditions, ES = E_0; hence v = V_M. V_M is not changed. However, the apparent K_M, K_{Mapp}, will change. We can use LaChatelier's principle to understand this. If I binds to E alone and not ES, it will shift the equilibrium of E + S → ES to the left. This would increase the K_{Mapp} (i.e. it would appear that the affinity of E and S has decreased.). The double reciprocal plot (Lineweaver-Burk plot) offers a great way to visualize the inhibition as shown in Figure 2.

$$v_0 = \frac{V_M S}{K_M\left(1+\frac{I}{Kis}\right)+S}$$

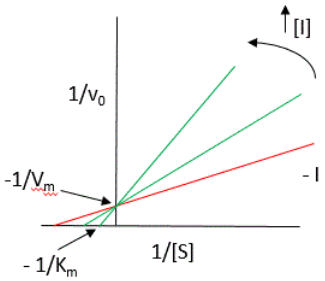
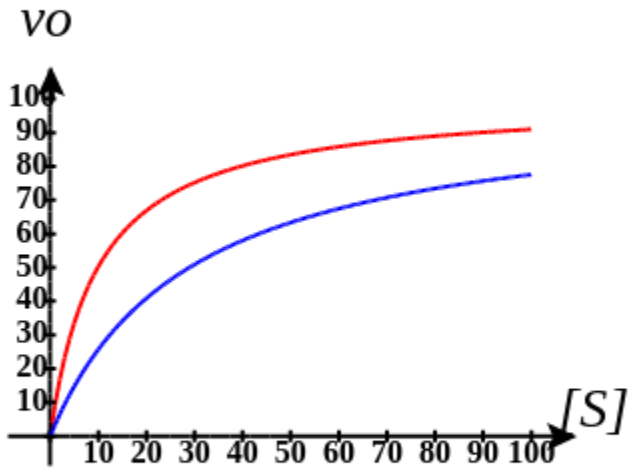


Figure 2: Competitive Inhibition: Lineweaver-Burk plots

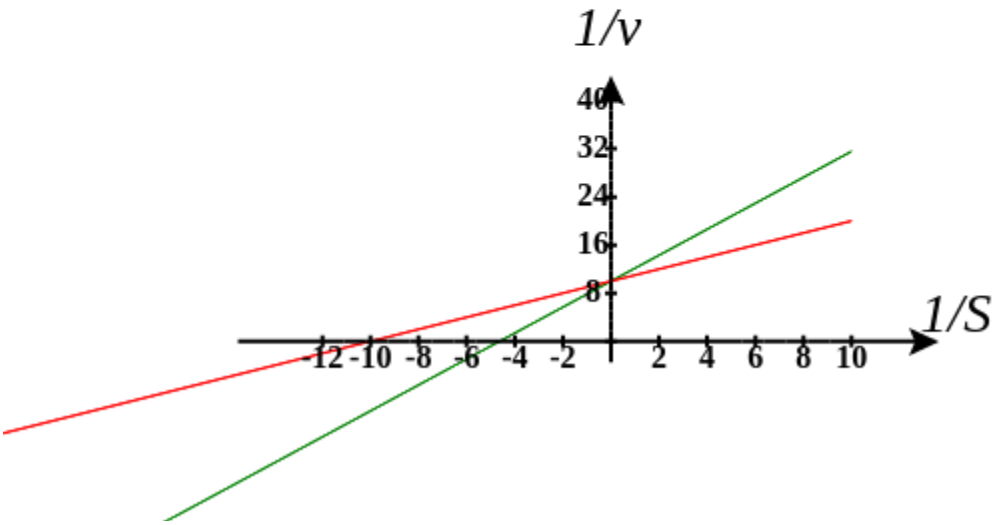
In the presence of I, V_M does not change, but K_M appears to increase. Therefore, $1/K_M$, the x-intercept on the plot will get smaller, and closer to 0. Therefore the plots will consist of a series of lines, with the same y-intercept ($1/V_M$), and the x-intercepts ($-1/K_M$) closer and closer to 0 as I increases. These intersecting plots are the hallmark of competitive inhibition.

Here is an interactive graph showing v_0 vs [S] for competitive inhibition with Vm and Km both set to 100. Change the sliders for [I] and Kis and see the effect on the graph.



Here is the interactive graphs showing $1/v_0$ vs 1/[S] for competitive inhibition, with Vm and Km both set to 10.

Competitive Inhibition



Note that in the first three inhibition models discussed in this section, the Lineweaver-Burk plots are linear in the presence and absence of an inhibitor. This suggests that plots of v vs S in each case would be hyperbolic and conform to the usual form of the Michaelis Menton equation, each with potentially different apparent V_M and K_M values.

An equation for v_0 in the presence of a competitive inhibitor is shown in the above figure. The only change compared to the equation for the initial velocity in the absence of the inhibitor is that the K_M term is multiplied by the factor $1+I/K_{is}$. Hence $K_{Mapp} = K_M(1+I/Kis)$. This shows that the apparent K_M does increase as we predicted. K_{IS} is the inhibitor dissociation constant in which the inhibitor affects the slope of the double reciprocal plot.

If the data were plotted as v_0 vs log S, the plots would be sigmoidal, as we saw for plots of ML vs log L in Chapter 5B. In the case of a competitive inhibitor, the plot of v_0 vs log S in the presence of different fixed concentrations of inhibitor would consist of a series of sigmoidal curves, each with the same V_M, but with different apparent K_M values (where $K_{Mapp} = K_M(1+I/K_{is})$, progressively shifted to the right. Enzyme kinetic data is rarely plotted this way. These plots are mostly used for simple binding data for the M + L ↔ ML equilibrium, in the presence of different inhibitor concentrations.

Reconsider our discussion of the simple binding equilibrium, M + L ↔ ML. For fractional saturation Y vs a log L graphs, we considered three examples:

1. $L = 0.01\ K_D$ (i.e. $L << K_D$), which implies that $K_D = 100L$. Then $Y = L/[K_D+L] = L/[100L + L] \approx 1/100$. This implies that irrespective of the actual [L], if $L = 0.01\ K_D$, then $Y \approx 0.01$.
2. $L = 100\ K_D$ (i.e. $L >> K_D$), which implies that $K_D = L/100$. Then $Y = L/[K_D+L] = L/[(L/100) + L] = 100L/101L \approx 1$. This implies that irrespective of the actual [L], if L = 100 Kd, then Y ≈1.
3. $L = K_D$, then Y = 0.5

These scenarios show that if L varies over 4 orders of magnitude ($0.01K_D < K_D < 100K_D$), or, in log terms, from $-2 + \log K_D < \log K_D < 2 + \log Kd$), irrespective of the magnitude of the K_D, that Y varies from approximately 0 - 1.

In other words, Y varies from 0-1 when L varies from log K_D by +2. Hence, plots of Y vs log L for a series of binding reactions of increasingly higher K_D (lower affinity) would reveal a series of identical sigmoidal curves shifted progressively to the right, as shown below in Figure 3.

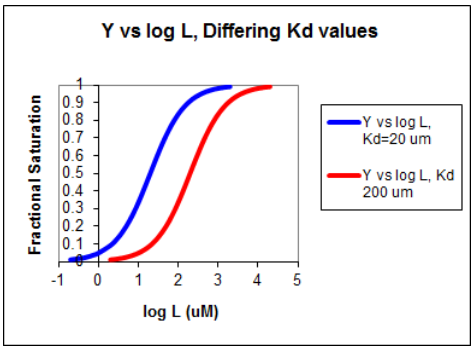


Figure 3: Plot of fractional saturation Y vs log L for two different K_D values.

The same would be true of v_0 vs S in the presence of different concentrations of a competitive inhibitor, for initial flux, Jo vs ligand outside, in the presence of a competitive inhibitor, or ML vs L (or Y vs L) in the presence of a competitive inhibitor.

In many ways plots of v_0 vs lnS are easier to visually interpret than plots of v_0 vs S . As noted for simple binding plots, textbook illustrations of hyperbolas are often misdrawn, showing curves that level off too quickly as a function of [S] as compared to plots of v_0 vs lnS, in which it is easy to see if saturation has been achieved. In addition, as the curves above show, multiple complete plots of v_0 vs lnS at varying fixed inhibitor concentrations or for variant enzyme forms (different isoforms, site-specific mutants) over a broad range of lnS can be made which facilitates comparisons of the experimental kinetics under these different conditions. This is especially true if Km values differ widely.

Now that you are more familiar with binding and enzyme kinetics curves, in the presence and absence of inhibitors, you should be able to apply the above analysis to inhibition curves where the binding or the initial velocity is plotted at varying competitive inhibitor concentrations at different fixed nonsaturating concentrations of ligand or substrate. Consider the activity of an enzyme. Let's say that at some reasonable concentration of substrate (not infinite), the enzyme is approximately 100% active. If a competitive inhibitor is added, the activity of the enzyme decreases until at saturating (infinite) I, no activity would remain. Graphs showing this are shown below in Figure 4.

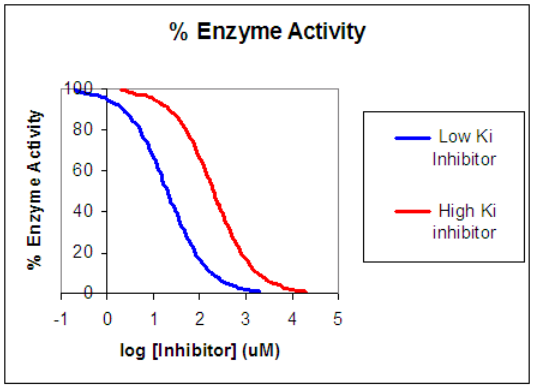


Figure 4: Inhibition of Enzyme Activity - % Activity vs log [Inhibitor]

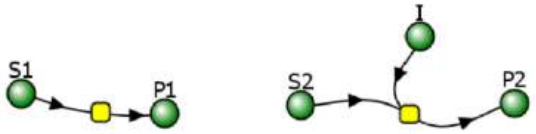
2.1: Progress Curves for Competitive Inhibition

In the previous section, we explored how important progress curve (Product vs time) analyses are in understanding both uncatalyzed and enzyme-catalyzed reactions. We are aware of no textbooks which cover progress curves for enzyme inhibition. Yet progress curves are what most investigators record and analyze to determine initial rates v_0 and to calculate V_M, K_M and inhibition constants, as described above. We will use Vcell to produce progress curves for reversibly inhibited enzyme-catalyzed reactions.

VCell MODEL

Competitive Inhibition with constant [I]:

No inhibition (left) and competitive inhibition (right)



Initial conditions for no inhibition

Name	Description	Global	
J	reaction rate	☐	$\frac{Vmax \cdot S1}{(Km + S1)}$
Km	Km (1/2 max)	☐	5.0
Vmax	max reaction rate	☐	10.0
S1	Species Concentration	☑	Variable

Initial conditions for competitive inhibition

Name	Description	Global	
J	reaction rate	☐	$\frac{Vmax \cdot S2}{(Km + S2 + \frac{Km \cdot I}{Ki})}$
Km	user defined	☐	5.0
Vmax	user defined	☐	10.0
Ki	user defined	☐	5.0
S2	Species Concentration	☑	Variable
I	Species Concentration	☑	Variable

I is fixed for each simulation (as it is not converted to a product) but can be changed in the simulation below.

Select Load [model name] below

Load NoInhib_CompInhib_MM_I

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed. For this model, select Vm, Km, Ki and I | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

The graphs from your initial run show the concentrations of S, P and I as a function of time for just the initial conditions shown above. In typical initial rate laboratory analyzes, of competitive inhibition, at least three sets of reactions are run with the same varying substrate concentrations and different fixed concentrations of inhibitor. In the analyses above, [I] is fixed at 5 uM.

Conduct a series of run at different values of I. Vary the K_I, the dissociation constant for the EI complex, as follows:

- I << K_I, the dissociation constant for the EI complex
- I >> K_I, the dissociation constant for the EI complex. Then download the data and determine the initial rate for each of the initial conditions.

Figure 5 shows an interactive iCn3D model of human low molecular weight phosphotyrosyl phosphatase bound to a competitive inhibitor (5PNT)

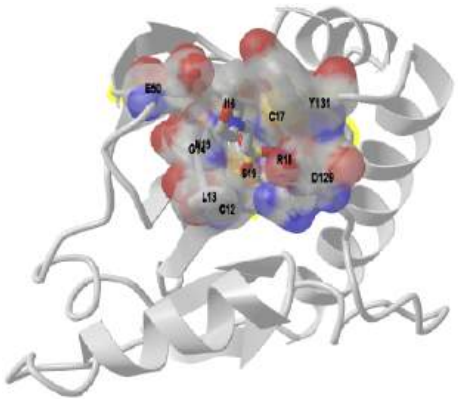


Figure 5: Human low molecular weight phosphotyrosyl phosphatase bound to a competitive inhibitor (5PNT). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...XsEacG2tixDDi9

The competitive inhibitor, the deprotonated form of 2-(N-morpholino)-ethanesulfonic acid (MES), is actually the conjugate base of the weak acid (pKa = 6.15) of a commonly used component of a buffered solution. It is shown in color sticks with the negatively charged sulfonate sitting at the bottom of the active site pocket. The amino acids comprising the active site binding pocket are shown as color sticks underneath the transparent colored surface of the binding pocket. The normal substrates for the enzyme are proteins phosphorylated on tyrosine side chains so the sulfonate is a mimic of the negatively charged phosphate group of the phosphoprotein target.

2.2: Two specials cases of competition inhibition

Product Inhibition

Let's look at an enzyme that converts reactant S to product P. Since P arises from S, they may have structural similarities. For example, what if GTP was the reactant and GDP was a product? If so, then P might also bind in the active site and inhibit the conversion of S to P. This is called **product inhibition**. It probably occurs in most enzymes, and when it does occur it will start bending downward the beginning part of the progress curve for P formation. If the product binds very tightly, it might cause a significant underestimation of the initial velocity (v_0) or flux (J_0) of the enzyme. Let's use Vcell to explore product inhibition. The model will explore two reactions:

- E + R ↔ ER → E + Q (no product inhibition)
- E + S ↔ ES → E + P (with product inhibition)

Note that the chemical equation above does not explicitly show the product P binding the enzyme to form an EP complex. An actual reaction diagram showing the inhibition of an enzyme by an inhibitor I and by the product P is shown in Figure 6 below.

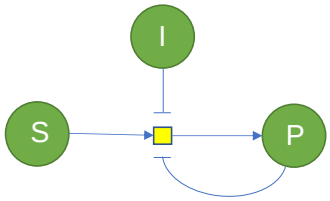


Figure 6: reaction diagram showing inhibition of an enzyme by an inhibitor I and by the product P

Vcell uses much simpler diagrams since it is most often used for modeling whole pathways or even entire cells. In the simpler Vcell reaction diagrams, the inhibitor is typically not shown since the inhibition is built into the equation for the enzyme, represented by the node or yellow square in the figure above.

Let's now explore product inhibition in Vcell. R and Q are the reactant and product, respectively, in the reaction *without* product inhibition. S and P are used for the reaction *with* product P inhibition.

VCell MODEL

Irreversible MM Kinetics - Without (left rx 1) and With (right, rx 2) Product Inhibition

Initial Conditions: No product inhibition

Name	Description	Global	
J	reaction rate	☐	$\frac{Vmax \cdot S1}{(Km + S1)}$
Km	Km (1/2 max)	☐	5.0
Vmax	max reaction rate	☐	10.0
S1	Species Concentration	☑	Variable

Initial Conditions: With product inhibition

Name	Description	Global	
J	reaction rate	☐	$\frac{Vmax \cdot S2}{(Km + S2 + \frac{Km \cdot P2}{Ki})}$
Km	user defined	☐	5.0
Vmax	user defined	☐	10.0
Ki	user defined	☐	5.0
S2	Species Concentration	☑	Variable
P2	Species Concentration	☑	Variable

Select Load [model name] below

Load ProdInhib_NoProdInhib

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

Inhibition by a competing substrate - the specificity constant

In the previous chapter, the specificity constant was defined as k_{cat}/K_M which we also described as the second-order rate constant associated with the bimolecular reaction of E and S when S << K_M. It also describes how good an enzyme is in differentiating between different substrates. If an enzyme encounters two different substrates, one can be considered to be a competitive inhibitor of the other. The following equation gives the ratio of initial velocities for two competing substrates at the same concentration is equal to the ratio of their k_{cat}/K_M values.

$$\frac{v_A}{v_B} = \frac{\frac{k_{cat\,A}}{K_A}A}{\frac{k_{cat\,B}}{K_B}B} \tag{2}$$

A derivation of the specificity constant for an enzyme with competing substrates

Here it is!

Derivation

$$v_A = \frac{V_A A}{K_A\left(1+\frac{B}{K_B}\right)+A} \quad v_B = \frac{V_B B}{K_B\left(1+\frac{A}{K_A}\right)+B} \tag{3}$$

$$\frac{v_A}{v_B} = \frac{\frac{V_A A}{K_A\left(1+\frac{B}{K_B}\right)+A}}{\frac{V_B B}{K_B\left(1+\frac{A}{K_A}\right)+B}} = \frac{\frac{V_A A}{K_A+\frac{K_A B}{K_B}+A}}{\frac{V_B B}{K_B+\frac{K_B A}{K_A}+B}} \tag{4}$$

Now in the above equation:

multiple the top half of the right-hand expression by

$$\frac{\frac{1}{K_A}}{\frac{1}{K_A}} \tag{5}$$

multiple the bottom half of the right-hand expression by

$$\frac{\frac{1}{K_B}}{\frac{1}{K_B}} \tag{6}$$

replace V_A with $k_{catA}E_0$ and V_B with $k_{catB}E_0$

This gives the following expression for v_A/v_B:

$$\frac{v_A}{v_B} = \frac{\frac{k_{catA}}{K_A}A}{\frac{k_{catB}}{K_B}B} \tag{7}$$

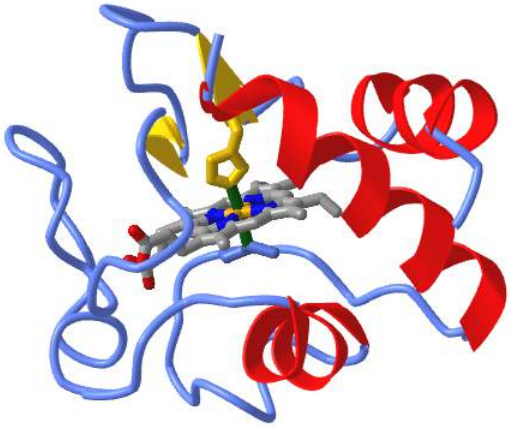
Figure x: Add caption

[ADD VCELL/SBML SIMULATION - REUSE]

*your text before and after insert as needed

3: New Heading

Figure x is an interactive iCn3D model of **[2B4Z cytochrome c]**



INSERT PNG ( ) OF YOUR iCn3D MODEL

Figure 6: [2B4Z]. Click the image for a popup or use this external link: [https://bio.libretexts.org/Learning_..._iCn3Ds_Helena]. (Copyright; author via source). iCn3D model made by [Helena Prieto]

Inserting an iCn3D model into a FOB Chapter section

(If you need a refresher, here is a link to an iCn3D workshop tutorial presented by BioMolViz at the BMB 2023 in Seattle, WA on March 25, 2023.)

A. Creating an iCn3D model in FOB

1. Open iCn3D, input a PDB code, and do the following:

   - **Color, Secondary, Sheets in Yellow**
   - **Style, Background, Transparent**
   - **File, Share Link**
   - Copy the Original URL with commands

2. Follow these instructions but use your code (open a new Word document for your code).

Here is a sample code for 1XWW:

> https://www.ncbi.nlm.nih.gov/Structu...0&command=load mmdb 1xww | parameters &mmdbid=1xww&bu=1; set thickness | linerad 0.1 | coilrad 0.3 | stickrad 0.4 | crosslinkrad 0.4 | tracerad 0.1 | ribbonthick 0.2 | proteinwidth 1.3 | nucleotidewidth 0.8 | ballscale 0.3; set background transparent; color secondary structure yellow|||{"factor":"1.000","mouseChange":{"x":"0.000","y":"0.000"},"quaternion":{"_x":"-0.006542","_y":"0.8578","_z":"-0.04771","_w":"0.5117"}}

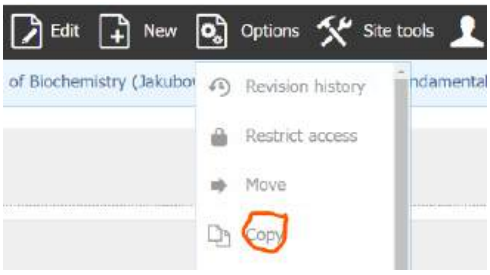
2. Copy all of the code ***BEFORE*** the ||| and paste it in a Word document. Then add the 3 small **red codes** in the example below, and add them in the corresponding positions in your sample code. It will be helpful to color the newly added code red to make sure it is correct. It will look like this but your code will replace the back code below.

> **template('EmbediCn3D/iCn3D',{config: '{"fullURL":** "https://www.ncbi.nlm.nih.gov/Structu...30430&v=3.25.0**&closepopup=1**&command=load mmdb 1xww | parameters &mmdbid=1xww&bu=1; set thickness | linerad 0.1 | coilrad 0.3 | stickrad 0.4 | crosslinkrad 0.4 | tracerad 0.1 | ribbonthick 0.2 | proteinwidth 1.3 | nucleotidewidth 0.8 | ballscale 0.3; set background transparent; color secondary structure yellow**"}'})**

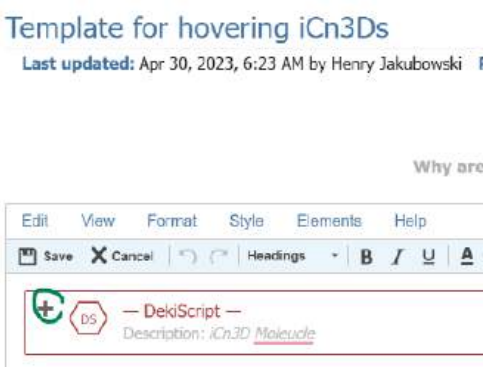
3. Now go to this page in a new window to create your own iCn3D page for FOB

4. Open the file **Template for Your Own Separate iCn3D Page**

5. **You must make a copy of this file**. From the top menu bar click **Options, Copy,** and give the file a name derived from the PDB web page. For example Human Low Molecular Weight Protein Phosphatase (1XWW). Don't overnight the original file with your new name.



6. Then from the top menu bar for your new file select **Edit**. Click the + button (green circle) on the DekiScript box.



7. Paste your entire modified URL code from your Word document (with the 3 inserted red codes) into the DekiScript box.

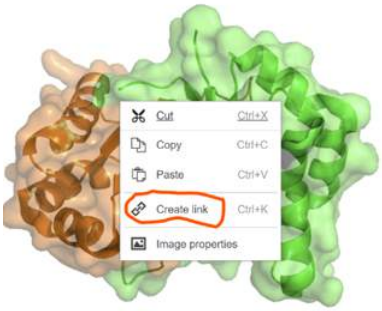
8. Choose the Save just above the DekiScript box and you should see a small version of the molecule in a workable iCn3D window.

B. Inserting your iCn3D model into your Chapter 33 Section

1. Open your Chapter 33 section and click **Edit**

2. Place your cursor on a new line where you would like to place the model to appear.

3. Go to this page in a new window, select Edit in the menu bar, and copy the material between the two horizontal lines into the desired location in your Chapter 33 section.

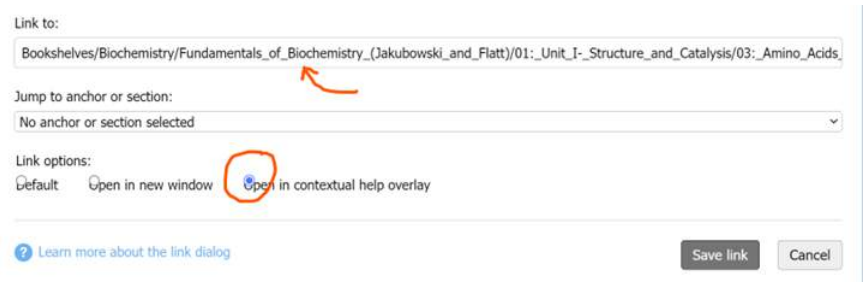


4. Take a screen snip showing the molecule in the iCn3D modeling window. Save it to your computer. Then insert the image in you Chapter 33 section where the text reads: INSERT PNG (just a screen snip) OF YOUR iCn3D MODEL. Use the top menu bar picture icon to select the file or drag the file from your computer into your section page.

5. Complete the captions for the image and make sure to include the PDB ID in your text as well as the iCn3D short link.

6. To make a hovering iCn3D model with a full menu within your chapter section (and not just from the external iCn3D short link), right-click the image you inserted in Step 4, and select **Create Link** (see image below).

7. Delete the link shown (red arrow) and replace it with the URL from the LibreText iCn3D file you just made in Part A above. Make sure that you choose **Open in contextual help overlay.**



8. Save the link.

Inserting an interactive mathematic graph into FOB

1. Navigate to this page in a new window
2. Select and view the graph of interest
3. Navigate to your Chapter 33 section and select Edit from the top menu bar
4. Move the cursor to the location you wish to insert the graph
5. From the top menu bar select **Elements, Content Reuse**
6. Navigate through the folder tree to get to the page with the interactive mathematical models:

    Home

        Learning Objects

            Visualization and Simulations

                Interactive Figures

                    Interactive Biochemistry Graphs

7. Click the file name for the model of interest to insert.
8. Save the page.

Inserting a mathematical simulation (SBML) into a FOB

1. Navigate to this page in a new window
2. Select and view the graph of interest
3. Navigate to your Chapter 33 section and select Edit from the top menu bar
4. Move the cursor to the location you wish to insert the graph
5. From the top menu bar select **Elements, Content Reuse**
6. Navigate through the folder tree to get to the page with the interactive mathematical models:

    Home

        Learning Objects

            Visualization and Simulations

                Progress Curve Analysis

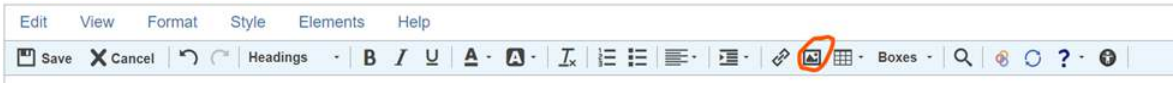
                    SBML Computational Models

7. Click the file name for the model of interest to insert.
8. Save the page.

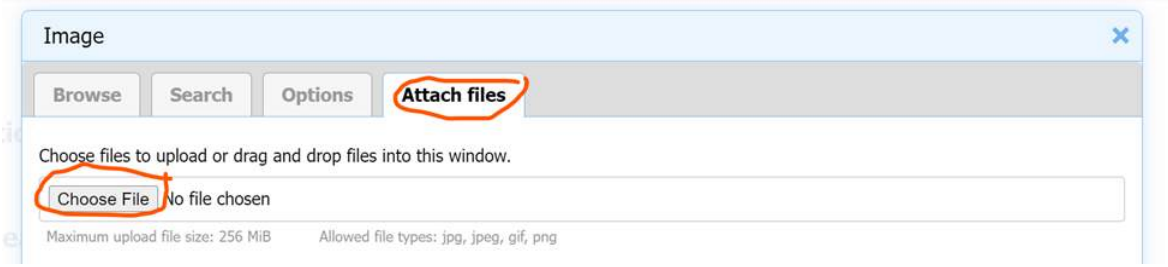
Insert an Image into FOB

Inserting an image into FOB

1. Do a Google search for an image with Creative Commons permission for reuse. The optimal permission category is CC BY.
2. Make a copy and save it to your computer.
3. Open your Chapter 33.x Libretext file and select **Edit** in the top menu bar
4. Navigate to ADD IMAGE, remove these words but leave the cursor there
5. From the Editing menu bar, select **Elements, Templates,** and from the drop-down Template:FigureCenterCenter.
6. Click on the Copy/Paste Placeholder image and **delete** it.
7. Recenter the cursor centered above the automatic figure legend, click the Image icon from the menu bar.



8. Select the Attach Files tab, followed by Choose File, and navigate to the image you wish to upload.



9. Select Save Image, and the image will appear centered above the caption.
10. Change the placeholder text in the caption to one of your choosing.

Insert a question with hidden answer and floating hint

Example 1

Which elements can form hydrogen bonds?

Solution

fluorine, oxygen, or nitrogen

Exercise 1

Can phosphorus form hydrogen bonds?

Here is a hint if you need one!

Answer

No. Phosphorus and hydrogen have almost equal values of electronegativity.

Box Question with floating Hint

1. On your Chapter 33 page, click Edit.
2. From the menu bar that appears, select Elements, Templates
3. From the dropdown menu select Box: Exercise. The following box will appear

Exercise

Example 1

Add example text here. Which elements can form hydrogen bonds?

Solution

Add example text here. **fluorine, oxygen, or nitrogen**

1

Add exercises text here. Can phosphorus form hydrogen bonds?

Answer

No. Phosphorus and hydrogen have equal values of electronegativity.

4. Edit the Box Exercise as shown below by clicking in the text areas and replacing the text with your own. You can add an image by scrolling to the bottom of this page and selecting **Attach a file**.

A question 1

Which does **NOT** describe sodium dodecyl sulfate (SDS). SDS ....

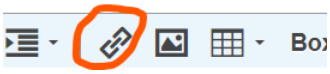
1. readily forms bilayers
2. readily form micelles
3. is a single-chain amphiphile
4. has (a) unsaturated acyl chain(s)

Here is a hint if you need one!

Answer

It does not readily form a bilayer since it is a single-chain amphiphile. Instead, it forms micelles.

5. The hint in this example is a file that has been uploaded (see the list of attached files below). Right-click on the SDS.png file below and choose **Copy the link address**. Then double-click the word "hint" in the above box to highlight it, choose the link icon from the top menu bar, and paste the link into the Link To box after replacing the default link in the box. Under link options, select **Open in Contextual Help Overlay** to get the hint to float above the page.



KP Procko Test

Authored by KP. Last update: 5/4/23

Date of origin

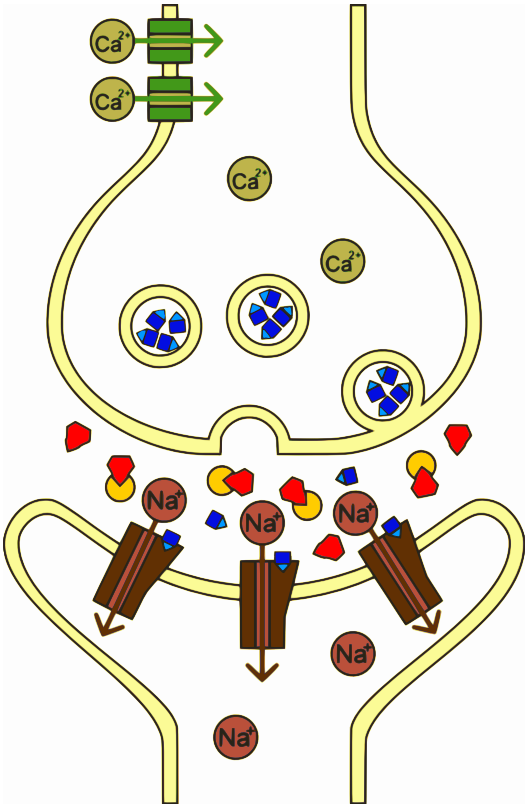
1: Introduction

[add content]

2: New Heading

[ADD TEXT]

your text



your text and image

[ADD iCn3D Model]

Figure x is an interactive iCn3D model of mammalian respirasome (5GPN)

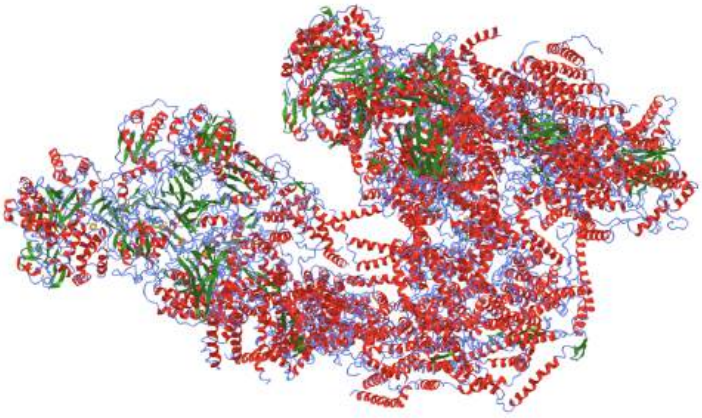
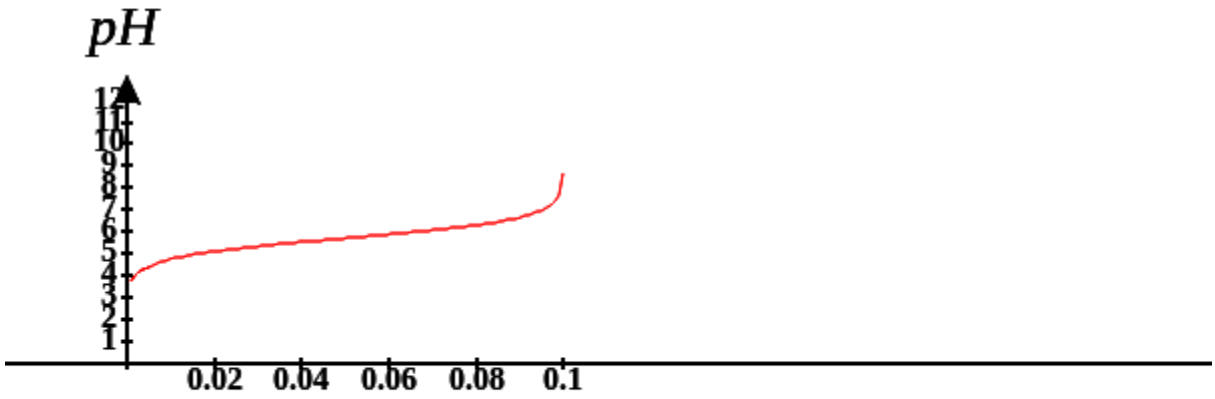


Figure 6: [INSERT THE PDB TITLE as above (INSERT PDB CODE)]. Click the image for a popup or use this external link: [INSERT Lifelong short URL from File, Share Link in iCn3D)]. (Copyright; author via source). iCn3D model made by [YOUR NAME]

your text and model

[ADD MATHEMATIC GRAPH - REUSE]



your text and graph

? Exercise 1

What is the pKa of acetic acid?

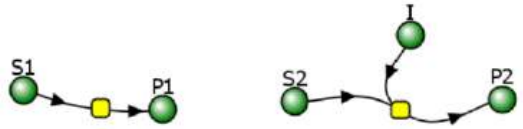
Answer

4.7

[ADD VCELL SIMULATION - REUSE]

VCell MODEL

No inhibition (left) and Uncompetitive Inhibition (right)



Note that the Vcell reaction diagram is the same as for competitive and uncompetive inhibition. It doesn't explicitly show that the mixed inhibitor binds to both free and substrate-bound enzymes. Those interactions are addressed in the mathematical equations for mixed inhibition.

Initial values No Inhibition

Name	Description	Global	
J	reaction rate	☐	$\frac{Vmax \cdot S1}{(Km + S1)}$
Km	Km (1/2 max)	☐	5.0
Vmax	max reaction rate	☐	10.0
S1	Species Concentration	☑	Variable

Initial values With Uncompetitive Inhibitor

Name	Description	Global	
J	reaction rate	☐	$\frac{Vmax \cdot S2}{(Km \cdot (1.0 + \frac{I}{Kc}) + S2 \cdot (1.0 + \frac{I}{Ki}))}$
Km	user defined	☐	5.0
Vmax	user defined	☐	10.0
Kc	user defined	☐	1.0
Ki	user defined	☐	1.0
S2	Species Concentration	☑	Variable
I	Species Concentration	☑	Variable

I is fixed for each simulation (as it is not converted to a product) but can be changed in the simulation below.

Select Load [model name] below

Load NoInhib_MIXEDInhib_MM

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

---

your tex

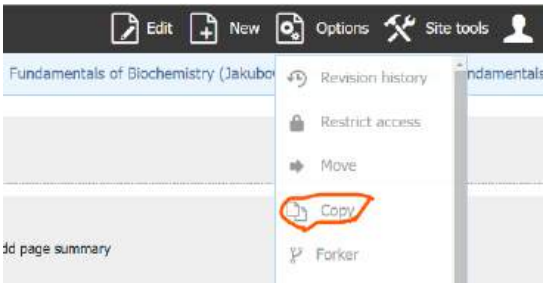
3: New Heading

Add what you want

Make your own Chapter 33.x Section

Make Your Own Chapter 33.x Section

1. Navigate to this page in a new window
2. Open the file named **Template to Copy**
3. On the top menu bar select Options, Copy



4. Add a page title with this format: FirstName LastName Short Descriptive Title and then chose Copy Page.

5. Editing a webpage is generally intuitive. Once you click the Edit button on the top left-hand side of the page, editing icons will appear, similar to a word processing program. The icons on the top menu bar next to the indent icon allow you to: create a link, add a picture directly, add a table, and search (magnifying glass).

Make Your Own Customized FOB: Assemble and Remix a Custom Book (Short Version)

A. Introduction

Remixes are texts created from existing OER content. Constructing Remixes on the LibreTexts platform is facilitated by the OER Remixer tool. The title of the Remix often starts with the campus acronym (e.g., the Chemistry 110A remix at the University of California is "UCD: Chem 110A Introductory Quantum Mechanics")

If using the OER Remixer to build a large textbook, it is advisable to construct a Remixing Map in a spreadsheet including to organize the sections you'll include. This content will be an effective Table of Contents. For this exercise, we'll make a remix with just 3 chapter sections from 3 different chapters so no spreadsheet is needed.

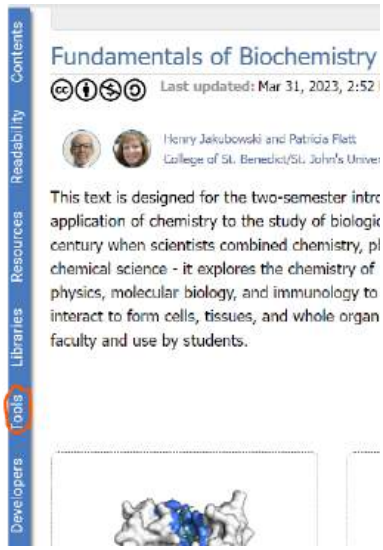
We'll use these chapter sections:

- Vol 1 – Chapter 6.5: Enzymatic Reaction Mechanisms (written by Henry and KP)
- Vol 2 – Chapter 14.5: Metabolism and Signaling: The Steady State, Adaptation and Homeostasis
- Vol 4 – Chapter 32.12: A Warmer World – Temperature Effects on Proteins

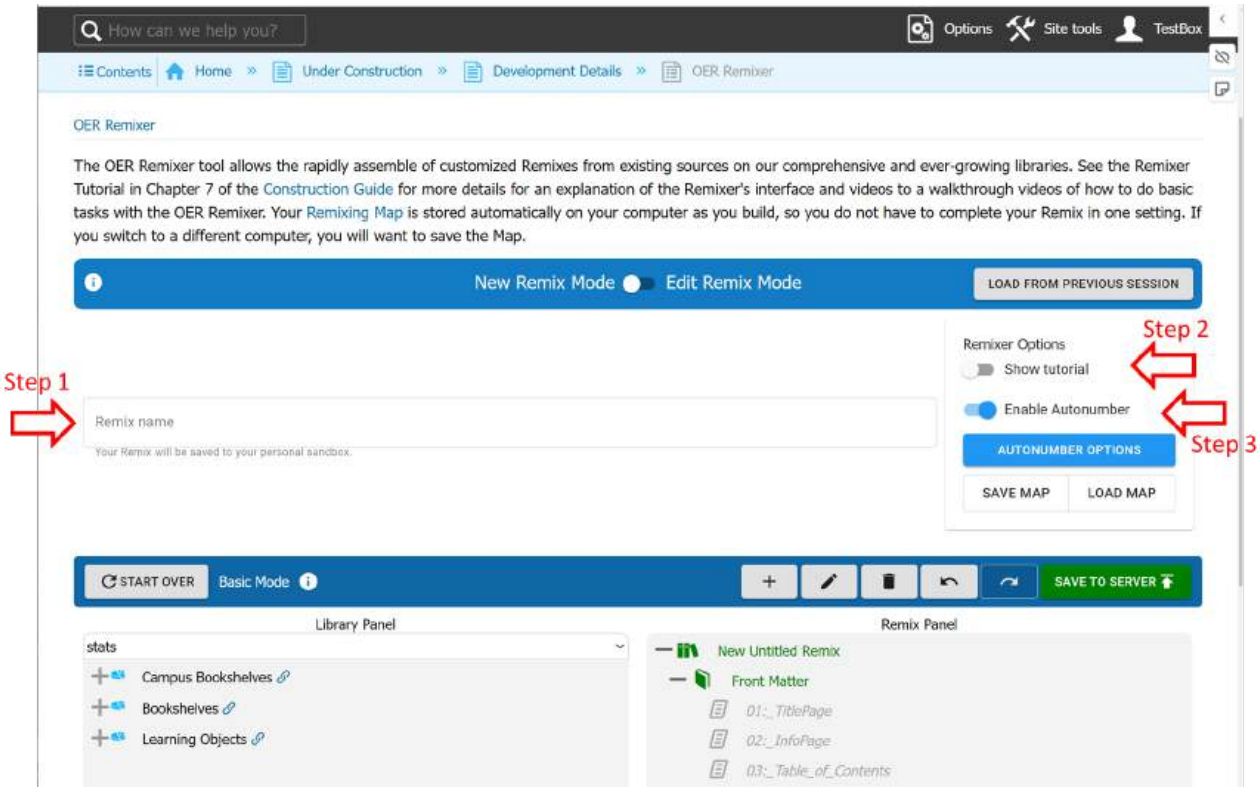
(Note: you could add content from any LibreText book as well.)

Instructions

1. Go to the main FOB page
2. An author with permission can see a blue vertical icon bar on the left side of the page.

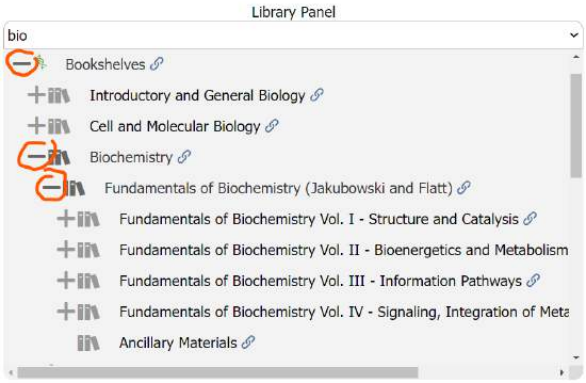


3. Select **Tools** and from the right side menu choose **OER Remixer**.
4. Complete the step indicated in the figure below



- ***Step 1:*** *Enter the name you want for your custom book (typically your name and the name of the class) in the box labeled "LibreText name" (Step 1 in Figure 7.3.57.3.5). Ex: Henry Jakubowski_TrialRemixFOB_1*
- ***For Step 2 and Step 3:*** *Accept the default*

5. In the left "Library Panel": Click the + by the bookshelves and continue to see Vol I, II, and IV (see image right). Keep expanding the list until you see Chapter 6 (Vol 1), then Chapter 14 (Vol 2), and finally Chapter 32 (Vol 4)

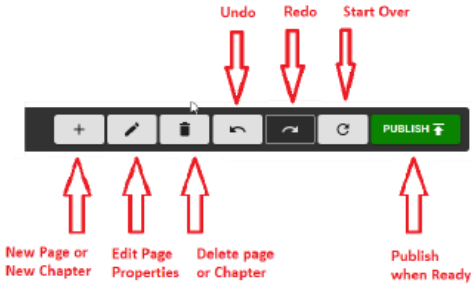


6. Insert the entire Chapter 6 by selecting it and dragging and dropping it to the right panel. Position it by moving it to the correct position (you will see a little blue line indicating the position)

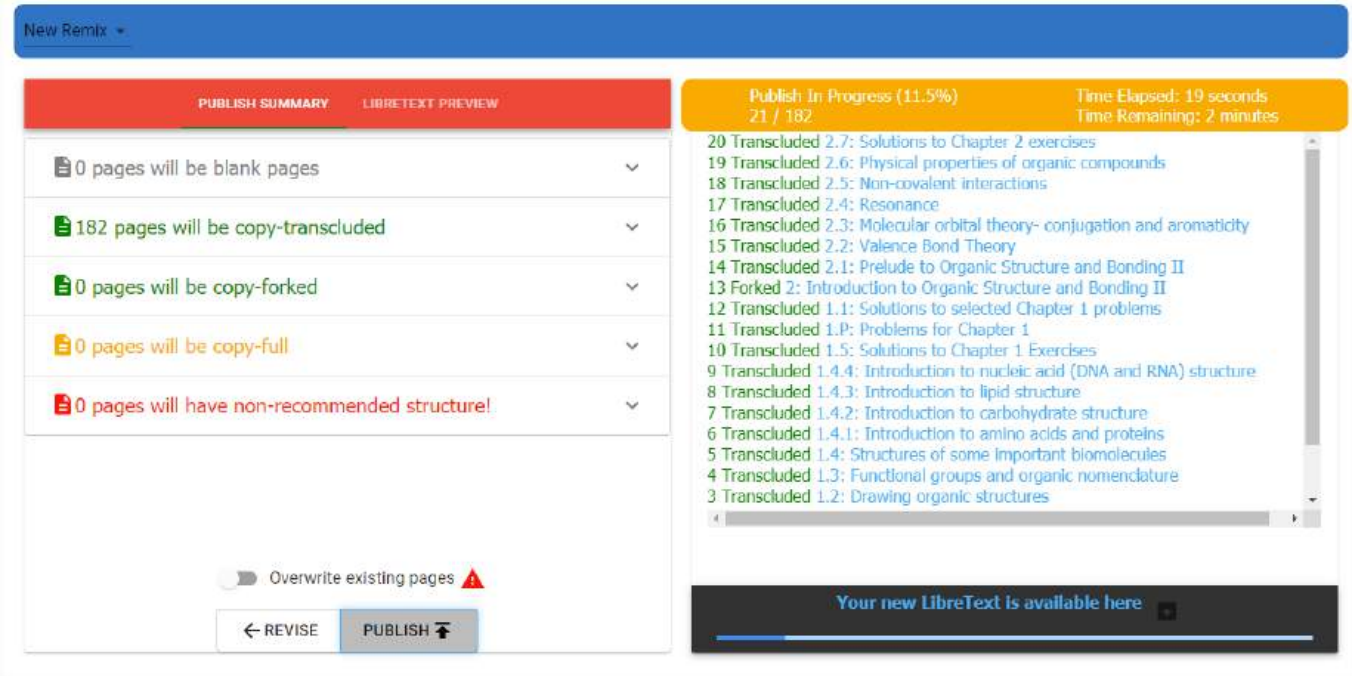
7. Repeat and move the entire Chapt 14 and Chapter 32.

8. Now expand the moved Chapters in the panel. Select the section you do **NOT** want and click the delete icon (recycle bin) in the menu bar in the right panel. Continue deleting until each Chapter has just the desired section.

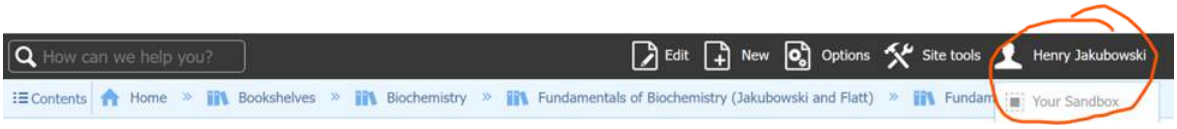
9. Select the Publish button (see image below)



10. In the next window select Publish again and wait until the remix is complete. Click the "Your new LibreText is available here to see your custom book (which is your Sandbox).

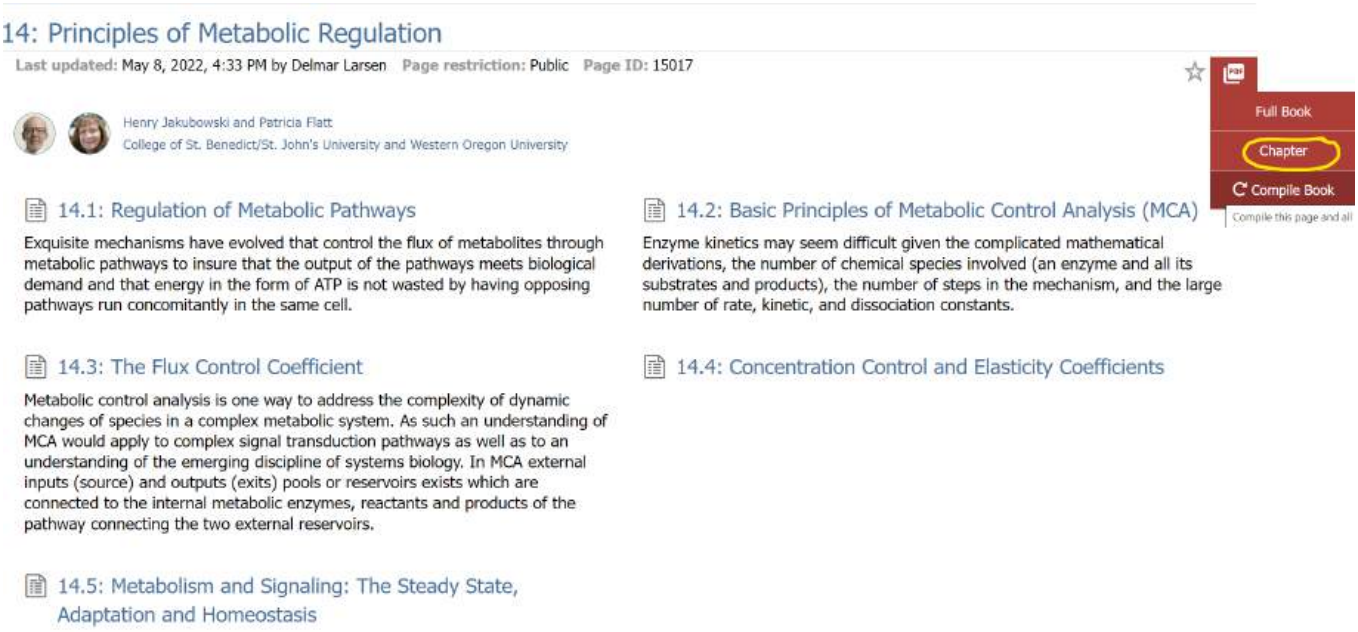


11. Your sandbox can be seen by clicking in your name on any LibreText page when you have logged in as shown below



B. Making a PDF

You can easily make a PDF of a chapter. For example, navigate to Chapter 14 and select PDF then Chapter and follow prompts.



Pallavi- Test

Authored by [YOUR NAME]. Last update: [FILL IN DATE]

Date of origin

1: Introduction

[ADD CONTENT]

2: New Heading

[ADD TEXT]

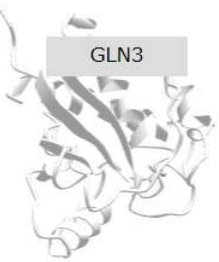
[ADD IMAGE] (saved to your computer and uploaded with picture icon from top menu bar or drag image file to location (required from svg image)

*

Figure x: [Add caption]

*Center Picture and Caption together using top menu ba

Figure x is an interactive iCn3D model of [GLN3]



INSERT PNG (just a screen snip) OF YOUR iCn3D MODEL

Figure 6: [INSERT THE PDB TITLE as above (INSERT PDB CODE)]. Click the image for a popup or use this external link: [INSERT Lifelong short URL from File, Share Link in iCn3D)]. (Copyright; author via source). iCn3D model made by [YOUR NAME]

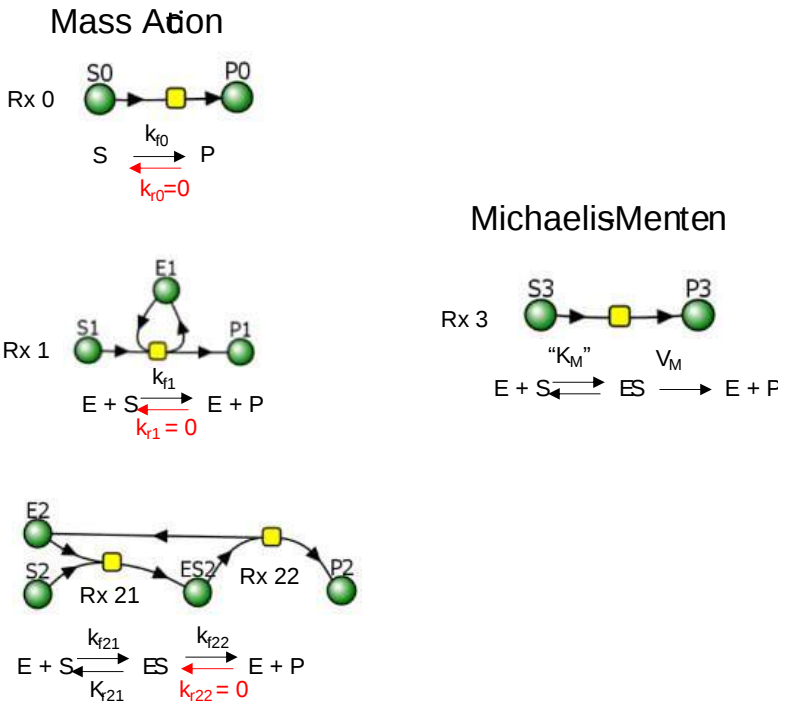
[ADD MATHEMATIC GRAPH - REUSE]

*your text and graph

Figure x: Add caption

[ADD VCELL/SBML SIMULATION - REUSE]

VCell MODEL



Initial Parameters

Reaction	Parameter	Value
r_0	k_{f0}	
	k_{r0}	
r_1	k_{f1}	
	k_{f1}	
r_2	k_{f21}	
	k_{f21}	
	k_{f22}	
	k_{f22}	
r_3	V_M	
	K_M	

Select Load [model name] below

Load Compare3MassAct_1MMKineticStoPIrrev

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

3: New Heading

Add what you want

Pam Mertz Chapter 33 Test

Authored by [YOUR NAME]. Last update: [FILL IN DATE]

Date of origin

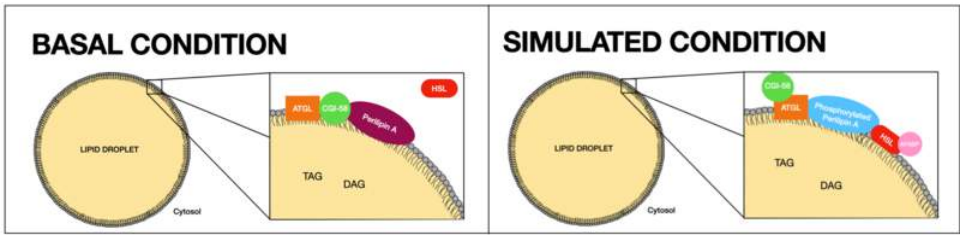
1: Introduction

Perilipin is a protein associated with lipid droplets that plays a key role in the regulated breakdown of triacylglycerols.

2: New Heading

[ADD TEXT]

your text



[ADD IMAGE] (saved to your computer and uploaded with picture icon from top menu bar or drag image file to location (required from svg image)

*Use the following under your picture:

Figure x: [Add caption]

*Center Picture and Caption together using top menu bar

Figure x is an interactive iCn3D model of Low Molecular Weight Phosphotyrosyl Phosphatase, 1xww

Figure 66: **[**INSERT THE PDB TITLE as above (INSERT PDB CODE)**]**. Click the image for a popup or use this external link: **https://structure.ncbi.nlm.nih.gov/i...7&t=1XWW(MMDB) in iCn3D**. (Copyright; author via source). iCn3D model made by Pam Mertz

[ADD MATHEMATIC GRAPH - REUSE]

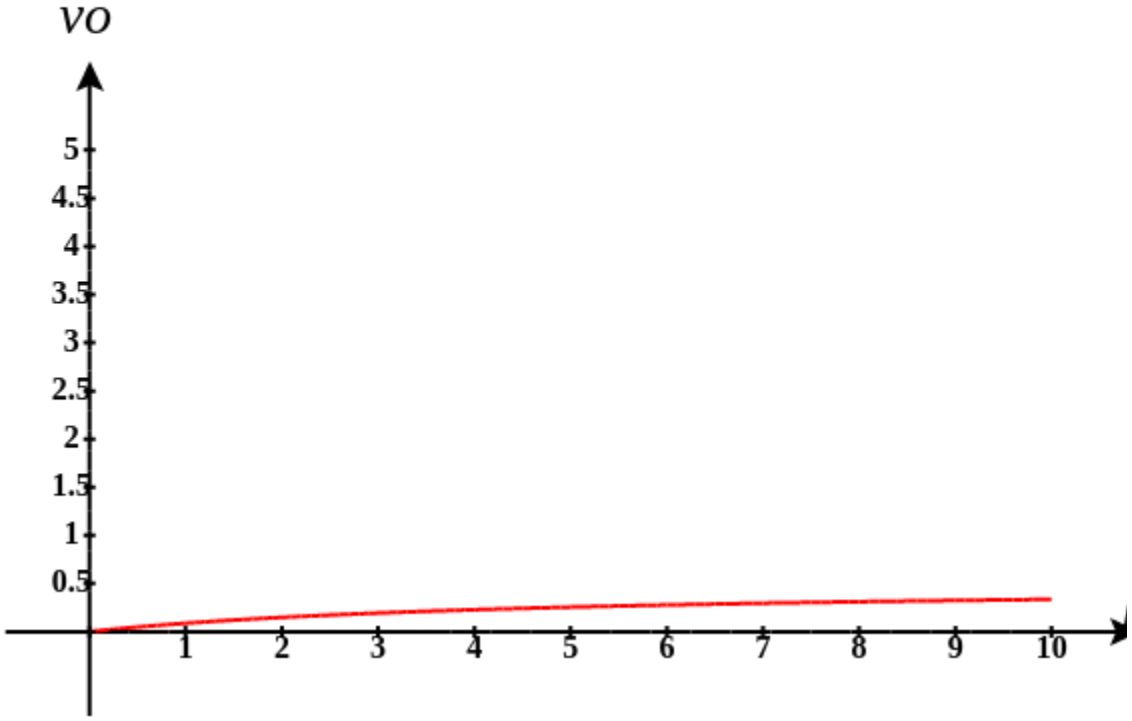
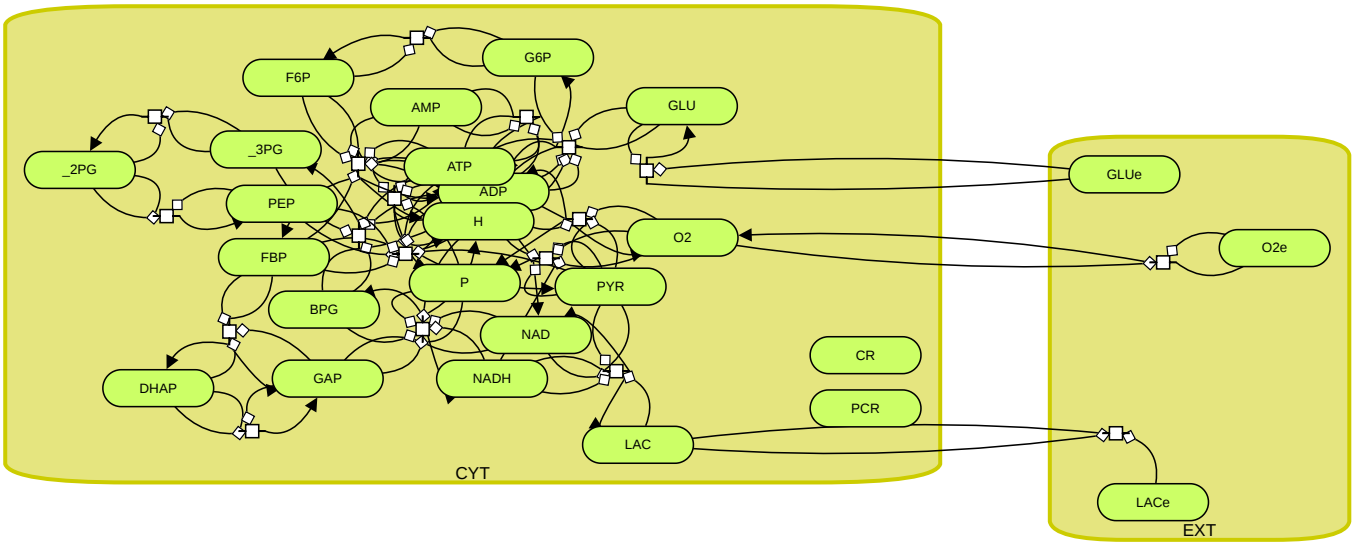


Figure x: Add caption

VCell MODEL

AerobicGlycolysis

Quantitative determinants of aerobic glycolysis identify flux through the enzyme GAPDH as a limiting step. .Shestov AA, Liu X, Ser Z, Cluntun AA, Hung YP, Huang L, Kim D, Le A, Yellen G, Albeck JG, Locasale JW. *eLife* , 7/ 2014 , Volume 3 , PubMed ID: 25009227. Biomodel MODEL1504010000



Select Load [model name] below

Load AerobicGlycolysis

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

---

*your text before and after insert as needed

3: New Heading

Add what you want

Rebecca Roberts Chapter 33 testing

Authored by Rebecca Roberts. Last update: 6/5/2023

Date of origin

1: Introduction

Sometimes proteins need to bind to something specifically, but sometimes a protein needs to recognize a variety of ligands. How does this happen?

2: New Heading

[ADD TEXT]

*Use the following under your picture:

Figure x: Major Histocompatability Complex Class II (MHC-II). The figure

*Center Picture and Caption together using top menu bar

Figure 3 is an interactive iCn3D model of **3DS8, a protein of unknown function**

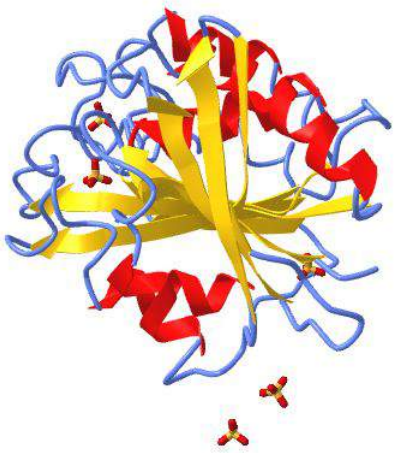


Figure 66: **[**INSERT THE PDB TITLE as above (INSERT PDB CODE)**]**. Click the image for a popup or use this external link: **[**INSERT Lifelong short URL from File, Share Link in iCn3D)**]**. (Copyright; author via source). iCn3D model made by **[**YOUR NAME**]**

[ADD MATHEMATIC GRAPH - REUSE]

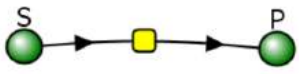
*your text and graph

Figure x: This is an interactive model for Michaelis-Menten kinetics

[ADD VCELL/SBML SIMULATION - REUSE]

VCell MODEL

Reversible reaction E + S ↔ ES ↔ EP ↔ E + P



Initial values

Name	Description	Global	
J	reaction rate	☐	$\frac{(\frac{S \cdot VmaxFwd}{KmFwd} - \frac{P \cdot VmaxRev}{KmRev})}{(1.0 + \frac{S}{KmFwd} + \frac{P}{KmRev})}$
KmFwd	Km forward	☐	10.0
VmaxFwd	max forward rate	☐	10.0
KmRev	Km reverese	☐	20.0
VmaxRev	max reverse rate	☐	5.0
S	Species Concentration	☑	Variable
P	Species Concentration	☑	Variable

Select Load [model name] below

Load EnzymeRevIsolated

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

---

*your text before and after insert as needed

3: New Heading

Add what you want

Rico Acevedo Testy

Authored by [YOUR NAME]. Last update: [FILL IN DATE]

Date of origin

1: Introduction

[add content]

2: New Heading

[ADD TEXT]

your text

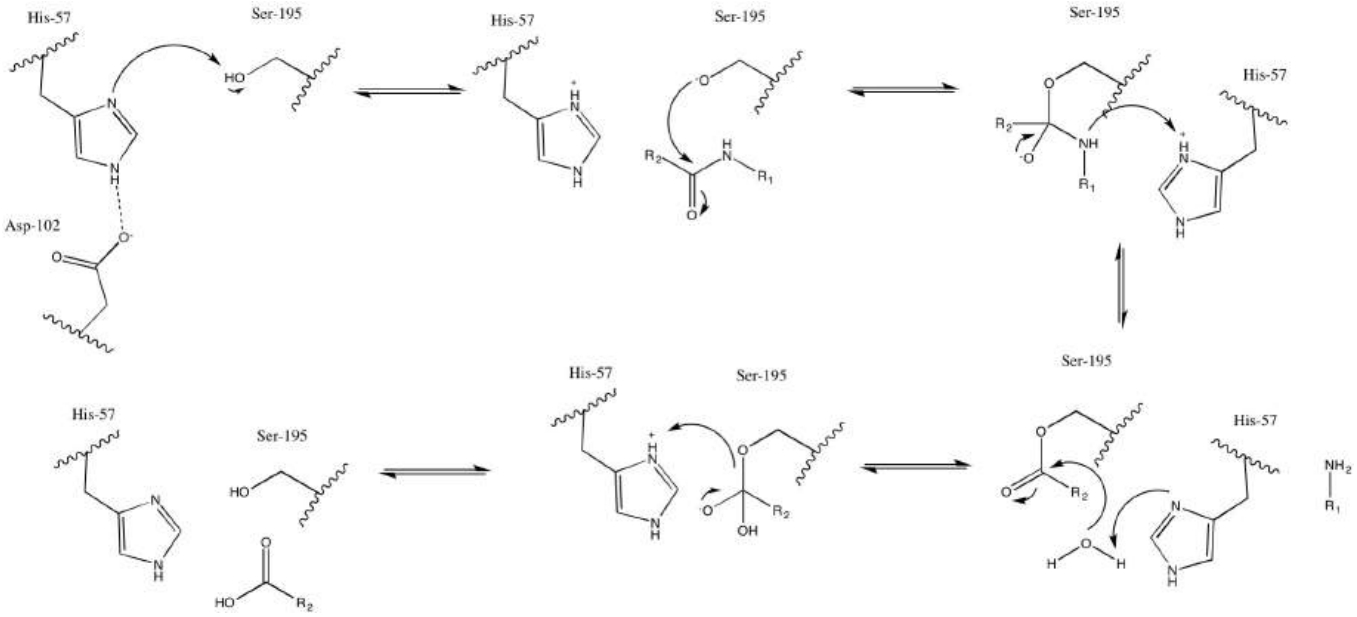
[ADD IMAGE] (saved to your computer and uploaded with picture icon from top menu bar or drag image file to location (required from svg image)

*Use the following under your picture:

Figure x: [Add caption]

*Center Picture and Caption together using top menu bar

[ADD iCn3D Model]



[ADD MATHEMATIC GRAPH - REUSE]

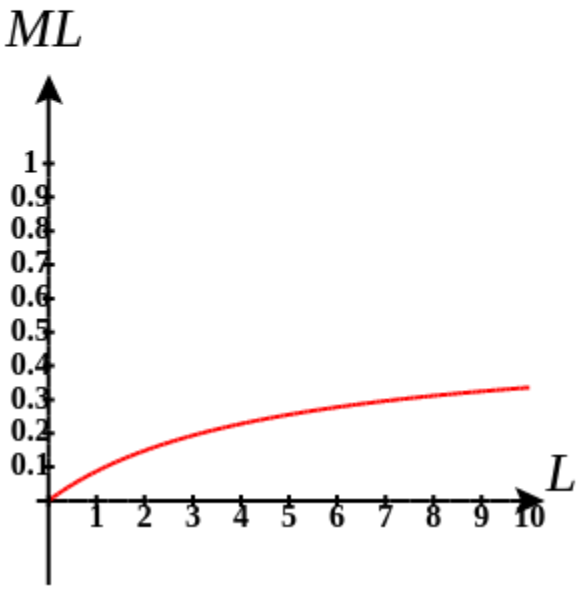
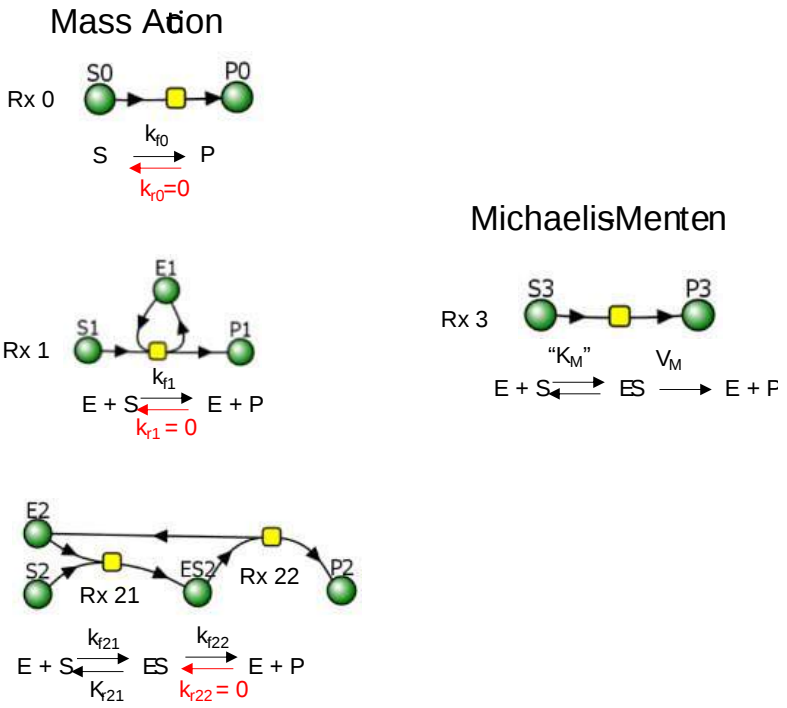


Figure x: Add caption

VCell MODEL



Initial Parameters

Reaction	Parameter	Value
r_0	k_{f0}	
	k_{r0}	
r_1	k_{f1}	
	k_{f1}	
r_2	k_{f21}	
	k_{f21}	
	k_{f22}	
	k_{f22}	
r_3	V_M	
	K_M	

Select Load [model name] below

Load Compare3MassAct_1MMKineticStoPIrrev

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

[ADD VCELL/SBML SIMULATION - REUSE]

*your text before and after insert as needed

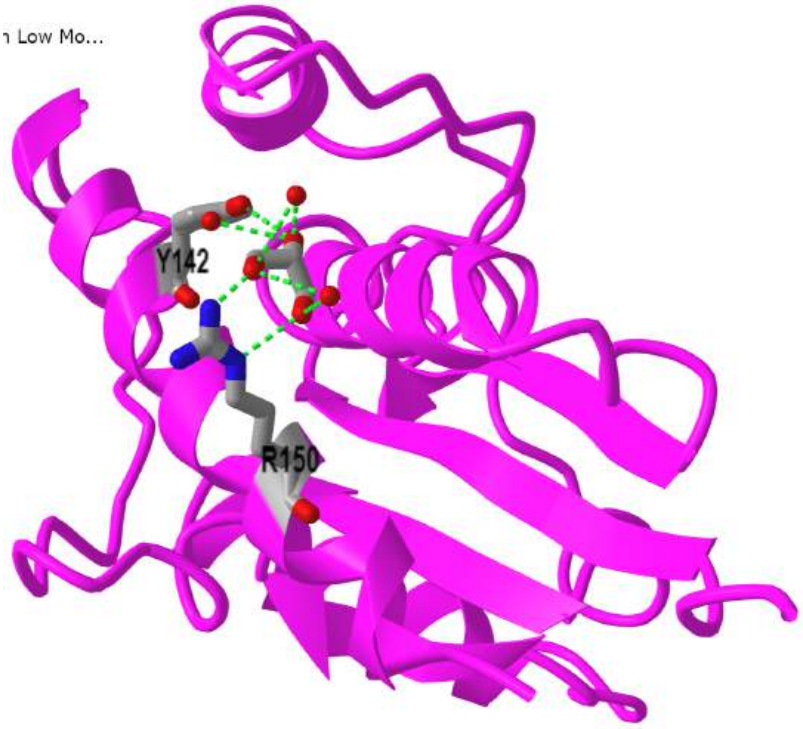
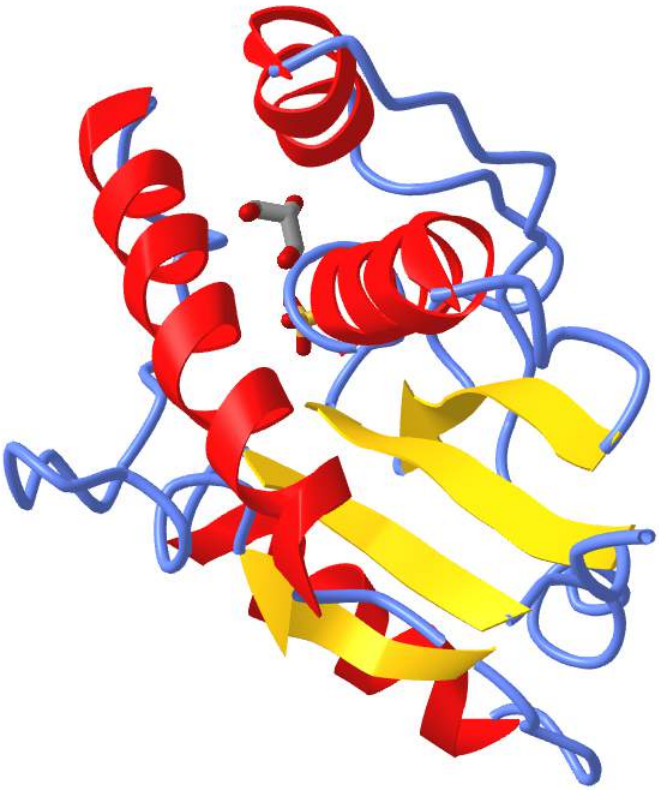


Figure x[illegible] is an interactive iCn3D model of **Low molecular weight 1XWW**



I L

NCBI iCn3D Figure 66: **Low molecular weight 1XWW**. Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...ssdBPpCB3pWSV7 (Copyright; author via source). iCn3D model made by **[**YOUR NAME**]**

3: New Heading

? Question 1 1

Which does **NOT** describe sodium dodecyl sulfate (SDS). SDS ....

1. readily forms bilayers
2. readily form micelles
3. is a single-chain amphiphile
4. has (a) unsaturated acyl chain(s)

Here is a hint if you need one!

Answer

1. readily forms bilayers.

Samantha Wilner Ch 33 Test

Authored by Samantha Wilner. Last update: June 5, 2023

Date of origin

1: Introduction

[ADD CONTENT]

2: New Heading

[ADD TEXT]

[ADD IMAGE] (saved to your computer and uploaded with picture icon from top menu bar or drag image file to location (required from svg image)

*Use the following under your picture:

Figure x: [Add caption]

*Center Picture and Caption together using top menu bar

Figure 1 is an interactive iCn3D model of a hydrolase (3H04)

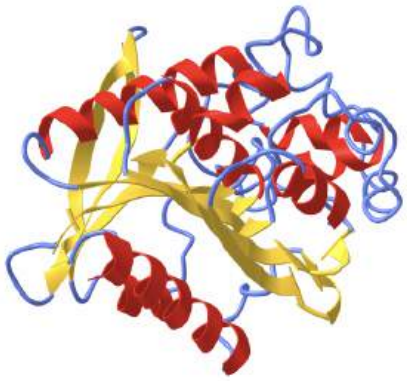


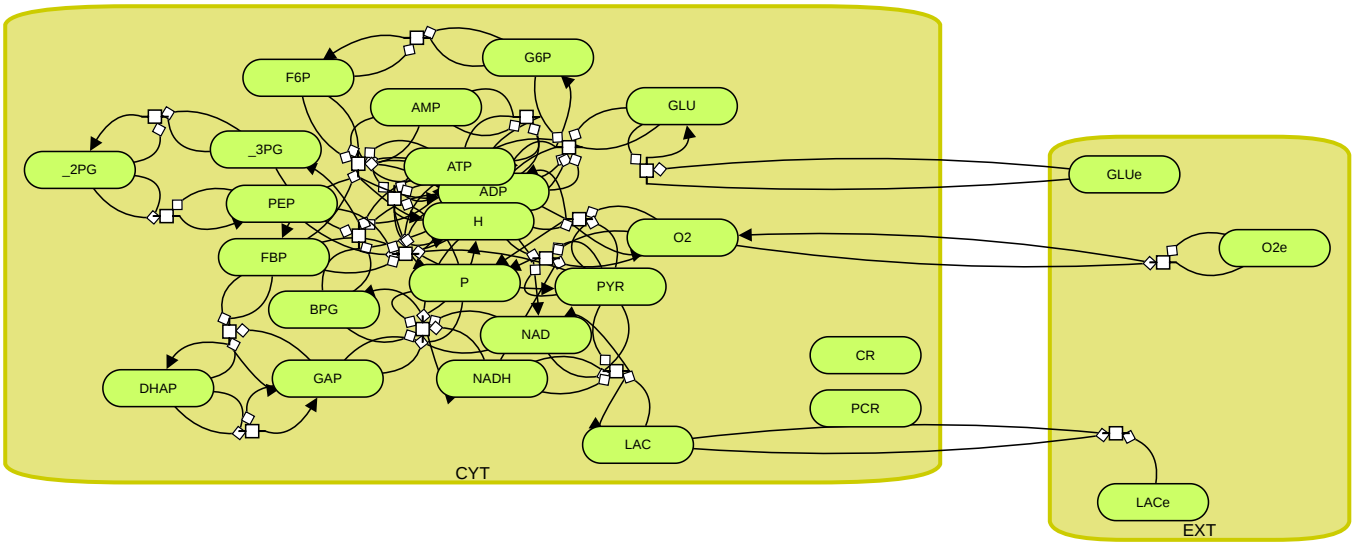
Figure 1: 3H04. Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...Eng4pm1AcyxZx8. (Copyright; author via source). iCn3D model made by Samantha Wilner

Figure 2: First and Second Order Reactions

VCell MODEL

AerobicGlycolysis

Quantitative determinants of aerobic glycolysis identify flux through the enzyme GAPDH as a limiting step. .Shestov AA, Liu X, Ser Z, Cluntun AA, Hung YP, Huang L, Kim D, Le A, Yellen G, Albeck JG, Locasale JW. *eLife* , 7/ 2014 , Volume 3 , PubMed ID: 25009227. Biomodel MODEL1504010000



Select Load [model name] below

Load AerobicGlycolysis

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

Figure 3: Aerobic Glycolysis

3: New Heading

Add what you want

Subhasish-test

Goal

After completing this how-to you will have ...

First Step

Begin by ...

Second Step

Then continue with ...

What's Next

This is what was achieved and what was omitted in this how-to.

Subhasish-TEST2

Authored by [YOUR NAME]. Last update: [FILL IN DATE]

Date of origin

1: Introduction

[ADD CONTENT]

2: New Heading

[ADD TEXT]

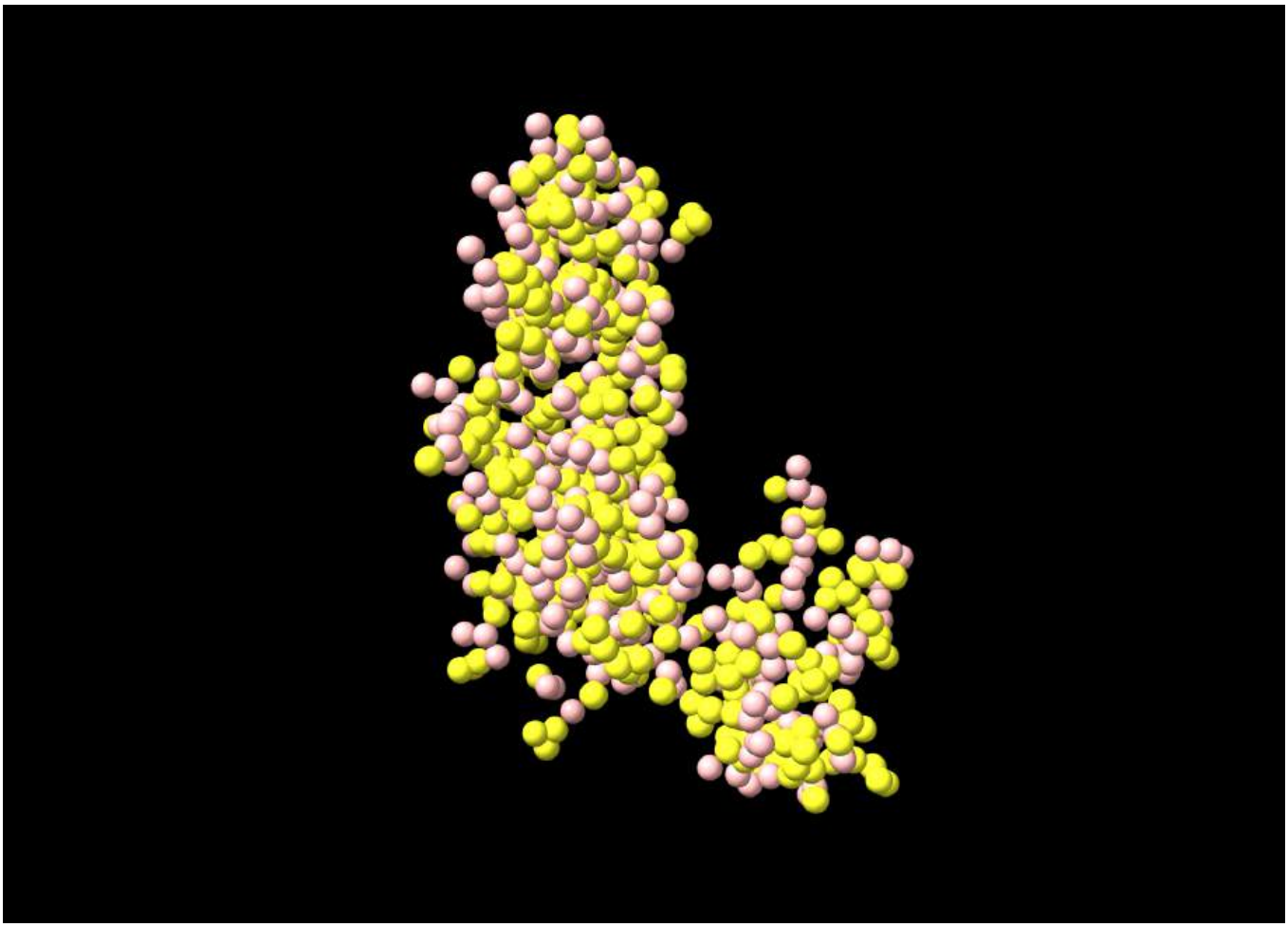
[ADD IMAGE] (saved to your computer and uploaded with picture icon from top menu bar or drag image file to location (required from svg image)

*Use the following under your picture:

Figure x: [Add caption]

*Center Picture and Caption together using top menu bar

[ADD iCn3D Model]



[ADD MATHEMATIC GRAPH - REUSE]

Figure x: Add caption

[ADD VCELL/SBML SIMULATION - REUSE]

*your text before and after insert as needed

3: New Heading

Add what you want